**Preface**

This thesis is the result of the graduation project for the Erasmus Mundus joint master’s degree programme in Big Data Management and Analytics (BDMA) at Eindhoven University of Technology (TU/e).

**Chapter 1 Introduction**

The current section aims to provide the motivation for the current thesis. This includes putting into context the magnitude of public procurement and the recurrent problems associated with it, as well as exposing the need for robust analytical methods that facilitate the detection and prevention of threats of corruption, collusion, and anti-competitive practices. Subsequently, this section describes how the emergence of Open Contracting Data Standard generates a new paradigm for the analysis of the transparency and efficiency of public procurement processes, and as such standard is conducive to the application of functional techniques of process mining and data mining. Finally, the research questions that guided this research effort are included.

**1.1 Motivation**

Every year, governments devote gigantic amounts of money to public contracts, from papers and pencils to large infrastructure projects such as airports, roads, schools, and hospitals. The size of public procurement represents a considerable segment of the global economy. In the case of the OECD, the volume of these transactions represents approximately 12% of GDP (OECD, 2017).

This significant volume of resources along with the complex interaction between public and private interests exposes the public contracting processes to multiple risks of waste, inefficiency, fraud, corruption, and uncompetitive practices throughout the whole procurement life-cycle. In a study dating from 2014, the OECD identified that slightly more than half of the cases of bribery (57%) are motivated by the intention to obtain public procurement contracts. In fact, public procurement represents the single greatest corruption risk within the governmental sector, according to multiple international organizations such as the UN, OECD, and the European Commission. Lack of transparency and corruption in procurement is a problem that affects both developed and developing countries, and the losses associated with this type of practices are estimated in the billions of euros every year.

In this sense, technology and especially data analytics can be used as a strategic weapon to reinforce the mechanisms of integrity and fairness promotion of public contracts. The possibilities that new technologies offer for the early detection of anomalous behaviors of bidding or contract allocation allow that cases of fraud and corruption can be avoided even before they materialize.

In the spirit of collaborating in the discussion on the new analytical methodologies that can be applied to ensure the transparency and integrity of public procurement, this thesis proposes a hybrid method that combines process mining and data mining for the analysis of public contracts.

**1.2 Open Contracting Data Standard**

Governments are active protagonists of the data explosion that in recent years has radically modified the way we live. The Open Data movement seeks that governments and public authorities around the world to publish machine-readable data under licenses that allow for unrestricted use, redistribution, modification, separation, compilation, propagation, and application to any purpose without charge and without discrimination against any person or group (Open Knowledge Foundation).

Open datasets address all kinds of issues, but if there are datasets that stand out for their importance, they are those that have to do with public finances. In fact, for many years the public administrations of different countries have published detailed information on their public contracts, with the problem that generally this information does not contain a complete vision of the whole life cycle of the public contracting process. In addition, the diversity of formats in which this data is published makes the task of comparing similar procurement processes in different countries very difficult, and even if comparisons are made between different bodies of the same public administration.

The recognition of these limitations led to the emergence of the Open Contracting Data Standard (OCDS), which enables the disclosure of relevant and comparable data at all stages of the contracting process by defining the common data model. “Open contracting involves the full chain of government deal-making, from concessions of natural resources to procurement of goods, works, and services for citizens. It starts at the planning stage, and covers tenders, awarding, and implementation of all public contracts” (Open Contracting Partnership).

The OCDS was created to support organizations to increase contracting transparency and facilitate deeper analysis of contracting data by a wide range of users. The fact that the opening of data on public contracts goes hand in hand with the implementation of data models such as OCDS builds trust and allows to create mechanisms to ensure that the astronomical sums of money spent by governments each year translate into better goods, services, and infrastructure projects for the direct benefit of citizens.

Structured and standardized data about public contracting can help stakeholders to:

* Deliver better value for money for governments,
* Create fairer competition and a level playing field for business,
* Drive higher-quality goods, works, and services for citizens,
* Prevent fraud and corruption,
* Promote smarter analysis and better solutions for public problems.

**1.3 Research questions**

How can the data model provided by OCDS be effectively used for the detection and prevention of corruption, collusion, or anti-competitive practices in public contracting processes?

This main question can be split into the following sub-questions:

1. What are the main indicators or red flags that suggest the presence of any irregularity in a public contracting process?
2. Which functional techniques of process mining can be used in the detection of which red flags in public procurement processes and what are their benefits?
3. In what characteristics do the conventional techniques of data mining used in the detection of fraud resemble the functional techniques of process mining?
4. What are the possibilities of integration of both approaches, process mining and data mining, in the detection of which red flags in public procurement processes?
5. How could a hybrid model be used to detect red flags in public procurement processes in accordance with the characteristics of the Open Contracting Data Standard?

**1.4 Thesis outline**

In Chapter 2, the preliminaries are presented. This includes a literature review on the state of the art in the definition and localization of indicators of corruption and fraud in public procurement. Then the state of the art of the functional techniques of process mining and data mining used in the detection of fraud is explored. And finally, the theoretical and practical efforts that have been used to combine both approaches are described. Next, Chapter 3 presents a hybrid method that combines the functional techniques of process mining and data mining in the detection of risks of corruption and fraud in public contracting that is coupled with OCDS specifications. Chapter 4 elaborates a use case. Just over 300,000 public procurement processes published under the OCDS are used to illustrate the implementation of the hybrid method. The results obtained are discussed in Chapter 5. Finally, this thesis is wrapped up in Chapter 6 with a reflection on the limitations of the method and some recommendations to further improve the OCDS, in addition to a section describing possible future work in this area.

**References**

OECD. (2014). *OECD Foreign Bribery Report: An Analysis of the Crime of Bribery of Foreign Public Officials*. OECD Publishing, Paris. <https://doi.org/10.1787/9789264226616-en>

OECD. (2012). *Progress Made in Implementing the OECD Recommendation on Enhancing Integrity in Public Procurement*. OECD Publishing, pp 14-15. <https://dx.doi.org/10.1787/9789264201385-8-en>

UN Office of Drugs and Crime. (2013). *Guidebook on anti-corruption in public procurement and the management of public finances*. Publishing and Library Section, United Nations Office at Vienna, p 1. Retrieved from: <https://www.unodc.org/documents/corruption/Publications/2013/Guidebook_on_anti-corruption_in_public_procurement_and_the_management_of_public_finances.pdf>

European Commission. (2014). Report from the Commission to the Council and the European Parliament EU Anti-Corruption Report. Retrieved from: <https://ec.europa.eu/home-affairs/sites/homeaffairs/files/e-library/documents/policies/organized-crime-and-human-trafficking/corruption/docs/acr_2014_en.pdf>