iMonitor

iMonitor 2.0

Enhancing law enforcement efficiency by bringing together public procurement data analytics and civil monitors

Project number: 101217271

Corruption Risk Indicators in Public Procurement: An updated opentender.eu framework

Risk assessment methodology for the iMonitor2.0 project

Deliverable nr. D2.2

Dissemination Level:	*PU - Public*
Nature of the Deliverable:	*R - Document, report*
Date:	*29/12/2025*
Work Package:	*WP 2*
Author(s):	*GTI*
Editor(s):	*GTI*
Reviewers:	*All Partners*
Contributors:	*All Partners*

Co-funded by the European Union

Corruption Risk Indicators in Public Procurement: An updated opentender.eu framework

Risk assessment methodology for the iMonitor2.0 project

Authors: Ágnes Czibik, Bence Tóth, Dániel Kofrán, Isabelle Adam, Mihály Fazekas, Mihály Nagy (Government Transparency Institute)

Date: 29/12/2025

Abstract

This risk assessment methodology report presents the tools that the iMonitor2.0 project uses for measuring corruption risks in the public procurement process, which will serve to select high-risk contracts for civil monitoring in the project. After introducing the public procurement process, the report focuses on the key aspects of measuring corruption, which starts with an actionable definition of corruption in public procurement and a reasonably broad overview of common corruption technologies in the different phases of the public procurement process. Based on this theoretical foundation, a thorough discussion on corruption measurement approaches and the characteristics of the data employed is introduced, highlighting issues related to its scope, depth, accuracy and accessibility. Moving on to the practical application of the framework on opentender.eu, the eleven individual integrity indicators and the Average Integrity Indicator Score available on the platform are presented in detail, specifying their connection to specific corruption risks and the relationship between corruption and integrity that underscores the conceptual background behind opentender.eu. The report concludes by discussing the usefulness as well as the limitations of such a risk assessment approach, touching upon how it must be complemented by further in-depth assessment and monitoring work, as developed in the iMonitor2.0 project.

Deliverable of the Enhancing law enforcement efficiency by bringing together public procurement data analytics and civil monitors (iMonitor2.0) project, led by Government Transparency Institute (Hungary) with the financial support of the European Commission (project no. 101217271).

Table of Contents

Executive Summary

Public procurement is the foundation for the activities of governments and state-owned enterprises, encompassing the purchase of goods, services, and works. In the European Union, approximately 14% of GDP is allocated annually through public procurement processes, playing a significant role in critical sectors like energy, transportation, health, and education services. Even marginal enhancements in this market can yield substantial savings, with a 1% efficiency gain potentially saving €20 billion annually.

The public procurement process comprises four main phases: Planning and advertisement, Selection and submission, Evaluation and award, and Contract management. The Planning and advertisement phase starts with detailed needs assessments and the development of a public procurement plan including specifications, implementation timelines, and procurement details. Here, budget estimation takes place and sufficient funding has to be secured for the project. The subsequent Selection and submission phase involves bidders expressing interest, pre-screening bids, and detailed specification review by bidders, allowing for questions and modifications to the tender documents. Crucially, this is when bidders submit their bids following the tender specifications and award criteria.

The Evaluation and award phase includes a minimum eligibility assessment and comprehensive evaluation, leading to the selection of the winning bidder(s). In the final phase, the contract is implemented. This implies goods delivery or project execution by the supplier, the buyer verifying deliverables, and payment. In some cases, contract renegotiation might take place, modifying contract value or the substantive content of the contract such as deadlines. Each phase of the procurement cycle faces specific corruption risks, underscoring the need for a comprehensive anti-corruption framework.

A widely accepted, general definition of corruption identifies it as the "abuse of entrusted public authority for undue private interest." While this definition emphasizes the deviation from serving the public interest, it is considered too broad for guiding quantitative measurement in public procurement. A more specific definition applicable to public procurement asserts that corruption aims to favor a connected bidder to the detriment of other competitors. This involves avoiding competition through unjustified sole sourcing, biased specifications, and other means. Corruption in this context is particularistic, institutionalized, and grand, often involving high-level politicians and business persons.

The subsequent sections of the report focus on corruption risks during the public procurement process. Examples of corruption during policy formation and monitoring stages are briefly introduced, underscoring the need for a holistic understanding. The report then presents a dictionary of typical corrupt schemes specific to different phases of the public procurement process.

Firstly, the Planning and advertisement phase can give rise to corrupt schemes such as tinkering with the advertisement period length, splitting public procurement processes into smaller tenders to circumvent competition, avoiding publication of the call for tenders, and selective information provision benefitting a connected bidder. These schemes manipulate the procurement environment, hindering fair competition and transparency. Secondly, biased product specifications and eligibility requirements are highlighted as examples of corruption

schemes in the Selection and submission phase, deliberately tailoring requirements to limit the pool of potential bidders.

Thirdly, corrupt techniques in the Evaluation and award phase may include the use of shell companies to evade proper eligibility assessment, strategically annulling tenders to initiate less open procedures, and unfair scoring, particularly using subjective evaluation criteria. Lastly, corruption during the Contract Management phase may involve "bogus" subcontracting, where subcontractors are used to obscure particularistic ties, and the delivery of substandard work or goods. These practices emphasize the importance of comprehensive quality checks.

The report outlines the methodology used to estimate or approximate such corruption schemes through big data analysis and also emphasizes the importance of high-quality public procurement data for such analyses. The primary data sources are national public procurement websites and the EU-wide Tenders Electronic Daily (TED). These different portals follow a range of publication rules leading to very different degrees of publication rate in each country (i.e. some countries publish more of their total public procurement activities on these portals than others).

The evaluation of data quality is a primary step in public procurement data analysis. Key considerations such as data scope, depth, accuracy, and accessibility are thoroughly examined for the seven iMonitor2.0 countries. Data scope, largely determined by national contract value thresholds, varies across countries, influencing the transparency of the sector and comprehensiveness of public procurement analyses. A breakdown of data depth reveals discrepancies in the coverage of the tender cycle across countries, with only few providing information on the contract implementation phase. The report highlights data accuracy levelsin iMonitor2.0 countries, which has greatly improved in recent years. Accessibility, the fourth dimension of data quality, is categorized into structured, semi-structured, not fully machine-readable, and no public database. The majority of analysed procurement systems fall into the structured or emi-structured category, underlining the importance of the data collection and standardization, as done by opentender.eu.

The report then moves to the practical application of the methodology on opentender.eu and introduces the eleven integrity indicators used on the platform to assess corruption risks in public procurement. In this framework, integrity is simply the absence of corruption, i.e. procurement procedures displaying a high level of integrity based on the indicators have a low level of corruption risk, and vice-versa. The uses and limitations of these indicators are also highlighted. On the one hand, the integrity indicators serve as tools for assessing the level of corruption risk associated with individual tenders by identifying potential corruption red flags, and provide a systematic way to evaluate and prioritize tenders based on their measured corruption risk. At the same time, corruption can be masked in many different ways, bringing with it the inherent challenge of distinguishing between clean and corrupt tenders and developing a 100% accurate indicator. To ensure the precision of corruption risk assessment, each indicator undergoes rigorous validation processes.

The report also emphasizes the importance of combining these indicators into a composite index, to enhance the robustness of assessing integrity. The Average Integrity Indicator Score, representing the composite index, is calculated as an arithmetic average of all available integrity indicators, ranging from 0 to 100. Validation against established corruption indices, like the Control of Corruption indicator of the World Bank, demonstrates the composite index's validity.

1. Introduction

Public procurement represents about one third of government spending across the European Union (EU), including in countries that are part of the iMonitor2.0 project. Such a high amount of spending is crucial for public services and the very functioning of modern governments. However, it is widely documented that public procurement is prone to corruption even in otherwise high integrity contexts. Curbing corruption in public procurement, typically in its high-level and well-organized forms, however, is challenging due to the difficulties of detecting it.

Building on its predecessor project iMonitor (project number: 101103267), the iMonitor2.0 project addresses that challenge by further developing an innovative public contract monitoring model that combines data-driven risk assessment with on-the-ground monitoring by trained civic monitors. The goal is to provide actionable information to public authorities to better identify potential corruption and fraud in public procurement. We implement our approach in 7 European countries/regions: Italy, Romania, Spain (Catalonia), France, Poland, Bulgaria, and Albania.

Supporting the goals of the iMonitor2.0 project, this report aims to introduce the corruption risk assessment methodology and resulting indicators used to analyse public procurement data, to inform civil monitors of the project and also other stakeholders who make use of opentender.eu, the main data analytics tool of the project. The report walks the reader through public procurement measurement concepts, offers a thorough review of data quality, and specifically describes the integrity indicators deployed on opentender.eu and used in civil society monitoring of public procurement.

The report is organised as follows: first, it introduces the reader to the field of public procurement, reviewing the public procurement process and its main steps. Second, it spells out the project's measurement approach according to the underlying corrupt behaviors the indicators aim to proxy. Third, the report discusses the data quality requirements for the approach, and presents an overview of public procurement datasets for the seven countries covered by the iMonitor2.0 project. Fourth, the integrity indicators calculated for the project countries are precisely defined. Finally, the report concludes with enumerating the main uses of the measurement framework and its limitations, and the appendices provide further detail on the data work behind the indicators.

2. What does the public procurement process entail?

"Public procurement refers to the purchase by governments and state-owned enterprises of goods, services and works," as defined by OECD[1]. In the EU, around 14% of GDP is spent through public procurement processes, annually. There are certain sectors, such as energy, transportation or provision of health or education services, where public procurement is responsible for a great proportion of the transactions between suppliers and buyers (i.e. contracting public authorities). Due to the sheer size of the public procurement market, even small improvements can add up to large savings: a "1% efficiency gain could save €20 billion per year" according to the European Commission[2].

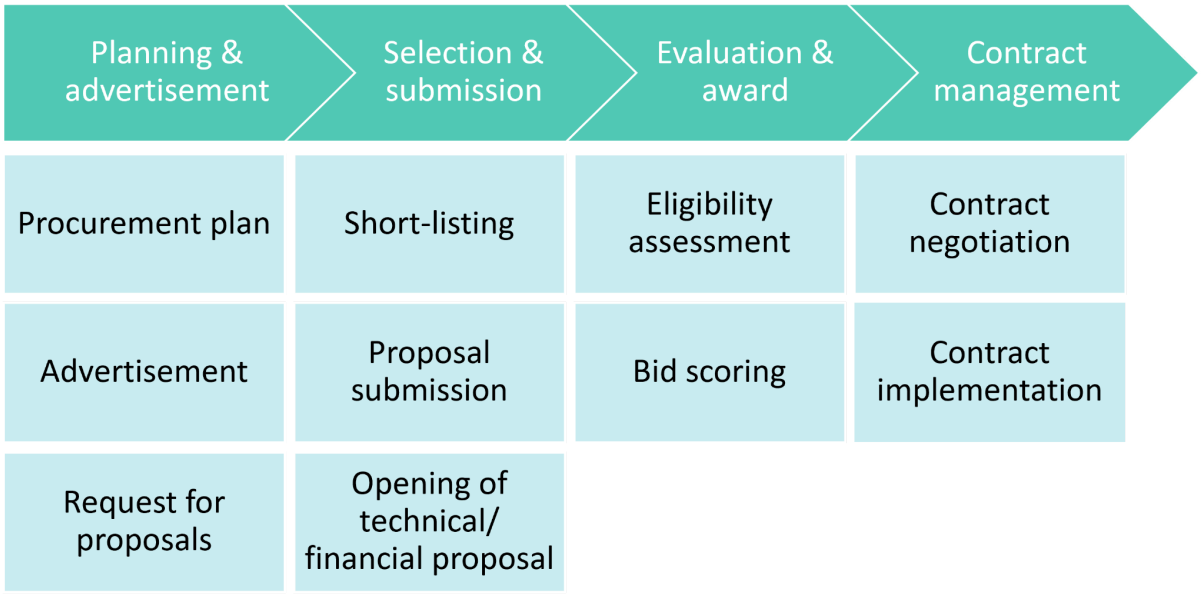
The public procurement process includes four main phases:

[1] https://www.oecd.org/governance/public-procurement/
[2] https://single-market-economy.ec.europa.eu/single-market/public-procurement_en

1. the planning and advertisement phase,
2. the selection and submission phase,
3. the evaluation and award phase, and
4. the contract management phase.

They all consist of smaller steps which follow each other in close proximity (see Figure 1).

Figure 1: The public procurement process



Source: own illustration

The **planning and advertisement phase** of any procurement procedure begins with the planning process, during which the contracting authority decides on the specifications of its purchase. This, among others, includes the initial price and volume estimates, the planning of the implementation timeline, the description of the subject matter of the procurement, the decision on the number of lots (separate portions or divisions of a larger contract that can be bid on and awarded independently) and, if relevant, the decision on the location of implementation. After the tender documentation is finished, the authority decides the method of advertisement (e.g. e-auction, non-electronic procurement) and bidding process (open, restricted/negotiated, or direct contracting) and the request for proposals begins.

The above process is followed by the **selection and submission phase** when bidders submit their initial expression of interest, which is used by the contracting authority to pre-screen bidders. Detailed technical and financial specifications are also made available, which usually further narrows the pool of eligible suppliers. At this stage, bidders usually have the opportunity to ask questions regarding the procurement and, to some extent, they may rectificate or modify their proposals.

The **evaluation and award process** starts with the minimum eligibility assessments, the goal of which is to filter down the list of bidders to the ones that – on paper – meet the requirements listed in the tender documentation. Then, the proposals are evaluated, compared and scored and the winning bidder(s) selected.

The final phase consists of **contract management**, including the execution of the project, the issuing of payments and the monitoring of deliverables, and the possible renegotiation of

the contract. Each phase displays its own corruption risks, which will be discussed in the next section.

3. Key steps of measuring corruption

In order to ensure that we develop valid and reliable corruption proxies, we review the main steps suggested by the literature[3]. This approach is used to measure the presence and intensity of corruption in public procurement. The necessary components that should be specified at this stage are the following:

1. The specific **definition** of corruption: to measure corruption, it is important to specify what is to be measured. A sufficiently specific definition of corruption should be used, that can be quantified and objectively evaluated.

2. The **dictionary** of corruption techniques: a reasonably broad repository of corruption strategies that are regularly being used in the procurement system helps the formulation of quantitative tools.

3. The **target population and sample**: the scale and scope of the procurement dataset, and level of the observations (e.g. tender, contract, or lot level dataset).

4. The **tailoring and validation steps** of quantifiable corruption risk indicators: the selection and validation of quantitative indicators that can correctly detect corruption techniques in the public procurement framework. Validation is an essential step to ensure that the indicator is an adequate measure of corruption. An inadequate measure can introduce measurement bias and can create a false picture about corruption risk in the procurement system.

The first two steps are described in the next two sections. Target population and sample are discussed in detail in section 5 (Data), and tailoring and validation are introduced in section 6 (Integrity indicators).

4. Understanding corruption in public procurement

This section follows the suggested steps of developing a corruption measurement framework in public procurement. Hence, first, a definition of corruption is introduced as the basis of the measurements and indicators discussed later. Second, some of the most typical corrupt schemes are presented to provide insights on the possible risks in the different phases of the public procurement process.

4.1. Defining corruption in public procurement

A widely used definition of corruption states that public corruption is the abuse of entrusted public authority for undue private interest. This definition assumes that universal interest is enshrined, hence public money should serve public interest. Consequently, deviation from this should be sanctioned as corruption leads to a social and economic loss.

This definition is too broad to adequately measure corruption in a public procurement setting. A more specific definition to be used in this analytical framework is that, in public

[3] Mungiu-Pippidi, Alina & Fazekas, Mihály (2020), “How to define and measure corruption”. In Alina Mungiu-Pippidi & Paul M. Heywood (eds.) *A Research Agenda for Studies of Corruption*. Ch. 2. Edward Elgar, Cheltenham.

procurement, “the aim of corruption is to steer the contract to the favored bidder without detection”[4]. This can be done in several ways, such as avoiding competition through direct contract awards or other non-competitive procedures, or favoring a specific bidder with tailored specifications or by sharing inside information, among other strategies.

The main assumption of this definition is that a corrupt procurement process has some form of manipulation in place to avoid competition, and this is used to favor a certain bidder. If the second half of the assumption is not applicable, it suggests that there is some other issue with the procurement process (e.g. the incompetence of the contractor), hence corruption is not intended, and the problem requires a different policy approach. Furthermore, corruption in public procurement must entail cooperation between public and private actors. In this sense, it can be distinguished from bidder collusion (e.g. a cartel), for instance, in which case suppliers coordinate among themselves to split a market and limit competition, without the necessary involvement of the public contracting authority (i.e. the buyer). Nor is there a necessity for law infringement, as procurement rules can be bent to allow corrupt practices, which means that corruption is not necessarily illegal.

Therefore, in the current setting, corruption is **particularistic, institutionalised and grand**. There is a particularistic (often personal) relationship between the actors involved in corruption and they try to exclude anyone who is not part of their interest group. It is institutionalised, indicating that it is recurrent, stable and systemic. Finally, it usually involves high-level politicians and business persons and involves a large amount of public funds.

4.2. Identifying typical corrupt schemes

The next sections of the report mainly focus on corruption risks during the public procurement process; however it is important to mention that corruption can already occur prior to the process, during the policy formation, as well as in the later monitoring stages. First, the section introduces some brief examples of corruption during these two stages to create awareness of the bigger picture. Then, the section showcases a dictionary of typical corrupt schemes that occur in the different phases of the public procurement process.

In terms of corruption in policy-making prior to the procurement process, adjusting procurement regulation can be an efficient way of limiting competition with the added benefit that economic actors do not have to break any rules in order to take advantage of the limited access to procurements. For example, persuading politicians to increase the value thresholds for non-open procurement procedures or to create special conditions for using negotiated ones exhausts the definition of institutionalized corruption which can only be dealt with at the highest levels. Regulatory entities and auditing bodies can also be corrupted to ignore complaints, partially settle disputes, or overlook the particularistic relationships between individual contractors and bidders. After the procurement procedure, corrupting the monitoring stage is also handy to create paper trails suggesting that everything went well during the procurement process. Although this form of corruption already requires breaking the law, it is often very difficult to detect, especially when the public procurement system is fraught with red tape.

---

[4] World Bank (2010). *Fraud and corruption awareness handbook: how it works and what to look for - a handbook for staff* (English). Washington, D.C.: World Bank Group.

.

During the actual procurement process, the contracting authority can specify needs to favor a certain bidder, it can misjudge the quality of other bids to suppress competition, or it can change the contract after the winner has been selected. In an entrenched system of corruption, a typical corrupt scheme might include all the stages of the procurement process. It can start by the supplier contacting a procurement project director (e.g., through a particularistic, personal tie) who is responsible for the management of some high-value projects. Then the director can instruct the technical evaluation committee to favor the bidder by, for example, overestimating its capabilities. In return, the supplier pays the project director and usually – as we defined corruption as being high-level (grand) and institutionalized – a high-ranking politician (such as a minister) who turns a blind eye on the process. It is important to mention that the payment does not have to be a bribe, rather a payment for a "consulting contract", or, in a well-oiled system, it can even take the form of a personal favor. Finally, the winning firm can also pay the accountant to sign-off the contract, and at the end of the implementation phase it might bribe the supervisor who evaluates the quality of the output.

After mapping the procurement process and specifying the corruption definition that is to be used for the analysis, it is important to identify common techniques that are being used to corrupt the procurement system. Finding well-documented examples of high-level corruption cases can help in the selection and formulation of objective and quantitative indicators. Therefore, the next part of this section outlines some of these techniques, while also highlighting the importance of substantive qualitative research. The cases are presented in association with the phases and steps of the public procurement process discussed in section 2.

4.2.1. Planning and advertisement phase

4.2.1.1. Shortened advertisement period

Most of the more developed procurement systems have a minimum advertisement period length in place, however, contracting authorities might be able to use policy loopholes to shorten it. A sufficiently short advertisement period makes competition impossible, because competitors will not have the time to obtain necessary documents, prepare the tender documentation, or to calculate their expenses to prepare their bids. Therefore, if a favored bidder receives insider information about the tender before it is advertised, it will most likely be the only one able to submit its expression of interest in time.

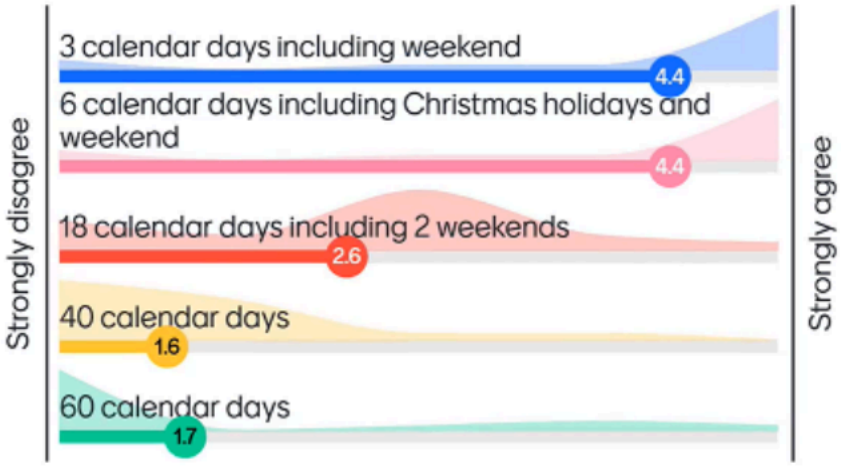
Imagine, for example, a road reconstruction project between two medium-sized cities where the winning bidder has to repair a 25 km, two-lane road. Before reading any further, let's try a brief exercise to think about what a sufficiently long advertisement period would be, based on Table 1 below.

Table 1: Sufficiently long advertisement period for a road reconstruction project

1.	3 calendar days including weekend
2.	6 calendar days including Christmas holidays and weekend
3.	18 calendar days including 2 weekends
4.	40 calendar days
5.	60 calendar days

This exercise was conducted with participants of the first specialized regional training for the Regional Good Governance Public-Private Partnership Platform[5]. They considered that options 1 and 2 are insufficiently short advertisement periods, whereas options 4 and 5 should be enough for a bidder to prepare all the necessary tender documentations (see Figure 2). However, opinions about option 3 (a 18-day advertisement period) already varied even among experts: a little more than half of them considered that it is adequately long. This highlights that a "sufficiently long" advertisement period can diverge across sectors and countries. In a country with a low level of red tape and a well-developed e-governance system, 18 days could be enough to obtain all the necessary documents, but in other, more bureaucratic systems, even up to 40 days may be too short. As the next sections will show, data analytics can help to decide upon thresholds below which procurements should be considered risky.

Figure 2: Experts' opinion on the sufficient advertisement period length for a road reconstruction project



Source: GTI.

[5] https://www.govtransparency.eu/first-specialised-regional-training-for-r2g4p-members/

4.2.1.2. Splitting procurement procedures to circumvent competition

It is simplest to illustrate splitting procurement procedures to circumvent competition through a hypothetical example: A municipal educational organization needs to purchase 100 computers (desktop units and monitors) in order to equip computer labs. The buyer's total budget for the tender is 100,000 euros. The average cost quoted by suppliers in the region that meet the buyer's requirements is at 700 euros per computer. The buyer intends to sign a contract with a specific supplier (the favored supplier) and purchase the equipment at an unreasonably high price. If a fair and competitive tender was to be held, there is a high probability that another supplier would win the contract.

To prevent this from happening, the buyer decides to split the contract into smaller purchases in order to be able to buy the goods from the favored supplier. While a single procedure with an estimated cost of 70,000 for all 100 computers would require a more competitive procedure, procurement regulations allow the buyer to directly award contracts below a lower threshold (e.g. 40,000 euros) directly, without a competitive selection.

Taking these circumstances into account, the buyer concludes four contracts with its favored supplier over the course of three months, each for the supply of 25 computers at 1,000 euros each. This way, no competitive tender is conducted and the buyer exhausts the available budget, allowing the selected supplier (and likely himself) to extract a good personal margin from the purchase. Of course, in reality such schemes may be more sophisticated to avoid easy detection, but at the same time, small volume contracts tend to face less scrutiny, and even a simple tactic like this might be overlooked by oversight bodies.

4.2.1.3. Avoiding publication of the call for tenders

The call for tenders can be published in different locations or publication can be avoided altogether. The most transparent place for publication is the Official Journal of the European Union or a national Public Procurement Bulletin which can guarantee the highest number of potential bidders get informed about the tender (e.g. there are for-profit providers who recycle and disseminate procurement notices to potential bidders). If buyers decide to only publish the call for tenders on their homepage it still can be considered as transparent, but much less so, as potential bidders may find it harder to monitor hundreds of individual homepages as opposed to a central national page of public procurement announcements. When no call for tender is published at all, but instead it is informally sent to selected bidders, the principle of transparency is violated most extensively.

In the EU, the location of publication is tightly regulated by the EU Public Procurement Directives as well as national Public Procurement Laws, typically requiring publication in the national Public Procurement for large tenders, while buyers can effectively choose the place of publication for smaller tenders. If they exercise this choice in favor of less transparency, leading to a lower number of bidders, it can be suspected that the buyer may have ulterior motives to limit competition, thus raising corruption risks.

External control of this corruption technique is very hard. When contract award announcements appear in the official government public procurement bulletin, their references to prior call for tenders can be checked but in European countries, tenders are hardly ever canceled due to a missing call for tenders.

4.2.1.4. Selective information provision

Communication between bidders and contracting authorities is heavily regulated in EU member states because it can have a decisive impact on competition. Getting better or more timely information on tenders can put some bidders in an unbeatable position. This is exactly what corrupt informal networks use to win public contracts in seemingly fair competition in a lot of countries. It is enough to informally give crucial information on specific aspects of the tender to one bidder while issuing a vague or erroneous tender specification, based on which unconnected bidders will be unable to prepare a valid or competitive bid.

The use of this corruption technique infringes on the principle of fair competition as well as transparency. There is little effective external control on any of these information flows as it is difficult, if not impossible, to monitor informal talks and information transmission through intermediaries.

This technique is related to the choice of procedure type as less transparent procedure types, such as negotiation, make it very easy to provide information to one bidder while concealing it from other bidders. Furthermore, shortening advertisement period works neatly in tandem with this technique as early informal 'heads-up' of the preferred bidder of a future call for tenders with very short submission deadline gives it a decisive competitive advantage (e.g. in extreme cases it can be the only one actually able to put together a valid bid).

4.2.2. Selection and submission phase

4.2.2.1. Biased product specifications

In order to restrict the number of short-listed bidders, a procurement procedure may include biased technical tender specifications, deliberately tailored to exclude bidders other than those with connections to the contracting authority. For instance, biased specifications for the goods or services being procured can intentionally define them in a way so that only one or very few connected companies are able to deliver them, thus restricting the pool of potential suppliers. Although this kind of corrupt scheme focuses on mainly manipulating the short-listing step in the Selection and submission phase, it also uses tools related to the Planning and advertisement phase to achieve its goal.

4.2.2.2. Biased eligibility requirements

Similarly, eligibility requirements can be manipulated in a tender, establishing excessively specific or restrictive criteria with regards to the technical or financial capacity (e.g. a certain amount of capital or level of revenue) that potential bidders must demonstrate to be pre-selected, also restricting the pool of potential bidders. Oftentimes, these requirements are intentionally defined to "match" the profile of the favored bidder. For example, a railway station construction project requiring 25 years of relevant experience probably won't have many contenders in an average size EU member state. Even if there is more than one competitor, defining what level of 'relevant experience' would be sufficient to involve a certain level of discretion, allowing such requirements to be used to exclude other suitable suppliers (e.g., a company that has 20 years of relevant experience could be a suitable bidder, but is not eligible based on such strict requirements).

4.2.3. Evaluation and award phase

4.2.3.1. Shell companies

Shell companies are firms that exist only on paper; they often have no activities and staff except for a formal manager and owner. These companies are often registered in tax havens, or in countries with an opaque registry system. They are not only used to avoid taxation, but also to disguise a conflict of interest situation in procurement or just to avert bad press. Thus, the company can evade a proper eligibility assessment in the Evaluation phase.

For example, the son of the president of the central bank winning a contract for the furnishing of bank property could raise accusations regarding a conflict of interest and can certainly induce bad press. However, if this company is registered in an opaque system, both the identity of the owner and the firm's qualifications can be hidden. If the procurement is then subcontracted to a firm with the required qualifications, the shell company's owners can pocket part of the contract price without any outsider noticing the particularistic tie between the contractor and the supplier.

4.2.3.2. Strategically annulling the tender

As mentioned above, contracting authorities are obliged to follow certain procedures based on the size and nature of the prospective procurement contract. However, if the first procedure is annulled, for example due to unforeseen circumstances, contracting authorities often have the right to re-launch the process, but using an accelerated and less open procedure (e.g. restricted or invitation procedure). This is the case in a range of countries. Annulment can be used strategically for corrupt purposes in at least two ways: first, in order to circumvent procedures requiring higher degrees of transparency and more open competition even though the awarded contract is high-value and no special exception would apply; second, in case other corrupt techniques of limiting competition fail and an unwanted firm must be awarded the contract, there is still an option of annulling the whole procedure and starting it all over again with a more effective arsenal of corruption techniques. For this corruption technique, annulment is the decision of the buyer. Of course, annulment can simply be due to incompetent planning by buyers.

As the potential reasons for contracting authorities to induce annulment is limited and annulments are fairly visible, external control may represent a strong obstacle to the application of this technique. In addition, it is also costly for the issuers as they have to re-run the same procedure multiple times. Hence, it is likely that this is used either as a 'solution of last resort' or as a blunt and expensive technique.

4.2.3.3. Unfair scoring

Scoring of competing bids takes place whenever bids are evaluated on the basis of other criteria in addition to the offered price. Contracting authorities are obliged to assemble an evaluation committee and keep records of their scoring. As the process of scoring is internal to the public organization and attaching scores to subjective criteria is difficult to effectively control from outside, scoring can be easily abused for the benefit of a corrupt network. If scores are given to benefit a 'desired' company it clearly infringes upon the principle of fair competition.

The process of scoring is difficult to control for external bodies in general, but especially in the case of subjective evaluation criteria. Bidders who did not win are likely to question the fairness of scoring; however, due to privacy law and trade secret considerations their chances of a successful challenge are rather low. This technique works in tandem with tailoring evaluation criteria, as subjective criteria defined already at the outset makes the application of this technique easier.

This technique can be exercised in a subtle way: the call for tenders defines evaluation criteria regarding non-quantitative performance and the evaluation committee scores these non-quantitative aspects in a barely visible unfair manner by giving the ‘desired’ bidder only one point in addition to the others. Nevertheless, these scores add up in the end just a little bit higher than the best ‘unwanted’ bidder. Due to the lack of publicly available detailed records of the scoring and evaluation processes inside the contracting agencies, the occurrence of this technique can be difficult to measure directly. The best approximation to detect this technique is to use indirect measures, such as the length of the decision period.

4.2.3.4. Unfairly excluding submitted bids

This corruption scheme focuses on manipulating the evaluation and award phase as a necessary fallback when earlier measures to restrict competition – such as biased specifications or selective information provision – fail to achieve a single bid. The theoretical operation of this scheme is straightforward: even if multiple suppliers submit bids, the corrupt contracting authority ensures the eventual winner is predetermined by administratively eliminating all competing offers. This manipulation is achieved through subjective interpretation of formal requirements, intentional misjudgment of bid quality, or manufactured technical non-compliance. The ultimate goal of this technique is to successfully reduce the number of valid bids to exactly one, thereby guaranteeing the contract award to the favored supplier. The "excluding all but one bid" indicator, which tracks the difference between the total number of bids received and the number of bids deemed valid, serves as a quantitative measure to detect this specific type of post-submission competition suppression.

4.2.4. Contract management phase

4.2.4.1. “Bogus” subcontracting

Information on subcontractors is often less transparent even in relatively developed procurement systems. Therefore, an otherwise qualified supplier is used to compete for the procurement which outsources the project to the subcontractor. The subcontractor could be registered in another country, so it could siphon off the funds and disappear without completing the project. In this setup, the particularistic tie lies between the contractor, the subcontractor and the final supplier, but the additional step – of subcontracting – creates a scheme which is hard to uncover. This corrupt scheme occurs during contract implementation in the Contract management phase.

4.2.4.2. Substandard work

Substandard work means providing goods, works, or services that do not comply with the specifications stipulated in the contract. This process may include corrupt officials or could be the result of a company taking advantage of poor contract management practices. In other cases, supervisors can be bribed or coerced to sign off on substandard work.

Substandard work typically becomes fraudulent, when the contractor recklessly or knowingly claims to have performed the work required to obtain payment. This type of scheme occurs during contract implementation in the Contract management phase and could be prevented by comprehensive quality checks; however it is often hardly feasible. Road construction projects for example, are the hotbed for substandard procurement projects, because the quality of roads is mostly visible only after several years of usage.

The examples above highlight the importance of in-depth qualitative assessments, such as the on-the-ground contract monitoring that will be conducted as part of iMonitor. Although the following sections mainly promote the importance of data analytics in public procurement, data limitations will often prevent the identification of more complex corrupt schemes. It is important to remember that results are only as good as the data being used, hence a comprehensive analysis should include both qualitative and quantitative assessments.

5. Data

Once the procurement process and its potential corruption risks have been outlined, the next step is to decide what data to analyse and how to obtain it in order to identify these risks. This section highlights the key dimensions of procurement datasets (scope, depth, accuracy, accessibility) and explains some of the usual data errors analysts should watch out for. As part of this section, these dimensions are illustrated for the data collected on the seven countries in the current scope of the iMonitor2.0 project.

5.1. Goals of creating a public procurement database

The main objective of data collection is to create a clean and comprehensive public procurement dataset, which makes the evaluation of procurement systems integrity feasible. Building procurement datasets is usually a very time consuming and often expensive process, since not many procurement authorities give up their data easily (even though these datasets should be publicly available). It is also important to note that the actual contracts are rarely available publicly; data is usually collected from the tender or contract award notices. In theory, this should correspond to the contract specifications. In reality, however, it is often impossible to verify whether this is the case, which could lead to corruption going undetected.

The data collection process behind opentender is described in detail in a methods note (Government Transparency Institute, 2025). First, if the available data is in a machine readable (e.g., HTML or searchable PDF) format, it can be scraped, and the data can be stored in a relational database. The second step is to parse, clean, and impute the raw data. Third is the matching of publication related to the same tender, and matching the same organisations (buyers, suppliers) together. Fourth, is creating a single tender record by normalising information appearing in different publications - such as saving a single buyer name or contract value. The last and arguably the most important step is the manual and automatic data validation. This consists of reviewing a random sample of the data to check data accuracy and correct potential errors (see Appendix 2) that could have occurred during the collection process.

For opentender.eu, data is collected from two main sources: national procurement platforms and the European-wide Tenders Electronic Daily (TED)[6]. TED was established by the European Commission as a response to the need for a unified platform to disseminate information about public procurement opportunities within the European Union and the European Economic Area. The initial goal was to create a centralized system that could streamline the dissemination of tender notices, making it easier for businesses to access information across borders and participate in public procurement processes internationally. EU member states are required to publish data on TED when a procurement falls above a set of specific value thresholds.

Data available from national sources varies in scope, depth, accuracy and accessibility due to the different national regulations and legislations. However, above EU threshold tenders have to follow the same publication practices in all member states. For the database behind opentender.eu, we combine TED data with national datasets. To avoid including duplicated tenders, TED and national sources are deduplicated using a simplified EU threshold rule (see Government Transparency Institute, 2025).

Generally, data collection from TED contributes to more comparable and complete datasets, however, it is important to highlight some exceptions. There are a few ways where tenders can be duplicated in the data collection process. First, there are tenders that are above both the national and the EU thresholds and are published on both the national website and the TED database. In addition, there is the case of voluntary reporting, when tenders below the EU threshold are also uploaded on TED. Duplicates created in either way need to be detected and deleted.

Furthermore, data quality can still vary greatly. For one, missing rates of different variables (e.g. missing rates of buyer and bidder IDs) are not evenly distributed among the member states. Another important aspect behind the variation of data quality is voluntary reporting. Some countries commonly provide voluntary publication of tenders under the mandatory thresholds, meanwhile in other countries voluntary reporting is minimal. There are also some special cases where the national thresholds are nearly the same as the EU thresholds, and also there are cases where only the TED database is accessible. In these cases, the TED database is the primary data source.

All in all, a significant level of comparability has been achieved using the combination of these different datasets and corresponding pipelines preparing the raw data for analyses. All data referred to in the report is publicly available on the opentender.eu portal.

5.2. Key aspects of procurement data quality

After the primary data source is established, it should be evaluated based on its scope, depth, accuracy and accessibility. The evaluation process is used to assess the overall adequacy of the dataset for corruption risk analysis. Using this preliminary assessment, the expectations about the data source can be adjusted, the research limitations can be outlined and, if necessary, the scope and depth of the analysis can be modified accordingly.

[6] https://ted.europa.eu

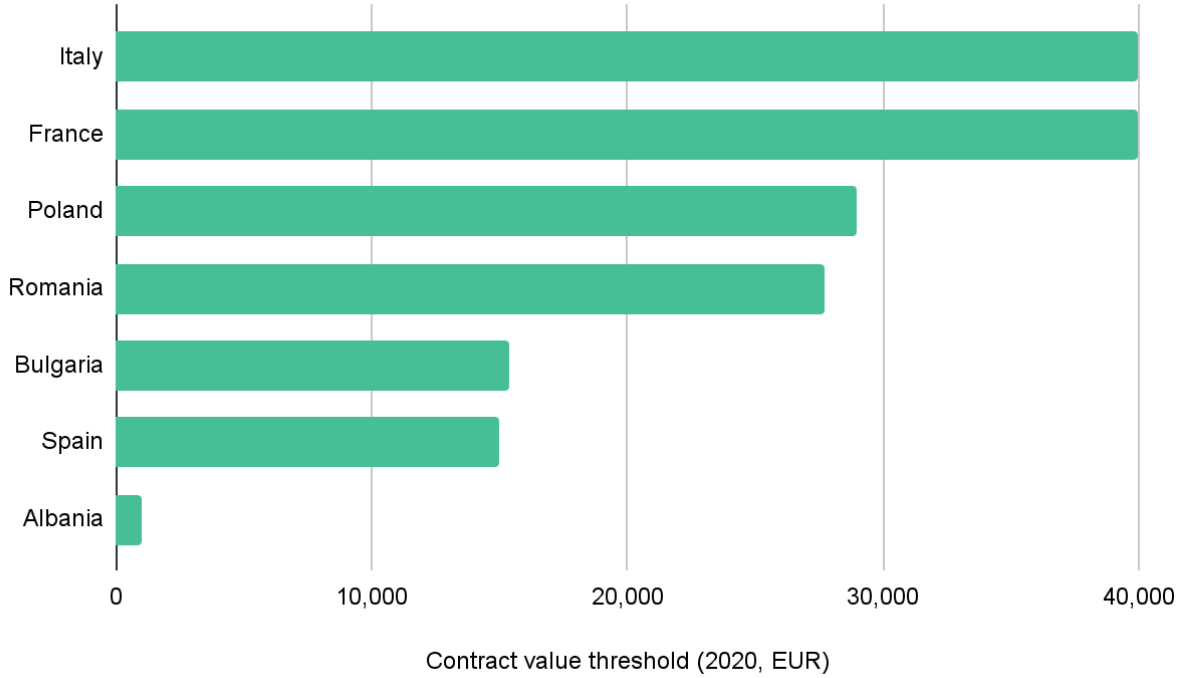
5.2.1. Data scope

In public procurement, we interpret data scope as the value threshold and timeline on which procurement data is available on the tendering website. Reporting thresholds are national contract value thresholds for mandatory publication of tenders on national or EU wide portals. Thresholds vary greatly across Europe and can have different scopes and regulations attached to them (for example, EU regulation distinguishes between different buyer types, such as central, subnational, utilities, or defence buyers). Figure 3 below shows that some countries have relatively high thresholds, while others require all contracts to be published. Procurements over the threshold usually have to comply with stricter rules, such as minimum length of advertisement period or publication of the scoring criteria, hence lower thresholds lead to more transparency. Furthermore, in countries with lower thresholds, larger portions of the procurement market can be analyzed leading to a more accurate analysis. It is important to take the scope of the analyzed data into consideration when forming policy recommendations.

Table 3 & Figure 3: Scope of public procurement databases for iMonitor2.0 countries - Minimum contract value for publishing supplies and services contracts (EUR, 2020) - based on EuroPAM 2020[7]

Country	**Contract value threshold (2020)**
France	EUR 40,000
Italy	EUR 40,000
Poland	PLN 130,000 [ca. EUR 29,000]
Romania	RON 135,060 RON [ca. EUR 27,700]
Bulgaria	BGN 30,000 [ca. EUR 15,400]
Spain	EUR 15,000
Albania	ALL 100,000 [ca. EUR 1050]

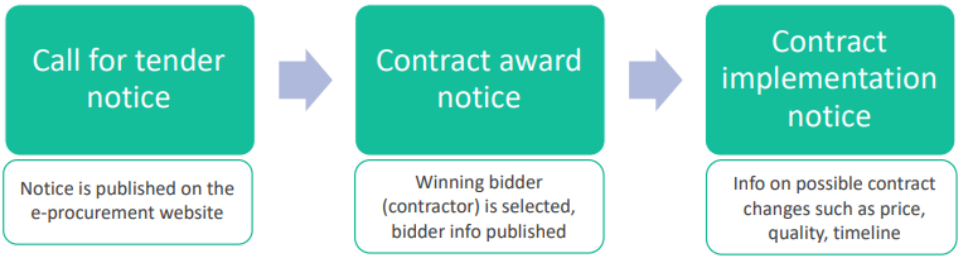
[7] Website: https://www.europam.org/index.php?module=overview
Dataset at: https://www.europam.org/data/mechanisms/PP/PP%202020_QualQuant.xls
The Albanian minimum threshold for conducting small value procurement procedure is 100,000 ALL, that is around 1050 EUR.



5.2.2. Data depth

Data depth includes the coverage of the tender cycle and the availability of indicator-relevant data. The tender cycle consists of the call for tender publication, the contract award publication and the implementation phases as shown on Figure 4 below (a more general version of Figure 1). Contract award related information is available for all contracts above the reporting threshold for all types of procurements, while call for tender notices are only available for projects with non-restricted procedure types, hence the latter dataset is usually a subset of the former.

Figure 4: Tender cycle



Source: GTI

There are relevant differences in the tender cycle coverage across countries. Most of the procurement systems in Europe only cover the advertising and the awarding phases. For example, TED typically covers prior information notices, contract notices, contract awards, modifications (even the modification of awards), and cancellations. Only a handful of EU countries' procurement systems disclose information on implementation - i.e. that would report data on the actual payments made (to allow assessing cost-overruns), work quality or implementation delays. Absent information on the implementation phase can give a false

picture about the quality of procurement, for example, if the contract is modified or the work is poorly implemented.

Furthermore, the depth of information within a cycle can also vary greatly due to different – and frequently changing – legislation. Some countries do not collect information on the number of bidders, hence the level of competition cannot be analysed. Other countries only publish the name and location of organizations without any unique identifiers, which makes over-time tracking impossible. The lack of IDs also forces data users to rely on probabilistic matching techniques that are prone to errors, such as string-matching or name-location matching.

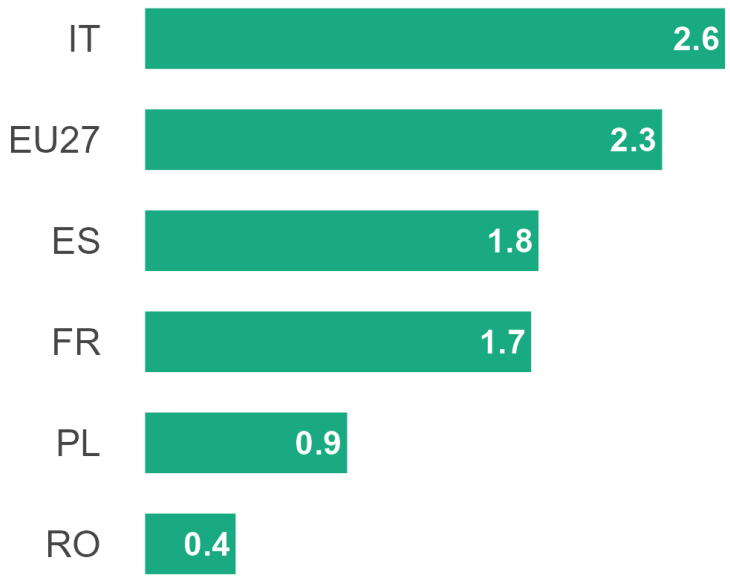
For the above reasons, this report offers a list of minimum required variables for comprehensive corruption risk assessment, shown in Appendix Table 1. These variables are necessary to calculate the integrity indicators that will be introduced in the next section, however, might not be sufficient for more in-depth country or sector specific studies. It must be mentioned that each research requires a specifically tuned dataset, hence this example should only act as a general guideline.

5.2.3. Data accuracy

Data accuracy should be examined both before and after the data collection process. Before data collection, it is useful to manually verify the accuracy of the most important variables. High missing rates or inadequate data in essential variables could require the modification of the initial research question or the use of another data source. Nonetheless, full-scale data coverage can only be tested after the data gathering phase is concluded. As a rule-of-thumb, a less than 10% missing rate should be considered as acceptable, however data analysts sometimes have to work with greater missing shares due to the lack of alternative public procurement data sources.

Figure 6: Percentage of missing fields in key tender information in TED (2018-2024)[8]

[8] The missing percentages are based on the following variables: buyer name, buyer city, buyer postcode, buyer NUTS region, bidder name, bidder city, bidder NUTS region, bidder country, contract award publication date, awarded contract value, procedure type, CPV code, EU funding status, subcontracting status. Framework agreement tenders and those with multiple lots are excluded. Single lot tenders account for around 72% of all tenders, and cover around 43% of awarded contract value between 2018 and 2024 based on above EU threshold tenders. Albania is excluded from the figure, as their tenders are not mandatorily published at TED.



Finally, database accuracy must be also checked and compared after the data collection process ends. In this step, the analyst should examine the raw data and compare it to the source to make sure that the collection process was flawless. This is especially important if the data is scraped or obtained in any other way that is not controlled by the official maintainer of the website. Appendix Table 2 lists a few typical data errors that can occur during data collection.

5.2.4. Data accessibility

Even if data accuracy meets the required standards, the ease of obtaining the data significantly varies across procurement systems. Accessibility can be categorised into four groups:

1. Structured format: Procurement data is stored in a relational database and can be downloaded (using a bulk download option or an API) into an csv/json file.
2. Semi-structured format (semi machine-readable): Information is available in a HTML format which can be scraped and parsed.
3. Not fully machine readable: Part of the data only accessible by manual cleaning (e.g., some of the documents are uploaded as scanned pdf files).
4. No public data publication.

The below table lists the national data sources that are behind the data published on opentender.eu.

Table 7: Data accessibility and machine readability in iMonitor2.0 countries

Country	**National data sources**	**Data format**

IT	The national source for public procurement data is the Open data portal of the National Anti-Corruption Authority (ANAC)[9]. Data is available from 2007 and is monthly updated on both the national portal and on opentender.eu. The structure of the data is standardized because the contracting authorities send the data to ANAC and they publish and standardize it. Furthermore, official unique tax IDs are provided, which are very helpful in identifying organizations. Italy is one of the few EU member countries where it is possible to match open data on public contracts with data on related investment projects. ANAC publishes a table connecting project IDs with contract IDs. All investment projects are included, regardless of the source of funding (EU, national, local, etc.).	Data is accessible in a structured format, downloadable both in CSV and JSON formats. Opentender.eu uses the CSV publications.
RO	Romanian public procurement data comes from e-licitatie.ro, which is the Electronic System for Public Procurement (SEAP). Historical data published before 2018 comes from http://istoric.e-licitatie.ro/. We regularly download all publication types in a JSON format – such as Simplified contract notice (SCN), Call for tender (RFQ), Contract notice (CN). Note, that https://data.gov.ro/en/ also publishes procurement data in structured xlsx and csv formats.	Data is accessible in a structured, machine-readable JSON format through an API.
ES	Spain has two national sources of data. One is available from 2011 (http://contrataciondelestado.es), and the other is available from 2013 (https://www.hacienda.gob.es). The two sources may contain overlapping tenders, hence tenders are deduplicated.The scope of the Spanish national data includes all tenders available on the above-mentioned websites, including data on minor contracts.	Data is accessible in a structured format. XML files are provided on both national datasets, with compressed (zip) files containing multiple tenders updated monthly on the newer website.
FR	French national data is collected from one source, the *Bulletin officiel des annonces de marchés publics* (BOAMP) live server that is continuously updated by the *Direction de l'information légale et administrative* (DILA) French authority. The source is composed of archived and currently updated materials that account for tenders starting from 2004 up to current day procedures.	Data is accessible as structured XML and HTML files.

[9] https://dati.anticorruzione.it/opendata

	Tenders that meet the TED publication criteria are also published on the BOAMP server; however, because they duplicate material already available on TED, we do not collect them.	
PL	Polish data is collected from three distinct sources, two of which are historical and have been discontinued, and one which is currently updated. The two historical sources are an FTP server (ftp://ftp.uzp.gov.pl) which contains downloadable tenders in compressed XML format ranging from 2007 to 2017, and a JSON API which ranges from 2017 to 2021 (http://websrv.bzp.uzp.gov.pl). The current source *Biuletyn Zamówień Publicznych* (BZP) comprises data from 2017 and is also in JSON format available on an API (https://ezamowienia.gov.pl).	Data is accessible in XML and JSON formats and accessible through an API.
BG	Bulgaria national data comes from two sources: a discontinued website (http://www.aop.bg//) where tenders awarded between 2007 and 2021 were published, and the currently active procurement website (https://app.eop.bg) maintained by the Bulgarian Public Procurement Agency (Агенция по обществени поръчки, AOP) from which tenders from 2021 onwards are collected.	Data are available in semi-structured HTML format.
AL	Albanian data is collected from a single source (https://www.app.gov.al) where tender data is published in downloadable CSV format by the *Albanian Public Procurement Agency* (Agjencia e Prokurimit Publik).[10] Data is available from 2010 onwards.	Data is available in structured, downloadable CSVs.

6. Integrity indicators

This section introduces the definition of integrity as applied in the report and the relationship between integrity and corruption. It then discusses the conceptual background of the individual integrity indicators, presenting different groups of risk indicators with their key features. Uses and limitations of these indicators are presented, followed by a brief description of the importance of validation and the validation process. Finally, the section describes eleven indicators used on Opentender and calculated to measure different

[10] Downloadable CSVs are available at the following link: https://www.app.gov.al/eksportimi-i-procedurave-te-publikuara/.

aspects of integrity specifically in the iMonitor2.0 countries, and a composite indicator (Average Integrity Indicator Score) designed to approximate the overall integrity level of a given tender.

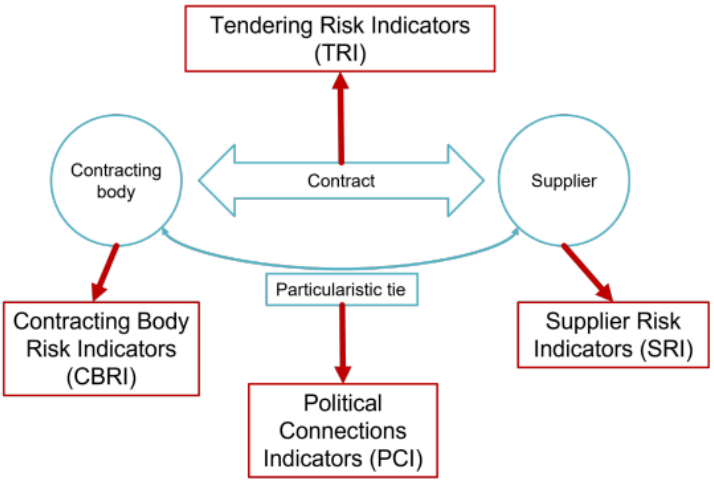
6.1. Relationship between integrity and corruption

In this report, we frequently address the underlying mechanisms as corruption. We refer to corruption definitions, risks and measures, which all aim to establish a better understanding of the conceptual background behind opentender.eu. The platform, however, uses a different approach: it aims at promoting integrity, instead of highlighting corruption. For this purpose, the term integrity is introduced, which is defined here as the opposite of corruption. In other words, when corruption is high, integrity is low; and vice versa.

6.2. Conceptualising risk indicators

Given that well-designed indicators can significantly reduce the costs of quantitative research, the next step is to outline the key features of valid and reliable corruption proxy indicators. The first step in this direction is to outline the framework in which corruption risk indicators can be developed. As mentioned in the previous sections, a corrupt public procurement process can be thought of involving the following elements: 1) a public body buying a service (or supply, or work) using public funds, 2) a supplier, who will provide this service in return for these funds, 3) a contract that describes the details and conditions of the agreement, 4) and a particularistic (informal) tie which makes the coordination of corrupt practices possible. The participants in the process and the relationships between them is shown in Figure 8.

Figure 8: Corruption indicator groups in the procurement process



Source: GTI

For each of these items a group of potential indicators can be developed:

- Tendering Risk Indicators (TRI) are a group of contract-specific, observable variables that can be steered to favor a certain supplier. These can include the procedure type, bidding period length, award criteria etc.

- Supplier Risk indicators (SRI) are supplier-level information which could indicate that it might be involved in corrupt practices. These variables can include the firm's registry date (e.g., if it was registered just before the high value contract was published), whether it is registered in a tax haven, or whether it is extremely profitable compared to other market participants.

- Contracting Body Risk Indicators (CBRI) are a group of variables indicating that the procurer might attempt to corrupt its purchases. These indicators are hard to measure because public organizations often don't have well-structured data. However, some publicly available information can be useful, such as the change in leadership (or employees) after a political regime change, or the proportion of adequately trained staff.

- Political/Personal Connections Indicators (PCI) describe the informal tie between the buyer and the supplier. This is also a hardly measurable area but indicators such as kinship, or previous political or business connections between the leaderships could indicate the presence of a particularistic tie. PCIs are good examples for risk indicators that are likely to underestimate corruption risk, since objectifying political or personal connections is often a difficult task.

Indicators in each of these groups must have a list of common qualities that make them adequate to measure procurement corruption risk. The following are the key features that analysts should consider when creating new indicators; they have to be:

- **objective**: they are based on factual data non-mediated by stakeholder's perceptions, judgements, or self-reported experiences,

- **de facto**: they describe actual behavior or events in contrast to legal prescriptions or expectations,

- **micro-level**: they are defined on the level of actors of corrupt exchanges (e.g., companies) or the transactions among them (i.e., contracts). They can nevertheless be aggregated at higher levels,

- **internationally comparable**: while defined on the micro-level, indicators should be comparable across countries or regions, due the same underlying theoretical concepts and measurement approach,

- **comprehensive**: they adequately capture corruption risks in a wide set of organizations performing comparable tasks,

- **time-series**: indicators are ideally measured and can be compared over time for at least 5-10 years.

Indicators that are not equipped with these qualities could bias the analysis and, therefore, are not to be used.

6.3. The uses and limitations of integrity risk indicators

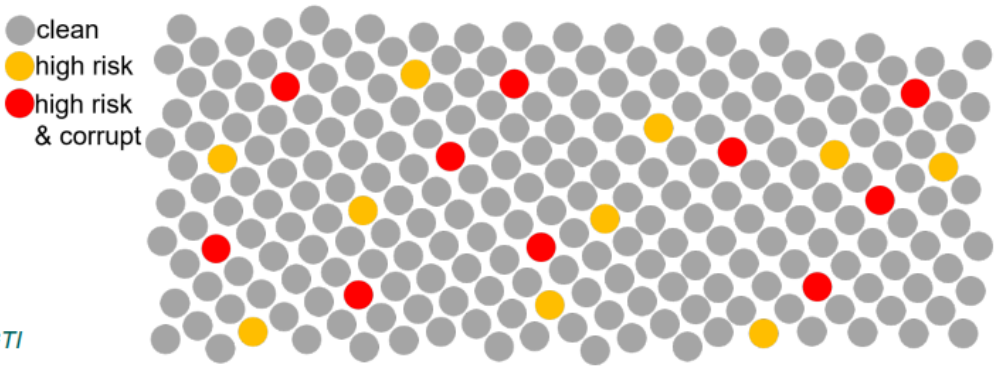
Let's consider the task of distinguishing clean vs. corrupt tenders, for example to investigate the extent and forms of corruption in a public procurement system. An initial idea might be to take a small sample of tenders for in-depth analysis, which would show that 1 in 20 tenders

could be considered corrupt. The problem with this approach is that 95% of the work put into the analysis is unnecessary, because the share of truly corrupt tenders is only 5%. Furthermore, while a randomly selected sample could potentially predict the share of corrupt tenders in the whole population (although a small sample with a sufficiently low share of corruption even hinders these estimates), it is inadequate to describe the exact characteristics of corruption in the system. For this, a researcher would need to analyze each tender in the procurement market of possibly thousands of tenders.

An alternative approach is to use risk indicators to find *potentially* corrupt tenders. A statistically developed and empirically tested indicator can predict each tender's corruption risk, hence theoretically - after adequately tuning the indicator - there is no need to manually check each tender in the whole population. The largest issue with this concept is that it is impossible to develop a 100% accurate indicator. It could both 'flag' tenders that are not corrupt (false positives), or miss procurements that are corrupt (false negatives). A poorly designed indicator might cause more harm than good, since it creates additional work by both having to find all non-flagged corrupt cases and remove flagged ones that are not corrupt.

Nevertheless, even a mediocre indicator, that flags false positives, can significantly reduce the time required for the risk assessment. Imagine an indicator that can label all the corrupt cases, but also flags an equal number of non-corrupt cases. Figure 9 below shows that in this scenario an analyst must manually check all 20 cases, 10 of which will be actually corrupt. Contrary to the first example, now all the corrupt tenders are found and 50% of the work was useful. However, this method will not work if an indicator understates the risks. False negatives are more dangerous, because they can only be found if the whole population is manually checked. Therefore, it is a better idea to start with 'strict' indicators that might initially flag non-corrupt tenders and refine them by testing on separate samples.

Figure 9: Example of corrupt and flagged cases in the total population of tenders



Source: GTI

Overall, the main goal of indicator building is to increase the overlap between the corrupt cases (red dots) and the high-risk cases flagged by the indicators (yellow dots). When testing these indicators, three issues should be considered:

- *False positives*: the indicator flags tenders that are not corrupt.
- *False negatives*: the indicator does not flag tenders that are corrupt.
- *Both*: the indicator is both missing some corrupt cases and flagging non-corrupt ones.

A perfect indicator would find each corrupt tender without flagging any additional ones, however, even a slightly imperfect indicator can speed up the research process. Furthermore, as it will be discussed in the next section, combining separate, independent indicators can increase the reliability of the final measurement tool.

6.4. Validating risk indicators

Validation is key when aiming to reliably measure the corruption risk and therefore the integrity of a selected tender. As discussed in previous sections, there are multiple ways corruption can happen during and beyond the public procurement phase. Yet there are limited ways to quantitatively measure corruption in these phases.

There are two prerequisites to ensure that integrity is measured in a reliable way while catching a wide range of suspicious activities. The first one is that each indicator has to be connected to at least one underlying corrupt scheme, to ensure that it makes sense to include the individual indicator in a composite index. Second, to select appropriate indicators for the final composite index, each of them have to be separately validated.

Single bidding, described further below, is broadly accepted as an appropriate indicator to measure potential corruption, as high correlation is frequently recorded between single bidding and other indices measuring corruption. Taking advantage of this attribute, the other individual indicators can be validated by comparing them to single bidding. In addition, during the validation process any new indicator is being compared to the previously validated indicators, to ensure that they have enough meaningful explanatory power on their own, thus ensuring individual validity. Furthermore, indicators used for comparative purposes across different countries are validated to verify their applicability to different contexts. For this, country-specific thresholds for indicator scores are calculated – these can be provided upon request.

6.5. Existing individual indicators on opentender

There are many potentially adequate indicators that measure slightly different aspects of corruption risk in public procurement. After rigorous validation, they can be combined into a composite score which increases robustness by reducing the amount of both false negative and false positive cases - thus increasing the overlap between the truly corrupt cases and the high-risk cases flagged by individual indicators. From a more technical point of view, combining different indicators can be thought of as removing confounding factors, hence creating a more robust relationship between tender level corruption and corruption risk measurement. This section describes the eleven individual indicators used by the opentender.eu portal. Table 2 at the end summarizes the indicators and the scoring criteria for each.

6.5.1. Single Bidding

One of the most widely used corruption risk red flags is the single bidding indicator. It indicates that a given tender only had one bidder during the procurement process, hence there was no competition for the tender. As explained earlier, the lack of competition is one of the main signs of corruption in the public procurement system. Even more conveniently, single bidding can be easily extracted from most of the available data sources, and it also holds the key qualities necessary for an adequate indicator.

Nevertheless, even a theoretically sound indicator must be tested before applied in quantitative research. The best scenario would be to test the indicator on 'labeled' data where each corrupt tender is already flagged. Unfortunately, only a handful of procurement systems disclose this information (e.g. court rulings) and even these datasets lack overall generalizability, since not all the corrupt tenders can be located by procurement authorities (many of them could stay hidden due to political and technical factors).

While none of these results necessarily demonstrate a causal relationship, both the theoretical (the lack of competition could signal corruption) and the empirical evidence suggests that single bidding might be a good starting point to identify corrupt tenders. Looking back at the corrupt schemes identified in section *4.2. Identifying typical corrupt schemes*, single bidding can be achieved (or at least approximated) by almost any identified scheme, as it would indicate the wanted result of employing such schemes, namely reducing competition. A sufficiently short advertisement period, for instance, hinders competition and may result in a single eligible bidder, while tailoring the specifications to an extreme extent may also result in only one bidder fulfilling the necessary requirements.

However, similarly to other indicators, single bidding cannot be used as an all-round tool for corruption measurement. Like many other indicators it has its weaknesses, such as overestimating corruption risk. Take for example a very specialized market where there are not enough companies. In this market, public procurement will be limited, and single bidding will be a measure of market imperfections (e.g. too high barriers to entry) rather than a measure of corruption. Another example could be the sudden increase in government spending, creating tight procurement markets where contracting authorities have to compete for suppliers. For instance, the current availability of EU Recovery and Resilience Facility funding has led to a sudden, significant increase in public investment in several EU member countries.

6.5.2. Call-for-tenders publication

Not publishing the call for tenders makes it less likely that eligible bidders gain knowledge of the bidding opportunity, weakening the competition and allowing the contracting bodies to more easily award tenders repeatedly to a well-connected company. As discussed in section *4.2.1.3. Avoiding publication of the call for tenders*, there is a very well established link between the mentioned corrupt scheme and the Call-for-tenders publication indicator. Furthermore, as the availability of data on whether a call for tender is publicly accessible, Call-for-tenders publication is a well suited choice to be included in the individual indicators' list.

This indicator can be in close relation with procedure type, as there are some specific procedure types that do not require a call for tender. The most interesting results lie in cases where there should be a call for tender, yet it is absent. Validation of the variable often includes controlling for procedure types, further strengthening the indicator.

6.5.3. Procedure type

Using procedure types which are less transparent and require less open competition can indicate the deliberate limitation of the range of bids received and to exclude bids as well as creating more opportunities for contracting bodies to repeatedly award tenders to the same well-connected company. Carefully exploiting characteristics of certain procedure types can

be a sophisticated way to conceal an attempt to deliberately hinder competition to favor a connected supplier.

For example, differentiating between the distribution of information between the favored bidder and the rest of the bidders through certain procedure types such as negotiation is a popular corrupt scheme, further discussed in *4.2.1.4. Selective information provision*. Public bodies may also refrain from using non-open procedures as their first intention, however, through the strategic annulment of a tender issuers often have the right to re-launch the process using an accelerated and less open procedure, as shown in section *4.2.3.2. Strategically annulling the tender*.

6.5.4. Length of advertisement period

A short advertisement (or submission) period, that is the number of days between publishing a tender and the submission deadline, leaves less time and thus makes it harder for non-connected companies to bid successfully, whereas a well-connected firm can use its inside knowledge to win repeatedly as the buyer can informally inform the favored bidder about the opportunity ahead of time. This phenomenon is described in further detail in *4.2.1.1. Shortened advertisement period.* This indicator is closely related to the mentioned corrupt scheme, providing a great basis for the inclusion of this indicator in the composite integrity index.

6.5.5. Length of decision period

An excessively short or long decision period, that is the number of days between the submission deadline and the contract award decision, can signal integrity risks. Snap decisions may reflect premeditated assessment, while long decision periods may signal extensive legal challenges to the tender, suggesting that the issuer attempted to limit competition. Extremely short decision periods may also be an indirect measure of unfair scoring (as discussed in section *4.2.3.3. Unfair scoring*), which adds more complexity to this indicator.

6.5.6. Tax haven

Awarding public tenders to companies registered in tax havens presents a risk that anonymous company ownership could be concealing a conflict of interest in the award of a tender to a politically connected beneficial owner. The founding of shell companies (see further in section *4.2.3.1. Shell companies*) and using them to bid for tenders where conflict of interest would otherwise arise is a popular corrupt scheme, seemingly hidden from plain sight. Using Tax haven as an indicator is a useful tool to adjust for this scheme.

The potential loss of tax revenue from the successful supplier through permissible tax avoidance or illicit tax avoidance is another risk. Thus raising the validity of the inclusion of Tax Haven even where conflict of interest is not a probable factor. While the definition of ‘tax haven’ (or secrecy jurisdiction) is still a matter of debate, this indicator uses an independent ranking by the Tax Justice Network[11] of countries’ legal frameworks with regards to banking and corporate secrecy.

[11] https://fsi.taxjustice.net/

6.5.7. Benford’s law

Benford's law is an observation about the leading digits of a naturally occurring collection of numbers. It states that the first digit is likely to be small: in sets that obey the law, the number 1 appears as the leading digit about 30% of the time, while 9 appears as the leading digit less than 5% of the time. If this indicator has a high value, it indicates that the price of the tender obeys Benford's law, thus it is similar to naturally occurring collection of numbers, and it is less likely that the price is manipulated. Meaning that tender values are more likely defined based on market cost structures rather than some deliberate manipulation, for example as a high-level figure reflecting corruptly agreed deals.

Therefore, Benford’s law is not directly related to any specific corrupt schemes, but could be an indirect proxy of any scheme where deliberate price manipulation is present – for instance, if a favored bidder receives a contract in a value that corresponds to the available budget of the purchase and not reflecting the price calculation procedure of the supplier. This makes Benford’s law a powerful complementing and corroborating element to the aforementioned indicators as it can provide a unique approach to capturing a wide range of corruption risks that might be reflected in manipulated contract values. Also, this indicator has the advantage of providing low overall false negatives when aggregated with the other indicators: a clean tender which by chance produces low scores in Benford’s law would still have an otherwise high overall integrity if other indicators do not identify additional risks.

6.5.8. Buyer spending concentration

A supplier’s share in a buyer's total spending in a given year can be used as a measure of market competitiveness and openness. A high share of supplier spending can signal that a supplier or a group of suppliers are part of a network, potentially leading to higher prices, and/or lower quality and value for money. When a buyer relies significantly (compared to other suppliers) or completely on a specific supplier, the probability of a deliberate and politically motivated decision making increases.

A good example of a corrupt scheme heavily dependent on this is *4.2.1.2. Splitting procurement procedures to circumvent competition procedures.* In this scheme the avoidance of more transparent procedures are achieved by splitting up the total tender value of an otherwise single tender into as little chunks that it is no longer required to be published or even to have competition involved (for more details see section 4.2.1.2.). While this scheme provides a great example of a deliberate attempt to award a tender to the desired contractor, high buyer spending concentration could also be a result of applying other corruption schemes (e.g. *4.2.1.4. Selective information provision* or *4.2.3.3. Unfair scoring*) over a longer period of time.

6.5.9. Distinct markets

The Distinct markets measure is designed to identify potential corrupt practices by assessing the diversity of markets a supplier operates in. This indicator is based on the idea that an unusually broad presence in different markets (unless in exceptional cases of few large conglomerates) may reflect an unusual pattern of supplier behavior and could be a sign of corruption, as it may be implausible for a single supplier to legitimately excel in many distinct sectors. The calculation involves determining the number of unique markets each supplier is involved in, dividing this by the total number of contracts per supplier, and checking whether

the deviation from the average is more closely related to the more fundamental corruption risk indicators such as single bidding or buyer spending concentration. The indicator identifies suppliers that deviate significantly from the average as having a high risk of corruption.

This variable is tailored to detect firms operating in an unexpectedly vast array of markets, flagging them as potential corruption risks. For instance, during the COVID-19 pandemic, a great number of firms ventured into the medical sector without relevant prior experience, and participated in high-value contracts. This scenario highlights a potential source of corruption that the Distinct markets indicator can effectively capture. To broaden its scope, further investigation should explore additional corrupt mechanisms that this indicator might unveil, ensuring its applicability across diverse contexts and industries. As with any corruption indicator, it is crucial to acknowledge its limitations, such as potential false positives in unique market scenarios or during periods of increased government spending, emphasizing the importance of a comprehensive approach to corruption risk assessment.

6.6 New indicators developed under iMonitor 2.0

As part of the innovation brought about by the iMonitor2.0 project, two new indicators are being developed: the tailoring of text-based field from procurement documentation, analysed by a natural language processing (NLP) model; and the exclusion of all but one bid.

6.6.1. Tailoring of eligibility requirements, product descriptions, and award criteria

The text-based indicator measures the extent to which tender documentation may have been subtly tailored to favor a specific bidder. While formal corruption strategies (such as restrictive procedures) are visible, manipulation of tender texts offers a hard-to-detect way to limit competition. Our methodology builds on insights from our working paper (Katona-Fazekas, 2024), ensuring that the approach is both theoretically and empirically validated. Three parts of the documentation are available for analysis: product descriptions, which may over-specify items so only one supplier can reasonably deliver the products; eligibility criteria, which can be worded to exclude unwanted bidders; and award criteria, where non-price scoring rules can discreetly benefit a preferred firm.

To detect these risks, we analyze six publicly available text fields from procurement announcements, covering tender titles and descriptions, personal/technical/economic eligibility requirements, and award criteria. Although these fields do not include the full documentation accessible only to registered bidders, they still contain the key information through which subtle restrictions can be signaled.

Our approach relies on two types of textual indicators:

1) N-grams (1–2 word combinations) derived from cleaned text.

2) Meta-characteristics, calculated on raw text, such as number ratios, references to laws, text length (normalized across sectors), and lexical diversity. These features capture broader structural patterns and compensate for information removed during preprocessing.

We vectorize the textual data using TF-IDF and incorporate these predictors into supervised learning models aimed at forecasting single-bidding, a core corruption risk indicator. After testing several algorithms (Katona-Fazekas, 2024), the best-performing approach is a random forest model. Random Forests is an ensemble model built on multiple decision

trees. Ensemble techniques let us train more accurate and more stable predictive models, by combining the output from a large number of individual models (in this case each individual decision tree), each of which have been estimated on a slightly different sample with somewhat different parameters. In Random Forest models each tree is built independently, and the final prediction is a majority vote of the individual tree predictions. With a higher number of trees we can achieve better performance but running a model with a high number of trees is very resource intensive (i.e. takes a long time to run on ordinary machines). An advantage of Random Forest models is that they handle imbalanced datasets effectively.

We used the built-in random forest models from the scikit-learn package in Python. We divided our data into two parts: earlier and a smaller subset of the observations belong to the train and more recent observations to the test set, we used the default settings of the model.

The resulting text-based risk score is a predicted probability between 0 and 1. The higher the probability the higher is the chance of receiving only a single bid based on the text of the tender publications. For instance, a predicted probability of 0.8 suggests an 80% likelihood of receiving one bid due to text-based competition restrictions. This means that out of all the decision trees in the random forest model, about 80% of the decision trees predict a single bid for that given public tender.

6.6.2. Exclusion of all but one bid

The exclusion of all but one bid indicator specifically targets corrupt practices occurring during the assessment phase of a public tender. Its suitability for measuring corruption lies in its ability to capture one of the core goals of institutionalized corruption: eliminating or severely restricting competition to favor a pre-selected company. This non-competitive outcome - only one valid bid - is achieved even if the initial attempt to deter bidders failed. This can primarily be achieved by formally or administratively excluding the bids of all unwanted competitors, ensuring that the 'well-connected' bidder remains the only valid one. Having a single valid bid is therefore heavily associated with corruption, functioning as a powerful signal of a tilted process.

The indicator is formally defined in the data as a scenario where only one received bid is NOT excluded. To quantify this in practice, the difference between the total number of submitted bids and the number of valid bids is examined. The indicator signals corruption risks whenever the total number of received bids is greater than one, but the number of valid bids is precisely one.

A significant limitation for cross-country analysis using this indicator is data availability. While basic information like the total number of bids received is available in the European Union's Tenders Electronic Daily (TED), the crucial metric - the precise number of valid bids - is typically missing from this cross-national platform. Consequently, the indicator can only be calculated by accessing more detailed national public procurement data, and even in those datasets is the number of valid bids not universally available. There are four countries within the target countries of iMonitor 2 where this indicator is possible to calculate, namely, France, Poland, Albania, and Italy.

Table 2: Summary table for the integrity Indicators

Integrity indicator	Category of indicator	Level of calculation	Values
Single bidder tender	Procurement process risk	Contract	100: more than 1 bid received 0: 1 bid received
Call for tenders is published	Procurement process risk	Tender	100: call for tender/prior information notice is published 0: no call for tender/prior information notice is published
Use of non-open procedure types	Procurement process risk	Tender	100: open; procedure type not considered a red flag for the country 50: limited; procedure type considered a mild red flag for the country 0: non-open; procedure type considered a red flag for the country
Length of adver-tisement period (time between tender adver-tisement and the submission deadline)	Procurement process risk	Tender	100: number of days between publication of call for tenders and submission deadline is in an interval not considered a red flag for the country 50: number of days between publication of call for tenders and submission deadline is in an interval considered a mild red flag for the country 0: number of days between publication of call for tenders and submission deadline is in an interval considered a red flag for the country
Length of decision period (time between submission deadline and tender award decision date)	Procurement process risk	Tender	100: number of days between submission deadline and the tender award decision is in an interval not considered a red flag for the country 50: number of days between submission deadline and the tender award decision is in an interval considered a mild red flag for the country 0: number of days between submission deadline and the tender award decision is in an interval considered a red flag for the country
Supplier is registered in a tax haven	Supplier risk	Supplier	100: supplier is not registered in tax haven country 0: supplier is registered in tax haven country
Benford's law	Procurement process risk	Buyer	100: tender price is less likely manipulated 0: tender price is most likely manipulated
Supplier's tender share of buyer spending on public procurement	Supplier risk	Supplier	100: the winner's share is close to 0% 0: the winner's share is 100% Continuous number between 0 and 100
Distinct markets	Supplier risk	Supplier	100: The supplier is present in a reasonably small number of markets compared to their total contract volume 0: The supplier is present in a suspiciously large number of markets compared to their total contract volume
Tailoring of eligibility criteria (NLP)	Procurement process risk	Tender	100: More than one valid bid (not excluded). 0: All bids except one are excluded.
Exclusion of all but one bid	Procurement process risk	Contract	Continuous predicted risk score (no threshold).

6.7. Average Integrity Indicator Score

To select appropriate indicators for the final composite index, each of them have to be separately validated. If, for example, single bidding is accepted as an appropriate indicator,

each additional measure can be validated by comparing it to single bidding and by comparing it to the other indicators. Furthermore, only indicators that are widely available make the list of indicators to preserve high comparability between countries.

For each separate procurement system, every indicator should be similarly tested before deciding on whether to use it as part of the composite score. Since not every countries' procurement system provides the same quality of information, the composite score can slightly vary across jurisdictions. The composite score, or the Average Integrity Indicator score, is the arithmetic average of each available integrity indicator. It can take a value between 0 and 100 where 100 indicates the highest integrity, that is the lowest corruption risk.

It is important to keep in mind that the individual indicators are proxies of actual corruption and may incorrectly indicate the underlying corruption for individual cases. One of the main advantages of using a composite score lies in its robustness; while individual indicators have a higher chance of mismeasuring corruption, if the majority of the indicators lean in the same direction, the probability of correctly measuring both high and low risk tenders greatly increases.

The Average Integrity Indicator Score can be also validated using established corruption indices such as the Control of Corruption (CoC) indicator of the Worldwide Governance Indicators of the World Bank. The composite index has a stronger negative correlation with the CoC than the single bidding rate has in itself. This illustrates that combining indicators can give a more complete picture about the corruption risks of a procurement system. The composite index is more robust to noise and changes in the underlying data and is supposed to have a lower false positive rate.

It is important to note that each of the indicators mentioned above are validated and tested using a more complex statistical procedure, the explanation of which is not part of this report. For more technical details on the indicator validation processes and the formulation of the Average Integrity Indicator Score's opposite, Corruption Risk Index please check the following studies: Fazekas and Kocsis (2015)[12], Fazekas et al. (2016a)[13], Fazekas et al. (2016b)[14].

7. Conclusions: How to take risk-based findings further?

The overall measurement framework and the specific indicators introduced in this report are meant to offer guidance for stakeholders aiming to assess where corruption may arise. They, by no means, offer definitive proof of corruption or claims for any wrongdoing. Many of the individual indicators, but also the composite score may suffer from false positives, that is indicating corruption risks or lack of integrity even though there was no corruption taking

[12] Fazekas, Mihály - Kocsis, Gábor, "Uncovering High-Level Corruption: Cross-National Corruption Proxies Using Government Contracting Data", Working Paper series: GTI-WP/2015:02, Budapest, 2015.

[13] Fazekas, Mihály - Cingolani, Luciana - Tóth, Bence, "A comprehensive review of objective corruption proxies in public procurement: risky actors, transactions, and vehicles of rent extraction", Working Paper series: GTI-WP/2016:03, Budapest, 2016a

[14] Fazekas, Mihály - Tóth, István János - King, Lawrence Peter, "An Objective Corruption Risk Index Using Public Procurement Data", *European Journal on Criminal Policy and Research*, Vol 22., No 01, DOI: 10.1007/s10610-016-9308-z, 2016b.

place. In many cases, the observed risky behaviors may arise from legitimate behaviors or problems little to do with corruption, such as low administrative capacity or time pressures of budget periods.

Hence, with a careful understanding of the strengths and weaknesses of the framework, it is best used as a starting point for further investigations. In particular, high risk tenders, buyers and suppliers can be identified for further in-depth assessment substantiating or disqualifying the suspicions raised by data and objective formulas. The iMonitor2.0 monitoring methodology is designed to gather further evidence on high risk cases, first by looking at online available information such as company profiles, second by on-site monitoring of contract implementation. Please see further information on this approach in our training manual **Monitoring Public Contracts**.

References

Katona, E., & Fazekas, M. (2024). Hidden barriers to open competition: Using text mining to uncover corrupt restrictions to competition in public procurement (GTI-WP/2024:01). Government Transparency Institute.
https://www.govtransparency.eu/hidden-barriers-to-open-competition-using-text-mining-to-uncover-corrupt-restrictions-to-competition-in-public-procurement/

Government Transparency Institute (2025) Public procurement data processing (Version 1.0). GTI-R/2025:06. Budapest: Government Transparency Institute.
https://www.govtransparency.eu/government-transparency-institute-2025-public-procurement-data-processing-version-1-0/

Appendix

Appendix 1: Minimum required information for comprehensive corruption risk assessment

Variable group	Variable
Buyer	Buyer's name, Buyer's unique ID, Buyer's address
Bidder/bids	Bidder's name, Bidder's unique ID/tax ID, Bidder's address, Number of bids submitted, Number of bids excluded, Bid price, Exact time of bid submission, Bid type (winner/loser bid), Beneficial owners
Tender/contract	Procedure type, Framework agreement, Estimated price, Procurement type (service, supply, work), CPV codes, NUTS codes, Status (cancelled, pending etc.)
Dates	Call for tender publication date, Bid submission deadline, Contract start and end dates, Publication date of contract award, Date of contract completion
Subcontracting	Subcontractor's name and unique ID, Subcontractor's share
Consortium	Consortium members' name and unique ID, Consortium member's unique ID
Contract performance	Contract performance end date, Was performed according to contract,Explanation in case of deferring from contract, Information on contract modification, Information on performance quality

Source: GTI

Appendix 2: Common error types

Error type	Description
Lexical error	The value provided is not consistent with the column name (e.g., country id column shows currency id).
Irregularity error	E.g, the unit of measurement differs from the other observations'.
Formatting error	E.g, date is in different time format leading to errors when data is loaded
Duplication error	There are duplicate observations in the data (each value is the same)
Contradiction error	Two columns measuring (almost) the same thing show different values for the same observation
Missing attributes	No information provided (not necessarily an error).
Outlier	Given variable for a given observation is significantly different from the others (not necessary an error, but usually should be dealt with)

Source: GTI