



Innovation
Procurement
Empowerment
Centre

A Connected Places
Catapult Initiative

Transparency on the reporting of public procurement information: lessons from handling compiled procurement information

Maria Fernanda Ibarra Gutierrez, Dr. Luis Ospina-Forero
and Prof. Elvira Uyarra

Research brief no. 10/2024 – 3

Publication Date: 10/24



UNIVERSITY OF
BIRMINGHAM

MANCHESTER
1824

The University of Manchester

Executive Summary

This research brief examines key challenges in ensuring transparency in UK public procurement data. The lack of unique identifiers, duplicated records, inconsistencies in contract dates, and incomplete data fields impede the ability to monitor procurement processes effectively. These challenges hinder transparency, accountability, and data-driven decision-making in public procurement. Key issues include:

- **Data Quality:** The absence of a centralised portal for procurement information leads to duplication and inconsistencies. Different platforms are used to publish tender notices, but they are not integrated, and there are no unique identifiers across platforms.
- **Missing Data:** Key procurement information such as contract dates and values is often incomplete, and there is no systematic tracking of suppliers during the bidding process.
- **Inconsistent Workflow Tracking:** It is difficult to trace the full life cycle of procurement from pre-information notices to contract payments due to unlinked data across stages.

The upcoming **Procurement Regulation Act** aims to address some of these issues by centralising data platforms and introducing supplier registration systems, which will improve data consistency and traceability. Improved data management is critical for enhancing competition, transparency, and public accountability in procurement processes.

1 | Context of public procurement information

Public procurement happens when the government or a public organisation (referred to as contracting authorities or buyers) intends to acquire work, supplies or services from the private sector or non-public parties (referred to as suppliers or bidders), and where this acquisition also needs to safeguard the public value of the public organisation (Grandia and Volker, 2023).

Some of the benefits of having good procurement data are:

- **Procurement markets and trends monitoring:** Governments can describe spending and time trends, and compare performance across entities, regions, contract types, etc.
- **Data-driven procurement policy-making:** Governments can assess efficiency gaps to identify areas for reform, monitor the impact of new policies, and understand potential trade-offs of different strategies.
- **Transparency and accountability:** Civil society can monitor the procurement system.
- **Integration of procurement functions with other public and private sectors:** When integrated with other administrative

data sources, procurement data can provide a more holistic picture of the procurement function and its links to other parts of government and the private sector (Cocciolo et al., 2023).

As the overall public procurement process occurs in sequential stages, different types of information can be captured depending on the stage of the process. As shown in Figure 1 (procurement workflow), four overarching types of information can be captured:

1. Pre-information notices are presented at the starting stage of the process in which contracting authorities are allowed to publish preliminary information about their commercial need to get marketing engagement.
2. The contracting authority calls for suppliers' participation by publishing a tender notice, which is a document that describes the process of awarding a public contract (Hrubý et al., 2018).
3. The contracting authority awards a contract to one or several suppliers.
4. Once a contract is formalised, the spending stage starts as contracting authorities make payments to suppliers.

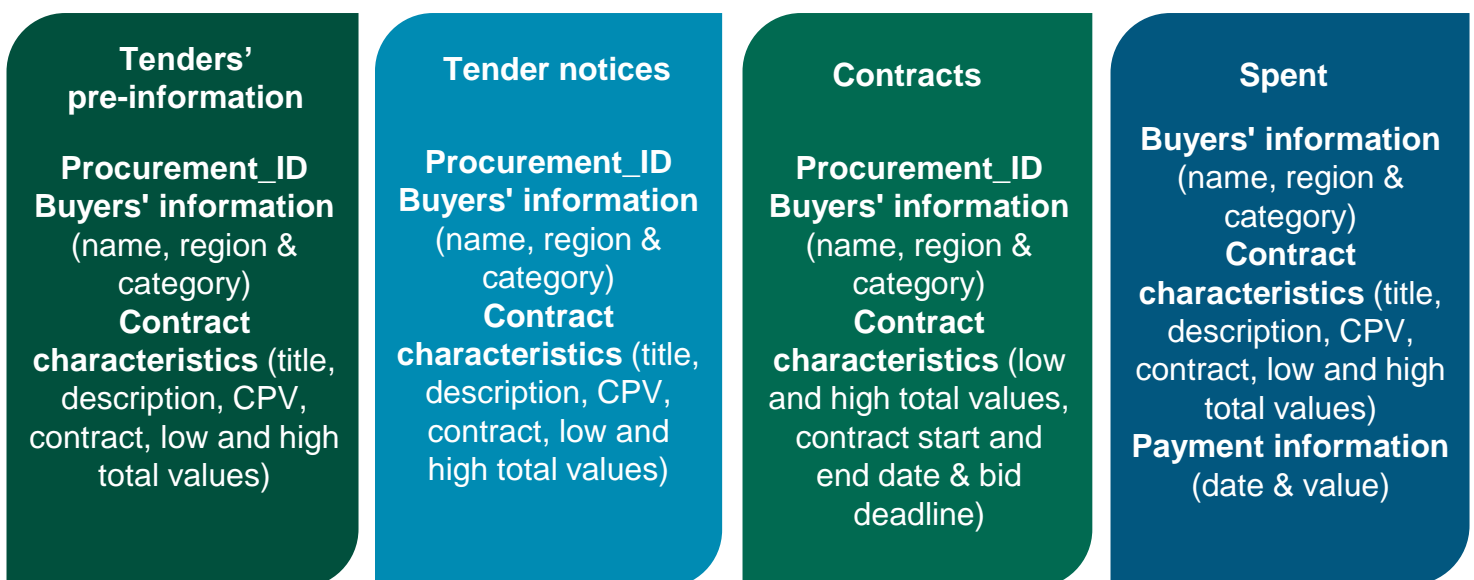


Figure 1: Public procurement process, and example of what incremental information can look like.

Datasets	Number of records	Number of variables
Tender Pre-Information	37,234	34
Tenders	250,235	34
Contracts	569,292	68

Table 1: Datasets about public procurement

2 | Data considered

On 11 May 2023, we downloaded the Tenders and Contract Awards data from Tussell. The datasets spanned a time range from 2012 to 2023.

3 | Key challenges of UK procurement data

Based on the exploration of the procurement datasets obtained, we have identified three key challenges that may hinder transparency and best practices in the reporting of public procurement information, as shown in Figure 2.

3.1 | Data quality challenges

Pre-information notices, tender notices and contract awards may be reported or published in more than one portal or more than one time in the same portal.

Currently, there is no central portal where UK tender and contract award publications are aggregated. Instead, information is collected from eight different portals. Six portals are domestic portals: *Contracts Finder* (CF), *Digital Market Place* (DMP), *Public Contracts Scotland* (PCS), *Sell2Wales* (S2W) and *In-Tend*; and the *Tenders Electronic Daily* (TED) is the official portal of the European Union (EU) which was substituted by the e-notification service *Find a Tender* (FTS) in 2020. Among these portals, CF and FTS are the most relevant portals.

In 2015 the Public Contracts Regulations (PCR 2015), started to set rules over the public procurement procedures.¹ Overall, contracting

authorities are allowed to publish tender opportunities in any portal, as long as those opportunities are also published in the CF portal. Similarly, according to the Procurement Policy Note 08/20, high-value opportunities must be published in both, FTS and CF portals (Cabinet Office, 2023).

As notices can be published independently in every portal without the requirement of a unique identifier, linking the same publication across different portals is not a standardised procedure. Furthermore, when a tender is updated, it may not be updated across all platforms. Thus, even if it were possible to link a tender, for example between the CF and FTS portals, it may not be clear which platform has the correctly updated information.

Access to historical data and historic comparability.

The procurement datasets obtained spanned a time range from 2012 to 2023 and provided records from the eight different portals previously mentioned. However, our exploratory exercise shows that not all historical information was available for all portals. Only procurement information from the CF and TED portals was captured from 2014 and 2012, respectively. Data from the other portal was collected from 2020, although these portals were launched before that year. For example, S2W was launched in 2008² and DMP was launched in 2014.³

Our exploration also highlights the potential inability to directly track and compare the evolution of the procurement system over time.

¹ Accessed on 17 December 2023 in <https://www.legislation.gov.uk/uksi/2015/102/contents/made>

² Accessed on 23 January 2024 in https://www.publiccontractsscotland.gov.uk/sitehelp/help_faq.aspx#1

³ Accessed on 23 January 2024 in <https://www.gov.uk/government/news/new-digital-marketplace-opens-forbusiness>

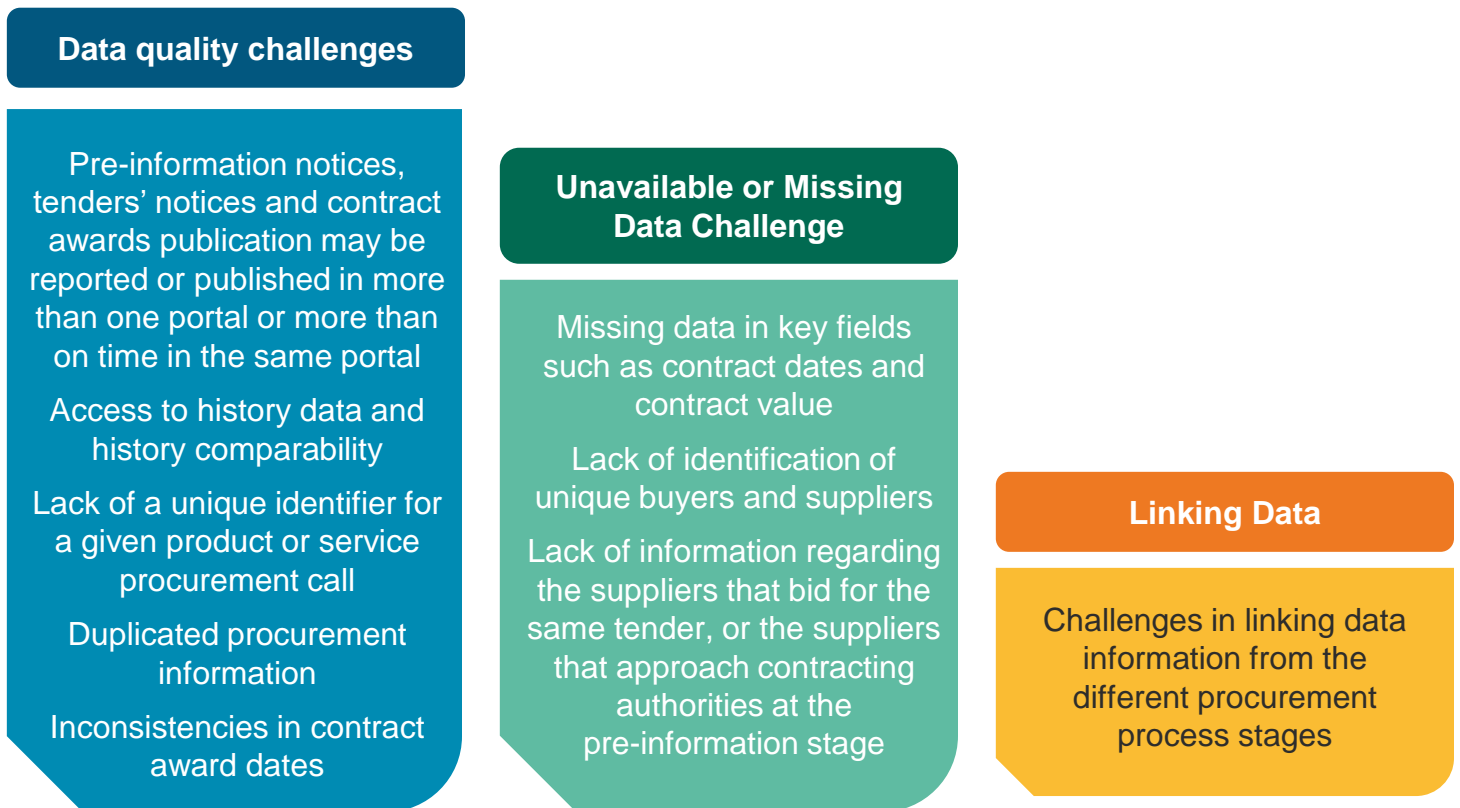


Figure 2: Public procurement process, and example of what incremental information can look like.

This constraint appears when, for example, new rules (such as the PCR 2015) come into place and alter the reporting protocols of the information, such as what should be reported, where it should be reported, what new cost thresholds apply, etc. In this regard, our data exploration provided indications that artificiality show different properties in the procurement system, which may simply be due to changes in how the data is reported.

Lack of a unique identifier for a given product or service procurement call.

Pre-information notices, tender notices and contract awards do not have a unique identification number (*Procurement_ID*) for each procurement call or notice; and to our knowledge, there is no regulation regarding setting a unique identification number per notice across portals or within the same portal. The lack of unique identifiers allows the possibility of duplicate records within the same portal or across portals, as one notice can be published more than once with the same information but with a different identification

number.

This problem is also present throughout the life cycle of procurement call (Figure 3), and along the spent data as well, where it may not be possible to directly track the “evolution” from pre-information notice, tender notice, contract and payments related to a contract. Furthermore, for large procurement projects, a notice can be split into several lots that are associated with it. However, every notice lot can have a different *Procurement_ID* at every stage of the process. Therefore, it may not be possible to identify the procurement call life cycle from a group or notices that belong to the same project. In such a situation, it may not even be possible to differentiate different lots from a lot that has been mistakenly duplicated in the system (and which can have a different ID or come from a complementary procurement portal).

In terms of transparency, this is relevant, as high-value projects can be split into multiple tender lots for which traceability throughout their life cycle will be hindered.

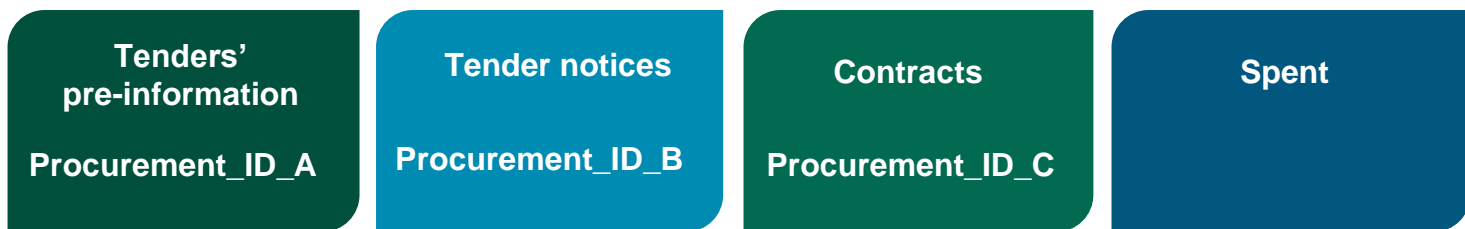


Figure 3: Example of public procurement process showing different identification numbers in each stage

Duplicated procurement information.

Pre-information notices, tender notices, and contract awards might be duplicated because, as previously mentioned, notices do not have a unique identifier across portals or within portals. Additionally, some notices may be published in more than one portal since, according to the PCR 2015, notices published in any national portal also have to be published in the CF and/or FTS portals (Cabinet Office, 2023). Also, there might be duplicated records of the same notice in the same portal, due to errors or corrections, but with a different identifier.

Inconsistencies in contract award dates.

Contracts reported datasets contain four dates that define their time frame: *contract awarded date*, *contract published date*, *contract start date* and *contract end date*. The PCR 2015 states that the contracting authority has 30 or 90 days to publish the contract in the CF portal, once it is awarded, depending on whether it is the central or local government (Cabinet Office, 2023). Therefore, we consider that contract awarded date should occur before contract published date. Similarly, a standard assumption is that contract start date occurs before contract end date and that contract published date occurs before or on the same day as the contract start date. Figure 4 thus shows the expected order these dates should follow.

In our exploratory analysis, we found some inconsistencies in the order of these contract dates. For instance, there were records across all years and all platforms, in which the contract start date was before the contract award date, suggesting the contract started before being awarded. Also, in some records across all years and all platforms, the contract's publish date was before the

contract's awarded date, which suggests the contract was published before being awarded. Finally, among all CF records, we observed that the PCR 2015 guideline which states that a contract must be published 90 days after being awarded is not met in all cases.

3.2 | Unavailable or missing data challenges

Missing data in key fields such as contract dates and contract value.

Our exploratory analysis also revealed potential missing data in procurement information. Specifically, pre-information notices, tender notices and contract data, across all portals and all years, show high levels of missing data in fields such as *Contract start date*, *Contract end date*, *Contract duration*, *Contract value low limit*, *Contract value high limit*, *Contract has EU funding*, *Contract suitable for SME* and *Contract suitable for VCO*.

While most of the contract information might not be completely defined at the pre-information stage, we argue that this information should be completely defined in the two following stages, tender notices and contracts. In terms of transparency, it is relevant to keep complete or close to complete records, especially in variables which are relevant to the contract such as contract amount and duration. This observation aligns with other researchers who have stressed the importance of having complete records to avoid contracting authorities systematically leaving sensitive information blank (Cocciolo et al., 2023).

Lack of identification of unique buyers and suppliers.

In our exploratory analysis, it was found that there is no true variable that can identify unique buyers or unique suppliers. The closest identification variable, the name of a supplier or buyer, is not good enough, as some buyers and suppliers can change their names across data sources or across time. This was mainly observed in suppliers who are more likely to change names over time, and different contracting authorities can register the same supplier with slightly different name variations (Committee of Public Accounts, 2023). However, there are some alternatives that can help to identify unique suppliers; for example, in the dataset explored, a company registration number could be used among the records that provided said information.

This problem has already been reported in other work such as Wachs et al. (2021). The lack of unique variables to identify unique buyers and suppliers leads researchers to follow different processes to de-duplicate buyers and suppliers, which may then impact research reproducibility. This problem might be reduced if accurate catalogues were used in the notices.

Lack of information regarding the suppliers that bid for the same tender, or the suppliers that approach contracting authorities at the pre-information stage.

In the public procurement process, the supplier can approach a contracting authority to offer or discuss the required services, work or supplies in two stages. First, when a contracting

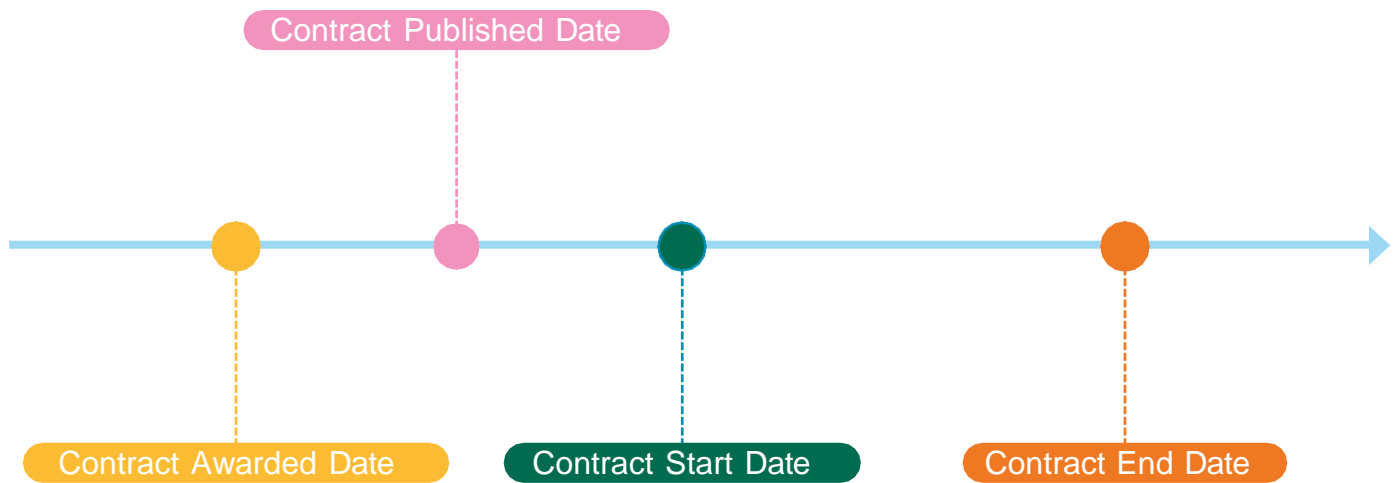


Figure 4: Considered contract time frame.

authority publishes a pre-information notice with broad information about their commercial needs. Second, when a contracting authority publishes a tender notice with a detailed description of the service, work or supply required. Currently, the data available does not store the information about the suppliers that approached contracting authorities in these two stages. We argue that for transparency, competition and anti-corruption purposes, it is important to record and disclose suppliers' participation information along the public procurement workflow (Cocciolo et al., 2023; Wachs et al., 2021).

Challenges in linking information from the different procurement process stages.

As previously mentioned, the public procurement process can be understood as a flow in which a pre-information notice can evolve into a tender notice, then into a contract award, and finally suppliers are paid as part of the contract. Similarly, the procurement process workflow can also involve several notices related to a single big project or recruitment process that evolve into several contracts. Currently, it is not possible to reproduce the public procurement workflow

using the available micro-data. This is mainly due to the challenges previously described.

Since notices do not have the same unique identification number (*Procurement_ID*) at every stage of the public procurement process (Figure 5), it is impossible to perfectly trace a notice's life cycle. Furthermore, procurement

spent data is collected from public invoices that do not have any information related to the contract they belong to, and therefore it may be impossible to link said payments to the respective tender notice with the information currently available. An even more challenging setup is present when a notice is broken down into multiple lots.

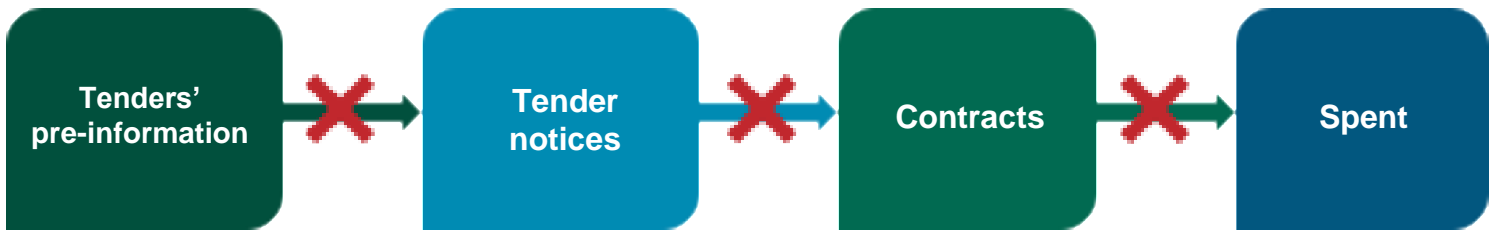


Figure 5: Example of public procurement process showing interrupted flow since data from every stage cannot be directly linked.

4 | Conclusion

In summary, we found three main challenges in the public procurement data. First, regarding data quality, we found that procurement does not have a centralised (or integrated) repository. Instead, there are several portals in which contracting authorities are allowed to publish a call, and according to the PCR 2015, in some cases have to publish the calls in the two main portals, CF and FTS. However, these portals are not integrated, thus it is not possible to identify unique notices. Other data quality issues found are related to the contracts' time frame inconsistencies, as well as the contracts that seem to have been published more than 90 days after the awarded date, breaking the current PCR 2015 regulation.

Second, regarding unavailable information or missing data, we found that relevant fields such as contract dates and contract values are sometimes left blank. Also, there is a lack of information about the suppliers that approach certain contracting authorities in pre-information stage and about the suppliers that bid for the same tender in the contract stage. The only information available is about the supplier which was awarded the contract.

Also, as buyers and suppliers can change names across the data and over time, we noted the unavailability of variables that allow for the direct identification of unique suppliers and buyers.

Third, the available procurement data cannot be linked to recreate the public procurement workflow, starting from pre-information up until payment/expenditure information. This is mainly a consequence of the elements previously found.

In conclusion, our report suggests that it is important to improve the collection, compilation, quality, storage and availability of procurement information. We argue that having optimal procurement data might help authorities to improve competition and ensure a transparent, traceable and easy-to-audit public procurement process. Furthermore, improving the public procurement data might allow researchers to propose innovations to the process and provide better measurements or indicators for monitoring and assessment.

The new Procurement Regulation Act 2023 (PCR 2023)⁴ will be implemented in October

⁴ Accessed on 26 January 2024 in <https://www.legislation.gov.uk/ukpga/2023/54/contents/enacted>

2024. The Act will address some of the problems mentioned in our report. For instance, there will be a platform for suppliers to register and store their details to avoid having different names for the same suppliers. Also, there are plans to join into a single database the information from portals CF and FTS, which is a big step towards centralizing procurement data (Davies, 2023).

References

Cabinet Office (2020). Procurement Policy Note – Introduction of Find a Tender. Action Note PPN 08/20.

Cabinet Office (2023). Guidance on the Transparency Requirements for Publishing on Contracts Finder (HTML).

Cocciolo, S., Samaddar, S., and Fazekas, M. (2023). Government Analytics Using Procurement Data. In Rogger, D. and Schuster, C., editors, *The Government Analytics Handbook: Leveraging Data to Strengthen Public Administration*. World Bank, Washington.

Committee of Public Accounts (2023). Competition in public procurement. Sixth Report of Session 2023–24. Technical report, House of Commons.

Davies, G. (2023). Lessons learned: competition in public procurement. SESSION 2022-23. Technical report, National Audit Office.

Grandia, J. and Volker, L., editors (2023). *Public Procurement: Theory, Practices and Tools*. Springer International Publishing, Cham.

Hruby, J., Pořepný, T., Krafka, J., Toth, B., and Skuhrovec, J. (2018). The Digital Whistleblower: Fiscal Transparency, Risk Assessment and the Impact of Good Governance Policies Assessed. D2.8 Methods Paper.

Wachs, J., Fazekas, M., and Kertész, J. (2021). Corruption risk in contracting markets: a network science perspective. *International Journal of Data Science and Analytics*, 12(1):45–60.



**Innovation
Procurement
Empowerment
Centre**



**UNIVERSITY OF
BIRMINGHAM**



The University of Manchester

Maria Fernanda Ibarra Gutierrez
PhD candidate.
Centre for Data Analytics and Society
The University of Manchester.
E-mail: MariaFernanda.IbarraGutierrez@manchester.ac.uk

Dr. Luis Ospina-Forero
Lecturer in Data Science
Alliance Manchester Business School
The University of Manchester
E-mail: Luis.Ospina-Forero@manchester.ac.uk

Prof. Elvira Uyarra
Professor of Innovation Studies
Manchester Institute of innovation Research
Alliance Manchester Business School
The University of Manchester
E-mail: Elvira.Uyarra@manchester.ac.uk

IPEC website: <https://www.ipec.org.uk/>

IPEC LinkedIn: <https://www.linkedin.com/company/93121184>

IPEC email: contact@ipec.org.uk