Guide to identify corruption risks in public procurement using data science







) ~ ()



The present "Guide to identify corruption risks in public procurement using data science" has been developed within the framework of the Project "Prevention of corruption through transparency and data analytics in public procurement systems", signed between the Organization of American States (OAS) via its Technical Secretariat of the Inter-American Network on Government Procurement (INGP) and CAF - Development Bank of Latin America, with technical support from Datasketch.

This guide was officially published on December 9, 2021, and is publicly available to INGP member countries to serve as a reference for the continued strengthening of their public procurement and contracting systems.

Prepared by:

Datasketch

Juliana Galvis, Director of Open Government Juan Pablo Marín, CEO and Co-Founder Juan Pablo Garnica, Research Assistant

Coordinators and reviewers:

Organization of American States (OAS) Department for Effective Public Management (DEPM) Helena Fonseca, Public Procurement Program Coordinator and Technical Secretary INGP Hugo Inga, Public Management Specialist

CAF - Development Bank of Latin America

Directorate for Digital Innovation of the State (DIDE, by its acronym in Spanish) Camilo Cetina, Main Lead for Integrity & Governance Practice.

Special thanks to:

Special thanks to the following national public procurement agencies: ChileCompra of the Government of Chile, the National Public Procurement Service (SERCOP, by its Spanish acronym) of Ecuador's Government, Colombia's National Public Procurement Agency —Colombia Compra Eficiente— and the National Public Procurement Directorate (DNCP, by its acronym in Spanish) of Paraguay for their collaboration and their contributions to this document.

Designed by: María Paula Lozano, Advertising Designer.

OAS Cataloging-in-Publication Data

Guide to identify corruption risks in public procurement using data science / [published by the General Secretariat of the Organization of American States and the Development Bank of Latin America, with technical support from Datasketch]. p. p.; cm. (OAS. Official Records; OEA/Ser.D/XXVII.10) ISBN 978-0-8270-7421-7

1. Public contracts--Latin America. Government purchasing--Latin America. I. Organization of American States. Department of Effective Public Management. II. Development Bank of Latin America. III. Inter-American Network of Government Procurement. IV. Prevention of corruption through transparency and data analytics in public procurement systems. V. Datasketch. VI. Series.

OEA/Ser.D/XXVII.10

This work's contents were adapted from work previously done by Datasketch <u>https://datasketch.co</u> in its learning section and published under a Creative Commons open license. All contents of this manual are licensed under a Creative Commons Attribution-NonCommercial-ShareAlike license. This work is licensed under a <u>Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License</u>.



Table of Contents

1. Glossary of terms and abbreviations.	4
2. Introduction	6
3.Background/State of the art	8
3.1 Data	8
3.2 Algorithms	10
3.3 Methodologies to estimate risk	12
4. Strategies to prevent corruption risks	15
4.1 Institutional Readiness	16
4.1.2 Institutional architecture	16
4.1.2.1 Monitoring and evaluation team	17
4.1.2.2 Risk management team	17
4.1.2.3 Technology innovation or information technology team	18
4.1.2.4 Contractual analysis and verification	18
4.1.2.5 Data science team	19
4.2 Information Flows	21
4.2.1 Open data recruitment sources	21
4.2.1.1 Open Contracting Data Standard (OCDS).	21
4.2.1.2 Infrastructure data standard	23
4.2.2 Complementary data sources	24
4.2.2.1 Open data platforms	24
4.2.2.2 Complementary sources associated with the procurement phases	25
4.2.3 Relationship with the institutional architecture.	25
4.2.4 Recommendations on information and data policies	27
4.3 Identifying Corruption Risks Using Data	30
4.3.1 Why identifying risks is useful	30
4.3.2 Red flags	31
4.3.2.1 What are red flags	31
4.3.2.2 Who identifies red flags	31
4.3.2.3 Which red flags are more useful	31
4.3.2.4 List of red flags	33
4.3.3 Risk matrixes	35
5. Architecture to implement an alert system.	39
5.1 Systems for early warning of corruption	39
5.2 Red flags identification and risks prioritization	42
5.3 Indicators selection.	45
5.4 Early warning system architecture.	46
5.4.1 Databases	47
5.4.2 Red flag collection	48
5.4.3 Indicator aggregation algorithm.	48
5.4.4 Risk visualization interface	49
5.4.5 Analysis and verification	50
6. Bibliography	.51
7. References	.53

1.

- Glossary of terms and abbreviations

Algorithm

It corresponds to the set of instructions destined to solve a problem step by step. This means that one instruction is carried out and then another follows sequentially. It is usually understood as the set of instructions given to a computer for it to perform a specific task.

Encryption algorithm

It is used to encrypt or encode data, such as passwords or sensitive information. This way, data is protected during transport between networks. As with passwords, the longer they are, the more secure they can be. An example of this is a hash.

Hash Algorithm

It is also known as hash and it is an algorithm that converts certain data to a series that will always be 40 characters long. For example, CAF's hash is 3908b5520c10387318acb3bf705a4f2a, and it changes completely if it is written in lowercase. Each hash is unique, so "OAS" gets a different one. Being a cryptographic function, what it does is encrypt or "key" a data block.

Matching algorithm

It is an algorithm that arises to solve graph theory problems. This involves creating connections over previously unconnected nodes. An example of the type of problems that these algorithms help to solve is that of assigning the best elements (according to some criteria) of a set A to elements of a set B.

API

Acronym for Application Programming Interface. These are protocols that allow applications to operate with each other through software development. It is understood as a form of information exchange between web services.

Red flags

In public procurement, red flags can be defined as warning signs, clues or indications of irregularities occurring in public procurement processes.



Construction Sector Transparency Initiative (CoST)

It is a global initiative that networks government, civil society and private sector in order to promote transparency within the construction sector. This is done by making available, validating and interpreting data around infrastructure projects.

Dashboard (interactive dashboard)

It is a tool that allows you to organize and present information in an easy and user-friendly way. With interactive dashboards it is possible to filter data and choose ways to visualize and download them.

Data dictionaries

This is the information that is used to expand the knowledge of one of the variables or others elements within a database by taking into account their definition, origin, use and format.

Open Contracting Data Standard (OCDS)

This is a tool created by the Open Contracting Partnership organization to publish structured information at all stages of a procurement process: from planning to implementation.

Machine Learning

The machine learning system is a branch of Artificial Intelligence (AI) that allows automatic learning from a continuously fed database. Therefore, it is capable of converting a data sample into a program that is able to draw inferences from new data sets for which it has not been previously trained.

Metadata

As its name indicates this is "data of data". That is, information that describes the data itself (resources or datasets) such as: titles, descriptions, sources, formats, among other characteristics. It serves to characterize the information and as a way to facilitate user access and information processing.

Open Contracting Partnership (OCP)

It is an international organization of cooperation that has proposed an ongoing dialogue between governments, companies, civil society and technology specialists to transform and transparentize public procurement.

Open Data Charter (ODC)

It is defined as collaboration between governments and civil society organizations to make data available so that it can be collected, shared and used efficiently to address socially relevant problems. In this sense, it works in areas such as anti-corruption, climate action and payment equity.

2.



Introduction

This document's objective is to present a guide to identify corruption risks in public procurement by taking into account the analysis of data collected by various public procurement agencies from different countries around the world. This document can function as a roadmap for procurement agencies to create mechanisms aimed at analyzing information on public procurement and contracting. With this guide, procurement agencies will possess more tools to detect corruption in this particular area of public services.

The document includes a background or "state of the art" section containing initiatives, projects and research that explore public procurement data, algorithms to detect corruption and some methodologies to estimate public procurement risks. It then discusses strategies to prevent corruption risks in public procurement through institutional readiness, information flows and the identification of corruption risks using red flag data and risk matrixes. Finally, it offers an architecture proposal for the implementation of alert systems. This proposal includes the identification of early warnings, red flags and risk prioritization, and, lastly, a brief roadmap to create an early warning system with databases, red flags, indicator aggregation algorithms, risk visualization interface, and information analysis and verification. The following is this guide's roadmap for public procurement agencies or transparency bodies around the world to identify corruption risks:



General outline of the Guide to identify corruption risks in public procurement

This general outline considers and takes into account its proposed combination regarding data science and the expertise that public procurement agencies have, particularly in cases where public procurement systems are not fully built on an electronic/digital basis.

3.

- Background/State of the art

As stated in the introduction, to begin this guide, we present the background or state of the art. This is intended to provide context on some projects, research and initiatives that explore public procurement data and develop standards, methodologies, algorithms and projects used to fight corruption in public procurement around the world.



3.1. Data

When talking about identifying corruption risks in public procurement, the first thing to consider is the type of information that is available for this work. This is why we reference public data and its processing, which are inputs to detect corruption in a country.

To facilitate the task of processing information and detecting procurement risks quickly and efficiently, organizations such as the Open Contracting Partnership (OCP) believe that *"it is impossible to manually track and monitor hundreds of thousands of procurement processes. Publicly accessible open data can make the task manageable, more reliable and credible when used to develop automated analyses to detect procurement processes that require further investigation."*

OCP is interested in raising awareness among governments, businesses, civil society and technology specialists about questions such as, among others: What are governments procuring and for what amounts? How well do procurements perform over time after the companies first bid? How well is a contract executed? Which were the documents provided by the company that won the bid?

To answer these questions, it is necessary to compare historical data from public procurement systems and in many cases complement that information with other data sources. That is why an Open Procurement Data Standard, (OCDS) was created. Based on standardized structures and fields within public procurement databases, it allows structuring, interoperating, reusing, querying and monitoring all stages of the procurement process.

For their part, researchers such as Mihály Fazekas, Professor at the School of Public Policy at Central European University, have repeatedly warned about corruption risks in public procurement and the ways in which they can be mitigated through the efficient use of procurement data in each country. One of corruption detection's major drawbacks comes from the data's scope, because many contracts are reported incompletely or not at all: "*Data quality is also an issue, we tend to find incorrect records (...). Then there are the accessibility challenges: how easy is it to download and analyze the data?*"

Contrastingly, there are civil society initiatives that promote procurement and contracting openness, publication and monitoring in their countries. This is the case of Cuentas Claras³, a Uruguayan initiative of CÍVICO and La Diaria, which systematically publishes journalism articles by using and analyzing open data on government procurement and contracting. Furthermore, this initiative has created the Public Procurement Performance Index, which reflects public agencies compliance in procurement and contracting processes. That index contains 8 indicators and allows to obtain a score on the performance of each public procurement process.

These projects reflect the importance that open data has as a precedent for the fight against corruption in national public procurement systems. Indeed, they allow increasingly detailed, reliable and transparent analyses because they record all the processes and stages of procurement. This way, they provide a broader picture of what has happened and is happening in the system.





3.2. Algorithms

The works on the use of algorithms to fight corruption have yielded several methodologies and specific projects that can be helpful for the local use of public procurement data.

From an academic perspective, Mihály Fazekas proposes some ways of detecting corruption cases in public procurement. In the article *Uncovering High-Level Corruption: Cross-National Corruption Proxies Using Government Contracting Data*⁴ he developed corruption detection measures based on a single tender in a competitive market, and a way of scoring red flags, which could be conceived as an unjustified access restriction to public procurement in order to favor a specific bid within a process.

Likewise, the article *Toolkit* for detecting collusive bidding in public $procurement^{5}$ states that, although there is no single approach to detecting corruption or collusion in public procurement, indicators and tools can be developed to adapt them to all procurement systems around the world. These indicators relate to:

- Increases in the relative values of a contract. This is because any collusive behavior results in an increase in non-competitive prices.
- The variations, range and skewness of the bid price distribution may indicate unusual behavior in competitive markets.
- The bids of the first and second highest bidder must be taken into account for the outcome of a process. The differences between the prices of their bids are particularly important.
- The concentration within a public procurement market of a single or a few companies that are adjudicated all the contracts, in the absence of competitors can become a collusive market.
- A stable market structure is not consistent with natural and competitive market outcomes. In a collusive situation, market shares are practically the same. This depends on the size and openness of the each country's particular economy.

• Success in a given bid is related to a company's individual characteristics. Consequently, there should be no relationship between a company's bidding history in a given market and the awarding of bids.

• The lack of bids from a company that was previously active in a given market may indicate collusive bidding.

• Competition in procurement markets can be distorted by the collusion of competitors who submit deliberately defective bids.

• The creation of a consortium may generate efficiency gains, but it may also reduce the number of competitors in a given market and encourage collusion.

• The existence of subcontracting indicates the active participation of subcontractors and a convenient form of revenue sharing between colluding parties; and as security instrument against possible losses in the contract.

Other initiatives that analyze risk indicators within particular countries are Funes in Peru⁶, Dozorro in Ukraine⁷, and Digiwhist (Digital Whistleblower) in European countries⁸.

In Peru, the Funes project was created by Peru's Ojo Público portal, and is the product of the labor of over 10 people in the course of a year. Funes is a search engine of more than 245,000 public contracts that yields corruption risk indicators in each contracting process by entering the name of a company or entity. In 2019, for instance, it allowed Ojo Público to detect that Peru's main milk supplier was the only bidder in procurement processes for 90% of the contracts it obtained and was awarded more than USD 70 million by the State.

Funes' system indicators are created through an algorithm methodology developed by Mihály Fazekas, which is fed from public procurement databases, campaign contributions, lists of suppliers and the country's Official Gazette. It provides a total of 20 parameters to be analyzed.

Similarly, Mihály Fazekas developed the Digiwhist project to define comparable corruption risk indicators across countries or regions in Europe, thus many of its theoretical concepts and measurement approaches are sought in order to be applied in procurement systems around the world. This project's key objective is to combine public procurement and public expenditure data with actionable governance indicators and a procedure that facilitates citizen feedback.

Finally, the Dozorro project is a model developed by Transparency International Ukraine that analyzes 40 automated risk indicators which are supported by citizen reports (citizens and companies can complain about non-competitive behavior, poor quality or check the efficiency of procurement). From this the Tinder for Tender tool was created, whose model is deployed by experts in public procurement to classify suspicious processes.

By gathering all citizen and company comments in one place, this project was able to collect and manage all complaints more efficiently and resolve them more quickly. It also became a project that tests machine learning with corruption risk detection in the field of public procurement.



3.3 Methodologies to estimate risks

In addition to algorithms, methodologies to estimate corruption risks have been created in the interest of increasingly automating public procurement data analysis in its relation to anti-corruption efforts around the world.

In Uncovering High-Level Corruption: Cross- National Corruption Proxies Using Government Contracting Data, Mihály Fazekas, describes several ways to extract indicators of corruption processes. Among such indicators are surveys to measure citizens perceptions and experiences of corruption, reports from public corruption watchdogs, and audits and investigations of individual cases.

Moreover, the existence of indirect indicators of corruption has also been mentioned. Some of these can be obtained by analyzing the unjustified restrictions that are usually applied in public procurement procedures. These unjustified restrictions seek to favor a specific supplier. Indirect corruption indicators measure corruption risks that may be hidden within the very corruption control mechanisms themselves.

To identify corruption both directly and indirectly at different stages of procurement, Fazekas created a Corruption Risk Index (CRI). This index identifies some components of public procurement according to the process stage. Some of these are:

- Short bidding periods.
- Low number of bidders.
- Low percentage of competitively awarded contracts.
- High percentage of amended contracts.

- Large differences between the award's value and the contract's final amount.

The CRI has been applied by Fazekas in various public procurement systems. It is used to analyze red flags (alerts) and warning signs. This index was applied in the study *Risk of Corruption in the procurement procedures of Petróleos Mexicanos and its Productive Companies*, authored by the civil societies Mexicans Against Corruption and Impunity and Mexican Transparency. The study reviewed a number of contracting processes from Petróleos Mexicanos (Pemex) and its subsidiary companies in order to evaluate the risk of corruption in each process by using risk metrics and risk values for each one.



On the other hand, the Open Contracting Partnership's published guide, *Red Flags for Integrity: Giving the green light to open data solutions*⁹, addressed all the questions related to red flag detection by taking into account expert voices from around the world who have investigated integrity threats in different contexts. This guide's result is a set of over 150 indicators of suspicious behavior or red flags that appear in all phases of public procurement. These indicators or red flags can be reviewed using the OCDS standard and allow for a tangible exploration of red flags in different countries around the world.

Thus, the detection of corruption risks is carried out by compiling red flags or corruption alerts in public procurement. Using these notifications, indicators are developed for each flag by using measurement parameters. Then, equations are mapped using database fields as proposed in the OCDS standard. Finally, the indicators are tested using a variety of data sources to measure integrity in procurement markets. Some of the red flags are¹⁰:

- Short time frame for bidders to submit expressions of interest or prepare bids for competitive processes. The sample threshold is less than 4 days.
- Low average number of bidders for competitive processes. The sample threshold is less than 3 bidders.
- Low percentage of bids awarded through competitive processes. The sample threshold is less than 50%.
- Large percentage difference between the contract award amount and the final contract amount. The sample threshold is a percentage difference greater than 30%.
- High percentage of amended contracts. The sample threshold is more than 25%.

There is another methodology to explore public procurement data used by the Mexican Institute for Competitiveness (IMCO, by its Spanish acronym). It constructs indicators to develop a risk index for each agency within the public procurement system. The first step relates to performing exploratory exercises to find risk patterns. These are based on general statistics produced over the years on the study of buyers, suppliers, types of procurement and also the amount and quantity of acquisitions. The second step is the application of the red flags methodology consisting of the creation of a corruption risk index in the public procurement system¹¹.

A key point among these methodologies is the existence of the red flag system. Such a system allows these methodologies to be used in different public procurement contexts around the world. As such, it provides an approach to measure risk in public procurement systems that is transversal to many countries.

In addition to the examples described above, an exemplary case is that of the Public Procurement and Contracting Directorate ChileCompra¹². It includes an anonymous forum for consultations within the bidding process' framework in which interested persons are allowed to ask questions and submit inquiries. This space has been particularly relevant in identifying patterns of inappropriate behavior, where suppliers can raise their doubts in regard to the practices of other suppliers. This way, a constant cross-review is achieved to safeguard the probity and transparency of the bidding processes, both from the institutional framework and from the suppliers.



4.



Strategies to prevent corruption risks

In order to identify corruption risks in public procurement, certain institutional and regulatory aspects must be taken into account. These will make it possible to have better tools for detecting corruption in countries' public procurement systems.

This section will address institutional readiness, the preparation of information flows, strategies to identify risk, and the creation of risk matrixes to prevent cases of corruption in public procurement systems. Thus, generating a road map to understand this proposal to prevent corruption risks, which starts from the largest scale of public administration, institutional architecture, and ends with a more concrete definition of internal indicators, such as red flags or risk matrixes.

Chapter "4. Strategies to prevent corruption risks" deals with the institutional and regulatory aspects that allow better tools to identify corruption risks in national public procurement systems.

First, the section on institutional architecture deals with the characteristics needed by key officials and teams to prevent corruption. Special attention is paid here to the data science team. Because of its relevance and pertinence in supporting other teams' work, proposals are made regarding the data science team's structure, composition and functions.

Secondly, it provides an in-depth analysis of information sources about public procurement as well as other complementary sources. On the one hand, the data standards proposed by international initiatives and associations are exemplified. On the other, we discuss complementary data sources for the detection of risks in procurement systems in accordance with the Open Data Charter's Anti-Corruption Data Openness Measurement Guide¹³. Such data includes among others: business and beneficial ownership registers; financial operations; tools such as surveys or indicators; and, databases or publishable datasets of actors involved in the procurement process along with their background, This section also analyzes open data's importance, the usefulness of open data platforms or the type of source complementarity in all phases of procurement. In this regard, we establish the applications of the aforementioned sources by teams dedicated to procurement processes to detect corruption risks and make recommendations on data policy according to various aspects.

Finally, the section on identifying corruption risks using data discusses various strategies to identify corruption risks in a public procurement process, such as red flags. Based on these, risk assessment matrixes can be created. Finally, examples of risk matrixes and indicators are presented, such as those of Transparency International and Moldova.

4.1 Institutional readiness

The first thing to think about when creating an institutional strategy to prevent corruption in public procurement is the very structure of these public procurement systems on a large scale. Organizational architecture mainly refers to the personnel and information flows within an institutional environment that will permit carrying out public policy tasks and defining internal processes to deal with alerts of irregularities in public procurement processes.

4.1.2 Institutional architecture

The institutional architecture of a public procurement agency includes those officials responsible in some phases of the procurement processes. Particularly those involved in information and data flows, from the definition and operationalization of procedures to information systems and analysis equipment management. Given the high-risks for corruption in public procurement, institutional architecture should include mechanisms to supervise and control the procurement cycle which promote integrity in contracting and generate valuable evidence on procurement performance and efficiency.

A general risk control cycle within a procurement agency consists of keeping information records throughout the procurement process. In case of suspicious observations, corruption prevention agents within the agencies may inspect the documents and prioritize them. Alternatively, or complimentary, the area in charge of quality management or control examines the prioritized documentation on a random or systematic basis to prevent instances of corruption.

It is important to take into account personnel limitations when supervising public procurement procedures that have alerts by using the "red flags" to prioritize the procedure's supervision rating.

Let's take a look at some key teams, within public procurement agencies, to prevent irregularities.

4.1.2.1 Monitoring and evaluation team

Within the teams of public procurement agencies, there may be monitoring and evaluation departments or internal control departments. These are responsible for increasing efficiency in the management of procurement processes by implementing risk models, tools and reports.

These instances serve various functions, such as validating methodologies; risk models; instruments for system control and monitoring, among others. Within the agency, they are almost always in charge of establishing and managing corruption risk systems or matrixes. Moreover, they have the technical expertise to generate methodologies to counteract risks and help control the system efficiently.



4.1.2.2 Risk management team

Risk management teams will allow directing, analyzing and generating preventive actions in contracts. In so doing, and through the use of national and international data analysis practices, they enable appropriate procurement procedures.

Among their functions is the collection of data to prepare statistics and indicators to manage corruption risks and risk mitigation in a country's public procurement. This team is made up, among others, of information verifiers or contract analysts, who perform their task by means of a manual control of the records.

4.1.2.3 Technological innovation or information technology equipment

In turn, technological innovation or information technology teams may exist. Their function is to plan and develop specific IT tools that contribute to the procurement system's maintenance of operations. Therefore, among their responsibilities are the administration of databases and information on public procurement processes, and managing this information's availability for other teams such as control and risk management teams.

Another instance of this teams' managerial responsibility relates to the automation mechanisms used to verify corruption risks within the procurement system. This way, both the risk control and management teams will possess better tools to mitigate or detect possible irregularities in procurement.



4.1.2.4 Contractual analysis and verification

This team is able to request information about any stage of contractual execution from the different bodies, agencies and municipalities that have signed or executed a procurement contract. Based on the above, these teams carry out specific compliance verifications on the contractual obligations of State suppliers or contractors.

These teams are mainly composed by experts in procurement's legal framework. Likewise, they are based on different risk models that are accompanied by qualitative assessments to determine the work structures and allocation of resources for the review of procurement processes. These teams use different methodologies to designate procurement processes susceptible to manual review by analysts or verifiers. In turn, analysts possess different levels of expertise for the verification of procurement processes and public contracts. From this it follows that the more complex processes must be reviewed by more experienced analysts, while standard processes are usually verified by junior analysts.

Agencies may receive hundreds of daily processes that could be subject to extensive review. In practice, having very large teams of analysts is not cost-effective for mitigating corruption risks in public contracting agencies. Therefore, incorporating data science tools will benefit the analyst team to process assignment flows for manual verification during different instances of indicator prioritization and monitoring systems.

For instance, by analyzing historical data on detected irregularities, artificial intelligence makes it possible to innovate the way in which automatic reviews of certain processes are performed to identify patterns. Even the assignment of these processes could be improved with the use of matching algorithms that take into account multiple metrics and attributes from the procedures and the analysts, their experience and historical performance.

4.1.2.5 Data Science Team

The constant improvement of public procurement information systems and their data flow management procedures, often shared by multiple areas within agencies, and the advances in data science and artificial intelligence technologies, have made it increasingly necessary to think about public procurement agencies' organizational structure and their ability to support a multitude of tasks from different areas in order to make the best possible use of their internally available information.

These data science roles start with teams of data analysts that are dispersed throughout different agencies. This offers the advantage of allowing analysts to assimilate knowledge within their specific area or directorates (e.g., planning, risk control or communications). However, its disadvantage lies in the fact that it does not promote a shared data infrastructure within the organization that would allow the use of all the multiple information sources available to the agency. Above all it discourages the creation of shared capabilities and infrastructure to incorporate technological advances in priority areas such as artificial intelligence.

It is common to see that information management functions are naturally centralized within an agency's technology teams. These teams focus on maintaining the entire transactional systems infrastructure of information flows. However, they may fall short in their mission due to a lack of resources dedicated to advanced analytics processes with historical information and the consolidation of complementary data sources, which often fall under other agencies' responsibilities.

An alternative may be the creation of the data science team dependent on the technology area. In this scenario it would be important to implement and invest in training the data science team to fully understand the procurement system's rules as well as the multiple processes within the procurement agency. In practice, this can be costly, as this type of training relies heavily on the particular experiences assimilated in the day-to-day operation of the agency's different areas.

Another alternative is to have a hybrid model in which a data science directorate exists and has a core team that determines the direction of innovative data science projects. At the same time, it also has analysts working in different agency areas with the objective of identifying and implementing impactful solutions that help to streamline processes. This model allows to align the objectives directly with the technology team in order to make efficient the implementation processes of any developments made by the data science team.

Regardless of whether a centralized or decentralized model is adopted for the science team, it is important to ensure that the data management system at least has the following roles:

- Data Science Manager
- Data Scientists
- Data Engineers
- Business Analysts

When we talk about data science projects, in practice they are related to multiple fronts in an agency's operating processes. Examples include:

• Support and joint work with technology to ensure the information's integrity and quality throughout the information flow phases within the agency: from algorithms to digitalized processes, to historical data integrity verification.

• Support to the communications area to streamline communication processes on data findings for internal collaboration with other areas, or externally with other entities and citizens by creating interfaces to search and access information.

• Support to the planning team in the design and implementation of management metrics and indicators.

• Support to the procurement process control and verification team in the design and implementation of risk indicators.

This way, the aforementioned roles help in the integration of work in which the collaboration mechanisms are data and their associated flows. Management is the instance in charge of the different data science projects with a long-term vision of their implementation's impact within the agency.

Data scientists are responsible for modeling and implementing multiple algorithms to support management. These can range from simple and practical developments that incorporate business rules, to the implementation of advanced Machine Learning methodologies in support of multiple processes.

Data engineers are essential to capture information directly via the transactional systems and databases available to the technology team. They are particularly necessary for direct integration with the agency's operational systems as feedback. At this point, the indicators and metrics defined by data scientists are incorporated into the decision-making processes in order for them to go beyond the data innovation pilot projects.

Finally, the analyst team is paramount to maintain consistency and test the data and algorithms according to the real needs of the other areas. It is the analysts who are meant to work directly with members from the agency's different areas. Consequently, they constantly help to identify opportunities for improvement through data use, visualizations or algorithms in the decision-making process.

This implies that the team has a transversal relevance within a public procurement control institution and therefore needs the means to communicate as efficiently as possible with the institutions' different areas.



4.2 Information flows

In addition to the composition of control, monitoring, management, innovation, analysis and data science teams, the institutional strategy to prevent corruption in public procurement must contain inputs for the analysis of procurement information and other complementary sources that can be associated to the different phases and variables of public procurement. This is because with these information sources it will be more feasible to identify corruption risks.

4.2.1 Open data recruitment sources

These are initiatives created in order to collect a growing amount of information on public procurement in countries, and to create more efficient efforts to fight corruption. One of them is the Open Data Charter's Guide for Measuring Open Data to Fight Corruption, which proposes the publication of a series of datasets¹⁴. Among them are those of public contracts, the list of government contractors, and tendering and procurement processes for use by civil society, public entities, and international agencies that may be involved in the fight against corruption in public service.

4.2.1.1 Open Contracting Data Standard (OCDS)

Open Contracting Partnership is an international cooperation organization that has proposed an ongoing dialogue between governments, businesses, civil society and technology specialists to make public procurement transparent and transform it worldwide¹⁵. To this end, they have placed effort into creating an Open Contracting Data Standard (OCDS), which supports organizations in increasing procurement transparency and enables deeper analysis of procurement data by a wide range of users.

This Standard aims at publishing structured information at all stages of a procurement process: from planning to implementation. Based on this information, data can be used to examine procurement documents to find red flags, i.e., misuse of funds. This organization created a system of public procurement stages that begins with the planning phase, in which -among others- budgets, plans, market studies and public hearings are defined. Then follows the bidding process, which contains the bidding notices, the process specifications, the items needed in the procurement, the bidding values and the process consultations.

The following phase is the awarding process, which includes process details, bidder information, bid evaluation, and contract values. Next is the contracting phase that includes final details, contract signing, amendments, and contract process values. Finally, contract implementation, including contractor payments, progress updates, contract extensions and amendments, and information on contract completion and termination.

For instance, the National Public Procurement Service of Ecuador (SERCOP, by its Spanish acronym) establishes three main stages: planning, pre-contracting and contracting. Planning is the stage in which only the procurement agency intervenes and includes, among others, market studies, plans, and budget definitions. The pre-contracting stage begins with the procedure's publication in the Official Public Procurement System of Ecuador (SOCE, by its acronym in Spanish) and finishes at the moment of awarding the contract to a supplier. Among others, this stage includes the phases of bids receipts, supplier qualification, questions and answers, and negotiation or reverse bidding. Finally, there is the contractual stage which occurs with the awarding of the procedure and lasts until its closing.



Phases of a contracting process. Open Contracting Partnership¹⁶.

Finally, there is the contractual stage, which takes place between the awarding of the contract and its closing. This stage involves public procurement's three actors: the procurement agency, the control entity and the supplier. At this stage, the identification of corruption risks is used to stop any potential corruption process already underway, in other words it is a corrective measure and not preventive as in the previous stages.

Data preparation to build models or red flags:

It would be useful to consider incorporating information about the stages preceding the construction of models or red flags, i.e., the preparation of the data that feeds them.

This document considers the convenience of standardizing data to the OCDS benchmark, as well as the problems that generally arise in respect to data quality where incomplete or erroneous data is common. However, no mention is made of any successful study or experience in which it was possible to overcome the barrier presented by obtaining data of inferior or insufficient quality.

In this sense, it would be positive to reference information sources, documents, papers or techniques that can be applied prior to the construction of models with the aim of obtaining clean data that can subsequently deliver better quality results.

4.2.1.2 Infrastructure Data Standard

Like the Open Contracting Partnership's Open Contracting Data Standard (OCDS), the Construction Sector Transparency Initiative (CoST) has also designed an Infrastructure Data Standard¹⁷. This Standard has been used in over 25,000 infrastructure projects to establish the nature of the data and information that should be disclosed at each stage of the project cycle, taking as variables budget, status, implementation time, among others. CoST has identified¹⁸ 67 key information points of disclosure for public works projects to help stakeholders monitor these infrastructure systems.



4.2.2 Complementary data sources

In addition to the primary sources of information on public procurement, the following can be considered other types of information that serve to contrast or complement the primary ones.

There are a number of developed tools to detect risks in procurement systems. Among them are some indicators about the system's integrity or the context, these include surveys or indicators that evaluate the judicial system's quality, the political system and bureaucracy, as well as transparency budget within the analyzed country. The latter takes into account standards related to access to information, conflicts of interest, and budget planning of each public entity involved in the procurement process.

Moreover, in most cases there are databases directly related to public procurement tables, such as: records of suppliers or bidders, records of public officials involved in procurement processes, records of goods and services for specific events or projects, sanctioned contractors, and product unit prices.

Regarding the Open Data Charter's Anti-Corruption Data Openness Measurement Guide¹⁹, one could consider the series of 30 datasets they suggest that countries should publish, as follows: register of lobbyists; declarations of interest; register of companies; register of charities, public servants involved in contracting, politically exposed persons; list of contractors, contracts, tenders and procurement, sanctioned or disqualified contractors; complaints in procurement processes; government advisory councils; political party financing, licensing, public-private partnerships, budget and spending, subsidies and international financing, audits, parliamentary votes, judicial decisions, government projects; record of meetings, regulatory changes, campaign promises; public records of property, taxes, and asset declarations.

4.2.2.1 Open data platforms

Open data platforms can be a way to connect public procurement data with other data sources through an overall government anti-corruption system. In many cases, such platforms have the capacity to interoperate information that serves as input for the analysis and investigation of corruption and, consequently, to enable increasingly sound public policy decisions. As an example, of these initiatives, as a result of the cooperation of different public institutions in Colombia, the PACO (Personal Analytics Companion) platform is being implemented, which seeks to strengthen and improve the processes to detect cases of corruption in the country.

In addition to the above, it is possible to propose control panels or dashboards that allow the constant identification of indicators that can be used to make public policy decisions in the fight against corruption.

4.2.2 Complementary sources associated with the contracting phases

There are complementary sources of information for each phase of the Open Contracting Partnership's proposed procurement phases, these could include:

• Planning: the procurement agency's action plans, allocated budget and justification of the need for contracting.

• Bidding: conflicts of interest of those involved in the procurement process, bidders legal background, implementation of previous contracting processes by bidders.

• Awarding and contracting: implementation of the winner's previous contracting processes, cost of bid items vs. average cost of items in previous contracting processes.

• Implementation: quality of the contractor's delivered products, contractor's products compliance according to the schedule, changes in product prices throughout the contract's implementation.



42.3 Relationship with the institutional architecture

These sources of information are useful for teams with responsibilities in procurement processes because they can use these information flows to detect corruption risks.

Within the monitoring and evaluation teams, public procurement data standards can be related to the efficiency of procurement processes management. This is because they allow to know the appropriate variables to establish and manage process follow-up systems or matrixes.

With the use of national and international data analysis practices such as those performed by OCP and CoST, risk management teams can use procurement standards to drive ever-improving procurement procedures. Implementing those standards, these teams will be able to gather statistics and indicators for corruption risk management. They will also be able to use other information sources to complement their risk matrixes and allow better parameterization of these cases.

Technological innovation or information technology teams can develop IT tools that contribute to the maintenance of procurement system operations based on already established initiatives at the international level. Since their responsibilities include managing procurement processes databases and information, they can rely on information standards such as those established in the OCDS to make this information available at a broader local level. In fact, with the implementation of these standards in other countries, international measurements or automation for risk verification executed by organizations such as OCP or CoST are feasible.

The contractual analysis and verification team may constantly require information regarding any stage of a contractual execution from the different agencies, entities and municipalities that have signed or executed a contract. Due to this reason, the team would constantly need to use the variables proposed in the standards, and even the recommended data sources for each stage of the process. Because agencies may receive many processes that could be subject to revisions, it is necessary to secure other information flows for a better analysis of the information by the data science tools.

For its part, to ensure data integrity and quality throughout the phases of the information flow within the agency, the data science team can include support: from algorithms to digitalize processes, to historical data integrity checks. Such support can be easily carried out with data standards at all stages of the procurement process. In addition to the above, is the development of tools to cross-reference official contracting information with other data sources to prevent and counteract corruption in the system.



42.4 Recommendations on information and data policies

To make progress in data science technologies it is advisable to incorporate a data policy and guidelines for its use within the procurement agency. In turn these guidelines should respond to and lead the work on information flows by taking into account the following aspects:

Organizational structure

Definition of who is in charge of the information, what types of data and at what times or phases are they involved in the process. With this in mind, the actions of the entity or unit, in terms of data, are oriented in accordance with its responsibilities and organizational mandates; at the operational level of human resources and at the technical level such as with technological information systems. The entity or unit's actions must consider the different actors that may participate in the processes, such as, among others, legal offices, technology teams and communications office.

Data life cycle

1. Definition of the data's lifecycle throughout the process of capture, management, and publication of information.

2. Identification of stakeholders and times of access and interaction with the information, so that information can be efficiently controlled to facilitate its quality and publication.

3. Mapping and delimitation of data structures and formats to create the institutional information catalog.

4. Incorporation of guidelines and best practices for standardization and data formats that are mapped to data capture, migration and archiving techniques to achieve a relevant data categorization that guides its appropriation.

5. Characterization of data with relevant descriptors within the metadata and a clear methodology to manage and update each dataset.

6. It is also important to have a segmentation that clarifies the data's nature. For example, to clearly differentiate processes, the transactional data from the archival data verified after processing the primary information sources.

Data preparation

It is of great importance to consider the stages prior to the construction of corruption risk identification models that contain the data preparation that feed them.

For this, the OCDS data review tool could be used to check the data's structure and format for compatibility with OCDS tools and comparability with other OCDS data. This tool reports any structural problems, and checks whether the data make sense.

Operability of publication and use of data

To ensure the technical and operational feasibility of information flows, it is important to consider metadata. Metadata is commonly defined as "data about data" and describes the content, quality, format and other characteristics of a resource or dataset. Metadata provides a mechanism to characterize information, as well as to access and processing paths for consumption by internal and external users.

Legal and regulatory aspects

Consider regulations such as the mandates and functions of entities and agencies. Relevant legislation on personal data protection or access to public information constitute the legal framework to support the agency's actions in regard to the publication and reuse of information on open procurements.

Among other things, this legal and regulatory framework allows to:

- Analyze and identify possible risks when opening personal data.
- Generate public confidence.
- Mitigate the agency's legal risks.
- Incorporate processes, if necessary, for database anonymization such as Hash algorithms, encryption, data perturbation, data reduction, etc.
- Incorporate appropriate licenses for the publication of open data with the assessment of different types of licensing.

Additionally, as an organizational mission and vision, compliance with international recommendations for the implementation of Open Government is evaluated, as well as the incorporation of online data projects in accordance with the Sustainable Development Goals (SDGs).

Technical Aspects

Technical considerations on the implementation of technologies and platforms for data reuse within the procurement agency and publication in open data formats. They also include specific recommendations to create metadata, apply programming interfaces (APIs), data dictionaries to facilitate interoperability with various systems, as well as data visualization tools to promote the reuse of information from interactive dashboards or user-friendly interfaces to create graphs and maps.

Usability

Identification of possibilities to improve the systems with the objective of complying with different accessibility criteria in accordance to specific populations: from young people to people with some type of disability. Evaluation of different implementation scenarios with specific recommendations to comply with international accessibility standards, as well as with legal provisions on access to information.

Identifying end users, both internal and external, is recommended to promote suggestions for data usability. Such suggestions range from formats to visualizations that allow the exploration of information in a simpler way and promote the use of public data itself within the organization. Finally, this point emphasizes the data's usability for machines, allowing the discovery of datasets and promoting the information's interoperability with other systems.

Appropriation strategy

At this point different techniques and recommendations are explored that generate a community or audience around the published data. The strategical objective is reusing information, as a branding strategy for the procurement agency itself in terms of transparency, but also as a strategy to reuse and integrate multiple internal databases within the organization. As it raises the awareness of other products or channels of the procurement agency, this strategy can complement other communications strategies by incorporating data-driven content and visualizations that consider the latest trends in data journalism.

Monitoring and evaluation of open data policy

This stage refers to the definition of metrics and indicators of use for the available data by using data analytics that allow monitoring and evaluation of the data's' impact within the organization in order to improve processes and external use with an extended community of users.





4.3 Identifying corruption risks using data

Now, with the suggested institutional and regulatory structure, what follows presents a series of strategies to identify corruption risks in public procurement data.

4.3.1 Why identifying risks is useful

In the public sector, risk could be thought of as the possibility of enduring economic or resource losses as a result of uncertainty, mismanagement or corruption in public procurement. There are measurable risks at each phase and stage of the process that help to understand the main areas of exposure faced by a procurement process. This helps to assess and plan the practical management and operational measures to be taken in order to mitigate those risks.

Therefore, to mitigate risks, public procurement agencies should perform risk identification analyses, particularly in the stage prior to the execution of contracts. This way, they will be able to identify the risks' origin, probability and magnitude. This will help to better focus on the issues that warrant greater attention because they already have national and international precedents, or simply because the regulations and the context make it clear. Moreover, risk analysis should be part of every stage of procurement planning and should be updated periodically.

Both foreseeable and unforeseeable risks may be determined in a procurement process. Foreseeable risks refer to those circumstances that may arise during the development and execution of a contract, and have the capacity to alter its financial equilibrium. On the other hand, the non-foreseeable risk is that which is not contemplated within the regulations, hence these would be circumstances that are not easily identifiable or quantifiable.

4.3.2 Red Flags

Almost every instance in which corruption risks in public procurement are discussed, references are made to the use and classification of red flags which indicate possible corruption risks in public procurement actions.

4.3.2.1 What are red flags

In public procurement, red flags can be defined as warning signs, clues or indications of possible irregularities. This does not imply that they exist, but it does mean that a particular procurement process requires special attention in order to rule them out or confirm them.

The presence of red flags should prompt staff and managers to increase vigilance: they should take the necessary steps to confirm or deny the existence of a corruption risk. In the meantime, procurement and managing authorities have the responsibility to clear up any doubts raised by the red flag.

4.3.2.2 Who identifies red flags

Procurement agencies in each country usually define some red flags that are applicable to their own countries or jurisdictions. Based on their experience and the experiences of other regions, they can create a series of warning indicators for each country's procurement systems, and thus make sound public policy decisions. These red flags are almost always set in accordance to the type of information that procurement systems can process, for instance, the number of bidders in a tender or whether there were additions to the contracting process.

Moreover, red flags can take various forms in relation to procurement agencies and their moment-tomoment needs. Accordingly, they may want to detect behavior that falls short of what the public would expect; behavior that is ethically or morally wrong; mismanagement through error or mistake; negligence or inadvertence; or misconduct that is intentional or deliberate and may be considered a crime. To this end, they look for different ways to detect the red flags that better serve their specific needs at a given moment. In some cases, flags can be incorporated directly into public procurement transactional information systems, which allow the definition of rules to publish and monitor processes. This way, they mitigate some of the most frequent red flags.

4.3.2.3 Which red flags are most useful

One of the main ways of knowing what type of red flags are most useful to detect corruption is by identifying those processes that do not resemble others with similar characteristics. This is particularly useful to know what type of irregularity may have existed in any of the procurement phases. On the other hand, some corruption schemes could be determined in order to define what possible irregularity may have occurred in each process.

Below are some corruption schemes to keep in mind:

• Collusion between contractors: an agreement intended to achieve a certain benefit at the expense of a third party through fraudulent means is rendered evident. This type of agreement may be obvious or hidden depending on the type of connection between contractors.

• Collusion between the contractor and the bidder: to favor the contractor over other bids (this fact has no international consensus, but is related to the impediment of competitive procurement processes).

- Irregularities in the pre-bidding process: including, among others, changed timing or little justification for bidding without competition.
- Bid rigging: sharing the bidding process in off-the-beaten-path locations, short bid opening time, few bidders, among others.
- Tenders with only one bidder.

Moreover, red flags must be indicators that can be contrasted in some way with public procurement data or other types of information. These, in turn, must have the quality to be compiled by procurement agencies or other public entities. Ultimately, the indicators are expected to measure corruption risks. To round up, with large numbers of contracts the means to determine the most useful and automatable red flags is to contrast them with existing variables in public contracts and with variables in other useful complementary data sources.

In its classification of red flags, the Open Contracting Partnership (OCP) proposes that the Open Contracting Data Standard (OCDS) be used to determine red flags' parameterization based on areas and procurement processes. Therefore, for each red flag identified by OCP, some columns and fields taken directly from the standard are indicated and recommended. These are published in a red flag mapping database during the phases of public procurement for their classification and identification²⁰.

National legislation

The easiest way to determine red flags in public procurement is by referring irregularities to each country's legislation. There may exist a general law on public procurement, or other laws aimed at strengthening public procurement, or even shielding sectors with high corruption risks such as infrastructure or public works.

Among the provisions normally contained in national red flag legislations are the following:

- Complete and efficient contract execution.
- Persons or companies disqualified from contracting.
- Evidence of deviations or abuses of power.
- Mechanisms for a control institution to intervene.

Administrative resolutions could be created in cases where there is no national legislation on the subject and the strengthening and regulation of public procurement is complicated. Resolutions of this type could allow progress to be made in the management of corruption risks in a country's public procurement systems. Incidentally, they generate evidence that helps to carry out more binding legislative processes at the national level.

4.3.2.4 List of red flags

In order to identify the greatest possible number of red flags that trigger corruption risk alerts, a list of these indicators has been made using various information sources that determine and characterize them. This list contains: the red flag's name and a description; information about which phase of the procurement process it belongs to; whether the indicator is related to the country's procurement regulations or whether it is deduced from the process itself; its association with corruption alert schemes; the source proposing this red flag; whether there is any kind of parameterization for the red flag; whether it has a unique identifier for Open Contracting Partnership; and, whether it has any columns related to the Open Contracting Data Standard (OCDS).

Some red flags established by the Open Contracting Partnership²¹ and directly related to OCDS are:

- The time between contract advertising and bid opening is short, an indicator that is directly linked to a time standardization for the different procurement methods.
- Similarities in the documents submitted by different bidders, these have to discard purchase specifications and instead focus on contractor ranges and suppliers.
- Adjudications to a single bidder in contravention to the procurement agency's acquisition planning.

• Division of the same procurement process to avoid the acquisition amounts thresholds for each process.

- Large number of contracts adjudicated to a single bidder.
- Payment of unjustifiably high prices in relation to the historical average.
- Large difference between the contract price and the winning bid price.
- Complaints of changes to bids after being received.
- Close relationship between a bidder and buyer.
- Certain assets that were never solicited in the past and/or will not be requested in the future remain on contracts.

Moreover, several international organizations and agencies have created some of the most important recommendations that are present in some of the following red flags:

• European Union (Hungarian standard)²²:

- The contract is renewed (several times or for a longer period).
- The bidder's technical capacity does not comply with national legislation.
- An accelerated procedure is used.

• OECD²³:

- Lack of an adequate assessment of needs.
- External actor's influence on officials' decisions.
- Poor procurement planning.

• European Commission²⁴:

• The document drafter organizes the process in such a way that there is no time to review the documents thoroughly before the end of the tender procedure launched.

• Procurement processes that were previously competitive are replaced by non-competitive ones.

• The bidding process was opened to a single bidder.

• World Bank²⁵:

- Use of sole sourcing or direct contracting when the procurement plan requires the use of more competitive methods.
- Inadequate or misleading documentation as required by the acquisition guidelines.

• Inadequate responses or clarifications from project officials to complaints from bidders on vague, ambiguous or incomplete specifications.

4.3.3 Risk matrixes

After the identification of red flags in the different stages of public procurement, a series of risk matrixes can be created. Their purpose is to determine some characteristics for corruption risk in procurement systems including the areas of greater exposure and some indicators to mitigate them.

In general, agency control teams have two ways of receiving these risks. In the first, complaints from bidders or citizens who report irregularities serve as direct input to coordinate the team of analysts or verifiers, who are able to add information with metrics counting verified complaints. The second way, which is presented in this section, uses the already established indicators and elements that trigger process alarms.

The areas that present higher risks could be determined by: sectors or activities (with higher volumes of contracts); by types of public institutions (national, regional, or sectorial); by contract value (in which contracts with higher costs tend to have higher risks), which increases the chances of corruption in procurement.

Some of the risks that a public procurement system may classify are:

- Legal risks: due to irregularities in the preliminary studies, in the terms and conditions of the contract.
- Financial and/or the economic risks: due to irregularities in the adjudicated amounts, reference budgets, among others.
- Market risk: due to irregularities in a process competitiveness, a special regime, single negotiation, among others.
- Linkage and effectiveness risk: due to cancelled, deserted, terminated processes, among others.

Most methodologies prioritize risks according to their potential impact and their probability of occurrence. Thus, risks are situated on a plane in which they are low when the potential impact and probability of occurrence is low, and high when these indicators are more probable. This is proposed by Transparency International in the document *Corruption Risk Assessment and Management Approaches in the Public Sector*²⁶.



Transparency International. Corruption Risk Assessment and Management Approaches in the Public Sector. Moreover, it is necessary to determine the level of each of these impact and probability indicators by determining how many and what levels of each indicator need to be measured, and deciding how to describe each indicator. Here is an example:

L	Probability level	Probability level description
	Almost certain	The risk is expected to occur within the normal course of events.
	Possible	The risk may occur at some point in the future.
	Rarely	The risk may occur only in exceptional or unlikely circumstances.

Transparency International. Corruption Risk Assessment and Management Approaches in the Public Sector.

Source: Adapted from CCECC no date

Impact level	Impact level description		
Low	The risk will have an insignificant effect on the organization's reputation or its ability to meet its objectives.		
Medium	If not stopped, the risk can have a significant effect on the organization's reputation or its ability to meet its objectives.		
High	Risk, by its consequences, can threaten the organization's stability and the fulfillment of its objectives causing significant economic damage, jeopardizing the success or efficiency of the organization's operations.		

Transparency International. Corruption Risk Assessment and Management Approaches in the Public Sector.

Source: Adapted from CCECC no date

Other impact and probability risk matrixes can be created according to each case's general perception and international experience. This is the case of Moldova²⁷, a country that has this risk matrix for possible corruption:

	Probability	Impact
Bribery and kickbacks	High	Medium
Conflicts of interest	High	Medium
Collusive bidding	High	High
Shell companies	Medium	Medium
Leaking dib data	Low	Medium
Unbalanced bidding	Low	Medium
Manipulation of the bidding procedure	Low	Low
Supply purchases	Medium	Low
Rigged specications	Medium	Medium
Excluding qualified bidders	Medium	High
Unnecessary purchases	Low	Medium
Implementation	High	Medium
Donation to political parties	High	High

Techniques for data collection and analysis for the relevant indicators. Probability and impact of corruption risks.



Similarly, there are some indicators that can be evaluated from different points of view and by different control and information verification institutions within the system. By way of example, in Moldova, mechanisms were created to identify corruption risks using matrixes such as the following:

	Red flag What is to be checked?		Who should check?	
	Red flags for bribery and kickbacks			
	Enhanced financial standing of public officials	 Spot checks and checks in suspicious cases (complaints, audit, findings, reports by NAC or CoA) For public officials and their relatives: Declarations to be submitted to NIC Bank accounts balance, transfers) Real estate registers Car registration Company register (existence of shares in companies) 	NIC	
	Intermediaries	Compete business information of contractor and subcontractors in order to identify shell companies (company registration, history of tax declarations and social security contributions) Recommendation: Subcontracts should only be allowed if they are indicated in the bid, and if the role of each of them is clearly justified.	NIC	
l	Repeated awards	Regular review of tender.gov.md to identify repeat awards to the same contractor, then review details (process level, implementation performance) of suspicious cases	PPA Civil society	
	Systemic issues	 Working group composition to be reported annually by all procurement authorities Analysis of duties segregation Recommendation: Rotation of members Collective procurement in contracting authorities with insufficient staff (particularly at local level) 	PPA	

Techniques for data collection and analysis for the relevant indicators. Probability and impact of corruption risks.

In a simpler way, templates containing information about possible risks in a procurement process or processes could be created, as follows:



This will allow for a larger corruption risks structure and the possibility of anticipating any of them.

5.



Architecture to implement an alert system

5.1 Systems for early warning of corruption

The system should be understood as a whole that incorporates the organizational, procedural and technological elements, which represent the bulk of the operation and are part of a public procurement agency's different strategies to fulfill their mandate and their national regulations. An organization's structure and workflow are fundamental for the efficient operation of public procurement agencies. This procedural work can be accompanied by information technologies to make processes more agile and efficient. The impact of processes and procedures in an agency's institutional reforms requires long term planning and implementation times. Conversely, supporting the already defined operation with technologies can represent a great opportunity in the application of corruption fighting procedures.

An example of how systematization can mitigate corruption risks has to do with integrating technological tools that incorporate red flags into the public procurement process flow itself. An instance of this are the transactional public procurement systems that have been implemented in several countries. These systems incorporate their business rules directly into the platform, which closes the door to tampering of parameterizable red flags, such as, for example: publishing tender documents exclusively online during the minimum number of days indicated in the regulations, or forcing suppliers to correctly upload all the process information in order to continue advancing in their applications. There are other types of red flags that could be parameterized in order to stop them from the very process of bid publications or supplier applications. However, in practice the information normally captured by the procurement agency represents integration challenges with external databases. That is the case of flags that require the cross-referencing of information with complementary sources, such as business records or sanctions from controlling bodies. In such circumstances, the flags could indicate issues external to the procurement agency's information control and thus represent new challenges related to the centralization and federalization of databases.

Often, however, there are more challenging red flags to analyze that cannot be parameterized directly in public procurement transactional systems. For example, forcing the number of bidders to be a minimum of 3, or ensuring that bid prices are set within the standard range for historical bid prices. Both options have different implementation issues. The minimum of 3 bidders would be immediately discarded in the case of CPCs with only one bidder at the national level. While historical data will continue to be updated constantly, therefore if a coalition of suppliers takes over a market they will be able to raise prices without raising alerts.

Moreover, red flags or risk indicators that arise from the historical analysis of big data can be incorporated. These can use machine learning algorithm technologies to provide insights for implementation with more sophisticated indicators. The principle of these indicators is to train a robot to learn how to identify common patterns, for example, with the history of hiring processes that are known to have had irregularities so that the robot can raise alarms about similar patterns occurring in new processes.

Different red flags represent different challenges to calculate. In practice, there are a large number of red flags that can be calculated and incorporated, regardless of the procurement agency's technological level. This was evidenced in the risk matrix methodology, where these systems can be implemented in a spreadsheet or even incorporated directly into transactional procurement systems. Red flag analysis incorporated directly as a system represents multiple opportunities to identify corruption risks. The relevant work then consists on the way in which the flags can be incorporated into the current flows and processes of publication, verification and analysis of public procurement processes in procurement agencies.

An early warning system using corruption red flag analysis methodology should have the following characteristics.

• Provide timely information: allowing timely detection of alerts.

• Offering quality information: to provide a true record of the reality of the situation and thus to be able to make better decisions.

• To be aligned with local regulations: to be based on the regulations of each public procurement system, and taking into account the explicit and implicit legal restrictions on certain acts.

• To be aligned with internal processes: alerts should aim to comply with the procedures, composition, regulations and systems within a country's procurement systems.

• Offer action alternatives: so that the agency can act on the information depending on its functions, or stop processes; also make recommendations or inform control bodies about possible irregularities.

A point to consider in terms of the implementation of a red flag early warning system for corruption is the responsible agency itself. In other words, a Centralized Procurement Body or each public agency may be in charge of the implementation process, which has different implications. Particularly, public agencies will have first-hand knowledge of the procurement and contracting processes that take place there. To this end, they may establish specific or tailor-made red flags.

An example is MercadoPúblico²⁸, the transactional platform for Chile's Public Procurement and Contracting Directorate ChileCompra, which is a decentralized and transparent system²⁹. This enables not only efficiency with each public agency carrying out its procurement and contracting, but also process transparency and equal conditions for suppliers given the free public access to information.



5.2 Red flags identification and risk prioritization

It is up to each agency to determine which red flags to implement. To this end, red flag prioritization exercises can be carried out according to the ease of their implementation. Mainly, this has to do with the logistical personnel resources, technological tools, available data, and the execution time needed to detect the flag efficiently. On the other hand, it must recognize the impact that these red flags can generate in the exercises of preventing and fighting against corruption.

The following images are sample exercises. A matrix could be created with a plane indicating greater and lesser ease for implementation, and another indicating the greater or lesser impact of red flags. In turn, these could be placed according to their different characteristics.



Matrix with axis graph: Ease of implementation vs. impact of red flags.



Example of matrix with axis graph: Ease of implementation vs. impact of red flags.

Likewise, valuation matrixes can be created according to ease and impact. Incidentally, scales from 1 to 10 can be determined on the Cartesian plane from lowest to highest. These two indicators are then added together to obtain the rating, so that red flags with higher ratings have a greater chance of being implemented.

Red flag	Easiness	Impact	Valuation
The need for the contract is not properly justified.	3	10	13
The bids were not included in the official procurement journal.	3	5	8
Contracts with a single bidder	8	9	17

Example of matrix with axis graph: Ease of implementation vs. impact of red flags.

This matrix can be created with the different potential and existing procurement risks. Existing risks can be understood as those consequences produced by irregularities in procurement. These have a particular and visible impact on the procurement system: delays in processes and extra costs arising from mismanagement or misuse within the system. This means that a number of certain red flags were found to be true.

This matrix can be defined using the greater or lesser frequency of occurrence for the event, and, at the same time, a greater or lesser impact for its implementation. This way, the risks that have a higher valuation can be prioritized to be systematized together with the selected red flags.

Risk	Facility	Impact	Valuation
Preventable process delays	3	10	13
Extra expenses in the processes	6	6	12
Closing the market. Bidding without competitive process	9	10	19

This exercise will allow us to have a better perspective on the possibilities and needs to prioritize each red flag and risk in procurement, and thus determine a system of alerts in case it is necessary to determine which are more relevant for each context.



5.3 Indicators selection

Once the red flags that are to be incorporated in the systems have been prioritized and an assessment is made of the most important risks to be covered, it is necessary to add a selection of indicators. At this point the red flags are incorporated with appropriate metrics that measure the risks that the procurement processes may have.

These indicators refer mainly to those characteristics that are parameterizable in the procurement processes and that will result in quantitative data that allows the differentiation of some procurement processes from others. For example, when there are certain agencies that, because of their characteristics, constantly resort to direct procurement, this should not necessarily raise a red flag, but it should for other agencies. Some examples of indicators include

- Number of processes per agency.
- Number of direct procurement processes per agency.
- Proportion of direct procurement processes by agency.
- Maximum values and services in a goods procurement process.
- Maximum values in a public works procurement process.
- Number of processes that have additions per agency.

When selecting the most appropriate indicators, the following can be taken into account:

• Data normalization: Data should consider appropriate normalizations so as to not give it more weight. For example, an agency that has many small contracts versus one that has few large contracts.

• Representation: It is important for samples to be balanced, so that an indicator is not underrepresented for one type of agency or one type of process.

• Contextual adjustments: there are contextual issues that need to be taken into account because certain procurement agencies or processes may raise red flags unnecessarily. For example, agencies that have a large amount of direct procurement or those that have high amounts of money destined for infrastructure or public works contracts.



5.4 Early warning system architecture

The following presents the architecture of an early warning system based on red flags intended to mitigate corruption risks. The system has different components, in which there are several sources of information that function as input for the collection of red flags and for indicators to be implemented. It also has a risk assessment module based on an algorithm. This allows it to gather the knowledge from the multiple red flag alarms in a single indicator, preferably with a graphic interface that classifies the riskiest processes susceptible to verification. Finally, actionable paths are shown based on the risks in order for the different entities or persons responsible for the processes to provide feedback to the system based on the findings.



5.4.1 Databases

As discussed in previous sections of this document, a first step in the architecture of early warning systems is to consider the sources of public procurement data which allow the tracking and monitoring of each of the procurement processes. These databases can be records from each process, with different variables related to the system, such as, the process's opening date, contract's object, and items to be procured.

In many cases, public procurement databases are available in electronic systems, which allow online consultation of the information and, if applicable, a closer look at a particular process.

Moreover, there are other types of data that can help enhance the fight against corruption in different countries around the world, such as those presented in the *Openness Guide: Using Open Data to Fight Corruption*³⁰ developed by the Open Data Charter. It proposes 30 priority datasets that can be used in different countries to fight corruption. The proposed datasets include:

- Registration of lobbyists
- Declaration of interests
- Company registration
- Registration of charitable organizations
- Public officials involved in procurement processes
- Politically exposed individuals
- Registration of public officials
- List of government contractors
- Government advisory councils
- Political party financing
- Budget
- Procurement processes
- Licenses
- Among others³¹

The Guide also proposes that data can be mutually interlinked and that international standards for different types of data, such as the OCDS standard for public procurement or the Fiscal Data Package created by the Open Knowledge Foundation, can also be used.

This list of data sets can be very useful for cross-checking information on public procurement with other information published by each country's agencies. This is intended to provide greater input in the construction of red flags for the detection of corruption risks in the country.

5.4.2 Red Flag Collection

The next step in the architecture of early warning systems refers to the collection of red flags and the methodology to prioritizing them. These are some steps that can be followed to create a collection or list of red flags to be systematized.

The first thing to think about is making a list of possible red flags that may alert irregularities in procurement processes. Among others, this list may contain the name, description and type of measurement. Based on this list, the flags that can be used are determined using the available information and taking into account the public procurement database's variables and other complementary sources. Once these red flags are prioritized, it is necessary to establish a methodology that determines the most relevant ones according to the public procurement system context. This way, it will be possible to have a battery of red flags that will serve to determine the alerts within the system.

It is important to clarify that, in addition to the red flags identified from the suggestions of multiple organizations such as OCP, OECD or the World Bank, there are other red flags that can be calculated by incorporating algorithms into the calculations. Such is the case, for example, of metrics that analyze the natural language of contracts to automatically extract relevant concepts that can be used to characterize risks. Another example are classification algorithms that complement the verification work. These incorporate risk alarms with historical data analysis trained from historical valuations or penalties on the bulk of contracts.

5.4.3 Indicator aggregation algorithm

Once the collection of red flags has been created, the next step is to implement them in the procurement system. One of the biggest hurdles when running the collection of flags has to do with the amount of information that a set of flags can yield per day. As an example, if you have 30 prioritized red flags, and every day you have 100 new processes to review, you would have to manually review 3000 indicators that may represent irregularities per day. This review would be impractical given its volume. Therefore, it is not enough to have good red flags identified, it is also important to group and consolidate risk information into simplified indicators.

The most efficient way to solve this challenge is to aggregate the different red flags into a series of risks grouped by different themes such as: foreseeable risks in the procurement process, extra expenses in the processes or market closures. These can be thought of as the consequences of improper actions in the procurement system. However, in the end they should be consolidated into a few indicators to streamline resource allocation processes.

To do this, there are methodologies that rank the risks from highest to lowest according to the total risk valuation and take into account that not all procurement agencies have the same variables because the behavior of each one is different. There are different methodologies to group multiple risk indicators into a single one as subjective indicators of corruption risk. These indicators can be constructed from total numbers and sums of red flags or as arbitrary cut-off points on the totals of contracted amounts and processes participations. There are even more elaborate ones, for instance the incorporation of methodologies such as Principal Component Analysis or the inclusion of Machine Learning algorithms to have more balanced indicators.

Some points to keep in mind in the creation of these indicators is that information availability is fundamental in order to have good aggregate indicators. In the end, these indicators are completely influenced by the selected red flags and correspond to general corruption metrics. In the best scenario, these metrics correspond to a lower limit on possible corruption risks because in some cases elaborate mechanisms can be used to avoid detection of irregularities by actors wanting to tamper with processes. Finally, another important challenge is that, over time, the nature of what can be considered as an act of corruption is determined by local and current regulations. In this regard, in some cases care must be taken with the incorporation of metrics that may vary over time due to regulatory changes or even changes in the institutional design of the procurement agencies themselves.

5.4.4 Risks visualization interface

One way of directly knowing a system's risks is to develop data visualization interfaces which allow to know the information's main findings graphically and grouped in order to make decisions on certain contracts or bidding processes.

Data visualization provides visual guides so that encoded information can be understood by the human eye. This allows for conclusions to be drawn quickly and easily by seeing graphic information such as a larger ratio, a peak or trough in a graphic plane, or a deviation in a constant measurement.

These visualizations can be carried out through information filters, for instance, by type of public entity, by departmental or municipal level, by type of procurement processes, and thus give much more context to the information being analyzed. Information interaction dashboards such as these make it possible to monitor particular types of procurement and expenditures for particular issues. For example, during the COVID-19 global pandemic, dashboards were created to monitor emergency procurement because these processes were susceptible to high corruption risks because of their uncompetitive nature within the procurement of goods and services.





It should be taken into account that the information obtained by analyzing the red flag system can become sensitive and directly affect the market if it becomes public knowledge. It can also allow groups that seek to abuse public procurement to obtain undue benefits and to take advantage of the system and evade the alerts. For this reason, it is necessary to define which alert information will be open to the public and what information will be reserved for the control institution's internal use.

5.4.5 Analysis and verification

Once it is possible to have an interface that identifies the information's main findings, the next step is to make the appropriate decisions according to the level of each actor's impact in the review of public procurement systems.

By stakeholders we refer to those persons or teams that act in relation to corruption risks and the verification of procurement processes. These may be:

- Analysis and/or verification teams within each procurement agency.
- Control bodies established in each country's regulations to verify, investigate and, if applicable and necessary, to sanction the procurement processes or actors involved in it.
- Other civil society actors such as NGOs, citizen watchdogs, and journalists who have their own systems to detect irregularities and processes to investigate and corroborate information.

These systems may be used as inputs to provide feedback to other indicators that raise different types of alerts, such as the number of sanctions a contractor has, whether the supplier has been previously disqualified, whether there exist conflicts of interest in the procurement system, among others.



Bibliography

Audit Scotland. Red Flags Procurement. https://www.audit-scotland.gov.uk/uploads/docs/um/fraud_red_flags_procurement.pdf

Bence Tóth, Mihály Fazekas, Ágnes Czibik & István János Tóth. 2015. Toolkit for detecting collusive bidding in public procurement. Corruption Research Center Budapest. Working Paper series: CRC- WP/2014:02. http://mihalyfazekas.eu/wp-content/uploads/2015/08/Toth-et-al_CRCB_WP_v2_150413.pdf

CÍVICO y La Diaria. Cuentas Claras. https://www.gub.uy/agencia-reguladora-compras-estatales/sites/agencia-reguladora-comprasestatales/files/2020-05/Datos%20Abiertos%20-%20C%C3%ADvico%20Uruguay.pdf

Comisión Europea. Detección de conflictos de intereses en los procedimientos de contratación pública en el marco de las acciones estructurales. https://ec.europa.eu/sfc/sites/default/files/sfc-files/guide-conflict-of-interests-ES.pdf

Dirección Nacional de Contrataciones Públicas (DNCP). Paraguay. Funciones. https://static.dncp.gov.py/PublishingImages/Lists/noticias/AllItems/FUNCIONES_PDF.pdf

Elizabeth Luengas Pinzón. 2015. Cómo estiman, tipifican y asignan los riesgos en la contratación las entidades sometidas al estatuto general de la contratación de la administración pública. Universidad Militar Nueva Granada. Diplomado en contratación estatal. https://repository.unimilitar.edu.co/bitstream/handle/10654/6690/RIESGOS%20EN%20LA%20

CONTRACION%20ESTATAL.pdf;sequence=1

Ernesto Cabal y Gianfranco Huamán. Ojo Público. "Precios de insumos e historial de empresas se deben fiscalizar para prevenir corrupción". https://ojo-publico.com/1812/covid-19-historial-de-companias-y-sobreprecio-son-banderas-rojas

European Commission. 2010. Risk management in the procurement of innovation. https://ec.europa.eu/invest-in-research/pdf/download_en/risk_management.pdf

Instituto de Finanzas de Castilla-La Mancha. Banderas rojas en la lucha contra el fraude en contratación pública. http://www.icmf.es/doc/antifraude/Banderas_Rojas_Contratacion.pdf

Ilse Schuster & Serghei Merjan. February 2016. Assessment Report of corruption risks in public procurement in the Republic of Moldova.

Interagency Procurement Working Group (IAPWG). UN Procurement Practitioner's Handbook. Chapter 4: Transverse Procurement Themes. https://www.ungm.org/Areas/Public/pph/ch04s01.html#sect_4 Manual Kibana OCDS. Introducción al Estándar de Datos de Contrataciones Abiertas. https://manualkibanaocds.readthedocs.io/es/latest/C1/Seccion1.html

Mexicanos Contra la Corrupción y la Impunidad y Transparencia Mexicana. 2020. Unidad de Investigación Aplicada. Riesgo de Corrupción en los procedimientos de contratación de Petróleos Mexicanos y sus Empresas Productivas.

https://contralacorrupcion.mx/wp-content/uploads/2020/11/riesgo-de-corrupcion-en-pemexmexicanos-contra-la-corrupcion-y-la-impunidad.pdf

Miguel Jorquera. 2019. Compras Públicas y Big Data: Investigación en Chile sobre índice de riesgo de corrupción. Espacio Público. Banco Interamericano de Desarrollo. https://www.repository.fedesarrollo.org.co/bitstream/handle/11445/3871/Repor_2019_Jorquera_ EP.pdf?sequence=1&isAllowed=y

Mihály Fazekas & Gábor Kocsis. 2015. Fazekas. Uncovering High-Level Corruption: Cross-National Corruption Proxies Using Government Contracting Data. European Research Centre for Anti-Corruption and State-Building Working Paper No. 46. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2711932

Nicolás Penagos y Óscar Hernández. 2019. Examinando con datos las «banderas rojas» de compras en América Latina. Open Contracting Partnership. https://www.open-contracting.org/es/2019/06/27/examinando-con-datos-las-banderas-rojas-decompras-en-america-latina/

Observatory of Public Sector Innovation. OECD. Dozorro. https://oecd-opsi.org/innovations/dozorro/ Open Contracting Partnership. Estándar de datos: Para Empezar. https://standard.open-contracting.org/latest/es/getting_started/

Open Contracting Partnership. RED FLAGS for integrity: Giving the green light to open data solutions. https://www.open-contracting.org/wp-content/uploads/2016/11/OCP2016-Red-flags-forintegrityshared-1.pdf

Organización para la Cooperación y el Desarrollo Económicos (OCDE). Preventing Corruption in Public Procurement. https://www.oecd.org/gov/ethics/Corruption-Public-Procurement-Brochure.pdf

Servicio Nacional de Contratación Pública (SERCOP). 2017. Ecuador. Estatuto Orgánico de Gestión Organizacional de Procesos.

https://portal.compraspublicas.gob.ec/sercop/wp-content/uploads/2017/11/Estatuto_org%C3%A1nico_aprobado2017.pdf

Transparency International. 2015. Corruption risk assessment and management approaches in the public sector. Anti-corruption helpdesk.

https://knowledgehub.transparency.org/assets/uploads/helpdesk/Corruption_risk_assessment_and_ management_approaches in the public sector 2015.pdf

Transparency International. Public Procurement: Topic Guide. https://knowledgehub.transparency.org/assets/uploads/kproducts/Public_Procurement_Topic_Guide.pdf

7.

References

019/06/27/examinando-con-datos-las-banderas-rojas-de-compras-en-america-latina/

- https://ojo-publice.com/10012/10010-historial-de-companias-y-sobreprecio-son-banderas-rojas
- 2. http://cuentasclaras.uy/
- 3. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2711932
- 4. http://mihalyfazekas.eu/wp-content/uploads/2015/08/Toth-et-al_CRCB_WP_v2_150413.pdf
- 5. https://ojo-publico.com/especiales/funes/
- 6. https://dozorro.org/
- 7. http://digiwhist.eu/
- 8. https://www.open-contracting.org/wp-content/uploads/2016/11/OCP2016-Red-flags-for-integrityshared-1.pdf
- 9. https://docs.google.com/spreadsheets/d/12PFkUIQH09jQvcnORjcbh9-8d-NnIuk4mAQwdGiXeSM/edit#gid=0
- 10. https://www.repository.fedesarrollo.org.co/bitstream/handle/11445/3871/Repor 2019 Jorquera EP.pdf?sequence=1&isAllowed=y
- 11. https://www.chilecompra.cl/
- 12. https://opendatacharter.net/
- 13. <u>https://airtable.com/shrwhz2OG3MYSIcUs</u>
- 14. Open Contracting Partnership. Quiénes sómos. https://www.open-contracting.org/es/about/
- 15. https://standard.open-contracting.org/latest/es/getting_started/contracting_process/#defining-a-contracting-process

16. Open Contracting Partnership. Manual de Contrataciones Abiertas para el Estándar de Datos sobre Infraestructura. https://standard.open-contracting.org/infrastructure/latest/es/

- 17. https://infrastructuretransparency.org/wp-content/uploads/2018/06/36 List of CoST Project Information.pdf
- 18. <u>https://opendatacharter.net/</u>
- 19. https://docs.google.com/spreadsheets/d/12PFkUIQH09jQvcnORjcbh9-8d-NnIuk4mAQwdGiXeSM/edit#gid=2027439485
- 20. https://docs.google.com/spreadsheets/d/12PFkUIQH09jQvcnORjcbh9-8d-Nnluk4mAQwdGiXeSM/edit#gid=2027439485
- 21. https://www.redflags.eu/files/redflags-summary-en.pdf
- 22. https://www.oecd.org/gov/ethics/Corruption-Public-Procurement-Brochure.pdf
- 23. https://ec.europa.eu/sfc/sites/default/files/sfc-files/Fraud%20in%20Public%20Procurement_final%2020.12.2017%20ARES%282017%296254403.pdf
- 24. https://docs.google.com/spreadsheets/d/12PFkUIQH09jQvcnORjcbh9-8d-NnIuk4mAQwdGiXeSM/edit#gid=2027439485
- 25. https://knowledgehub.transparency.org/assets/uploads/helpdesk/Corruption_risk_assessment_and_management_approaches_in_the_public __sector_2015.pdf

26. Ilse Schuster & Serghei Merjan. February 2016. Assessment Report of Corruption Risks in Public Procurement in the Republic of Moldova.

- 27. https://www.mercadopublico.cl/Home
- 28. https://www.chilecompra.cl/que-es-chilecompra/
- 29. https://docs.google.com/document/d/182USj4La896XgMZ-LTw7A3HYbA5i7aP0p2JLBL8a-24/edit#heading=h.5tohb79d3s3d
- 30. Lista completa en https://airtable.com/shrwhz2OG3MYSIcUs/tbIHKFQsAdVNa1TC9



INTER-AMERICAN NETWORK ON GOVERNMENT PROCUREMENT

Guide to identify corruption risks in public procurement using science of data





