FGV EBAPE

RAP
BRAZILIAN JOURNAL OF PUBLIC ADMINISTRATION

Forum: Practical Perspectives

Information technologies for countering corruption: analysis of the Costa Rican public procurement system

Dagoberto José Herrera Murillo [1]
Gabriel Valerio-Ureña [2]
Gabriel Silva Atencio [1]
Jorge Asprón [2]
Alejandra Álvarez Alfaro [1]

[1] Universidad Latinoamericana de Ciencia y Tecnología / Escuela de Ingeniería Informática, San José – Costa Rica
[2] Tecnológico de Monterrey / Escuela de Humanidades y Educación, Monterrey – Mexico

Corruption has disastrous consequences for the well-being of societies. In particular, public procurement is a highly vulnerable factor due to the high value of its transactions. This research analyzes the role that information technologies can play in the fight against corruption. We developed a case based on open data on public procurement in Costa Rica to identify the main concepts associated with the underlying information system. The case shows the potential of technological tools for social network analysis to counter the complexity of corruption networks. It also shows the need for a rich ecosystem of anti-corruption data – preferably in an open format.
Keywords: corruption; transparency; public procurement; socio-technical system; network analysis.

Tecnologias da informação para a luta contra a corrupção: análise da contratação pública costarriquenha

A corrupção é um fenômeno com consequências desastrosas para o bem-estar econômico, social e político das nossas sociedades. Em particular, a contratação pública é um fator altamente vulnerável devido ao alto valor econômico de suas transações. A pesquisa analisa o papel que as tecnologias da informação podem desempenhar no combate à corrupção. Além disso, desenvolvemos um caso baseado em dados abertos sobre compras públicas na Costa Rica, a fim de identificar os principais conceitos associados ao sistema de informação subjacente. O caso mostra o potencial das ferramentas tecnológicas de análise de redes sociais para ir de encontro a complexidade das redes de corrupção. Mostra também a necessidade de um ecossistema diversificado e sustentável de dados anticorrupção – preferencialmente em formato aberto.
Palavras-chave: corrupção; transparência; procuração pública; sistema sócio técnico; análise de rede.

DOI: http://dx.doi.org/10.1590/0034-761220220126x
ISSN: 1982-3134

Article received on April 19, 2022 and accepted on September 30, 2022.
[Translated version] Note: All quotes in English translated by this article's translator.
Editor-in-chief:
Alketa Peci (Fundação Getulio Vargas, Rio de Janeiro / RJ – Brazil)
Associate editor:
Mauricio Dussauge Laguna (Centro de Investigación y Docencia Económicas, Ciudad de México – México)
Reviewers:
Felipe Fróes Couto (Universidade Estadual de Montes Claros, Montes Claros / MG – Brazil)
Oliver David Meza Canales (Centro de Investigación y Docencia Económicas, Ciudad de México – México)
One of the reviewers did not authorize the disclosure of their identity.
Peer review report: the peer review report is available at this URL.

Tecnologías de información para lucha anticorrupción: análisis de la contratación pública costarricense

La corrupción es un fenómeno con consecuencias desastrosas para el bienestar económico, social y político de nuestras sociedades. En particular, la contratación pública es vulnerable debido al elevado valor económico de sus transacciones. La presente investigación analiza el rol que las tecnologías de información pueden jugar en la lucha anticorrupción. Desarrollamos un caso basado en datos abiertos de contrataciones públicas en Costa Rica, con la finalidad de identificar los conceptos principales asociados al sistema de información subyacente. El caso muestra el potencial de las herramientas tecnológicas de análisis de redes sociales para contrarrestar la complejidad de las redes de corrupción. Asimismo, se evidencia la necesidad de disponer de un ecosistema rico en datos anticorrupción, preferiblemente en formato abierto.

Palabras clave: corrupción; transparencia; contratación pública; sistema sociotécnico; análisis de redes sociales.

1. INTRODUCTION: CORRUPTION AND GOVERNMENT

Understanding corruption begins with acknowledging that it takes many forms and behaviors: it can occur anywhere, it can involve anyone, it frequently occurs underhandedly with the collaboration of professional enablers, and it adapts to different contexts in response to factors such as legislative and technological changes.

When discussing corruption, it is difficult not to mention the situation of public procurement. Government procurement accounts for 12% of the Organization for Economic Cooperation and Development's gross domestic product and 29% of public spending (Organization for Economic Cooperation and Development, 2017). Such a significant volume of resources exposes general procurement processes to high risks of inefficiency, mismanagement, and corruption.

The following section of this article provides an outline of the socio-technical system that must be taken into account when incorporating new technology into an anti-corruption approach. At the same time, the case study section illustrates relevant technology concepts using data from public works awards in Costa Rica.

2. ANTI-CORRUPTION SOCIO-TECHNICAL SYSTEM

Technological artifacts are not isolated elements with a life of their own. Their adoption occurs in a context of use mediated by human factors -as or more relevant to the success of their adoption than the performance of the technology itself-. Budgetary limits, rigid financing schemes, low-risk tolerance, a lack of innovative processes, and in innovations due to political changes must all be considered when applying new technology to anti-corruption in the public sector (Gartner, 2019).

Understanding the socio-technical system of reference and the social structure that operates on a technological basis is essential for successfully implementing technology. A complicated interaction between societal infrastructure and human behavior occurs inside these systems. The roots of socio-technical design are found in general systems theory (Bertalanffy, 1968) and, as such, can be modeled in the form of layers or levels that allow us to understand their characteristics better. In socio-technical systems, four levels can be identified: mechanical, which refers to the physical and technological infrastructure (hardware); informational, which refers to logical applications (software); personal, which refers to the relationship with users and computational interfaces; and community, which refers the social structures created around the use of technology (Whitworth & Ahmad, 2013). The logic is cumulative, which means that each level incorporates the parts of the preceding level, and the final layer comprises all of the needs at the hardware, software, and human-computer interface levels.

Mechanical level. Hardware refers to the tangible physical support of a computer system containing electromechanical components. Any attempt to integrate information technology into anti-corruption operations assume that regardless of the procurement model, the persons in charge of its execution have sufficient computer equipment to store vast volumes of data and run analysis procedures and the results.

Informational level. This depict the components of the information system enable the execution of specific tasks, as well as their linkages. Data is the fuel of the coating. Once the data is available, the next step is identifying the appropriate analysis techniques to detect signs of corruption. The information layer must provide end-user interfaces for consuming, discovering, and interpreting data results via products such as query engines, reports, and visualization dashboards.

Personal level. In the context of anti-corruption systems, the unique factor will be marked by the sector of origin and the level of digital literacy of the user in question. Public officials, who ensure the strict application of the law for the prevention, detection, and control of corruption within institutions; journalists, who keep an eye on how the system is operating in light of public opinion; academics, who conduct scientific study on corruption issues and educate future decision-makers; citizens, who assess how the government conduct itself; and private representatives, who contribute by making financial contributions, are the most pertinent audiences for this purpose.

Social level. Social structures, roles, and community norms are at the apex of socio-technical ecosystems. Collective action is now acknowledged as more successful than the efforts of isolated individuals. If this layer is well articulated, she brings ingredients such as freedom, synergy, productivity, privacy, fairness, transparency, and conflict resolution (Whitworth & Ahmad, 2013). Anti-corruption systems are the ultimate manifestation of community anti-corruption organizations. These are formal collaborative structures that bring together all players in the anti-corruption ecosystem (institutions, media, academia, citizens, private corporations, and third-sector groups) to establish a coalition motivated by shared aims of transparency and integrity. Ecosystem members share information, influence, and resources. To succeed, these alliances must attract government leaders who have a strong commitment and sense of ownership (Byrne, Arnold, & Nagano, 2010).

3. THE STUDY OF CORRUPTION NETWORKS

Corruption, like other illegal acts, has social network dynamics (Diviák, Dijkstra, & Snijders, 2018). In corruption networks, multiple actors coordinate to plan, execute, and overlap an action. The foregoing highlights the necessity for analysis methodologies that can go beyond individual records and assist us in understanding the linkages between network players (Luna-Pla & Nicolás-Carlock, 2020). Social network analysis is a mechanism for understanding these complex structures.

Network science emerged as a discipline at the beginning of the 21st Century, drawing on the contributions of sociology and graph theory. Barabasi (2016) defines four main characteristics of this area of knowledge: an interdisciplinary approach to tackling complex problems; an empirical focus on reliable data analysis; a quantitative nature that employs formalisms from disciplines such as graph theory and statistical physics; and finally, substantial use of computer tools such as databases and algorithms. When applied to the investigation of social networks, this field allows for the realistic modeling of leadership, communication, and human organization phenomena.

Multiple researchers have used social network analysis to investigate corruption cases in the last decade. For instance, Srivastava and Singh (2018) analyzed the emblematic case of the Panama Papers. The Neo4j software was utilized in this study to represent the entities and connections involved in the case. This enabled the network to be seen and facilitated the application of relevant techniques to determine the importance of the various participants in the network. This method has also been used to examine public procurement systems in the European Union (Villedieu, 2016) and Brazil (Van Erven, Holland, & Carvalho, 2017).

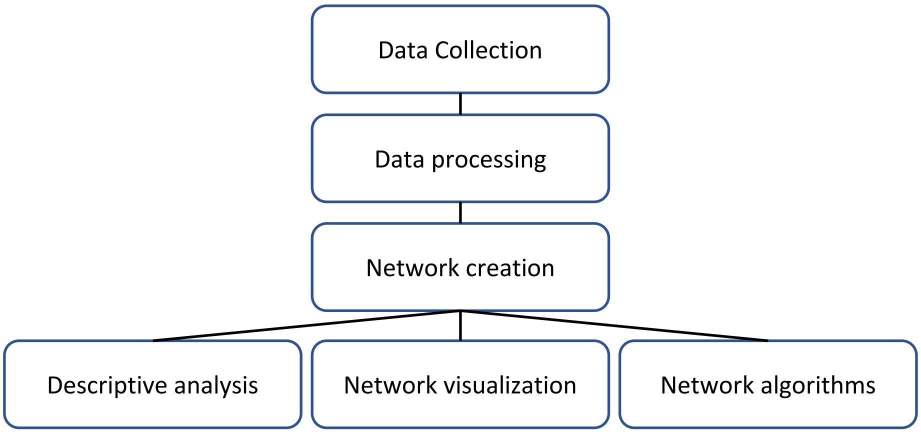
4. CASE STUDY, DATA, AND METHODOLOGY

Costa Rica is ranked 42nd out of 180 countries in the most current Transparency International's Corruption Perceptions Index (2020). This puts the country in a good position; it is third in Latin America, trailing only Uruguay and Chile and far ahead of the other Central American countries.

Owing to its journalistic, legal, and political ramifications, the Cochinilla case became a symbol of corruption in 2021. Some of the country's top construction corporations allegedly gave different incentives (money, vehicles, and sexual services) to public officials in exchange for preferential treatment in the distribution of public works, particularly highway building and maintenance of highways. The irregularities also involved inspectors, who endorsed the quality of defective materials. The fact that Costa Rica's paved roads are among the worst in Latin America adds to the significance of the event (World Economic Forum, 2019).

The Cochinilla case points explicitly to two construction companies, which we will refer to hereafter as suppliers A and B. The case's goal will be to demonstrate the power of social network analysis in diagnosing the operation of these two enterprises within the public works network. Figure 1 depicts the basic steps of the case development methodology: The first step is to acquire data from relevant and trustworthy sources; the second step is to analyze that data to extract nodes and relationships; and the creation of the network according to the specifications of a model. Finally, the formed network allows for descriptive analysis, visualization, and algorithm application to the network. The source code used to develop the case is presented in the supplemental material section.

FIGURE 1 **THE FLOW OF THE METHODOLOGY FOLLOWED IN THE CASE**

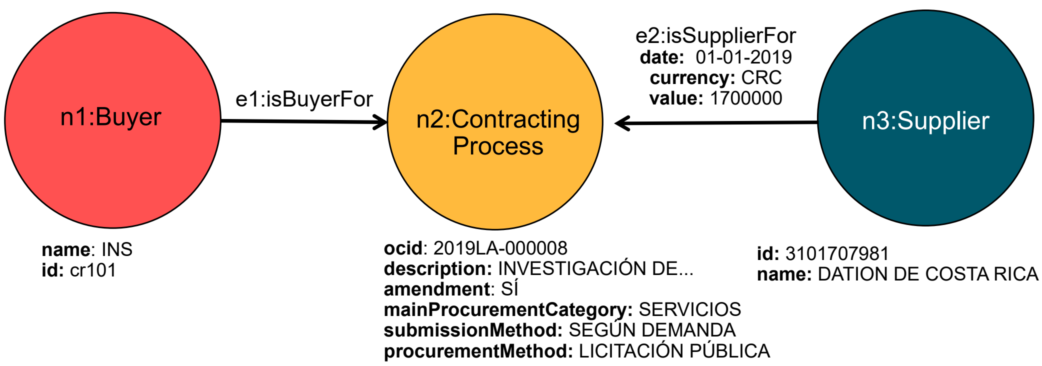


Source: Elaborated by the authors.

The dataset is from the Integrated Public Procurement System (Sistema Integrado de Compras Públicas [SICOP], 2022). SICOP is a digital platform that the central administration must use to process administrative procurement procedures. The original dataset contains103,617 procurement operations spanning ten years from 2011 to 2020.

The next step in the process was creating a data model capable of capturing the essential characteristics of the problem studied. The value of the produced discoveries (including queries, reports, and visualizations) is proportional to the quality and specifications of this initial model. The formal fields of the Open Contracting Partnership Data Standard (Open Contracting Partnership, 2020) and ontology for public procurement suggested by the European project TheyBuyForYou (Soylu et al., 2019) were used as a reference for the model's design, resulting in the graph shown in Figure 2.

FIGURE 2 **A NETWORK MODEL FOR PUBLIC PROCUREMENT NETWORK**



Source: Elaborated by the authors using the fields defined by the Open Contracting Partnership Data Standard (Open Contracting Partnership, 2020).

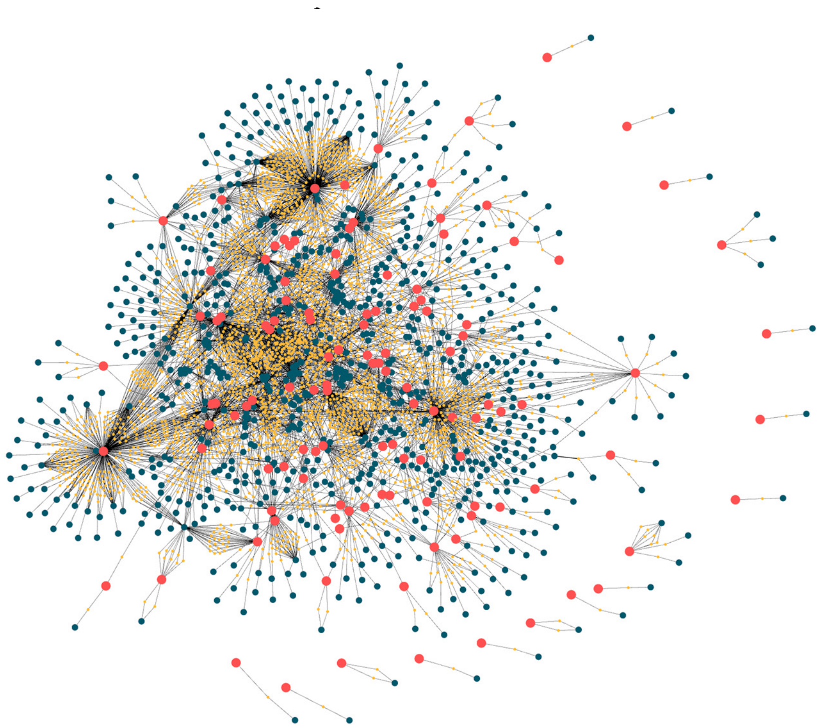
The node, which is represented by a circle and its labels, is the basic unit of the network. Arcs are directional lines connecting pairs of nodes. Nodes and arcs can have attributes that help to describe them. This specification is used to build a network model with three sorts of nodes: Institution, Contracting, Process, and Supplier. Arcs of type isBuyerFor are projected from the nodes representing the institutions toward the procedures that they processed.

Similarly, from the suppliers' nodes, arcs (identified with the supplier for the label) are directed toward the awarded procedures. Contracts from several providers might be produced from the same procurement procedure. Note that the names of the titles are in English as specified by the standard. The characteristics of the arcs of type are Supplier For containing information about the particular contracts. The generalities of the model were validated by a journalist specializing in corruption cases in public procurement.

Even though there are other graph database providers, the exact technology used for implementing the model was Neo4j, an open-source database. This platform has various features that make it an excellent choice for simulating a public procurement network. On the one hand, it can handle vast volumes of data efficiently; it features user-friendly interfaces and visualization engines that allow non-specialized audiences to edit the network to investigate dynamic relationships between nodes. On the other hand, it has libraries to implement network algorithms and artificial intelligence operations.

There are four types of contracting procedures available: products, services, goods and services, and public works. Public works, with which we collaborated in this case, account for 2.3% of total procedures yet use somewhat more than a fourth of total contracts. On average, a public works procedure is 32 times greater than one for services and nine times greater than one for goods. The above-mentioned data model was loaded into Neo4j, resulting in 2,996 public works contracting procedures processed by 105 institutions and contracts given to 786 (See Figure 3).

FIGURE 3 THE PUBLIC WORKS NETWORK



Note: Institutions (red), suppliers (blue), and contracting procedures (yellow).
Source: Elaborated by the authors.

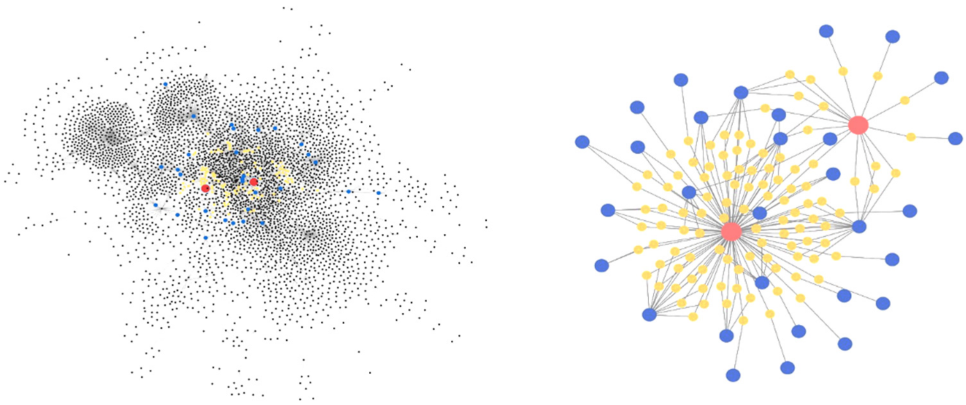
The first finding that jumps out is the dominance of a significant component that contains 98.3% of the nodes. The details are network segments in which it is feasible to travel from one node to any other node by ignoring the orientation of the arcs. Around them orbit some smaller components with few nodes. This structure implies that certain vendors supplied several institutions and that some contracting processes benefited multiple suppliers at the same time. Hence, there is the possibility of tracing direct or indirect links between most of the agents of the public works market in the country.

Suppliers A and B, who were singled out in the Cochinilla case, are active participants in this network. Supplier A was awarded 90 contracts, whereas Supplier B was awarded 16 (one in consortium with another company). In the case of the public works network, the 786 bidders received only one contract on average. Supplier A's 90 contracts are the maximum value reported, i.e., Supplier A is the largest recipient of contracts; Company B's contracts place it up to the 95th percentile. The aggregate of the A and B contracts for suppliers places them first and third as the major receivers of funds, respectively. These totals are 78 and 40 times the average total amount received by the other companies.

Another relevant factor is the level of competition according to the contracting. While the other companies received 44% of their contracts through a bidding process, suppliers A and B received 79% and 88%, respectively. In principle, tenders are conducive to awarding more significant works through a competition where the quality and price of the bids come into play. This is in sharp contrast to direct allocations, which are typically better justified in lower-value projects.

Figure 4 depicts the key network component separating suppliers A and B (red), the contractual procedures in which they participated (yellow), the institutions that processed those procedures (blue), and the rest of the network actors (nodes in black). Simultaneously, the part on the right focuses on the nodes in the environment of providers A and B, leaving out the other actors.

FIGURE 4 A AND B SUPPLIERS IN THE PUBLIC WORKS NETWORK



Source: Elaborated by the authors.

A brief examination of the network shows that suppliers A and B are at the core of the procurement network rather than in a remote place. However, beyond that visual interpretation, networks provide a wide range of algorithms to measure the level of importance attributable to each actor objectively. The results of three such algorithms applied to the organizations in question are shown in Table 1. The degree is the number of neighbors a node has. A high degree can be considered evidence of status. Intermediation refers to nodes that act as bridges between different parts of the network. Higher betweenness suggests that a node controls the flow of information and resources in the network. The distance between a node and all other nodes in the network is measured by closeness. Nodes with greater closeness can disseminate information or resources within the web faster. Typically studies on criminal networks characterize those participating in terms of these indicators. Increased visibility implies greater visibility, whereas greater intermediation and closeness indicate nodes that interact and communicate with the network. Leaders of criminal networks often have a low degree of discretion and compensate for this by having high values of the other variables (Diviák et al., 2018).

TABLE 1 MEASURES OF THE CENTRALITY OF THE PUBLIC WORKS CONTRACTING NETWORK

	Grade	Intermediation	Proximity
Provider A	90	1,412,249	0.22
Provider B	16	22,359	0.15
Average (n = 786)	4	12,451	0.14
Median (n = 786)	1	0	0.14

Source: Elaborated by the authors.

According to the findings, providers A and B have high scores for all three parameters. Provider A is the network node that comes first for all of them. As a result, they are visible nodes in all dimensions of centrality. In practice, this would indicate that they are actors whose behavior should be closely observed.

5. DISCUSSION AND CONCLUSIONS

Without being thorough in its implications and implementation, the case developed presents a generic structure to aid in the investigation of indicators of corruption in public procurement utilizing information technology. The advantages of displaying information in the form of a graph are apparent. This database offers functional query interfaces for technical and non-technical users. Adoption of these technologies in the field will be dependent on sufficiently resourcing the solution at the mechanical and informational levels, understanding the demands of various audiences at the personal level and translating their results into tangible actions at the community level.

The results also show the importance of having data available so that the fight against corruption is based on evidence and not on mere speculation. Even if it were the best information technology platform, it is difficult to anticipate excellent performance without relevant, fresh, abundant, and high-quality data. Moreover, open data formats ensure that anyone can access, use, reuse, share, and compare your findings. The example given is limited to a single source of information; however, additional data sources are required to populate the graph more effectively, which could not be located in the Costa Rican situation, at least in an open format.

In terms of future study, it would be desirable to experiment with more complex implementations of a knowledge graph that truly capture the diversity of information sources. Possibly other Latin American countries better positioned in the rankings of open data availability, such as Brazil, Mexico, or Colombia, offer favorable conditions for this. It would also be beneficial to assess socio-technical implementation from the standpoint of the difficulties encountered by the public policy designer in articulating a solution at the personal and community levels. Among these problems is the necessity for regulatory foundations for the use of anti-corruption technology, as well as the design of mechanisms for interaction between specialists and non-specialized audiences.

REFERENCES

Barabasi, A. L. (2016). *Network Science*. Cambridge, UK: Cambridge University Press.

Bertalanffy, L. (1968). *General System Theory: Foundations, Development*. New York, NY: George Braziller.

Byrne, E., Arnold, A. K., & Nagano, F. (2010). *Building Public Support for Anti-Corruption Efforts: Why Anti-Corruption Agencies Need to Communicate and How*. Washington, DC: UNODC, CommGAP, and The World Bank. Retrieved from https://www.unodc.org/documents/corruption/Publications/StAR/CorruptionWhitePaperpub31110screen.pdf

Diviák, T., Dijkstra, J. K. & Snijders, T. A. B. (2018). Structure, multiplexity, and centrality in a corruption network: the Czech Rath affair. *Trends in Organized Crime*, *22*, 274-297. Retrieved from https://doi.org/10.1007/s12117-018-9334-y

Gartner. (2019). *2020 CIO Agenda: Resilience During Disruption*. Retrieved from https://www.gartner.com/en/publications/2020-cio-agenda

Luna-Pla, I., & Nicolás-Carlock, J. R. (2020). Corruption and complexity: a scientific framework for the analysis of corruption networks. *Applied Network Science*, *5*(13), 1-18. Retrieved from https://doi.org/10.1007/s41109-020-00258-2

Open Contracting Partnership. (2020). *Open Contracting Data Standard*. Retrieved from https://standard.open-contracting.org/latest/en/

Organization for Economic Co-operation and Development. (2017). *Government at a Glance - 2017 Edition*. Retrieved from https://stats.oecd.org/Index.aspx?DataSetCode=GOV_2017

Sistema Integrado de Compras Públicas. (2022). *Datos Abiertos Adjudicaciones*. Retrieved from https://www.sicop.go.cr/index.jsp

Soylu, A., Elvesæter, B., Turk, P., Roman, D., Corcho, O., Simperi, E., … Lech, T. C. (2019). Towards an Ontology for Public Procurement Based on the Open Contracting Data Standard. In I. O. Pappas, P. Mikalef, Y. K. Dwivedi, L. Jaccheri, J. Krogstie, & M. Mäntymäki (Eds.), *Digital Transformation for a Sustainable Society in the 21st Century*. (Lecture Notes in Computer Science, Vol. 11701, pp. 230-237). Cham, Switzerland: Springer. Retrieved from https://doi.org/10.1007/978-3-030-29374-1_19

Srivastava, S., & Singh, A. K. (2018). Graph-Based Analysis of Panama Papers. In *Proceedings of the 5º International Conference on Parallel, Distributed and Grid Computing*, Himachal Pradesh, India. Retrieved from https://doi.org/10.1109/PDGC.2018.8745785

Transparencia Internacional. (2020). *Corruption Perception Index*. Retrieved from https://www.transparency.org/en/cpi/2020/index/nzl

Van Erven, G. C. G., Holanda, M., & Carvalho, R. N. (2017). Detecting Evidence of Fraud in the Brazilian Government Using Graph Databases. In A. Rocha, A. Correia, H. Adeli, L. Reis, & S. Costanzo (Eds.), *Recent Advances in Information Systems and Technologies* (Advances in Intelligent Systems and Computing, Vol. 570, pp. 464-473). Cham, Switzerland: Springer. Retrieved from https://doi.org/10.1007/978-3-319-56538-5_47

Villedieu, J. (2016, January 04). Exploring €1.3 trillion in public contracts with graph visualization. *Linkurious*. Retrieved from https://linkurious.com/blog/exploring-e1-3-trillion-in-public-contracts-with-graph-visualization/

Whitworth, B., & Ahmad, A. (2013). Socio-Technical System Design. In Interaction Design Foundation (Ed.), *The Encyclopedia of Human-Computer Interaction* (Chap. 24, 2a ed.). Retrieved from https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/socio-technical-system-design

World Economic Forum. (2019). *Quality of roads*. Retrieved from https://reports.weforum.org/pdf/gci-2017-2018-scorecard/WEF_GCI_2017_2018_Scorecard_EOSQ057.pdf

SUPPLEMENTARY MATERIAL

The source code used to develop the case can be downloaded at the following repository: https://github.com/gvaleriou/tecnologias-de-informacion-para-lucha-anticorrupcion

Dagoberto José Herrera Murillo

https://orcid.org/0000-0002-6331-5929
Master in Big Data Management & Analytics; Professor at the School of Informatics of the Universidad Latinoamericana de Ciencia y Tecnología. E-mail: dherreram752@ulacit.ed.cr

Gabriel Valerio-Ureña

https://orcid.org/0000-0002-4446-6801
Ph.D. in Educational Innovation; Director of the Doctorate in Educational Innovation at the Tecnológico de Monterrey. E-mail: gvalerio@tec.mx

Gabriel Silvia Atencio

https://orcid.org/0000-0002-4881-181X
Ph.D. in Business Administration; Professor at the School of Computer Science, Universidad Latinoamericana de Ciencia y Tecnología. E-mail: gsilvaa468@ulacit.ed.cr

Jorge Asprón

https://orcid.org/0000-0001-9332-0553
Ph.D. candidate in Humanistic Studies at the School of Humanities and Education of the Instituto Tecnológico y de Estudios Superiores de Monterrey. E-mail: a00884529@tec.mx

Alejandra Álvarez Alfaro

https://orcid.org/0000-0003-3156-2084
Degree in Computer Science with an emphasis in Technology Resources Management from the Universidad Latinoamericana de Ciencia y Tecnología (Latin American University of Science and Technology).
E-mail: aalvareza773@ulacit.ed.cr

AUTHOR'S CONTRIBUTION

Dagoberto José Herrera Murillo: Conceptualization (Lead); Data curation (Lead); Formal Analysis (Lead); Software (Lead); Visualization (Lead); Writing - original draft (Lead); Writing - review & editing (Equal).

Gabriel Valerio-Ureña: Conceptualization (Supporting); Funding acquisition (Lead); Writing - original draft (Supporting); Writing - review & editing (Equal).

Gabriel Silva Atencio: Conceptualization (Supporting); Funding acquisition (Supporting); Visualization (Supporting); Writing - original draft (Supporting); Writing - review & editing (Equal).

Jorge Asprón: Writing - original draft (Supporting); Writing - review & editing (Equal).

Alejandra Álvarez Alfaro: Data curation (Supporting); Formal Analysis (Supporting); Visualization (Supporting).