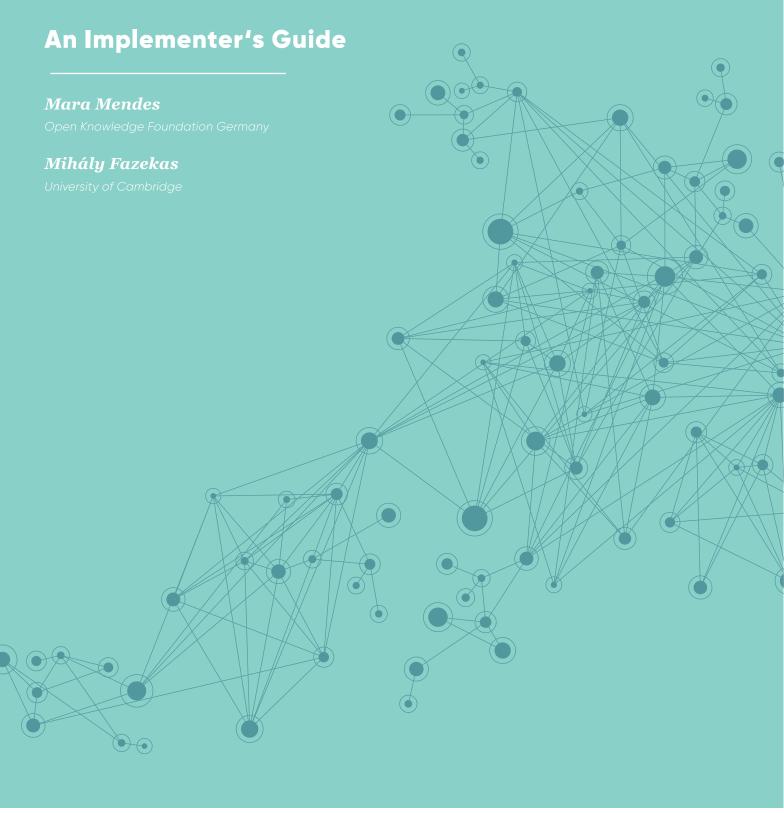
DIGIWHIST Recommendations for the Implementation of Open Public Procurement Data









I. PUBLICATION FORMAT AND LOCATION

1. Governments should set up a comprehensive central public procurement portal

Currently, 23 of 34 DIGIWHIST countries already have a single national public procurement portal publishing all regulated tenders and contracts, albeit information content, usability, and reliability varies greatly. In some cases, national portals only publish below EU threshold contracts, with the EU-wide Tenders Electronic Daily (TED) portal publishing the above EU threshold contracts1. At the EU level, TED serves as a central platform publishing tenders of EU institutions as well as national governments above the EU thresholds². DIGIWHIST analysis has found that eight countries3 have multiple national and sub-national portals which often have overlapping content, generating confusion and unnecessary complexity in public procurement markets. Having multiple portals is often rooted in a federal administrative structure which nevertheless creates barriers to market entry, hence economic inefficiency and corruption risks. Unfortunately, in three countries (Austria, Luxembourg, and Sweden) there is no public portal at all, making public procurement markets especially non-transparent4.

Providing comprehensive public procurement information free of charge in an easy-to-use format to all interested parties is expected to increase market transparency, decrease transaction costs, and facilitate government accountability⁵. Hence, a well-functioning central public procurement platform should contribute to achieving value for money in public procurement as well as increase integrity throughout the public sector. The DIGI-WHIST portal opentender.eu featuring all of the above functionalities will be launched in 2017, filling the gap where source data quality allows.

^{1 -} Cingolani, L., Fazekas, M., Kukutschka, R. and Tóth, B. (2016). Towards a comprehensive mapping of information on public procurement tendering and its actors across Europe.

 $^{{\}tt 2-The\ Tender\ Electronic\ Daily\ (TED)\ can\ be\ accessed\ here: http://ted.europa.eu/TED/misc/chooseLanguage.do}$

^{3 -} This is the case in Belgium, France, Germany, Iceland, Italy, Malta, Spain, and the UK.

^{4 -} In such countries, typically various private companies step in to fill the informational void, however in return for levying an access fee while effectively preventing governments from controlling data quality and content

⁵ - This has also been recommended by the G20. For more information see: G20 Principles for promoting integrity in public procurement. Available at: http://g20.org.tr/wp-content/uploads/2015/11/G20-PRINCIPLES-FOR-PROMOTING-INTEGRITY-IN-PUBLIC-PROCUREMENT.pdf

2. Governments should commit to publishing procurement data by default in an open and simple data format

Currently, the format and accessibility of public procurement data greatly varies across Europe.6 In terms of diversity of publication templates used for publishing call for tenders, modification or cancellation of tenders, and contract awards, five countries7 have no publicly defined publication standard, while all the other countries use on average more than eight templates. Having no publicly defined publication template essentially disposes governments of the capacity to directly control data content and quality. On the upside, complexity of reporting standards is close to the minimum possible in 13 of the examined countries with a defined standard, as they use three or fewer templates for the three types of announcements (Figure 1). Having so few reporting formats makes the understanding and processing of public procurement information the least cumbersome and thus represents the optimal scenario.. Quite problematically, several countries as well as the EU-wide TED portal use a large number of different templates - more than 15 or even 20 - which adds an unnecessary layer of complexity to understanding and reusing the data.

In terms of the accessibility of public procurement data, that is, the machine-readability of data, the current situation in Europe is quite problematic leading to fundamental barriers to data reuse.⁸ In only three countries⁹ and the EU-wide TED is

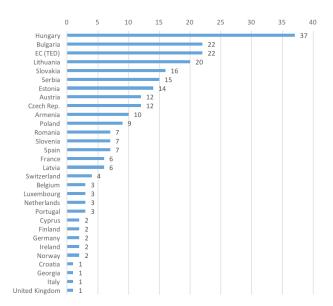


Figure 1.

Reporting format complexity

Number of national announcement types used for call for tenders, contract awards, and contract modifications on the main national portal is optimal only in about half of European countries

public procurement data fully machine readable, and even in these cases downloading data is made difficult by unnecessary hurdles contradicting basic open data principles (e.g. Polish national public procurement data is downloadable in XML format, but files take the form of self extracting RAR archives, which are hard to automatically collect). In 26 countries, data is only semi machine-readable, meaning it is possible to automatically collect the information, but it is not consistently and clearly structured in most cases. This makes database building a costly, lengthy, and error-prone endeavour. Meanwhile, in five countries data is not even semi machine-readable, meaning that

^{6 -} Cingolani, L., Fazekas, M., Kukutschka, R. and Tóth, B. (2016). Towards a comprehensive mapping of information on public procurement tendering and its actors across Europe.

^{7 -} These countries are Denmark, Greece, Iceland, Malta, and Sweden.

^{8 -} Cingolani, L., Fazekas, M., Kukutschka, R. and Tóth, B. (2016). Towards a comprehensive mapping of information on public procurement tendering and its actors across Europe.

 $^{{\}bf 9}$ - These are Belgium, Poland, and the UK.

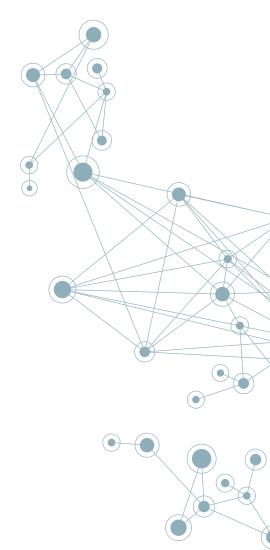
^{10 -} Czibik, Á., Tóth, B., & Fazekas, M. (2015). How to Construct a Public Procurement Database from Administrative Records? With examples from the Hungarian public procurement system of 2009-2012. Government Transparency Institute: Budapest. Available at: http://www.govtransparency.eu/wp-content/uploads/2015/12/GTI_publicprocurement_techreport_1512221.pdf

^{11 -} These are Armenia, Finland, Iceland, Malta, and Sweden.

by and large only manual data collection is reliable enough for building a public procurement database. For example, in these cases data may be stored in scanned PDFs which are difficult for algorithms to parse.

Publishing public procurement data in a timely, simple, and easy-to-understand format and publishing information as machine readable data are essential for lowering the barriers to data use and reuse by all stakeholders. Such elementary standards are expected from EU accession countries12 in spite of most member states not living up to them. As recommended by international civil society organisations like the Open Knowledge Foundation, the Sunlight Foundation or the Open Contracting Partnership, governments should adhere with machine-readable file formats such as CSV, JSON, and XML to ensure usability. Users should be also able to download data in bulk either as .csv or through an Application Programming Interface (API)13. The number of data publication forms should be kept to the very minimum in order to minimize complexity, facilitating stakeholder engagement with the data. This implies that the benefits of different formats due to different regulatory and legacy requirements (e.g. unique rules applying for public utilities) should be set against the high costs they impose on non-specialist stakeholders who wish to understand and use the data. Crucially, public procurement is a highly technical domain all across Europe and is already proving challenging to many small and medium-sized enterprises, journalists, and civil society, as well as government agencies in some cases.

To ensure data format and accessibility meet user needs, governments should establish a monitoring, evaluation, and learning process involving data users as well as data producers. This is also backed by the recommendations set forth in the Open Government Partnership (OGP) commitments and should be implemented by all OGP members¹⁴.



^{12 -} OECD/Sigma. (2014). The Principles of Public Administration. Paris: OECD/Sigma.

^{13 -} See also the Open Data Handbook by the Open Knowledge Foundation. Available at: http://opendatahandbook.org and the Procurement Open Data Guidelines by the Sunlight Foundation. Available at: https://sunlightfoundation.com/procurement/opendataguidelines/.

^{14 -} Open Government Guide. Public Procurement. Disclose key documents and data. Available at: http://www.opengovguide.com/commitments/publish-key-documents-and-data/ Those principles are also set forth in more general in the International Open Data Charter. Available at: http://opendatacharter.net/principles/

II. REGULATORY SCOPE AND DATA COVERAGE

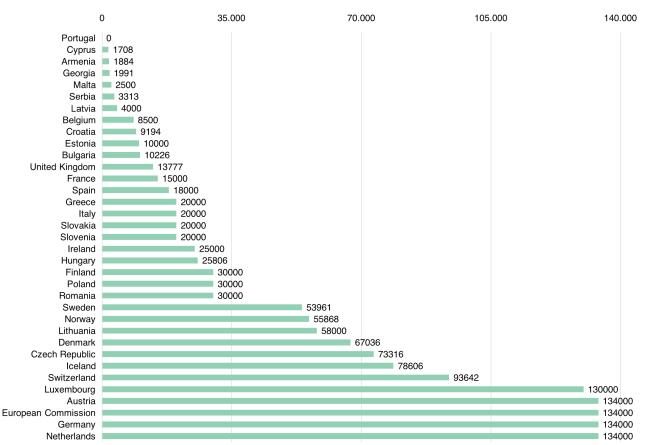
3. Governments should require low reporting thresholds with the same regulatory framework for all public bodies and spending areas

Transparency requirements are largely determined by monetary thresholds mandating the application of certain procedural and reporting rules if contract values exceed them. Below such thresholds specified in the EU Public Procurement Directives, national governments can apply their own regulations with the corresponding contract value thresholds, leaving contracts falling under these national thresholds largely unregulated (e.g. direct award without competition is permitted). A small

Figure 2.

Scope of public procurement databases

Wide variation across Europe in terms of the minimum contract value threshold for publishing supplies and services contracts (2015, EUR) group of European countries apply low monetary thresholds establishing transparency throughout most of their public procurement spending (Figure 2), with Portugal going as far as requiring all contracts published at the national portal. At the other end of the spectrum, four countries—Austria, Germany, Luxembourg, and the Netherlands—have no national mandatory threshold below the EU thresholds, implying they essentially do the bare minimum in terms of public procurement regulatory scope. This is highly problematic not only in terms of government accountability, but also economic efficiency. Major corruption scandals and

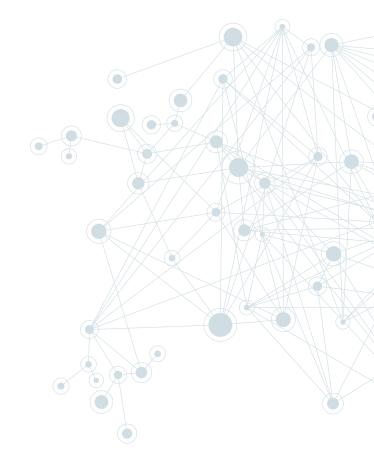


failed investments in countries like Germany¹⁵ or the Netherlands¹⁶ amply demonstrate the risks of not establishing a sufficient degree of transparency in an area covering up to one third of government spending.

Exempting some public bodies, such as public utilities, or some markets, such as defence, can be justified on the grounds of administrative burden or national security concerns; however, exceptions are often abused for corrupt reasons.¹⁷ Exceptions abound throughout Europe,¹⁸ though no thorough evaluation exists on the degree of abuse and the impact on public integrity.

Governments should implement low monetary publication thresholds and apply public procurement rules to all public bodies and spending areas. Ideally, monetary thresholds should be between o€ and 5,000€ so that most public spending through public procurement systems is transparently published and regulated. In order to carefully balance the demand for transparency and the associated administrative burden, a lighter regime could be applied to the smallest value contracts, with full procedural and transparency rules applied for higher value contracts starting from about 20-40,000€.

The scope of public institutions outside the remit of public procurement law and publication requirements such as public utilities or local governments and sectoral exceptions such as defence and national security should be minimised in order to reap full benefits of market transparency and minimise gaming of exceptional rules.



 $^{{\}bf 15}$ - See for example the infamous Berlin-Brandenburg airport.

^{16 -} Van Den Heuvel, G. (2006). The Parliamentary Enquiry on Fraud in the Dutch Construction Industry Collusion as Concept Between Corruption and State-Corporate Crime, Crime, Law and Social Change, 44(2), 133–151.

^{17 -} OECD. (2007). Integrity in Public Procurement. Good Practice from A to Z. Paris: OECD. Available at: http://www.oecd.org/development/effectiveness/38588964.pdf

^{18 -} For full data on legally mandated exceptions see: http://europam.eu/

III. DATA DEPTH

4. Governments should increase the depth of procurement data published, in particular publishing contract implementation data, reliable data on corrected and failed tenders, and organisational IDs by default.

Procurement processes typically span over a long period of time and include various stages, such as call for expression of interest or pre-announcement; a formal call for tenders; announcements of the winning bidder(s); contract implementation; and ex-post assessments of performance. Any of these stages may also be subject to modifications, cancellations, or legal challenge. As public procurement typically implies a highly structured and complex procedure, there are many variables or bits of information which should be reported at each stage in order to provide sufficient information to interested bidders and civil society.

First, it is quite problematic that almost all European public procurement systems fail to provide any information on contract implementation.

Regulatory frameworks as well as publication systems are set up as if public procurement stopped at contract award. Naturally, contract implementation information is recorded in other administrative systems and a different set of rules apply, such as public financial management rules; however, using such disconnected systems makes the monitoring and adequate understanding of public procurement performance and corruption risks very hard.

Only eight out of 35 jurisdictions actually publish contract performance information as part of their public procurement systems, with Hungary stopping this practice in 2011 (Figure 3).¹⁹

Figure 3.

Coverage of the full tender cycle

Only a few countries publish information on contract implementation (full coverage was only available until 2012 in case of Hungary)

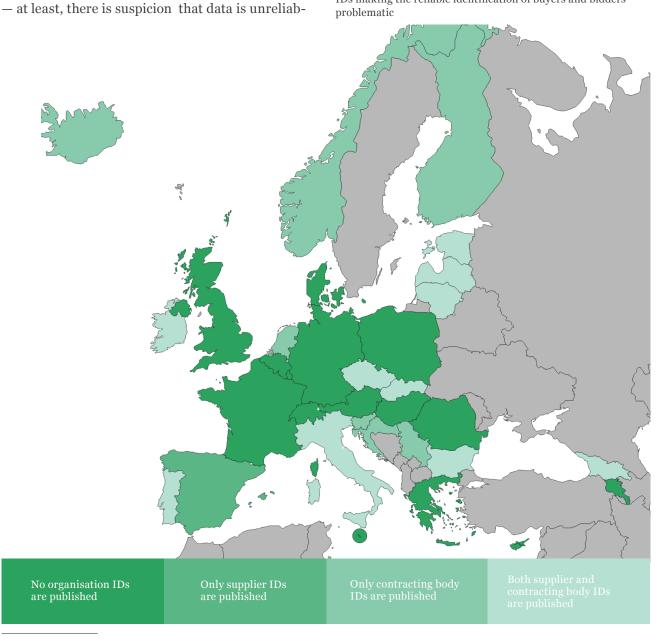


^{19 -} Cingolani, L., Fazekas, M., Kukutschka, R. and Tóth, B. (2016). Towards a comprehensive mapping of information on public procurement tendering and its actors across Europe.

Second, due to complexity, a high degree of regulation, and the inherent uncertainty of purchasing complex goods and services, corrections and failed tenders are frequent even in countries with high administrative capacity. Where approximately reliable data is available in countries like Sweden or Hungary, the rate of failed tenders ranges between 5-13% of tenders launched. Such figures underline the importance of better monitoring and understanding of failed tenders, which impose considerable costs both on buyers and bidders. However, in many public procurement data systems, modifications and failed tenders are not adequately logged

le in this respect. There is no data point indicating tender failure, making failed tenders look like tenders with incomplete information — two scenarios with different consequences. Moreover, administrative corrections and modifications of already published announcements are typically recorded and published in unstructured text fields, making it close to impossible to automatically identify the corrected information and attach it to the original announcement.²⁰

Figure 4. **Reporting unique organisational identifiers**Most European countries fail to report unique organisational IDs making the reliable identification of buyers and bidders



^{20 -} Crucially, the new TED standard forms under public consultation until January 2017 aim to address this issue by requiring the corrected information to be republished in the same format as the original incorrect information making corrections tractable.

Third, even after ensuring that publication requirements span through the full tender cycle and information modifications are appropriately dealt with, a crucial question hinges in the air: which data fields are currently reported, and which ones should be reported? Currently, there is a wide array of country practices across Europe,21 with some countries (such as the UK) reporting only a handful of variables in their national public procurement systems, while many others (such as Poland) use an extensive data template, going even further than the EU-wide TED templates. Important gaps nevertheless persist across countries. For example, only five countries²² report information on losing bidders and their bids. Bidder information is essential for understanding market dynamics and open access, as well as for developing indicators of collusion in public procurement. Another crucial variable which is typically missing in most European data systems is organisational identifiers, which allow for assessing organisational performance by reliably and uniquely identifying buyers and bidder (Figure 4.).

Governments should increase the depth of public procurement data publication in order to allow for the comprehensive monitoring of public procurement processes underpinning good government, efficient competition, and government accountability. In particular, we recommend that governments:

- publish public procurement data relating to the whole procurement cycle on existing public procurement platforms, including at a minimum: call for tenders, contract awards, and contract completion/ implementation announcements.
- publish information on amendments, modifications, and failed tenders in a structured and reliable format so that up-to-date information is available on all tenders.
- publish at least a minimum set of variables essential for government accountability and transparency of bidding (Table 1.). Crucially, unique organisational identifiers linkable to external registries such as company registries is a fundamental precondition to monitoring organisational performance. DIGIWHIST proposes an ideal list of variables and their publication location for governments. This list is close to the Open Contracting Data Standard²³ to connect to already existing international standards but contains some specific variables for the European context, underpinning effective monitoring across the continent.

^{21 -} Cingolani, L., Fazekas, M., Kukutschka, R. and Tóth, B. (2016). Towards a comprehensive mapping of information on public procurement tendering and its actors across Europe

^{22 -} These countries are Bulgaria, Georgia, Hungary (until 2012), Malta, and Sweden.

^{23 -} More information on the Open Contracting Data Standard including variables can be found here: http://standard.open-contracting.org/latest/en/schema/

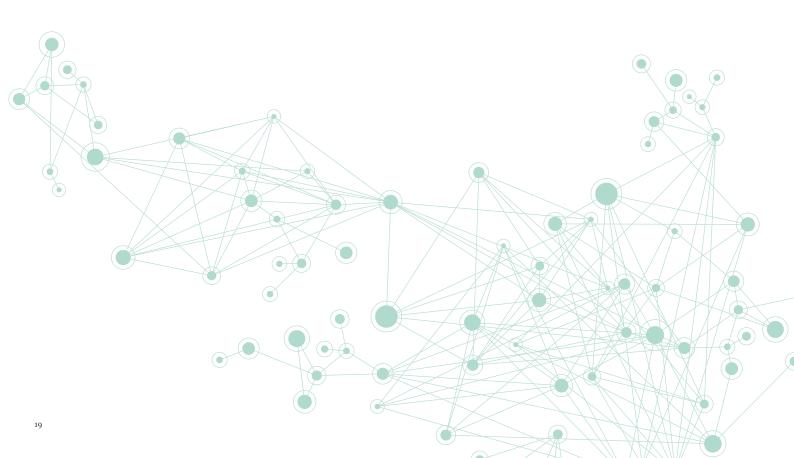
A parsimonious list of key variables is able to support government accountability and transparency of bidding

Variable group	Variable	Inclduded in the announcement		
		call for tender	contract award	contract im- plementation
Buyer	Buyer's name	•	•	•
	Buyer's department/office	•	•	•
	Buyer's unique ID	•	•	•
	Buyer's address	•	•	•
	Buyer's type	•	•	•
Bidder / bids	Bidder's name		•	•
	Bidder's unique ID/tax ID		•	•
	Bidder's address		•	•
	Number of bids submitted		•	
	Number of bids excluded		•	
	Bid price (details on total and unit prices)		•	•
	Exact time of bid submission		•	
	Bid type (winner/loser bid)		•	
	Beneficial owners		•	•
Tender / contract	Tender unique ID	•	•	•
	Procedure type	•	•	
	Framework agreement (1st/2nd stage)	•	•	
	Award criteria	•	•	
	Threshold (below/above EU thresholds?)	•	•	
	Estimated price (details on total or unit prices)	•	•	
	Procurement type (service, supply, work)	•	•	•
	CPV codes (% contract value per product)	•	•	•
	NUTS code(s) of contract implementation	•	•	•
	Status (cancelled, pending, etc.)	•	•	•
Dates	Call for tender publication date	•	•	•
	Bid submission deadline	•		
	Contract start and end dates	•	•	•
	Publication date of contract award		•	
	Contract signature date			
	Publication date of contract completion			•
Subcontracting	Subcontractor's name and unique ID (tax ID)		•	•
	Subcontractor's share		•	•
Consortium	Consortium members' name and unique ID (tax ID)		•	•
	Consortium members' share		•	•
Contract performance	Contract performance end date			•
	Was performance according to the contract			•
	Explanation in case of deferring from contract			•
	Information on contract modification			•
	Information on performance quality			•

5. Governments should facilitate the link between public procurement data and further datasets

Currently, in every European public procurement system reviewed by DIGIWHIST, public procurement data is stored in a data and publication system disconnected from other databases describing the key actors of public procurement transactions. Linking these datasets is not possible, at least not without considerable investment, even if a government desired to do so because of the aforementioned missing organisational identifiers. Storing public procurement data separately from linked datasets is problematic as public procurement is a cross-cutting government function with numerous links to organisational financial performance, such as public body budget deficit, and to sector-specific outputs, such as hospital mortality or road usage.

We recommend that governments establish the link between public procurement data and related datasets describing organisational behavior and performance through the use of common organisational and contract IDs across different data systems, such as public procurement, payments, company registry, or court rulings. Sectoral public sector outputs such as quality of roads is what ultimately matters for citizens, hence increasing the trust in complex public procurement systems and minimizing corruption risks should build on linked data. Linking datasets would not only allow for better understanding how public money is spent but also for better risk assessment of public procurement processes. Data on company ownership and the data on individuals involved in public procurement could be linked in order to measure and eventually control the risks of favouritism.



6. Governments should link public procurement announcements to original procurement documents by default

Typically, public procurement data as published in public procurement platforms only contain summary information of the original full procurement documents like full tender specification or the signed contract. In 22 jurisdictions out of 35, public procurement announcements contain the links to the original procurement documentation by default, while in 13 jurisdictions such link is not a mandatory part of announcement templates. ²⁴ Directly linking announcements to the original full documentation greatly decreases transactions costs and decreases the probability of corruptly providing crucial tendering information to selected bidders.

Such direct linking of public announcements and the underlying original full documentation such as the signed contract allows for easily verifying the content of public procurement announcements as well as for further investigations into any data point worthy of public interest due to corruption risks, for example. This is also backed by an OGP recommendation, which says to open up other relevant datasets (including international aid, budgets, financial management, corporate registers and beneficial ownership of companies, project

management, political lists, and sector-specific performance data) and to create tools to link and analyse these datasets to inform evidence-based decision making²⁵.

Hence, governments should link procurement announcements to all relevant original documents. Those should include the full tender documentation and maps, plans, etc. Ideally, signed contracts should also be linked and easily available. Here, information on sub-contractors as well as contract amendments, invoices, and completion reports submitted should be linked to the dataset. Submitted bids, or at least parts of them, may be exempt from these stringent transparency rules for protecting commercially sensitive information or privacy of individuals.

^{24 -} Cingolani, L., Fazekas, M., Kukutschka, R. and Tóth, B. (2016). Towards a comprehensive mapping of information on public procurement tendering and its actors across Europe

^{25 -} Open Government Guide. Public Procurement. Link open contracting data with other sources of data. Available at: http://www.opengovguide.com/commitments/link-open-contracting-data-with-other-sources-of-data/

^{26 -} Šípoš, G., Samuek, S., & Martin, K. (2015). Not in force until published online. What the radical transparency regime of public contracts achieved in Slovakia. Bratislava: Transparency International Slovakia. Available at: http://www.transparency.sk/wp-content/uploads/2015/05/Open-Contracts.pdf

IV. DATA QUALITY

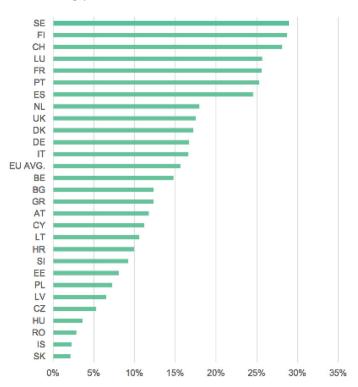
7. Governments should introduce control mechanisms to ensure data quality is maintained

Missing, incomplete, and erroneous data is one of the fundamental problems in European public procurement data systems.²⁷ Even in countries with comprehensive templates for data reporting and a central public procurement authority overseeing the data generation process, the quality of the administrative data is low, predominantly due to poorly designed online platforms and the lack of enforcement.28 The most basic measure of data quality, the amount of missing information, shows how badly most EU countries fare, especially the better governed and richer member states like Sweden or Germany (Figure 5). Crucially, investment in data infrastructure pays: new member states with more comprehensive public procurement data systems and more central control fare comparatively better, as is the case with Slovakia and Romania, for example.

Accordingly, we recommend that existing data reporting requirements are adequately enforced and data quality is increased to the legally mandated minimum throughout Europe. To ensure all required data fields are filled out with truthful information, governments should introduce centralised control mechanisms and penalties for

Extent of missing information

Data quality is low throughout Europe with 15% of mandatory fields empty in the EU-wide TED data in 2009-2015



non-compliance. Punishment of non-compliance in extreme cases can include the freezing of the tender or interrupting payments until records are corrected, practices employed in Slovakia, for example²⁹. But more light-touch enforcement may also produce the desired data quality, such as sending automatic clarifying questions back to the

^{27 -} Fazekas, M. and Toth, B. (2016) Assessing the potential for detecting collusion in Swedish public procurement. KKV: Stockholm. Available at: http://www.konkurrensverket.se/globalassets/publikationer/uppdragsforskning/forsk_rapport_2016-3.pdf

^{28 -} More detailed discussion on data quality issues see: Czibik, Á., Tóth, B., & Fazekas, M. (2015). How to Construct a Public Procurement Database from Administrative Records? With examples from the Hungarian public procurement system of 2009-2012.

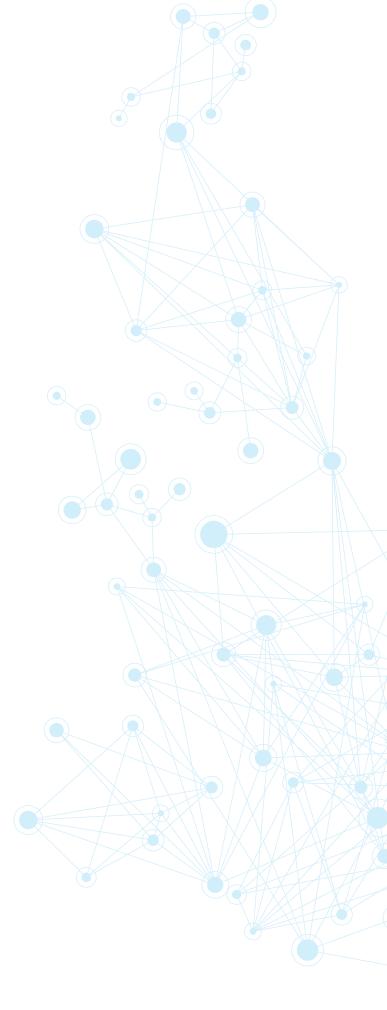
^{29 -} Šípoš, G., Samuek, S., & Martin, K. (2015). Not in force until published online. What the radical transparency regime of public contracts achieved in Slovakia.

reporting public bodies until all missing information and inconsistencies are resolved. While data completeness and adequate publication may not appear crucial for public procurement outcomes, more information in an easily accessible format leads to better competition and better outcomes.³⁰

8. Governments should lower bureaucratic burden by linking publication systems to tender, contract and payment management systems

Currently, public procurement data is only rarely automatically filled in from linked administrative databases such as corporate registries (see above on data linking), as such links are not established to start with. This means that in most European countries, standard elements of public procurement announcements and records such as winning company name and address have to be manually entered, which creates additional unnecessary work and increases the potential for error.

If administrative datasets are linked, the officially verified (and supposedly correct) information should be automatically added to public procurement records to lower public procurement administrators' administrative burden as well as minimizing the risk of erroneous data entry. This is also recommended by the G20 and should be followed by the EU as a guideline³¹.



^{30 -} Coviello, D., & Mariniello, M. (2014). Publicity requirements in public procurement: Evidence from a regression discontinuity design. Journal of Public Economics, 100, 76–100.

^{31 -} G20 Principles for promoting integrity in public procurement.

V. DATA USE

9. Governments should encourage the regular use of public procurement data both inside and outside governments

As there is very little structured data, there are only few government agencies and non-governmental organisations which actually make use of public procurement databases in a substantive way. In many countries such as Slovakia or Hungary, civil society has stepped in to create public procurement datasets and build portals that make the data easily understandable and actionable. Reuse of public data has the capacity to contribute to greater competitiveness and more accountable government if stakeholders can understand it, act on it, and pursue change.

Governments who are the principal data guardians should promote the use of public procurement data within government and facilitate data reuse by non-governmental stakeholders such as civil society watchdogs and data provider firms. The use of public procurement data should be facilitated first by creating direct feedback mechanisms throughout the entire procurement process cycle (i.e. planning, tendering, awarding, implementation) involving all stakeholders from within and outside

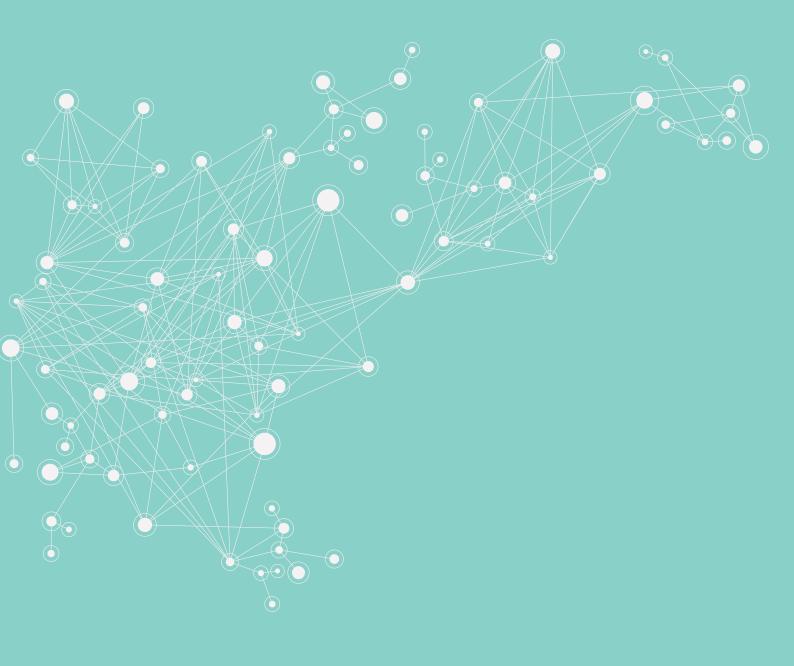
government. Second, governments should support non-governmental organisations which monitor, analyse, and investigate issues in the process as a friendly ally to both procuring entities and monitoring bodies such as prosecutors.

There is plenty of global good practice of engaging stakeholders in verifying, monitoring, and acting on public procurement data and analytics. For example, Integrity Pacts in Europe³² and Social Witness programs in Mexico³³ provide potentially effective ways of strengthening the importance of user views and civil society monitoring. In addition, stakeholders can be given the opportunity to provide feedback to existing procurement processes in the case of irregularities, for example in the implementation phase, but also on the data published (for instance, by marking incomplete datasets). Some countries like the UK have already started working on stakeholder engagement practices, with one example being the construction of Heathrow Terminal 5.34

 $³²⁻European\ Commission\ (2017)\ Integrity\ Pacts.\ Available\ at: http://ec.europa.eu/regional_policy/hu/policy/how/improving-investment/integrity-pacts/$

^{33 -} Open Government Guide (2017) Country Example. In Mexico "social witnesses" oversee public procurement. Available at: http://www.opengovguide.com/country-examples/in-mexico-social-witnesses-oversee-public-procurement/

 $^{34 -} OECD\ (2016)\ Public\ Procurement\ Toolbox.\ Country\ case:\ Stakeholder\ engagement\ during\ the\ construction\ of\ Heathrow\ Airport\ Terminal\ 5.\ Available\ at:\ https://www.oecd.org/governance/procurement/toolbox/search/stakeholder-engagement-during-construction-heathrow-airport-terminal-5.pdf$



WWW.DIGIWHIST.EU

Authors: Mara Mendes, Mihály Fazekas Design: Bela Seeger, Nadine Stammer