SMARTER CROWDSOURCING
ANTI-CORRUPTION

Smarter Crowdsourcing for Anti-corruption

A Handbook of Innovative Legal, Technical, and Policy Proposals and a Guide to their Implementation

Beth Simone Noveck
Kaitlin Koga
Rafael Aceves Garcia
Hannah Deleanu
Dinorah Cantú-Pedraza

APRIL 2018

IDB
Inter-American Development Bank

GOVLAB

anticorruption.smartercrowdsourcing.org

Acknowledgments

We acknowledge the contribution to the execution of this project to the following individuals from the Inter-American Development Bank (IDB):

- **Nicolás Dassen**, *Senior Specialist, Modernization of the Public Sector (ICS/IFD)*
- **Arturo Muente**, *Senior Specialist, Modernization of the Public Sector (ICS/IFD)*
- **Mario Sanginés**, *Principal Specialist in Mexico, Modernization of the Public Sector (ICS/IFD)*
- **Darinka Vásquez Jordan**, *Consultant, Innovation in Citizen Services (ICS/IFD)*
- **Michelle Marshall**, *Consultant, Knowledge Management and Open Innovation (KNM/KNL)*

Contents

FOREWORD

The rise in cybercrime and the proliferation of so-called “fake news” on social media have soured many people on the digital revolution. However, the same novel technologies that enable these distressing trends can also help us combat corruption much more effectively.

There are several encouraging examples in Latin America, a region recently swept by major corruption scandals, where digital tools are becoming a powerful means of rooting out fraud in government and foiling unscrupulous officials and their accomplices.

CUTTING RED TAPE

Complex bureaucratic procedures requiring a myriad of steps and approvals typically set the stage for corruption. Opportunities to demand or offer bribes tend to multiply when transactions between citizens and officials happen in person, as is still often the case in our region.

Recognizing this problem, governments across Latin America and the Caribbean are aggressively cutting red tape by simplifying and digitizing their administrative processes. By automating services and putting them online, governments leave corrupt officials less room to make arbitrary decisions.

Uruguay, for example, is close to its goal of enabling citizens to initiate 100% of their government transactions online. Panama is embarking on an ambitious program to put 450 key services online, and Argentina is also moving to become a paperless government.

Cutting red tape in these and other countries has proven to be a good thing. Apart from limiting corruption, simplifying and streamlining bureaucracy can make economies more competitive and strengthen trust in public institutions – two goals that our region must pursue with a passion. Ultimately, these efforts seek to deliver better services to meet the ever-rising expectations of our digital citizens.

LEVERAGING BIG DATA

The digital revolution is generating massive amounts of data that can be leveraged by governments, businesses and civil society organizations to make better informed decisions, improve services, and become more transparent and accountable.

By opening access to their data, governments are enabling citizens to track more closely how their taxes are spent. In Brazil, the Public Expenditure Observatory uses big data analytic tools to detect potential fraud in procurement. In 2015 it scrutinized more than 120,000 contracts, raising red flags in more than 7,500 cases involving $104 million in business. One of its filters, for example, identify when big contracts are split into smaller deals to avoid more competitive bidding processes.
The IDB is supporting several open data projects in the region, ranging from an anti-corruption package in Mexico, to local initiatives in Argentina and training Costa Rican officials in big data analysis. Countries are also carrying out their own experiments, such as the Mexican government’s testing of smarter crowdsourcing based on a methodology developed by New York University’s GovLab.

TRACKING HIGH-RISK AREAS

Since antiquity, public sector investments have been particularly vulnerable to corruption. But online data visualization platforms can be strong bulwarks against venality.

Colombia has created a platform called MapaRegalías where anyone can track how royalties paid by mining companies are spent on public works. The central government uses it to police compliance. Last

year it suspended transfers to 125 municipalities for not reporting data adequately. Similar platforms are being developed for Paraguay, Costa Rica and Peru.

While excessive controls can paralyze governments, the smart use of technological solutions can help make audits work more nimbly. Chile's Integrated Audit Control System provides a simplified and risk-based auditing process. The portal Contraloría-Ciuadadano is an effective communication channel between auditors and citizens, allowing the public to participate in the process.

Technology-based integrity solutions can also help fight money laundering. Secure digital identity systems can make it easier for banks to comply with "know-your-customer" regulations. Jamaica is introducing a national identity system that will facilitate such compliance. Beneficial ownership information not only ensures financial sector integrity, but also helps fight tax evasion and money laundering.

TESTING TRUST TECHNOLOGIES

New digital technologies are emerging and proliferating rapidly. One example is blockchain, which creates secure distributed ledgers guarded by many parties, protecting transactions from tampering.

Blockchain is widely known as the technology behind bitcoin, a digital currency. But it is also being tested to register property transactions and to combat blood diamond trafficking. The IDB is currently working with the Argentine customs agency and MIT's Media Lab on a project to use blockchain technology to identify and track cargo shipments crossing borders.

None of the benefits of the digital revolution will amount to much unless our countries embrace change. Technology is not a magic wand we can wave to fix everything. It is part of the solution, but its adoption must go hand in hand with a broader effort to strengthen institutions and rebuild civic capital.

We are optimistic on that score, too. Next year, Peru will host the Summit of the Americas, and the World Economic Forum on Latin America will take place in Sao Paulo Brazil. Democratic governance against corruption will be a central focus for the heads of state and government at those meetings. The private sector will also be represented, and has been involved in a dialogue with our region's governments fostered by initiatives such as the World Economic Forum's Partnership Against Corruption Initiative (PACI).

Coupled with political resolve, the digital revolution can help us get to our goals of greater transparency and accountability in government much faster and efficiently than we ever imagined possible.

Luis Alberto Moreno
President, The Inter-American Development Bank

This article was first published on the World Economic Forum's Agenda website as part of a World Anti-Corruption Day series curated by the World Economic Forum's Partnering Against Corruption Initiative (PACI).

PREFACE

There is a broad consensus among international organizations and academia that corruption and lack of transparency erodes public trust and weakens public institutions and the rule of law. It also has a negative effect on the economy. It erodes the efficient allocation of public spending, encourages tax evasion, generates additional costs in financing and transactions, discourages investment, and negatively impacts productivity.

The Government of Mexico approached the Inter-American Development Bank for support to identify innovative solutions to tackle corruption from different angles, including the prevention and detection of corruption. In search of those innovative and practical strategies, we collaborated with the NYU GovLab to carry out a Smarter Crowdsourcing exercise with various scholars, civil society representatives, and practitioners from around the world. This report contains the result of the requested assistance, with proposing solutions accompanied by practical insights describing what is required to implement them.

Designed and implemented by the NYU GovLab, Smarter Crowdsourcing relies on the use of collective intelligence to harness the diverse and combined insight of 100+ global experts around a specific challenge. Through this process, we were able to move beyond the traditional agenda of transparency and anticorruption, and move toward open government.

An open government promotes a completely different relationship between the State and its citizens. It aims to strengthen democracy and build trust in government, while improving the efficiency, effectiveness, and transparency of public services through the effective use of new technologies. It's also about implementing a citizen-centric approach by engaging citizens in collaborative schemes to co-design and co-implement public policy decisions, and by opening governments to public scrutiny to make officials more accountable. In this sense, "collaboration" is an inspiring principle of open government initiatives, and Smarter Crowdsourcing is probably one of the main methodologies that captures that spirit.

This is not just crowdsourcing but "expert sourcing." Although political will can be a significant roadblock in the fight against corruption, equally challenging and important are the difficulties many public institutions face when attempting to implement cutting-edge technological solutions in legacy institutions which sometimes lack the infrastructure, personnel or culture of innovation necessary for success.

For this reason, it is my desire to share the findings of the report, as a contribution to the countries of Latin America and the Caribbean, of which the IDB is proud to support in their efforts to promote transparency and strengthen their institutional systems of integrity. Sharing the results of this process will allow any public or private institution to access the most innovative ideas along with the practical know-how to implement them.

Nicolás Dassen
Senior Modernization of State Specialist, The Inter-American Development Bank

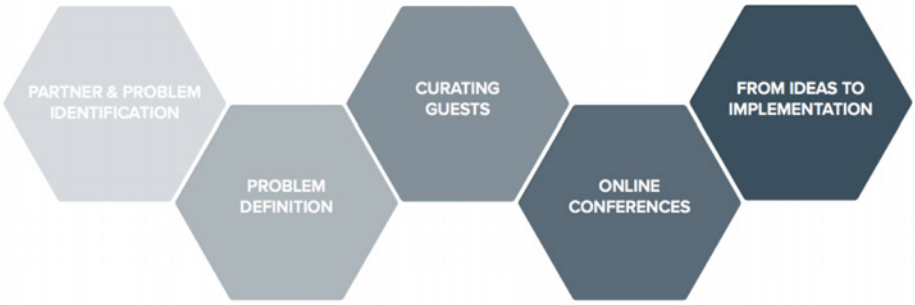
EXECUTIVE SUMMARY

Corruption presents a fundamental threat to the stability and prosperity of Mexico and combating it demands approaches that are both principled and practical. In 2017, the Inter-American Development Bank (IDB) approved project ME-T1351 to support Mexico in its fight against corruption using Open Innovation. Thus, the IDB partnered with the Governance Lab at NYU to support Mexico's Secretariat of Public Service (Secretaría de la Función Pública) to identify innovative solutions for the measurement, detection, and prevention of corruption in Mexico using the GovLab's open innovation methodology named Smarter Crowdsourcing.

The purpose of Smarter Crowdsourcing was to identify concrete solutions that include the use of data analysis and technology to tackle corruption in the public sector. Although written at the behest of and for the Mexican context, the recommendations and plans for their implementation developed in this report could be adapted for use in other countries.

This document contains 13 implementation plans laying out practical ways to address corruption. The plans emerged from "Smarter Crowdsourcing Anti-Corruption" (2017). The Smarter Crowdsourcing method is an agile process, which begins with robust problem definition followed by online sourcing of global expertise to surface innovative ideas and then turns them into practical implementation plans.

FIGURE 1. SMARTER CROWDSOURCING METHODOLOGY



Smarter Crowdsourcing Anti-Corruption (2017) focused on six specific challenges:

- **Measuring corruption and its costs:** Where and with what frequency do acts of corruption occur? How much does corruption cost and how else does it impact people? How do the costs of corruption change over time? How can we develop rigorous and informative measures of corruption and its associated harms?
- **Strengthening integrity in the judiciary:** How can we increase judicial transparency and oversight without undermining judges' independence and autonomy?
- **Engaging the public in anti-corruption efforts:** How can we include citizens in the policy-making process in ways that are meaningful, inclusive, and visible?

- **Whistleblowing:** What deterrents prevent whistleblowers from coming forward? How can we develop systems that protect tippers' anonymity while also ensuring the reliability and legal admissibility of such evidence?
- **Effective prosecution:** What internal and external factors prevent more successful prosecutions from being brought against corrupt officials? How can we ensure that prosecutors are equipped with the necessary tools to hold corrupt officials to account?
- **Tracking and analyzing money flows:** What kinds of technological and data-driven approaches can government or civil society use to identify and preempt misappropriation of government resources?

Summary of Solutions

The report includes implementation plans for thirteen data and technology-based solutions to these six specific corruption problems. Informed by research and in-depth interviews, each plan includes a detailed description of the problem and the solution, outlines the necessary resources and assets needed to implement the solution successfully and identifies the relevant personnel and the skills they must have. Each also includes references to global experts who can further advise on the topic.

PROBLEMS	SOLUTION RECOMMENDATIONS
1. MEASURING CORRUPTION AND ITS COSTS How can we improve our ability to measure corruption and identify when and where it occurs?	**1.1 DATA COLLABORATIVES** - Establish public-private partnerships ("data collaboratives") to utilize private as well as public sources of data to improve its estimation, drawing on the example of UNICEF's use of data collaboratives.
	1.2 MEMEX - Use MEMEX open-source search tools for advanced online search to improve the extraction of information relevant to uncovering hidden corruption in public procurement. MEMEX is a research program initiated by the US Defense Advanced Research Projects Agency (DARPA), whose goal is to create software that can better search the world wide web and is in widespread use in US law enforcement.
	1.3 RED FLAGS - Develop red flags to build a predictive analytics tool capable of revealing incidents and patterns of corruption and of estimating corruption and its costs in government contracting data---as done in Brazil.
2. STRENGTHENING INTEGRITY IN THE JUDICIARY How can we use data and technology to increase transparency and improve oversight over the judicial system?	**2.1 OPEN COURTS** - Develop a national "open courts" portal to publish data openly and online about judicial and court performance---as done in Argentina, Slovakia and Lithuania.
	2.2 PUBLIC MONITORING - Integrate public oversight into judicial selection and evaluation by disseminating data on judicial appointments and judges' assets, interests and networks as open data.

PROBLEMS	SOLUTION RECOMMENDATIONS
3. ENGAGING THE PUBLIC IN ANTI-CORRUPTION EFFORTS Can technology enable the public to play a greater role in reporting corruption and making corruption policy?	**3.1 COMPLAINT REPORTING** - Use mobile technology, human-centered design and prize-backed challenges to update and improve the existing online reporting platform to increase the number and quality of reports received---as done in India and Pakistan.
	3.2 CITIZEN ENGAGEMENT - Design and launch an online mechanism for individuals and organizations to contribute to the Federal Administration's anti-corruption policy-making efforts and foster ongoing conversation with a more diverse audience---inspired by the best of crowdlaw programs around the world.
4. WHISTLEBLOWING What are the innovative strategies for increasing corruption reporting by employees in government and companies?	**4.1 ONLINE REPORTING** - Optimize online reporting systems with enhanced security and accountability, utilizing encryption technologies to protect anonymity (like the SEC) and develop co-accountability for follow-up with civil society organizations.
	4.2 TRAINING - Train managers on how to create an environment conducive to reporting and how to follow up on reports in a quick and effective manner---as done by the best companies.
5. EFFECTIVE PROSECUTION How can technology be employed to enhance prosecutorial accountability and effectiveness?	**5.1 COMMUNITY OF PRACTICE** - Build an "Open Justice" community of practice of anti-corruption prosecutors supported by an online learning platform.
	5.2 OPEN DATA - Create a website to enable the public to monitor the progress of corruption prosecutions.
6. TRACKING AND ANALYZING MONEY FLOWS What kinds of technological and data-driven approaches can government or civil society use to identify and preempt misappropriation of government resources?	**6.1 BENEFICIAL OWNERSHIP** - Develop financial incentives as well as legal requirements to encourage voluntary disclosure of comprehensive corporate beneficial ownership and other legal entity data. Adopt Global Legal Entity Identifier.
	6.2 BLOCKCHAIN - Map the demand for blockchain technology in the procurement process and publish a primer on the implementation of blockchain in the procurement process as a first step toward exploring blockchain implementation to enhance integrity.

ISSUE AREA 1

Measuring Corruption and its Costs Using Big Data and New Technologies

New technology and new sources of data offer the potential to innovate and improve upon previous methods of measurement that can help governments estimate more accurately the impact, frequency, and cost of corruption, particularly in public procurement. Based on the priorities articulated by Mexico's Secretariat of the Civil Service (*Secretaría de la Función Pública* or "SFP") this memo makes the following three recommendations, which we will discuss at length in this memorandum:

- Establish public-private partnerships ("data collaboratives") that will constitute a hub of public and private data sources to address corruption and improve its estimation.
- Use MEMEX open-source software for advanced online search to improve the extraction of information relevant to uncovering hidden corruption in public procurement.
- Use red flags to build a predictive analytics tool capable of revealing incidents and patterns of corruption and of estimating corruption and its costs.

By building on international experience in governments in Europe and Latin America, where these techniques have been tested and significant academic research has been conducted, we believe that an institution with comparable capacity to SFP would be able to accomplish these projects in one year. The implementation by SFP will depend upon either the development of an in-house data analytics capacity or collaboration with university or private sector partners external to the government, as well as partnership with relevant civil society groups such as the Open Contracting Partnership. If successful, these projects will create the analytical and human capital infrastructure needed to monitor and rout corruption from public procurement. This infrastructure can then be repurposed and applied to spotting patterns of corruption in other contexts.

I. PROJECT BACKGROUND

On June 6, 2017, 24 experts from nine countries joined officials representing the Mexican government, the Inter-American Development Bank, and members of the GovLab in a two-hour online conversation to identify data- and tech-enabled strategies that the Mexican government and civil society leaders can use to measure corruption and its costs. This conference took place in the context of landmark legislative and institutional reforms that Mexico's National Anti-Corruption System aims to implement. Crucial to the success of this ambitious reform is a rigorous set of measures of corruption and its associated harms. These will allow agencies and the public to track the progress of the National Anti-Corruption System, and will serve as the necessary empirical foundation for crafting and adjusting anti-corruption policies.

Subsequent discussion focused on innovative measurement strategies that take advantage of public information and new sources of data to:

1. Estimate accurately the impact, frequency, and cost of corruption, while
2. Increasing the government's responsiveness in pinpointing incidents of corruption in the public sector.

The conference resulted in seven concrete recommendations. This document provides the broad implementation outlines of the three recommendations selected by Mexico's Secretaría de la Función Pública. The content of the document, including its recommendations, is the sole responsibility of the GovLab and does not represent the IDB's official position or view on this matter, or an endorsement of any individual or firm to perform activities related to the recommendations.

II. PROBLEM OVERVIEW

Mexico's National Anti-Corruption System[1] aims to **establish**, **articulate** and **evaluate** policies to eradicate systemic corruption across local, state, and federal governments. Public procurement is one of the largest areas of public spending, and expert and public perceptions worldwide suggest that it is the governmental function most vulnerable to corruption[2]. At the same time, the importance of public procurement for every member of society cannot be overestimated.[3] Countering corruption in public procurement will not only reduce the waste of public money, but can also prevent the purchase of unnecessary or poor quality goods and services.[4] Crucial to the success of Mexico's anti-corruption reforms is **a rigorous set of measures of corruption and its associated harms** that can be used both to track the progress of the National Anti-Corruption System and to serve as the necessary empirical foundation for crafting and adjusting anti-corruption policies.

Current measurements of corruption, in Mexico as well as other countries, are inadequate to the task of pinpointing where corruption occurs.

---

[1] A summary about Mexico's National Anti-Corruption can be found here.

[2] Rose-Ackerman, Susan. "**The challenge of poor governance and corruption.**" *Especial 1 DIREITO GV Law Review* (2005), pp. 207-266, available here.

[3] Ibid.

[4] Tanzi, Vito and Hamid Davoodi. "**Corruption, public investment, and growth**." *The Welfare State, Public Investment, and Growth*. 1998, pp. 41-60.

Problems With Estimation

Most corruption estimates are based on **perceptions of experienced corruption rather than data collected directly from the field.**[5] For example, each year, Transparency International creates a Corruption Perception Index (CPI): a composite index of surveys created from polls of analysts, businesspeople, and other experts in countries around the world. This index is assumed to reflect the actual incidence of corruption in each country. The goals of these estimates are to connect citizens' perceptions of corruption to the forms of governance that can address those complaints, and to push for reforms through naming-and-shaming practices.[6]

In 2016, Mexico's National Institute of Statistics and Geography (INEGI) conducted a survey revealing that more than half of the surveyed Mexican companies reported paying a bribe to expedite business procedures. On the basis of this survey, INEGI estimated that the cost of corruption in 2016 exceeded $80 million.

But several limitations hamper current estimation techniques:

- Corruption perception indices cannot reveal the locus of corruption or help evaluate the effectiveness of anti-corruption policies.[7] Broad spectrum victimization surveys do not identify corrupt practices, culprits, or corruption networks, and do not help government agencies tackle corruption in real time.
- These indices rely on the fundamental assumption that perceptions reflect reality, yet experience shows that they are sensitive to and influenced by media reports.[8]
- The surveys also have large margins of error by relying on flawed assumptions: that the experiences of the questioned sample reflect the experiences of the people involved in public procurements, that there is truthful reporting, etc.[9]

---

[5] TI's surveys are increasingly focused on the public services known to be sensitive to corruption - e.g., issuance of driving licenses, school enrollment, and medical care. Similarly, the WEF's Executive Opinion Survey asks companies to report the fraction of the contract value solicited in bribes.

[6] Braithwaite, John. (2006). "**Responsive Regulation and Developing Economies**". *World Development*, vol. 34, no.5, pp. 884-898.

[7] Mungiu-Pippidi, Alina. (2017). "**The time has come for evidence-based anticorruption**", *Nature Human Behavior*, vol. 1, no. 0011, DOI: 10.1038/s41562-016-0011.

[8] Nguyen, T. (2017) Anti-corruption Media Coverage and Corruption Perception, *Indiana University Working Paper*, available here.

[9] For instance, the Global Corruption Barometer of Transparency International asks citizens whether they paid a bribe when interacting with a public servant working in, among others, the education services, police, and the judiciary. This measure is not adequate to measure corruption in public procurement because it relies too much on the assumptions that respondents involved in public procurements are responding truthfully, and that randomly selected citizens have experience with public procurements. The EBRD and World Bank's Business Environment and Enterprise Performance Survey, which is a confidential questionnaire sent to businesses about their experiences in dealing with government officials, is a general survey of doing business. Although it asks about corruption and bribery, it focuses neither on these issues nor on the region and is not designed for the purpose of fixing corruption.

- Estimates remain abstract, unable to trigger responses from the responsible authorities either to uncover and redress current incidents of corruption or to prevent future incidents.[10]

Measuring corruption is different from measuring other social phenomena because it is an iceberg where the majority of the activity remains fundamentally hidden. Corrupt agents obviously wish to remain unobserved and therefore will not self-report corruption or will report untruthfully. Similarly, ordinary citizens will have a hard time perceiving the extent of corruption, and where and how it is carried on.

Problems With Data Collection

Several problems with existing data on corruption make responses ineffective. On the positive side, granular data about public procurement, which is relevant for measuring corruption, exists in open, accessible and standardized databases.[11] On the other hand, the data is not strategically connected and often does not allow government institutions to timely, effectively, and efficiently gauge the size and scope of corruption. Moreover, it is widely agreed that the databases lack indicators that could provide more specific or more timely estimates of corruption, including predictions about which tenders are likely to be corrupt in the future. Observers cannot use the data to trace public contracts, public servants, and their relationships to the bidding companies from the inception of the procurement to its full completion.

Instead, there is a need for **accurate, timely, and objective data** to monitor public contracts, public servants, and their relationships to the bidding companies from the inception of procurement to its full completion.

For an extended discussion of the problems at issue, see full problem brief here.

III. SUMMARY OF SOLUTIONS

This memo proposes that the Government of Mexico generate mechanisms to tap into relevant and readily available open data sets and utilize predictive analytics to identify and measure corruption in public procurement. Responses to corruption can take place at different stages:

- *Before it happens* – by predicting the probability that certain companies, people, and public servants are susceptible to corruption
- *As it happens* – by revealing the incidences of corruption and allowing law enforcement and supervisory organizations to start an investigation
- *After it happens* – by providing the evidence necessary to prosecute corrupt deeds successfully and by estimating the incidence and costs of corruption in certain contexts

Mexico can achieve this by (1) establishing collaborations with public and private entities in order to **unlock the potential of readily available open data sets;** (2) **using advanced data-mining tools** to improve content discovery, information extraction, and information retrieval relevant to corruption; and (3) using **advanced analytics** on the available open datasets to **measure the incidence of corruption** in public procurements

---

[10] Mungiu-Pippidi, Alina. (2017). "**The time has come for evidence-based anticorruption"**, *Nature Human Behavior*, vol. 1, no. 0011, DOI: 10.1038/s41562-016-0011.

[11] In 2016, Mexico ranked 16 in the Global Open Data Index, available here.

and its costs in a timely, efficient, and granular way. These solutions have been articulated in accordance with the priorities of the National Anti-Corruption System[12] and are based on the latest developments in informational technology and governance as well as the latest criminological research.

Countering corruption in public procurement starts with the triggering of a suspicion. While open contracting data helps analysts, auditors, and the public investigate public expenditure and triggers suspicions where appropriate, it is not the only source of information. For example, the Mexican Financial Intelligence Unit (FIU) receives suspicion reports from a broad array of economic agents – financial entities, lawyers, notaries, traders in high value goods, exchange offices, etc.– on the grounds that relying on a variety of sources of information can best trigger justified suspicions instead of false positives (instances where the analytical tool wrongly suspects a procurement of being corrupt).[13]

Recent research shows that social media is another rich source of information. Academics, lawyers, law enforcement and other white-collar crime analysts from across the world are beginning to understand that social media data is a low-hanging fruit that can be exploited for criminal investigations and associated trends. Social media may contain suspicions of corruption that carry a timestamp, can be geographically located, and refer to a particular incident type of corruption or corrupt individual.

Similarly, suspicion of corruption in public procurement can be triggered by **searching the dark web**. Developments in informational technology funded by the US Defense Advanced Research Projects Agency may allow for the discovery of complex patterns of corruption that are otherwise invisible to the naked eye. These deep web tools have already been successfully applied to searching the dark web for suspicion of human trafficking. The algorithms searched through immense amounts of data for human trafficking patterns and were able to correctly identify human trafficking ads. Monitoring these ads, the frequency of their occurrence, and their geolocations helped law enforcement effectively measure human trafficking activities in the US and their developments, in real time.

Finally, research conducted by the GovLab indicates that open data has been effectively used across the world to reduce fraud, waste, and inefficiency. Open data initiatives that were targeted at countering corruption in public expenditure have amplified accountability and curbed corruption in federal

---

12 In the context of the National Anti-Corruption System, a priority of the Citizens' Participation Committee is to "promote the collaboration with public institutions dedicated to the battle against corruption, so that measuring and evaluation tools are developed with the goal of preventing, detecting and fighting corruption." Similarly, the Coordinating Committee must "establish adequate tools and indexes for periodic evaluation of the anticorruption policy" and "create the necessary alliances and collaboration programs in order to access datasets and information required to carry out its duties." The priority of the Ministry of Public Function is to "create digital platforms that contribute to intelligence gathering." Finally, the Superior Audit Office must "institute covenants with other public institutions, to facilitate the exchange of information, between the offices and their public officials, so that it can access any data it needs to carry out its functions."

13 Mexico's Federal Law to Prevent and Identify Transactions Involving Resources Illegally Obtained requires obligated entities (e.g. banks, insurance companies, exchange offices, lawyers, notaries, etc.) to identify their customers when entering a contractual relationship and to report suspicion of money laundering or the predicate offence.

expenditures in Brazil, public procurements in Slovakia, and expenditures on international development cooperation in Sweden.[14]

A 2015 Organization for Economic Co-operation and Development (OECD) assessment places Mexico above the OECD-average and well ahead of Sweden (see Figure 1 in Appendix I). Not surprisingly, although Mexican governmental data is highly rated in terms of openness and relevance to countering corruption, the availability of open data by itself has not led to performance improvements. Open data, in and of itself, does not solve problems without a plan to use it. A 2016 OECD report on Mexico reveals that Mexico's open data initiative has delivered few of the promised improvements.[15] To this end, our third solution proposes the development of a tool that analyzes this wealth of open data in order to develop indicators of corruption that are applicable to Mexico's psychographics and that are able to predict instances of corruption in real-time. Such predictive tools have the capacity to help law enforcement, auditors, and risk managers estimate the probability of corruption that occur in decisions, contracts, or procurements.

1. Increase sources of data by establishing public-private partnerships (data collaboratives) that will constitute a hub of public and private data sources to address corruption and improve its estimates

What needs to be done and why?

Data collaboratives are "*a new form of collaboration, beyond the public-private partnership model, in which participants from different sectors – including private companies, research institutions, and government agencies – can exchange data to help solve public problems."*[16] By serving as frameworks for safe and responsible cross-sector data sharing, they allow the existing vast stores of privately held data to be used for public good. Across a number of sectors, including humanitarian and anti-poverty efforts, urban planning, natural resource stewardship, health, and disaster management, national authorities are using data collaboratives to harness private sector data, such as automotive sales data, telecommunications records, and social media data in order to improve current measurements of corruption in public procurements.

Why social media data? **Social media data can be used to enable governmental authorities to identify and analyze new suspicions of corruption** - e.g., suspicions of personal enrichment through the acceptance of kickbacks that are otherwise not captured through surveys and that are not reported to the relevant authorities, personal confessions of corruption and bribery, or other expressions of power and

---

[14] Graft, A. et al. (2016). Brazil's Open Budget Transparency Portal: Making Public How Public Money is Spent, *GovLab report,* available here; Clare, A. et al. (2016). Open Contracting and Procurement in Slovakia: Establishing Trust in Government through Open Data, *GovLab report,* available here; Clare, A. et al. (2016) OpenAid in Sweden: Enhanced Transparency and Accountability in Development Cooperation, *GovLab report*, available here.

[15] The report found that despite the ambitious implementation of the National Open Data Policy (e.g., the issuance of the Open Data Executive decree, the deployment of a fully functional open governmental data portal, the establishment of regulatory and technical support bodies and guidelines, and the implementation of several initiatives to promote collaboration among public institutions on solutions to public issues), open data was not able to make a significant impact on Mexican economy and society. In particular, the data has insufficiently engaged the public, open data was not reused, and there were no metrics for evidence-based policy making. (OECD, 2016, p. 13).

[16] Stefaan G. Verhulst, Introduction of Data Collaboratives, available here.

untouchability that reveal corrupt incidents, practices, individuals, or networks of individuals. Social media has previously been used to track undeclared assets that cannot be explained by income level. Undeclared wealth is not readily available to the agencies charged with monitoring public expenditure, and just as with testimonies of witnessed corruption, they can escape the attention of the relevant agencies. While victimization surveys and citizen participation apps may capture a significant portion of testimonies of witnessed corruption, they may be effectively supplemented by social media reports (e.g. bragging about or posting a picture of one's Ferrari). Criminals do not report their crimes to the Government, but they often brag about "getting away with it" to their friends, family, and sometimes social media followers.[17] While seemingly improbable, many books, series, and documentaries tell the story of how criminals want to get recognition for their ability to commit the crime and get away with it. As Alex Gibney, director of *The Armstrong Lie*, says: *"This is not a story about doping. It's a story about power. And the story became hanging onto that power."* Tracking these acts of confession are fundamental for detecting and measuring corruption, as they give away both actors and *modi operandi*.

Establishing a Data Collaborative

Twitter already provides most of its data to select partners through the "Firehose API," including to the MIT Media Lab, UN Global Pulse, Dataminr, Datasift, Gnip, Lithium, and Topsy. Governments can partner with Twitter, with one of those organizations, or with a university to access, mine and analyze Twitter data, but users need a plan for making use of it effectively, the details of which we discuss below.

Suggested plan: Research conducted by the GovLab suggests the following steps need to be completed in setting up a data collaborative.

CORE ACTION 1

FORMULATE THE GOAL AND VALUE PROPOSITION OF THE DATA COLLABORATIVE

Task description: This task aims to: (1) formulate the overarching societal problem addressed through the data collaborative, (2) describe the reasons for choosing a data collaborative for achieving the earlier set goal, and (3) enumerate the assumptions needed to reach the goal. This needs to be sufficiently precise to ensure a well-targeted collaborative.

Who: The Coordinating Committee of the National Anticorruption System must formulate the goal for which the data collaborative is being built, the reasons why, and the assumptions needed in order to reach the goal.

How: In order to formulate the goal of the data collaborative, we recommend that the Coordinating Committee use the Problem Solving Canvas[18] developed by the GovLab. This tool helps organize the discussion in order to:

---

[17] LexisNexis Risk Solutions, 2012. Survey of Law Enforcement Personnel and Their Use of Social Media in Investigations, available here, shows how social media is effectively used to collect evidence, locate suspects, and identify criminal networks.

[18] In designing the problem to be solved, useful resources can be found here and here.

1. Describe the issue to be addressed;
2. Identify the beneficiaries;
3. Articulate why it matters now;
4. Take stock of the underlying assumptions; and
5. Take stock of potential counterarguments and risks.

In formulating the value proposition[19] of the data collaborative, we recommend answering the following questions:

1. What is the intended societal benefit of the data collaborative?
2. How will the collaborative be designed and implemented?
3. How will the collaborative be assessed and evaluated?

Why: Having a goal that is well formulated and easy to understand increases the chances of a well-functioning data collaborative. Having the Coordinating Committee formulate the goal and the value proposition of the data collaborative ensures that the key stakeholders are consulted through their representatives on the Coordinating Committee. Furthermore, the discussion of how a data collaborative should help the Government better measure corruption in public procurement needs to reflect on the risks this method entails. These risks should be balanced against the expected gains.

As examples of expected gains, an increased ability to track corrupt deeds in public procurement helps the Secretaría de la Función Pública better monitor some public servants, review existing procurement policies, and promote practices aimed at inhibiting corruption. Furthermore, suspicions raised through this data collaborative may help the Superior Audit Office focus its verifications on the suspected sectors and individuals and conduct more targeted investigations - e.g., by looking for kickbacks whose values approach 10% of the value of a contract in which a public official known as "Mister 10%" was involved.

CORE ACTION 2

INVENTORY EXISTING DATA AND DATA GAPS

Task description: The goal of this task is to understand what data the government has, as well as data the government does not have but could benefit from having to better measuring corruption in public procurement.

Who: The Coordinating Committee must discuss and review private data sources that may be relevant for measuring corruption and complement what the government already holds.

How: According to Warwick Graco (Senior Director Operational Analytics at the Australian Tax Office), identifying the needed data is a simple concept that is hard to put into practice, as there are lots of data from which to choose. In order to identify the relevant data to be gathered through the collaborative, the government should undertake a step-wise process of inventorying its data.

[19] Resources that support the building of a compelling value proposition are available here and here.

These steps are the necessary basics for setting up a data collaborative, undertaking further network analysis, or creating a challenge using the data. To translate open data into performance improvements, it is necessary to know what one is trying to accomplish, what data one has, and how to make sense of it.

The following figure illustrates the different components of a thorough analysis of open data, each component requiring a different set of expertise:

A step-wise approach to a data-driven analysis of Mexican open datasets:

Although the order of the components can vary, the analysis may start from:

1. Asking **what questions the data should answer.** (e.g., "Which of the money flows are likely to signal kickbacks in public procurements?") There is ample **theory to suggest how corruption takes place**, what are 'red flags,' and what anti-corruption policies work. There is also ample **theory on how to construct and interpret predictive models**, distinguish significant results, and remove noise.
2. **Cleaning and cleansing the available data**, and displaying it in the form of summary statistics or data visualizations.
3. Sharing the **analyses and the results** with the broad public.
4. **Searching for new datasets to link to the original open data** to enhance the analyses and better enable the detection of corrupt networks and patterns of corruption. This fourth piece might involve a data collaborative or another network analysis project.

STEP A DEVELOP THE QUESTIONS THAT THE OPEN DATA MUST ANSWER

Task description: Assess what the Mexican public expects open data about public procurements to provide.

Who: The Coordinating Committee should compose a list of questions through desk research (e.g., what the World Bank, OCP, and other non-profit organizations have found out by asking the public and the experts what types of open data could counter corruption) and ask the Mexican public for feedback on it.

How: Launch a **call for feedback** page and ask citizens to read the set of questions found in the desk research and rank them according to their urgency. Additionally, allow citizens to complete this list with their own relevant questions.

Why: A robust initial set of questions offers a benchmark to judge the answers that open data provides. Asking the "crowd" to brainstorm allows for the emergence of diverse questions and has multiple benefits: it empowers citizens, reveals the most pressing questions, publicizes efforts to answer them, and legitimizes further efforts to open more data.

Costs: The cost should be minimal. The construction of the call for feedback page (the desk research, the web-page design) can be completed within 2 weeks by one person working full time. Experience needed includes web design and knowledge of the literature on corruption. The actual call for feedback should last no more than 6 weeks.

STEP B FORMULATE A RESEARCH DESIGN TO ANSWER THE QUESTIONS THAT THE OPEN DATA SHOULD ANSWER

Task description: The task aims to construct the **research design** on the basis of which network analysis tools and machine learning algorithms will be applied to answer the questions identified during Core Action 1.

Who: The Coordinating Committee should reach out to relevant epistemic communities - e.g., members of academia with expertise in corruption, crime, public procurement, and data science - and stimulate them to formulate testable hypotheses and models to answer the questions identified during Core Action 1.

How: A public agency such as the SFP can launch an **ideation challenge** and support ideas proposed by multidisciplinary teams.

- Through **ideation challenges** researchers can be stimulated to formulate research designs for questions identified as relevant during Core Action 1. The best ten models should be awarded symbolic cash prizes and the funding to execute the design. Using a similar model, the Alfred P. Sloan Foundation awarded $15,000 for good ideas on what to fund in connection with research for the White House Smart Disclosure Initiative.

- **Matchmaking** can be used to form multidisciplinary teams with complementary expertise who can design research protocols that can later be applied on the open datasets. Experts can originate from the relevant fields: data science, criminology, economics, law and psychology. Using a similar design, the North Atlantic Tourism Association issues grants to support tourism and cultural exchange projects in Greenland, Iceland, and the Faroe Islands and requires applicants to work across borders and to find business partners and new ideas using their own matchmaking tool.

Why: Developing and registering the theoretical premises for an analytical exercise is a robust way to ensure that the risk of researcher bias, otherwise observed with quick-and-dirty analyses, is eliminated.

Costs: Cash prizes need not be great because reward mechanisms are built into the challenge's design. Multidisciplinary teams offer great scope for knowledge creation and dissemination. Setting-up an ideation challenge can take between 7 - 9 weeks, divided as follows: (1) Allowing multidisciplinary teams to be formed can take up to 2 weeks; matchmaking tools have been developed for scientific purposes as well as for B2B interactions with costs of no more than $1,000; (2) The submission of proposals may take another 4 weeks - e.g., Your City, Your Data initiative allowed 4 weeks for proposal submissions; (3) judging the proposals and awarding the winners can take another 2 weeks.

STEP C PREPARE THE OPEN DATASETS FOR ANALYSIS

Task description: This task aims to cleanse and clean the open data that is needed to apply the theoretical models developed during Core Action 2.

Who: The Secretaría de la Función Pública can source the data preparation to the same epistemological community identified during Core Action 2.

How: Launch a **data cleaning crowdsourcing initiative** on the data identified in the proposals submitted in the ideation challenge during Core Action 2. Allow students to take part in this initiative and offer outcome-based financial rewards.

Why: An empirical exercise requires carefully prepared data. Mistakes, omissions, and other inaccuracies may have significant consequences for the analysis. Data cleaning could be done manually (e.g., scientists at NASA have turned to "citizen scientists" (volunteer hobbyists, amateur science buffs, and space enthusiasts) to classify the images the Hubble Space Telescope recorded of the Milky Way and other galaxies according to their shape: elliptical, spiral, lenticular, irregular[20]) or with the help of machine learning algorithms (e.g., coders at Code4Romania employ machine learning algorithms to transform the vast data on the wealth of politicians into usable data, while the AMPLab at UC Berkeley has developed several machine learning and human hybrids that help clean large datasets fast).

[20] Noveck, Beth Simone. "**Smart Citizens, Smarter State: The Technologies of Expertise and the Future of Governing**" (pp. 4-5). Harvard University Press. Kindle Edition.

Costs: Depending on the size of the datasets to be cleaned, this exercise can take 3 - 12 months. Crowdsourcing the microtasks can be done using various crowdsourcing platforms - e.g., Amazon Mechanical Turk, Clickworker, Microworkers, Kaggle, etc.

STEP D VISUALIZE THE CLEANED DATASETS

Task description: This task aims to meaningfully visualize the data cleaned at Core Action 3.

Who: The Secretaría de la Función Pública should reach out to members of the epistemological community identified during Core Action 2 and stimulate them to build data visualizations of the data cleaned during Core Action 3.

How: Launch a **visualization challenge** to stimulate researchers to apply advanced graphic tools to the cleaned open data. Reward the best visualizations with a symbolic reward and post the visuals on the Mexican Open Data Portal.

Why: Visualizations are crucial to understanding complex social problems, as recognized when Eyebeam: a leading US art and technology center, launched the Data Viz Challenge to engage citizens to explore the question "where are our tax dollars being spent?" Additionally, visualizations can help generate new hypotheses, reveal which actionable data are really also action-worthy and provide an additional review of the research designs generated earlier.

Costs: Setting-up a visualization challenge can take between 5 - 7 weeks, divided as follows: (1) The submission of visualizations may take 4 weeks; (2) judging the proposals and awarding the winners can take another 2 weeks.

STEP E DISSEMINATE THE ANALYSES AND THE RESULTS

Task description: This task aims to collect and disseminate in a coherent manner the results of the research designs that were executed on the cleaned open data.

Who: The Secretaría de la Función Pública should support the multidisciplinary teams that executed the research in disseminating their results.

How: Fund a dissemination conference and a conference proceedings publication. The conference should allow the research teams to present and discuss their results. The conference should aim to attract a broad and diverse audience, including members of the scientific community, advocacy groups, law enforcement, public servants, and journalists, and should be publicly broadcasted.

Why: GovLab's research shows that open research best serves the interests of the public. Conferences allow the public to inquire into the details of the research: assumptions, margins of error, implications, etc. Depending on the diversity of the research designs, conferences can be used to meaningfully integrate the different research into one big effort to use open data to counter corruption in public procurement.

Costs: Organizing a conference may take one full time employed event planner and a part time employed scientist between 4 - 9 months. According to the Institute of Electrical and Electronics Engineers, the costs of an academic conference with 150 participants may range from $100,000 to $150,000. Costs can be covered through conference fees and sponsorships.

STEP F DEVELOP A LIST OF COMPLEMENTARY DATASETS

Task description: This task aims to generate a list of private or public data that could be added to the existing open datasets and meaningfully support current efforts to utilize big data and machine learning algorithms to detect, prevent, and counter corruption in public procurement.

Who: The Secretaría de la Función Pública can outsource the finding of complementary datasets to the crowds.

How: The Coordinating Committee can launch an **ideation challenge** to stimulate citizens to identify datasets that are open or could be opened up and that could meaningfully add to the analyses disseminated later during Core Action 5. The best proposals should be awarded symbolic cash prizes.

Why: A complex, intertwined set of data offers a richer exploration of the corruption patterns, possibilities, and networks that are present in public procurements. Asking the "crowd" to brainstorm allows for the emergence of diverse ideas and offers legitimacy and publicity to the efforts to link this new data to the existing ones.

Costs: Setting-up this ideation challenge can take between 7-9 weeks, divided as follows: (1) designing the challenge and posting it online may take two weeks; (2) the submission of proposals may take four weeks; (3) judging the proposals and awarding the winners can take another two weeks. Our interviews show that designing an ideation challenge may involve hiring a full time web developer and acquiring a domain name, or using already developed online platforms and tools.

Why: Creating a data inventory and an assessment of gaps helps to keep track of data resources already held and new resources that are needed as the data collaborative unfolds. This can be used for planning and implementing novel data management measures throughout life of the data collaborative.

CORE ACTION 3

SELECT THE MOST PROMISING DATA PROVIDERS

Task description: The goal of this task is to identify the data providers who hold and could provide the missing data identified during Core Action 2, and to **select the most promising data providers.**

Who: The Coordinating Committee must identify the relevant data providers and select the ones that are willing to join the partnership under reasonable conditions.

How: In order to identify the companies that are best suited to provide the needed data, the government should:

1. **Take stock of the social media and business intelligence companies** that are operating in Mexico and may have the necessary data. Facebook, Google Hangouts, Instagram, Twitter, LinkedIn, and Bloomberg all operate in Mexico and have been used by analogue governmental agencies for detecting misconduct. In analogous situations, the Australian Tax Office uses social media to spot tax evasion, avoidance, and fraud, while the Office of Community Oriented Policing Services in the US published a report suggesting tactics through which law enforcement can make use of social media to prevent and investigate crimes. But social media companies are not the only sources. For example, Ferrari proved to be an invaluable source of data about car purchasers in Mexico.

2. **Identify the motivations that data providers could have to join the data collaborative.** There are many ways to probe the willingness of corporations to join a collaborative: interviewing their representatives, organizing consultation series, conducting desk review, sending out a questionnaire, etc. We recommend conducting a desk review[21], as many corporations have already expressed a position towards working with governments and law enforcement to fight crime. For instance, Facebook has in place an information request protocol and publishes yearly reports on how it replies to government requests. The same holds for Google and Twitter. Research conducted by the GovLab reveals several potential motivations:

    - **Research & Insights:** sharing data may generate new insights on how criminals behave online and may enable companies to identify new niches for activity - e.g., BlueJay was a low cost Law Enforcement Twitter Crime Scanner tool. The tool had the potential to be economically feasible until Twitter tapped into this line of business and, on its side, law enforcement chose non-intermediated access to the data.
    - **Responsibility and corporate philanthropy:** sharing data may be considered a sign of corporate social responsibility or philanthropy. Countering crime can be seen as a public good to which corporations may contribute by sharing their data with the relevant authorities.
    - **Reputation and public relations:** sharing data for public good may enhance a firm's corporate image and reputation.
    - **Revenue generation:** companies may charge the government for sharing their data - by replacing services such as BlueJay.

3. **Identify the most promising data providers.** We recommend starting with the least risk-averse companies in order to reap the highest benefits at the lowest costs. To this end, we recommend focusing on Twitter data[22], as most Twitter account information is public: profile name, pictures, and Tweets. Although Twitter's policy on its API usage has become more conservative with respect to surveillance, Twitter is still working together with law enforcement and government authorities to counter crime.

---

[21] Helpful resources on how to engage in a desk review can be found here.

[22] Twitter is a real-time global information network that lets users create and share ideas and information instantly through 140-character messages. Twitter does not require the use of a real identifier, but users can provide one. Next to tweets, account holders can also share their bio, personal website and geolocation.

Why: Creating an inventory of suppliers and their motivations to join a data collaborative helps keep track of the existing opportunities and keep the collaborative successful, safe, and flexible.

CORE ACTION 4

SELECT THE OPTIMAL DESIGN FOR THE DATA COLLABORATIVE

Task description: The goal of this task is to select the optimal design for the data collaborative, considering the data providers selected during Core Task 3.

Who: The Coordinating Committee must select a design for the data collaborative that is informed by the data providers identified earlier.

How: There are many forms in which a data collaborative can be organized, each of which is better positioned for a certain problem or date type. Research conducted at the GovLab suggests choosing between the following:

- **Trusted Intermediary** - companies share data with a limited number of known partners. (For instance, anonymized mobile phone data is shared with researchers to create applications for social impact through the Dalberg Data Insights initiative.)
- **Prizes or Challenge** - companies launch a competition, sharing data with qualified applicants and providing incentives to compete. (For instance, companies and researchers post data on Kaggle to allow statisticians and data miners compete to produce the best predictive and analytical models.)
- **Research Partnership** - companies share data with universities and other research organizations. (For instance, Twitter shared data with MIT's Media Lab to provide insights and research into a number of fields.)
- **Intelligence Products** - companies share data that provides general insight into market conditions, customer demographic information, or other broad trends.
- **Application Programming Interfaces** (APIs) - companies allow developers and others to access data for testing, product development, and data analytics. (For instance, Facebook's Open Graph Search allows users to mine social graphs for data from search queries, "likes," and multimedia to improve product marketing.)
- **Corporate Data Cooperatives or Pooling** - various data providers group together to create "collaborative databases" with shared data resources. (For instance, the US National Institutes of Health has created the Accelerating Medicines Partnership to organize data pooling among the world's largest biopharmaceutical companies to better fight diseases.)

We suggest that the ideal type of a collaboration with Twitter would be use of Twitter's Public API (Application Programming Interface), which allows developers to retrieve data for analysis and build applications that reveal incidences of corruption in public procurements or public service delivery. The API is the best manner in which the Mexican Government can cast a wide net, in ways similar to what other law enforcement authorities have done. Where additional data or assistance is required, Twitter's Chief Media Scientist Professor Deb Roy, runs a lab at MIT that can provide additional advice.

Why: A data collaborative introduces a number of risks and challenges for the stakeholders. Using an API design to use Twitter data for the purpose of measuring corruption in public procurement reduces the risks and ethical concerns of the data collaborative.[23]

CORE ACTION 5

DEFINE THE OPERATIONAL ASPECTS OF THE IMPLEMENTATION

Task description: Armed with a clear understanding of the scope and value of the data collaborative, the data required, the data providers willing to share it, and the conditions under which data will be shared, this task puts the implementation in place.

Who: The Technical Secretary of the National Anticorruption System must define and ensure the execution of the operational aspects of the data collaborative in accordance with the decisions of the Coordinating Committee.

How: The Technical Secretary must create the infrastructure needed to download, store, clean, integrate, and analyze the new data. To this end, the following steps are required (based on the Twitter example):

1. Create a Twitter developer account and connect to a Twitter API (or partner with an organization with access).
2. Agree on a data plan with Twitter and, if necessary, purchase access to its Historical PowerTrack.
3. Apply sentiment analysis techniques to the tweets provided by the Historical PowerTrack in order to identify incriminating tweets.
4. Store, order, and integrate incriminating tweets on the Mexican Open Data Portal.
5. Allow analysts to analyze the tweets, visualize them, and conduct reverse geo-coding in order to identify the Tweets that provide strong suspicions that corruption has taken place in public procurement.
6. Allow analysts to use the information contained in the tweets as well as metadata about the tweets to estimate the incidence of corruption or collusion in public procurement and their attendant costs.

Why: Defining the operational aspects of the implementation ensures that the data collaborative is correctly and successfully launched. Research conducted at the GovLab suggests that it is crucial to formulate the steps of the project, from data sharing to measuring corruption in public procurement, to transform this idea into reality.

Estimated Timing and Costs of These Actions: Using Twitter data to measure corruption in public procurements can be much less expensive than creating the mechanisms to generate and collect this data. Nevertheless, there are human and capital costs of setting-up a data collaborative, running it, and extracting the content relevant to the analysis.

The formulation of the goal and value proposition of the data collaborative involves members of the Coordinating Committee and can be achieved through consultation in 2 - 4 weeks, depending on the availability of the members and the frequency of the meetings. Similarly, given the high level of expertise

[23] A more ample discussion of the risks involved in setting-up a data collaborative is presented here.

concentrated in the Coordinating Committee, identifying the relevant data to be gathered through the collaborative can be achieved through consultation in 2-4 weeks. Moreover, conducting the groundwork for the selection of the most promising data providers and the design of the data collaborative can be achieved by a social media analyst hired full time for a period of 2-4 weeks. According to Andrew Young, Knowledge Director at the Governance Lab, implementing the initiative could take 6 months – 1 year. The time costs are calculated assuming a team of 2 data scientists employed full time. Additionally, there are infrastructure costs such as the purchase of additional software (e.g., GNIP, which involves corporate fees) that supports this type of targeted scraping and server space.

What are the risks?

There are challenges to be overcome in building a data collaborative. In particular, we advise addressing the following:

- **The Twitter Terms of Service could be violated.** The Coordinating Committee should ensure that Twitter is in agreement with the ways in which its API will be used and to what purpose, in order to protect the credibility (and prevent liability) of the national authorities in charge of constructing the data collaborative.
- **Tweets may be "read out of context."** We advise the Technical Secretary to ensure that tweets are analyzed in a rigorous academic way and that relevant experts are consulted to develop the methodology for analysis.
- **State surveillance.** The Coordinating Committee must ensure that the tweets will be used exclusively to measure corruption in public procurements, not for surveillance.
- **Evolve as criminals always adjust.** Twitter users that tweet about their ill-gotten gains or crimes of corruption may refrain from using social media or may choose to do so under pseudonym. Using a pseudonym is easy and withdrawing from social media can be done quietly with few ill effects. There could be a learning effect that makes social network analyses no longer reflect true behavior, but a combination of true behavior and behavior 'according to expectations.' The Coordinating Committee should consequently monitor risks and mitigation strategies on a constant basis and remain engaged in a conversation on this matter with Twitter.

2. Use MEMEX open-source software for advanced online search to improve the extraction of information relevant to uncovering hidden corruption in public procurement

What needs to be done and why?

MEMEX offers a set of tools that, when combined with open data, can help to reveal patterns of corruption. Measurements of corruption start from a "suspicion of corruption." This suspicion can take the shape of a **perception of corruption** (e.g., the Corruption Perception Index) and can be informed by a **personal experience** (e.g., victimization surveys, WEF surveys on corruption). Aggregating these suspicions or experience reports forms most of the traditional estimates of corruption in Mexico.[24] Alternatively, suspicion

[24] In 2016, Mexico's National Institute of Statistics and Geography (INEGI) conducted a survey revealing that more than half of the surveyed Mexican companies reported paying a bribe to expedite business procedures. On the basis of this survey, INEGI estimated that the cost of corruption in 2016 exceeded $80 million.

can be derived by **applying state of the art algorithms to open contracting data to look for anomalies** - i.e., evidence of behavior that is out of the ordinary. These anomalies can reveal themselves as very simple patterns in the data (such as too many teachers having the same date of birth[25]) - or as more complex changes in the patterns of the distribution of the data (such as when the distribution of the first digits of national statistics on money laundering unexpectedly do not follow Benford's law[26]).

Owing largely to the availability of big data, the algorithmic way of raising suspicions is becoming more popular and more widespread. It can uncover leads complementary to those generated by traditional means at a comparably lower cost and in a more precise way. According to Bart Baesens - scientist at KU Leuven - *"Although classical [...] approaches, as discussed before, are still in widespread use and definitely represent a good starting point [...] a shift is taking place toward data-driven or statistically based fraud-detection methodologies."*[27]

The difficult part of extracting suspicion from statistical analysis is understanding the meaning of anomalies. If not placed in the proper context, these outliers are not likely to lead to an investigation of corruption and, therefore, using them for the purpose of measuring corruption would only add noise. Developing the context is a resource-intensive action that requires understanding, among others, the motivation behind the outlier (e.g., does the selection procedure for teachers require them to be born on the same date?) and of the individuals that have generated the outlier (e.g., it may be normal for private schools to pay their top teachers more than the salary of the President of Mexico). The context informs the development of the suspicion, i.e., the further investigation of the outlier to detect the predicate offence.

In order for analysts to add context in a fast and reliable way, we suggest using the MEMEX tools.

MEMEX is a research program initiated by the US Defense Advanced Research Projects Agency (DARPA), whose goal is to create software that can better search the World Wide Web. The MEMEX research community (MEMEX is actually a series of grants to researchers working on the common problem of developing better search tools) creates software to *"improve content discovery, information extraction, information retrieval,* and *user collaboration."*[28] MEMEX tools designed by computer scientists at New York University, Arizona State and elsewhere have been successfully used by law enforcement to track human traffickers and to provide evidence for their crimes by focusing on the temporary ads that were used to recruit and sell the victims. MEMEX's success has made it an integral part of every recent human trafficking case conducted by the New York County District Attorney's Office.[29]

---

25 Mejora tu Escuela, created by El Instituto Mexicano para la Competitividad (IMCO) using open government data, created the ecosystem and environment in which open government data would be scrutinized, allowing for the discovery that 1512 teachers on the payroll all had the same birthdays and all earned higher salaries than the President of Mexico.

26 Research shows that countries that faced incentives and opportunities to misinform the international community about their efforts to combat money laundering may have manipulated their statistics in order to conform to the pressures of the international community.

27 Baesens, Bart, Veronique Van Vlasselaer, and Wouter Verbeke. (2015) Fraud analytics using descriptive, predictive, and social network techniques: a guide to data science for fraud detection. John Wiley & Sons, p. 17.

28 Shen, Wade. "Memex" Defence Advanced Research Projects Agency, available here.

29 Greenemeier, Larry. "Human Traffickers Caught on Hidden Internet" Scientific American (Feb. 2015) available here.

SUGGESTED PLAN

How can it be implemented?

CORE ACTION 1

DESIGN A STRATEGY TO DETECT OUTLIERS IN PUBLIC PROCUREMENTS

Task description: The goal of this task is to design a strategy to generate suspicions by applying algorithms to open contracting data.

Who: The Secretaría de la Función Pública should design a strategy to detect outliers in data on public procurements.

How: There are many algorithms that can be used to generate suspicion of corruption using data on public procurements - e.g., red flags, Benford's law, benchmarking. We suggest conducting a benchmarking study to assess the performance of public procurement and to set a baseline from which to measure improvements.

In order to design a benchmarking study, the Secretaría de la Función Pública needs to:

1. Decide on the scope of the comparison

a. **External benchmarking** - compare public procurement data in Mexico with public procurement data of another country where corruption is less pervasive.
b. **Internal benchmarking** - compare public procurement data across the different administrative units of Mexico (e.g., data from Mexico City to data from Guadalajara), assuming that one administrative unit is less corrupt.

2. Decide what the benchmarks are

a. **Select the country/sector to use as benchmark**

Consult with experts to see what is the appropriate benchmark. Consideration should be given to the following questions: (1) *Which indicators of corruption must be used in the selection of the benchmark (e.g., countries with the lowest CPI score, sectors considered least corrupt in victimization studies)?; and* (2) *What contextual factors must be kept "equal" to ensure a meaningful comparison (e.g., procurement selection procedures, safety requirements for publicly procured goods, urgency with which the public acquisition must be made)?*

b. **Select the right benchmark**

Defining the "gold standard" is crucial to the identification of outliers that are meaningful in the uncovering of corruption in public procurements. If the bar is set too high, too many outliers will be created, thereby diluting the resources that can be allocated to analyze them. If the bar is set too low, too few outliers will be observed and a great deal of corrupt practices will go unabated. Consequently, the choice of "golden standard" significantly impacts the measurements of corruption in public procurement.

c. **Ensure the comparability of data**

Using standardized data is at the core of any benchmarking exercise. In order to achieve comparability, the World Bank's Benchmarking Public Procurement 2017 study employs case studies and hypothetical scenarios to test the differences in legal and regulatory frameworks across countries.

3. Decide what constitutes an outlier

It is important to explain what constitutes an outlier for the purpose of predicting corruption. To this end, the Secretaría de la Función Pública should consult with national financial experts to understand the acceptable risk premiums. This is a difficult decision that most commercial benchmarking services refrain from making. For example, GFTrade - a tool constructed by Global Financial Integrity - enables customs officials to determine whether goods are priced outside the usual range for comparable products. While the tool reports the difference in prices, it makes no judgement on what is "too far off price." In 2013 the European Commission funded a study to develop a benchmarking methodology for the costs of publicly procured contracts. However, this study also refrained from drawing clear thresholds to distinguish outliers.

Why: While benchmarking is a technique that has long been employed to track illicit flows of money and to follow the tracks of white-collar criminals[30], it has only recently been used to measure corruption and its associated costs. Despite its novelty, the method is easy to conduct from a statistical point of view. Outliers are easy to visualize and attribute to an individual (the company that provides the highest priced public goods[31]), to a geographical location (the country with the most expensive publicly procured goods[32]), or to a moment in time (the years before or after a legislative change when publicly procured goods were very expensive[33]).

Required Resources and Costs: According to Dr. Ferwerda, a researcher at Utrecht University who participated in the study "Identifying and Reducing Corruption in Public Procurement in the EU," this task can be completed by 3 full-time researchers within 8 weeks.

CORE ACTION 2
CONDUCT A BENCHMARKING PILOT

Task description: The goal of this task is to conduct a comparative analysis of public procurements in Mexico for the purpose of discovering the outliers.

---

[30] E.g. The FATF recommends that all transactions above the equivalent of $15,000 be reported to the Financial Intelligence Unit, to be further analyzed, with the belief that some of them would be useful leads for uncovering money laundering.

[31] After facing allegations of collusion and bribery financed by imposing higher than normal prices for water supply in France's Marseille, Veolia is facing similar charges in Romania, where the price paid by the citizens of Bucharest is alleged to have increased by 125% from 2008 - 2015.

[32] In Romania highway construction costs are three times more expensive than in Bulgaria, under similar topographic conditions.

[33] Wiehan reports that the construction cost per kilometer for Milan's subway dropped $130 million from 1991 to 1995 as a result of a series of anti-corruption investigations.

Who: The Secretaría de la Función Pública should conduct a benchmarking pilot.

How: There are many aspects on which public procurements that can be compared, such as the procurement process, the complaint review mechanism, the cost of publicly procured goods, etc. Experts suggested focusing on the cost of publicly procured goods.[34] To this end, the Secretaría de la Función Pública should:

1. Construct a database of Mexican public procurements and their costs as well as a database of public procurements and their costs from the benchmarked country. In doing so, the strategy designed during Core Action 1 should be followed.
2. Employ adequate statistical tools[35] to separate the outliers. Having considered the appropriate risk premiums that businesses supplying public goods should be able to charge without raising "suspicions," the focus should fall on identifying prices that exceed this threshold.

Why: A benchmarking pilot delivers a first set of anomalous prices that can be further analyzed for the purpose of uncovering corruption. The pilot will expose the difficulties of conducting a benchmark and offer an initial set of solutions to handle them.

Costs: According to Dr. Ferwerda, the human costs of a benchmarking pilot depend on the quality of the available data and on the desired sample size. Constructing a sample of 200 public procurements with granular cost data may take 4 full-time researchers 10 weeks to complete. Where this data is already open, available, and machine-readable, however, this significantly reduces the required effort. Collaboration with the **Open Contracting Partnership and its staff and partners might help to accelerate progress toward this goal.**

CORE ACTION 3
CUSTOMIZE THE MEMEX TOOLKIT

Task description: The goal of this task is to customize the MEMEX tools to discover relevant context for the outliers generated through the benchmarking study.

Who: The Secretaría de la Función Pública should support the customization of the MEMEX tools for the purpose of creating a meaningful context to the outliers generated during Core Action 2.

OUTLIER: A company run by Vladimir Kudyakov is winning public contracts with bids that are 30% higher than the benchmark.

MEMEX: MEMEX is made up of several powerful search tools that are especially designed for creating context around a suspicion. While the page-indexing algorithms that Larry Page and Sergey Brin have developed have brought them fame and fortune, they are unfit for advanced data discovery necessary to counter corruption in public procurement.[36] Alternatively, the MEMEX explorer allows for queries to be performed on a regular basis, recognizes spelling errors, saves **the** query results, organizes the results

---

34 This suggestion was made by Wade Shen.

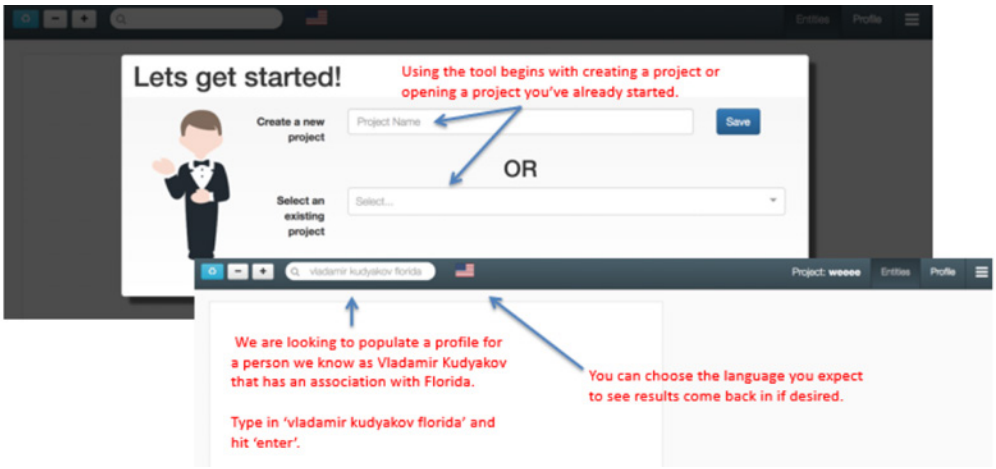
35 Useful resources to understand the standard unit cost approach can be found here and here

36 Commercial search engines are able to capture only **a small fraction** of the websites that are available via the HTTP standard internet protocol, as they face a tradeoff between in-depth indexing (i.e., capturing information within a webpage) and broad indexing (i.e., indexing information of as many different pages as possible).

logically, and includes information from the Deep Web—where content is not indexed by standard commercial search engines and is therefore often ignored—and meta-data such as content shared across web pages.

The “Butler” tool is the MEMEX interface for information retrieval, pictured below.

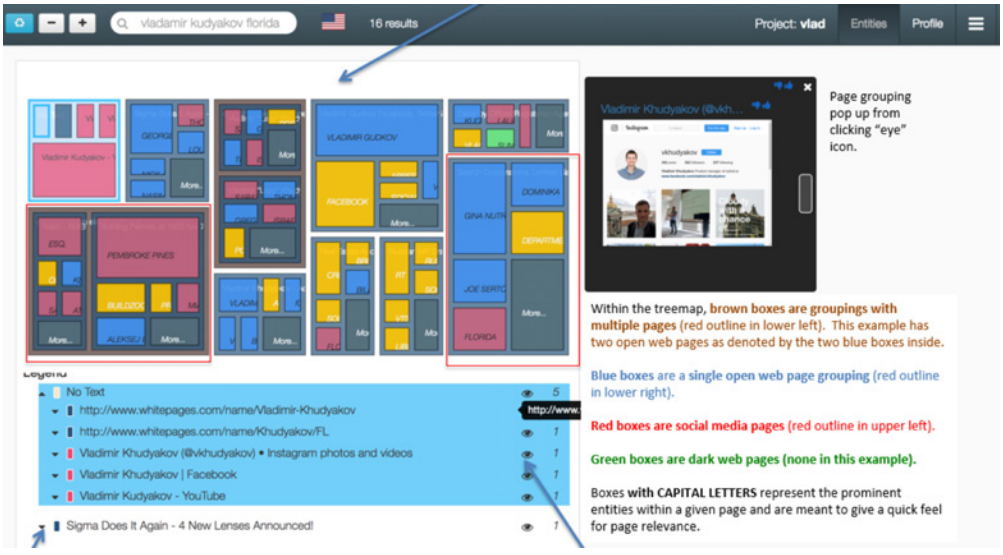
INPUTTING OUTLIERS IN BUTLER



After querying the Deep Web, the information retrieved is organized as shown in Figure 1.2. Butler organizes this information by source:

1. Information retrieved from the Open Web (the web that is indexed by most commercial search engines) is colored in blue in the left hand side matrix
2. Information retrieved from social media websites is colored in red
3. Information retrieved from the Dark Web is colored in green
4. Images are recorded separately (on the right hand side of Figure 1.2)

BUTLER DASHBOARD



CONTEXT: The analyst can use this dashboard to record her preferences (e.g., removing entries that are clear false leads, upgrading relevant information through "likes"). This allows the MEMEX tool to learn and to improve. As an example, the analyst can see whether Vladimir has purchased a house with a total value that is similar to excessive markups he charged for supplying the public good and then whether that house was subsequently sold to a public servant through an intermediary (such as a Facebook-friend of Vladimir) who played a role in assigning the contract to Vladimir. The analyst can then search for more information on the person who intermediated the sale of the house to see which other deals he has intermediated in a similar fashion. The analyst stops once enough evidence or leads have been gathered or when it appears that the system has failed to find enough leads.

Once this suspicion is confirmed, the extra markup can count towards the costs of corruption and this can count as an incidence of corruption in public procurements.

How: The Secretaría de la Función Pública should launch a **crowdsourcing initiative** to allow the search algorithms of MEMEX to be calibrated to the Mexican context and to the Spanish language. Open this data annotation task to the public.

Why: The query algorithms of the MEMEX tools are calibrated in English. Since most of the text required for this exercise is in Spanish, the algorithms should successfully find content and retrieve information from texts written in Spanish. Without calibration, the tools will yield many irrelevant results. Crowdsourcing is a good way to achieve this task because text annotation is a low skill, high labor task, the same criteria that led, for example, scientists at NASA to turned to "citizen scientists" (volunteer hobbyists, amateur science buffs, and space enthusiasts) to classify the images the Hubble Space Telescope recorded of the Milky Way and other galaxies according to their shape: elliptical, spiral, lenticular, irregular.[37]

Costs: According to Wade Shen at DARPA, assuming the worst-case scenario (the algorithms need to be trained from scratch), the task would require the equivalent of 1.5 person-years, a task which, when crowdsourced, could be accomplished in a matter of a few weeks.

3. Build a predictive analytics tool capable of revealing incidents and patterns of corruption and of estimating corruption and its costs

What needs to be done and why?

Mexico opens large amounts of data on public procurements.[38] As mentioned earlier, an Organization for Economic Co-operation and Development (OECD) assessment places Mexico above the OECD average but reveals that **Mexico's open data is underutilized.** This has significant negative consequences:

1. Federal and state **resources are spent on collecting and publishing data** that is not utilized. These funds are effectively mismanaged.

---

[37] Noveck, Beth Simone. "**Smart Citizens, Smarter State: The Technologies of Expertise and the Future of Governing**" (pp. 4-5). Harvard University Press. Kindle Edition.

[38] The Mexican Government created a central open data portal that accumulates data from public sector institutions, both at the central and the local level. The portal is backed by a mobile application that aims to support the usage of the portal by citizens.

2. **Collected data** is not **put to the test**.
3. **Relevant data** that should be collected **is missing** and corruption goes unabated.

Predictive analytics refers to the use of statistical algorithms to identify the probability of future outcomes based on the available data. Predictive analytics rely on capturing relationships between explanatory variables (red flags) and the past occurrences of the predicted variables, and exploiting them.

The goal is to understand what has happened in the past and to go beyond that and provide a prediction of what will happen in the future. Predictive analytics have been used thoroughly to help governments improve safety in the workplace and police conduct[39] and to predict which buildings should be inspected to prevent fires. In the European Union, analytics were successfully used in 2013 to measure corruption in public procurement and its costs and to predict which tenders are likely to be corrupt. Recently, the NGO *Mexicanos contra la Corrupción* (Mexicans against Corruption) used its own red flagging technique to identify the Attorney General of Mexico as himself corrupt. The technique involved using vehicle registration data to spot Ferraris registered to suspicious addresses, namely homes that were not of comparable value to the cars. In so doing, they uncovered a massive corruption scandal.[40]

Similarly, the Open Contracting Partnership together with Development Gateway have applied red flags to open contracting data in an effort to develop a functional tool that governmental agencies, researchers, and the public could use to monitor the integrity of public acquisitions. This builds on earlier experience with Brazil's Public Expenditure Observatory, which began using data science and red flags to monitor the use of public funds, detect anomalies and cases of fraud, and generate savings beginning in 2008-9.[41] Brazil's red-flagging system, which cost a little more than half a million dollars to build, is reported to generate 60,000 annual instances of warnings from the computer-assisted audit tracks used by the Office of the Comptroller General of the Union to identify possible procurement irregularities.[42]

Predictive analytics can help exploit the Mexican open data in order to generate more granular and timely measures of corruption in public procurement.

---

[39] Kane, R. J., & White, M. D. (2009). **Bad cops: A study of career-ending misconduct among New York City police officers**. Criminology & Public Policy, 8, 737–769.

[40] Claudio Gonzalez Caraza, **La Modesta Privada de los Ferraris, Mexicanos contra la CorrupciÓn y la Impunidad**, September 4, 2017.

[41] See Transparency Fund Annual Report 2015, p. 13-14. Brazil's Public Expenditure Observatory has adopted a red flag analysis program. It has detected problems in more than 7,500 procurement processes in 2015 alone, for a value of US$104 million. "One of the filters applied by the ODP helps detect cases in which high-value procurements are fractioned into smaller ones in order to stay below thresholds for competitive processes. This analysis was applied in Santa Catarina, one of two Brazilian states that is replicating the federal ODP model. More than 2,300 cases of fractioning were found, for a total of US$11.2 million. In the state of Bahia, the local ODP pilot detected several cases in which procurement thresholds had not been respected, resulting in excessive costs for the state in the order of US$7 million." The experience of Brazil's ODP is currently being replicated in Colombia, where the project Institutional Strengthening of the Comptroller General of the Republic to Improve Fiscal Control (CO-T1389) was approved in 2015." See also Graft, A. et al. (2016). Brazil's Open Budget Transparency Portal: Making Public How Public Money is Spent, *GovLab report,* available here.

[42] Public Spending Observatory in Brazil, available at https://www.oecd.org/governance/procurement/toolbox/search/public-spending-observatory-brazil.pdf.

However, it is worth noting, care must be taken to assess the availability of data for the project. For example, Development Gateway, an international nonprofit social enterprise aimed at reducing poverty by improving aid effectiveness, governance, and transparency through IT, analyzed open contracting data from six African countries. In spite of the large volumes of data, errors in collection, coding and publishing, badly aggregated data, and missing data made drawing results a difficult exercise. A study on the costs and effectiveness of public procurement in Europe found that missing data was conspicuously absent from the European Tender Electronic Database and that cost measurements had to rely on advanced statistical techniques. With this in mind, the Government of Mexico, as well as other governments, face the urgent task of testing the quality of data they have already collected.

Assuming the availability of comprehensive data (although there are other methods that can be applied to sampled or incomplete data), a four-step approach needs to be followed to establish a predictive analytics program:

1. An inventory of red flags the literature considers relevant;
2. An inventory of relevant data sources, their coverage, and quality;
3. The testing of the relevance and applicability of the red flags; and
4. The estimation of the incidence of corruption in public procurement in real time.

How can it be implemented?

ENABLING CONDITIONS

Mexican public agencies including the Superior Audit Office and the Secretaría de la Función Pública must support the construction of a predictive analytics tool to be applied on Mexican open data for the purpose of revealing the incidence and costs of corruption in public procurement. Additionally, experts from the Special Anti-corruption Prosecutor, the representative of the Federal Administrative Court, the representative of the Federal Judiciary, and the Technical Secretary should aid in the construction of the tool.

SUGGESTED PLAN

In order to build a robust predictive analytics tool, the indicators of corruption relevant to the Mexican public procurement sector must be applied to the corresponding data in a scientifically sound manner. Attention must be paid to ensure that the tool is able to correctly distinguish risky situations, individuals, or procurements in order to ensure its wide adoption. Consultation with experts at the Inter-American Development Bank and public officials with experience in Brazil and Colombia is recommended.

CORE ACTION 1

COMPILE A WELL-STRUCTURED LIST OF RELEVANT RED FLAGS

Task description: The goal of this task is the creation of a comprehensive list of red flags, a direct or indirect indication of potential corruption, collusion, fraud or other illicit behavior that could be applied to reveal incidences of corruption in public procurement in Mexico. Although many red flags are in common across jurisdictions, defining what constitutes relevant red flags is very context specific.

Who: The Coordinating Committee should compile the list by consulting international and national experts.

How: The list can be generated by:

A. **Identifying** the most commonly used red flags for corruption in public procurement.[43] The literature on "red flags" used to signal corruption in public procurements is based on generic case studies and investigations, expert interviews, and academic works. For example, the OCP and Data Gateway have sifted through 150 known red flags to settle on 36 red flag indicators that are easily tracked when data uses the Open Contracting Data Standard and an additional 75 flags that could be calculated with minor additions to the OCDS schema.

B. Organizing the identified red flags according to **the different stages of the procurement process,** as each stage contains a specific set of vulnerabilities:

    A. The decision to contract
    B. The definition of contract characteristics
    C. The contracting process
    D. The contract award
    E. The contract implementation and monitoring

C. Consulting with Mexican experts (in addition to learning from Brazil, Colombia, and elsewhere) to **identify the red flags that are relevant in the Mexican context**. Experts should be knowledgeable on the following topics:

    A. The public procurement process in Mexico - in order to eliminate the red flags that do not apply due to the rules of engagement that are in place in Mexico
    B. White-collar crime modi operandi - in order to select the red flags that fit into the operational patterns of criminal networks that are observed in Mexico
    C. Social psychology - in order the identify the red flags that are likely to reveal the socio-cultural customs of Mexico as well as patterns of corruption

Why: A comprehensive list of red flags is needed for the construction of a predictive analytical tool. Taking a funnel approach to the creation of the list allows the consultation to consider as many indicators as possible. In general, having numerous sources of suspicion is considered beneficial in the identification of instances of white-collar crime.[44] Nevertheless, in order to reduce the likelihood of false positives, it is important to select meaningful red flags - i.e., indicators that could identify corrupt procurements in the framework of the Mexican procurement process and in the context of local criminal patterns. Moreover, it is

---

43 See Open Contracting Partnership and Development Gateway, RED FLAGS for integrity: Giving the green light to open data solutions, available at https://www.open-contracting.org/wp-content/uploads/2016/11/OCP2016-Red-flags-for-integrityshared.pdf. Helpful resources on how to engage in a desk review can be found here.

44 Mexico's Federal Law to Prevent and Identify Transactions Involving Resources Illegally Obtained obliges many financial and non-financial professions (banks, insurances, exchange offices, lawyers, notaries, etc.) to report suspicion of money laundering or the predicate offence.

important to understand which of the red flags may capture socio-cultural customs that are not crime-related, for the purpose of operationalizing and calibrating them (see Core Action 3).

Costs: According to Dr. Ferwerda, researcher at Utrecht University, who participated in the study "Identifying and Reducing Corruption in Public Procurement in the EU", the construction of a list of relevant red flags for Mexico involves one full-time researcher working for 2-3 weeks.

CORE ACTION 2

GENERATE AN INVENTORY OF RELEVANT DATA SOURCES AND THEIR CHARACTERISTICS

Task description: The goal of this task is to identify the data needed to test the red flags that were identified during Core Action 1, the location of this data, whether it is openly available, and whether it is of sufficient quality.
Who: The Coordinating Committee should generate an inventory of the relevant data sources.
How: In order to generate a meaningful inventory of data sources, the following steps should be taken:

1. **Identify the currently available open data** that is created in the public procurement context. Start from the platforms that already centralize this data, such as the Portal of Mexico City, Mexico's Open Data Portal, information centralized on the National Digital Platform, etc. The inventory[45] should clearly track:
    - What the data means
    - How and when it was created
    - Who owns the data
    - Who has access, use, and editing rights
    - Who is responsible for managing and securing the data
    - Whether the data can be shared and under which conditions

2. Kaufmann and Kraay argue that new indicators of corruption can have a meaningful impact only when **attention is paid to the assumptions and error terms** that underpin them. Most of the error terms are due to the poor quality of the used data. It is therefore important to describe **the quality of the data[46]** along the following dimensions:
    - **Completeness** - the fraction of the data against the potential "100% complete"
    - **Uniqueness** - how often a data point is recorded only once
    - **Timeliness** - the degree to which data represent reality from the required point in time
    - **Validity** - that the units, format, and other characteristics of the data are recorded in conformity to its definition
    - **Accuracy** - the degree to which data correctly describes reality
    - **Consistency** - that there is no difference when comparing two or more representations of the same event

[45] A useful tool for compiling the data inventory can be found here.

[46] A useful tool for compiling a data quality assessment can be found here.

Why: Effective research data management is necessary for conducting responsible research. Compiling an inventory enables data sharing, future reuse, and archiving. The inventory can be used as a data management tool and can help inform decisions regarding future updates.
Costs: The data inventory can be done by one full-time researcher within 2-3 weeks.

CORE ACTION 3
DESIGN THE PREDICTIVE ANALYTICS TOOL

Task description: The goal of this task is to create the roadmap for estimating corruption using the red flags and the data identified earlier.
Who: The Coordinating Committee should design a predictive analytics tool customized to detect corruption in public procurements in Mexico.
How: In order to build a robust analytical tool, the following steps should be taken:

1. Transform the data to allow for the application of the red flags.

Data should be transformed so that it can be used to test the incidence of red flags. (e.g., if a red flag is *"high inertia of the commission evaluating the bids,"* data must be transformed into a binary format where "1" illustrates a team that does not and "0" marks a team that often changes).

It is crucial to note and disclose all the assumptions and the transformations made to the data. Following the earlier given example, it is important to disclose how this red flag is operationalized for the purpose of predicting corruption (e.g., through the ratio of "new name" to "recurring names" in the commission; by counting how often a member has been sitting in a commission; or by counting how often the commission has had the same exact composition). These metrics are not obvious and the ad-hoc choice of a metric can greatly impact the accuracy of the predictive tool as well as the outcome of the prediction. Consequently, as recommended by Kaufmann and Kraay, all metrics and transformations must be disclosed and the disclosure must explain the ultimate choices.

2. Construct control and treatment groups.

It is important to test the predictive powers of the red flags. Academic rigor has often been neglected in applying red flags to open data, with potentially large and negative consequences. This is a particularly pressing problem in criminal research. To this end, we suggest taking the following steps:

A. **Select from the data a group of procurements where corruption was known to have happened**. Call this group the (corruption) *treatment group*, following terminology commonly used in analogous medical trials. The EU conducted a study entitled *"Identifying and Reducing Corruption in Public Procurement in the EU"* where it grouped together 96 procurements across 8 EU Member States where corruption was identified. These cases were identified as corrupt by national and international experts.

B. **Select an equally large group of procurements where no corruption was found to take place.** Call this group the *control group*. The EU study grouped together 96 procurements across 8 EU Member States where corruption did not take place. National and international

experts were asked to review these cases in order to ensure that no evidence or suspicion of corruption could be found.

It is important to create sufficiently large groups to allow for a meaningful comparison and sufficiently similar groups so that differences can be attributed only to corruption.

3. Test the predictive power of the red flags

Having correctly constructed a sample of corrupt and non-corrupt cases, the predictive power of the red flags can be put to the test through the use of standard econometric techniques[47] such as traditional OLS, prohibit regressions, and forecasts used in the EU study. A red flag that predicts corruption should reveal itself only in the procurements in the treatment group and not in procurements in the control group. This method also reveals how relevant each red flag is to correctly predicting corruption.

4. Identify the most powerful predictive red flags

The analysis above allows red flags to be classified into three groups:

A. **No-impact red flags.** They cannot distinguish among corrupt and non-corrupt procurements. These red flags should be discarded.
B. **Low-impact red flags.** They can distinguish among corrupt and non-corrupt procurements, but the strength of their presence on the final outcome of the procurement (i.e., whether the procurement is corrupt or not) is low. These red flags should be kept.
C. **High-impact red flags.** They can distinguish among corrupt and non-corrupt procurements, and the strength of their presence on the final outcome of the procurement is high. These red flags should be kept.

Knowing which red flags are able to distinguish corruption, the probability of corruption can be estimated by monitoring the incidence of red flags in future procurements, as well as past procurements where it is not clear whether corruption has taken place.

Why: For a predictive analytics tool to be trusted and broadly used, it must be sufficiently accurate. The accuracy of the predictive tool crucially depends on the selection of the indicators of corruption and on their ability to distinguish among procurements where corruption takes place and procurements where corruption does not take place. Research shows that few of the traditional red flags had these qualities.

Costs: According to Dr. Ferwerda, the construction of the control and treatment groups was the most time-consuming in the European case. The construction of groups of 200 or more data points requires 4 full time researchers working for approximately 3-4 months. In total, this task can be achieved by 5 full time researchers (4 researchers, and 1 senior researcher) that are employed full time for 4 months.

---

[47] A useful resource for understanding how to correctly conduct such econometric exercises is Jeffrey Wooldridge's book, available here.

CORE ACTION 4

BUILD THE PREDICTIVE ANALYTICS TOOL

Task description: The goal of this task is to construct a predictive analytics tool that applies the powerfully predictive red flags identified earlier to new data and allows for a real-time monitoring of the corruption risks in Mexico's procurements.

Who: The Coordinating Committee should arrange the construction of a predictive analytics tool that follows the design agreed upon during Core Action 3.

How: In building the tool, the Coordinating Committee has several options:

1. Partner with a university or a research institute with experience in constructing predictive analytics tools.

Academic and research organizations such as the Applied Mathematics School (EMAp FGV) together with the Law School (FGV Direito Rio) in Brazil and researchers at Utrecht University in the Netherlands have used predictive analytics to estimate the likelihood that public procurements involve corruption and to estimate the cost of corruption in public procurement. In this case, the Coordinating Committee needs to coordinate with the external researchers and agree on the terms of any grant as well as key milestones and deliverables to ensure timely completion of work.

2. Hire a data science company to develop a customized predictive analytics tool.

Companies such as SafetyNet develop customizable predictive analytics. This form of collaboration may require an open tender if the costs are above the direct award threshold. However, this affords the advantage of potentially faster delivery.

3. Develop the predictive analytics tool in-house.

This requires relying on in-house computer science knowledge or the hiring of a specialist that has the necessary technical background to customize a predictive analytics tool. Additionally, this requires providing the data scientist with the software (e.g., Stata v10+, SPSS) and the hardware (e.g., computers powerful enough to support the calculations) required to develop and test the tool. Although this approach eliminates the need to negotiate an agreement with an external party and can help build institutional knowledge and capacity for future projects, it relies heavily on a well-managed in-house capability.

A well-functioning predictive analytics tool should have the following:

1. A functional back-end

The back end should contain a **data repository** that can be manually populated with raw data. This process can be automatized at a later stage to allow for the automatic update of the database once data suppliers make the data available. It should also contain a **set of algorithms** that transform the raw data, apply the red flags, and store the results according to the design formulated during Core Action 3.

2. A user-friendly front-end

The front end should provide clear instructions as to which data to import, in which customized fields, and in which format. It should clearly describe which red flags that should be provided as well as the ways in which these red flags are operationalized. For example, SafetyNet, OCP and DataGateway, (the EU predictive tool), all have freely downloadable white papers explaining their methodology, error sources, and data sources. The front end should clearly display the results of the analysis and allow them to be used in visualizations (e.g., Corruption Risk Dashboard).

Why: A predictive analytics tool that can correctly distinguish corrupt procurements and that is constructed in a user-friendly way can be a useful tool for monitoring corruption in real time and for improving measurements of corruption in public procurement. Such a tool allows for more flexible measurements - i.e., depending on the input data, the tool can help measure corruption at a granular level, before and after an anti-corruption policy has been enacted, or in real-time.

Costs: According to Dr. Serban, a data scientist at Accenture, the basic version of such a predictive analytics tool requires 2 data scientists, 1 data architect, and 1 data engineer hired full-time for a period of 2 months. Additionally, budgets should include storage space and a domain. For comparison, the basic SafetyNET package that allows up to 10 users to run predictions costs $2400 a year.

What are the risks?

- **Badly calibrated predictive models.** Forecasting models are notoriously hard to calibrate. The Coordinating Committee should dedicate sufficient resources to the design of the predictive analytics tool and to its updating on a continuous basis.

- **Equating predictions with actual occurrence of corruption.** Although forecasts and estimates are useful for conducting targeted audits and verifications, it is important for the Coordinating Committee to discourage the use of predictions as pseudo-evidence by the public and the media.

ISSUE AREA 2

Tech-Enabled Strategies to Reduce Corruption in the Judiciary

A critical focus of Mexico's recent anti-corruption reforms is improving the judiciary, which 80% of Mexican respondents perceive to be corrupt or extremely corrupt.[48] When one considers that, according to CIDAC, a Mexican think tank, 96% of crimes went unpunished between 1996 and 2003, it is clear that the Mexican public needs better tools and tactics to strengthen both judicial independence and accountability, and provide the means for ongoing monitoring of judicial integrity.[49] Drawing on international experience, this implementation plan outlines two short-term initiatives designed to help Mexico advance transparency, increase openness and improve oversight. Over the next six months, Mexico should:

- Develop a national online "open courts" portal to publish data openly and online about judicial and court performance and thereby improve the effectiveness of the federal judiciary through comparative evaluation.
- Integrate public oversight into judicial selection and evaluation by disseminating data on judicial appointments and judges' assets, interests and networks and thereby reduce judicial impunity and the politicization of judicial selection.

While these initiatives are focused on distinct outcomes, the former concerned with improving performance of the federal courts and the latter focused on securing the independence of judges, both involve

---

48 International Bar Association & Basel Institute on Governance, "The International Bar Association's Judicial Integrity Initiative: Judicial Systems and Corruption" (May 2016), 52.

49 Michael Reid, Forgotten Continent: The Battle for Latin America's Soul (Yale University Press, 2007), 250.

publishing already-collected or easily obtained data on existing websites and securing public participation in the scrutiny of that data. Therefore, this memo outlines the common steps to take in service of either or both initiatives, which can be accomplished in the next six months and show results within a year. If successful, Mexico will catch up to other Latin American countries and demonstrate its commitment to open government and open data. These open justice webpages will allow Mexico to demonstrate progress against strategic goals and show how its policies and practices are contributing to greater efficiency and effectiveness of the courts.

I. PROJECT BACKGROUND

On June 20th, 2017, 21 experts from 9 countries joined officials representing the Mexican government, the Inter-American Development Bank, and members of the GovLab, in a two-hour online conversation to identify novel strategies that the Mexican government and civil society leaders can use to address corruption in the judiciary.

The conference resulted in seven concrete recommendations from which la Secretaría de la Función Pública selected two high-impact projects that form the basis of this implementation plan. They were selected based on a combination of their potential importance, efficacy, and feasibility to undertake in 2017.

This plan contains:

- An overview of the problem which these solutions address
- An outline of the solution proposed including
  - Overview and value proposition
  - Examples of similar initiatives around the world
  - Driving questions and considerations to define the solution at a high level
  - Steps and resources needed to implement the solution, including a staged roll-out plan that enables the project to be launched without waiting to have data from every court
  - Metrics to evaluate the success of implementation

The content of the document, including its recommendations, is the sole responsibility of the GovLab and does not represent the IDB's official position or view on this matter, or an endorsement of any individual or firm to perform activities related to the recommendations.

II. PROBLEM OVERVIEW

See full problem brief here.

Impunity in the Mexican judicial system is one of the most prominent symptoms of corruption. A study by the Inter-American Commission on Human Rights found that **98% of crimes in Mexico "fail to result [in] convictions."**[50] Although some of that number can be attributed to unfair arrests and innocent defendants, in many cases the failure to prosecute, especially public officials, is the result of corruption and judicial impunity. As one Mexican scholar characterizes it, in Mexico's judiciary, "impunity has become the rule and prosecution and punishment the exception."[51] The prevalence of corruption practices resulting in high impunity within Mexico's justice system has undermined the performance and fairness of the judiciary.

---

[50] Monica Ortiz Uribe, Mexico's Justice System Battles Its Own Reputation to Build Trust (June 12, 2016).

[51] Gabriel Ferreyra, The Michoacanazo: A Case-Study of Wrongdoing in the Mexican Federal Judiciary (2015), 29.

While the spectrum of corruption practices in the judiciary is broad and varies according to the level of court or tribunal, **bribery** is one of the most common forms. In a report published by the International Bar Association's Judicial Integrity Initiative (JII), 82% of survey respondents reported believing there is a high incidence of bribery in the Mexican judicial system, while **55% reported paying a bribe** to the judiciary.[52] While judges are often the recipients of bribes, JII also highlights that there are corruption networks in the judiciary, operating complex schemes to obtain bribes through judges, court staff, and other intermediaries. Additional common forms of judicial corruption include **lack of impartial appointment of judges to cases, undue political influence, misuse of funds, and impropriety in interactions between judges and intermediaries.**

As frequently emphasized, one of the core causes of corruption is the **lack of transparency in the judicial system**. In a workshop led by the GovLab in Mexico to define the problem of judicial corruption, experts noted how this lack of transparency comes from a number of sources. Although rulings in cases can be requested, judges have no mandate to release their decisions publicly. Additionally, while the federal judiciary publishes court statistics, the majority of states lack strong commitments to open data. Finally, there is a lack of news coverage from independent media. According to Leticia Bonifaz Alfonzo, Directora General de la Dirección General de Estudios, Promoción y Desarrollo de los Derechos Humanos, "In Mexico, strong investigative journalism in judicial matters is practically non-existent. For example, the media does not follow the appointment of judges or share the profiles or qualifications of candidates."[53]
This lack of transparency is directly related to the history of Mexico's judicial system, which formed as a common law "inquisitorial" system in which judges typically examine all evidence in a closed session and deliver their verdict in writing, often without detailed explanation.

While many countries follow an inquisitorial system, Mexico's system had become so corrupt that by 2008, "mistaken arrests, bungled investigations and confessions extracted under threats and torture have become common in Mexico."[54] That year, the government announced a constitutional overhaul to "replace its closed proceedings with [oral] public trials in which defendants are presumed innocent [and] legal authorities can be held more accountable."[55] Although this was seen as a landmark step to reducing opacity, in 2016 – when the reforms were due to be fully rolled out – only 24 of Mexico's 32 states had adopted the new system and CIDAC, an independent Mexican think tank, anticipated the country needed another 11 years for the reforms to be fully implemented.[56]
Aside from the lack of transparency in judicial processes, there is also a **lack of data about the workings of the system (i.e., how many complaints are filed, the time to disposition, final verdicts) that is needed both to monitor and improve efficiency and potentially to spot problems**. The full extent of judicial corruption in Mexico is unknown because it is extremely difficult to measure: no official or unofficial data exists on the topic, and any reported data would likely be understated as corrupt behaviors happen in

---

52 International Bar Association & Basel Institute on Governance, 52.

53 Smarter Crowdsourcing: Anti Corruption Problem Definition Workshop on Judicial Integrity, April 21st, 2017

54 Watson, Julie. "How Mexico Is Overhauling Its Legal System." SFGate, Associated Press, 14 Sept. 2008,

55 International Bar Association & Basel Institute on Governance, 52.

56 Michelle Mark, and Reuters. "Mexico Has Spent 8 Years Overhauling Its Dysfunctional Justice System, but It May Need 11 More to Fix the Mess." Business Insider, Business Insider, 7 May 2016,

secret.[57] Nonetheless, judicial corruption is widely understood to be a real and perceived problem, seriously undermining faith in public institutions. Survey responses from public officials in the Mexican Federal Judiciary, scholars, and attorneys suggest that up to 10% are actually corrupt,[58] and a 2014 Gallup poll indicates that 47% of Mexicans do not have confidence in Mexico's judicial system.[59]

III. SOLUTION

Value Proposition

The Consejo de la Judicatura Federal (Federal Judiciary Council) should publish federal courts data on judicial/court performance, judge backgrounds and interests, and judicial selection processes on its National Statistics website and appoint staff and outside experts to scrutinize the data and recommend improvements.

Mexico is part of a broader movement toward publishing data in formats that allow its reuse and that are machine-readable, also known as "Open Data." As former Chair of the Open Government Partnership and signatory to the Open Data Charter, Mexico has committed to "opening up government data . . . [to] allow innovation, justice, transparency, and prosperity to flourish, all while ensuring civic participation in public decisions and accountability for governments.[60] Research by the GovLab indicates that publishing information collected by government about its own workings has already helped the executive and legislative branches to reduce inefficiency, spot waste and fraud, and empower citizens to mobilize for change. The driving hypothesis underlying this project is that publishing easily accessible, open data on the judicial sector will have an analogous effect by increasing visibility of the information needed for the judiciary itself to prevent, identify, and reduce corruption and for the public to hold the judicial system accountable for such improvements.
There are two main ways open data published via open justice webpages on the National Statistics website can achieve these goals.

First, publishing comparative open data **provides greater insight to enable performance improvements**. For instance, in the healthcare sector, studies found that publishing comparative data on clinical outcomes in heart surgeries in the UK reduced the number of deaths by up to 1,000 annually.[61] Similar results have been seen in Germany and Sweden, where public reporting of health data has significantly improved a variety of health outcomes, including recovery following aortic aneurysm repair, and cardiology-related ailments. Similarly, publishing courts data, including performance data as well as judicial qualifications, exam results and interviews of judicial candidates during the selection process, may lead to a reduction in corrupt behavior, as courts and judges can no longer operate without some level of scrutiny, including by the court system itself. With open and comparative data available, selection committees may be less likely to choose a candidate due to political or financial influence as opposed to merit.

---

57 Ibid.

58 Ibid.

59 Shawnette Rochelle and Jay Loschky, "Confidence in Judicial Systems Varies Worldwide" Gallup

60 International Open Data Charter

61 Verhulst, Stefaan and Beth Noveck et al, "The Open Data Era in Health and Social Care." The Governance Lab (May 2014), p. 30

Second, government agencies, infomediaries and civil society organizations can **apply algorithms to raw judge and court performance data to identify corruption**. For example, by calculating average case review time, organizations can identify cases either dismissed too quickly or which drag on indefinitely, two characteristics potentially indicative of judicial malfeasance or inefficiency.

The potential benefits of reducing the information asymmetry between the courts and the public are numerous. From the citizen perspective, availability of information can lead to **increased knowledge** of the judiciary and consequently **increased power to monitor corruption and demand improvements.** Citizens may also experience **more just and positive experiences** with the judiciary as a result of a decrease in such corrupt practices as bribery.

For government, an open data portal, especially one that takes advantage of existing website infrastructure, can be an invaluable tool to **combat corruption** in the judiciary, **build legitimacy** among citizens, and **improve performance** of courts and judges. Governments also have the possibility to **realize tremendous cost savings by improving courtroom efficiency.**

EXAMPLES OF EXISTING GLOBAL AND LATIN AMERICAN INITIATIVES

It is well-understood and documented how to create such open justice data websites. There are numerous existing examples around the world that suggest what to do and not to do to achieve results. On the whole, these projects demonstrate that it is possible to build and launch such an initiative in a few months. At the same time, they showcase the importance of setting goals, identifying the data needed to achieve those goals, and cultivating the community to scrutinize the data and translate insights into improvements.

- **Datos Justicia Argentina (Argentina)** – In 2016, Argentina's Ministry of Justice opened up more than 30 administrative datasets about the ministry to the public. The Ministry has also committed to publishing performance and administrative data for all of the federal and provincial courts. The data portal is part of a larger citizen engagement initiative called Justicia 2020, an "online platform as well as an in-person collaborative space where civil society and the government can debate and participate in the design, implementation and evaluation of public justice policy." In its first year, nearly 20,000 people registered in Justicia 2020 and submitted over 3500 contributions to 118 active debates.[62]
- **Atviras Teismas (Lithuania)** Transparency International (TI) Lithuania built this website in 2015 after successfully advocating for a commitment from the National Courts Administration to open up their data to the public. The portal looks at key performance indicators for courts and judges such as length of proceedings and the number of cases resolved. The portal also documents the annual budget of courts and the experience of judges working in courts. The portal calculates averages for all metrics, allowing users to compare courts and judges to peers. Currently, the site has received 24,000 unique visitors which amounts to nearly 1% of Lithuania's citizens.
- **Otvorené Súdy (Slovakia)** Transparency International (TI) Slovakia created this website in 2013, with the goal to "increase pressure on the quality and efficiency of Slovak judiciary by utilization of open-data."[63] It started as a project by two university students at the Slovak University of Technology in

---

62 "Argentina's Judicial Data" TICTeC Conference.

63 Molnár, Samuel, and Pavol Zbell. "Frequently Asked Questions · Open Courts." · Open Courts

Bratislava initially analyzing court performance and has grown to include data on judges, hearings, decrees, proceedings and judicial selection processes. TI Slovakia has also developed its own statistical methodology that produces a score and rank for each judge in order to compare performance. Each week, the site receives several thousand unique visitors. This project demonstrates the importance of starting small and growing as more data becomes available.

- **Observatório da Estratégia da Justiça Federal (Brazil)** – Beginning in 2009 the federal judiciary in Brazil began publishing an annual written report documenting performance data on the federal courts, particularly their progress toward annual strategic goals. Since 2014, they have published that data online. The portal replicates the report's data, but in a more timely manner, publishing progress monthly instead of annually. The portal also allows users to more easily manipulate the data by comparing courts against peers or across time. The Federal Justice Council also holds an annual event that reveals the data hosted on the portal in a large public event. High performing courts receive prizes while broad media coverage encourages competition and puts pressure on lower performing judges to improve their score.
- **Justiciapedia (Argentina)** Asociación Civil por la Igualdad y la Justicia (ACIJ) and Chequeado launched this website in 2015 to promote accountability of judges and prosecutors in Argentina by creating profiles with information about their activity as public officers and their academic and work records. The profiles document when and how the official was appointed, their declarations and tax reports, the most important cases and decisions made in office, and any information regarding potentially corrupt rulings. The profiles also include network maps that show the judge's connections to individuals and organizations based on publicly available open data, which aims to help conflicts of interest detection. Shortly following the publication of Justiciapedia, the Supreme Court committed to publishing similar data on its website.
- **Concursos Transparentes (Argentina)** was developed in 2010 to monitor judicial appointment competitions in Buenos Aires in response to landmark legal reforms that opened up a number of new positions in the judiciary. The site monitored these competitions by publishing information about the process and candidates, including the CVs, interviews and scores of the candidates in each stage of the competition. As a result of ACIJ's work and recommendations, the judiciary made important modifications to the regulations of the competitions and opened their data, publicly posting information such as candidate exam scores, CVs and interviews. Following this project, ACIJ launched the National Edition of Concursos Transparentes by analyzing regulations and laws of all the Argentinean provinces.

QUESTIONS FOR CONSIDERATION

To design an effective judicial data portal, there are four main questions to consider, each of which is summarized below and then discussed at greater length:

1. What data?
2. What is the availability of the data?
3. What is the quality of data?
4. Who is the user and what are the user needs?

The **first question** focuses on identifying what data to publish. To improve the integrity of the judiciary, there must first be a theory of change that identifies what information is needed to identify, prevent and act on corruption.

The **second question** focuses on determining what data is currently collected. Existing data is often the most practical source of information and the fastest to go up online. Collecting new data requires investment and training to enable staff across the court system to supply this information. Thus, there needs to be a clear vision as to what data is needed and why.

The **third question** highlights the need to understand the quality of existing data. This is a critical consideration as incomplete and inaccurate data will not produce the desired performance improvements. Deciding and openly disclosing the level of timeliness, accuracy and comprehensiveness of data are essential to predict and analyze corruption meaningfully and to increase the legitimacy of the project.

Finally, **the fourth question** considers who the users of the platform will be and their needs. In order for data to be used to create actionable changes in society, it must be useful and accessible for the intended actors. Thus, in developing a portal, organizations must consider which users are most primed to act on the data and how the portal will meet their needs.

WHAT DATA?

Building an open justice website assumes that publishing data will result in a decrease in corruption. It requires identifying both what data is collected or could be collected and the theory of change by which such disclosure will lead to changes in behavior and a strategy for testing those outcomes. Thus, the first consideration is: what data should be published and why?
Other countries currently collect and publish four categories of data to combat corruption and increase public engagement in oversight of the judiciary: court and judge performance data, judge profile data, and judicial selection data.

DATA CATEGORY	EXAMPLES OF TYPES OF DATA COLLECTED	THEORY OF CHANGE	PORTAL EXAMPLES
Court Performance	▸ Number of judges working ▸ Number of cases reviewed ▸ Average time to review a case ▸ Average workload of judges ▸ Judges' average number of years of experience ▸ Annual funds allocated to the court	Documenting performance through standardized metrics enables users of the portal to benchmark or rank an individual judge or court against peers. **By calculating average performance, users can flag negative outliers that**	▸ Federal Justice Council of Brazil: **Observatório da Estratégia da Justiça Federal** ▸ Transparency International Lithuania: Atviras Teismas ▸ Transparency International Slovakia: Otvorené Súdy

DATA CATEGORY	EXAMPLES OF TYPES OF DATA COLLECTED	THEORY OF CHANGE	PORTAL EXAMPLES
Judge Performance	▸ Number of cases ▸ Time to review a case ▸ Workload of judge ▸ Number of days working in a year ▸ Years of experience	**may indicate instances of corruption** while studying positive outliers to establish best practices for the judiciary.	
Judge Profile	▸ Background – education, CV ▸ Selection process ▸ Assets and declarations ▸ Network maps – connections to organizations and individuals ▸ Involvement and rulings in high profile / controversial cases	Documenting conflicts of interest enables users to **identify cases** where a judge may have been or may be subject to **undue influence.** Applying analytics to the data could identify **conflicts of interest** in **case assignment.**	▸ Asociación Civil por la Igualdad y la Justicia in Argentina: Justiciapedia ▸ Transparency International Slovakia: Otvorené Súdy

Judicial Selection	▸ Type of process
	▸ People involved in selection / selection committee
	▸ Candidate data
	  ▸ Education
	  ▸ Work History
	  ▸ Age / Years of Experience
	  ▸ Oral / Written Exams and Scores
	  ▸ Interviews
	  ▸ Scores for each step in the process
		Publishing selection data **increases transparency of the process** and allows users to **evaluate the fairness** of the competition.
		
		Applying analytics to the data could provide an objective analysis of candidates and **flag if outlier high performing candidates do not advance or low performing candidates are ultimately selected.**
			▸ Asociación Civil por la Igualdad y la Justicia: Concursos Transparentes (Buenos Aires – monitored specific competitions)
			▸ Asociación Civil por la Igualdad y la Justicia: Concursos Transparentes (Argentina – documented process for every region)
			▸ Transparency International Slovakia: Otvorené Súdy

Table 1: Types of Data on Anti-Corruption Judicial Websites and their Rationale

The table above summarizes the types of data that are published on open justice websites and explains why (theory of change) they are disclosed.

The project's architects will still need to decide what data they want and have access to and articulate the rationale for collecting and publishing such data to achieve improved outcomes, including increased

efficiency for the courts, more and more rapid dispositions, and a reduction in impunity. As we shall discussed at greater length in the section on workplan, we recommend: 1) publication of all data that is collected about court and judge performance; 2) selected publication of key data about judicial personnel and their selection; 3) a roll-out plan to gather and publish additional data over time; 4) extensive public and stakeholder consultation to gain buy-in for the disclosures and to recruit collaborators to scrutinize the data; and 5) follow-up evaluation to assess the impact of these disclosures over time and evolve the project.

WHAT IS THE AVAILABILITY OF DATA?

The second factor to consider is the availability of data: what is collected, by whom, with what frequency and, to the extent to which there is a desire to publish data that is not yet collected, what would be involved in gathering that data. While it is possible to obtain data that is not currently being collected, it is often a labor-intensive process that includes developing collection processes, building political buy-in and training staff and developing publication mechanisms.

More often than not, existing data is often the most practical source of information. Thus, we recommend establishing a process for inventorying available data and its sources and identifying that which is missing and to assess the burden of collecting additional data. That inventory can be done in a distributed fashion with different agency components reporting back in a short time frame about data availability or by using domain crawling software to find data hiding in plain sight.

In Mexico, there is already a wide array of data published in machine-readable format from which to create open justice websites. However, that data is currently: 1) not available from a single place, 2) cannot easily be compared, and 3) does not yet translate into performance improvements or a reduction in corruption.

According to Eduardo Bohórquez, Executive Director of Transparency International Mexico, there are four main places where data is or will be collected and published on the judiciary:

1. La Plataforma Nacional de Transparencia (PNT)
2. La Plataforma Digital Nacional (PDN)
3. Federal and State Judiciary Council sites
4. Instituto Nacional de Estadística y Geografía (INEGI)[64]

The first portal, La Plataforma Nacional de Transparencia (PNT or National Transparency Portal) , was created to comply with the Ley General de Transparencia y Acceso a la Información Pública (LGTA, General Law on Transparency and Access to Public Information). The portal allows the public to request access to information as well as view all data on a portal called Sistema de Portales de Obligaciones de Transparencia (SIPOT, or Transparency Obligation Portals System) that all government agencies, including the Consejo de la Judicatura Federal are mandated to publish. This data includes general administrative data such as governance structure, finances, organizational goals and key indicators as well as specific judicial data such as public versions of sentences, weekly publications and agreements from el Semanario

64 Smarter Crowdsourcing: Anti Corruption Online Conference on Judicial Integrity, June 20th, 2017

Judicial de la Federación (The Judicial Weekly of the Federation) or in el Semanario Gaceta (the weekly Gazete) of tribunal administrative courts, and judicial selection processes.[65]

The second portal, La Plataforma Digital Nacional (PDN) is still in development as of the writing of this report (August 2017) and is being designed for the purpose of "sharing, concentrating, systematizing and analyzing relevant information for the prevention, detection and punishment of acts of corruption."[66] However, one of the main purposes of the PDN, as mandated by the Ley General de Responsabilidades Administrativas (LGRA), is to publish the taxes, declarations and interests of all public officials, including members of the judiciary.

The third source of data are the state and federal portals hosted on the respective Consejo de Judicatura (Judiciary Council) websites. The Federal portal lists all of the transparency obligations mandated by the LGTA and publishes them on a transparency data portal. Additionally, the site publishes a judicial statistics portal called la Dirección General de Estadística Judicial (the General Directorate of Judicial Statistics) that maintains statistical information about jurisdictional activity, including case filings, correspondences and administrative records.[67]

Finally, the national statistics office of Mexico, el Instituto Nacional de Estadística y Geografía (INEGI), publishes public safety and justice statistics, including the number of criminal cases filed, number of staff in judicial organizations, and perceptions of the performance of the judiciary. But there is no easy way to track cases filed with eventual outcomes.

Overall, these four data sources contain hundreds of data sets mandated by law. They are published with varying frequency – PNT and INEGI commonly aggregate their data sets on a yearly basis while CJF publishes more recent data, such as rulings, as often as daily. The data published through these sites covers many of the categories represented in Table 1 above. The largest gaps in data are seen in the detailed profiles of federal judges and officials, including their networks and background, as well as candidate data about judicial selection processes.

For data that is currently not being collected by an agency but may be critical inputs for anticorruption objectives, organizations should consider the value of starting a new initiative to source the data. On the benefits side, owning the sourcing process makes it possible to define the data points that are the most useful proxies for corruption and set standards for quality. On the cost side, collecting more data may require investments in technology and staff to develop and implement data collection systems. Organizations may also have to do considerable work to develop buy-in from relevant stakeholders such as judges or courts that may be wary or fearful of the burden of collection and the consequences of publication of new data sets.

---

65 For a complete list of data, see articles 70, 73, 77, 81 and 82 of the Ley General de Transparencia y Acceso a la Información Pública

66 Santillán, Oscar. "Perfilan Plataforma Digital Nacional Para El Sistema Nacional Anticorrupción" Publimetro México

67 Consejo De La Judicatura Federal

WHAT IS THE QUALITY OF DATA AND HOW SHOULD THE DATA BE PROCESSED?

"The quality of data is the most crucial piece. If it is not complete, if it has mistakes, if it's not giving you the actual information you need, you cannot even start," says Matej Šimalčík, Legal Counsel at Transparency International Slovakia. While data quality can be defined in many different ways, the Open Data Charter describes several common indicators of open data quality. The charter advocates for data that is **timely and comprehensive**, meaning it is accurate, released within a reasonable time frame, containing as much unmodified information as possible and, if relevant, disaggregated to the lowest levels of administration. Data should also be **accessible and usable**, meaning it is easily downloadable, in open formats that can be read by machines and humans, and used effectively by the widest range of users.[68] Building an open justice portal that will allow users to evaluate and monitor the judiciary meaningfully requires defining the frequency with which data is updated and setting parameters for what it means in this context for data to be accurate, complete and easily manipulated to predict or analyze corruption patterns. Although Mexico already publishes a great deal of data, including about the judiciary, much of it is not of high quality as it is published as PDFs or embedded tables, charts and graphics that cannot be easily downloaded or manipulated. It is also unknown how accurate the current data is and how often it is audited. Thus, it will be important to consider ways of periodically evaluating data and its quality.

WHO IS THE USER AND WHAT ARE THE USER NEEDS?

The production of data absent someone to scrutinize and analyze it will not result in behavior changes. For example, in Muskingum County, Ohio in the United States while access to clean water was available to most of the predominantly white county, residents of the largely African-American cities in the county were systematically cut off from clean water lines for nearly 50 years. After years of legal battles, the issue was resolved only when a demographic specialist, Dr. Arnold Parnell, meaningfully combined long-available geographic information system data, household survey information, and water bills to create a map clearly highlighting the water access discrimination for the jury.[69] The most important lesson learned from earlier examples of open justice websites around the world is that it is as important to define and target the intended audience for this data and what resources they need to make use of it as it is to publish the data, keeping in mind that that audience might be internal to government. We recommend developing user stories – narratives to describe how different users will take advantage of the data to produce the desired changes – in order to ensure that raw data is available in adequate quantity for professionals and selected infographics and data stories for the general public.

In other countries, there are five primary users of open court data:

- Government agencies with the power to select, evaluate and dismiss judges and who can use the data to identify and investigate judges potentially involved in corruption (in Mexico, the *Consejo de la Judicatura*)
- Media organizations that can use the data to support investigative articles on the judiciary
- Infomediaries with the technical skills to clean, analyze and make sense of the data for everyday citizens

---

68 International Open Data Charter

69 Rogawski, Christina, et al. "Kennedy vs. the City of Zanesville, United States." Open Data's Impact - The GovLab,

- Prosecutors that can use the data to develop cases against corrupt members of the judiciary or politicians/corporations exercising undue influence on justice issues
- Citizens and civil society organizations that can use the data to generate evidence that supports their advocacy or reform efforts

Undertaking the project should include a stakeholder analysis to determine the types of users to target for the open data portal. These users should be incorporated in every stage of the project design and implementation to optimize the usability of the final product.

While the portal should be customized to the target users, a key demographic of any open justice website should be the average citizen. Experts around the world agree that a primary goal of publishing judicial data is to enable the citizenry to understand everyday performance and corruption in the judiciary in order to mobilize for change. To this end, organizations should consider how to process data and visualize it and with what frequency to be accessible to the everyday user.

To tell a story with the data for the benefit of the general public, for judges and for non-technical users, the project needs to include some investment in data processing and visualization. There are two main types of processing that organizations have utilized to tell a story about legal information: *grouping* and *analytics*.

GROUPING

For performance data, appropriate grouping or combining of data into relevant categories requires identifying the characteristics that define peer groups, where performance comparisons make sense. In the National Justice Council of Brazil, courts are grouped by level (i.e.,Federal, State), specialization (i.e., labor justice, electoral justice), and size. Transparency International Slovakia groups judges based on specialization and geography. As TI Slovakia explains on their open courts platform, depicting performance data: *"is most useful for comparisons of judges at the same or similar courts, and those who decide cases in the same or similar agenda. Specific courts and agendas have an impact on the overall performance of judges. Simply put, [a] commercial judge from Bratislava and criminal judge from Rožňava decide uncomparable cases."*[70]

Data can also be grouped thematically to allow information that is usually stored in separate databases to be viewed in a single location. To develop judge profiles, *Asociación Civil por la Igualdad y la Justicia (the Civil Association for Equality and Justice)* combined data sourced from physical records in the Senate, personal outreach to judges, online public databases, official government websites, press releases and media coverage on a single page to provide a more holistic picture of judges' interests.

ANALYTICS

There are several algorithmic or statistical methods that can be applied to analyze raw data. For performance data, **benchmarking** based on means can allow users to quickly spot outliers that may be indicative of inefficiency or assist in identifying corruption. For example, Transparency International Lithuania depicts every court and judge performance statistic against the group average which allows users to immediately visualize irregularities. Organizations can also utilize **scoring** – Transparency International Slovakia developed a statistical methodology to produce a score and rank for each judge that is published on their portal. Finally, organizations can develop predictive algorithms to identify when there are red flags that may indicate corruption. Two applications of **flagging** include analyzing judge profile data

70 From example judge profile: https://otvorenesudy.sk/judges/4?l=en

against case assignments to determine if there are conflicts of interest and comparing predicted top candidates against those actually advancing in the judicial selection process.

In addition to how the data is processed for target users, the organization should also consider how the user will interact with the data through the design interface. Portals should be designed with continuous user feedback to check for ease of navigation, ability to search, find and download key information, and visual appeal. Organizations can amplify the user experience with the use of graphics and charts that can visually tell a story about the data.

EXAMPLES

- Brazil's Federal Justice Council utilized Microsoft Power BI visualization software to make a number of infographics that visually demonstrated court progress toward strategic benchmarks.
- Transparency International Slovakia developed visuals that allow users to compare judges on four aspects of quality and identify whether they are above/below the mean. The example below indicates that the judge (marked in purple) outperforms the mean (marked in gray) on two axes, while underperforming on the remaining two axes.

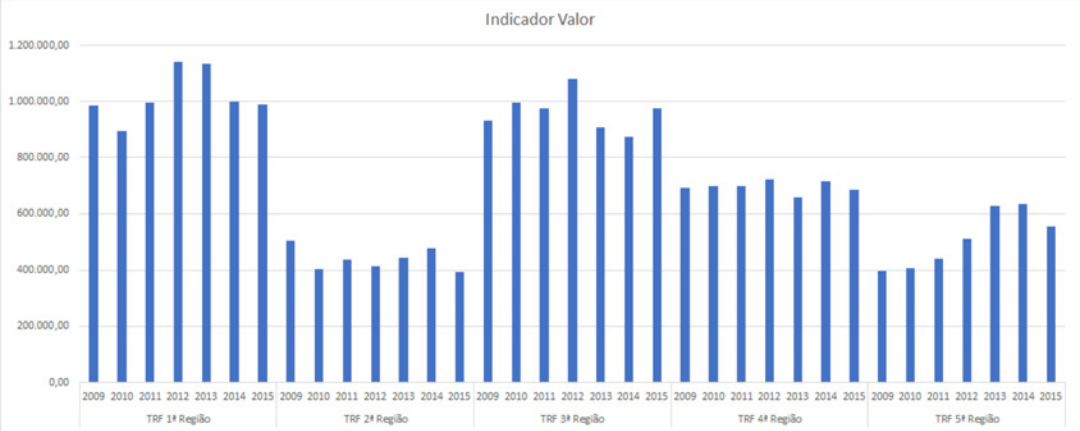


Image 1: *Comparison of the five National Justice Council courts in Brazil from 2009-2015 on remanded/ dismissed cases (***source***: Federal Justice Council of Brazil)*

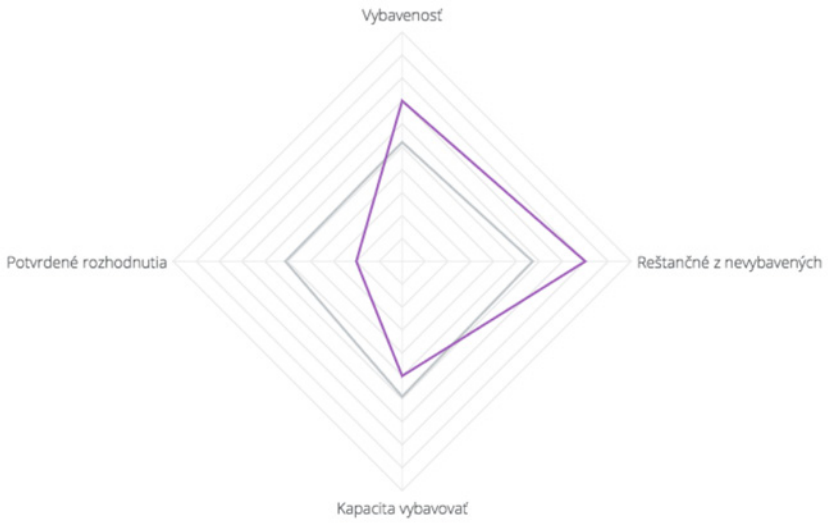


Image 2: *Visual of individual judge performance (purple) on four metrics, mapped against the mean (gray) (***source***: Transparency International Slovakia)*

- Asociación Civil por la Igualdad y la Justicia created network maps that enable viewers to see the connections between a judge and affiliated individuals and organizations. The example below describes a network map for one of the judges in Argentina.[71]

Image 3: *Network map of a judge (***source***: Asociación Civil por la Igualdad y la Justicia)*

[71] From example judge profile on Justiciapedia: http://chequeado.com/justiciapedia/profiles/alberto-agustin-lugones/#

IMPLEMENTATION STEPS AND RESOURCES

Based on our comparative evaluation of other open justice projects, we estimate that an organization with access to the right data – be it a government agency working alone or in cooperation with a civil society organization or university that can accelerate the analysis of the data – can create an open justice website with a staff of 1 full-time employee to manage the project and 2 to program it within 6 months and grow the project over 18 months.

The greatest impediment to achieving success in this timeframe will be the political difficulties created by those afraid of the burden created by data collection or who fear what transparency might reveal. According to Paulius Murauskas, who helped manage the open courts project at Transparency International Lithuania, "buy in from judges is essential – you need to understand whether they will be comfortable with the platform, if they see how it could be beneficial for them or whether they would find it intrusive."[72] Thus the project owner(s) need to have enough clout to drive the project to fruition and have the backing of senior leadership at the highest levels as well as access to boots-on-the-ground help with translating the raw data into insights and improvements. Also end users of the portal must be actively involved in the design, development refinement and outreach processes. The biggest pitfall is designing a website that is not accessible or actionable for its intended users.

This section will walk through the relevant steps and resources needed to implement this solution.
At a high level, in general developing any data platform involves identifying the problem the portal seeks to solve, identifying and sourcing the data that will address the selected issue, working with stakeholders to design and build the platform, launching a communication strategy and evaluating and refining use.

1. **Define and get agreement on the project goals (2 weeks) -** *Identify the specific challenges within the judiciary that the data portal will address* - For example, in Brazil, the National Justice Council conducts an annual process for defining the strategy and performance expectations of each court. The Federal Justice Council court portal was designed to address the issue of accountability toward these strategic plans by measuring court progress against articulated performance goals. Some of the specific challenges Mexico may choose to address are: opacity in the judicial decision-making process, undue political influence in rulings, and high impunity rates. In addition to scope, the problem statement should also address scale. In Argentina, the Ministry of Justice chose to address the opacity of the judicial system across national and provincial courts, while the Federal Justice Council in Brazil opted to focus exclusively on the performance of federal courts. Due to availability of data, Mexico should focus on piloting the initiative at the federal level. This implementation plan can be used to craft a memo and deck to get buy in and agreement on the goals of the project from senior officials and other key partners.
2. **Develop a strategy for involving key people and groups in the project (2 weeks)** – *Identify and convene stakeholders at each stage of the process* – Critical stakeholders will include government agencies and organizations that supply political, financial or information resources to build the portal, as well as end users such as accountability agencies, watchdog organizations, info-mediaries, journalists and citizens who will use the portal. While the latter set of stakeholders is essential for

[72] Interview with Paulius Murauskas, Open Courts program coordinator, Transparency International Lithuania on 26th July, 2017

developing the vision and ultimate end product, the former set is necessary to scope the practical parameters for what type of data can be published. Organizations should identify which individuals and organizations best represent key stakeholder groups and how to convene these groups for input and co-creation throughout the project. It is important to convene partners throughout the process.

3. **Identify and source the data (3 months)** – *Determine which data sets can be utilized to address the problem and their quality*
    a. Conduct an analysis of the existing data landscape to determine the scope of data available to publish. To speed up this process, ask others to respond to a questionnaire rather than trying to do all the research on one's own.
    b. Identify criteria to prioritize which data sets to source – examples:
        i. **Usefulness** in addressing the problem – How directly can the data be used to generate solutions to the defined challenges?
        ii. **Ease of access** – Will institutions sourcing the data easily provide access?
        iii. **Ease of publication** – Is the data already formatted for publication and analysis?
        iv. **Timeliness of data** – How regularly is the data updated?
        v. **Quality** – Is the data set complete, accurate, and machine readable?
        vi. **Political importance** – Does the release of the data carry political weight in establishing the credibility and commitment of the government?
        vii. **Citizen input** – Is there articulated or solicited citizen demand for the data?
        viii. **Compliance with transparency commitments** – Are there data sets that are mandated to be published that are currently inaccessible to the public?
    c. If data is not already published, create a list of priority data set candidates to publish
    d. Conduct legal diligence to ensure that identified data meets security, privacy and sensitivity requirements as mandated by law
    e. Use the data inventory to develop a roll-out plan for the project by deciding which data to start with, even if incomplete, and how to evolve the project over time.
    f. **Source the data.** The data sourcing process is highly dependent on the scale of the data being collected, the technical capabilities and political will of the sourcing agencies and the standards for formatting and submitting data. For example, in Brazil all federal court data is stored on an electronic system to which the Federal Justice Council has access, thus making it relatively easy to source. Conversely in Argentina, the Ministry of Justice developed a year long process to visit all 23 provinces in order to secure buy-in from the heads of the judiciary and the prosecutors' offices and assess local data collection systems to provide technical assistance for submitting data.

4. **Design and build the website (1-3 months)** – *Identify desired user experience* – Most open data portals are not databases but simply websites that point to data published elsewhere and that include additional visualizations and discussion features to build a community around the data. Thus the technical aspects need not be complicated.

a. Utilizing human centered design[73] or A/B testing[74], develop experiments to design the platform interface with target users. Paulius Murauskas described this part of the process as "essential and iterative"[75] – throughout their portal development, they conducted several rounds of stakeholder meetings with government officials, judges and citizens which allowed users to test functions of the portal and offer feedback on their experience.
b. Create a plan for sourcing data and publishing data on an ongoing basis
    i. Define standards for formatting and submitting data – examples:
        1. The Ministry of Justice in Argentina required that court data be submitted in CSV formats through a Secure File Transfer Protocol (SFTP).
        2. The Federal Justice Council of Brazil piloted a standardized data collection system but switched to Excel files when the system proved to be overly complex. They have since developed a system using PHP and MySQL.
    ii. Based on the technical competency of team and user requirements, select open source coding languages, software packages, database management systems and frameworks – examples:
        1. Argentina's Open Data portal was developed entirely through the open source CKAN.
        2. Brazil's Federal Observatory was developed through a wide variety of languages and software packages, including R, Tableau, Microsoft Power BI, MySQL, PHP, Bootstrap and Pentaho.
c. Develop the website – There are countless examples of good government open data websites. For advice about basic development principles, see the UK Government Digital Services Design Principles and use the Network of Innovators to ask questions of and learn from other open data designers. Network of Innovators also includes a 20-question checklist for planning an open data project.
d. Design key charts, graphs and visualizations to tell a story with data that responds to user needs.

5. **Launch the platform (1 month**) – *Create a communication and outreach strategy for engaging target users* – A communication strategy may focus on simply getting the word out about the platform or holding training or events that provide guidance on how to use the website. For example, the Ministry of Justice in Argentina connected the open data portal with its Justice 2020 initiative, which enables civil society and government officials to meet in forums to debate and co-design public justice policy. The Ministry also organized hackathons and datathons, and launched a media and social media campaign to spread the words about the initiative. In Brazil, the Federal Justice Council engaged over 1,000 citizens to provide feedback on the performance targets presented on the platform. They also launched a series of videos such as this one, which explains how the data portal is serving the public. After launching the national version of Concursos Transparentes, ACIJ reached out to local activists to provide assistance in understanding the portal and organizing for reform. In one province, civil society organizations successfully lobbied to reduce the weight of oral examinations – one of the most arbitrary measures of fitness – in the judicial selection process.

---

73 For resources on human centered design concepts, methods and case studies, consult Ideo's Design Kit

74 For more info on A/B testing concepts, methods and case studies, consult Optimzely's Primer on A/B Testing

75 Interview with Paulius Murauskas, Open Courts program coordinator, Transparency International Lithuania on 26th July, 2017

6. **Evaluate portal and refine (ongoing)** – *Measure site usage and performance outcomes, adjust design and engagement strategy as necessary* – The website should be regularly evaluated against progress toward addressing the defined problem. Based on the performance of the portal, government should continuously refine the platform and the strategy for engaging its target users. Metrics might include:

   - Outputs to track usage of portal:
     - Number of visitors, disaggregated by type of user (e.g., citizen, civil society org)
     - Number of data sets downloaded
     - Number of citizen engagement events or initiatives utilizing open data platform

   - Outcomes to track progress toward addressing identified issues
     - If the key issue is opacity in the judicial decision-making process leading to impartial rulings, measure changes in:
       - Citizen beliefs in the transparency of the decision making process
       - Citizen reporting on paying bribes to the judiciary to influence rulings
       - Number of appeals filed on rulings deemed impartial (for example, the judge used insufficient or low-quality evidence to support ruling)
     - If key problem is to reduce bias in the judicial selection process, measure changes in:
       - Competitiveness of processes (for example – average exam scores of judges increase)

ISSUE AREA 3

Increasing public engagement in anti-corruption efforts using new technology

Strong social pressure drove recent landmark legislative anti-corruption reforms in Mexico.[76] Civil society organizations working in collaboration with the national government played a key role in designing the reforms. Thus, not surprisingly, the new legal framework calls for increased participation by and collaboration with the public in preventing, detecting and investigating corruption.[77] But paying lip service to public engagement is not enough. The public expects **meaningful interactions that lead to measurable outcomes and more effective policymaking and service delivery.** To this end, this implementation plan builds on original data collected about the SIDEC platform and develops two practical recommendations to foster and strengthen public engagement in anti-corruption efforts:

- **Recommendation 1. Revamp Mexico's Online Corruption Reporting Platform To Increase Public Engagement**: Update and improve a set of features of the existing online reporting platform to increase the number and quality of reports received.
- **Recommendation 2. Set Up An Ongoing Open Policy Discussion Platform On Anti-Corruption And Public Integrity:** Design and launch an online mechanism for individuals and organizations to contribute

---

[76] The Mexico Institute. Mexico Wins: Anti-Corruption Reform Approved. Forbes. Published July 18th, 2016. Available at: https://goo.gl/NxaH9D

[77] A Citizens Committee will head Mexico's National Anti-Corruption System. This independent body will be in charge of identifying opportunities for public engagement, as well as channeling demands and proposals from the public to the government agencies that are part of the system. Chapter 3, Ley General del Sistema Nacional Anticorrupción (2016). Available at :http://www.diputados.gob.mx/LeyesBiblio/pdf/LGSNA.pdf

to the Federal Administration's anti-corruption policymaking efforts and foster ongoing conversation with a more diverse audience.

Together, these two projects, which build upon well-documented international examples and best practices, involve straightforward improvements to existing websites that can be accomplished over the next year. If successful, they will improve the legitimacy and effectiveness of the government's anti-corruption efforts and lead concretely to a doubling of the number of citizen reports, an improvement of the quality of those reports and, done right, an increase in the number of prosecutions and actions taken to correct problems and reduce the culture of impunity.

I. REVAMP MEXICO'S ONLINE REPORTING PLATFORM TO INCREASE PUBLIC ENGAGEMENT AND IMPROVE THE CAPACITY TO ADDRESS CORRUPTION

This implementation plan outlines a six-month process to increase incentives and reduce barriers to enable the public to report corruption online through SFP's "*Sistema Integral de Quejas y Denuncias Ciudadanas*" (SIDEC, Comprehensive Denouncement and Citizen Complaint System), as well as to make a strategic use of the reports received to act on corrupt behaviors within the Federal Administration. (Please note: this implementation plan focuses on public engagement and reports. See also the related Implementation Brief on Whistleblowing.)

By (1) increasing SIDEC's capacity to capture actionable information, (2) designing and publishing a new generation of open data and data analytics based on citizen reports that yield greater insight into corruption within the administration and where it occurs, and (3) implementing a proactive communication strategy to raise public awareness about the platform, SFP will be able to increase both awareness and use of the platform and grow the number and quality of reports received, enhancing the ability of the government to route corruption in the public sector.

Successful reporting mechanisms depend on their ability to raise awareness among and attract those diverse audiences affected by corruption and to create incentives for these different individuals to file an information-rich complaint. In addition to simple accessibility and usability, paramount among those incentives is the belief that the complaint will be acted upon. Thus, these recommendations for specific enhancements to existing tools are designed to improve the experience for both the public and the officials who must act on their complaints to ensure a more effective and actionable reporting process.

A. THE CHALLENGE

In an effort to detect and address public corruption based on the public's participation, Mexico's Ministry of Public Administration (SFP) renewed and launched a platform to report corrupt behaviors and other misconduct by government officials, called "Sistema Integral de Quejas y Denuncias Ciudadanas" (SIDEC-Integrated System for Complaints and Citizen Denouncements). The platform is intended to invite users to file a written report electronically on 36 different issues ranging from bribery to misuse of public resources to delays in service or disrespectful treatment by a public servant. The platform enables users

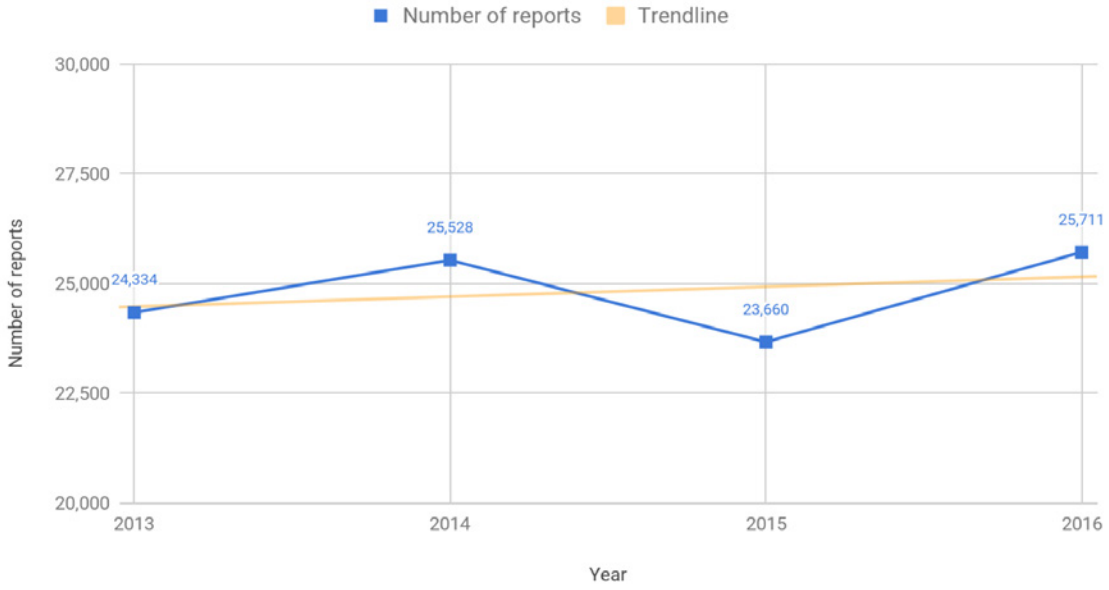
confidentially to provide an open-ended description about the incident as well as submit relevant evidence to the SFP, such as documents or photographs.

However, **participation levels on the platform are low** and fall far short of the potential to detect and, ultimately, prevent corruption as a result of the enhanced ability to spot and prosecute misconduct.

An analysis of the data about the system's performance identified **four problems** that are negatively affecting the relevance and impact of the system and the incentives for the public to report corruption. These are:

1. **Lack of Usage: SIDEC is failing to attract and sustain enough users.** Despite a series of efforts and actions to inform the public about its existence, the number of reports received through SIDEC has stagnated.[78] The number of reports received last year (2016) is similar to its first year of operation (2013). On average, the platform receives 24,808 reports annually, and it has never exceeded 26,000 reports during its four years of operation (see graph 1). As an additional reference, the mobile application[79] of SIDEC that SFP launched this year has not been downloaded by a critical mass of users. Whether because of a lack of awareness, a lack of accessibility, or a lack of credibility, the app has only been downloaded fewer than 1,000 times since it was published in January 2017 and we have no data as to how often it has been used.[80]

Graph 1. Number of reports received through SIDEC by year



Source: *The Governance Lab with data provided by Mexico's Ministry of Public Administration*

---

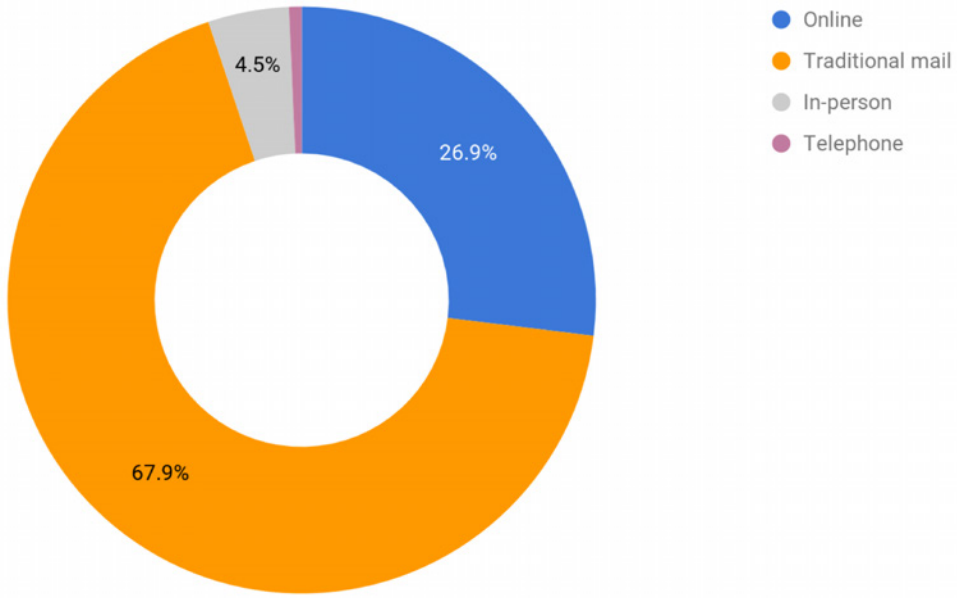
78 SFP has only provided data about reports received since December 2012, when the current Administration took office. For the purposes of this report only data from January 2013 to December 2016 has been reviewed.

79 The mobile application is available for Android, iOS, and Windows Phone.

80 Denuncia la corrupción, App description at Google Play. Available at https://play.google.com/store/apps/details?id=sfp.denuncia_la_corrupcion&hl=es. See also https://itunes.apple.com/mx/app/denuncia-la-corrupci%C3%B3n/id1171271018?mt=8

2. **Paper-Based Reporting: The public is not submitting reports electronically.** Despite the presumed simplicity and ease of online reporting, **1 in 4 (26.9%) of the reports that SIDEC received in the last four years were submitted electronically.** Among the different channels,[81] submission through traditional mail remains the main channel to report corruption. It is unknown whether the preference for paper reflects a lack of access to the Internet or the desire to preserve anonymity or a problem of usability (the digital option includes 30 different types of reportable misconduct but with very little explanation of the choices) but, regardless of the cause, the persistence of paper-based reporting is slowing SFP's capacity to route complaints to the proper authorities for action (see graph 2).

Graph 2. Reports received from 2013 to 2016 through SIDEC



Source: *The Governance Lab with data provided by Mexico's Ministry of Public Administration.*

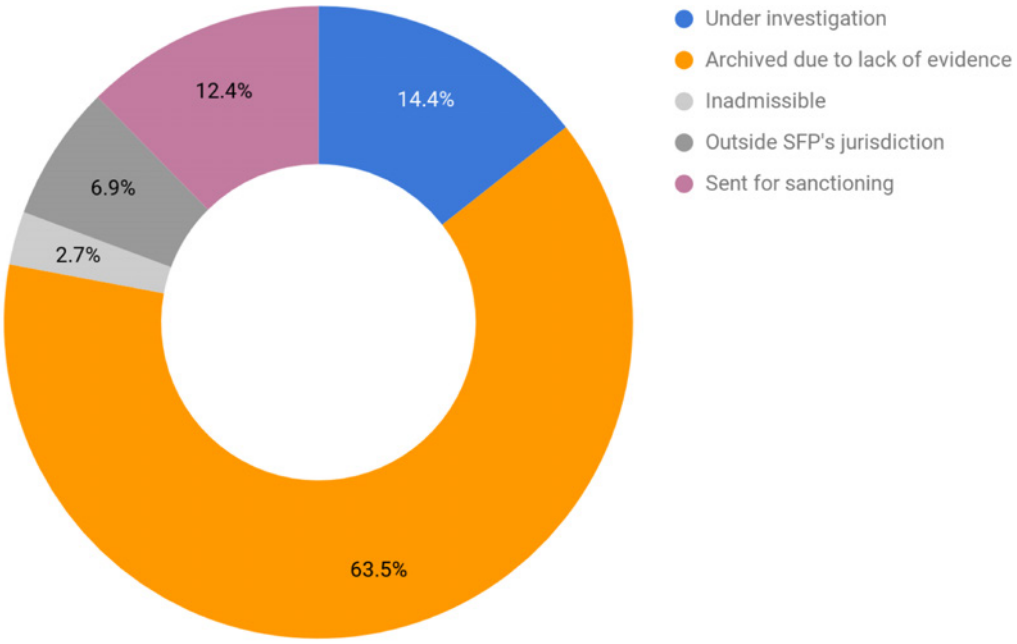
3. **Information Quality and Impediments to Data Sharing: Reports do not allow SFP to identify and investigate corruption.** SFP dismisses approximately seventy-three percent of the reports received, mainly because of the lack of evidence or information that would allow them to initiate any action. In the four years of the program **26.8% of the reports provided SFP with enough elements to start an investigation** (see graph 3). Regarding the investigations that were opened, it is not possible to identify which ones concluded with sanctions since the unit that manages SIDEC (Dirección General de Denuncias e Investigaciones) does not receive updates from the unit responsible for following up the investigations opened (Dirección General de Responsabilidades y Situación Patrimonial). As such, the reports that SIDEC claims led to an investigation could either have ended in a penalty against a government official or been archived due to lack of evidence or inaction from the investigative units (Órganos Internos de Control). Additionally, SFP is not

[81] There are four available channels to send a report: (1) telephone, (2) online (web platform, web chat and mobile application), (3) traditional mail and ,(4) face to face.

analyzing the reports received to detect patterns or outbreaks of corruption within the Federal Administration and, consequently, adopt preventive measures.

4. **No Open Data: Lack of transparency and accountability.** Although knowledge of the platform and its impact is integral to attracting more users, SIDEC does not currently provide public information about the reports received or the results. Obtaining this kind of information requires the submission of an access to information request. Furthermore, government agencies lack information about the platform that could be used to structure training programs or provide staffing to respond to complaints.

Graph 3. Status of reports received through SIDEC from 2013 to 2016



Source: *The Governance Lab with data provided by Mexico's Ministry of Public Administration*

To improve the process and the tools, it will be useful to: 1) understand who the current users are and how they perceive the experience, 2) survey those who do not use the digital options to document why they prefer paper over the electronic options, 3) survey those who do not use the system at all using direct questionnaires and indirectly using sentiment analysis with social media data to understand why they do not participate at all, and 4) endeavor to connect complaint data to actions subsequently taken.

B. THE OPPORTUNITY

During the Smarter Crowdsourcing online **conference** held on June 27th, 2017, that brought together **20 experts from 10 countries**, as well as officials representing the Mexican government, the Inter-American

Development Bank, and members of the GovLab, participants discussed different ideas and experiences to foster public engagement with anti-corruption efforts through the use of data and digital tools.

Further research and in-depth interviews have also contributed to identifying key strategies and approaches to strengthen public engagement with online reporting tools and increase the relevance and usefulness to prevent, detect, and investigate corruption. Although the conversation did not focus on the SIDEC suite of tools per se, the experiences of global experts with designing and operating online engagement platforms elicited instructive and intuitive insights for improving the current features and incentives of SIDEC and transforming it into a strategic tool to fight corruption in Mexico. The design recommendations stemming from that conversation with global experts can be summarized as follows:

- **Offer a useful service that responds to needs**. In anticipation of scaling up SIDEC to become the national reporting platform within the National Anti-Corruption System,[82] SIDEC should make it easy to report misconduct. As stated by the **UK's Government Digital Service**, a government platform is more than a website. To be useful for the public, "the digital world has to connect to the real world, so we have to think about all aspects of a service, and make sure they add up to something that meets user needs."[83] Thus, as we shall explore, it is imperative to make the platform as good as it can be. That means designing tools that are easy-to-use by the public but also by the institutions that must use the information. It is not enough to optimize the experience for the individual participant with clear directions, respect for privacy, and tiered authentication. But it needs to serve the needs of governing institutions and civil servants better.

- **Use human-centered design to improve the platform**. Reporting corrupt behaviors through SIDEC should be easy, simple, and intuitive. To achieve that, SIDEC should incorporate a human-centered approach, using ethnographic and participatory design techniques to craft a process that is well-designed. Its scope, instructions, and content should be clear and easy to understand for a general audience, and design should support and reinforce such objective. As Eduardo Bohórquez from Transparencia Mexicana mentioned, "instead of talking about laws and regulation, documents that are distant from citizens, we should look for simple things [for the user] to review and confirm in their reality."[84] The goal consists of reducing barriers to participation and allowing any user to provide valuable input to detect corruption.

- **Make the platform relevant and designed to trigger institutional reaction.** SIDEC must allow SFP to initiate either a legal or a policy action. The capacity to trigger a response from the government is one of the most significant incentives to foster public engagement with an online platform or civic initiative. The public is more likely to participate – and keep participating – when it can foresee a tangible benefit or change. This attribute can be strengthened by improving the quality of the reports received, as well as reviewing and improving the flow of information along the process and creating a more seamless

[82] Título cuarto (page 18), Ley General del Sistema Nacional Anticorrupción. Available at: www.diputados.gob.mx/LeyesBiblio/pdf/LGSNA.pdf

[83] Design Principles. Government Digital Service, Government of the United Kingdom. Available at: https://www.gov.uk/design-principles

[84] Participación Ciudadana Guarderías IMSS (2015) Mexican Social Security Institute. Available at https://www.youtube.com/watch?v=Qv2gaM4gmBA

connection between the work of SIDEC and Dirección General de Responsabilidades y Situación Patrimonial. Thus, changes in workflow more than just technology are integral to the reform effort.

- **Use data analytics to improve service delivery.** SIDEC should incorporate new methods, such as hotspotting and red-flags, to make a better and more agile use of the data that is received through the reports (as well as data from social media and other sources) and use these insights to target actions and improvements. SFP can take advantage of data generated by citizen complaints to yield new insights and evidence of corruption by incorporating simple data analysis to identify where the most complaints occur, who reports them, what kind of complaints are most often reported. A simple comparison of reports over time can improve the ability to detect patterns. Although the volume of data is still very small, with more data eventually it could become possible to spot patterns and identify members of a corruption network operating within a government agency. Such tools might be useful to predict corrupt behavior in the same way, for example, that government agencies have used data about past practice to enhance future performance.

- **Enhance trust and legitimacy using open data and feedback loops.** Public engagement does not end after the public uses the platform. Informing the public about the outcome of their participation can influence ownership of the initiative, foster continued or future interactions, and sustain public interest in the platform and the legitimacy of the larger program. In addition to responding to an individual about the specifics of his or her case, the government needs to publish the number of investigations and penalties imposed, SFP could also inform the public about actions that were taken to prevent or identify corruption, based on advanced analysis of the reports received.

C. THE ACTION PLAN

1. Strengthen SIDEC's capacity to capture strategic data to detect corruption by revising the reporting questionnaire and testing user's experience

Today the questionnaire is currently the **core component of the reporting tool** since it is the medium to capture the information needed to identify and address a problem. (As we shall discuss, there are other potential mechanisms for data intake, including the use of social media as well as developing new indicators for spotting corruption indirectly.) The structure and content of the questionnaire is thus critical as it affects the participation of the public as well as the government's capacity to respond and undertake corrective measures.

Finding the right balance between a simple questionnaire and gathering enough information to trigger governmental action will require experimentation and iteration. There is no standard in terms of the number or type of questions as people will share even a great deal of information, if they have incentive to do so. Still, there is a general consensus that reporting should be as effortless as possible, prioritizing questions that are easy for the user to observe or experience, balancing free-text options with multiple-choice questions, avoiding asking for unnecessary information and asking for information in stages.
Two different platforms offer instructive illustration of the value of good information architecture, especially transparency, for revising the SIDEC questionnaire:

The **first is FixMyStreet.com**, launched in 2007 by My Society (a UK charity), that reached one million reports this year. FixMyStreet (like its US equivalent See Click Fix or Brazil's Colab.Re) aggregates complaints for municipalities, such as reports of downed street lights or potholes. It is a lightweight, web-based alternative to a 311call center. Forty percent of those million reports have been reported as fixed. "Users send over 12,000 reports to UK councils, with more than 50,000 people viewing the site"[85] according to MySociety. However, what makes this tool unique is that regardless if the user knows who's the authority in charge of fixing the problem, it "matches the exact location of the issue, and its category, with the council responsible for getting it fixed. It sends a report by email to that council, and also publishes it online for all to see."[86] Reporting a problem can be done by answering only 5 questions (see image 1):

1. Mapping the problem (georeference)
2. Selecting the type of problem (one of only twenty categories)
3. Summarizing the problem (free text)
4. Providing (photograph)
5. Explaining the problem (free text)

The descriptions of the questions and the twenty categories use simple language, avoiding technical terms, and reducing the opportunity for ambiguity.

The **second example**, Mexico's **Social Participation Mechanism in Day Care Units,** designed by Transparencia Mexicana and managed by Mexico's Social Security Institute (IMSS), illustrates what happens when a site offers a relevant service, where the public is clearly invested in the outcomes and has an incentive to participate. The mechanism allows parents to review and report breaches in safety measures in the infrastructure and operation of public day care units. In contrast to FixMyStreet.com, users have to answer 91 yes-or-no questions. Yet over 42,000 parents have participated since 2010 because of the strong incentive to do so. Safety measures have improved considerably as a result of the reporting, further strengthening the reason to respond.[87] The instructions to complete the questionnaire also avoid the use of technical language and further descriptions are given in case of complex questions. In the words of Eduardo Bohórquez, Transparencia Mexicana's Executive Director, they did "a reverse-engineering exercise that translated safety regulations, written in legal terms, to a set of measures that were easy for the parents to observe and confirm in reality. There was no room for doubt."

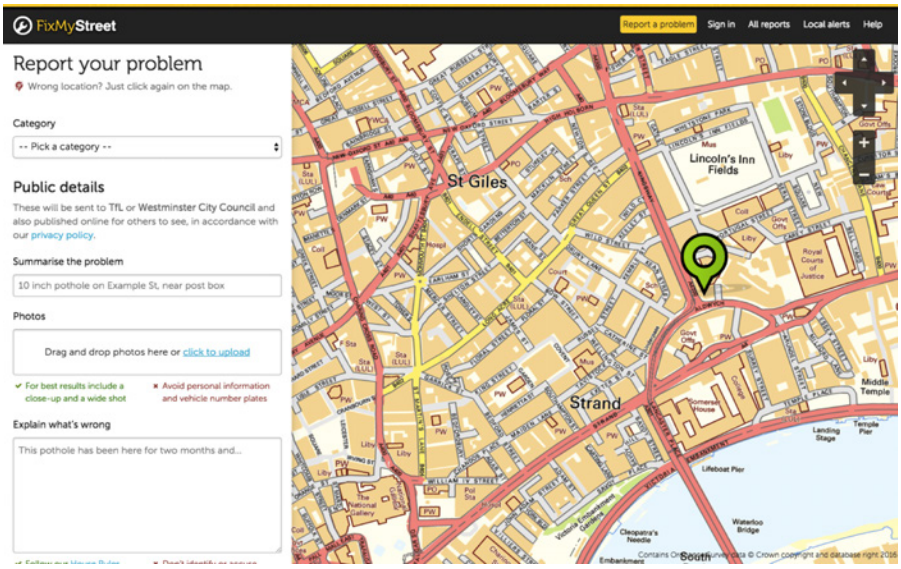
These are but two examples. The New York District Attorney, for example, facilitates complaints using a WhatsApp hotline to preserve anonymity while SMS Frontline uses SMS to enable citizen feedback without the complexity of an app. Ushahidi, for instance, asks the crowd to find and report instances of urban disrepair or ethnic violence, which governments then use to allocate city resources or deploy peacekeepers. Another example is PulsePoint, an app that alerts nearby CPR-trained first responders to come to the rescue for faster emergency response and that simply uses geolocation to trigger a response. For more on mobile feedback systems, see How Mobile Crowdsourcing Can Improve Workplace Safety.

---

85 My Society (2017). FixMyStreet.com case study. Available at: https://www.mysociety.org/better-cities/fixmystreet-in-the-uk/

86 Ibid.

87 The Open Government Partnership (2015) Open Government Award Booklet, 2015 edition. Available at: https://www.opengovawards.org/data/OGPBooklet2015.pdf



Source: *Reporting questionnaire of FixMyStreet.com by MySociety*

These examples suggest three concrete learnings that SFP can incorporate to strengthen the capacity to capture strategic data through SIDEC while reducing barriers and generating incentives for the public to use the tool:

1. **Capture the necessary information but not more.** SIDEC's questionnaire must provide SFP with the minimum data and evidence necessary to start an investigation. As such, it should be tailored to the different types of misconduct that can be reported as well as to the needs of the investigation units within SFP and the Federal Administration. Using responsive questionnaires and expert systems, the response to a first question should channel the user to answer only the necessary questions, making it possible to limit the number of questions to ten or fewer.
2. **Make the platform easy to use but specific.** SIDEC's questionnaire should guide the complainant and describe the minimum information that is needed to support a report. Tailoring guidance according to each misconduct, simplifying descriptions, and using predefined answers are techniques that improve the user's experience. For instance, the use of predefined answers reduces the margin of error while advising the denouncer about the kind of information that should be gathered to support the report. It also allows faster processing of the reports. Although demographic information about the complainant is also desirable such information collection should be separated from and come after the complaint itself.
3. **Simplify instructions and language.** The use of non-technical language is critical to broaden the scope of users, to non-specialized audiences. The questionnaire titles and descriptions should avoid the use of legal terms. Mouse-overs and links can provide the user who wants more detail with additional information.

ACTION ITEMS

1.1. Carry out an inventory and journey mapping of the reporting requirements to identify information needs

SFP should document the legal and practical requirements governing citizen complaints and then convene the three units that investigate administrative responsibilities to get consensus on this list.[88] This list of information requirements should distinguish between must have information and nice-to-have information.
SFP needs both to capture the requirements of the new General Law on Administrative Responsibilities, which now provides a more specific description about each misconduct (*falta administrativa*), to capture and identify precisely that information needed to act upon a complaint.

In addition, SFP should work with a user-experience designer to conduct a "journey map" documenting the information flow from complainant to recipient to those responsible for taking action and back to the citizen in order to identify what information is needed when and by whom and to compare against the design of the current system.[89] The journey map goes beyond a mere listing of categories of information to describe who uses the information and how in order to uncover opportunities for improvement.
This preliminary work should also, as discussed above, include 1) a survey of current users to understand how they perceive the experience, 2) a survey of those who do not use the digital options to document why they prefer paper over the electronic options, 3) information collection from those who do not use the system at all using direct questionnaires and indirectly using sentiment analysis with social media data to understand why they do not participate at all, and 4) analysis of prior reports to gain further insight into who users are, most frequently used processes and gain insight to inform the redesign.
Both the information requirements and the journey map should be published openly for comment.

1.2. Design and build a new expert system

A multidisciplinary but small team, comprising civil servants and experts on survey design, environmental psychology, strategic communication, and public engagement experts, should revamp SIDEC's questionnaire in order to respond to the requirements documented in step 1.1 as well as redesign the information flow to streamline how complaints get routed to those who can act upon them and the outcomes back to users.
This planning effort should produce a revised version of SIDEC's questionnaire and the design of an expert system (Q&A trees). Expert system tools like Screendoor and Typeform are relatively cheap, easily scaled, provide analytics and reports and are customizable as are more sophisticated expert system tools such as Neota Logic's expert system authoring tools. It should also produce documentation of the workflow. Even the best complaint or report websites are worthless if there is no clear process by which information ends up in the hands of people who have the authority, ability and incentive to take action.
The plan and new platform should be presented to a working group comprised of representatives of civil society organizations and academia specialized in public engagement for rapid feedback as well as to civil service members for feedback and testing. Prior users of the system should also be invited to serve as beta-testers for any new system.

1.3. Test the expert system with the public and the civil service

---

[88] This may include: Dirección General de Denuncias e Investigaciones, Dirección General de Responsabilidades y Situación Patrimonial and Coordinación General de Órganos de Vigilancia y Control.

[89] Sample Journey Map: "Government whistleblowing policies", UK National Audit Office, Fig 4, p13, January 2014. Note: A journey map is a visual aid which describes the various steps in a process, in this case that of acting on citizen complaints. A sample journey map is shown in the above report and describes the whistleblowing process in the United Kingdom.

Even the best designed website, if not well-enough integrated into institutional practice, will lead to an explosion in the volume of information received without any concomitant improvement in the quality of outcomes. In addition to being easy to use for the public, SIDEC needs to serve the needs of governing institutions better. Thus, SFP should run a series of tests with potential users, including government officials working in the corruption-sensitive area of procurement, as well as users of government services, like government contractors, and, of course, members of the public. The tests should include the use of ethnographic and observational techniques to watch users interacting with the site (as complainants and recipients) as well as questionnaires to survey users about their experience.

1.4. Shift from Pull to Push: Go to where the public is rather than waiting for them to come

Do not wait for people to complain. Instead, develop a plan to incorporate phone and text message surveys of recipients of government services to ensure wide demographic participation. For example, in the Punjab region of Pakistan the district coordination officer (DCO) of Jhang district "realized he could build credibility for the government by reaching out to citizens rather than waiting for disgruntled citizens to lodge formal complaints through an impersonal system" and thus launched an experiment. When citizens received services (regardless of whether corruption of misconduct was present), officials noted the cell phone numbers of those using government services. Then the DCO called a random sample of them after-the-fact to inquire about their experience.[90] Such a system of randomly calling citizens could be part of the input and design feedback process to ensure participation by diverse members of the public, rather than only the self-selected few. It also involves more distributed civil servants in the process.

The success of the Punjab project resulted in its evolution from a calling program to solicit feedback later to an automated texting platform to *push* requests to citizens for feedback on their experience receiving services. Compare the low level of usage of SIDEC's platform against Punjab's Citizen Feedback Monitoring Program, which sent 1.5 million text message surveys across the province of 100 million residents of the state of Punjab just in its first year of operation. The CFMP is an automated system that sends text messages to citizens' cell phones and request feedback on the experience of using government services. Citizens reply by texting. "The texting system had a significant cost advantage over the call-center system and provided comparable reach. A 160-character text message cost 80 paisa (less than 1 cent), while an average three-minute agent call cost 5 rupees (about 5 cents).

Convenience and effectiveness also favored the texting system. Because text messages did not require an immediate response, unlike live calls, citizens had time to consider their answers and discuss the messages with others. In addition, text replies could be stored more easily and accurately for analysis than audio recordings of phone conversations." In 2014 the World Bank commissioned a study of the program. A phone survey of users showed that: "55 percent of those surveyed said overall service delivery had improved, 71 percent said staff attitude had improved, and 63 percent said timeliness of service delivery had improved. Although only 30 percent of those surveyed said the program had reduced corruption, 76 percent of the respondents said they still believed it would help reduce the problem in the future."[91] Romania and Albania have now copied the same system.

---

90 Mohammad Omar Masud, Calling In Against Corruption, Foreign Policy, June 30, 2015, http://foreignpolicy.com/2015/06/30/calling-in-against-corruption-pakistan-punjab/.

91 Mohammad Omar Masud, Calling Citizens, Improving the State: Pakistan's Citizen Feedback Monitoring Program, 2008 – 2014, Innovations for Successful Societies, Princeton (Feb 2015), 14-15.

1.5. Run Experiments to Test What Works

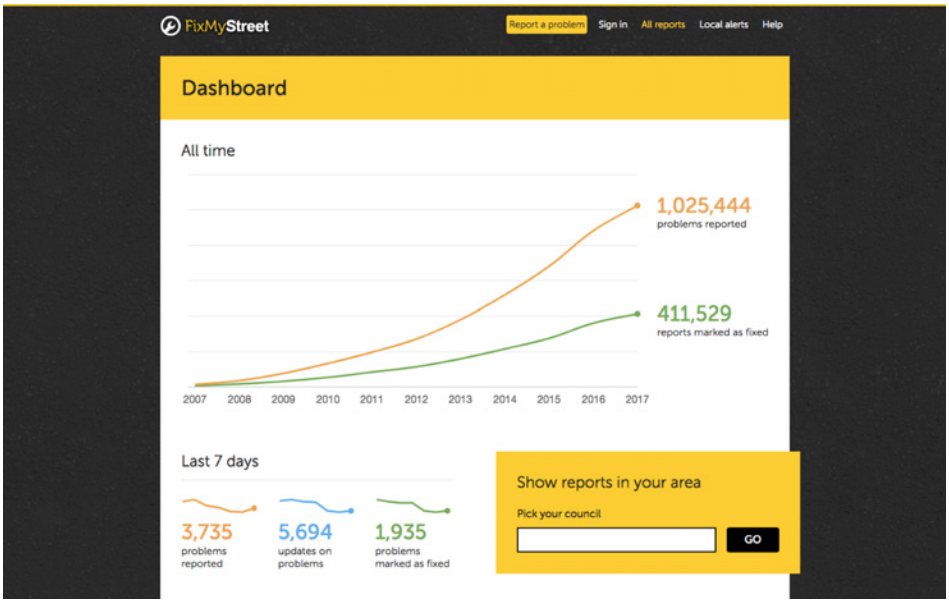
Work with social scientists to develop experiments, including randomized controlled trials to test design improvements and their impact on participation. In an effort to diagnose why more people do not use the SIDEC tools to complain, it will be important to test more than one design. Such experiments could include, for example, testing the hypothesis that changing the "priming" or the language used to invite people to participate may change participation rates. The research literature shows that the nature of the prime or language used to invite people to participate might have a significant impact on people's willingness to respond. Thus you could flesh out a series of research experiments designed to test different primes, such as texting the users of government services and asking them for feedback to improve their own experience versus performing a public service, for example. You could also test different designs for the questionnaires such as the use of open-ended questions versus multiple choice questions. You could test the use of Internet advertising and different communications strategies to see whether this impacts participation. Such experiments are designed by articulating the hypothesis and then designing a method for operationalizing the test, using digital technology to randomize the creation of two or more groups, including a control, to try to gain greater insight. Further experiments might test whether people are more likely to use the tools for different types of complaints or to target different types of users. You may also want to test the impact of providing a feedback loop on participation.

2. Design and publish a new generation of data analytics for SIDEC through collaboration with civil society organizations and academia

Data gathered through a reporting tool has usefulness beyond registering a problematic situation. For example, many reporting platforms have dashboards that contain statistics about the reports received and indicators about the performance of the authorities in charge of addressing the problems reported. **FixMyStreet.com** has a specific section that allows users to review the number of reports received since the tool was launched, the number of reports already addressed, and the five most responsive councils, among other features. This information can generate positive incentives for users to report, as well as for government agencies to respond to the reports received.

Generating this kind of information does not have to be expensive, elaborate, or time-consuming. There are several experiences where communities of data scientists, designers, developers, and civic organizations have been mobilized to analyze and visualize data for social purposes. For example, **DataKind** organizes *DataDives*, or "high energy, marathon-style events where mission-driven organizations work alongside teams of volunteer data scientists, developers, and designers to use data to gain insight into their programs, the communities they serve and more all in 48 hours or less."[92] In Mexico, the President's Office has organized several "**Public Challenges**" through the Open Data website of the country on issues such as climate change, health, and corruption.

[92] DataKind, What are DataDives? Available at: http://www.datakind.org/datadives



Source: *FixMyStreet's dashboard by MySociety*.

Among the strategies that can improve SFP's ability to make a better use of the reports received and take advantage of them to detect patterns or outbreaks of corruption, we have identified the following:

1. **Publish aggregated data in open formats.** The publication of SIDEC's database or at least certain parts of it can allow civil society organizations or research centers to generate independent studies about corruption challenges in the Federal Administration, especially if such participation is fostered through hackathons or prize-backed challenges to encourage use of the data. For example, New York City issues monthly reports about NYC 311 complaint hotline to the City Council, the Public Advocate, community boards, and the public, providing actionable insight into citizen concerns and people use it, especially universities. Opening the corpus of complaint data about food-borne illnesses provided Chicago with the supply of information it needed to create better algorithms to manage scarce enforcement resources and make how it conducts restaurant inspections more efficient and effective.[93] The most important customer for this open data is government itself. The ability to monitor SIDEC's data could help SFP improve and target its work and the public to hold it accountable. (For more about evidence-based policymaking and how governments are using their own data to improve service delivery, see GovLab's case studies on "data labs" here.)
2. **Generate a set of data-oriented insights to detect corruption**. Mining the complaint data could provide more accurate insights about corruption in the Federal Administration. At this stage, of course, the data is so little that this project will have to be evolved over time with the growth of the platform. The Dirección General de Denuncias e Investigaciones –the unit that manages SIDEC– has acknowledged that it does not analyze the reports received. For instance, simple calculations could be useful to detect patterns, outbreaks or networks of corruption. Eventually, with more

[93] Julian Spector, *Chicago Is Predicting Food Safety Violations. Why Aren't Other Cities?*, Atlantic City Lab (January 7, 2016), http://www.citylab.com/cityfixer/2016/01/chicago-is-predicting-food-safety-violations-why-arent-other-cities/422511/?imm_mid=0df22b&cmp=em-data-na-na-newsltr_20160120 [https://perma.cc/94VJ-JQ4L].

data, more sophisticated algorithms and modeling can be employed to spot trends. When combined with complaint data about corruption from Twitter and social media, often called sentiment analysis, government and civil society leaders have a rich source of information to mine for trends about corruption that can be used to target reform efforts (see, this article and p. 109 of Smarter Crowdsourcing against Zika, for more resources on using sentiment analysis to understand public perceptions). Those insights need to be translated into stories and visuals on an ongoing basis to enhance awareness. For more about the use of data analytics, see Implementation Brief 1.

3. **Generate a public dashboard within SIDEC**. To increase public relevance, SFP can design[94] and publish a new section within SIDEC that provides statistics about the reports received, as well as the actions taken by the SFP to investigate them. Depending on the information to be displayed, a dashboard could generate incentives for more users to report and could also increase accountability from the units responsible for investigating the reports. The US federal government's Consumer Financial Protection Bureau shares its database of over a million consumer complaints as downloadable open data as well as sharing reports and visualizations.

ACTION ITEMS

2.1. Anonymize SIDEC's reports received and publish them as open data.

SFP's Dirección General de Denuncias e Investigaciones and the Unit on Information Technologies should clean and anonymize SIDEC's database consistent with Mexican privacy law to generate a version that is available to the public as open data. Additionally, taking a page from the Open311 playbook, which sets standards and provides an Application Programming Interface (API) for publishing municipal complaint data, SFP should insure that data is updated regularly and available in standard machine-readable formats for other software platforms to use it and shared with other international anti-corruption open data efforts, such as those promoted by the Open Data Charter.

2.2. Launch a public challenge to visualize SIDEC data.

Once SIDEC's database has been published as open data, SFP should consider launching a competition to create analytical models and visualizations that inform the public about the reports received, their status and follow-up (e.g,. total number or percentage of reports received by agency, region or type of misconduct). The competition could be organized using Mexico's Public Challenges platform and inviting civil society organizations and volunteers to participate. SFP could also provide an award to increase incentives for participation or commit to involve the winning team in the implementation of its proposal. There are multiple precedents for such prize-backed challenges in the public sector, including this one to visualize public expenditures or this one to visualize obesity data or IDB's 2015 cartoon contest to create

---

94 "Smarter Crowdsourcing Zika: Recommendations-Communication and Behavior Change", The GovLab, 2016. Note: Designing the dashboard or generating the analytic reports can be done using prize-backed challenges which have been used to solve needs ranging from specific technology gaps to community-level behavior change. These public challenges have become a popular method to drive rapid innovation and investment in specific issue areas by offering incentives (financial or non-financial) as rewards for solving a clearly defined problem. More details about designing prize-backed challenges is available in the recommendations made by the GovLab for using prize-backed challenges as an incentive for public behavior change in the effort to combat the Zika epidemic.

an image that embodies the fight against corruption.[95] In July 2017, Transparencia Mexicana (with support from Mexico City, Bloomberg Associates, and the Open Contracting Partnership) launched a challenge entitled "Your City, Your Data" to encourage use of available open contracting data to realize operational improvements. SFP's Unit on Information Technologies can integrate winning visualizations into the website.

3. Implement a proactive communication strategy

Once the questionnaire has been readjusted and the public dashboard designed and published, SFP should implement a communication strategy to announce the improvements made to the platform and increase public awareness about SIDEC. According to SFP, there have been communications efforts but mainly targeted at government officials. Over 8,000 brochures have been handled in government offices and places like Mexico City's Airport, among other actions.[96]

ACTION ITEMS

3.1. Do Social Media Marketing

The Citizens Foundation in Iceland reports that to promote citizen engagement on their Better Reykjavik platform they "invested heavily in online advertising via Facebook and some with Google Adwords. Spending about 13,000 Euros they reached just about everyone in Iceland. They tested and tracked different civic messaging and know what led to wasted paid visits and what led to actual deep participation." Taking into account that it is critical to attract new users to report electronically, it could be useful to design a proactive communication strategy on social media in addition to the calling strategy discussed above. Both Facebook and Twitter have publicly expressed their commitment to social causes and have carried out initiatives to inform their users about specific issues in the past. For example, through the use of Twitter's API, SFP could identify users tweeting about corrupt misconducts in Mexico and proactively send an invitation to report it through SIDEC. Specifically, Twitter has launched **#Promote**, an online platform to submit marketing challenges. SFP could seek a partnership with Twitter and make use of this initiative to launch a challenge to connect users reporting corruption on Twitter with SIDEC as well as using Google's non-profit ad-words program to send targeted advertising about the platform to search engine users.

3.2. Announce the revised version of SIDEC

When the implementation of the previous activities is about to conclude, SFP should start preparing a communication strategy to announce what has been done to revamp SIDEC. Announcing achieved results yields higher impact than future results. The strategy should contemplate at least two objectives:

- Introduce the public to the new features and sections of SIDEC and invite people to report corrupt behaviors and misconducts of government officials using the online platform.

---

95 See IDB News Release "Colombian Cartoonist Wins Contest for Best Drawing against Corruption," http://www.iadb.org/en/news/news-releases/2015-12-07/cartoon-contest-on-corruption-in-latin-america,11352.html. The event was a joint effort of the Bank in partnership with the Colombia-based Gabriel Garcia Marquez Foundation for the New Ibero-American Journalism (FNPI). The contest was advertised mostly through the social media and in a regional seminar on Cartoons and Opinion Journalism, organized by the FNPI. By November 12, the Bank had received 206 cartoons from 18 of the IDB´s 26 borrowing member countries.

96 Interview with Jesus Antonio Suárez Hernández, Director General de Denuncias e Investigaciones, SFP. August 24th, 2017 in Mexico City.

- Support the invitation to report corruption through SIDEC with an energetic commitment of a top authority, such as the President or the head of SFP, to follow up every report received until its last consequences.

This can be achieved by the use of multimedia content, such as infographics and video, as well as by issuing an official communication which announces the launch of the revised version of SIDEC and the commitment to investigate corruption reports.

SUMMARY OF SOLUTIONS 01: REVAMP MEXICO'S ONLINE REPORTING PLATFORM TO INCREASE PUBLIC ENGAGEMENT AND IMPROVE ITS CAPACITY TO ADDRESS CORRUPTION	
MAIN ACTIVITES	1. Strengthen SIDEC's capacity to capture strategic data to detect corruption by revising the reporting questionnaire, documenting the workflow, testing user's experience and designing a proactive push mechanism for the site 2. Design and publish a new generation of data analytics for SIDEC through collaboration with civil society organizations and academia, including the use of prize-backed challenges 3. Design and implement a communication strategy to increase public awareness about SIDEC on social media
ESTIMATED LENGTH	▸ Activity 1 — 22 Weeks ▸ Activity 2 — 14 Weeks ▸ Activity 3 — 6 Weeks
RESOURCE INTENSITY	MODERATE INVESTMENT (50K - 100K STARTUP, 1 FULL-TIME HIRES)

ESTIMATED CRONOGRAM TO IMPLEMENT SOLUTION 01

II. SET UP A PERMANENT OPEN POLICY DISCUSSION PLATFORM ON ANTI-CORRUPTION AND PUBLIC INTEGRITY

A. THE CHALLENGE

The head of SFP has proposed, as part of her working plan, to **boost participation and collaboration with the public to design, review and assess public policies on corruption**. Since the current Secretary took office on October 27th, 2016, SFP has established several forums and multi-stakeholder mechanisms to review policies on public procurement, civil service, internal control, the implementation of the National Anti-Corruption System, among others. Without detracting from these efforts, there is room for consolidating them as permanent institutional features.

Currently, the multiple policy discussions are dispersed and fragmented and, beyond the information that can be found at SFP's website, it is hard to keep track of each conversation and its results. Throughout the world, several experiences with citizen engagement in the discussion, drafting and annotation of laws and policies – dubbed "crowdlaw" by the GovLab – can provide inspiration to **transform this institutional priority into an institutional practice**, through the use of technology. An online platform would not only centralize active policy discussions, but it would also broaden the spectrum of participants and keep an open record of the discussions, the suggestions received, the evidence provided, and the results achieved.

B. THE OPPORTUNITY

As a way to **establish a permanent channel to dialogue and collaborate with the public in the creation and monitoring of anti-corruption policies** at the Federal level, the Ministry of Public Administration (SFP) can develop a platform that:

1. Gathers in one single place the information and documents published on every federal anti-corruption regulation, program or policy, and **allows the public to review, assess, and discuss its content**.
2. Enables the public to **propose new policies or initiatives** to strengthen transparency, openness, and integrity at the Federal level, as well as to **introduce new evidence and data** that support anti-corruption policy discussion (e.g., indexes, surveys, etc.).
3. Disseminates information about **opportunities to involve and collaborate** with the Federal Public Administration in special events, forums and initiatives related to anti-corruption efforts (e.g., public challenges, hackathons).

A platform of this nature can be a complementary vehicle for the Ministry of Public Administration (SFP) to fulfill its new anti-corruption mandate, summarized as exerting internal control to prevent, detect and investigate corruption within the Federal Administration. By opening the existing anti-corruption instruments and policies – including their revised procedures – this platform could leverage SFP's independence to exert internal control over different agencies while gaining support and feedback from the public. Its effects could contribute to securing the correct application of the Law as well as regaining public trust in SFP.

In an effort to increase public support and to harness the insight of "the crowd," governments around the world have developed initiatives to collaborate with the public by means of this kind of platform. The

objectives, tools, and means of these initiatives are diverse. While some of them have managed to improve contact and communication between government officials and the public, others have tackled knowledge barriers, allowing new audiences to engage in law and policy design. Mexico's platform could incorporate learnings and experiences of at least three different participatory models[97] when developing its platform:

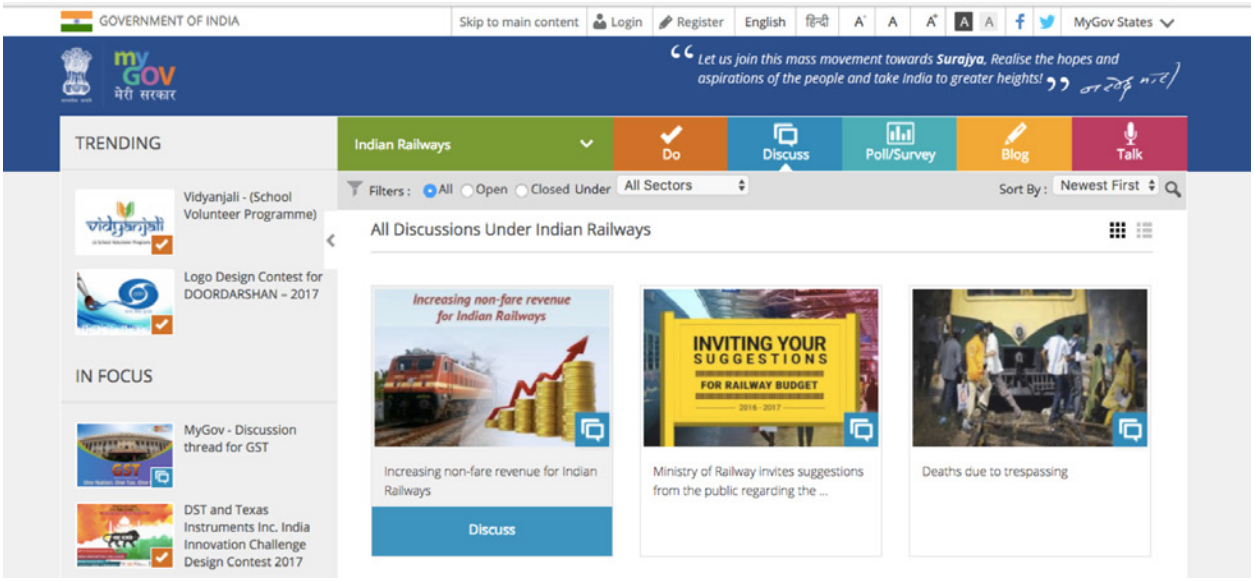
- **Participatory models for policy proposals** – These kinds of platforms seek to capture the public's knowledge and ideas on an issue. The results of such platforms are generally non-binding on public institutions, but can be useful for generating a large number of ideas. A platform may follow an "open call" model, where citizens propose ideas of their own initiative, or they may pose a specific issue for citizens to help address. **Decide Madrid**, an online portal launched in 2015, is an open call platform that allows Madrid residents to submit proposals for draft ordinances. When a proposal obtains the support[98] of 1% of Madrid's residents (approximately 27,000 endorsements) it is officially proposed to Madrid's Government and sent to a voting stage. If the proposal gets a favorable vote over certain period of time, Madrid's Government assumes the proposal and carries it out. The software behind the Decide Madrid platform is an open source project called Consul, openly available for reuse through the IDB's "Code for Development" initiative. On the other hand, harnessing citizen input may be more suited to a model like France's Parlement & Citoyens (and Parlement & Gouvernement), which set out specific issues for citizens to engage on together with public officials. Through the platforms, a public official or agency owns a three- to five-step process for defining and solving a specific issue. Citizens and public officials collaboratively define the problem to be addressed through policy or legislation, and then citizens propose solutions. A final report at the conclusion of the process highlights how citizen input was used and what future action will be taken on the issue.

- **Participatory models for policy discussion** – To ease communications and broaden opportunities for engagement in policy design, certain platforms have focused on generating an online space for the public to discuss the content of government legal documents, such as law bills or ordinances. In some cases, these platforms have become central for discussing critical policies or legislation. Recently, the Government of India launched **MyGov.in**, a platform for discussing policies across the different agencies. Although policy discussions can only be initiated by a government agency, a broad array of discussions has been opened. The platform has over four million registered users who have provided approximately 3.75 million comments in 728 discussions.

---

97 The following information is part of the report "Crowdlaw: Online Public Participation in Lawmaking", by The Governance Lab and the Governance Innovation Clinic at Yale Law School.

98 To learn more about how Madrid fostered an open government culture from zero using the Decide Madrid platform, visit [https://blogs.iadb.org/abierto-al-publico/2018/01/23/consul-ciudad-abierta-codigo-abierto/].



Example of policy discussions at MyGov.In

- **Participatory models for gathering data for policy design** – Discussion of complex public issues demands not only open forums but also solid evidence that can be taken into account to support or repel solutions. In the United Kingdom, the Parliament's Science and Technology Committee launched the Evidence Check Programme in 2016 which invites individuals and organizations to provide evidence underlying proposed policies on topics ranging from gender wage gaps to healthcare technology. This modest program has funneled academic and scientific evidence into specific policy discussions. These "checks," which exist in a simple forum model with uploading capabilities, create a pipeline between public institutions and private and academic researchers working in policy-relevant fields.

C. THE ACTION PLAN

1. Generate the first repository of active anti-corruption instruments available at the Federal level

The first stage in the implementation of this solution consists of gathering every regulation, policy or program that can be considered an anti-corruption instrument within the Federal Administration (e.g., codes of conducts, procurement regulations, hiring guidelines, etc.). This information will provide SFP with a clear understanding of the type of anti-corruption instruments available at the Federal level. SFP's legal unit can lead this effort. To ease the work, it can conduct an online survey to ask other legal units within the Federal Government about anti-corruption regulations, policies or programs currently in place in each agency. Afterwards, it should organize and categorize the information gathered. Such classification will be the basis for structuring the platform and its content.

2. Design, develop, and test a platform to access and discuss anti-corruption instruments of the Federal Government

The second stage of this solution consists of designing and developing the platform itself, with its three core sections:

- **Repository of anti-corruption legal instruments.** This section should allow the user to access and download the different anti-corruption regulations, policies and programs that are active at the Federal level. The availability of a search engine could improve the user's experience.

- **Mechanism for proposing and discussing specific policies**. This section should allow any user to submit ideas and suggestions to create new policies or discuss those already in existence. It should allow the ability to register discussions among several users on the same proposal.

- **Mechanism for submitting new evidence on corruption.** This section should allow any user to submit recent evidence about corruption in Mexico or globally, such as indices, surveys, databases, or reports. This information can be structured to highlight those reports with more relevance for the work done by SFP.

SFP should evaluate whether its internal unit on ICT has the ability to develop the platform by itself or if it should utilize the expertise of other areas of the Mexican Government[99] or engage external firms. In the case of MyGov.In, the platform was developed by the ICT unit of the Indian Government. Also, SFP can take advantage of software available for free under open licenses, such as **Consul** which is the one that was used by the Government of Madrid to develop Decide Madrid.[100]

Once the platform has been designed and developed, SFP can organize a focus group with key civil society leaders and members of the National Anti-Corruption System. The comments received during focus group meetings should be taken into consideration in the adjustment of the platform.

3. Launch the platform along with an engagement strategy

In addition to a presentation event, the launch of the platform should include an accompanying series of online discussions on key policies to activate the platform. SFP can invite target users –such as scholars or civil society organizations– to review and discuss a set of anti-corruption policies or programs. For example, there could be a challenge for reviewing public procurement regulation, the federal government's code of conduct, or civil service policies. SFP can also invite other agencies to use the platform, including the ones that are part of Mexico's National Anti-Corruption System.

---

99 It has been identified that open data platforms have been developed internally by the National Digital Strategy Unit at the Office of the President of Mexico and Infotec.

100 The IDB is actively encouraging public entities to take advantage of open software including Consul through its "Code for Development" initiative. See https://code.iadb.org/en

SUMMARY OF SOLUTIONS 02: SET UP A PERMANENT OPEN POLICY DISCUSSION PLATFORM ON ANTI-CORRUPTION AND PUBLIC INTEGRITY	
MAIN ACTIVITES	1. Generate the First Repository of Active Anti-Corruption Instruments Available at the Federal Level 2. Design, Develop, and Test a Platform to Access and Discuss Anti-Corruption Instruments of the Federal Government 3. Launch the Platform Along With an Engagement Strategy
ESTIMATED LENGTH	▸ Activity 1 – 20 Business Days (4 Weeks) ▸ Activity 2 – 40 Business Days (8 Weeks) ▸ Activity 3 – 60 Business Days (12 Weeks)
RESOURCE INTENSITY	SOME INVESTMENT (10K - 100K STARTUP, 0-1 FULL-TIME HIRES)

ESTIMATED CRONOGRAM TO IMPLEMENT SOLUTION 02

ISSUE AREA 4

Creating a Culture of Whistleblowing through Website Enhancements and Training

Whistleblowing is a critical input for prosecutors and other investigative authorities to build strong anti-corruption cases. As part of landmark legislative anti-corruption reforms, Mexico has recently enacted the General Law on Administrative Responsibilities, which mandates the creation of "the investigating authorities will establish areas of easy access, so that any interested party can submit complaints for alleged administrative offenses, in accordance with the established criteria."

But according to a 2014 report by the Blueprint for Free Speech, Mexico received the *lowest possible score as it relates to legal protections for whistleblowers*. To comply with the new law and translate its principles into practical changes that lead to a stronger culture of ethics within public and private organizations, Mexico urgently needs to take steps to protect whistleblowers and act on their disclosures.

This memo outlines the implementation plan for two initiatives designed to improve whistleblowing culture and policies in the public sector. Mexico should:

- **OPTIMIZE ONLINE REPORTING SYSTEMS WITH ENHANCED SECURITY AND ACCOUNTABILITY** – By improving existing platforms to protect anonymity and involve civil society, the federal government can make whistleblowing safer and easier

- **DEVELOP TRAINING FOR PUBLIC MANAGERS ON RESPONDING TO DISCLOSURES** – By training and creating incentives for public officials in management positions to follow up reports in a quick and effective manner, this training program can dramatically increase the number and quality of whistleblowers in the federal civil service

By relying on well-developed international resources and the help of accomplished experts, both projects can be accomplished in less than a year and result in an immediate five-fold increase in the number of whistleblowers coming forward to report on corruption in the public sector.

I. PROJECT BACKGROUND

On July 11th, 2017, 22 experts from 11 countries joined officials representing the Mexican government, the Inter-American Development Bank, and members of the GovLab, in a two-hour online conversation to identify novel strategies that the Mexican government and civil society leaders can use to protect and support whistleblowers.

The conference resulted in seven concrete recommendations from which la Secretaría de la Función Pública (SFP) selected two projects to flesh out into implementation plans. They were selected based on a combination of their potential importance, efficacy, and feasibility to undertake in 2017.

This plan contains:

- An overview of the problem
- An outline of each solution proposed, including
  - Overview and value proposition
  - Examples of existing initiatives
  - Driving questions and considerations to define the solution at a high level
  - Steps and resources needed to implement the solution
  - Metrics to evaluate the success of implementation

The content of the document, including its recommendations, is the sole responsibility of the GovLab and does not represent the IDB's official position or view on this matter, or an endorsement of any individual or firm to perform activities related to the recommendations.

II. PROBLEM OVERVIEW

See full problem brief here.

Effectively dismantling corruption networks and prosecuting their members requires a pipeline of promising leads for investigative authorities to review and take action upon. Although leads can come from predictive analysis of data, building cases often requires tips from those "on the inside," such as civil servants who witness the malfeasance of colleagues and supervisors or members of the public who are victims of corruption. Mexico's Manifesto on the National Anti-Corruption System called for the development of an effective whistleblower program that provides sufficient incentives and adequate protection mechanisms to enable the public to "exercise active social surveillance over the actions of public servants and dishonest individuals acting in a manner contrary to law." Additionally, Mexico recently enacted a new General Law on Administrative Responsibilities (LGRA) with a strengthened whistleblowing protection framework.

While the opportunity to capitalize on the political momentum spurred by these reforms is tremendous, Mexico's whistleblowing commitment has historically been very weak. According to a 2014 report by the Blueprint for Free Speech, Mexico received the *lowest possible score as it relates to legal protections for whistleblowers*. In particular, Mexico received low marks on anonymity, quality and availability of reporting

channels, availability of remedies, and transparency. In order to move from government commitment to effective implementation, Mexico must develop robust systems that create incentives for reporting and hold agencies accountable for follow up.

III. SOLUTION

Optimize online reporting systems with enhanced security and accountability

Value Proposition

The Dirección General de Denuncias e Investigaciones should improve SFP's existing online reporting system, known as SIDEC (Sistema Integral de Quejas y Denuncias Ciudadanas) for the specific purpose of increasing and responding to legitimate disclosures of corruption.

A critical challenge of any whistleblowing initiative is creating incentives for people to come forward. According to Dana Gold, Director of Education at the Government Accountability Project (GAP), "whistleblowers stay silent for two primary reasons – fear of reprisal and fear of futility."[101] Whistleblowers often take on tremendous risk in reporting – "they may be fired, sued, blacklisted, arrested, threatened, or in extreme cases, assaulted or killed."[102] Even when whistleblowers feel protected from retaliation, they may not make the effort to disclose information if they do not believe there will be actionable follow-up from their report.

Our driving hypothesis is that online reporting platforms designed to mitigate, if not overcome, fears of reprisal and futility will accelerate the identification of corruption and accelerate its prosecution. Effective online platforms offer several advantages to traditional in-person or telephone forms of reporting. First, online platforms **can be accessed anywhere, making reporting more convenient as well as more confidential** than visiting an office located in an organization's headquarters, for example. Online platforms can further assure confidentiality through anonymous web browsing software like TOR that **can more thoroughly protect the identity of whistleblowers** to prevent reprisal. Second, online platforms can develop secure communication channels that enable whistleblowers to **track their reports in real time**, offer supporting follow-up evidence and see if and how action is taken. Finally, online platforms can offer a **central place to connect whistleblowers to multiple anti-corruption agencies**, including government offices, investigative journalists and civil society organizations, increasing the reach and potential impact of their reporting.

For government agencies, online reporting may not only lead to an increase in critical disclosures for fighting corruption, but also has the potential to increase the **legitimacy and accountability of whistleblowing initiatives**. Online platforms can be designed to share reporting and communication channels with civil society organizations that citizens may be more likely to trust as third parties. Digitalizing reports also enables the government to have more easily accessible data on the filing and progress of reports, which can be published and tracked for accountability.

EXAMPLES OF EXISTING INITIATIVES

---

101 Interview with Dana Gold, Director of Education, Government Accountability Project on 16 August 2017

102 "Transparency International. International Principles for Whistleblowing Legislation." 2016: 2

Our recommendations for the changes Mexico needs to make to the SIDEC system draw upon lessons learned from global examples, including:

- **MéxicoLeaks (Mexico)** was founded in 2015 by civil society organizers to enable whistleblowers "at the service of Mexican society to reveal documents of public interest."[103] The platform allows users to upload documents anonymously relating to corruption and send them to civil society and journalist organizations. Once documents are received, they are checked for accuracy and interest, and cleaned to redact identifying information or metadata before they are published. MéxicoLeaks also publishes information through its Twitter account, which has 127,000 followers. The website demonstrates the importance of easy to use tools. At the same time, because MexicoLeaks is disconnected from official institutions, it has limited use as a whistleblowing platform as the only possible follow up is disclosure.
- **Curiamo La Corruzione (Italy)** was developed in 2016 as a partnership between GlobaLeaks (software provider), Transparency International (TI) Italy (civil society organization) and 10 public hospital systems in Italy to address whistleblowing in the healthcare sector – largely considered the "most corrupt sector" in Italy.[104] Users of the platform can fill out a detailed questionnaire about the corruption they witnessed and send it to the office responsible for corruption prevention in the hospital system where the incident occurred. Users can also opt to send the report to TI Italy. The system received between 30-40 detailed reports in its first year in operation and many additional hospital systems have approached TI Italy to develop a partnership on the platform.
- **Buzón Ético y de Buen Gobierno (Spain)** was developed in 2017 by the Oficina para la Transparencia y las Buenas Prácticas in Barcelona for the purpose of reporting actions of municipal actors that are "contrary to the principles of good governance, as a way to reinforce public management."[105] The platform is reserved for misconduct of municipal government actors that is contrary to laws, principles or ethics; general complaints and suggestions for improvements are collected on a separate form. The questionnaire allows users to describe the incident and upload supporting documents that get sent directly to the government agency charged with investigating whistleblowing reports in Barcelona.
- **Securities and Exchange Commission (SEC) Tips, Complaints and Referrals (United States)** was created by the SEC following the passage of the landmark 2010 Dodd-Frank Wall Street Reform and Consumer Protection Act, which mandated the creation of structures and incentives for increased whistleblowing protection and support. The portal provides a structured form, combined with open-ended responses to "enable members of the public to provide information to the SEC regarding possible violations of the federal securities laws.[106] The site also informs users about the SEC's unique whistleblower award program, which provide bounties for information that enables the SEC to prosecute with large monetary sanctions. According to Sean McKessey, former Chief of the SEC Office of the Whistleblower, this platform, combined with other reporting avenues, "led to 15,000 tips with over one billion dollars of fines levied against corrupt actors and institutions and $150 million paid out to whistleblowers."[107]

---

103 MéxicoLeaks

104 Interview with Davide Del Monte, Executive Director, Transparency International Italy. on 1st August, 2017.

105 "¿Qué es?", Buzón Ético y De Buen Gobierno, 19 Dec. 2016

106 MéxicoLeaks

107 Smarter Crowdsourcing: Anti Corruption Online Conference on Ensuring Whistleblower Support and Protection, July 11th, 2017

- **New York County District Attorney's (DA) Office WhatsApp reporting line (United States)** was launched to enable construction workers in Manhattan to report unsafe labor conditions confidentially to the DA's office. The office encourages workers to use the free WhatsApp messaging application to share "photos or videos depicting the conditions; photos of work permits; the names of the contractor, subcontractor, and supervisors on the project; the project address; and the date and time of any incidents."[108] The DA's office selected WhatsApp because it easily facilitates sending media like photos or videos which can be rich sources of evidence for prosecuting corruption. Additionally, the platform appeals to the DA's target whistleblower – construction workers who are often recent immigrants to the US and feel comfortable and safe communicating on the international and encrypted platform.[109]

QUESTIONS FOR CONSIDERATION

To design an effective whistleblowing platform (or a platform that can be replicated for multiple agencies), there are three main questions that you must consider:

- What kind of reporting will the platform solicit?
- Who are the users and what functions will they require from the platform?
- After the initial report, how do platforms enable secure ongoing, two-way communication?

The **first question** addresses what problem the organization is seeking to solve. This is essential as the scale and scope of the corruption being addressed will affect all other design considerations of the portal, including the target users and the information necessary to take action.

The **second question** focuses on who will use the platform and what functions are necessary for the platform to be accessible and actionable. This memo focuses on three main users – whistleblowers who report corruption, report receivers who must act on the information and respond to the whistleblower, and subsequent actors with the power to investigate and in some cases, prosecute. It is critical that all three types of users are identified and incorporated into the design process in order to increase reporting of and responding to legitimate acts of corruption.

Finally, the **third question** focuses on how reporters and receivers communicate once the disclosure has been sent through the platform. Having secure and transparent follow-up communication channels increases accountability by allowing the whistleblower to track their disclosure, while also enabling investigatory agencies to request critical additional information for taking action.

We now discuss each of these three questions in turn.

1. WHAT KIND OF REPORTING WILL THE PLATFORM SOLICIT?

The type of corruption that the government is seeking to identify and address through whistleblowing will drive the primary design considerations for the platform. Thus, the first consideration is: what kind of reporting does the platform seek to solicit and why?

There are three prototypical user scenarios of reporting. **The first such "use case"** is the generic collection of any type of citizen complaint or report. Such a "horizontal" platform is agnostic as to the nature of the disclosure. For example, Mexico's current platform SIDEC is designed to collect any report for over 30

---

108 New York Country District Attorney's Office, "DA Vance: Construction Company Foreman Convicted at Trial In Connection with Worker's Death"

109 Interview with Michael Sachs, Executive Assistant District Attorney and Chief of the Investigation Division, New York Country District Attorney's Office, August 3rd, 2017

issue areas across sectors, ranging from "being disrespectful when providing a service" to "diversion of public resources."[110]

The second is a "vertical" platform that focuses on issues within a specific sector (e.g., health services) or specific set of organizations (e.g., large financial institutions). In the UK, the Network for Police Monitoring developed a platform called Netpoleaks specifically for reporting oppressive policing and surveillance. Its form is custom-designed anticipating specific types of reports such as "policing of fracking protests" or "information on the government's Prevent 'counter-radicalisation' programme."[111]

Finally, a platform can be designed to attract corruption complaints or whistleblowing about a specific organization or office rather than on a specific subject. The global audit and tax firm KPMG has established its own company specific EthicsPoint portal for employees to surface issues of illegal, unethical or improper conduct within the company.

As mentioned, SIDEC currently functions as a horizontal platform. While this type of reporting is appropriate for some forms of accountability – for example, journalistic whistleblowing organizations that may be broadly seeking leads to investigate – the "one size fits all" interface does not enable the government to ask detailed questions. Thus, it is a reporting site without being a whistleblowing site that elicits the information needed to follow up on specific areas of corruption. What's needed is to go beyond the open-ended fill-in-the-box to offer a user experience tailored to anticipated use cases and designed to provide the security to the complainant and the information to the official to be able to take action.

By developing reporting platforms that are more tailored to a specific vertical issue, organizations can potentially receive better and more actionable leads. According to Fabio Pietrosanti, Founding Member and President of the Hermes Center for Transparency and Digital Human Rights, which developed GlobaLeaks software[112]:

> **When you are working in a specific vertical sector, you only want to receive information of higher quality. One important lesson learned in partnerships between GlobaLeaks and Transparency International Italy is the need to define very detailed questionnaires that mimic in a digital way the interview process. We start from the perspective that a motivated whistleblower will spend 10-15 minutes to fill out a questionnaire, giving to the whistleblowing initiative analyst or investigator the required information to tackle the case. It is a structured report of 40-50 questions, with information that would facilitate identifying fake reports and making the report actionable, especially considering that the resources [to read and follow up on reports] are often scarce. You need to have an electronic filter that reduces the noise and increases quality.**
>
> *— Fabio Pietrosanti, Founding Member and President of the Hermes Center for Transparency and Digital Human Rights*

110 SFP, "Registro de Queja o Denuncia"

111 Netpoleaks (link is a ".onion" web address and can only be accessed through a TOR browser)

112 Interview with Fabio Pietrosanti, President and Co-Founder, Hermes Center on 1st August, 2017.

La Dirección General de Denuncias e Investigaciones, in partnership with la Dirección General de Responsabilidades y Situación Patrimonial de la Secretaria de la Función Pública, should identify critical verticals within the constellation of anticorruption issues that could benefit from more targeted reporting platforms. Currently, the ten agencies that receive that largest number of reports through SIDEC are:

- IMSS – Instituto Mexicano del Seguro Social (8%)
- ISSSTE – Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (8%)
- SAT– Servicio de Administración Tributaria (6%)
- AFSEDF – Administración Federal de Servicios Educativos en el Distrito Federal (5%)
- SEP – Secretaría de Educación Pública (5%)
- PF – Policía Federal (5%)
- PGR – Procuraduría General de la República (4%)
- SEDENA – Secretaría de la Defensa Nacional (4%)
- INM – Instituto Nacional de Migración (3%)
- SAGARPA – Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (3%)

The government should adopt a two-pronged strategy and both develop reporting platforms for these verticals as well as testing whether the creation of a more targeted form increases reporting about agencies for which complaints are not currently received. In both case, the goal is to see whether customizing the form leads to a higher quality of complaint than before.

2. WHO ARE THE USERS AND WHAT FUNCTIONS DO THEY REQUIRE FROM THE PLATFORM?

Improving existing reporting platforms assumes that confidential, secure platforms with clear and transparent accountability mechanisms will increase the volume of and improve the response to legitimate disclosures of corruption. This theory requires three sets of actors to be involved as users of the system – those who have information to report, those who receive and sort the information for follow up, and those with the power and political will to take action on the information. Thus, the second consideration is: who are the target users for each category of actors and what do they need to successfully report and follow up on corruption?

Organizations should identify key users through stakeholder analyses and engage them in designing the system to meet their needs. At a high level:

- Actors reporting on corruption may be citizens, employees, relatives or accomplices in the corruption network. To create incentives for these users to report, the platform must assure confidentiality will be protected. This may be through secure browser technology such as TOR, which blocks the tracking of IP addresses, the option to report anonymously, or cyber security technology that assures data cannot be accessed, hacked or deleted. Government or civil society organizations can also form the platform in partnership with a third party who receives information to increase both perception and actual protection of confidentiality.

- Actors receiving information may be government organizations, civil society organizations or journalist organizations. For those who receive reports, the platform should allow the right users to read, download and sort information so that it can be easily accessed, analyzed and flagged for follow-up. Thus, any platform needs to route information to the right person.

- Actors who follow up on reports may be supervisors, government officials, civil society groups or journalists. The most important function these actors require is access to sufficient data to investigate, confirm, publish and potentially prosecute acts of corruption. But to ensure follow up, the website needs always to route, if not all information, then notice of the complaint to more than one party, preferably in more than one organization. According to Fabrio Pietrosanti, having multiple actors involved can benefit all parties: "when reports go to multiple actors, such as an NGO and to a public office, there is more trust and accountability. The whistleblower trusts the process more when they know there is an NGO involved and the public agency is also happy because they can share the burden of responsibility to take action."[113]

The table below outlines several different use cases that illustrate options for security, submission fields and recipients, that should be customized based on the needs of users:

PORTAL NAME & USE CASE	SECURITY	SUBMISSION FIELDS*	RECIPIENTS
MéxicoLeaks (Mexico) **Horizontal Investigative leaks**	▸ Only accessible through TOR browser ▸ Terms and conditions advise against disclosing personal info or submitting from work computer ▸ 16-digit code to access report after submission	▸ **Report Details:** 3 questions about the contents of submitted documents ▸ **Documents:** Area to upload documents ▸ **Terms**: Checkbox to accept terms	▸ User can select 1-8 recipients ▸ 3 civil society research orgs ▸ 5 media / journalist orgs
Buzón Ético y de Buen Gobierno (Spain) **Horizontal General Corruption Reporting for Municipal Govt**	▸ Option to access through TOR browser ▸ Advises not to disclose details in form that could be identifying ▸ 16-digit code to access report after submission	▸ **Report Details:** 2 questions on details regarding the unethical or unlawful behavior observed ▸ **Documents**: Area to upload documents ▸ **Identity:** Area to optionally provide name, passport number, email and phone ▸ **Terms:** Checkbox to accept terms	Automatically sent to 1 recipient – Oficina para la Transparencia y las Buenas Prácticas de Barcelona

[113] Smarter Crowdsourcing: Anti Corruption Online Conference on Ensuring Whistleblower Support and Protection, July 11th, 2017

PORTAL NAME & USE CASE	SECURITY	SUBMISSION FIELDS*	RECIPIENTS
Curiamo La Corruzione (Italy) **Vertical** **Health Sector Corruption Reporting**	▸ Option to access through TOR browser ▸ 16-digit code to access report after submission	▸ **Terms:** Checkbox to accept terms ▸ **Preliminary Info**: 4 questions on how user found tool, relationship to the org and experience reporting to the org (if applicable) ▸ **Identity:** Area to optionally provide name, gender, age, city, email, telephone, title ▸ **Report Details:** 8 questions on details and scope of incident, economic value of the incident, people involved, parties reported to, action taken and sources for verification ▸ **Documents**: Area to upload documents	Automatically sent to 1 recipient – the office responsible for corruption prevention within the selected hospital system Option to send civil society organization (Transparency International Italy)
KPMG (United States) **Company Specific**	▸ Only accessible through external EthicsPoint portal ▸ Code to access report after submission	▸ **Terms**: Checkbox to accept terms ▸ **Identity:** Area to optionally provide name, phone, email, best time for communication relation to company ▸ **Report Details:** 16 fields on people involved, details of incident, date of incident, length, witnesses, affected groups or people, client involvement, management involvement, reporting history, and willingness to engage in confidential call or chat ▸ **Documents:** Area to upload documents	Automatically sent to the Chief Compliance Officer of KPMG who assigns an investigatory team based on the nature of the request

** Listed in the order they appear on the respective portals*

3. AFTER THE INITIAL REPORT, HOW DO PLATFORMS ENABLE SECURE ONGOING, TWO-WAY COMMUNICATION?

In order to create the incentives for whistleblowers to utilize the platform as well as provide additional information to actors following up on their reports, platforms need to be able to have a communication mechanism between reporters and actors who follow up. Thus, the final consideration is: how can systems develop secure communication channels that enable whistleblowers to track their cases as well as agencies to follow up with requests for additional information?

One method, that has been used by a number of online platforms is issuing a 16-digit numerical key that allows whistleblowers to access their report, provide additional information, message with the receivers of the disclosure and track the progress of the case. This key ensures that the user can choose to remain fully anonymous throughout the investigation while still having a method to track follow up.

Currently, users of SIDEC receive a report number and a password to monitor the progress of their case. When users log back into the portal, they can view the details of what they submitted (narrative, date submitted, etc.), actions taken to investigate the report and status/result of the investigation. But two issues exist with the current system. First, in order to view the report status, users must submit an email address where they are mailed login information. If the user wishes to remain anonymous, this process entails creating a pseudonymous or anonymous email address but this creates an undue burden for the average user. Second, there is no way for users to update reports and no way for the agencies who receive reports to request follow up information or clarification. La Dirección General de Denuncias e Investigaciones should learn from the example of other successful whistleblowing and complaint websites to offer users the option to securely track their reports and communicate with investigators in both directions.

Furthermore, the filing of the complaint should be tracked by a third-party who provides oversight. Even if the contents of the complaint remain anonymous to protect against the harm of spurious complaints, every complaint should be given a tracking number and the correspondence on that matter tracked through to a successful resolution to ensure a timely response. This is a common feature on Customer Relationship Management (CRM) websites that track follow up on sales calls. It is not difficult to attach and openly publish dates associated with complaints to ensure timely response and follow up.

IMPLEMENTATION STEPS AND RESOURCES

Since what is required here, namely to 1) customize the complaint forms, 2) offer greater anonymity in reporting, and 3) ensure timely follow up by publishing case tracking information and case analytics are all enhancements to an existing site, the changes can be made with a small budget, 1-2 full-time employees and 6-10 months of time. This section will walk through the relevant steps and resources needed to implement this solution.

At a high level, optimizing a current online reporting system requires clarifying the key purpose of the system, conducting stakeholder analyses to diagnose current challenges, developing strategic partnerships to participate in the redesign and management of the new system and launching a communications strategy to engage the public in use. As with all change initiatives, the process is highly iterative, involving regular evaluation and refinement to continuously improve the platform.

Throughout this process, it is critical to regularly involve the intended users of the portal in the design and communication strategy. The biggest pitfall in designing a reporting portal is creating a system that a) does not adequately create incentives for whistleblowers to use it due to poor confidentiality or follow-up mechanisms; and/or b) does not enable investigatory agencies to receive the information they need to take action on legitimate reports.

Image 1: *Portal homepage for Barcelona's Ethics and Good Governance Mailbox (source: Buzón Ético y de Buen Gobierno)*

A second challenge, inherent to all whistleblowing initiatives, is making sure the government can adequately protect whistleblowers who use the platform. It is essential that throughout the reporting process, the government manages the risk of investigating or publishing leads which may have identifying information in order to protect whistleblowers from reprisal.

(Please note that Implementation Memo 3 on Citizen Engagement also addresses improvements to the SIDEC platform designed to improve general reporting. Although there is some invariable overlap between these two proposals, one of which focuses on increasing complaints, and the other specifically on whistleblowing, we distinguish between them in the event there is greater appetite for one than the other. Ideally, they should be read in tandem and both undertaken.)

1. Diagnose the main challenges with the existing platform in consultation with civil society and experts (6 weeks)

By consulting with end users and expert designers, identify the specific challenges that improvements to the reporting platform will address

Our recommendation is to focus on the three problems discussed above, namely the failure to get usable information, the failure to create a mechanism for anonymous reporting and the failure to follow up and take action in a timely fashion. In response, we recommend three improvements that differentiate SIDEC from best of breed sites elsewhere, in particular, customizing the forms to elicit information needed to transform SIDEC from a complaint website into a whistleblowing tool; bolstering effectiveness by

enhancing anonymity using third-party tools that protect against manipulation and abuse; and accelerating follow up by ensuring transparent reporting of analytics and information sharing with civil society.

a. In order to **refine the system to be better tailored to specific verticals**, evaluate the platform with input from the public and whistleblowing experts to determine what fields or features of the system could be improved to achieve increased reporting and more actionable responses. Questions may include: Do whistleblowers believe the platform is designed to confidentially facilitate action? Do the fields ask the right questions to provide relevant useful information for following up on whistleblowing? To improve legitimacy, this study should be conducted by a trustworthy third-party such as a university or a civil society organization and the results openly published.

b. Regarding **increasing the security of the system**, consult with digital security and encryption experts outside the government to diagnose how easily the platform can be compromised, how easily anonymity can be broken and the best ways to ensure communication while protecting privacy. Questions may include: How are the identities of whistleblowers protected? How are communication channels between the whistleblower and the government secured? How does the system protect against hacking, deleting or altering existing records currently? Especially in light of recent scandals involving government surveillance of civil society leaders, anonymity must be more than simply assured by the government, it has to be backed up with secure, third-party tools and techniques designed by civil society and the technology community.

c. Regarding **developing partnerships with civil society organization**, diagnose where third party partnerships are needed to make the project viable. Given the lack of trust in government, reliable outside partners need to participate throughout the re-design process to ensure that the site meets user needs. Third parties may need to play a role in making the system work. As discussed, we recommend ensuring that complaints are followed up and acted upon by publishing the date and time and nature of any complaints logged or even sharing those complaints with an outside ombudsman from civil society. But to secure outside partners requires documenting some of the basic information about current workflow and workload, such as how many whistleblowing reports the platform has received. What are the types of complaints that whistleblowers have submitted? What issues are not being reported? What types of users is the system engaging and what types of users are not accessing the system? Are the current agencies assigned to follow up the correct ones for providing actionable response?

2. Identify and draft plan for platform improvements (1-2 months)

Use research and partnerships to create a strategy for improving existing platform

Draft a strategy for improving the relevance, security and accountability of the reporting platform in collaboration with relevant partners.

It is also urgent to create a **clear and graphical explanation, including a video,** of the whistleblowers rights under current law and directions about how to use the site.

3. Identify and broker strategic partnerships (1-2 months)

Based on key areas of improvement, determine key partners for platform development and accountability

a. Identify strategic partnerships based on organizations that have convening power with key demographics, public credibility and/or authority or methods to follow up on whistleblowing reports. Examples include:
    i. Anti-corruption organizations (e.g., Transparency International)
    ii. Citizen engagement organizations
    iii. Other government agencies (e.g., Procuraduría General de la República)
    iv. Journalist / media organizations
b. Determine the terms and scope of the partnerships. This may range from co-leading changes in design to receiving copies of the existence of report or of reports themselves to having a formal role in following up on reports by taking on communication, investigation or support functions.
c. Draft and sign memoranda of understanding (MOUs) formalizing terms of partnership.

Then there are the usual things that must be done when designing any website. We will not describe these in detail here but will, instead, provide separate guidance and resources on tech development.

4. Design and implement changes to the portal (1-2 months)

Program the capabilities needed to make platform updates and changes

a. Source programmers internally or externally to create updates and changes to the portal
b. If the design and strategy team is separate from the developer team, institute regular communication and feedback loops to ensure that the improvements are translated accurately.
c. Consider hosting the servers for the website outside the country or outside the control of the central government to enhance the project's legitimacy.

5. Test platform (ongoing)

Ensure all new functions of the platform are working by asking users for feedback

Working with partner organizations, test platform to ensure that the platform is functioning correctly for whistleblowers and all receivers of disclosures. Ensure that the website offers a feedback form to enable users to suggest improvements and that there is a workflow process for making those improvements in a timely fashion.

6. Launch the platform (1 month)

Create a communication and outreach strategy for engaging target users

The President's communications office should work in tandem with relevant civil society organizations to organize a broad-scale campaign for the re-launch of the site. People need to be made aware of the website, what it does, how to use it and the outcomes. This might include a speech by the President or a blog post showing the President using the new site. Consider hosting a prize-backed challenge to create a new logo for the site or to design an infographic or a video to explain whistleblowing to the general public. The United States health ministry (HHS), for example, held a competition to create a 2-minute public service announcement to encourage hand-washing to fight disease, offering a $2500 prize and received

200 entries, including this one by the "rapping doctor" or the Department of Labor's prize-backed challenge to explain labor law to the public. The prize in that case was a meet and greet with senior government officials and no cash. Given the low levels of trust in government, the government needs to work closely with civil society to develop the strategy for launch. The communications strategy should include plans to celebrate actions taken by government officials to act on complaints and thereby extol and promote worthy civil servants rather than simply focusing on punishing the bad ones.

7. Evaluate portal and refine (ongoing)

Collect and publish metrics on platform outputs and outcomes and adjust design and engagement strategy as necessary

The platform should be regularly evaluated against progress toward combatting corruption. Based on the performance of the platform, government should continuously refine the platform and the strategy for engaging its target users. In order to build a whistleblowing culture, it is essential to share information about the project continuously, sharing with people, as the SEC does in the United States, for example, the stories of actions taken as a result of whistleblowers' complaints.

EVALUATION AND IMPACT

In the short term, metrics should focus on usage of the platform for the purpose of whistleblowing. They may include:

- Number of visitors
- Number of reports filed
- Diversity of agencies about whom they file reports
- Number of communications with whistleblowers
- Number of reports requiring follow-up
  - Of these, number of reports that are investigated and, if relevant, prosecuted.

Although in the first instance the object will simply be to double the number of people using the site in the first six months, the goal in the first year should be to increase the number of successful actions taken, including prosecutions as a result of disclosures via the website. Ideally, over time, those disclosures should increase and then, in the long run, decrease as the culture of whistleblowing gets stronger and leads to self-regulation and a decrease in corrupt behavior.

2. DEVELOP A PROGRAM OFFERING IN-PERSON HALF-DAY MANAGER TRAININGS ON RESPONDING TO DISCLOSURES IN GOVERNMENT AGENCIES

Value Proposition

> **There's an essential shift in organizations that has to happen around whistleblowing. We have to move from a 'speak up' culture that is ubiquitous to a 'listen up' culture, shifting the burden from employees to management.**
>
> *— Dana Gold, Director of Education at the Government Accountability*

The Coordinación General de Órganos de Vigilancia y Control should develop a program offering half-day in-person manager trainings on how to create an environment conducive to reporting and how to follow up on reports in a quick and effective manner. The training program should be both scalable to reach multiple government entities as well as customizable to address the specific structural and cultural issues within each agency.

According to Dr. Nerisa Dozo, business manager of the Whistling While They Work 2 project at Griffith University, 10+ years of research has indicated that mishandled reporting by managers is one of the largest obstacles to having an effective whistleblowing program. While managers sometimes ignore disclosures to protect their own interests, more often than not "they simply have no idea what to do and the report disappears into the ether."[114]

Our hypothesis is that by training middle managers on how to respond appropriately to disclosures, protect confidentiality of whistleblowers, flag the relevant investigatory agencies, and keep whistleblowers abreast of developments in their case, reporting and responding to instance of corruption will increase. Additionally, as SFP works with government agencies to understand and comply with the new LGRA, training programs can be designed to amplify other reform efforts around whistleblowing. The Inter-American Development Bank cites training as a key input to strengthening anticorruption projects: "Innovative tools are often complemented by the preparation of diagnostics, action plans and training, which serve to deepen and sharpen the overall impact."[115]

For organizations, having managers who effectively respond to whistleblowing can be "one of the best risk management systems" according to Dana Gold[116]. Managers are often the first and most accessible people for employees to report serious wrongdoing. "Employees on the ground are uniquely able to see serious forms of corruption – graft, violations of laws, rules, or regulations, gross mismanagement of funds, and issues that pose substantial danger to public health and safety. In an environment where managers establish safe cultures for whistleblowing, employees will be more likely to report and ultimately accept the results of the investigation if they experience a process that feels procedurally fair, Additionally, management benefits from having an early warning system for issues that have the potential to grow into much larger problems for the organization."[117]

QUESTIONS FOR CONSIDERATION

To design an effective management training program, there are three main questions that organizations must consider. We summarize and then discuss each of these in turn.

1. Which organizations should be the target audience for a managerial training program?
2. What are the major issues preventing managers from being effective at responding to whistleblowing?

---

114 Interview with Nerisa Dozo, Survey and Business Manager, Whistling While They Work 2, Griffith University on 2nd August, 2017

115 IDB, "Transparency Trust Fund (AAF) Annual Report 2015"

116 Interview with Dana Gold, Director of Education, Government Accountability Project on 16th August, 2017

117 Interview with Dana Gold, Director of Education, Government Accountability Project on 16th August, 2017

3. What additional supports or resources are necessary to make a training program effective?

The **first question** focuses on identifying the recipients of the training. This consideration is critical because agencies must have certain whistleblowing policies in place for a managerial training to be useful. The government may also select agencies based on how impactful an improved managerial response could have on reducing corruption.

The **second question** focuses on diagnosing the major issues with managerial response for each agency. Effective trainings are heavily customized and adapted to the context of the organization. Thus, it is essential for government to understand what systemic and cultural issues are affecting the ability of managers to protect and support whistleblowers.

Finally, the **third question** focuses on resources outside of the training that can support the impact of the program. This is a critical consideration as training can be greatly amplified when combined with other initiatives such as advice, resources or services to whistleblowers and managers.

1. WHICH ORGANIZATIONS SHOULD BE THE TARGET AUDIENCE FOR A MANAGERIAL TRAINING PROGRAM?

In order to maximize the impact of a training initiative, governments must identify organizations or agencies that are the best candidates to receive training and who within those agencies to train. Thus, the first question is: which organizations would be a good strategic fit as audiences for managerial whistleblowing training?

The first and most important consideration is that the organizations have existing legal and cultural infrastructure to support and protect whistleblowers. According to John Devitt, Executive Director of Transparency International (TI) Ireland, "Managerial training cannot exist in a vacuum. It's essential that it works in a holistic way to ensure that systems extraneous to the organization and inside the organization, as well as the culture externally and internally, are conducive to acting on concerns. You need to be aware of the risks associated with speaking up. It's unfair to put someone in a position where they won't get support from peers, where management won't support them. You could put that person in serious danger – it could end their career and affect their livelihood."[118]

The new LGRA general law mandates the creation of "adequate whistleblower and reporting systems, as well as disciplinary procedures for employees who contravene [a] company's policies or Mexican law" for public agencies. For this reason, the government should focus this initiative on public agencies as they are better positioned and have an incentive to take advantage of a managerial training compared to private or nonprofit organizations that have no such legal mandate. Also successfully modeling what to do can help model for the private and nonprofit sectors what to do and help them follow suit.

The government should also consider the scale of the training program. Carla Miller, President of City Ethics, who worked on developing a national training program for the government of Greece, advised

[118] Interview with John Devitt, Chief Executive, Transparency International Ireland on 29th August, 2017

starting small: "Trying to roll out a training program across many different regions at once is often setting yourself up for failure. Start by taking a mid-sized municipality or region and use it as a pilot to test the training."[119] We recommend starting with 1-2 of the agencies that have the most interaction with the public as one place to start.

Finally, in choosing a set of agencies to target, the government should develop a strategy to identify offices where managerial training would have the most impact on revealing and addressing corruption. For example, the government may choose to target agencies where corruption is most prevalent and increased incentives to disclose information is critical to reducing misconduct. Alternatively, the government could focus on agencies that have a larger percentage of new hires and new managers to have a stronger imprint on shaping and defining whistleblowing culture.

2. WHAT ARE THE MAJOR ISSUES PREVENTING MANAGERS FROM BEING EFFECTIVE AT RESPONDING TO WHISTLEBLOWING?

Effective trainings tailor their content to the specific context and needs of the target organization. Therefore, in order to design an effective training program, the second question to consider is: what are the major issues impacting managers' abilities to effectively respond to reporting?

As part of a training program, the government should develop and administer a survey to assess the whistleblowing systems and culture of the organizations targeted for training. They may also choose to outsource this assessment to a third party.

One program that has developed robust, free, research-based whistleblowing assessment surveys is Whistling While They Work 2, a project led by Griffith University's Centre for Governance & Public Policy. The first survey – the Survey of Organisational Processes and Procedures – identifies the whistleblowing processes that are in place and how robust those processes are across five support and protection components:

1. Incident reporting and tracking
2. Support strategy
3. Risk assessment
4. Dedicated support
5. Remediation[120]

Essentially, this initial diagnostic measures the organization's formal commitment to whistleblowing support and protection. The second survey – Integrity@WERQ – looks at an organization's more informal, cultural commitment to responding to whistleblowing. The survey assesses the strength of eight key indicators:

1. Aspects of the organization's ethical culture and climate
2. Standard of ethical leadership
3. Awareness and knowledge of reporting / 'speak up' processes

---

119 Interview with Carla Miller, President and Founder, CityEthics on 29th August, 2017

120 Brown, A J, and Sandra A Lawrence. "Strength of Organisational Whistleblowing Processes- Analysis from Australia & New Zealand" Griffith University, July 2017

4. Staff trust and confidence in the organization's responsiveness to concerns
5. Recent/current levels of reporting, silence or other action in response to actual wrongdoing concerns
6. Investigation rates and outcomes for internally reported concerns
7. Quality of responses to internally reported wrongdoing
8. Support, treatment and welfare of staff who have raised wrongdoing concerns[121]

Alternatively, Carla Miller has taken a behavioral ethics approach to diagnosing organizational culture. She argues that one of the foundational elements to understanding how to shift reporting norms is understanding and improving levels of hope: "Employees will only respond to training if they have any hope that their organization can change. There are tests that measure hope in an organization, and if you have low hope levels, you need to have programs to change that. Without assessing and addressing behavioral elements like hope, the foundations for building the whistleblower program isn't in place."[122]

3. WHAT ADDITIONAL SUPPORTS OR RESOURCES ARE NECESSARY TO MAKE A TRAINING PROGRAM EFFECTIVE?

In addition to training, there are other programs and resources organizations can offer to help improve the managerial response to whistleblowing. Thus, the third question is: what supports or resources outside of the training can amplify the impact of the program?

One model for thinking about this question is Transparency International Ireland's Integrity at Work (IAW) program. While managerial training is a core component of the program, IAW also offers resources and support that complement and enhance the training. These include: facilitating a commitment from the highest level of leadership to promote whistleblowing in the organization, promoting free legal advice centers or hotlines where whistleblowers can receive advice, providing assessments that measure use of systems and efficacy, and providing monitoring guidance to management on their whistleblowing response. Describes John Devitt: "We have established a helpline and free legal advice center for staff that raise concerns. We also provide support to staff as part of our Integrity at Work program who report that the organizational response was inadequate or that they suffered as a response of reporting, so we can draw the organization's attention to any systemic failures or individual failures in responding."[123]

In addition, it will be important to decide the level of commitment and resources that can be invested in creating training with high production value. In other words, trainings might be delivered in person or via video or might rely on interactive simulations and role-playing conducted via an immersive, massively multiplayer online environment such as a videogame. Online words are increasingly used to simulate workplace conditions and improve the quality and outcomes of training in professions ranging from firefighting to emergency medicine to the military. Virtual-reality based training offers the ability to control and simulate social conditions in a controlled environment. Its highly visual approach may improve employee engagement and learning outcomes at lower cost and at greater scale than live training.

---

121 "Integrity @WERQ.", Griffith University, 2017

122 Interview with Carla Miller, President and Founder, CityEthics on 29th August, 2017

123 Interview with John Devitt, Chief Executive, Transparency International Ireland on 29th August, 2017

While ideally all of these supports are in place to enhance the efficacy of a management training program, resource constraints may require that organizations prioritize which ones they implement. The government should also actively seek out partnerships with civil society organizations that may be providing these services for free or at subsidized costs.

IMPLEMENTATION STEPS AND RESOURCES

The Coordinación General de Órganos de Vigilancia y Control can develop a program that will provide an assessment of whistleblowing culture for 5-10 agencies and deliver a half-day training for 200 managers customized to address the specific issues of their agency with a medium budget, 1-2 full time employees and 4-9 months.[124]

At a high level, developing an effective managerial reporting training program requires for each agency: diagnosing issues with the current culture around whistleblowing, developing content that addresses the systemic challenges keeping managers from being effective, developing norms that embed a pro-whistleblowing culture into regular routines, and, finally, evaluating changes in whistleblowing culture and practice.

Throughout this process, it is critical to be aware of two major risks. **The first** is that training – particularly ethics and compliance training – is often viewed as a "box checking" activity by staff. Often this means that employees will, at best, tune out the training and, at worst, be actively resentful for wasting time. Thus, throughout the process it is essential for leadership to regularly communicate the importance of the training and make sure it is actually actionable for its audience.

The **second risk** is that the training is effective in increasing reporting but whistleblowers suffer reprisal due to the lack of organizational protections in place. Within a company, whistleblowers can face severe consequences for reporting malpractice, thus it is essential that organizations ensure that they can protect those that come forward.

1. **Diagnose the current state of whistleblowing in the targeted organizations (1-2 months)**

   Within each organization, designate an internal committee to lead an evaluation of whistleblowing culture and diagnose areas for improvement

   a. Name an "ethics officer" in each agency who is charged with leading the committee and overseeing the training program. Carla Miller sees this as an essential step: "You need someone who is dedicated to thinking about the ethics of the organization. It is crucial to developing structures that live on past the training."[125]

   b. Recruit a diverse committee for each target agency. Bill English, assistant professor of Strategy, Economics, Ethics, and Public Policy at McDonough School of Business explains: "Organizations need to make it a specific mandate of some task force or group to catalogue what we can improve on, catalogue problems and make recommendations on how to improve how things operate. There's a lot of insider knowledge with people who have been at an organization for many years and have a sense of what's dysfunctional, what's been a problem in the past, and what's a liability for the future. At the same time, newer people will have a different perspective, so it's useful to

---

124 This estimate is based on TI Ireland's staffing and budget model for their Integrity at Work program, which provides assessment, a one-day training and supporting resources to 20 government agencies with 10-20 managers per training per year with 6.5 people and a budget of about half a million Euros. It is scaled down from the TI Ireland model to function as a pilot.

125 Interview with Carla Miller, President and Founder, CityEthics on 29th August, 2017

have staff who are at different levels in the organization and have been there for different time frames be tasked with doing a systematic review."[126]

c. In partnership with the ethics officer and committee, conduct internally or contract with a firm to anonymously survey the organization's managers and employees. Possible survey/interview questions may include:
    i. Do managers know what the organization's whistleblowing policies are?
    ii. Do managers feel comfortable receiving disclosures?
    iii. Do managers understand the resources available to whistleblowers and do they feel prepared to share them with employees?
    iv. Do managers know how to follow up on whistleblowing by sharing reports with relevant parties or acting on them themselves?
    v. Do employees know they can disclose wrongdoing to their managers?
    vi. Do employees feel their managers are receptive to whistleblowing reporting?
    vii. For employees who have reported wrongdoing, how comfortable were they disclosing to their manager?
    viii. For employees who have reported wrongdoing, how satisfied were they with how their complaint was handled? Do they feel that the issue they raised was ultimately resolved?

2. **In partnership with whistleblowing experts, develop a library of managerial trainings on responding to disclosures (2-4 months)**

Design and implement a series of 30-90 minute modules based on regulatory requirements and general best practices as well content that addresses the specific issues identified in target agencies

Experts agree that there is no "one size fits all" training that can be applied across agencies. Legally, agencies may need to comply to different standards – for example, an agency that works with the environment may be subject to very different regulations than one that works with financial institutions. As previously mentioned, agnostic of agency function, organizational culture will also vary from office to office and heavily impact the type of training needed to create the desired behavioral changes. Therefore, a good training program will have a library of a series of short modules that can be adapted and pieced together to form longer trainings customized to each agency.

To develop the trainings, the government may want to consult with anticorruption training experts, like the International Anti-Corruption Agency, which provides tailored trainings on whistleblowing protection or the IDB's Office of Institutional Integrity which provides educational ethics training and support. The government may also opt to create the library in-house, utilizing free and open source tools such as Carla Miller's City Ethics website and examining different technologies for delivering those modules (text, video, audio, virtual reality).

While the types of modules that are developed will depend on the needs articulated by the assessment surveys, there may be certain trainings around whistleblowing regulation or culture that apply to many or all of the agencies. For example, in Massachusetts, Bill English helped develop a module outlining conflict of interest laws that over half a million government employees were mandated to learn. In Mexico's case, the Coordinación General de Órganos de Vigilancia y Control should identify what regulatory modules may be applicable to all managers, particularly in light of the new agency requirements outlined in the LGRA. In addition to general legal regulations, there may be general best practices for addressing organizational

126 Interview with Bill English, Assistant Professor of Strategy, Economics, Ethics, and Public Policy at McDonough School of Business on 25th August, 2017

culture that the government may want to develop modules for. Examples include how to talk about whistleblowing policies in organizations, what supportive responses look like, and how to track and manage follow up to reporting.

For these more general modules, the government may want to invest in creating a video that can be easily used to reach large numbers of people. Drawing on his experience with the Massachusetts government training, Bill English advised: "Especially with regulatory information, a video can be much more interactive, compelling and scalable than a PowerPoint. If you are strategic, you can boil down a 1-hour highly resented training into a powerful, useful 5-minute video that in a clever, thoughtful but also heart-gripping way illustrates certain pathologies of corruption and outlines standards, boundaries or rules of behavior."[127]

3. **Run the trainings (1 month)**

Develop and deliver tailored trainings for target agencies

Utilizing the menu of training modules developed in step 2, create customized half-day training modules for each of the target government agencies and determine who will be the trainer. According to Dana Gold an outside trainer is often ideal – "you need someone who is not part of the institution but has outside lived experience, has a degree of legitimacy and can bring a level of candor to the discussion."[128] Carla Miller agrees that external experts are essential to getting started, but also advocates that the ethics officer and committee should be involved in order to build in-house expertise to start developing internal programs. When running the training, there are a few general effective practices:

a. **Start with the "why".** According to Dana Gold: "You need to spend time up front explaining why compliance is important and through the lens of institutional corruption; this is not just about checking the box of ethics training; it's about the bigger picture – graft makes us ineffective, corruption undermines employee morale, it erodes public confidence."[129] Carla Miller also emphasizes: "You need to tie training to the purpose of the particular governmental unit and the mission and daily work that would be affected by corruption."[130]

b. **Emphasize the complexity of ethics.** Bill English explains, "When surveyed, most people will say this stuff is obvious and they would do the right thing. We know from actual research that that isn't the case." In his trainings, John Devitt takes time to discuss how, in a given scenario, people will utilize different ethical theories and make different decisions about what to do. "Some won't speak up unless they are given assurance about confidentiality. For others, even if you give them such assurances or opportunities, they won't speak up unless it makes sense to them from a moral perspective. Others will speak up if someone has broken rules. Others will internalize their decision – they will ask if they can live with themselves or sleep at night if they don't speak up. It's important to help people understand that colleagues are different and have different ethical drivers that affect their behavior."

c. **Present real-life situations and role play the managerial response.** Carla Miller believes role-playing is one of the most effective practices in management training: "Come up with a situation based on a problem you've discovered in the organization. Have an employee communicating it to their supervisor and then have that manager practice what they would say

---

127 Interview with Bill English, Assistant Professor of Strategy, Economics, Ethics, and Public Policy at McDonough School of Business on 25th August, 2017

128 Interview with Dana Gold, Director of Education, Government Accountability Project on 16th August, 2017

129 Interview with Dana Gold, Director of Education, Government Accountability Project on 16th August, 2017

130 Interview with Carla Miller, President and Founder, CityEthics on 29th August, 2017

in response. Oftentimes, people get tongue-tied, so it becomes an opportunity to coach and script it, to model what an effective response looks like."[131]

4. **Embed training takeaways into organizational structures (1-3 months)**

Help develop mechanisms within each agency that allow managers to utilize their training as part of organizational routines

The 70-20-10 model for training describes that only 10% of learning comes from formal education like training, with 20% coming from interactions with others and 70% coming from on-the-job experience.[132] Work with agencies to develop a strategy for embedding training takeaways into daily routines so that managers regularly reference and practice their learning. For example, organizations could make it a part of monthly manager-employee meetings for managers to inquire about ethical conduct in the workplace, as part of other routine check-ins.

5. **Evaluate training initiative and refine (ongoing)**

Collect metrics on outputs and outcomes and adjust strategy as necessary

Survey and interview data collected from the initial assessment of whistleblowing climate and culture can be used as a benchmark for assessing improvements as a result of a training policy. The organization should re-survey its staff every year to document progress and adjust training accordingly.

EVALUATION AND IMPACT

Every management training initiative should start with an assessment of the present whistleblowing culture to diagnose target training areas. Ideally, this survey serves as a baseline for whistleblowing culture and can be repeated on a regular basis to evaluate training impact.

Transparency International Ireland's Integrity at Work program is aimed at promoting positive cultural change in organizations and considers several attitude metrics to measure changes in whistleblowing culture:

- **Gap between employee and employer confidence** – According to John Devitt, typically 90% of employers feel confidence in the organization's ability to respond to concerns, while only 50% of employees share that confidence. Organizations should seek to measure and narrow that gap when initiating a training program.
- **Attitudes toward whistleblowing** – TI Ireland conducts word association tests through prompted and unprompted questions. For example, they ask staff to describe words they associate with "whistleblower" and document whether there are positive attitudes (for example "hero" or "conscientious") or negative attitudes (for example "snitch" or "traitor"). Over time, organizations can see if the proportion of positive to negative attitudes shifts.[133]

Carla Miller also advocates measuring changes in attitudes, particularly the level of hopefulness that employees have regarding the organization's ability to address wrongdoing. Whatever the metrics, Carla advises on keeping the evaluation simple – "start with five to eight questions that reveal attitudes toward hopefulness, fairness and fear in reporting. Measure changes over a year, publish your findings and most importantly have a conversation about it with staff to see how to improve."[134]

---

131 Interview with Carla Miller, President and Founder, CityEthics on 29th August, 2017

132 The 70:20:10 Model for Learning and Development

133 Interview with John Devitt, Chief Executive, Transparency International Ireland on 29th August, 2017

134 Interview with Carla Miller, President and Founder, CityEthics on 29th August, 2017

ISSUE AREA 5

Mechanisms and tools to support the effective prosecution of corruption cases

Improving Mexico's capacity to prosecute corruption[135] will require more than legal changes, it demands a sustained investment in strengthening the capacity of prosecutors and anti-corruption prosecutors, in particular, to investigate, coordinate, and prosecute cases effectively and efficiently. To support this goal, this implementation plan develops two practical recommendations to support the effective and accountable prosecution of corruption cases and the reduction of impunity in the prosecution of all cases:

1. **Build an "Open Justice" community of practice of anti-corruption prosecutors supported by an online learning platform**
   To support the new generation of anti-corruption prosecutors, create a prosecutors online learning network that enables them to learn and discuss best practices, as well as to access historical data about corruption investigations and review it in order to build more solid cases for future investigations.
2. **Create a website to enable the public to monitor the progress of corruption prosecutions**
   Design and publish an online platform to release key open data for the public to follow up the development of current corruption prosecutions and alert prosecutors on upcoming deadlines and major events.

---

135 Exporting corruption: Progress report 2015 (2015), Transparency International. Available at: https://www.transparency.org/exporting_corruption/Mexico

These two projects, which build upon well-documented international examples and best practices, involve straightforward steps towards improving the abilities of prosecuting authorities to investigate corruption crimes, based on openness, transparency, and collaboration.

1. BUILD AN "OPEN JUSTICE" COMMUNITY OF PRACTICE OF ANTI-CORRUPTION PROSECUTORS SUPPORTED BY AN ONLINE LEARNING PLATFORM

A. THE CHALLENGE

With the goal of turning around the prevailing levels of impunity, Mexico has undertaken different reforms and measures over the last years. Specifically for prosecuting corruption, in 2013 the Constitution was amended to create an autonomous and specialized Anti-Corruption Prosecutor's Office, both at the federal level and in each of the 32 states of the country. The reform aimed at creating independent bodies to investigate corruption crimes with increased efficiency and effectiveness. However, bringing these legal changes to fruition has turned out to be a rougher path than expected.

Prosecuting bodies in Mexico, including anti-corruption prosecutors, are lacking the primary resources and skills to build robust investigations. For instance, while the General Prosecutor's Office (PGR) opened 63,000 cases between November 2014 and December 2016, only 34.4% of them (21,728 cases) resulted in a prosecutorial decision (i.e., making a legal decision regarding pursuing further investigation or presenting charges). The rest are still pending decision.[136] Moreover, a report by the Mexican Judiciary revealed that 85% of cases made by PGR and presented before a judge were for crimes committed in flagrante and not due to an investigatory process.

Several organizations have highlighted that at the core of all of these issues is a lack of training and resources to support a robust investigatory process. After reviewing 811 files for robbery and homicide crimes committed between 2005 and 2015, the Mexican Institute for Competitiveness (IMCO) found that bungled investigations derailed 60% of these cases.[137] Of particular importance, IMCO concluded that prosecutors frequently gave imprecise instructions to police investigators; that they did not adequately review evidence gathered by police and experts or request follow up to inadequate information; that they omitted key steps in directing investigations; and that they generally were "passive" in conducting investigations, leading to prolonged periods of inaction.[138]

The creation of specialized and multiple anti-corruption prosecutors offices has been recognized as a step in the right direction. However, Mexico still needs to invest heavily in building the capacities that these bodies need to manage investigations and support them with robust evidence. The scarcity of resources will demand an investment in mechanisms and tools that can provide better and more effective training, as well as a more efficient transfer of knowledge. The use of technology platforms have not only been useful

---

136 Sin resolver 7 de cada 10 casos en nuevo sistema por incapacidad para investigar, admite la PGR (July 12th, 2017) Animal Político. Available at: http://www.animalpolitico.com/2017/07/nuevo-sistema-incapacidad-investigar-pgr/

137 "The Crisis of Impunity in the Mexican Criminal Justice System", Hernandez A., Mexican Institute for Competitiveness, January 2017.

138 Ibid.

to generate repositories of information, but also to build networks of practitioners and reduce the cost of exchanging ideas and best practices.

B. THE OPPORTUNITY

The new Anti-corruption Prosecutor's Offices should explore and leverage technologies of collaborative expertise – platforms and methodologies that facilitate skill-sharing, knowledge transfer and capacity-building – to create an online community of practice. Imagine if state or federal prosecutors and their teams had access, not only to reports and research, but also to a talent bank of individuals with relevant experience in the latest investigative techniques, the uses of data science to spot fraud, legal know-how about winning arguments, strategies for tapping into social media to build an evidentiary cases, knowledge of open data, ideas for how to collaborate with university researchers, and experience in bringing similar cases. In particular, there is an absence of personnel in these administrative and legal institutions with technical knowledge let alone the use of predictive analytics or data modeling. To strengthen capacity in the prosecutors' offices, there needs to be a way for people to share data, facts, opinions, advice, and insights about their experiences trying new approaches, what works and what doesn't. There is a tremendous amount of local innovation – pockets of experimentation and data collection and use taking place – but unless people know about what others are doing those experiments will never scale.

Such "conversational" infrastructure to enable professionals to tap the best know-how on an ongoing basis to undertake the difficult work of delivering criminal justice administration and services are making it possible to make more searchable what people know and to match the supply of expertise to the demand for it in government.

Identifying who knows what inside government — let alone outside — is hard. Within the bureaucracy, titles such as "prosecutor" and "assistant prosecutor" disclose little about skills and lived experience. Government projects, especially legal cases, are branded as institutional achievements, making it difficult to identify individual contributions. There are no well-developed taxonomies to describe human capital in new domains such as data science, which are increasingly relevant to the practice of law.

However, such a skill-sharing network can be the key to providing support among and between anti-corruption prosecutors and their teams. In particular, technology can help to supplement infrequent meetings and conferences by:

1. Providing a place for people to ask and answer one another's questions.
2. Enabling better peer-to-peer training.
3. Making it simple to tap distributed expertise across jurisdictions.
4. Finding those with relevant skills, especially tech and data science skills, who might be hard to find simply based on title or organizational affiliation.
5. Identifying what people know and want to know more about to enable the design of better training programs that match those needs.
6. Allowing for breaking out of silos and fostering collaboration across the agencies involved in the fight against corruption.
7. Connecting talented individuals across institutions to support one another and share and discuss similar challenges under the implementation of the new anti-corruption legal framework.

There is possibly no more important public issue today than how to evolve our governing organizations to make them better able to recognize and implement innovative and effective ways of solving problems. Policymakers can and will continue to debate which policy is better and more likely to lead to a desired outcome. But this may be of secondary importance to reinventing the processes by which we make policy in the first place to take account of what people know. No previous generation of humanity has had to tackle the myriad and complex challenges — such as eradicating corruption —society will confront over the next decades. To succeed, we have to run our institutions differently. Instead of governing behind closed doors, governments must learn to make use of the full gamut of expertise — formal scientific knowledge and credentials, but also skills, experiences and craft know-how — to solve problems.

Using technology in ways that marshal human capital more effectively and help to create a culture of continuous improvement needed to support the growth of a professional anti-corruption prosecutorial function. Additionally, and in the case of Mexico, this kind of network could be useful to **accelerate the processes under which public officials can meet their responsibilities** outlined by the recent reform that creates the National Anti-corruption System by connecting and allowing government officials facing similar challenges to identify potential solutions.

Such a knowledge network would build on existing global examples. Examples of existing global initiatives:

- **The World Bank: Skillfinder** - The World Bank's SkillFinder tracks expertise and experience across three dimensions: technical expertise, geography and clients, and business processes. Technical expertise includes items such as primary and secondary specialization, languages spoken, and publications. Geography and clients shows where someone has worked, with whom, and toward what end. Business processes includes products, project life-cycle, activities and roles. Through this pilot, the World Bank is testing if the ability to target and pinpoint expertise and to match people to problems based on their skills makes a difference for projects such as a water quality study in Punjab, India, or an assessment of access to sanitation in low income communities in Tanzania.

- **Aristotle** - Developed by the Air Force Research Lab (AFRL) and launched in the fall of 2008, Aristotle was a Department of Defense expert network aimed at increasing expertise awareness across the Department's agencies. The key challenge Aristotle wanted to address was finding a single place to consolidate the expertise within the Department of Defense. Essentially, the DOD viewed the platform as a tool that would provide more information about "What you need to know, when, and how you need to know it." Before the program concluded in 2013, the Aristotle pilot had 12,000 active users and 26 million links.[139]

- **MAX.gov** - Developed by the US Office of Management and Budget (OMB), MAX.gov began modestly in 2007 to solve the problem of needing to pass budget information securely back and forth between OMB and agencies and to facilitate document collaboration during the budgeting process. Because it uses strong, two-factor authentication relying on the Common Access Card (CAC) and Personal Identity

---

139 Department of Defense: Aristotle. Making Skills and Know How More Searchable at the Department of Defense. The GovLab. February 10, 2016. Read full case-study here: http://www.thegovlab.org/static/files/smarterstate/aristotle.pdf

Verification (PIV) cards issued to every federal employee, MAX has evolved into the de facto government-wide collaboration platform. It is the one place that is both open to every federal agent (and authorized contractors) and closed to everyone else. At the same time, MAX's permissions system enables the creation of secure "enclaves," where designated outsiders can be invited without access to the rest of the platform. Hence it is a flexible, practical way to collaborate and share securely.[140]

- **EPA Skills Marketplace** - To remedy the problem of silos, the Environmental Protection Agency (EPA) Skills Marketplace program encourages managers to post projects and recruit volunteers to the end of fostering collaboration across the 12 headquarters offices and 10 regional offices of the agency. Employees have an incentive to sign up for projects; to learn about other parts of the organization; to find work assignments that match with their skills and professional development goals; and to increase the organization's productivity. In just over a year of agency-wide operation, managers have posted close to 400 projects with more than 70% of projects receiving applications from 627 employees, 388 of whom were selected to work on the projects.[141]

- **OECD Observatory for Public Sector Innovation** - Developed and maintained by the Organisation for Economic Cooperation and Development (OECD), the Observatory for Public Sector Innovation (OPSI) aims to be "your gateway to a library of innovation experiences from around the world" that is "tailored to your needs." The platform was launched in 2014 by the OECD's Public Governance Committee (PGC), which "helps countries strengthen their capacity to govern by improving policy-making systems and the performance of public institutions." The central aim of OPSI is to connect individuals to the expert knowledge and codified best practices of innovators from around the world.[142]

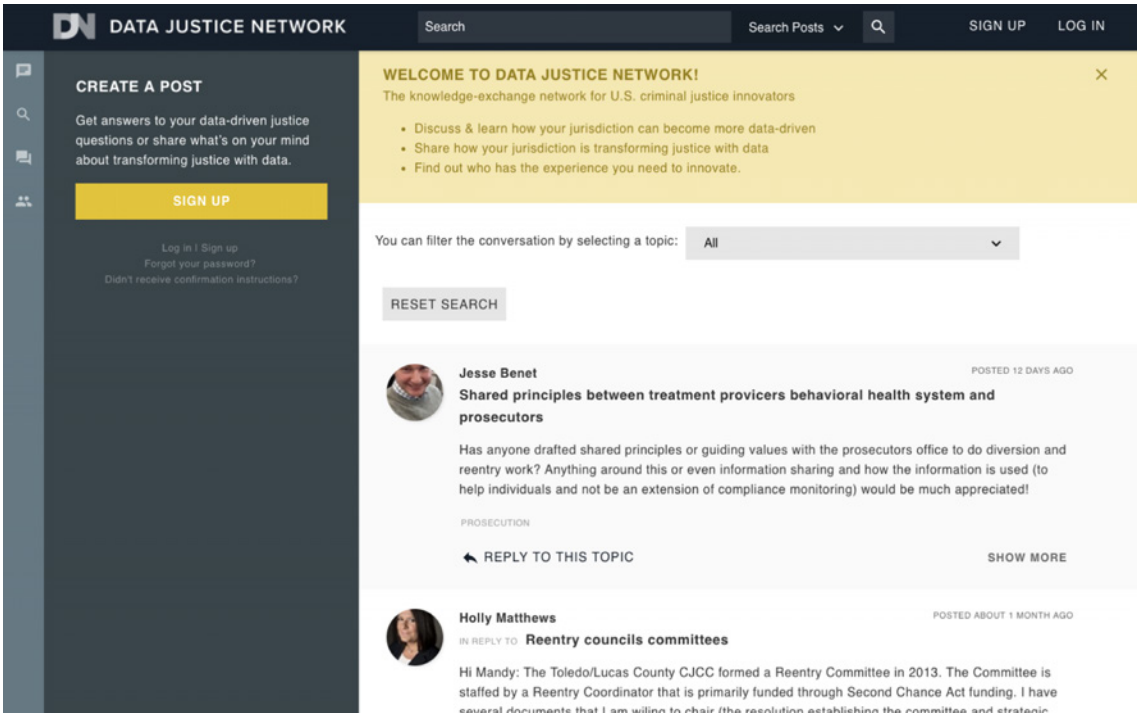
- **Data Justice Network** - Developed by The GovLab in collaboration with The Justice Management Institute using open source software, the website fosters peer-to-peer learning among criminal justice practitioners and policymakers and helps officials get fast and comprehensive answers to their questions about how to make better use of data to reduce incarceration and crime. Topics can be used to search for people with relevant experience. They describe the issue areas that people are working on and for which they seek knowledge, such as: Data Strategy, Regulatory & Legal, Data Policies, Standards & Platforms, Analytics & Visualization Recidivism, Community Corrections, Re-entry, Mental Health, Technology and more.

---

[140] Office of Management and Budget: Max.Gov. The government collaboration site whose robust authentication system enables working and sharing knowledge across agency boundaries. The GovLab. February 10, 2016. Read full case-study here: http://www.thegovlab.org/static/files/smarterstate/MAX.pdf

[141] Environmental Protection Agency: One EPA Skills Marketplace. The GovLab. February 10, 2016. Reaf full case-study here: http://www.thegovlab.org/static/files/smarterstate/epa.pdf

[142] Organisation for Economic Cooperation and Development (OECD): Observatory for Public Sector Innovation. The GovLab. February 10, 2016. Read full case-study here: http://www.thegovlab.org/static/files/smarterstate/OPSI.pdf



Source: *Data Justice Network, developed by The Governance Lab*

C. THE ACTION PLAN

1. Establish the purpose

Convene representatives of the community to agree on the mission: Accelerate corruption reform by connecting distributed practitioners with the knowledge of how to prosecute cases using innovative legal, technological and other approaches; enable people to learn from people rather than PDFs by asking direct questions about what to do and what not to do; create learning opportunities between conferences and workshops, which is how people learn today.

2. Decide on the target audience

Professionals working to prosecute corruption cases. Will the website be open to all prosecutors? Their staffs? Academics? Technologists? Will anyone be able to join or who will control invitations?

3. Document how people will use the site

Possible use cases:

- One state is trying to learn how to build a stronger investigative unit and wants to know what others have done,
- Another state wants to know what software to use to track the progress of cases and wants a recommendation.
- A federal prosecutor is struggling to know how to publish open data about case filings and is looking for advice.
- A prosecutor's office hires a person whose job it is to train people in the new legal framework and wants to know how to get that person up to speed.

4. Articulate key user benefits

- Avoid starting from scratch by learning from others
- Know what works, what doesn't and how to implement
- Learn from people, not just PDFs
- Build your network to help you solve problems

5. Convene potential users to agree on design principles

Document what people want. For example, what will the average user want the website to be. Make sure to ask the users, who may not be tech-savvy, rather than tech-professionals

- **EASY:** Must be easy to use, intuitive, simple, not overwhelming.
- **INFORMATION SHARING OPTIONAL**: Minimize need to fill out many fields
- **PRESERVE PRIVACY**: Enable private as well as public information sharing
- **MATCH PEOPLE**: Enable the ability to connect with people not just with content
- **ITERATIVE**: Take a phased and iterative approach

6. Find out the basic features people want and what's most important to them

For example:

- **PEOPLE:** User Profiles – ability to show what you know and what and how you want to be consulted
- **QUESTIONS: Q&A** – ability to ask, answer and search for questions
- **CONTENT:** Library - document libraries
- **SEARCH:** Search by topic, by person (including by any field in a personal profile, such as county size), by question, by keyword
- **COMMUNICATIONS:** sending private messages/emails or publicly responding to questions, private discussion groups

Take time to also find out what people don't want. For example, they might not want the site to be public but instead to have required logins and authentication of everyone on the website instead of an open platform. Such concerns should be taken into account in the design.

7. Decide on a roll-out plan

Determine what features are needed at the outset and which are optional and can be added later. Also develop a plan to get started with a small number of users who are willing to "kick the tires" and work out the problems with the site before rolling out to a wider audience. Keep in mind that lawyers are not used to interacting online and, thus, there must be no impediments to ease of use.

8. Other considerations

A. THE Q&A COMPONENT

Creating a place where people can ask and answer questions of one another is easily done with free, open source tools like Discourse. An example of such a Q&A site already in use to connect criminal justice professionals is DataJustice.US. Setting this up took a team of two a few weeks to design, build and launch.

But it requires deciding on a scheme for tagging and labelling content. Therefore, identify the priority topics: Investigative integrity? Data-driven investigation? Working with witnesses? Filing

cases? Tracking cases? Case strategy? Agree on and start with no more than ten tags/labels to make questions and answers easy to sort and search.

B. EXPERT NETWORKING COMPONENT

Beyond asking and answering questions, the platform would be improved with personal profiles that enable people to search for and find those with the experience they want.

Creating such an expert network requires first identifying the relevant taxonomies for describing human capital, including skills, experiences, credentials, and interests. Taxonomies can be borrowed by descriptors already in use by the bar association or among data science professionals, depending on the skills to be catalogued.

With the right schema for organizing data about what people know in this field, people's expertise becomes easily searchable.

Information can be captured using a simple questionnaire or mined automatically from existing sources of data.

The reason for collecting a great deal of rich personal information from individual participants as a condition for participation would be to make it easier to know whom to target with specific questions.

Rather than merely asking users for their credentials, the questionnaire would require members to complete a profile, describing in detail their relevant skills, experiences, and interests, which could include, for example, cases prosecuted or experience with specific software packages.

But the actual build-out of a profile would not have to rely exclusively on user contributions. Rather, people could correct profiles that were automatically generated using existing sources of online data, such as publications and case filings, and supplement it by completing the questionnaire.

Drawing on the insights of those who have studied recommender systems and “pyramiding,” the network or directory would also rely on third-party recommendations of who-knows-what about various topics. Third-party endorsements and recommendations could be automated, as is done on websites like Health Tap, where doctors endorse one another for their expertise. Twitter might be called upon to help identify those who talk a lot or are perceived by others to know a lot about a specific kind of policy based on discourse analysis of their social media activity. LinkedIn might share its know-how as well as its data to provide a further source of likely candidates for inclusion.

Anything that illuminates relevant expertise would be valuable to building upon a rich schema that accurately describes people and their profiles. The system would acquire data through a combination of self-reporting (a user filling out a profile), referral (a user suggesting someone else), and feeds (datasets coming from other directories or social sites).

Importantly, the system would not simply be a stand-alone directory — one more neglected database among many. It would be a directory of directories that enables federated searching of experts across other, existing networks of experts, such as the communities already in existence through interest groups like the bar association.

C. DECIDE ON FIELDS FOR THE PROFILES

My info

Name (e.g., Juan Alvarez)

Title (e.g., Senior Investigator)

Role (e.g., Investigator)

Organization (e.g., Police Department of XYZ/Prosecutor's Office of Chihuahua)

Location (city, state)
State – auto-populate
Integrate my LinkedIn profile with one click
Contact information – email, phone(s), text – check preferred mode of contact; if phone is checked, indicate preferred time of day

My Interests
My Office's Major cases – case name, summary, topics
What are the areas of expertise to be tracked?

Remember, instead of asking people to state simply that they are an expert, they might check in response: 1) My office has done this, 2) I can explain this topic, 3) I can refer you to others, 4) My jurisdiction is interested in exploring this. This design is in use on datajustice.us and networkofinnovators.org.

D. **CODE OF CONDUCT**
Members should be asked to sign a terms of service agreement requiring adherence to standards of conduct.

2. CREATE A WEBSITE TO ENABLE THE PUBLIC TO MONITOR THE PROGRESS OF CORRUPTION PROSECUTIONS

While underscoring the secrecy and confidentiality needed to conduct robust criminal corruption investigations, a certain degree of transparency and openness can be a powerful tool to bolster prosecutorial action. According to experts, the publication of specific information can create incentives for prosecutors to investigate, prevent undue political influence, and increase accountability.

The following section outlines the plan for implementing an 8-12 month project for a prosecuting authority – such as the General Prosecutor's Office or the future Special Anti-Corruption Attorney – to develop a mechanism to proactively inform the public about the status of corruption cases, while at the same time generate an internal mechanism to alert about future deadlines and major events in current investigations. The project itself is relatively fast to implement. However, we anticipate the need for an extended period to discuss and gain buy-in among stakeholders to undertake the project.

By implementing this proposal, the investigative authority should:

1. Prove its commitment to the independent investigation of corruption cases.
2. Increase public awareness about the legal actions carried out to prosecute corruption cases.
3. Prevent delays in investigations and guide top authorities in the prioritization of investigative efforts in corruption cases.

A. THE CHALLENGE

In the past, Mexico has struggled with the implementation of a major reform intended to transform and improve criminal procedure. Among several changes, the reform has introduced oral trials and other

measures designed to simplify and speed up the process while increasing openness and fairness in the criminal justice system. However, in reality, its implementation has represented a bigger challenge, making evident some of the core gaps and weaknesses in the system. Most of the investigative authorities throughout the country lag behind in its implementation and, more importantly, in achieving the expected outcomes.

In the case of Mexico's Attorney General (PGR), a recent internal report revealed that between November 2014 and December 2016 – under the new criminal system – **seven out of ten investigations never went to trial**. This was mainly due to lack of evidence or insufficient reason for further investigation. For cases that did enter the trial phase, 84% ended up in the same situation.[143] Aside from insufficient evidence, another major issue for PGR is the lack of capacity to deal with an increasing number of investigations. The same report highlights that while PGR opened 63,000 cases, only 34.4% of them (21,728 cases) resulted in a prosecutorial decision. The rest are still pending decision.[144]

This situation has severely affected public trust in PGR, especially in the perceptions about its commitment and ability to prosecute corruption. While it is true that addressing these problems will be a complex task and require multiple and coordinated actions, transparency and openness can offer a simple alternative to trigger change within PGR, as well as to turn around negative perceptions. There are global experiences that offer a solid basis for planning a strategy of such nature. Additionally, there are certain precedents within PGR that indicate it may be possible to carry it out. For instance, within the current Open Government Partnership's action plan, PGR committed to release certain information about investigations under the name "Carpeta Abierta." Since then, PGR has already undertaken concrete actions to revise its internal management systems and analyze which information could be released to the public.[145] Moreover, PGR also committed in 2016 to release data about foreign bribery investigations, under the OECD Convention. The investigation of one case has already been published.[146]

B. THE OPPORTUNITY

"The affirmative case for expanding the use of analytics starts with recognizing the extraordinary power of prosecutors," writes law Professor Jason Kreag.[147] As a way to **establish a mechanism that proactively informs the public about the status of corruption investigations,** a corruption prosecuting authority – such as the PGR – can develop a platform that:

1. Publishes in one single place socially-relevant data about active investigations of corruption and allows the public to understand their status and the prosecutorial actions undertaken.
2. Provides aggregated statistics about the status of corruption cases, their progress, and the performance of the investigative units.

---

143 Sin resolver 7 de cada 10 casos en nuevo sistema por incapacidad para investigar, admite la PGR (July 12th, 2017) Animal Político. Available at: http://www.animalpolitico.com/2017/07/nuevo-sistema-incapacidad-investigar-pgr/

144 Ibid.

145 La política de transparencia y apertura de la PGR (2016). PGR. Available at: https://www.gob.mx/pgr/articulos/la-politica-de-transparencia-y-apertura-de-la-pgr?idiom=es

146 Cohecho Internacional, Investigaciones en Curso sobre el Cohecho Internacional, PRG. Available at: https://www.gob.mx/pgr/acciones-y-programas/cohecho-internacional

147 Jason Kreag, Prosecutorial Analytics, __ Wash. U. L. Rev. __ (2017).

3. Enables an alert to the General Prosecutor and other top positions about upcoming deadlines and critical events in the investigation of corruption cases (e.g., hearings or trials).

A platform of this nature could incorporate learnings and experiences of at least three different initiatives intended to monitor the evolution of corruption cases and inform the public. These initiatives offer unique insights about the goals set for justice monitoring platforms:

- In **Argentina, the Asociación Civil por la Igualdad y la Justicia (ACIJ)** recently launched the "Observatory of Corruption Causes" (**Observatorio de causas de corrupción**) which monitors the most relevant corruption cases of the country and their evolution, both during investigation and during court. The organization has been gathering information by following directly each criminal procedure or through the media the main legal events. When ACIJ has accessed public documents, they have also been published in the platform. Additionally, ACIJ has engaged in a dialogue with the Ministry of Justice and other institutions to start obtaining data automatically, through the use of technology such as an Application Programming Interface (API). It is expected that they will be able to increase the information available through this service.

- **Transparency International's chapter in Hungary** generated a national database on corruption cases and kept records – through the use of information published by the media – about the length of each of the stages of the prosecution process. The platform allowed users to monitor the length of corruption cases under investigation and track the average days of each stage of the criminal procedure. Unfortunately, the **platform** is no longer being updated due to the challenge of accessing official information.

- In the **United Kingdom**, different justice institutions have all agreed to release different statistics about key areas related to the criminal justice system, their performance and management. They provide "information for the latest 12 months (April 2016 to March 2017) with accompanying commentary, analysis and presentation of longer term trends."[148] Additionally, the Ministry of Justice has committed to publish a set of socially-relevant justice statistics under the name **Justice Data Lab**. This initiative provides a service to organizations that request it, who in turn receives tailored reports and data. The goal of this initiative is "to help organisations to assess the impact of their work on reducing reoffending" while also helping to "develop a collaborative understanding of effective rehabilitation."[149]

These experiences suggest that corruption investigative authorities in Mexico can revise their engagement strategy with the public, while transforming information into an input to hold accountable their investigators and authorities. The main references can be summarized as followed:

- **Provide socially-relevant data.** The public relevance of this kind of platforms is highly linked to the value of the data perceived by social actors. While robust investigations demand secrecy and confidentiality, it will be necessary that Mexico's investigative authorities identify data –such as description of the case, the names of the authorities involved, a timeline containing the main events or access to specific public legal documents– that allows the public to understand the status and progress

148 Criminal Justice System statistics quarterly: March 2017. Ministry of Justice. Published August, 2017. Availbale at: https://www.gov.uk/government/statistics/criminal-justice-system-statistics-quarterly-march-2017

149 Accessing the Justice Data Lab service. Ministry of Justice

of an investigation without compromising its outcome. Such understanding can be critical to trigger increased interest in the evolution of corruption crimes and generate political incentives to build robust cases.

- **Collaborate with key social actors.** To identify socially-relevant data, investigative authorities could benefit from establishing a dialogue with specialized organizations and academics with regard to the prosecution of corruption and the criminal justice system. These actors will not only provide parameters and references about the data that matters, but will also contribute in building trust and legitimacy around this effort. Mexico's investigative authorities should seek to build a partnership with local actors to understand their needs, address questions about sensitive information, and agree on the minimum data that should be published about corruption investigations.

- **Communicate proactively.** The lack of information in the public sphere likely raises questions about the work performed by any government institution. Transparency and openness have become two essential values in today's governance. Despite the efforts and mechanisms in place to secure the right of access to information, proactive disclosure of information can increase the credibility about the commitment of an authority to a specific cause. Mexico's investigative authorities could benefit from opening their work to the public to prove an active commitment to investigate and penalize corruption crimes, regardless of their political background.

C. THE ACTION PLAN

1. Map and analyze data generated along the criminal procedure that allows the public to monitor status and progress

The prosecuting authority should undertake an analysis to map the information, especially the meta-data, produced during the criminal process, from the moment an investigation starts to the end when a sentence is issued and confirmed. This exercise should be done by a joint team comprised of different units involved in the investigation of corruption crimes. It should also review which data are generated and archived through the use of electronic systems and which are only physically available. This exercise should conclude in a systematized record of the information and data that is produced during a criminal procedure, in its different possible scenarios. It should include a clear description of the characteristics of the data and information found. The analysis will be the basis for identifying which and when certain data can be published without affecting the result of an investigation, while allowing the public to track its development.

To complement this action, a similar exercise to the Justice Data Lab initiative can be undertaken. For example, the prosecuting authority could invite organizations and the public in general to submit a proposal of the data they would need about an active corruption investigation and publish a report on the main findings. The proposals received can be used to:

Match the data requested by social actors to the data that is generated by the prosecuting authority. This action will allow for identifying expectations and setting a benchmark should PGR consider publishing data on investigations.

Identify data that is not being generated but that is requested by social actors. This action will allow revisions if the prosecuting authority could generate such data by the use of statistical methods or anonymization techniques.

Address the expectations of the public and provide justification for safeguarding sensitive information and explaining when this information can be available to the public.

In Mexico, several organizations and think-tanks have already undertaken significant efforts to support the implementation of criminal reform. For instance, organizations like CIDAC have conducted research to assess the implementation of the reform, and also communicate to the public the main changes of the criminal procedure.

Source: *Summary of the new criminal procedure in Mexico. Proyecto Justicia, CIDAC (2016)*

2. Design and develop a platform for visualizing data on the status and progress of corruption investigations

Once a set of data to be published has been identified, the prosecuting authority should develop an interface that allows the public to (1) know the number of active investigations for corruption, (2) review their current status, and (3) access information that describes its progress since the investigations were opened. The prosecuting authority should evaluate whether its internal unit on ICT has the ability to develop the platform by itself or if it should utilize the expertise of other areas of the Mexican Government[150] or engage external firms.

[150] It has been identified that open data platforms have been developed internally by the National Digital Strategy Unit at the Office of the President of Mexico and Infotec.

Alternatively, the prosecuting authority could launch a public challenge to create and visualize a set of statistics based on the data that will be published. The competition should focus on creating visualizations and statistics that inform the public about the status and progress of corruption investigations. The competition could be organized using Mexico's Public Challenges platform and inviting civil society organizations and volunteers to participate. The prosecuting authority could also provide an award to increase incentives for participation or commit to involve the winning team in the development of its online platform.

3. **Develop an alert mechanism for identifying future events and preventing delays in investigations**

Based on the data that the prosecuting authority has identified, it is possible to develop a series of algorithms to alert the head prosecutor or attorney or top aides about future events or deadlines in the investigation of corruption crimes. The use of red-flags systems have become more popular, especially in areas such as public procurement. For example, these systems alert on potential risks or failures in terms of competition or corruption in public tenders. Similarly, a series of alerts could be programmed based on the available data and provide investigative authorities with early warnings to prevent delays in investigations and guide them in the prioritization of investigative efforts in corruption cases.

To do so, it will be necessary to match the legal deadlines and other time-sensitive requirements within the criminal procedure with the data that is generated by the prosecuting authority, and ideally kept in internal management systems. For this purpose, a multidisciplinary team of data scientists, mathematicians, and legal experts will have to be created.

4. **Launch the platform along with an engagement strategy**

The launch of the new platform can be an opportunity to renew the commitment of a prosecuting authority to the effective investigation of corruption crimes. The announcement should stress the effort to release data about active investigations and frame it as an invitation for the public to participate in holding investigative authorities accountable.

SUMMARY OF SOLUTIONS 02: DEVELOP A MECHANISM TO INFORM AND MONITOR THE DEVELOPMENT CURRENT CORRUPTION PROSECUTIONS	
MAIN ACTIVITES	1. Map and analyze data generated along the criminal procedure that allows the public to monitor status and progress 2. Design and develop a platform for visualizing data on the status and progress of corruption investigations 3. Develop an alert mechanism for identifying future events in and preventing delays in investigations 4. Launch the platform along with an engagement strategy
ESTIMATED LENGTH	▸ Activity 1 – 60 Business Days (12 Weeks) ▸ Activity 2 – 40 Business Days (8 Weeks) ▸ Activity 3 – 40 Business Days (8 Weeks) ▸ Activity 4 – 20 Business Days (4 Weeks)
RESOURCE INTENSITY	MODERATE INVESTMENT (100K - 500K STARTUP, 1-2 FULL-TIME HIRES, IMPLEMENTATION OF A SMALL PROGRAM/PILOT)

ESTIMATED CRONOGRAM TO IMPLEMENT SOLUTION 02

ISSUE AREA 6

Tracking and Analyzing Money Flows

Mexico has now replaced South Africa as the nation which is most anxious about the direction their country is heading. Only 8% in Mexico think their country is going in the right direction and the issue that worries Mexicans? Corruption. To assuage the public's **severe mistrust of the government and concern about corruption,**[151] it is urgent to improve the tracking and analyzing of money flows in public procurement, in particular. This memo makes the following recommendations, to be discussed at length, which could increase the integrity and the efficiency of public acquisitions:

1. **Develop financial incentives** to encourage voluntary disclosure of corporate beneficial ownership and other legal entity data.
2. **Map the demand for blockchain technology in the procurement process** and publish a primer on the implementation of blockchain in the procurement process as a first step toward exploring blockchain integration

These projects, which can be accomplished in under a year, are designed to enhance the quantity and quality of data about public procurement. They go in tandem with the projects proposed in Implementation Plan #1 on Measuring Corruption. That brief focuses on how to use data to spot and predict corruption. This brief focuses on how to create more of that data in the first place. Both point to the need for data science trained personnel to execute on these projects and significant changes to how government does business. Thus, pressure may need to be applied externally to drive through these changes.

---

151 This 2016 Ipsos survey showed *"corruption"* as the second biggest worry of the average Mexican

I. PROJECT BACKGROUND

On July 25th, 2017, 12 experts from 6 countries joined officials representing the Mexican government, the Inter-American Development Bank, and members of the GovLab,in a two-hour online discussion concerning mechanisms, tools and strategies that could help track and analyze money flows to detect, prevent and counter corruption in public procurement. This conference took place in the context of landmark legislative and institutional reforms that Mexico's National Anti-Corruption System[152] aims to implement. Crucial to the success of this ambitious reform is a set of **tools and strategies that can better track and analyze money flows related to public procurements to identify and prevent potential acts of corruption or collusion.**

The conversation focused on increasing the government's ability to respond effectively to corruption in public procurement by 1) identifying innovative tactics that would support the identification of the ultimate beneficial owner who controls the money flows; and (2) identifying novel instruments that take advantage of publicly available data to identify the corrupt flows of money.

The conference resulted in five concrete recommendations. This document provides the broad implementation outlines of the three recommendations selected by Mexico's Secretaría de la Función Pública. The content of the document, including its recommendations, is the sole responsibility of The GovLab and does not represent the IDB's official position or view on this matter, or an endorsement of any individual or firm to perform activities related to the recommendations.

II. PROBLEM OVERVIEW

In recent years, a series of major scandals has demonstrated the pervasive nature of corruption in Mexico across all levels of governance: local, state, and federal.[153] Mexico's National Anti-Corruption System[154] aims to implement legislative and institutional reforms to eradicate this pervasive corruption.

Public procurement is one of the largest areas of public spending across all levels of governance. Expert and public perceptions worldwide suggest that public procurement is the governmental function most vulnerable to corruption.[155] According to the World Bank, inefficiencies accounted for 10 - 14% of procurement budgets in Mexico in 2009, while a report by Americas Market Intelligence estimates that procurement abuse can account for 15 - 30% of an infrastructure project's total cost.[156] Countering corruption in public procurement will not only reduce the waste of public money, but can also prevent the purchase of goods and services that are unnecessary or poor quality.[157] It is therefore no coincidence that

---

152 A summary about Mexico's National Anti-Corruption System can be found here.

153 President Peña Nieto has come under scrutiny for his involvement in several controversial real estate transactions in both Miami and Mexico City. According to a report by the New York Times, local governments in Mexico are complicit in corruption schemes with drug cartels. The former Governor of the state of Veracruz, Javier Duarte, was recently arrested in Guatemala, after fleeing the country to avoid prosecution for allegedly embezzling hundreds of millions of dollars.

154 A summary about Mexico's National Anti-Corruption System can be found here.

155 Rose-Ackerman, Susan. "The challenge of poor governance and corruption." *Especial 1 DIREITO GV Law Review* (2005), pp. 207-266, available here.

156 This is not unique to Mexico. In Europe, an estimated 13% of the total value of public procurements is lost to corruption, amounting to a staggering 1.4 to 2.2 billion Euros in 2010.

157 Tanzi, Vito and Hamid Davoodi. "Corruption, public investment, and growth." *The Welfare State, Public Investment, and Growth*. 1998, pp. 41-60.

two of the six implementation plans written by the GovLab focus on countering corruption in the public procurement sector.

Recognizing these vulnerabilities, Mexico has undertaken several steps to shield its public procurement sector from corruption:

1. **Increase the transparency of the procurement process to allow for an easier identification of the ultimate beneficial owner[158] who controls the contracts and the money.**
   Mexico's Competition Commission committed to adopting and implementing the OECD Competition Committee's Guidelines for Fighting Bid Rigging in Public Procurement, thereby obliging procurement agencies to collect information on potential suppliers and market conditions to understand whether public money awarded to a variety of different contract winners ultimately flows to the same supplier. Mexico also committed, through the Open Contracting Partnership, to explore open contracting in the health procurement and energy sectors and committed to implement the Open Contracting Data Standard for the construction of Mexico City's new international airport. Finally, Mexico City published a contracting information primer on the planning, tendering, awarding, and implementation stages in standard data formats. Open contracting data informs the public as to which companies won the contracts and under which circumstances, and is the bedrock for further investigation into where this money ultimately flows.

2. **Increase the likelihood that suspicion of corrupt money flows are triggered, are reported, and lead to an investigation.**
   In 2014, after a lengthy vetting procedure, the Financial Action Task Force (FATF) deemed Mexico sufficiently compliant in the context of the 3rd Round of Mutual Evaluations. Mexico achieved this approval by showing progress in setting up sufficient customer due-diligence requirements to prevent anonymous or pseudonymous usage of the financial sector, and by showing an increase in the quantity of reports on suspicion of money laundering sent by obliged financial institutions to the Financial Intelligence Unit for analysis and investigation.[159]

Despite these measures undertaken by the Mexican Government with the support of international organizations, little tangible progress has been made. For example, while Mexico identifies the corporate beneficial owner, according to the Financial Secrecy Index of the Tax Justice Network (an independent research, analysis, and advocacy group working on international tax avoidance and evasion), **information on ultimate beneficial owners of companies is not recorded in the company's official documents.[160]** Furthermore, Mexico **does not require public, online posting of details about company ownership and**

---

[158] The literature on money laundering and corruption for the past 20 years touts financial tools that shield the identity of the ultimate beneficial owner as the main vehicle for criminal money. Anonymous corporations are considered to be the vehicle through which money illicitly flows back to the public servant from the bidding company, in the form of kickbacks. The Financial Action Task Force (FATF), the intergovernmental body which sets global anti-money laundering standards, defines a beneficial owner as *'the natural person(s) who ultimately owns or controls a customer', where reference to ultimate ownership or control refers to situations in which 'ownership/control is exercised through a chain of ownership'*. In this report we follow this definition.

[159] The 7th Follow-up on the Mutual Evaluation Report of the Financial Action Task Force, available here.

[160] Research shows that legal entities are required to maintain official documents attesting legal ownership but not necessarily documentation attesting the individuals who ultimately control the legal entity.

company accounts and **does not require companies to publish country-by-country financial reports.** Mexican authorities are well equipped to identify the legal person that controls the money but not the natural person.

Since 2015 Mexico has also been working on implementing the Legal Entity Identifier (LEI) system[161] - a unique identifier of legal entities that companies can apply for and that can show **which legal entity owns which other legal entity -** on all corporations operating in the public sector. Although integrating this identifier with the payment system and payment messages is thought to increase the transparency of the money flows in the public procurements sector, the LEI **does not contain information on private individuals owning shares.** While LEI may seem adequate to uncover corruption because it allows for a better monitoring of money flows in public procurements, it does not identify the ultimate beneficial owner of the money flows.[162]

According to the G20 Anti-Corruption Working Group of the World Bank, ultimate beneficial owners of companies that are incorporated in Mexico must be identified by Public Notaries and Public Brokers and disclosed in the Public Registry of Commerce (see Table A6.1 in Appendix VI). But the Working Group reveals that the registry does not have any sanctions or enforcement power to ensure that the data is correct, nor are their standards for timely incorporation of updates into the registry. Furthermore, the Public Registry of Commerce cannot be checked online, and information relating to the identity of the ultimate beneficial owner is not available to the general public.

Finally, as in many other countries, there are insufficient checks on the accuracy of the data in the Public Registry of Commerce, so a single corrupt public notary or financial entity is sufficient to hide the identity of a company's real ultimate beneficial owner. Although data-driven tools could identify the subtle links between the money flows, the predicate offence, and the ultimate beneficial owner, these have not been sufficiently applied.
For a full description of the problem, see the problem brief here.

III. SOLUTIONS

This implementation plan proposes that the Government of Mexico generate mechanisms to better identify, analyze, and monitor money flows, and better identify the ultimate beneficial owners who control the money that flows within the public acquisitions sector. To this end, Mexico can unlock the potential of open data in several ways:

1. Create incentives for companies engaging in public contracts to self-disclose the identity of the ultimate beneficial owners;
2. Test the applicability of blockchain technology as infrastructure for storing open contracting data and for conducting public procurements.

---

161 The LEI system was developed after the 2007-2008 Financial Crisis to ensure that members of the financial industry meet their Know-Your-Customer obligations in real time, and that regulators can enforce international standards.

162 In its 2016 Guidance on how to effectively conduct customer due diligence, FinCEN does not recommend the LEI as the UBO identifier. Their decision is based on several observations: the authorized bodies assigning the code do not require the beneficial owner to be a **natural person**, use a 50% ownership threshold, and do not verify the identities of beneficial owners of legal entities.

These recommendations have been developed in the context of the current political pressures to counter corruption, and are based on the latest research on criminal psychology as well as the latest developments in information technology. This memorandum lays out the three solutions in detail.

Identifying the ultimate beneficial owner of the money that flows through public procurements is considered the *"Holy Grail"* in every country in the fight against white-collar crime.[163] Having an open record of beneficial ownership answers the question of *"Where does this money come from and where is going to?"* Although forcing companies to disclose such information has been met with resistance, and some argue against the disclosure of beneficial ownership as an invasion of privacy, **positive incentives** may **promote self-disclosure** and focus the attention of the public and regulators on those companies that have chosen not to disclose.

Data-driven analytical techniques can substantially increase investigators' reach by adding an active process of uncovering abuse to the current passive existence of data. Currently, Mexican law enforcement and the Financial Intelligence Unit (FIU) have access to all the records of the Public Registry of Commerce (see Table A6.1). The FIU can investigate the registry and look for evidence of money laundering once it is tipped off, which typically happens through a suspicion report.[164] Such reports are the principal mechanism used by financial entities to signal their suspicion or knowledge of a client engaging in or facilitating money laundering, and can also be submitted by ordinary citizens. Applying data analysis to the datasets would allow, among other things: (1) identification of new suspicions;[165] (2) confirmation of existing suspicions; and (3) review of known "red flags" for corruption. Supporting structured analyses conducted by people who hold advanced statistical knowledge and social science knowledge can cut down greatly on false positives.

But Mexico has now replaced South Africa as the nation which is most anxious about the direction their country is heading. Only 8% in Mexico think their country is going in the right direction and the issue that worries Mexicans? Corruption. To assuage the public's **severe mistrust of the government and concern about corruption,**[166] introducing a blockchain infrastructure may significantly help.[167] Blockchain technology deploys a shared, synchronized, distributed ledger of transactions, guaranteeing privacy and security; integrity of data and increased trust by providing a permanent record of who accessed ledgers and what they did. Such Distributed Ledger Technologies (DLTs), including Blockchain, have disruptive potential beyond innovation in products, services, revenue streams and operating systems within industry.

Blockchain technology was popularized in 2008-2009 by Bitcoin, a payment system that functions on the blockchain infrastructure and that proposed the redesign of the financial system in the aftermath of the

---

163 At the 2014 Brisbane Summit, the G20 leaders adopted High-Level Principles on Beneficial Ownership Transparency, describing financial transparency as a "high priority" issue.

164 Mexico's Federal Law to Prevent and Identify Transactions Involving Resources Illegally Obtained requires obliged entities to identify their customers when entering a contractual relationship.

165 The website Mejora tu Escuela, created by El Instituto Mexicano para la Competitividad (IMCO) using open government data, went beyond publishing data about school spending to create an ecosystem and environment in which that data would be scrutinized, allowing for the discovery that 1512 teachers on the payroll all had the same birthdays and all earned higher salaries than the President of Mexico.

166 This 2016 Ipsos survey showed *"corruption"* as the second biggest worry of the average Mexican

167 Maltaverne, Bertrand (July 2017) Blockchain: what are the opportunities for procurement?

Great Financial Crisis. Redesigning the public procurement process to be based on the blockchain infrastructure may allow real-time monitoring of procurements, while preserving more privacy. Yet for all the enthusiasm, we in fact know very little about how blockchain technologies can impact social change through the creation of a trusted identity — what kinds of applications serve what needs, what technological attributes matter the most, what risks are involved and under what conditions DLTs have maximum impact.

1. DEVELOP FINANCIAL INCENTIVES TO ENCOURAGE VOLUNTARY DISCLOSURE OF CORPORATE BENEFICIAL OWNERSHIP AND OTHER LEGAL ENTITY DATA

WHAT NEEDS TO BE DONE AND WHY?

Ultimate beneficial ownership identification effectively links the money that flows in the public procurement sector to the individuals that control that money. This link is crucial to the integrity of the public procurements process, as it allows citizens to see that their contributions are spent wisely, and that the contracts are awarded to the best providers, not to the better bribing providers.

The support of the business sector is crucial to identifying the ultimate beneficial owners to which money is directed. For years, it has been thought that companies could be motivated to reveal the beneficial owners voluntarily through the dual approach of appealing to the societal benefits of eliminating corruption and a practice of sanctions[168] and naming-and-shaming.[169] But plenty of evidence suggests that these self-regulatory approaches have not been effective.[170] Alternatively, more tangible positive incentives are now believed to help engage the business sector in the efforts to counter corruption.[171] Similar incentive structures have been used in the US to motivate health-care providers to offer higher quality care to their patients, and to motivate grant recipients to open their data and results to public scrutiny. Table 6.1 presents different incentive schemes and illustrates the contexts in which they have been applied.

To identify ultimate beneficial owners, **gating** is an effective means to remove companies that do not conform. This option, however, allows only a binary separation of bidding companies and its implementation is difficult. If the bar is set too high, the business community may not support it,[172] and if set too low, the entire exercise is undermined (one current example being the ineffective FATF blacklists).[173]

---

168 In a seminal article, Gneezy and Rustichini showed that fines do not deter misbehaving corporations, who tend to see them as prices to be paid for "deviant behavior."

169 Unger, Brigitte, and Gregory Rawlings. "Competing for criminal money." *Global Business and Economics Review* 10.3 (2008): 331-352.

170 Findley, Michael G., Daniel L. Nielson, and Jason Campbell Sharman. Global shell games: Experiments in transnational relations, crime, and terrorism. Vol. 128. Cambridge University Press, 2014.

171 Unger, Brigitte, and Daan Van der Linde, eds. *Research handbook on money laundering*. Edward Elgar Publishing, 2013.

172 When companies that received public funds were asked to disclose various levels of information, corporate councils started lobbying against the initiative, which ended up being vetoed by Mexico's President.

173 Unger, Brigitte, and Frans Van Waarden. "How to dodge drowning in data? Rule-and risk-based anti money laundering policies compared." *Review of Law & Economics* 5.2 (2009): 953-985.

INCENTIVE SET-UP	REWARD TYPE	OTHER CONTEXTS WHICH UTILIZE THIS INCENTIVE SCHEME
Gating	Only companies that disclose ultimate beneficial ownership can place a bid.	In Mexico, companies that have been sanctioned cannot place a bid. The registry of suppliers and providers that are sanctioned is considered helpful in countering corruption in public procurements.
Compensation	Bidders that disclose ultimate beneficial ownership will receive higher compensation.	In the US, the Centers for Medicare and Medicaid Services started to pay healthcare providers for outcomes instead of just the healthcare transactions. Incentives are paid on top of the standard fee-for-service compensation if the healthcare provider meets or exceeds pre-established expectations.
Outcome	Bidders that disclose ultimate beneficial ownership have a higher chance of winning the tender.	The Directorate of Research and Innovation of the European Union actively supports applications from gender balanced research teams. At the evaluation stage, gender balanced teams will be favored over gender non-balanced teams when the quality of the proposals is similar.

Table 6.1: Incentives to motivate the disclosure of the identity of the ultimate beneficial owner

Alternatively, outcome-based rewards are more flexible. Applied in the context of ultimate beneficial ownership disclosures, they may:

1. Allow for a one-on-one conversation between members of the business community and the public authorities on what it means to disclose the ultimate beneficial owner and how to do so. This can help breach the "no disclosure" status-quo observed by most companies. In 2014 the US National Science Foundation ran a pilot in which it tested a new policy of giving preference in grants to those who disclosed underlying scientific data. For one year the NSF asked members of the science community applying for grants how they would share the results of their research or why they were not able to do this prior to refining and instituting the policy. The pilot results showed that most researchers endorsed open access to research results, even though they were not providing such access.
2. Distinguish among bidding companies and allow for partial disclosures where full disclosures are not possible or desirable for such reasons as economic espionage, state security, etc. Privacy concerns are not always ill-founded, so it is important to restrict disclosure requirements to information that matters for countering corruption. In the Netherlands, for instance, although ultimate beneficial ownership must be publicly disclosed, personal information such as home address and citizen IDs is protected.
3. Tap into the different motivations that companies have for complying. For example, in tight competitions, firms may want to increase their competitive advantage, while in less competitive settings, the motivation may have to come from other sources, such as financial rewards, positive publicity, etc.

How can it be implemented?

CURRENT SYSTEM

CompraNET is an internet-based platform that manages and centralizes every step of the procurement process in Mexico. CompraNET was made the sole mandatory system by the Federal Law of 2009 and is administered by the Secretaría de la Función Pública. CompraNET offers companies that are willing to offer their services a well-organized map of information - i.e., requirements, assessment forms, deadlines, etc. The CompraNET flow therefore follows the succession of the tender from the perspective of a company (see Figure 6.1).

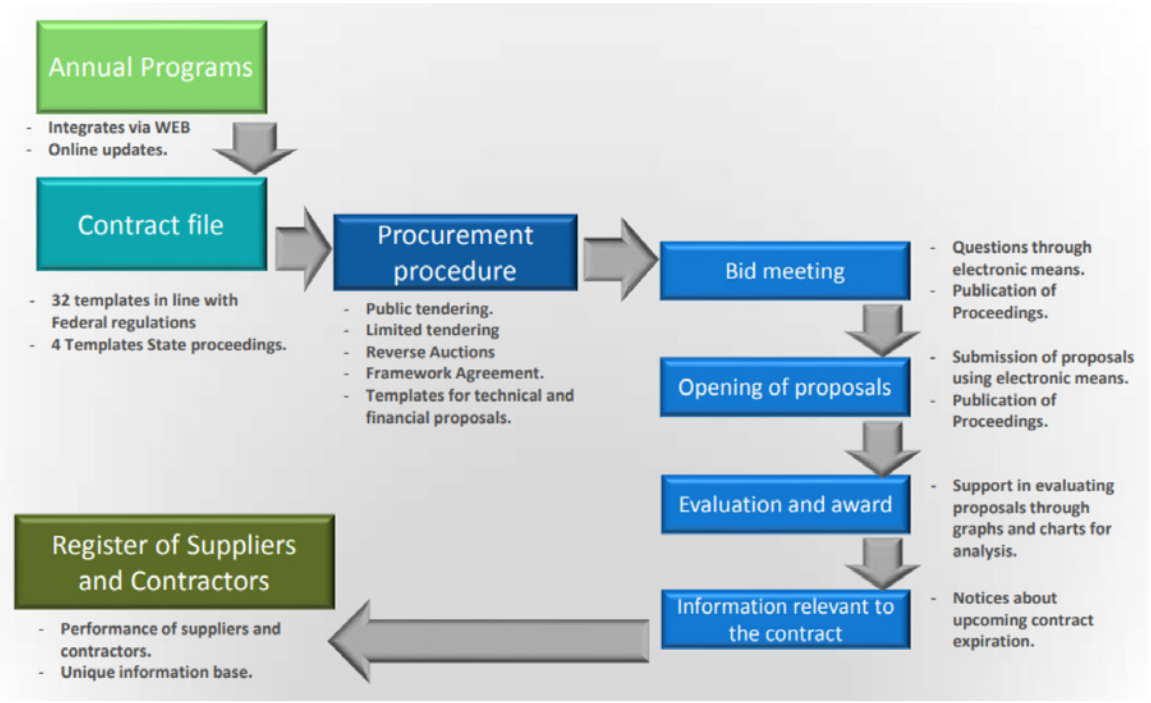


Figure 6.1: CompraNET information. **Source:** CompraNET: e-Procurement as an effective tool to limit opportunities for corruption. Available here

Companies that have contracts with partial or total federal funding **can voluntarily be** registered in a searchable and downloadable Register of Suppliers and Contractors (RUPC). The Secretaría de la Función Pública reveals the incentives for companies to register on the RUPC:

1. **Visibility and publicity.** The CompraNET portal, which is linked to the RUPC portal, receives 35,000 visits per day. Public servants can identify through the RUPC registry companies that have experience providing high quality goods and services.
2. **Lower costs in subsequent applications.** Companies that have good track records on their contracts can submit less evidence of the quality of their work. Also, companies registered in the RUPC can replace certificates documenting the legal existence of their company with the RUPC identity code and a written testimony that all information is complete and up-to-date.

The RUPC contains information on, among others, legal ownership, tax ID, and web address, but **not on ultimate beneficial ownership** (see Table A6.2 in Appendix VI). Data consumers can use the search feature shown in Figure 6.2 to retrieve information on companies registered in RUPC.

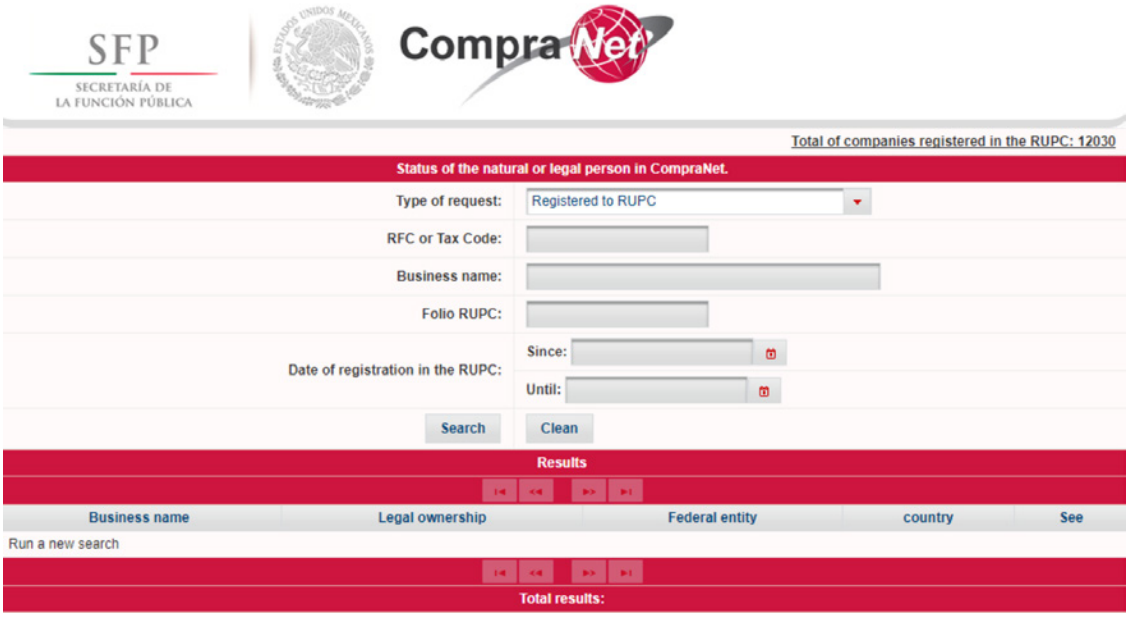


Figure 6.2: Snapshot of the search page in the RUPC. **Source:** *https://cnet.funcionpublica.gob.mx/servicios/consultaRUPC.jsf*

After obtaining information from the RUPC, interested parties can search for more information in Mexico's Public Registry of Commerce. But several limitations on that registry render it ineffective for public inquiries into suspected corruption. Although the law requires Public Notaries and Public Brokers to record the ultimate beneficial owner of companies incorporated in Mexico and disclose it in the Public Registry of Commerce, **this information can be accessed only by law enforcement** (see Table A6.1 in Appendix VI). Moreover, the registry does not check the accuracy of the data and does not impose sanctions for late updates. Finally, the registry cannot be checked online.

ENABLING CONDITIONS

CompraNET (through the RUPC) can be used to gather disclosures on ultimate beneficial ownership that can be readily accessible to the public. But several conditions must be met before enabling incentives for companies that engage in public procurements to disclose ultimate beneficial ownership:

- **Institutional support:** The Secretaría de la Función Pública must design and support a viable reward system for the voluntary disclosure of the ultimate beneficial ownership.
- **Widespread buy-in:** A critical mass of the business community must be willing to accept the incentives that the Secretaría de la Función Pública can propose.
- **Technical support:** The Secretaría de la Función Pública must introduce the necessary changes into CompraNET.

SUGGESTED PLAN

As owner of CompraNET, the Secretaría de la Función Pública should:

CORE ACTION 1

Consult members of the business community to propose disclosure policies that increase the transparency of the ultimate beneficial owner of the money flows, while preserving the safety and the economic interests of the parties involved.

TASK DESCRIPTION

A consultation with the aim of revealing, in an inexpensive and timely manner, the business community's willingness to disclose ultimate beneficial ownership.

HOW

Of the many forms of consultation (e.g., questionnaires, interviews, conferences) we recommend using a questionnaire that measures the disclosure preference of the business community. Interviews with academics can help shape the questionnaires. For example, the US National Science Foundation asked grant applicants to express their standpoints in a two-page document.

The questionnaire should be sent to a randomly selected representative sample of companies participating in public procurements. The questionnaire should aim to understand which of the incentive schemes described in Table 6.1 are preferred by the community - e.g., is a financial reward most motivating or do firms prefer scoring higher by disclosing the identity of their ultimate beneficial owner (or even by being registered on RUPC)? The results should be summarized and published, so that both the international and national academic and business communities can draw important lessons from this exercise.

WHY

A random representative sample of businesses participating in public procurements provide assurance that the consultation correctly reflects the mindset of the business community. Survey results are easy to aggregate and interpret in a timely manner.

COSTS

Constructing and interpreting a robust survey requires knowledge of the ultimate beneficial ownership literature and qualitative data analysis skills. Designing, executing, and interpreting the survey should take 1 person 4-6 weeks, depending on the sample size. High-quality questionnaire platforms (e.g., SurveyMonkey, Google Forms) are available for free. Alternatively, the Secretaría de la Función Pública can hire a commercial polling service, such as Ipsos or Gallup, which may increase the costs. Stakeholder consultations during the Public Access Initiative of the US National Science Foundation was budgeted for $1,250,000. The use of questionnaires, however, can greatly reduce these costs.

CORE ACTION 2

Establish policies for disclosing and rewarding the disclosure of the ultimate beneficial owner both for new companies and for companies already registered on the RUPC.

TASK DESCRIPTION

Utilizing the findings from the consultation exercise of Core Action 1, create a reward roadmap for companies that voluntarily disclose ultimate beneficial ownership.

HOW

Determine first how the data on ultimate beneficial ownership will be stored and disclosed. The two options are *open* access, as in the OpenOwnership model, and *pre-registered* access, such as the Public Registry of Commerce allows for viewing most of its information. We recommend open disclosure[174] as is the case in the UK.

Secondly, determine administration of the rewards.[175] Depending on the incentive system chosen at Core Action 1 (increasing the chances of receiving the contract award or increasing the financial benefits received upon the award), arrange the technical and financial implementation of these rewards in the CompraNET flow (see Figure 6.1).

Finally, release an information memo on the rewards and their application needs through the News section on the CompraNET portal and the media.

WHY

Creating a roadmap helps build trust among the stakeholders and increases the likelihood that the ultimate beneficial ownership will be disclosed.

COSTS

To create the roadmap for its Public Access Initiative, the US National Science Foundation budgeted less than $50,000.

CORE ACTION 3

Execute the necessary changes in the IT architecture of the CompraNET and of the RUPC to allow for: (1) the recording, storing, and disclosure of ultimate beneficial ownership data, and (2) the rewarding of the companies that disclose.

TASK DESCRIPTION

---

174 The UK has published its rules on the disclosure of ultimate beneficial ownership and how this fits the G20 principles.

175 For designing and implementing a reward strategy, a helpful resource can be found here.

Following the roadmap established at Core Action 2, modify the architecture of the CompraNET system to ensure that disclosures of ultimate beneficial ownership can be easily submitted by companies and smoothly retrieved by the public, and that rewards are automatically and justly assigned to the disclosing companies.

HOW

Input: Send every company currently on the RUPC list a request to update their RUPC information, explaining the procedure for updating and their potential benefits. Add an extra "Ultimate Beneficial Ownership" field in the RUPC list of information (see Table A6.2 in Appendix VI) and use this to collect disclosures. **Output:** Change the RUPC visualization page to contain a homologue "Ultimate Beneficial Ownership" field in a way that corresponds to the roadmap developed at Core Action 2. **Link to rewards:** Link the "Ultimate Beneficial Ownership" field to rewards established later in the procurement process as designed in the roadmap.

WHY

Update requests will inform every member of the RUPC list how to disclose and what to expect when disclosing. The "Ultimate Beneficial Ownership" fields will receive the company disclosures and store them within the general repository of CompraNET. Applying the permissions discussed in the roadmap, these disclosures will be exported to the RUPC visualization page under the same "Ultimate Beneficial Ownership" field. Simultaneously, the new inputs will activate a set of decisions about rewards that can be visualized by the company as it progresses through the procurement stages.

COSTS

Developing a functional and integrated IT infrastructure would likely require the involvement of a data architect and a system engineer who could implement the changes to the system. Depending on the complexity of the roadmap and of the reward system, the architectural changes could take between 2 - 6 months, according to our interviews.

What are the risks?

- Consulting parties might agree on a wrong definition of ultimate beneficial ownership.[176] To avoid this, the Secretaría de la Función Pública must make clear in the consultation and in the roadmap that the FATF definition will be applied.

- Voluntary disclosures might be untruthful. To avoid this, the Secretaría de la Función Pública must make clear that the even harsher sanctions apply for perjury in addition to penalizing failure to disclose.

- It is often said that regulations aimed at reducing white-collar crime capture only "small fish." Disclosure of beneficial ownership is one step in a process of building the evidentiary basis for subsequent prosecution. By itself, without other measures, it is not enough,

- The costs of covering the positive incentives may be high. To control costs, the Secretaría de la Función Pública **must conduct and publish a cost-benefit** analysis and set a cost ceiling and prepare a

[176] Not all countries adhere to the FATF definition of ultimate beneficial ownership. Mexico does.

document that justifies these costs. Since we know that 15 - 30% of public expenditure is lost to corruption and collusion, if voluntary disclosures of ultimate beneficial ownership helps deter 1 in 10 corrupt companies, it reduces white-collar crime waste by 1 - 3% of public expenditure. Consequently, positive incentives that amount to 1% of public expenditure will not only increase the integrity of the procurement process but also reduce costs.

3. MEXICO SHOULD MAP THE DEMAND FOR BLOCKCHAIN TECHNOLOGY IN THE PROCUREMENT PROCESS AND PUBLISH A PRIMER ON THE IMPLEMENTATION OF BLOCKCHAIN IN THE PROCUREMENT PROCESS AS A FIRST STEP TOWARD EXPLORING BLOCKCHAIN INTEGRATION

What needs to be done and why?

In public procurements, tracking and analyzing money flows for the purpose of detecting, preventing and countering corruption relies on **accurate records**. Accurate records of ownership (e.g. information, decisions, money) reveal who decided on the basis of which information whether a public good is needed, how a contract should be designed, who should be awarded the money. This information is the bedrock for understanding whether these decisions were done in good faith, whether the information they were based on was correct and thus whether the money flows they generated ultimately landed in the pockets of corrupt companies or not. Consequently, making these **ownership records fraud-proof, easy to register and transparent,** is a huge leap for the integrity and efficiency of the public procurement sector.

The original **solution to avoiding fraud, corruption and collusion in public procurements** was the design and implementation of **checks and balances.**[177] **Their efficiency crucially depended on** (1) the ability of the government to keep **granular records of the procurement process (including of records of money flows and of their transacting parties),** and (2) the ability of **trusted parties to verify and certify** the accuracy and the honesty of these records. This solution has often failed however, as records are notoriously easy to forge and costly to check.[178]

Opening contracting data has effectively improved the integrity of the procurement process because it lowered the costs of checking the accuracy and honesty of public procurement contracts. The problem is, it is still very cheap and easy for data providers to changing their data ex-post to cover-up corruption.[179] Stakeholders and citizens often found themselves asking: *"Who watches the watchers?"* And rightfully so, as there is ample evidence that government statistics are vulnerable to manipulation and that these

---

177 Rose-Ackerman, Susan. (2005) "The challenge of poor governance and corruption." *Especial 1 DIREITO GV Law Review*, pp. 207-266, available here.

178 Kranacher, Mary-Jo, Richard Riley, and Joseph T. Wells. (2010). *Forensic accounting and fraud examination*. John Wiley & Sons, available here.

179 Anyone could use open data to check whether fakes names such as "John Doe" are listed as representatives of the winning companies. After noticing that corrupt deals are revealed through this method, there is little to prevent data providers from changing these entries. While this may show-up as "mistakes" in the data, it is hard to see who made these mistakes and whether there is a pattern within these mistakes that is not random at all.

statistics will be manipulated more, the more pressure is put on gathering them and the more they will be used to measure the performance of governments or to sanction public servants.[180]

Now **distributed ledgers known as blockchains may help increase the costs of forging data and guarantee greater integrity in the contracting process.** But blockchain is not well-understood and Mexico, as Britain's Scientific Advisor did in 2016, should create and release a primer on Blockchain for the Mexican government to educate public servants, explain the value proposition, map the demand for blockchain, and create the momentum for implementation of blockchain in government processes.

Blockchain is a relative new technology based on the age-old idea of distributed trust. In 2008, an unknown source calling itself Satoshi Nakamoto released a paper named "Bitcoin: A Peer-to-Peer Electronic Cash System" which introduced blockchain technology. Blockchain is a novel system that uses a distributed ledger to record transactions and ensure compliance. Blockchain technology relies on an ability to act as a vast, transparent, and secure public database. It has since gained recognition as a tool to transform governance by creating a decentralized system to manage and protect identity, trace and track; and create incentives for smarter social and business contracts. Blockchain is already in use to create greater integrity in the contracting process.[181] For more on What is Blockchain, see Appendix 1 (Selected Readings) and 2 (Basics)

After years of successful use with Bitcoins, **blockchains** have been recognized as **viable platform for financial transactions, as it is very difficult if not impossible to forge financial transactions in the Bitcoin blockchain** (see Figure 6.4) even in the absence of a trusted third party. Bitcoin effectively capitalized on the lack of trust people had in the third parties, after seeing how they had used their position to factually forge records (e.g. LIBORI scandal, Lehman Brothers, AIG, Citigroup fraud, etc.)

The idea of a secure interaction platform that is easy to check and hard to forge has quickly gathered proponents. NYU scholar David Yermack argues that the **cost, speed and data integrity** improvements that blockchain brings relative to the traditional methods of proving ownership can explain the large investments by venture capitalists and by established players in the financial services industry to test their functionality. Consequently, banks are piloting using blockchain to improve **records of interbank payments and regulator supervision.**[182] Australia is piloting **speeding-up its settlement systems** through blockchain.[183] In Argentina, there are plans to use blockchain to instantly and incorruptibly **record and update commercial registries.** In Kenya, pilots are run to track the **identity of refugees and individuals in extreme poverty zones.** The UN Commission on International Trade Law has adopted a Model Law on Electronic Transferable Records, based on blockchain, to **secure the accuracy of shipping documents** that

---

180 The forensic accounting literature reveals numerous examples of countries manipulating official statistics - e.g. to gain better access to foreign flows of capital, join an economic union, avoid international sanctions.

181 The Governance Lab, Selected Readings on Blockchain Technology and Its Potential for Transforming Governance, available at http://thegovlab.org/the-govlab-selected-readings-on-blockchain-technology-and-its-potential-for-transforming-governance/.

182 Australian developed Red Belly Blockchain allows payment ledgers to synchronize securely and instantaneously, and replaces the central bank clearing house system.

183 The Australian Securities Exchange considers replacing its Clearing House Electronic Subregister System with a private blockchain.

accompany the goods imported and exported and the US state of Vermont allows blockchain technology to authenticate legal evidence.

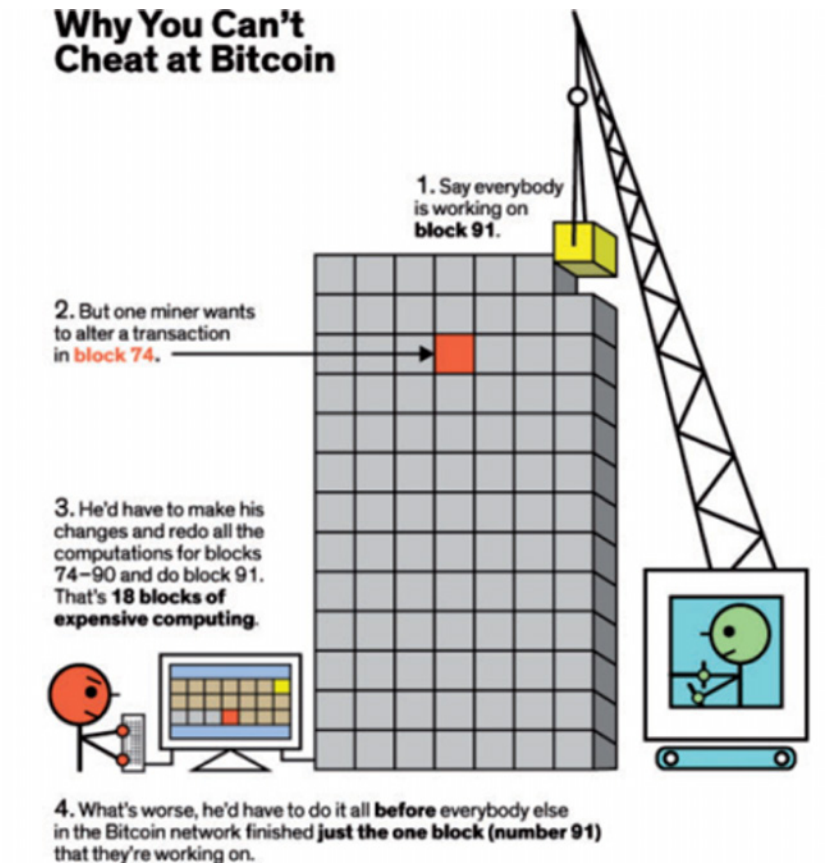


Figure 6.4: The fraud-proof distributed Blockchain Source: Yermack, D. (2017) Corporate Governance and Blockchains, *Review of Finance*, 21(4), available here.

New research shows that blockchain can also increase trust in government administration by:

1. **Improving the integrity of ownership records and of transactions of ownership.**
   The Republic of Georgia has piloted securing land ownership through blockchain, thereby effectively reducing the time and costs of buying or selling land (from a one-day process and $50 to $200 in notary fees, to several minutes and 5 to 10 cents in transaction costs). Japan is planning to unify all property and land registries across urban, farmland and forested areas in a single ledger using blockchain. The UK published a report in 2015 stating that critical civil infrastructure should be protected from malicious cyberattacks using blockchain.

2. **Embedding anti-corruption incentives in the procurement process.**
   Smart contracts are algorithms that read records of ownership in a blockchain infrastructure and **once a set of conditions is met, they transfer or allow transfers (of ownership, money, information, etc.) to happen**. Anti-corruption incentives can be introduced in the form of conditions. Companies such as Provenance use blockchain to **track resources and materials across the supply chain** and to allow customs officials to spot illegal trades by using portable DNA barcode scanners. An example of an anti-

fraud incentive is to not allow the sale of products that do not have all the necessary border-control checks written down in their ledger. Similarly, Codify helps startups **codify Know-Your-Customer protocols** on the blockchain to ensure that business relationships cannot be commenced before the compliance checks are conducted. More recently, the UK Department for Work and Pensions completed a pilot on the use of blockchains to **distribute welfare support** more efficiently. Linking payments to the completion of all the justifying documents is yet another example of embedding incentives in the blockchain. Finally, Japan is piloting processing government tenders through a blockchain-based system.

3. **Increasing contracting transparency while securing higher levels of privacy**.

When companies and governments do not wish to share data for privacy reasons, a transition towards blockchain may allow **the data to be scrutinized without privacy compromises**. For example, through coloured coins information on real world assets - sensitive data - can be encrypted on the blockchain. Its movements, its owner at a certain moment, and its creator would all be visible while other aspects, such as the content of the information and the ability to transfer ownership could be restricted by password. FollowMyVote uses blockchain to **increase election transparency** without compromising voter privacy. Using their private key, a voter can cast and follow their vote into the ballot box. Auditors can see where the votes were cast, which types of ID were given, where they were checked, but they do not have the passwords to view the actual vote.

While there are many pilots being developed across the world with encouraging results, there is still a lot of uncertainty as to how blockchain should be implemented to meaningfully improve the accuracy of ownership in public procurements. Studies in the UK and Australia have already discussed the implementation of blockchain technologies to increase integrity and efficiency of their national administrations. However, they have not focused on procurements. Given its potential to improve the integrity and efficiency of the public procurement sector - by reducing administrative bureaucracy, removing untrustworthy middlemen, and supporting economic development - a primer on how to meaningfully implement blockchain technology to public procurements is needed. Since there is great interest in blockchain throughout the region, there may be appetite for collaboration on such a project and wider distribution. We focus here, however, on developing an understanding of blockchain's use in context of Mexico's procurement process.

How can it be implemented?

Suggested plan. We suggest starting with the core problem - the datasets that are easy to manipulate, difficult to check, and that pose the highest perils for public procurements. Identifying them, the type of information they contain, as well as, their providers and consumers can help inform a decision on whether blockchains can be successfully applied and how.

CORE ACTION 1

IDENTIFY PUBLIC PROCUREMENT DATA THAT SHOULD BE PLACED ON A BLOCKCHAIN.

TASK DESCRIPTION

This task aims to identify which registers of data relevant to the public procurement system are at risk of being manipulated by a single entity with significant negative consequences.

WHO

The organization in charge of the primer should identify data registers that have the highest urgency of being secured through the blockchain infrastructure.

HOW

Of the many ways in which vulnerable datasets can be identified, we suggest taking a funnel approach in order to ensure that the focus is well placed and that a bedrock is set for future pilots.

1. **Identify all datasets** relevant to tracking ownership in the public procurement sector.
2. **Interview the relevant experts** (e.g., cybersecurity experts, risk managers, procurement officers) in order to **rank each of these datasets** with respect to the following features[184]:

   A. The relevance for the healthy functioning of the procurement infrastructure
   Ask the relevant experts to answer questions such as: *"If the dataset would be taken offline or corrupted, what would be the damage to functioning of the public procurement sector?"*

   B. The data protection mechanisms
   Ask the relevant experts to answer questions such as: *"How easy is it for any one party to single-handedly manipulate the data without immediately triggering checks and balances?"*

   C. The speed with which data can be read
   Ask the relevant experts to answer questions such as: *"How fast can data consumers read the data?"*

   D. The fluency with which data consumers can be added
   Ask the relevant experts to answer questions such as: "*How easy can data consumers or other datasets link to the data?"*

   (For all the questions above, employ qualitative data analysis tools such as Likert scale, and Delphi method to analyze and aggregate the answers).

   E. Data Ownership and **Data privacy**
   Ask the relevant experts to answer questions such as: "*Is the data proprietary information?"* "*Does the dataset contain sensitive information?" Allow only Yes/No answers.*

Research conducted by Data 61 suggests that the **most cost effective solution for open non-sensitive datasets is to place them on a public blockchain.** Alternatively, datasets containing sensitive information

[184] Research conducted by Data 61 suggests that these features are relevant to understanding how blockchains can help limit the vulnerabilities and inefficiencies of traditional registers.

or proprietary information can be placed on a different blockchain infrastructure to protect the information from being divulged - e.g., in consortium or private blockchains.

3. Select the databases that have the **optimal mix of relevance and risk.**

WHY

In order to know where to best apply blockchain, it is crucial to know where blockchain can have the largest positive impact. A 2016 report conducted by the UK Chief Scientific Adviser argued that critical civil infrastructure (railways, bridges, energy installations) had a high risk of being erased or maliciously modified in the course of a cyberattack and should be protected using blockchain technology. Successfully placing these on a blockchain would ensure the highest return on investment for this pilot.

COSTS

An overview of the relevant data and a clear understanding of the risks each database is exposed to is an estimated 3-month project.

CORE ACTION 2

EXPLORE PLACING A RISKY RELEVANT DATA ON A BLOCKCHAIN

TASK DESCRIPTION

The goal of this task is to map out how to place a relevant dataset that is at risks of being manipulated on blockchain. For parsimony reasons, we assume that the relevant data is openly available on CompraNET.

WHO

The provider of data must encrypt their data on the blockchain - using metadata.[185]

HOW

Currently, CompraNET operates as a unified internet platform that centralizes data from individual data providers (see Figure 6.5). Different data providers link their metadata to CompraNET and export their data to their assigned fields. The data consumers interact with the CompraNET portal and retrieve the data stored on CompraNET.

Alternatively, the data provider could integrate their metadata **on a public blockchain** and maintain the raw data on their portals. Instead of having the data stored on CompraNET and independent copies possibly stored (for a finite period of time) on local/ federal databases, the data would be permanently stored on a public blockchain, with updated copies being available to consumers directly. The consumers would query the blockchain, discover relevant datasets in the **metadata,** download the data from their providers, and analyze it.[186] Additionally, they could also program data analytic tools to periodically review the data.

---

[185] Through metadata, one can encrypt records of ownership of more than just money on a virtual coin, in a traditional blockchain: it can be ownership of ideas, decisions, property, rights, etc.

[186] Instead of constructing and securing a new blockchain (which can be very expensive), companies such as Factom and Counterparty propose encrypting their entire data on a small message (called metadata) that is then broadcasted publicly by being attached to a Bitcoin transaction. If the source dataset is large, it will not likely be encrypted in the metadata, just identifiers of the data within will be encrypted.

Figure 6.5: Data modules in CompraNET
Source: CompraNET presentation, available here.

WHY

In judging the benefits of the transition towards blockchain, we recommend performing a comparative analysis of the two systems using the features described at Core Action 1.

Table 6.2: Grounds for comparison of the data designs

	DATA INTEGRITY	AVAILABILITY	INTEROPERABILITY	READ LATENCY	ADDING DATA PROVIDERS
Open data portal	▸ No check of the data providers; ▸ No check of the repository	Single point of failure	Uni-directional exports	Fast	Easy
Public blockchain	▸ No check of the data providers;	No single point of failure	Shared data infrastructure	Slower	Medium

Note. Colours are metrics of vulnerabilities: red - high; yellow - moderate; green - low.

DATA INTEGRITY

In the original design, new data entries are validated solely by CompraNET. Using a public blockchain, the creation of new data would require the private key of the data providers and the validation of the majority

of the members of the public blockchain.[187] This implies that: (1) **no alterations can be made to the data through the open portal** as the consumers hold a local copy of the blockchain through which they access the raw data; (2) while consumers directly see who is the data provider, they can not see manipulation taking place before placing the data on the blockchain. Nevertheless, if discrepancies between one or more data providers occur, they can be directly traced back to their owners and they would be directly signaled.

DATA AVAILABILITY

The open platform is a single point of failure for all data consumers and data providers. Blockchain effectively reduces this risk for all stakeholders.

INTEROPERABILITY

Data providers link their data to the open portal through metadata, and then export the raw data in the designated fields. Research suggests that through blockchain *"different registries can more easily interact with each other."*

READ LATENCY

Reading data in the current design is fast. Since the data consumer reads data locally from their own blockchain copy, speed depends on their computing power and can often be done slower. As opposed to financial transactions that need to happen fast, It is not necessary for this data to be read within seconds. Hence the entry is marked yellow (not red) in Table 6.2.

EASE OF ADDING NEW DATA PROVIDERS

The open data portal must register a new data provider by giving them an account. Through blockchain, data providers can independently join by creating a new pair of public and private keys. Authentification of the data providers can be done on the blockchain or off the blockchain, and data providers should themselves run a blockchain node.

COSTS

Research conducted by Data61 suggests that running a registry of approximately 100,000 data entries, including organizations, packages and resources, on an Etherium blockchain amounts to: (1) **Fixed costs** (i.e., the amount for blockchain transaction itself and the cost for allocating an address on the blockchain); (2) **Marginal costs: to deploy the registry** ($.90 - $1.30 depending on the complexity of infrastructure of the registry); **to add new data to the registry** ($.10 - $.50 depending on the complexity of the complexity of infrastructure of the registry). Considering that there are 11,000 businesses registered on CompraNET - **transferring the RUPC data would cost $6,000 - $ 7,000,** assuming that the data infrastructure would be kept complex to allow for the execution of more granular smart contracts. Compared with traditional databases, the costs of adding records is higher, but this data will be globally replicated and exist possibly indefinitely, at no additional cost. Additional consideration to training data suppliers and data providers to be able to use the infrastructure must also be given.

---

187 The identification of additional risks embedded in more complex blockchain IT architecture can be outsourced to specialized companies.

CORE ACTION 3

EXPLORE PLACING THE PUBLIC PROCUREMENT PROCESS ON A BLOCKCHAIN

TASK DESCRIPTION

This task aims to explore how placing the procurement process on a blockchain could improve the public procurement process.

WHO

Data providers that are meaningful to the procurement process and have decided to join the blockchain consortium should track their actions within the public procurement process on the blockchain.

HOW

Data providers should encode their actions and decisions on a consortium blockchain[188] using the following stamps/ identifiers:

1. **Who?** - an identifier of the author of the event recorded
2. **What?** - an identifier of the relevant object
3. **Why?** - business reason for this event
4. **When?** - a timestamp for when the event took place

All blockchain nodes have access to the information disclosed on the blockchain, but aspects of the information shared that are sensitive and meant for a select audience only (e.g., the technical contents of a submitted bid) are further encrypted. Thus, all fields except for *"Who"* and *"What"* could be encrypted with the key shared with the appropriate audience. While this does not allow the broad public to see what the business reason for a decision in a given procurement has been, they see in real time who decided on what part of the procurement process.

WHY

Blockchains can help improve the integrity and the efficiency of the procurement process by allowing unprecedented levels of transparency in the procurement sector. Participants in the blockchain can retrieve relevant information by (1) **crawling the blockchain** and looking into past events to identify competitors, public servants and evaluate the quality of the decisions taken - given the information that was known at the time of the decision; or (2) **analysing the data "on the fly"** with the help of predictive analytics tools to understand where potential risks of corruption or collusion exist.

Secondly, they allow the introduction of anti-corruption incentives in the procurement process. When data is stored using more complex blockchain architectures, it is possible to **include "smart contracts"** - i.e., transactions, payments, or events that self-execute - once a predetermined set of conditions is met.[189] Smart contracts can act as doors that open to allow movements to a secondary step once the conditions are right. They can be programmed, among other things, to:

1. Open doors only when the information they receive is complete (e.g., a bid will not be registered unless all required documents, including UBO disclosure, have been submitted);
2. Keep doors closed when the information is sent at the wrong time, such as late submission of bids;

---

188 Research conducted by Data61 suggests that the most cost effective way to implement such a supply chain solution is to rely on a private (consortium) blockchain.

189 Omohundro, S. (2014) Cryptocurrencies, smart contracts, and artificial intelligence. AI Matters, Vol 1, No 2, pp. 19–21.

3. Keep doors closed when the information is sent by a participant not authorized to send this information.

Finally, evaluating the benefits of placing the public procurement process on a blockchain needs to be done in comparison with the current system (see Table 6.3).

Table 6.3: Grounds for comparison of the supply chain designs

	INTEGRITY	SCALABILITY	INTEROPERABILITY	LATENCY	CONFIDENTIALITY
Open data portal	**No check of the portal**	**Much info is not used**	**Relies on data standards**	**Not real time**	**Little confidential data is shared**
Consortium Blockchain	**Data change is visible**	**Retains only useful info**	**Relies on data standards but wide application of blockchain increases network effects**	**Within minutes**	**Metadata can reveal identities over time**

Note. Colours are metrics of inefficiencies: red - high; yellow - moderate; green - low.

INTEGRITY

CompraNET unifies information that is relevant to the procurement process in Mexico. By forcing data providers to standardize the data they collect and to submit it in a machine readable form to CompraNET, the centralised portal creates a risk as a single point of failure - if it is not working or if it is able to manipulate the data. Such manipulations can have a low chance of detection. Blockchains can remedy this. While large parts of the data may be stored offline, to ensure latency, its hashes immediately signal manipulations.

SCALABILITY

All the information disclosed is placed on the open data portal. Blockchains allow stakeholders to prioritize the information that is published and get notified of changes in the data relevant to them.

INTEROPERABILITY

The public procurement process can be streamlined across different institutions and organizations only when data records are standardized. Tracking and analyzing decisions and money flows can only yield meaningful scenarios, risk profiles and statistics when all data providers follow the same single standards. The more blockchain gathers momentum, the more likely there will be network benefits to be gained from utilizing this standard.

LATENCY

The open data portal is currently updated but not in real time. Blockchain may allow for every decision to be broadcast within minutes. This time frame is sufficiently fast in the low latency field of public procurements.

CONFIDENTIALITY

CompraNET unifies information that is relevant to the procurement process in Mexico. If stakeholders share sensitive information, they must trust the portal and the Secretaría de la Función Pública not to abuse their

trust and share the information in way that would harm them. Alternatively, blockchains could use encryption to encode this sensitive information and delegate some of the checks performed on this information to algorithms. There are scholars who criticize this approach, warning about the dangers of immutable records[190] and encrypting sensitive data with technology that could be insufficiently advanced in the era of quantum computing. Finally, there are companies that specialize in assessing the privacy security risks of a blockchain-based solution, and to offer remedies.

COSTS

While cost measurements are rare, there is a general belief that blockchain will yield significant cost reductions in data and supply chain management. Proponents of blockchain suggest that the significant consequences of corruption (e.g., the high uncertainty, transaction costs, monitoring and enforcement costs) may justify implementation of blockchain regardless of the high initial implementation costs: *"While this would be most beneficial in countries where corruption is high to begin with, it could improve trust everywhere."*[191]

What are the risks?

It is often assumed that if blockchain is beneficial, it will be adopted; yet this is a cumbersome cost/benefit analysis and there is no one clear path to adoption. The primer might therefore draw too optimistic a picture for blockchain-based solutions in order to encourage its implementation. To avoid this, one needs to *pilot and test*, allowing the proposed solution to run in parallel to the existing solutions for a longer period of time. The resilience of the Bitcoin blockchain is known because the solution has been tested for almost 10 years, and has helped remove some myths:

- "Bitcoin is a universal panacea." Because of several limitations,[192] blockchains are amenable to **some categories of assets of value** that require greater security and transparency to shield from misappropriation and to guarantee their worth, such as currency, property rights and votes.[193]
- **Blockchain removes the need for trust**. Nodes trust a (collectively designed) blockchain software, and the data providers when they first link their data to the blockchain.
- **Smart contracts are legal contracts.** There is still great uncertainty with respect to whether smart contracts are really "smart" or "contracts."
- **Blockchain is energy intensive**. While they are more energy intensive than non-synchronizing databases, alternative consensus mechanisms are more energy efficient and becoming more widespread.

---

190 Referring to the decision of the UK Department of Work and Pensions to place welfare payments on blockchain, Jeni Tennison of the Open Data Institute argued "Experimenting with putting highly personal data in immutable data stores is fraught with danger."

191 Hanson RT, Reeson A, Staples M (2017) Distributed Ledgers, Scenarios for the Australian economy over the coming decades, Canberra.

192 Companies that register numerous transactions per minute (e.g. Uber) are not yet attracted to the traditional blockchains, as the validation of a transaction in the blockchain is too slow. Moreover, the size of the blockchain ledger is already large enough to discourage some small nodes from participating, unless they rely on a third party.

193 New America. The Missing Piece of the Internet.

Table A6.1: Public Registry of Commerce data, by type of access

	PUBLIC ACCESS* (AVAILABLE IN PERSON OR VIA PRE-REGISTRATION)	ACCESS ONLY TO LAW ENFORCEMENT
Name of Legal Entity	✓	
Entity Number	✓	
Type of Legal Entity	✓	
Date of Incorporation	✓	
Current Status (active etc)	✓	
Principal Address of Business		✓
Principal Purpose of Business	✓	
Registered Capital	✓	
Registered Agent Information		✓
Officer/ Director Information (including the power of representation)**		✓
Shareholder/ Member Information**		✓
Memorandum		✓
Articles of Incorporation	✓	
Application/ Certificate of Formation	✓	
Annual/ Biennial Reports**		
Shareholder Register	✓	
Register of Charges		✓
Bank Account Information		✓
Payments Records		✓
Historical Records (i.e. past annual filings)		✓

Note: ** registry cannot be consulted on-line; ** information that may allow for the identification of the ultimate beneficial owner(s) of the corporation.* **Source:** *Mexico's Public Registry of Commerce, available here.*

Table A6.2: CompraNET corporate registration fields, with description of the field.

FIELD NAME	DESCRIPTION
RUPC (UNIQUE NUMBER)	The Registration number of the supplier or contractor to the RUPC. Registration is done once and is permanent.
RFC (UNIQUE TAX NUMBER)	For national companies: the "Federal Taxpayer Registration". For foreign companies: the unique tax roll, unique tax register, tax ID number, tax ID card, tax ID, Business license, etc.
COMPANY NAME	Name or business name of the physical or moral person enrolled in the RUPC
COUNTRY	Acronyms that correspond with the country where the company is located.
FEDERAL ENTITY	The States of the Federation according to Article 43 of the Political Constitution of the United Mexican States
STRATIFICATION	Declared by the supplier or contractor to the date of issuance of the report.
TYPE USER	User type declared by the supplier or contractor to the date of issuance of the report: contractor, provider or both.
SECTOR	Main economic activity of the supplier or contractor (e.g. Industry, Commercial, Service, Agriculture, livestock, forestry, Communications and transport, Construction, Energy, Mining, or Manufacturing).
TURN	Brief description of the corporate purpose of the supplier or contractor
CONTRACTS	The agreement of wills to create or transfer rights and obligations, and through which the acquisition or lease of movable property, the provision of services, public works or services related to them are formalized.
DATE INSCRIPTION RUPC	Date on which the unit or entity registers its supplier or contractor and that the Secretaría de la Función Pública assigns the registration number to the RUPC
CONTRACTS EVALUATED SUPPLIER	Agreement of wills to create or transfer rights and obligations, and through which the acquisition or lease of movable property and the provision of services
GRADE COMPLIANCE SUPPLIER	Average rating of supplier contracts entered into with any dependency or entity after registration with the RUPC and under the coverage of the Public Sector Acquisitions, Leases and Services Law.
CONTRACTS EVALUATED CONTRACTOR	Agreement of wills to create or transfer rights and obligations, and through which public works or services related to them are formalized.

FIELD NAME	DESCRIPTION
GRADE COMPLIANCE CONTRACTOR	Average rating of contractor contracts entered into with any dependency or entity after registration with the RUPC and under the coverage of the Law on Public Works and Services Related to the RUPC.
WEBSITE	Vendor or contractor website

Source: *CompraNET - Single Registry of Suppliers and Contractors (RUPC) available here.*

APPENDIX I

OPEN DATA IN MEXICO: INTERNATIONAL RANKINGS, PORTAL, CONTRIBUTING ORGANIZATIONS

Mexico's open data efforts rank high in the international rankings. Figure A1.1 shows that Mexico's government data is more open, useful and reusable than the OECD average.

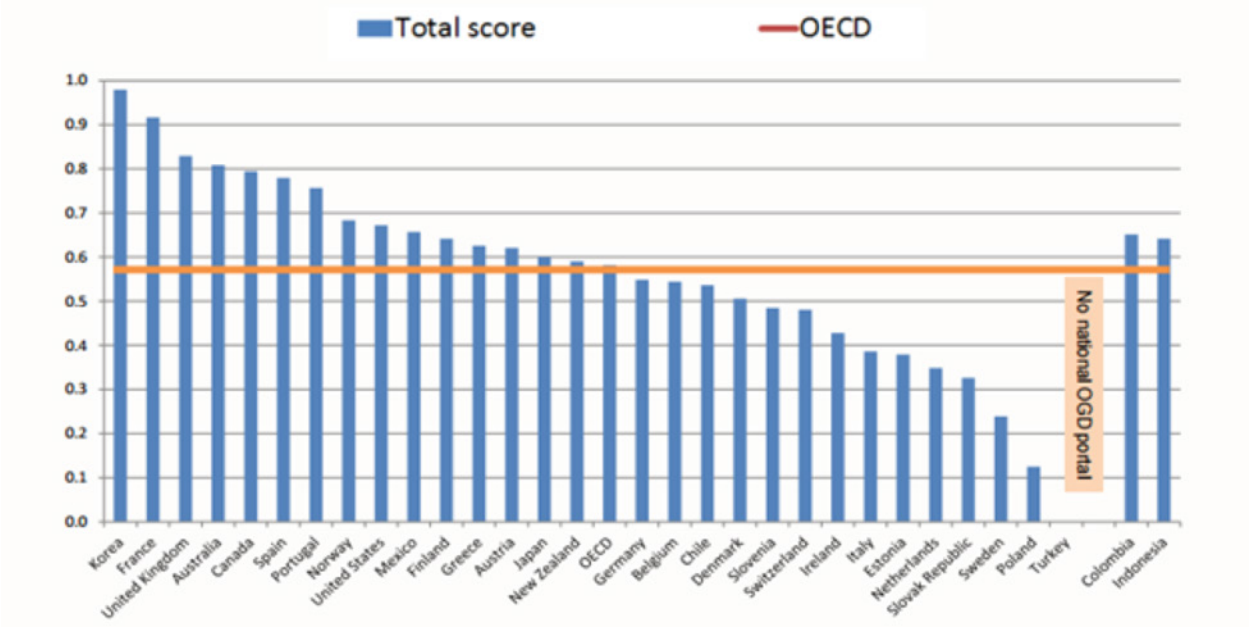


Figure A1.1: OURdata Index – composite index on how open, useful and reusable government data is (from 0 – lowest to 1 – highest). **Source:** OECD (2015). Government at a Glance 2015.

Created by the Mexican Government, datos.gob.mx is the central open data portal which accumulates data from public sector institutions, both at the central and the local level. The portal is backed by a mobile application that aims to support the usage of the portal by citizens. Furthermore, Mexico has put in place an Open Mexico Network that integrates data from several strategic partners – states, municipalities and autonomous bodies – as shown in Table A1.1. The Open Mexico Network portal integrates public, social, and private data in order to share knowledge, increase trust in government, and better evaluate public policies. The portal provides training on how to collect poll data and use open data, and provides tools to add, analyze, and mash existing open data.

Table A1.1: Organizations of the Open Mexico Network, by type and amounts of resources contributed

NAME PARTNER INSTITUTIONS	TYPE	# PUBLISHED RESOURCES
Solidarity, Quintana Roo	Municipality	0
Merida Yucatan	Municipality	0
Minatitlán, Veracruz by Ignacio de la Llave	Municipality	16
Veracruz of Ignacio de la Llave	State	58
Veracruz, Veracruz by Ignacio de la Llave	Municipality	2
Xalapa, Veracruz by Ignacio de la Llave	Municipality	106
Puebla, Puebla	Municipality	82
Mexico	State	140
National Autonomous University of Mexico	Autonomous Body	0
National Institute of Statistic and Geography	Autonomous Body	2133
Morelos	State	27
Jiutepec, Morelos	Municipality	0
Colima	State	15
Jalisco	State	381
Superior Audit of the State of Guanajuato	Autonomous Body	6
Victoria, Tamaulipas	Municipality	5
Reynosa, Tamaulipas	Municipality	16
Torreón, Coahuila de Zaragoza	Municipality	11
Coahuila de Zaragoza	State	14
Ahome, Sinaloa	Municipality	392
Superior Audit of the State of Chihuahua	Autonomous Body	7
San Pedro Garza Garcia, Nuevo León	Municipality	2
Sonora	State	6

NAME PARTNER INSTITUTIONS	TYPE	# PUBLISHED RESOURCES
Baja California	**State**	**17**
Juárez, Chihuahua	**Municipality**	**17**
Mexicali, Baja California	**Municipality**	**0**

Source: *compiled from the MXabierto portal, available here.*

APPENDIX II

The GovLab Selected Readings on Blockchain Technology and Its Potential for Transforming Governance

As part of an ongoing effort to build a knowledge base for the field of opening governance by organizing and disseminating its learnings, the GovLab Selected Readings series provides an annotated and curated collection of recommended works on key opening governance topics. In this edition, we explore the literature on the **Blockchain Technology**.
To suggest additional readings on this or any other topic, please email biblio@thegovlab.org.

The readings below represent selected readings on the applications for governance. They have been categorized by theme – *Governance Applications, Identity Protection and Management*, *Tracing and Tracking*, and *Smart Contracts.*

SELECTED READING LIST

GOVERNANCE APPLICATIONS

- **Atzori, Marcella – The Center for Blockchain Technologies (2015)** Blockchain Technology and Decentralized Governance: Is the State Still Necessary? **–** Aims to investigate the political applications of blockchain, particularly in encouraging government decentralization by considering to what extent blockchain can be viewed as "hyper-political tools." The paper suggests that the domination of private bodies in blockchain systems highlights the continued need for the State to remain as a central point of coordination.
- **Boucher, Philip. – European Parliamentary Research Service (2017)** How blockchain technology could change our lives **–** This report commissioned by the European Parliamentary Research Service provides a deep introduction to blockchain theory and its applications to society and political systems, providing 2 page briefings on currencies, digital content, patents, e-voting, smart contracts, supply chains, and blockchain states.
- **Boucher, Philip. – Euroscientist (2017)** Are Blockchain Applications Guided by Social Values? **–** This report by a policy analyst at the European Parliament's Scientific foresight unit, evaluates the social and moral contours of blockchain technology, arguing that "all technologies have value and politics," and blockchain is no exception. Calls for greater scrutiny on the possibility for blockchain to act as a truly distributed and transparent system without a "middleman."
- **Cheng, Steve; Daub, Matthew; Domeyer, Axel; and Lundqvist, Martin –McKinsey & Company (2017)** Using Blockchain to Improve Data Management in the Public Sector**–**This essay considers the potential uses of blockchain technology for the public sector to improve the security of sensitive information collected by governments and as a way to simplify communication with specialists.
- **De Filippi, Primavera; and Wright, Aaron –Paris University & Cordoza School of Law (2015) Decentralized Blockchain Technology and the Rise of Lex Cryptographia –** Looks at how to regulate blockchain technology, particularly given its implications on governance and society. Argues that a new legal framework needs to emerge to take into account the applications of self-executing blockchain technology.
- **Liebenau, Jonathan and Elaluf-Calderwood, Silvia Monica. – London School of Economics & Florida International University (2016)** Blockchain Innovation Beyond Bitcoin and Banking. **–** A paper that

explores the potential of blockchain technology in financial services and in broader digital applications, considers regulatory possibility and frameworks, and highlights the innovative potential of blockchain.

- **Prpić, John – Lulea University of Technology (2017)** Unpacking Blockchains **–** This short paper provides a brief introduction to the use of Blockchain outside monetary purposes, breaking down its function as a digital ledger and transaction platform.
- **Stark, Josh – Ledger Labs (2016)** Making Sense of Blockchain Governance Applications **–**This CoinDesk article discusses, in simple terms, how blockchain technology can be used to accomplish what is called "the three basic functions of governance."
- **UK Government Chief Scientific Adviser (2016)** Distributed Ledger Technology: Beyond Blockchain **–** A report from the UK Government that investigates the use of blockchain's "distributed leger" as a database for governments and other institutions to adopt.

IDENTITY PROTECTION AND MANAGEMENT

- **Baars, D.S. – University of Twente (2016**) *Towards Self-Sovereign Identity Using Blockchain Technology.* **–** A study exploring self-sovereign identity – i.e. the ability of users to control their own digital identity – that led to the creation of a new architecture designed for users to manage their digital ID. Called the Decentralized Identity Management System, it is built on blockchain technology and is based on the concept of claim-based identity.
- **Burger, Eric and Sullivan, Clare Linda. – Georgetown University (2016) E-Residency and Blockchain. –** A case study focused on an Estonian commercial initiative that allows for citizens of any nation to become an "Estonian E-Resident." This paper explores the legal, policy, and technical implications of the program and considers its impact on the way identity information is controlled and authenticated.
- **Nathan, Oz; Pentland, Alex 'Sandy'; and Zyskind, Guy – Security and Privacy Workshops (2015) Decentralizing Privacy: Using Blockchain to Protect Personal Data –** Describes the potential of blockchain technology to create a decentralized personal data management system, making third-party personal data collection redundant.
- **De Filippi, Primavera – Paris University (2016)** The Interplay Between Decentralization and Privacy: The Case of Blockchain Technologies **–** A journal entry that weighs the radical transparency of blockchain technology against privacy concerns for its users, finding that the apparent dichotomy is not as at conflict with itself as it may first appear.

TRACING AND TRACKING

- **Barnes, Andrew; Brake, Christopher; and Perry, Thomas – Plymouth University (2016) Digital Voting with the use of Blockchain Technology –** A report investigating the potential of blockchain technology to overcome issues surrounding digital voting, from voter fraud, data security and defense against cyber attacks. Proposes a blockchain voting system that can safely and robustly manage these challenges for digital voting.
- **The Economist (2015), "Blockchains The Great Chain of Being Sure About Things." –** An exploratory article that explores the potential usefulness of a blockchain-based land registry in places like Honduras and Greece, transaction registries for trading stock, and the creation of smart contracts.
- **Lin, Wendy; McDonnell, Colin; and Yuan, Ben – Massachusetts Institute of Technology (2015) Blockchains and electronic health records. –** Suggests the "durable transaction ledger" fundamental to blockchain has wide applicability in electronic medical record management. Also,

evaluates some of the practical shortcomings in implementing the system across the US health industry.

SMART CONTRACTS

- **Iansiti, Marco; and Lakhani, Karim R. – Harvard Business Review (2017)** The Truth about Blockchain **–** A Harvard Business Review article exploring how blockchain technology can create secure and transparent digital contracts, and what effect this may have on the economy and businesses.
- **Levy, Karen E.C. – Engaging Science, Technology, and Society (2017)** Book-Smart, Not Street-Smart: Blockchain-Based Smart Contracts and The Social Workings of Law. **–** Article exploring the concept of blockchain-based "smart contracts" – contracts that securely automate and execute obligations without a centralized authority – and discusses the tension between law, social norms, and contracts with an eye toward social equality and fairness.

ANNOTATED SELECTED READING LIST

Cheng, Steve, Matthias Daub, Axel Domeyer, and Martin Lundqvist. "Using blockchain to improve data management in the public sector." McKinsey & Company. Web. 03 Apr. 2017. http://bit.ly/2nWgomw

- An essay arguing that blockchain is useful outside of financial institutions for government agencies, particularly those that store sensitive information such as birth and death dates or information about marital status, business licensing, property transfers, and criminal activity.
- Blockchain technology would maintain the security of such sensitive information while also making it easier for agencies to use and access critical public-sector information.
- Despite its potential, a significant drawback for use by government agencies is the speed with which blockchain has developed – there are no accepted standards for blockchain technologies or the networks that operate them; and because many providers are start-ups, agencies might struggle to find partners that will have lasting power. Additionally, government agencies will have to remain vigilant to ensure the security of data.
- Although best practices will take some time to develop, this piece argues that the time is now for experimentation – and that governments would be wise to include blockchain in their strategies to learn what methods work best and uncover how to best unlock the potential of blockchain.

"The Great Chain of Being Sure About Things." The Economist. The Economist Newspaper, 31 Oct. 2015. Web. 03 Apr. 2017. http://econ.st/1M3kLnr

- This is an exploratory article written in *The Economist* that examines the various potential uses of blockchain technology beyond its initial focus on bitcoin:
  - It highlights the potential of blockchain-based land registries as a way to curb human rights abuses and insecurity in much of the world (it specifically cites examples in Greece and Honduras);
  - It also highlights the relative security of blockchain while noting its openness;
  - It is useful as a primer for how blockchain functions as tool for a non-specialist;
  - Discusses "smart contracts" (about which we have linked more research above);
  - Analyzes potential risks;
  - And considers the potential future unlocked by blockchain

- This article is particularly useful as a primer into the various capabilities and potential of blockchain for interested researchers who may not have a detailed knowledge of the technology or for those seeking for an introduction.

Iansiti, Marco and Lakhani, Karim R. "The Truth About Blockchain." Harvard Business Review. N.p., 17 Feb. 2017. Web. 06 Apr. 2017. http://bit.ly/2hqo3FU

- This entry into the *Harvard Business Review* discusses blockchain's ability to solve the gap between emerging technological progress and the outdated ways in which bureaucracies handle and record contracts and transactions.
- Blockchain, the authors argue, allows us to imagine a world in which "contracts are embedded in digital code and stored in transparent, shared databases, where they are protected from deletion, tampering, and revision", allowing for the removal of intermediaries and facilitating direct interactions between individuals and institutions.
- The authors compare the emergence of blockchain to other technologies that have had transformative power, such as TCP/IP, and consider the speed with which they have proliferated and become mainstream.
  - They argue that like TCP/IP, blockchain is likely decades away from maximizing its potential and offer frameworks for the adoption of the technology involving both single-use, localization, substitution, and transformation.
  - Using these frameworks and comparisons, the authors present an investment strategy for those interested in blockchain.

IBM Global Business Services Public Sector Team. "Blockchain: The Chain of Trust and its Potential to Transform Healthcare – Our Point of View." IBM. 2016. http://bit.ly/2oBJDLw

- This enthusiastic business report from IBM suggests that blockchain technology can be adopted by the healthcare industry to "solve" challenges healthcare professionals face. This is primarily achieved by blockchain's ability to streamline transactions by establishing trust, accountability, and transparency.
- Structured around so-called "pain-points" in the healthcare industry, and how blockchain can confront them, the paper looks at 3 concepts and their application in the healthcare industry:
  - **Bit-string cryptography:** Improves privacy and security concerns in healthcare, by supporting data encryption and enforces complex data permission systems. This allows healthcare professionals to share data without risking the privacy of patients. It also streamlines data management systems, saving money and improving efficiency.
  - **Transaction Validity:** This feature promotes the use of electronic prescriptions by allowing transactional trust and authenticated data exchange. Abuse is reduced, and abusers are more easily identified.
  - **Smart contracts:** This streamlines the procurement and contracting qualms in healthcare by reducing intermediaries. Creates a more efficient and transparent healthcare system.
- The paper goes on to signal the limitations of blockchain in certain use cases (particularly in low-value, high-volume transactions) but highlights 3 use cases where blockchain can help address a business problem in the healthcare industry.
- Important to keep in mind that, since this paper is geared toward business applications of blockchain through the lens of IBM's investments, the problems are drafted as business/transactional problems, where blockchain primarily improves efficiency than supporting patient outcomes.

Nathan, Oz; Pentland, Alex 'Sandy'; and Zyskind, Guy "Decentralizing Privacy: Using Blockchain to Protect Personal Data" Security and Privacy Workshops (SPW). 2015. http://bit.ly/2nPo4r6

- This technical paper suggests that anonymization and centralized systems can never provide complete security for personal data, and only blockchain technology, by creating a decentralized data management system, can overcome these privacy issues.
- The authors identify 3 common privacy concerns that blockchain technology can address:
  - **Data ownership:** users want to own and control their personal data, and data management systems must acknowledge this.
  - **Data transparency and auditability:** users want to know what data is been collected and for what purpose.
  - **Fine-grained access control:** users want to be able to easily update and adapt their permission settings to control how and when third-party organizations access their data.
- The authors propose their own system designed for mobile phones which integrates blockchain technology to store data in a reliable way. The entire system uses blockchain to store data, verify users through a digital signature when they want to access data, and creates a user interface that individuals can access to view their personal data.
- Though much of the body of this paper includes technical details on the setup of this blockchain data management system, it provides a strong case for how blockchain technology can be practically implemented to assuage privacy concerns among the public. The authors highlight that by using blockchain "laws and regulations could be programmed into the blockchain itself, so that they are enforced automatically." They ultimately conclude that using blockchain in such a data protection system such as the one they propose is easier, safer, and more accountable.

Wright, Aaron, and Primavera De Filippi. "Decentralized blockchain technology and the rise of lex cryptographia." 2015. Available at SSRN . http://bit.ly/2oujvoG

- This paper proposes that the emergence of blockchain technology, and its various applications (decentralized currencies, self-executing contracts, smart property etc.), will necessitate the creation of a new subset of laws, termed by the authors as "Lex Cryptographia."
- Considering the ability for blockchain to "cut out the middleman" there exist concrete challenges to law enforcement faced by the coming digital revolution brought by the technology. These encompass the very benefits of blockchain; for instance, the authors posit that the decentralized, autonomous nature of blockchain systems can act much like "a biological virus or an uncontrollable force of nature" if the system was ill-intentioned. Though this same system can regulate the problems of corruption and hierarchy associated with traditional, centralized systems, their autonomy poses an obvious obstacle for law-enforcement.
- The paper goes on to details all the possible benefits and societal impacts of various applications of blockchain, finally suggesting there exists a need to "rethink" traditional models of regulating society and individuals. They predict a rise in *Lex Cryptographia* "characterized by a set of rules administered through self-executing smart contracts and decentralized (and potentially autonomous) organizations." Much of these regulations depend upon the need to supervise restrictions placed upon blockchain technology that may chill its application, for instance corporations who may choose to purposefully avoid including any blockchain-based applications in their search engines so as to stymie the adoption of this technology.

APPENDIX III

BLOCKCHAIN BASICS

What is Blockchain and how does it work?

Blockchain is an IT protocol that replaces double-entry bookkeeping with n-entry bookkeeping by ensuring a transparent, secure and democratic synchronization of the *n* copies, on a regular basis. Every participant in the blockchain network (see the colored nodes in Figure A6.1 below) gets a complete copy of the books as soon as it joins the network. The copy – also called the blockchain - is a sequential chain of blocks, where every block contains a set of transactions that are cryptographically validated and time stamped.[194] **The blockchain thus contains the complete details of the nodes, their balances and transactions, from the very first block to the latest,[195] and the Blockchain is shared among all the participants to the Blockchain network.**

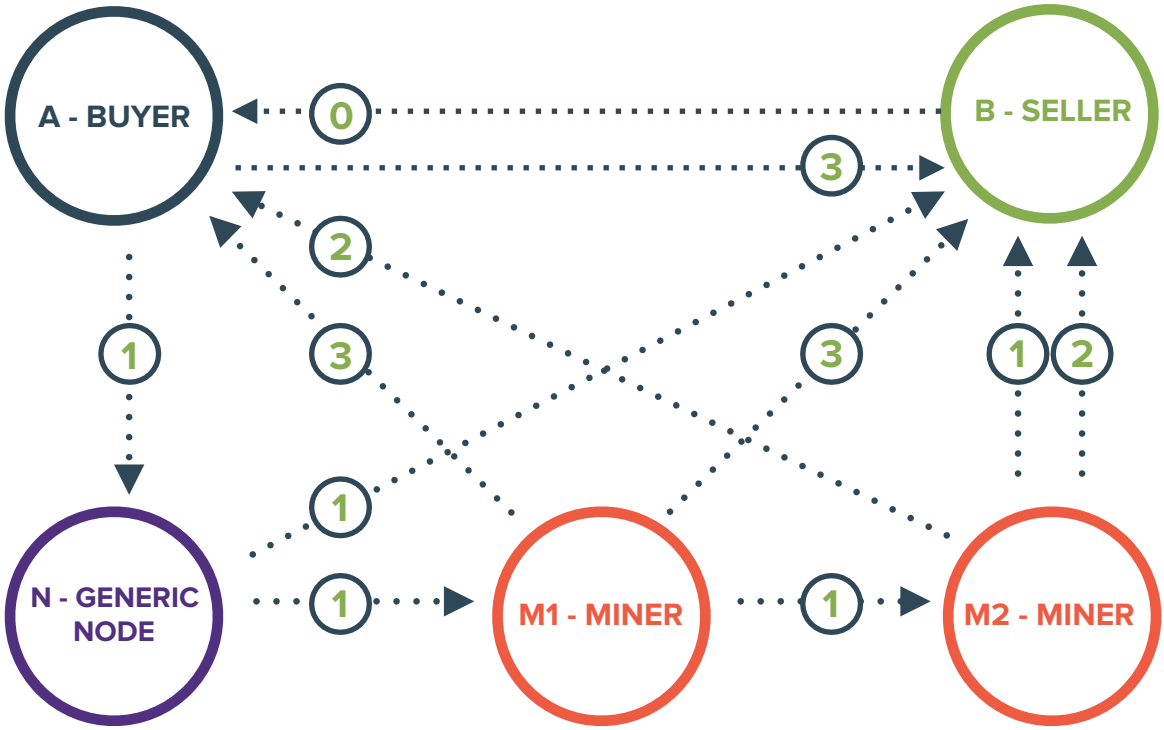


Figure A6.1: BITCOIN transactions from initiation to its permanent recording

Figure A6.1 shows how a new BITCOIN transaction occurs and is recorded in the Blockchain. Several types of participants are illustrated – the two transacting parties (the buyer and seller), two competing miners,[196]

---

[194] For an intuitive introduction to BITCOIN, visit Khan Academy's video series.

[195] Scott, Tony. Blockchain: Blueprint to Dissecting The Hidden Economy!- Smart Contracts, Bitcoin and Financial Technology (Kindle Locations 105-107), Kindle Edition.

[196] Miners are computers that update the general ledger by encrypting the new transactions that everyone has heard of, and by share an updated ledger with all other users in the networks. Miners are compensated with new Bitcoins for their efforts to validate and encrypt the new ledger. (Nakamoto, 2008)

and a generic user node. The communication between nodes is illustrated through black arrows, while the red arrow illustrates the transfer of the P BITCOINs.

The sequence of decisions and communication is the following:

1. *At time T0*: Node B broadcasts its wallet address and its products price P BTC. This is often done "in the real world" by listing the price of a good or service in BTC.
2. *At time T1:* Node A broadcast to her active blockchain connections "Transaction xxx01". A factually declares to her connections "I am paying to B amount P of BTC". Her friends (denoted by the Generic Node for parsimony reasons) begin to gossip about this new transaction. The gossip propagates until everyone knows of it;
3. *At time T2 (approx. 10min later):* Transaction xxx01 is confirmed by mining node M2 and the Block (N+1) with transaction xxx01 added to blockchain is created;
4. *At time T3 (e.g. n x 10min later):* Transaction xxx01 confirmed by another mining node M1 and the block (N+1) as generated by the fastest miner (Mining node M2) is now part of the blockchain forever and M2 is to be remunerated for its proof of work. At this point, the wallet of A has P BTC less than at T0 and the wallet of B has P BTC more.

What can it solve?

The Blockchain protocol is based on computers gossiping about eventual new changes that are new vis-à-vis the ledger that they hold. Consequently, there is **no need for a trusted third-party.** Blockchain can keep accurate and timely records of any form of information that are **safe from being forged**. Since a distributed Blockchain relies on computers algorithmically updating their ledgers by gossiping with each-other and choosing to record only majority-accepted historical updates, it is only a majority of computers (or one computer with more than 50% of the total computing power) that can "fraudulently" change history. Moreover, the change in history (called "fork") is visible and can be traced to the computers that validated it, and it is very costly to implement.

APPENDIX IV

Comprehensive List of Supporters and Participants

EXPERT PARTICIPANTS AND SUPPORTERS					
#	NAME	ORGANIZATION	ROLE	COUNTRY	CONFERENCES ATTENDED
1	Adam Foldes	Transparency International	Legal Advisor	Germany	Prosecuting Corruption
2	Alejandro Ponce	World Justice Project	Chief Research Officer	United States	Measuring Corruption
3	Alexandra Rusu	Code4Romania	Founder	Romania	Money Flows
4	Alexis Bushnell	Queen's Management School	Research Fellow	UK	Whistleblowing
5	Alfonso Hernandez	National Anti-Corruption System's Citizen Participation Committee	Member	Mexico	Measuring Corruption
6	Ana Lorena de la O	Yale University	Scholar	United States	Measuring Corruption
7	Andres Knobel	Tax Justice Network	Consultant	United Kingdom	Money Flows
8	Andrew Hoppin	Global Integrity	Chair	United States	Money Flows
9	Andrew Young	GovLab	Associate Director of Research	USA	Measuring Corruption
10	Ania Calderon	International Open Data Charter	Executive Director	Mexico	Money Flows

EXPERT PARTICIPANTS AND SUPPORTERS

#	NAME	ORGANIZATION	ROLE	COUNTRY	CONFERENCES ATTENDED
11	Anjali Nayar	Timby	Founder	Africa	Measuring Corruption
12	Carlos E. Jimenez-Gomez	IEEE e-Government Initiative	International Consultant	Spain	Judicial Corruption
13	Carolina Pozo	Wonder Lab	Co-Founder	Canada	Measuring Corruption
14	Cristiano Ferri Faria	Brazilian House of Representatives	Project Manager	Brazil	Citizen Participation
15	Dan Wasser	Sentinel Visualizer	Director of Business Development	USA	Measuring Corruption
16	Dana Gold	Government Accountability Project	Director of Education & Strategic Partnerships	USA	Whistleblowing
17	Daniel Tanis	University of Cambridge	Research Assistant	United Kingdom	Measuring Corruption
18	Danielle Denny	Yale University	Visiting Researcher	Brazil	Measuring Corruption
19	Danny Sierra	Universidad Nacional de Colombia	Professor	Colombia	Citizen Participation
20	David Bringle	Whistle Lake Consulting	CEO	USA	Measuring Corruption
21	David Jancsics	Rutgers University	Post-Doctoral Associate	United States	Measuring Corruption
22	David Yoon	Timby	Technical Advisor	USA	Citizen Participation

EXPERT PARTICIPANTS AND SUPPORTERS

#	NAME	ORGANIZATION	ROLE	COUNTRY	CONFERENCES ATTENDED
23	Davide Del Monte	TI Italy	Executive Director	Italy	Whistleblowing
24	Eduard Martin-Borregon	PODER	Coordinator for Transparency Technology	Mexico	Whistleblowing
25	Eduardo Bohorquez	Transparencia Mexicana	Director	Mexico	Measuring Corruption
26	Eliza Keller	Abdul Latif Jameel Poverty Action Lab	Senior Policy Associate	United States	Citizen Participation
27	Elsa Peraldi	Global Integrity	Manager, Africa Integrity Indicators	United States	Judicial Corruption
28	Fabiano Angelico	Transparency International Brazil	Program Consultant	Brazil	Citizen Participation
29	Fabio Pietrosanti	Hermes Center	President and Co-Founder	Italy	Whistleblowing
30	Fabro Steibel	Institute of Technology and Society	Executive Director	Brazil	Citizen Participation
31	Fernando Bracaccini	Asociacion Civil por la Igualdad y la Justicia (ACIJ)	Coordinator, Strengthening Democrcy and Anti-Corruption	Argentina	Judicial Corruption
32	Florencia Coelho	La Nacion	New Media Research and Training Manager	Argentina	Citizen Participation
33	Francesca Recanatini	World Bank	Senior Public Sector Specialist, Governance	United States	Measuring Corruption
34	Francisco Sanchez Lay	General Secretariat of the Presidency	Advisor for the Commission for Probity and Transparency	Chile	Whistleblowing

EXPERT PARTICIPANTS AND SUPPORTERS

#	NAME	ORGANIZATION	ROLE	COUNTRY	CONFERENCES ATTENDED
35	Fredric I. Lederer	Center for Legal and Court Technology	Director	United States	Judicial Corruption
36	Gaurav Dwivedi	MyGov, Government of India	CEO	India	Citizen Participation
37	Georg Neumann	Open Contracting Partnership	Senior Communications Manager	USA	Money Flows
38	Gretta Fener	Basel Institute on Governance	Managing Director	Switzerland	Prosecuting Corruption
39	Grzegorz Makowski	Batory Foundation	Program Director, Public Integrity	Poland	Whistleblowing
40	Guojon Idir	Citizens Foundation	Director of Operations at the Citizens Foundation	Iceland	Citizen Participation, Whistleblowing
41	Hera Hussain	Open Corporates	Community and Advocacy Manager	United Kingdom	Money Flows
42	Ines Arroyo Quiroz	Centro Regional de Investigaciones Multidisciplinarias (CRIM), UNAM	Postdoctoral researcher	Mexico	Measuring Corruption
43	Inigo Fernandez	Ministry of Public Administration	Head of the liaison unit with the National Anti-Corruption System	Mexico	Measuring Corruption
44	Jan van Zyl Smit	Bingham Center for the Rule of Law	Research Fellow	England	Judicial Corruption
45	Jane Ellis	International Bar Association	Director, Legal Policy & Research Unit	England	Judicial Corruption
46	Johannes Tonn	Global Integrity	Director, Programs & Partnerships,	USA	Measuring Corruption

EXPERT PARTICIPANTS AND SUPPORTERS

#	NAME	ORGANIZATION	ROLE	COUNTRY	CONFERENCES ATTENDED
47	John Pettus	Fiskkit	CEO	USA	Citizen Participation
48	Joras Ferwerda	University of Utrecht	Assistant Professor	Netherlands	Measuring Corruption
49	Jose Andres Chavez Legarreta	Bayonet.IO	Co-Founder	Mexico	Measuring Corruption
50	Jose Marin	Transparency International	Public Sector Integrity Programme Coordinator	Germany	Citizen Participation
51	Juan David Martin	Veeduria Distrital de Bogota	Veedor Delegado	Colombia	Whistleblowing
52	Juanita Riano	Inter-American Development Bank	Senior Integrity Officer in the Office of Institutional Integrity	USA	Measuring Corruption
53	Kartik Kumar	Satsearch.co	Co-Founder	Netherlands	Measuring Corruption
54	Leah Ambler	OECD	Legal Analyst, Anti-Corruption Division	France	Prosecuting Corruption
55	Leonardo Oliveira	Federal Justice of Brazil (Rede InovaGov)	Technical Coordinator	Brazil	Judicial Corruption
56	Lorraine Martin	Open Democracy Advice Centre	Manager, Whistleblowing Programme	South Africa	Whistleblowing
57	Madhumita Bhattacharyya	IBM	Partner at the Cognitive & Analytics Deparment	United States	Money Flows
58	Manuel Garrido	Innocence Project	Chair	Argentina	Judicial Corruption

EXPERT PARTICIPANTS AND SUPPORTERS

#	NAME	ORGANIZATION	ROLE	COUNTRY	CONFERENCES ATTENDED
59	Manuel Garrido	Innocence Project	Chair	Argentina	Prosecuting Corruption
60	Marco Konopacki	Institute of Technology and Society	Project Coordinator	Brazil	Citizen Participation
61	Mariana Niembra	Incorruptible	CEO	Mexico	Whistleblowing
62	Mat Tromme	Bingham Centre for the Rule of Law	Senior Research Fellow	United Kingdom	Measuring Corruption
63	Matej Simalcik	Transparency International Slovakia	Judiciary Program Coordinator	Slovakia	Judicial Corruption, Prosecuting Corruption
64	Michael Sachs	New York County District Attorney's Office	Chief of the Investigation Division	United States	Prosecuting Corruption
65	Mirte Postrama	Stanford Law School	Fellow for Human Rights, Criminal Justice and Prison Reform in the Americas	United States	Judicial Corruption
66	Monica Wills Silva	Behavioral Insights Team	Home Affairs and International Programmes	United Kingdom	Measuring Corruption
67	Mortaza S. Bargh	Ministry of Security and Justice	Research and Documentation Centre, Statistical Data and Policy Analysis Division	Netherlands	Judicial Corruption
68	Mukelani Dimba	Open Government Partnership	Co-Chair	South Africa	Whistleblowing
69	Nerisa Dozo	Griffith University	Survey and Business Manager, Whistling While They Work 2	Australia	Whistleblowing

EXPERT PARTICIPANTS AND SUPPORTERS

#	NAME	ORGANIZATION	ROLE	COUNTRY	CONFERENCES ATTENDED
70	Olga Savran	OECD	Anti-Corruption Network Manager	France	Prosecuting Corruption
71	Orlando A. Rojas	Observatorio del Gasto Fiscal de Chile	Project Manager	Chile	Money Flows
72	Oscar Jaimes Bello	INEGI	Director for Government Information Development	Mexico	Measuring Corruption
73	Oscar Mendez	The Economic Institutions, Behavior, and Performance program at the Alfred P. Sloan Foundation	Program Associate	USA	Money Flows
74	Pablo Collada	Ciudadano Inteligente	Executive Director	Chile	Citizen Participation
75	Paula Perez	Asociacion Civil por la Igualdad y la Justicia (ACIJ)	Responsable de Desarrollo de Programas	Argentina	Judicial Corruption
76	Paulius Murauskas	Transparency International Lithuania	Open Courts Program Coordinator	Lithuania	Judicial Corruption
77	Paulo Pandolfi	Colab.re	Founder	Brazil	Citizen Participation
78	Philip Langebroek	Utrecht University	Professor	Netherlands	Judicial Corruption
79	Rafael Velasco	Fundacion Getulio Vargas	Researcher, Public Transparency Program	Brazil	Measuring Corruption
80	Renata Baptista	The Brazilian Federal Prosecution Office	Prosecutor	Brasil	Citizen Participation

EXPERT PARTICIPANTS AND SUPPORTERS

#	NAME	ORGANIZATION	ROLE	COUNTRY	CONFERENCES ATTENDED
81	Renzo Lavin	Asociacion Civil por la Igualdad y la Justicia	Director	Argentina	Prosecuting Corruption
82	Robert Palmer	International Open Data Charter	Partnerships and Communication Director	United Kingdom	Citizen Participation
83	Rocio Paniagua	International Bar Association	Senior Legal Advisor, Legal Policy & Research Unit	England	Judicial Corruption, Prosecuting Corruption
84	Rodrigo Mora Ortega	General Secretariat of the Presidency	President of the Citizen Defense and Transparency Commission	Chile	Whistleblowing
85	Ronald Meijer	Ministry of Security and Justice	Research and Documentation Centre, Statistical Data and Policy Analysis Division	Netherlands	Judicial Corruption
86	Sandor Lederer	K-Monitor	Director	Hungary	Measuring Corruption
87	Sandra Elena	Ministry of Justice	Coordinator of Open Justice Program	Argentina	Judicial Corruption
88	Sean McKessy	Phillips and Young	Attorney	United States	Whistleblowing
89	Sebastian Acevedo	Inter-American Development Bank	Consultant	USA	Measuring Corruption
90	Sergejus Muravjovas	Transparency International Lithuania	Director	Lithuania	Judicial Corruption
91	Sophia Lin	ICAR	Legal & Policy Associate	United States	Prosecuting Corruption

EXPERT PARTICIPANTS AND SUPPORTERS

#	NAME	ORGANIZATION	ROLE	COUNTRY	CONFERENCES ATTENDED
92	Stephan Wolf	Global Legal Entity Identifier Foundation (GLEIF)	CEO	Germany	Money Flows
93	Talia Hagerty	Institute for Economics and Peace	Research Fellow	Australia	Measuring Corruption
94	Ursula Indacochea	Due Process Law Foundation	Senior Program Officer	Washington D.C.	Judicial Corruption
95	Victor Alistar	Transparency International Romania	Director	Romania	Judicial Corruption
96	William (Bill) Nichols	Office of Financial Research	Associate Director of Strategic Data Support	USA	Money Flows
97	Yago Bermejo Abati	Laboratorio de Inteligencia Colectiva para la Participacion Democratica	Coordinador de Proyecto MediaLabICs	Spain	Citizen Participation
98	Zosia Sztykowski	OpenOwnership	Project Coordinator	United Kindgom	Money Flows
99	Wade Shen	DARPA	Program Manager	USA	Measuring Corruption
100	Nicolás Dassen	Inter-American Development Bank	Senior Modernization of the State Specialist, Innovations for Citizen Service Division	USA	Measuring Corruption, Money Flows, Citizen Participation, Judicial Corruption, Whistleblowing, Prosecuting Corruption
101	Michelle Marshall	Inter-American Development Bank	Open Innovation for Development, Knowledge and Learning Sector	USA	
102	Arturo Muente Kunigami	Inter-American Development Bank	Senior Modernization of the State Specialist, Innovations for Citizen Service Division	USA	

EXPERT PARTICIPANTS AND SUPPORTERS

#	NAME	ORGANIZATION	ROLE	COUNTRY	CONFERENCES ATTENDED
103	Mario Sangines	Inter-American Development Bank	Modernization of the State Principal Specialist, Innovations for Citizen Service Division	Mexico	Measuring Corruption, Money Flows, Citizen Participation, Judicial Corruption, Whistleblowing, Prosecuting Corruption
104	Dante Preisser Rentería	Secretaría de la Función Pública	Head of the Contact Unit for the Anti-Corruption System Secretaría de la Función Pública, México	Mexico	
105	Javier Berain Garza	Secretaría de la Función Pública	Deputy Director-General for Studies and Policies on Transparency and Accountability	Mexico	
106	Aline Guardado Estrella	Secretaría de la Función Pública	Deputy Director of Anti-Corruption Conventions	Mexico	
107	Beth Simone Noveck	The GovLab	Project Director	USA	
108	Dinorah Cantú-Pedraza,	The GovLab	Project Manager	USA	
109	Hanna Deleanu	The GovLab	Research Fellow	USA	
110	Rafael García Aceves	The GovLab	Research Fellow	USA	
111	Kaitlin Koga	The GovLab	Research Fellow	USA	
112	Jesse Marks	The GovLab	Research Fellow	USA	