

Modelo para definir índices de corrupción en convocatorias de contratación estatal en Colombia basada en Big Data y Procesamiento del Lenguaje Natural



Examen de Valoración

Ing. Julio César Luna Ortega

Director: PhD. Carlos Alberto Cobos Lozada
Codirectora: PhD. Martha Eliana Mendoza Becerra

Universidad del Cauca
Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Sistemas
Grupo de I+D en tecnologías de la información (GTI)
Área de interés en Gestión de la Información
Popayán, Septiembre de 2020

TABLA DE CONTENIDO

1	INTRODUCCIÓN.....	3
2	ESTADO Y AVANCE DEL TRABAJO	3
3	RESULTADOS ESPERADOS EN LAS SIGUIENTES ETAPAS.....	4
4	CONCLUSIONES PRELIMINARES	5
5	REFERENCIAS	5

LISTADO DE TABLAS

Tabla 1	Cronograma de actividades.....	3
----------------	--------------------------------	---

[illegible]

será llevado a cabo mediante 3 ciclos del PII, pero en la fase de desarrollo se usará SCRUM como metodología.

- Se espera adelantar la estancia de investigación en la Universidad de Granada (España), para lo cual se realizarán los trámites de consecución de apoyos económicos para viajes al exterior.
- Continuamente durante el tiempo de ejecución del proyecto se llevará a cabo la escritura de la monografía y finalmente un artículo, que plasme la investigación final. Se espera que este artículo sea publicado o quede en turno de publicación en una revista nacional o internacional indexada mínimo en categoría C según el PUBLINDEX de COLCIENCIAS.

4 CONCLUSIONES PRELIMINARES

Con respecto al trabajo desarrollado hasta el momento se puede concluir que:

- Se ha llevado a cabo la revisión del estado del arte relacionada con el tema de investigación, dando por cumplido el primer objetivo específico de la tesis. Este trabajo se realizó siguiente un proceso que es repetible.
- Se han elaborado dos artículos como resultados de la revisión del estado del arte, el primero, un mapeo sistemático que se encuentra en proceso de evaluación en la revista Guillermo de Ockham. Y el segundo, una revisión sistemática de la literatura aceptado para su presentación en el evento de las JIISIC 2020 – XV Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento, y para su publicación en la Revista Investigación e Innovación en Ingenierías de la Universidad Simón Bolívar. Ambas revistas son categoría B hasta diciembre de 2020 según el PUBLINDEX de MINCIENCIAS.
- Los resultados obtenidos en la revisión sistemática demuestran que el uso de técnicas estadísticas para solventar el problema de la identificación temprana de la corrupción es una idea que en los últimos años ha tenido una buena acogida, así como la combinación de estas técnicas con tecnologías de la información. Lo anterior sumado al incremento en los casos de corrupción global reportados, llevan a la necesidad de seguir en el desarrollo de modelos que permitan solventar de manera más eficaz y eficiente el problema de la detección de la corrupción en procesos de contratación. Los estudios muestran que, dada la complejidad del problema, se hace necesario contar con un conjunto (batería) de herramientas especializadas para atacar los diferentes delitos que ocurren en contratación.

5 REFERENCIAS

- [1] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, "Systematic literature reviews in software engineering--a systematic literature review," *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, 2009, doi: <https://doi.org/10.1016/j.infsof.2008.09.009>.
- [2] K. S. Pratt, "Design Patterns for Research Methods: Iterative Field Research," in *AAAI Spring Symposium: Experimental Design for Real-World Systems*, 2009.

Identificación y detección de corrupción en contratación pública: Un mapeo sistemático

Julio-César Luna-Ortega, Carlos-Alberto Cobos-Lozada, Martha-Eliana Mendoza-Becerra

jluna@unicauca.edu.co, ccobos@unicauca.edu.co, mmendoza@unicauca.edu.co

Grupo de I+D en Tecnologías de la Información (GTI), Universidad del Cauca, Popayán, Cauca, Colombia

DOI: 10.17013/risti.n.pi-pf

Resumen. La corrupción como un fenómeno propio de la contratación estatal es uno de los problemas económicos y sociales con mayor trascendencia en la sociedad actual. Buscando combatir esta problemática, algunas entidades aplican estrategias anticorrupción, pero desafortunadamente, la detección manual de la corrupción es intrínsecamente difícil, es por eso por lo que la idea de aplicar modelos automáticos para resolver este problema surge como una alternativa prometedora. El objetivo del presente mapeo sistemático es recopilar el conocimiento actual en relación con modelos que apliquen estrategias anticorrupción, con el propósito de determinar cuáles son las características más importantes a la hora de desarrollar un modelo automático anticorrupción. Los resultados del mapeo muestran que son pocos los estudios en donde se hace uso de técnicas de tecnologías de la información, sin embargo, los resultados de estos estudios han sido muy relevantes. Por otra parte, existe la necesidad de establecer unos índices de efectividad de los modelos anticorrupción, así como, seguir en el desarrollo de modelos automáticos que permitan la detección de los diferentes delitos de la corrupción en la contratación.

Palabras clave: Corrupción, Contratación, Modelo, Mapeo sistemático.

Identification and detection of corruption in public procurement: A systematic mapping

Abstract: Corruption as a phenomenon of state contracting is one of the most important economic and social problems in today's society. Seeking to combat this problem, some entities apply anti-corruption strategies, but unfortunately, the manual detection of corruption is intrinsically difficult, which is why the idea of applying automatic models to solve this problem arises as a promising alternative. The objective of this systematic mapping is to gather current knowledge in relation to models that apply anti-corruption strategies, with the purpose of determining which are the most important characteristics when developing an automatic anti-corruption model. The mapping results show that there are few studies where information technology techniques are used, however, the results of these studies have been very relevant. On the other hand, there is a need to establish rates of effectiveness of anti-corruption models, as well as to continue in the development of automatic models that allow the detection of the different crimes of corruption in contracting.

Keywords: Corruption, Hiring, Model, Systematic Mapping.

1. Introducción

La corrupción puede definirse como una práctica deshonesta llevada a cabo por individuos, que abusan del poder público para su beneficio privado. A nivel mundial, el promedio de pérdida sufrida por las entidades víctimas de corrupción es aproximadamente de \$250,000 USD por caso [1]. Para Transparencia Internacional, el problema de la corrupción sigue siendo alarmante y considera que la

lentitud de la Justicia para castigar a los corruptos y las repetitivas prácticas de abuso del poder en todos los niveles del sector público se reflejan en las bajas calificaciones que obtienen algunos países en el Índice de Percepción de Corrupción (IPC) [2].

Un país, como un estado social de derecho debe primar por la satisfacción del interés general, así pues, el Estado debe contar con una serie de herramientas en instituciones a su servicio para cumplir con dicha finalidad. Una de ellas es la Contratación Estatal, en donde el Estado por medio de una serie de principios y reglas que orientan el proceso, busca la satisfacción de dicho interés. La corrupción como un fenómeno propio de la contratación es uno de los problemas sociales con mayor trascendencia en la sociedad actual. Cuando ocurre un caso de corrupción, el proceso de contratación deja de estar dirigido a la satisfacción del interés general y se enmarca en la satisfacción de intereses personales o de pequeños grupos de agentes del proceso contractual, lo que afecta el bienestar de la comunidad.

Buscando combatir esta problemática algunas entidades aplican estrategias anticorrupción, pero desafortunadamente, la detección manual de la corrupción es intrínsecamente difícil dado que, según [3] el crimen se oculta por su propia naturaleza, por esta misma razón, la aplicación de modelos estadísticos y/o tecnológicos para resolver este problema, es una idea que parece prometedora.

Este estudio se busca identificar las estrategias prácticas en contra de la corrupción (anticorrupción) que se han utilizado y reportado en la literatura. Se espera que el presente mapeo sistemático sea útil al sintetizar el estado actual del conocimiento en modelos prácticos anticorrupción, que además facilite la posterior definición de un modelo para identificar los posibles delitos que se están presentando en un proceso de contratación y los esquemas de solución a estos, entre otros.

El resto de este documento está organizado de la siguiente manera: en la Sección 2 se describe el protocolo de investigación llevado a cabo para realizar el mapeo sistemático. La Sección 3 muestra los resultados obtenidos de dicho mapeo. En la Sección 4 se discuten las observaciones principales de los resultados, así como las limitaciones e implicaciones del mapeo en el campo de investigación. Y, finalmente, en la Sección 5 se presentan las conclusiones y el trabajo futuro que espera realizar el grupo de investigación en el área.

2. Protocolo de Investigación

El presente mapeo sistemático se ha llevado a cabo siguiendo las guías provistas en los trabajos presentados en [4, 5], realizando tres etapas principales: Planificación, Ejecución y Discusión. Las actividades relativas a las dos primeras etapas se describen en las siguientes subsecciones y la etapa de discusión se presenta en la Sección 3 con los resultados por pregunta de investigación y la Sección 4.

Etapas de Planificación

En esta etapa se llevaron a cabo las siguientes actividades:

1. Definición de la pregunta de investigación.
2. Definición de la estrategia de búsqueda.
3. Establecimiento de los criterios de selección de los estudios primarios.
4. Establecimiento de los criterios de evaluación de calidad.
5. Definición de la estrategia de extracción de datos.
6. Selección de los métodos de síntesis.

2.1.1 Preguntas de Investigación

Esta investigación tiene como objetivo principal identificar las técnicas, métodos, modelos o similar, utilizadas para la identificación de riesgos de corrupción en la contratación pública. Para ello se han establecido las preguntas de investigación que se presentan en la **Tabla 2**.

A través de estas preguntas se recopila y categoriza la información existente acerca de la detección de corrupción en contratación, con el fin de realizar un mapeo general del tema, identificando las carencias existentes y con esto, en un próximo paso plantear una posible solución a la problemática expuesta.

Tabla 2. Preguntas de investigación

Preguntas de Investigación	Motivación
Q1. ¿Qué estudios existen sobre la detección o identificación de corrupción en contratación?	Determinar el número de publicaciones actuales y la tendencia a lo largo de los últimos años en relación con el campo de la identificación de corrupción en contratación.
Q2. ¿Qué tipos de soluciones se han implementado para la identificación de corrupción?	Determinar cuáles son las técnicas, métodos, modelos o similares que se han utilizado para la detección de corrupción en contratación y como son los informes que estas realizan.
Q3. ¿Qué tipos de corrupción en contratación se han trabajado?	Determinar qué tipos de corrupción o fraude son normalmente a los que se presta más atención en el proceso de contratación.
Q4. ¿En qué dominios de aplicación se ha tratado el problema de la identificación de la corrupción?	Determinar cuáles son los objetos (contratos, negocios, etc.) en donde se ha trabajado sobre el problema de la identificación de corrupción.
Q5. ¿Qué indicadores se utilizan en las mediciones de corrupción o fraude?	Determinar cuáles son los indicadores que se suelen utilizar a la hora de evaluar un contrato y cuál es el dominio de valores que estos pueden tomar para alertar un posible caso de corrupción.
Q6. ¿Qué problemas y desafíos identifican los autores en relación con la corrupción?	Determinar cuáles son los problemas y retos identificados por los autores que trabajan en el proceso de identificación de la corrupción en la contratación.

Para realizar la búsqueda de información se utilizaron las bases de datos Scopus, ACM Digital library, ScienceDirect e IEEE Xplore en las que se introdujo las cadenas de búsqueda. La **Tabla 3** muestra la cadena de búsqueda final, la cual se obtuvo de un proceso iterativo de mejora con el análisis de los resultados reportados por cada versión de esta. La metodología utilizada para la construcción de la cadena se basó en [6], estructurando los términos de población, intervención, comparación y resultado. La cadena se presenta dividida por conceptos, para su mayor comprensión. Se aplican operadores booleanos OR para unir los términos y sinónimos en cada una de las partes, y AND para unir las partes entre sí.

La investigación se realizó aplicando la cadena de búsqueda sobre el título, el resumen (abstract) y las palabras claves (keywords) de cada artículo.

Tabla 3: Cadena de búsqueda

Concepto	Estructura	Términos Alternativos y Sinónimos	Descripción
Técnicas	Resultado	("model" OR "algorithm" OR "framework" OR "technology" OR "method" OR "technique") AND	Se delimita lo que se está buscando: algoritmos, técnicas, modelos y similares.
Corrupción	Población	("*corruption" OR "collusion" OR "fraud") AND	Se define el problema a nivel general: corrupción y similares.
Tipos de delitos	Comparación	("bribery" OR "buys influence" OR "conflicts of interest" OR "cronyism" OR "embezzlement" OR "extort*" OR "favoritism" OR "gift-giving" OR "graft" OR "influence peddling" OR "kickback" OR "nepotism" OR "parochialism" OR "patronage" OR	Se delimitan los tipos específicos de delitos de corrupción.

		"political interference" OR "rigging" OR "splitting" OR "theft") AND	
Objeto	Población	("governm*" OR "agreement" OR "contract" OR "commitment" OR "arrangement" OR "settlement" OR "understanding" OR "bond" OR "deal" OR "bargain") AND	Se menciona el objeto que tiene el problema: contratación gubernamental y palabras relacionadas con contrato.
Acciones	Intervención	("detect" OR "avoid" OR "prevent" OR "warn" OR "caution" OR "identify" OR "understanding" OR "avert" OR "stop" OR "alert" OR "perceive" OR "distinguish" OR "notify" OR "inform" OR "recognize" OR "establish")	Se relaciona con el tipo de soluciones que buscan identificar, evitar, prevenir (o similares) la corrupción.

*Los asteriscos significan que cualquier carácter o caracteres pueden ser incluidos en la palabra, de esta manera se consiguen permutaciones en los términos de búsqueda en los que se utiliza.

2.1.3 Criterios de Selección de los Estudios Primarios

Con el objetivo de determinar la inclusión a la lista de estudios primarios de los resultados entregados por los motores de búsqueda, se realizó y aplicó una lista de criterios de aceptación, que los documentos (basado en título, resumen y palabras clave) debieron cumplir. Los criterios fueron los siguientes:

- Estudios escritos en el idioma inglés.
- Estudios completos publicados en revistas, conferencias, congresos o talleres de prestigio con revisión por pares.
- El estudio propone un modelo, técnica, algoritmo, framework o similar.

Luego, se eliminaron los documentos que cumplieron con alguno de los siguientes criterios de exclusión:

- El artículo es un estudio duplicado.
- El artículo no presenta un diseño de investigación bien definido.
- El artículo no presenta una pregunta de investigación bien definida.
- El artículo corresponde a un estudio de debate, o solo está disponible en forma de resumen o presentación.
- El artículo es un estudio secundario, es decir, corresponde a un reporte de resultados de otras investigaciones.
- El artículo trata el tema de la corrupción de manera superficial, sin hacer énfasis en el tema como tal o en un delito en específico.
- La contribución principal del artículo no se relaciona con alguna modelo/técnica/método/algoritmo de identificación de corrupción
- El artículo no propone una modelo/técnica/método/algoritmo, o que lo hagan, pero la propuesta no pueda ser implementada con ayuda del computador.
- El artículo presenta un modelo de índole psicosocial.

Luego de aplicar los criterios de selección, sobre todo el documento, en aquellos estudios identificados como relevantes, se hizo uso de la técnica de snowballing sobre estos, con el objetivo de ampliar el

horizonte de búsqueda, Además, se incluyeron algunos estudios que fueron identificados como relevantes en la etapa de literatura gris, documentos a los que también se les aplicaron los criterios de selección.

2.1.4 Criterios de Evaluación de Calidad

Además de los criterios generales de selección, el mecanismo de evaluación de la calidad, generalmente aplicado en revisiones sistemáticas [7], fue aplicado en este trabajo para evaluar la confiabilidad de los estudios primarios. Esta evaluación es necesaria para limitar el sesgo en la realización de este estudio empírico, para comprender mejor las posibles comparaciones y para guiar la interpretación de los hallazgos. Para esto se desarrolló una prueba con un sistema de puntuación de tres valores (-1, 0, 1). La prueba está conformada por las siguientes preguntas:

1. El estudio presenta una descripción detallada sobre los tipos de corrupción o fraude que se presentan en un objeto de población (ejemplo contratos).
2. El estudio contiene un modelo/técnica/método/algoritmo/similar aplicada.
3. El estudio tiene como finalidad aplicar acciones (identificar, detectar, prevenir, etc.) sobre un problema de corrupción.
4. El estudio expone de manera clara y detallada los resultados obtenidos tras aplicar un modelo/técnica/método/algoritmo/similar sobre un problema de corrupción en un objeto de población.
5. A qué cuartil pertenece la revista, conferencia o congreso en donde fue publicado el estudio.
6. El estudio ha obtenido dos o más citas por año desde su publicación (según índice de citas de Google).

La **Tabla 4** presenta la puntuación que se le asignó a cada documento de acuerdo con las respuestas obtenidas. La puntuación total otorgada por la prueba es la suma de cada uno de los criterios, con ello, los documentos evaluados obtuvieron una puntuación que estaba en el dominio [-6, 6], estas puntuaciones no fueron utilizadas como criterio de exclusión, solo con el objetivo de brindar una calificación de la calidad de los documentos resultantes del mapeo.

Tabla 4: Puntuación criterios de evaluación

No	Puntuación		
	-1	0	1
C1	No	Parcialmente	Si
C2	No	Parcialmente	Si
C3	No	Parcialmente	Si
C4	No	Parcialmente	Si
C5	Q4 O NO	Q3	Q2 O Q1
C6	Promedio de citas menor a 1 por año	Por lo menos 1 cita por año	2 o más citas por año

2.1.5 Extracción de Datos

Para la clasificación de los documentos se utilizaron los datos extraídos de los documentos junto con las respuestas encontradas en estos a las preguntas de investigación fijadas en la definición del protocolo de investigación. Las preguntas de investigación, como sus posibles respuestas se muestran en la

Tabla 5.

Tabla 5: Esquema de clasificación

Preguntas	Respuestas	
Q1. ¿Qué estudios existen sobre la detección o identificación de corrupción en contratación?	a. Encuesta b. Estudio de caso c. Mapeo sistemático d. Propuesta	
Q2. ¿Qué tipos de solución se han implementado para la identificación de corrupción?	a. Big Data b. Machine learning c. Marco de gobernanza d. Métodos de ciencia de redes e. Métodos de cognición	f. Random forests g. Redes neuronales h. Regresión i. Screen j. Sistemas de gestión k. Técnicas estadísticas l. Teoría de conjuntos difusos
Q3. ¿Qué tipos de corrupción en contratación se han trabajado?	a. Abuse of power b. Bid rigging c. Bribery d. Cartels e. Collusion f. Conflicts of interest	g. Embezzlement h. Extortion i. Fraud j. Money laundering k. Múltiples delitos l. Trading in influence
Q4. ¿En qué dominios de aplicación se ha tratado el problema de la identificación de la corrupción?	a. Contratos b. Corrupción pública c. Licitaciones	
Q5. ¿Qué indicadores se utilizan en las mediciones de corrupción o fraude en contratos?	a. Banderas rojas b. Binarios c. Diferencia entre ofertas corruptas y competitivas d. Índice de probabilidad	e. N/A f. Niveles de correlación
Q6. ¿Qué problemas y desafíos se identifican en los autores con relación hacia la corrupción?	a. Debe realizarse de forma discreta, ya que, si algún cartel se entera del proceso, podría destruir evidencia. b. Empresas corruptas pueden aplicar acciones que ayuden a evitar la identificación de corrupción. c. Falta de compromiso de las entidades. d. Falta de información. e. La corrupción en los proyectos es un problema complejo para el que no hay una fácil o única solución. f. La corrupción puede tomar muchas formas. g. No tener una definición clara del objeto a abordar.	

2.1.6 Métodos de Síntesis

En el presente mapeo se aplicó una síntesis de datos cuantitativa basada en:

- Establecimiento y representación a través de tablas y/o gráficos del número y/o porcentaje de los estudios seleccionados clasificados según sus posibles respuestas en cada una de las preguntas de investigación.
- Representación a través de tablas y/o gráficos de los estudios seleccionados clasificados según los resultados de las evaluaciones de calidad realizadas.

2.1.7 Calendario del Mapeo Sistemático

El presente mapeo sistemático inició en septiembre de 2019 y finalizó en noviembre de mismo año.

2.2 Etapa de Ejecución

Una vez aplicado el protocolo de revisión de literatura mencionado en la etapa anterior, se procedió con la etapa de ejecución, esta misma fue guiada mediante las siguientes fases principales:

1. En la primera fase se aplicó la cadena de búsqueda establecida en la **Tabla 3** sobre las bases de datos científicas Scopus, ACM Digital Library, ScienceDirect e IEEE Xplore, los resultados obtenidos se muestran en la **Tabla 6**.

Tabla 6: Resultados cadena de búsqueda

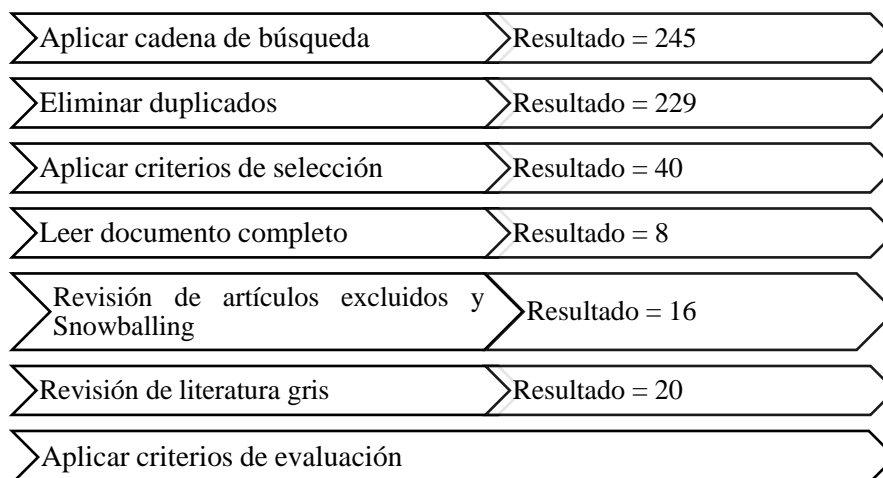
Base de datos	Número de documentos
Scopus	212
ACM Digital Library	12
ScienceDirect	5
IEEE Xplore	16
Subtotal	245
Duplicados (Auto detectados)	13
Duplicados (manualmente)	3
Total	229

2. Una vez concluida la primera fase se obtuvieron 229 documentos diferentes, con los que se procedió a aplicar los criterios de selección sobre el abstract de cada uno de ellos, una vez finalizada esta segunda fase, se obtuvo un total de 40 estudios relevantes.
3. Durante la tercera fase, se llevó a cabo la aplicación de los criterios de selección sobre los 40 estudios relevantes, pero en este caso sobre el documento completo. Tras esto, se obtuvo un total de 8 estudios primarios.
4. En la cuarta fase se realizó una revisión de los artículos excluidos, así mismo se aplicó las estrategias de snowballing hacia adelante y hacia atrás, con ello se obtuvo un total de 16 estudios primarios.
5. En la quinta fase se incluyeron artículos de literatura gris, a los que también se les aplicó los criterios de selección, obteniendo así 4 nuevos estudios, a los cuales se les aplicó estrategias de snowballing hacia adelante y hacia atrás sin obtener nuevos estudios primarios.
6. Finalmente, en la última fase se llevó a cabo la evaluación de calidad de cada uno los 20 estudios primarios resultantes de la revisión. La **Tabla 7** presentan los resultados de la evaluación de los estudios de acuerdo con los criterios de evaluación.

Tabla 7: Resultados de la evaluación de los estudios primarios

Criterio	Referencia/Año																			
	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]	[26]	[27]
	1993	2001	2003	2012	2012	2018	2019	2019	2016	2018	2019	2012	2013	2015	2017	2018	2017	2018	2020	2009
C1	1	1	1	1	1	1	1	1	0	0	1	0	1	-1	1	1	1	0	1	-1
C2	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	0	0
C3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
C4	1	1	1	1	1	1	1	1	1	1	-1	1	1	1	1	-1	-1	1	-1	-1
C5	1	1	1	1	1	1	1	1	1	1	1	-1	-1	1	0	1	-1	-1	-1	-1
C6	1	1	1	1	1	1	1	1	1	1	1	1	0	1	-1	1	1	-1	1	1
Puntaje	6	6	6	6	6	6	6	6	5	5	5	3	3	3	3	3	1	1	1	-1

La **Gráfica 1** muestra en resumen el proceso realizado en la etapa de ejecución.



Gráfica 1: Etapa de ejecución

3. Resultados

A continuación, se muestran los resultados obtenidos en cada una de las preguntas de investigación, definidas en la sección del protocolo de investigación, sobre los estudios relevantes los cuales fueron debidamente referenciados con el fin de posibilitar el estudio posterior en profundidad por parte del lector.

3.1 Pregunta Q1. Estudios existentes sobre detección o identificación de corrupción en contratación

El 95% de los estudios, es decir 19 de los 20 estudios seleccionados son propuestas donde los autores pretenden poner a prueba un nuevo modelo, bien sea para detectar o prevenir un delito de corrupción, de estos estudios se destaca [11] en donde, además, se realiza un mapeo sistemático sobre los métodos utilizados para la detección de colusión. Por otra parte en [24] llevan a cabo un estudio de caso en donde aplican unas técnicas de detección de fraude en un ámbito en particular.

3.2 Pregunta Q2. Tipos de solución implementadas para la identificación de corrupción

Solamente el 15% de los estudios (3) aplican una solución que no se relaciona con tecnología de la información o técnicas estadísticas ([23, 24, 27]), en estos estudios se despliegan una serie de estrategias para prevenir o detectar la corrupción. Por otra parte en 11 estudios (55%) se aplican técnicas estadísticas tales como regresión ([11, 16, 20]), pantallas ([13, 14]), teoría de conjuntos difusos ([15, 21, 22]) entre otras ([8, 10, 19]). De estos 11 estudios se destacan algunos en donde las técnicas estadísticas aplicadas fueron combinadas con técnicas de tecnologías de la información tales como Big Data [16], para desarrollar un indicador compuesto de corrupción a nivel de contrato y Machine Learning [14] como técnica de aprendizaje automático para detección de ofertas colusorias en contratos de construcción. Finalmente, en los últimos años se han venido aplicando técnicas de inteligencia artificial para la detección o identificación de índices de corrupción, es el ejemplo de [12, 17, 25] en donde se usaron técnicas de minería de datos, tales como reglas de asociación [12], algoritmos de clustering como K-means [25] y Redes Neuronales [17]. En el último grupo de estudios, equivalente al 15% ([9, 18, 26]) donde aplican diferentes técnicas a las mencionadas anteriormente, para detectar posibles situaciones de corrupción en diversos ámbitos.

3.3 Pregunta Q3. Tipos de corrupción en contratación que han sido tratadas

Respecto a los tipos de delitos de corrupción que se han tratado dentro de los estudios identificados como primarios, se tiene que todos los estudios en donde se trabaja uno varios tipos de delitos (el 65% equivalente a 13 estudios) coinciden en abordar el tema de la colusión, creación de carteles o manipulación de ofertas, de estos se destaca el 23% ([22-24]) de ellos, en donde, además, tratan el tema del soborno. Finalmente, el 35% de los estudios restantes (7) no enfatizan en un delito en específico.

3.4 Pregunta Q4. Dominios de aplicación donde se ha tratado el problema de la identificación de la corrupción

La mayoría de los estudios (85% equivalente a 17 estudios) trata el tema de la identificación o detección de corrupción sobre contratos, de ellos cerca del 71% lo hacen en el área de la construcción, esto guarda relación con lo planteado en [28], donde se afirma que la industria de la construcción es considerada "uno de los sectores más corruptos". Los estudios restantes de este grupo abordan el tema sobre contratos en entidades gubernamentales ([10, 23, 25]), salud [24] y educación [18]. Por otra parte, el 15% de los estudios restantes, abordan el problema sobre corrupción pública ([12, 13, 20]).

3.5 Pregunta Q5. Indicadores utilizados en la medición de corrupción o fraude

Quince estudios, equivalente al 75% del total respaldan su investigación con un indicador para medir sus resultados, entre estos se destacan los métodos binarios ([12, 14, 15, 25]) para informar si un caso en particular presenta o no delitos de corrupción, niveles de correlación para identificar variables que tiene más peso a la hora de predecir o detectar corrupción ([11, 20]), índice de probabilidad para informar que tan probable es que el objeto de estudio presente corrupción ([10, 17-19]) y un indicador ampliamente conocido en este campo de investigación, llamado "banderas rojas" ([16, 23, 24]) las cuales se pueden describir como señales de alarma, pistas o indicios de posible fraude. Finalmente, los 5 estudios restantes (25%) no mencionan haber utilizado un indicador para medir la corrupción o fraude.

3.6 Pregunta Q6. Problemas y desafíos identificados con relación con la corrupción

En la literatura analizada se observa que la identificación o detección de posibles delitos de corrupción conlleva a enfrentar ciertos desafíos que dificultan el proceso. A continuación, se listan los desafíos encontrados:

- Empresas corruptas pueden aplicar acciones que ayuden a evitar la identificación de corrupción [10, 20, 27].
- Falta de compromiso de las entidades [8, 12, 17, 22, 24-26].
- Falta de información [8-10, 13, 18, 19, 21, 24].
- La corrupción es un problema complejo para el que no hay una fácil o única solución [9, 11, 13, 15, 16, 19, 20, 27].
- La corrupción puede tomar muchas formas [11, 13, 18].
- No tener una definición clara del objeto a abordar [26].
- Las estrategias anticorrupción deben realizarse de forma discreta, ya que, si algún cartel se entera del proceso, podría destruir evidencias [14, 20, 23].

3.7 Resultados del Mapeo

Tras analizar cada una de las preguntas en cuestión, como resultados generales del mapeo se puede determinar lo siguiente:

- Aunque la mayoría de los estudios evaluados no se preocupan por redactar una definición global del concepto de corrupción, con lo recopilado se puede afirmar que la corrupción se define como una práctica deshonesta llevada a cabo por individuos, que abusan del poder público para su beneficio privado.

- La mayoría de los estudios son propuestas sobre la detección o identificación de corrupción, en donde aplican un modelo práctico, bien sea estadístico, tecnológico o una mezcla de ambos. En los últimos años, hay un crecimiento de los modelos que aplican técnicas informáticas para abordar el problema que se está estudiando (ver **Gráfica 3**). Estos trabajos se realizan en diferentes países, mostrando la importancia del tema en diferentes latitudes.
- El delito que más ha sido abordado para su identificación o detección es el de colusión, cartel o manipulación de ofertas, así mismo en la inmensa mayoría de estudios, el dominio de aplicación trabajado fueron los contratos en diferentes áreas, siendo la construcción al que más se le ha prestado atención.
- La mayoría de los estudios concuerdan en que los principales desafíos en aplicar técnicas anticorrupción son la falta de compromiso de las entidades, la falta de información y que el problema de corrupción es muy complejo para el que no hay una fácil o única solución.

A continuación, se presentan algunos resultados bibliométricos de los 20 artículos seleccionados:

La **Gráfica 2** muestra los estudios primarios realizados por país, discriminados por área, con esta gráfica se logra apreciar que el área de mayor interés a la hora de aplicar modelos de anticorrupción es la de construcción.

Similarmemente, la **Gráfica 3** muestra los estudios primarios realizados por año, discriminados por la técnica utilizada, se logra apreciar que en los últimos años hay un crecimiento de los modelos que utilizan técnicas computacionales.



Gráfica 2: Estudios primarios por área y país



Gráfica 3: Estudios primarios por año

4. Discusión

4.1 Observaciones Principales

El objetivo del presente mapeo sistemático es el de conocer el estado actual del campo de la identificación o detección de delitos de corrupción, con el fin de determinar las características más importantes que se deben tener en cuenta para la creación de un modelo que permita, precisamente la detección temprana de la corrupción en algún campo de aplicación específico. Luego de analizar los resultados se puede deducir lo siguiente:

- Las investigaciones sobre la identificación de posibles delitos de corrupción se han venido trabajando desde hace un tiempo, pero en los últimos años se ha abordado este tema con técnicas estadísticas que pueden ser llevadas al campo computacional, para una mayor optimización de los resultados, así mismo, en los últimos años ha crecido el interés por abordar este tema desde el área de las tecnologías de la información.
- La detección de la corrupción es muy compleja, entre otras causas porque es común que cada caso involucre muchos delitos. La mayoría de los estudios que se han realizado usando técnicas de tecnología de la información enfatizan la investigación de un o un conjunto delimitado de delitos, con el propósito de cerrar el problema a abordar.
- Si bien se han utilizado algunos índices para medir los resultados, no hay un estándar general que permita, evaluar si una solución propuesta es efectiva.
- El tema ha sido abordado por diferentes investigadores y en diferentes zonas geográficas, lo que demuestra, que, si bien no hay un grupo de personas empeñadas en el problema como tal, la solución a este, es de suma importancia, puesto que ha sido estudiado ampliamente en diversos sectores.

4.2 Limitaciones del Mapeo Sistemático

La principal limitación del presente mapeo sistemático se basa en los motores de búsqueda utilizados y la cadena de búsqueda como tal, la inclusión de estudios solo en inglés puede significar que los estudios relevantes en otros idiomas no se tengan en cuenta, pero, como es sabido, este idioma es el más usado a la hora de publicar artículos científicos.

4.3 Trascendencia para la Investigación y la Práctica

El presente mapeo sistemático es importante para aquellos interesados (investigadores, profesionales y entidades/organizaciones/empresas) en desarrollar investigaciones sobre estrategias para la detección o identificación de delitos de corrupción en un campo de aplicación específico.

Para los investigadores es un área muy interesante, ya que, como se ha visto, se trata de un campo que si bien ha sido trabajado hace un buen tiempo, en los últimos años se ha abordado desde técnicas más automáticas.

Gracias al desarrollo de este campo, los investigadores podrán crear nuevos modelos/técnicas/métodos/algoritmos que permitan crear índices de detección de corrupción o, así mismo, crear técnicas que combinen algunas de las estrategias ya utilizadas y que reportan mejores resultados.

Para las organizaciones, el presente trabajo les permite identificar los principales estudios relacionados con la identificación y detección de corrupción en contratación realizados a la fecha. Hasta el momento las pocas organizaciones que han desarrollado e implementado alguna técnica para prevenir la corrupción utilizan, en su inmensa mayoría, estrategias motivacionales, las cuales no ayudan a detectar la corrupción de forma temprana.

5. Conclusiones y Trabajo Futuro

La corrupción tiene enormes consecuencias negativas tanto en lo económico como en lo social. En los últimos años, el problema de la identificación temprana de la corrupción se ha trabajado desde el área de la estadística y más recientemente, desde las tecnologías de la información. Los resultados obtenidos en este mapeo sistemático demuestran lo dicho anteriormente, así como la necesidad de establecer unos índices de efectividad de los modelos anticorrupción. Por ello, es de suma importancia seguir en el desarrollo de modelos que permitan llevar a cabo esta idea. Respecto al trabajo futuro, el grupo de investigación pretende trabajar en lo siguiente: 1) El desarrollo de un modelo que integre diversas técnicas estadísticas y de computación, que permita la identificación temprana de la corrupción en un dominio específico de aplicación y 2) La elaboración de uno o más índices que permitan evaluar el modelo desarrollado.

Referencias

- [1] ACFE, "Report to the Nations of Occupational Fraud and Abuse: 2012 Global Fraud Study," *Association of Certified Fraud Examiners Inc.*, 2012.
- [2] T. por Colombia, "Sobre la Corrupción: Radiografía de la corrupción—La corrupción en Colombia—radiografía y retos," *Recuperado de <http://transparenciacolombia.org.co/index.php>*, 2013.
- [3] N. Ehlermann-Cache, *Bribery in public procurement: methods, actors and counter-measures*: OECD, 2007.
- [4] D. Budgen, M. Turner, P. Brereton, and B. A. Kitchenham, "Using Mapping Studies in Software Engineering," in *PPIG*, 2008, pp. 195-204.
- [5] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson, "Systematic mapping studies in software engineering," in *Ease*, 2008, pp. 68-77.
- [6] J. Bailey, D. Budgen, M. Turner, B. Kitchenham, P. Brereton, and S. Linkman, "Evidence relating to Object-Oriented software design: A survey," in *First International Symposium on Empirical Software Engineering and Measurement (ESEM 2007)*, 2007, pp. 482-484.
- [7] M. Bastidas, C. Pardo, and C. Ardila, "Risk-Based Testing: Preliminary Findings Obtained From A Systematic Mapping Study of the Literature."
- [8] R. H. Porter and J. D. Zona, "Detection of bid rigging in procurement auctions," *Journal of political economy*, vol. 101, pp. 518-538, 1993.
- [9] P. Bajari, "Comparing competition and collusion: a numerical approach," *Economic Theory*, vol. 18, pp. 187-205, 2001.

- [10] P. Bajari and L. Ye, "Deciding between competition and collusion," *Review of Economics and statistics*, vol. 85, pp. 971-989, 2003.
- [11] R. Chotibhongs and D. Arditi, "Detection of collusive behavior," *Journal of Construction Engineering and Management*, vol. 138, pp. 1251-1258, 2012.
- [12] C. G. Ralha and C. V. S. Silva, "A multi-agent data mining system for cartel detection in Brazilian government procurement," *Expert Systems with Applications*, vol. 39, pp. 11642-11656, 2012.
- [13] D. Imhof, Y. Karagök, and S. Rutz, "Screening for Bid Rigging—Does It Work?," *Journal of Competition Law & Economics*, vol. 14, pp. 235-261, 2018.
- [14] M. Huber and D. Imhof, "Machine learning with screens for detecting bid-rigging cartels," *International Journal of Industrial Organization*, vol. 65, pp. 277-301, 2019.
- [15] E. K. Owusu, A. P. Chan, and E. Ameyaw, "Toward a cleaner project procurement: Evaluation of construction projects' vulnerability to corruption in developing countries," *Journal of cleaner production*, vol. 216, pp. 394-407, 2019.
- [16] M. Fazekas, I. J. Tóth, and L. P. King, "An objective corruption risk index using public procurement data," *European Journal on Criminal Policy and Research*, vol. 22, pp. 369-397, 2016.
- [17] F. J. López-Iturriaga and I. P. Sanz, "Predicting public corruption with neural networks: An analysis of spanish provinces," *Social Indicators Research*, vol. 140, pp. 975-998, 2018.
- [18] J. Wachs and J. Kertész, "A network approach to cartel detection in public auction markets," *arXiv preprint arXiv:1906.08667*, 2019.
- [19] S. T. Parente, B. Schulte, A. Jost, T. Sullivan, and A. Klindworth, "Assessment of predictive modeling for identifying fraud within the Medicare program," *Health Manag Policy Innov*, vol. 1, pp. 8-37, 2012.
- [20] I. Morozov and E. A. Podkolzina, "Collusion detection in procurement auctions," *Higher School of Economics Research Paper No. WP BPR*, vol. 25, 2013.
- [21] M. Shan, A. P. Chan, Y. Le, B. Xia, and Y. Hu, "Measuring corruption in public construction projects in China," *Journal of Professional Issues in Engineering Education and Practice*, vol. 141, p. 05015001, 2015.
- [22] K. A. Belokrylov and L. P. Runova, "Localization of Subjective Assessment Risks in the Public Procurement System Based on Fuzzy Logic," *Journal of Advanced Research in Law and Economics*, vol. 8, pp. 411-417, 2017.
- [23] D. Peltier-Rivest, "A model for preventing corruption," *Journal of Financial Crime*, vol. 25, pp. 545-561, 2018.
- [24] D. Peltier-Rivest, "The prevention and detection of corruption in pharmaceutical companies," *Pharmaceuticals Policy and Law*, vol. 19, pp. 17-31, 2017.
- [25] M. T. Islam and M. A. Yousuf, "Development of a Corruption Detection Algorithm using K-means Clustering," in *2018 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)*, 2018, pp. 1-4.
- [26] J. A. Dorofeeva and M. Tokmakov, "The Use of Identical IP Address by Tender Participants as an Indication of Cartel Scheme," in *Digital Transformation of the Economy: Challenges, Trends and New Opportunities*, ed: Springer, 2020, pp. 244-252.
- [27] C. Stansbury, "The global infrastructure anticorruption centre," *Leadership and Management in Engineering*, vol. 9, pp. 119-122, 2009.
- [28] C. P. Index, "Transparency international," URL: http://www.transparency.org/news/feature/cpi_2013_now_is_the_time_for_action, 2010.

Una revisión sistemática de los métodos de identificación y detección de corrupción en contratación pública

Julio-César Luna-Ortega ^{1,2}, Carlos-Alberto Cobos-Lozada ¹, Martha-Eliana Mendoza-Becerra ¹

jcluna@unicauca.edu.co, ccobos@unicauca.edu.co, mmendoza@unicauca.edu.co

¹ Universidad del Cauca, Departamento de Sistemas, Calle 5 N° 4-70, 190003, Popayán, Colombia

² Clúster CreaTIC, Grupo de Investigación CreaTIC, Cl. 20N N° 7a-34, 190003, Popayán, Colombia

DOI: 10.17013/risti.n.pi-pf

Resumen: Recientemente, se ha observado un notable crecimiento de los casos de corrupción en procesos de contratación. Esto ha motivado a que las entidades promuevan estrategias anticorrupción, la mayoría de estas estrategias son manuales y de índole social. Para poder predecir o detectar efectivamente (eficiente y eficazmente) la corrupción se hace necesario realizar procesos automáticos. La presente revisión busca establecer los conceptos fundamentales del estado de arte relacionados con soluciones que apliquen estrategias anticorrupción y con base en esto, analizar la viabilidad de crear un modelo de detección e identificación de corrupción en el proceso de contratación pública. Los resultados muestran que combinar técnicas estadísticas con tecnologías de la información (inteligencia artificial, big data, minería de datos, entre otros) es una línea de investigación prometedora en esta área de aplicación. Además, que se hace necesario desarrollar modelos automáticos específicos para la detección de los diferentes delitos de corrupción en contratación.

Palabras clave: Corrupción, Contratación, Modelo, Revisión sistemática.

A systematic review of the methods for identification and detection of corruption in public procurement

Abstract: Recently, notable growth has been observed in corruption cases in public procurement processes. This has motivated entities to promote anti-corruption strategies, most of these strategies are manual and social in nature. In order to effectively (efficiently and efficiently) predict and detect corruption, it is necessary to carry out automatic processes. This review seeks to establish the fundamental concepts of the state of the art related to solutions that apply anti-corruption strategies and based on this, analyze the feasibility of creating a model for detecting and identifying corruption in the public procurement process. The results show that combining statistical techniques with information technologies (artificial intelligence, big data, data mining, among others) is a promising line of research in this area of application. Furthermore, it is necessary to develop specific automatic models for the detection of the different crimes of corruption in contracting.

Keywords: Corruption, Hiring, Model, Systematic review.

1. Introducción

El creciente descubrimiento de casos de corrupción en procesos de contratación pública que se ha dado en los últimos años se ha convertido en una oportunidad para la implementación de sistemas que analicen los datos y permitan detectar y prevenir nuevos casos de corrupción. A nivel mundial, el promedio de las pérdidas sufridas por las entidades víctimas de corrupción es aproximadamente de USD\$250,000 por caso.

Según diversas investigaciones el proceso más permeado por actos corruptos es el de contratación pública, lo cual, en una grave contradicción, puesto que la Contratación Estatal se presenta como una herramienta utilizada por un estado social de derecho para garantizar la satisfacción del interés general.

Cuando el proceso de contratación se ve afectado por un caso de corrupción, este deja de estar enfocado en la satisfacción del interés común para enmarcarse en el de terceros, afectando el bienestar de la comunidad y en el caso del Estado, alterando significativamente las finanzas éste.

El aumento de los casos de corrupción ha hecho que diversas entidades estén aplicando variadas estrategias anticorrupción, desafortunadamente la mayoría de estas son técnicas manuales y de enfoque social, que buscan prevenir, más que detectar o identificar posibles índices de corrupción. Si bien las técnicas manuales pueden ser útiles como estrategias anticorrupción con enfoques sociales, en estrategias con enfoque predictivo se hace necesario realizar un proceso automático (como modelos estadísticos y/o tecnológicos), dado que, como se menciona en (Ehlermann-Cache, 2007) la detección manual de la corrupción es intrínsecamente difícil puesto que esta se encubre por su propia naturaleza.

Esta revisión sistemática busca establecer los conceptos fundamentales del estado de arte relacionados con modelos y/o estrategias utilizadas en contra de la corrupción y con ello analizar la viabilidad de crear un modelo sistemático de detección e identificación de corrupción en el proceso de contratación pública.

La guía definida en (Kitchenham & Charters, 2007) fue utilizada como base para el proceso de la revisión sistemática utilizado en este estudio. El resto de este documento está organizado de la siguiente manera: en la sección 2 se presenta la planeación de la revisión con su protocolo de investigación, más adelante en la sección 3 se muestran los resultados obtenidos de dicha revisión, luego en la sección final se presentan las conclusiones y trabajos futuros que el grupo de investigación espera realizar.

2. Planeación de la Revisión Sistemática de la Literatura (RSL)

Los documentos relacionados con la temática de investigación se buscaron en Scopus, ACM Digital Library, IEEEExplore y ScienceDirect. Posterior a esto, se aplicaron los criterios de inclusión/exclusión para la selección de los estudios primarios y finalmente se realizó una evaluación de la calidad de los estudios incluidos.

2.1. Preguntas de investigación

Las preguntas de investigación de esta revisión buscan recopilar y categorizar la información existente acerca de la detección de corrupción en contratación pública, con el fin de realizar una revisión general del tema, identificando las carencias existentes y con esto, en un próximo paso plantear una alternativa de solución a la problemática expuesta (ver Tabla 8).

Tabla 8 - Preguntas de investigación.

Preguntas de Investigación	Motivación
Q1. ¿Qué estudios existen sobre la detección o identificación de corrupción en contratación?	Determinar el número de publicaciones actuales y la tendencia a lo largo de los últimos años en relación con el campo de la identificación de corrupción en contratación.
Q2. ¿Qué tipos de soluciones se han implementado para la identificación de corrupción?	Determinar cuáles son las técnicas, métodos, modelos o similares que se han utilizado para la detección de corrupción en contratación y como son los informes que estas realizan.
Q3. ¿Qué tipos de corrupción en contratación se han trabajado?	Determinar qué tipos de corrupción o fraude son normalmente a los que se presta más atención en el proceso de contratación.
Q4. ¿En qué dominios de aplicación se ha tratado el problema de la identificación de la corrupción?	Determinar cuáles son los objetos (contratos, negocios, etc.) en donde se ha trabajado sobre el problema de la identificación de corrupción.
Q5. ¿Qué indicadores se utilizan en las mediciones de corrupción o fraude?	Determinar cuáles son los indicadores que se suelen utilizar a la hora de evaluar un contrato y cuál es el dominio de valores que estos pueden tomar para alertar un posible caso de corrupción.
Q6. ¿Qué problemas y desafíos identifican los autores en relación con la corrupción?	Determinar cuáles son los problemas y retos identificados por los autores que trabajan en el proceso de identificación de la corrupción en la contratación.

2.2. Desarrollo del protocolo de revisión

A continuación, se resume el protocolo seguido para llevar a cabo la RSL.

Cadena de búsqueda de estudios primarios: Los términos de la cadena de búsqueda fueron contruidos a través de un proceso iterativo de mejora con el análisis de los resultados reportados por cada versión previa de esta. Se hizo uso de la estrategia planteada en (Bailey et al., 2007), la cual consiste en derivar los términos de las preguntas de investigación, identificando y agrupándolos en población, intervención, comparación y resultado (ver Tabla 9). Los operadores booleanos OR incorporan palabras alternativas y sinónimos en cada una de las partes, y los operadores AND para vincular los términos principales entre sí.

Tabla 9 - Cadena de búsqueda usada para iniciar la revisión.

Concepto	Estructura	Términos Alternativos y Sinónimos	Descripción
Técnicas	Resultado	("model" OR "algorithm" OR "framework" OR "technology" OR "method" OR "technique") AND	Se delimita lo que se está buscando: algoritmos, técnicas, modelos y similares.
Corrupción	Población	("*corruption" OR "collusion" OR "fraud") AND	Se define el problema a nivel general: corrupción y similares.
Tipos de delitos	Comparación	("bribery" OR "buys influence" OR "conflicts of interest" OR "cronyism" OR "embezzlement" OR "extort*" OR "favoritism" OR "gift-giving" OR "graft" OR "influence peddling" OR "kickback" OR "nepotism" OR "parochialism" OR "patronage" OR "political interference" OR "rigging" OR "splitting" OR "theft") AND	Se delimitan los tipos específicos de delitos de corrupción.
Objeto	Población	("governm*" OR "agreement" OR "contract" OR "commitment" OR "arrangement" OR "settlement" OR "understanding" OR "bond" OR "deal" OR "bargain") AND	Se menciona el objeto que tiene el problema: contratación gubernamental y palabras relacionadas con contrato.
Acciones	Intervención	("detect" OR "avoid" OR "prevent" OR "warn" OR "caution" OR "identify" OR "understanding" OR "avert" OR "stop" OR "alert" OR "perceive" OR "distinguish" OR "notify" OR "inform" OR "recognize" OR "establish")	Se relaciona con el tipo de soluciones que buscan identificar, evitar, prevenir (o similares) la corrupción.

*Los asteriscos significan que cualquier carácter o caracteres pueden ser incluidos en la palabra, de esta manera se consiguen permutaciones en los términos de búsqueda en los que se utiliza.

Proceso de búsqueda: Las búsquedas se realizaron sobre Scopus, ACM Digital Library, IEEEExplore y ScienceDirect. Siguiendo el proceso propuesto en (Bailey et al., 2007), la búsqueda se organizó en fases; en la primera se identificaron los estudios candidatos derivados de las palabras claves de la cadena de búsqueda, una vez esto que se procedió a aplicar los criterios de selección sobre el abstract de cada uno de ellos, posteriormente los mismos criterios fueron aplicados a todo el documento de los estudios resultantes de la fase previa, finalmente se aplicó la técnica “bola de nieve”, la cual consiste en revisar las referencias de cada estudio primario buscando identificar más estudios candidatos, así mismo se hizo una revisión de la literatura gris. La Tabla 10 resume el proceso de búsqueda realizado.

Tabla 10 – Proceso de búsqueda.

Actividad	Resultados
Aplicar cadena de búsqueda	245
Eliminar duplicados	229
Aplicar criterios de selección (inclusión y exclusión)	40
Leer documento completo	8
Revisar artículos excluidos y Snowballing (bola de nieve)	16
Revisar literatura gris	19

Criterios de inclusión y exclusión: Los criterios para incluir un estudio primario fueron los siguientes: 1) Estudios escritos en el idioma inglés; 2) Estudios completos publicados en revistas, conferencias, congresos o talleres de prestigio con revisión por pares; y 3) El estudio propone un modelo, técnica, algoritmo, framework o similar. Los criterios para excluir un estudio primario fueron los siguientes: 1) El artículo no presenta un diseño de investigación bien definido; 2) El artículo no presenta una pregunta de investigación bien definida; 3) El artículo corresponde a un estudio de debate, o solo está disponible en forma de resumen o presentación. 4) El artículo es un estudio secundario, es decir, corresponde a un reporte de resultados de otras investigaciones; 5) La contribución principal del artículo no se relaciona con alguna modelo/técnica/método/algoritmo de identificación de corrupción; 6) El artículo no propone una modelo/técnica/método/algoritmo, o que lo hagan, pero la propuesta no pueda ser implementada con ayuda del computador; y 7) El artículo presenta un modelo de índole psicosocial. Por último, vale la pena mencionar que, debido a la poca existencia de investigaciones que abordan el problema presentado en esta investigación, no se tuvo en cuenta el año de publicación del artículo como un criterio para incluir o no el trabajo en la presente revisión sistemática.

Evaluación de calidad de los estudios primarios: Adicional a la selección (criterios de inclusión y exclusión) se realizó una evaluación de la calidad de los estudios primarios (Bailey et al., 2007) para ponderar la importancia de cada estudio seleccionado. Esta evaluación permitió comprender mejor las posibles comparaciones entre los estudios y guiar la interpretación de los hallazgos. Con este objetivo se desarrolló una prueba con un sistema de puntuación de tres valores (-1, 0, 1) en 6 preguntas (ver Tabla 11). La puntuación total máxima corresponde a la suma de las calificaciones parciales y el puntaje total está en el dominio [-6, 6]. El puntaje de calidad de cada estudio se presenta en la Tabla 12, donde 6 es el máximo valor posible.

Tabla 11 - Prueba de evaluación de la calidad.

No	Criterios	Puntuación		
		-1	0	1
C1	El estudio presenta una descripción detallada sobre los tipos de corrupción o fraude que se presentan en un objeto de población (ejemplo contratos)	No	Parcialmente	Si
C2	El estudio contiene un modelo/técnica/método/algoritmo/similar aplicada.	No	Parcialmente	Si
C3	El estudio tiene como finalidad aplicar acciones (identificar, detectar, prevenir, etc.) sobre un problema de corrupción	No	Parcialmente	Si
C4	El estudio expone de manera clara y detallada los resultados obtenidos tras aplicar un modelo/técnica/método/algoritmo/similar sobre un problema de corrupción en un objeto de población	No	Parcialmente	Si
C5	A qué cuartil pertenece la revista, conferencia o congreso en donde fue publicado el estudio	Q4 O NO	Q3	Q2 O Q1
C6	El estudio ha obtenido dos o más citas por año desde su publicación (según índice de citas de Google, Promedio anual)	menos de 1 cita por año	1 cita por año	2 o más citas por año

Tabla 12 - Resultados de la evaluación de los estudios primarios ordenado por año y puntaje.

Puntaje	Referencia
6	(Porter & Zona, 1993)
6	(Bajari, 2001)
6	(Bajari & Ye, 2003)
3	(Parente, Schulte, Jost, Sullivan, & Kindworth, 2012)
6	(Chotibhongs & Arditi, 2012)
6	(Ralha & Silva, 2012)
3	(Morozov & Podkolzina, 2013)
3	(Shan, Chan, Le, Xia, & Hu, 2015)
5	(Fazekas, Tóth, & King, 2016)
1	(Peltier-Rivest, 2017)
3	(Belokrylov & Runova, 2017)
1	(Islam & Yousuf, 2018)
3	(Peltier-Rivest, 2018)
5	(López-Iturriaga & Sanz, 2018)
6	(Imhof, Karagök, & Rutz, 2018)
5	(Wachs & Kertész, 2019)
6	(Huber & Imhof, 2019)
6	(Owusu, Chan, & Ameyaw, 2019)
1	(Dorofeeva & Tokmakov, 2020)

3. Resultados de la Revisión Sistemática de la Literatura (RSL)

Haciendo uso del protocolo mencionado se encontraron 19 estudios que cumplieron con los diferentes criterios de calidad establecidos. Estos estudios se analizaron y organizaron en 5 grupos (A, B, C, D y E) conforme se presentan en la Tabla 13. En esta tabla, además, se muestra para cada estudio las respuestas a cada pregunta de investigación previamente formulada. Para determinar el grupo al que pertenecía cada estudio se hizo un análisis de las respuestas obtenidas por cada uno de ellos en las diferentes preguntas de investigación, siendo la pregunta **Q2** la de mayor peso en esta decisión.

Tabla 13 - Agrupamiento de los estudios

Id	Ref.	Preguntas
A	(Porter & Zona, 1993)	Q1: Propuesta.
		Q2: Caso de estudio de un cartel con información de cuentas debidamente detalladas respecto a sus actividades y prácticas de licitación. Técnicas estadísticas. Procedimientos de prueba econométricos.
		Q3: Colusión. Aparejo de ofertas (práctica ilegal en la que los competidores se confabulan para elegir al ganador de una licitación).
		Q4: Contrato.
		Q5: No utiliza.
		Q6: Falta de información de contratación pública.
	(Bajari & Ye, 2003)	Q1: Propuesta.
		Q2: Técnicas estadísticas. Enfoque para identificar y probar la manipulación de ofertas en subastas de adquisición. Los autores destacan la importancia de la implementación de herramientas computacionales que puedan ser utilizadas para detectar la manipulación de ofertas.
		Q3: Aparejo de ofertas.
		Q4: Contrato.
		Q5: Índice de probabilidad.
		Q6: Falta de información de contratación pública. Empresas corruptas pueden aplicar acciones que ayuden a evitar la identificación de la corrupción.
	(Chotibhongs & Arditi, 2012)	Q1: Propuesta.
		Q2: Técnicas estadísticas. Método paso a paso para detectar las ofertas colusorias mediante el análisis de los datos históricos de ofertas proporcionados por el propietario de una construcción. El método se basa en un modelo de regresión e implica identificar a los posibles postores del cartel utilizando la prueba residual y la prueba de estabilidad de la estructura de costos.
		Q3: Colusión. Aparejo de ofertas.
		Q4: Contrato.
		Q5: Niveles de correlación.
		Q6: La corrupción en los proyectos es un problema complejo para el que no hay una fácil o única solución. La corrupción puede tomar muchas formas. El modelo no puede detectar las ofertas colusorias en el momento en que sucede; se usa para detectar posible colusión en el pasado y solo si hay acceso a información histórica detallada sobre un conjunto de variables independientes.
	(Morozov & Podkolzina, 2013)	Q1: Propuesta.
		Q2: Técnicas estadísticas. Método econométrico de detección de manipulación de ofertas.

	<p>Q3: Colusión. Aparejo de ofertas.</p> <p>Q4: Contrato.</p> <p>Q5: Niveles de correlación.</p> <p>Q6: La corrupción en los proyectos es un problema complejo para el que no hay una fácil o única solución.</p>
(Shan et al., 2015)	<p>Q1: Encuesta. Identificación de elementos importantes de medición de corrupción.</p> <p>Q2: Técnicas estadísticas. Modelo sistemático para medir la corrupción potencial en un proyecto de construcción pública.</p> <p>Q3: No enfatiza un delito.</p> <p>Q4: Contrato.</p> <p>Q5: No utiliza.</p> <p>Q6: Falta de información de contratación pública.</p>
(Imhof et al., 2018)	<p>Q1: Propuesta.</p> <p>Q2: Técnicas estadísticas. Método para detectar la manipulación de ofertas mediante la aplicación de pantallas (índices estadísticos específicos). El método es particularmente adecuado para detección de la colusión parcial (colusión que no involucra a todas las empresas y/o todos los contratos en un conjunto de datos específico). Se logra aislar a un grupo de empresas "sospechosas" de ser un cartel local de manipulación de ofertas.</p> <p>Q3: Colusión. Aparejo de ofertas.</p> <p>Q4: Contrato.</p> <p>Q5: Índice de probabilidad.</p> <p>Q6: Falta de información de contratación pública. La corrupción en los proyectos es un problema complejo para el que no hay una fácil o única solución. La corrupción puede tomar muchas formas.</p>
(Parente et al., 2012)	<p>Q1: Propuesta.</p> <p>Q2: Técnicas estadísticas. Inteligencia artificial.</p> <p>Q3: Fraude.</p> <p>Q4: Contrato.</p> <p>Q5: Índice de probabilidad.</p> <p>Q6: La corrupción en los proyectos es un problema complejo para el que no hay una fácil o única solución. Falta de información de contratación pública.</p>
(Fazekas et al., 2016)	<p>Q1: Propuesta.</p> <p>Q2: Técnicas estadísticas. Inteligencia artificial. El método hace uso de Big Data para crear un indicador de corrupción a nivel de contrato.</p> <p>Q3: No enfatiza un delito.</p> <p>Q4: Contrato.</p> <p>Q5: Banderas rojas. Propone indicador de corrupción a través de Big Data.</p> <p>Q6: La corrupción en los proyectos es un problema complejo para el que no hay una fácil o única solución.</p>
(Belokrylov & Runova, 2017)	<p>Q1: Propuesta.</p> <p>Q2: Técnicas estadísticas. Algoritmos computacionales. Propone un modelo basado en lógica difusa.</p> <p>Q3: Aparejo de ofertas. Sobornos.</p> <p>Q4: Contrato.</p> <p>Q5: No utiliza.</p> <p>Q6: La corrupción en los proyectos es un problema complejo para el que no hay una fácil o única solución. Falta de compromiso de las entidades.</p>
B	
(Huber & Imhof, 2019)	<p>Q1: Propuesta.</p> <p>Q2: Técnicas estadísticas. Inteligencia artificial. Combinan técnicas de aprendizaje automático con pantallas estadísticas.</p> <p>Q3: Aparejo de ofertas.</p> <p>Q4: Contrato.</p> <p>Q5: Binarios. Clasificación binaria que informa si una oferta es marcada como posible colusoria, clasificando como 1 si la probabilidad de colusión de la oferta es superior a 0,5 y 0 de lo contrario.</p> <p>Q6: Debe realizarse de forma discreta, si algún cartel se entera del proceso, podría destruir evidencia. Falta de compromiso de las entidades. Información de costos es difícil de obtener.</p>
(Owusu et al., 2019)	<p>Q1: Propuesta. Muestra una revisión de la literatura con 21 actividades que se deben realizar en las cuatro etapas del proceso de adquisición.</p> <p>Q2: Técnicas estadísticas. Algoritmos computacionales. Modelo que usa un enfoque de computación flexible (técnica de evaluación sintética difusa).</p> <p>Q3: No enfatiza un delito.</p> <p>Q4: Contrato.</p> <p>Q5: Índice de probabilidad.</p> <p>Q6: La corrupción en los proyectos es un problema complejo para el que no hay una fácil o única solución.</p>

C	(Ralha & Silva, 2012)	Q1: Propuesta. Se enfoca en extraer información de las bases de datos del proceso de adquisiciones federales brasileñas para identificar carteles. Q2: Inteligencia artificial. Minería de datos: Agrupamiento y reglas de asociación. Q3: Carteles. Q4: Contrato. Q5: Binario. Q6: Falta de compromiso de las entidades.
	(Islam & Yousuf, 2018)	Q1: Propuesta. Q2: Inteligencia artificial. Se ha propuesto un modelo que usa el algoritmo K-means. Q3: No enfatiza en un delito. Q4: Corrupción pública. Q5: Índice de probabilidad. Q6: Falta de compromiso de las entidades.
	(López-Iturriaga & Sanz, 2018)	Q1: Propuesta. Q2: Inteligencia artificial Propone sistema de alertas tempranas basado en red neuronal (mapas autoorganizados, SOM). Q3: No enfatiza un delito. Q4: Corrupción pública. Q5: Binarios. Q6: Falta de compromiso de las entidades.
D	(Bajari, 2001)	Q1: Propuesta. Q2: Algoritmos computacionales. Los algoritmos computacionales desarrollados en este documento pueden usarse para la evaluación de los daños causados por la manipulación de ofertas. Q3: Colusión. Aparejo de ofertas. Q4: Contrato. Q5: Diferencia entre ofertas corruptas y competitivas. Q6: La corrupción en los proyectos es un problema complejo para el que no hay una fácil o única solución. Falta de información de contratación.
	(Wachs & Kertész, 2019)	Q1: Propuesta. Q2: Ciencia de redes. Propone un marco basado en análisis de redes (grafos) para identificar grupos de empresas que interactúan intensamente en un mercado. Q3: Colusión. Aparejo de ofertas. Q4: Contrato. Q5: Índices de probabilidad. Q6: Falta de información de contratación pública. La corrupción puede tomar muchas formas.
	(Dorofeeva & Tokmakov, 2020)	Q1: Propuesta. Q2: Inteligencia artificial. Análisis de las direcciones IP de los miembros de una licitación pública. Q3: Aparejo de ofertas. Q4: Contrato. Q5: No utiliza. Q6: Falta de compromiso de las entidades. Falta de claridad en la definición del objeto a abordar.
E	(Peltier-Rivest, 2017)	Q1: Estudio de caso. Q2: No relacionado con tecnología o estadística. Propone un marco de gobernanza, haciendo uso de la teoría del diamante del fraude. Q3: Aparejo de ofertas. Sobornos. Conflictos de interés. Extorsión. Q4: Contrato. Q5: Banderas rojas. Q6: Falta de información de contratación pública. Falta de compromiso de las entidades.
	(Peltier-Rivest, 2018)	Q1: Propuesta. Muestra aspectos interesantes para tener en cuenta en el proceso de detección de la corrupción. Q2: No relacionado con tecnología o estadística. Introduce un modelo derivado de las teorías criminológicas (como el triángulo de fraude). Q3: Aparejo de ofertas. Sobornos. Q4: Contrato. Q5: Banderas rojas. Q6: Debe realizarse de forma discreta, si algún cartel se entera del proceso, podría destruir evidencia. Falta de compromiso de las entidades.

3.1. Análisis de los grupos resultantes

A continuación, se muestran un análisis detallado de la conformación de cada uno de los grupos con relación a las preguntas de investigación planteadas.

Grupo A: Los estudios pertenecientes a este grupo tienen un enfoque de solución relacionada con técnicas estadísticas como Test econométricos ((Morozov & Podkolzina, 2013; Porter & Zona, 1993)), pantallas ((Imhof et al., 2018)), regresión ((Chotibhongs & Arditi, 2012)) y teoría de conjuntos difusos ((Bajari & Ye, 2003; Shan et al., 2015)). Los tipos de delitos de corrupción abordados por este grupo de estudios son en su mayoría la colusión y el aparejo de ofertas, así mismo las métricas de corrupción utilizadas tienden a ser análisis estadísticos como es el caso de un índice de probabilidad. Es de destacar que los estudios pertenecientes a este grupo se presentan desde 1993 hasta 2018.

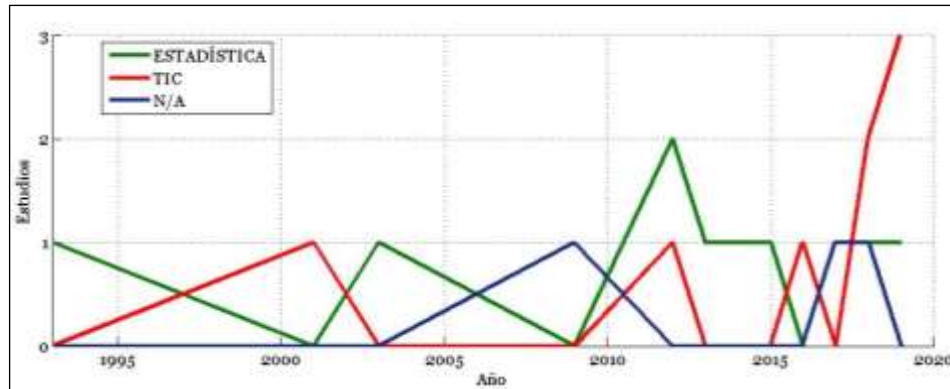
Grupo B: La principal característica de los estudios pertenecientes a este grupo es que el modelo desarrollado involucra técnicas estadísticas con métodos de Inteligencia Artificial, destacándose la utilización de la teoría de conjuntos difusos con algoritmos de agrupamiento (Belokrylov & Runova, 2017) buscando medir el riesgo de corrupción en el proceso de contratación o técnicas de evaluación difusas con el objetivo de evaluar los niveles de vulnerabilidad con respecto a las etapas de corrupción (Owusu et al., 2019), así mismo la combinación del análisis de regresión con algoritmos predictivos para la detección analítica de fraude (Parente et al., 2012) o con algoritmos de big data (Fazekas et al., 2016) para el desarrollo de un indicador compuesto de corrupción y finalmente la combinación de pantallas con aprendizaje de máquina (Huber & Imhof, 2019) con el objetivo de predecir la colusión a través de carteles de manipulación de ofertas. La mayoría de estos estudios no enfatizan su solución a un delito en particular, y los que los hacen se enfocan en el aparejo de ofertas. Por otro lado, similar al grupo anterior, la tendencia de las métricas de corrupción utilizadas se encuentra en un marco estadístico. Estos trabajos son relativamente nuevos, teniendo su primera publicación en 2012, así mismo, se puede apreciar que la tendencia de este tipo de artículos ha venido incrementando en los últimos años.

Grupo C: Los integrantes de este grupo son estudios en donde se aplicó una solución tecnológica basada en Minería de Datos, haciendo uso de algoritmos de agrupamiento y reglas asociación en (Ralha & Silva, 2012), donde se obtiene un modelo que permitió detectar las ofertas colusorias mediante el análisis de los datos históricos de ofertas. En (Islam & Yousuf, 2018) hacen uso del algoritmo K-means (agrupamiento) el cual permitió crear diferentes grupos de acuerdo con el nivel de corrupción de un grupo de entidades. Finalmente en (López-Iturriaga & Sanz, 2018) se usó un enfoque de red neuronal, mapas autoorganizados (agrupamiento), para predecir casos de corrupción en diferentes espacios temporales, en una región específica de un país. Los estudios pertenecientes a este grupo por lo general no se enfocan en un delito en particular, solo uno de ellos lo hace, enfocándose en la organización de carteles. Por otro lado, en este grupo prima un indicador binario como métrica de corrupción. Finalmente, y, como era de esperarse, similar al anterior grupo este tipo de soluciones es relativamente nueva teniendo su primera publicación en 2012.

Grupo D: Los integrantes de este grupo, si bien hacen uso de técnicas computacionales, difieren del anterior en que dichas técnicas no están enmarcadas dentro de la Inteligencia Artificial. Destacándose soluciones que hacen uso de algoritmos computacionales (Bajari, 2001) para encontrar diferencias entre ofertas competitivas y colusorias, o la utilización de análisis de redes (grafos) (Wachs & Kertész, 2019) para la identificación de grupos de empresas que interactúan intensamente en un mercado y detectar su potencial colusorio. Y finalmente el análisis de las direcciones IP de los miembros de una licitación pública para la identificación de posible presencia de un cartel en un proceso de licitación. (Dorofeeva & Tokmakov, 2020). Este grupo se centra en la detección de colusión y aparejo de ofertas, teniendo como principal métrica los índices de probabilidades. Finalmente, y similar a los anteriores grupos que hacen uso de técnicas de tecnologías de la información, la tendencia de este tipo de investigaciones es relativamente nueva.

Grupo E: El último grupo de investigaciones está conformado por los estudios cuyo modelo de solución no está relacionado con tecnología o estadística, en estos estudios se despliegan una serie de estrategias para prevenir o detectar la corrupción, haciendo uso de la teoría del diamante del fraude (Peltier-Rivest, 2017, 2018). Los delitos que se abordan en este grupo son la colusión y el aparejo de ofertas, teniendo como métrica el uso de banderas rojas.

Por último, en la Gráfica 4 se muestran los estudios primarios realizados por año, discriminados por la técnica utilizada, se logra apreciar que en los últimos años hay un crecimiento de los modelos que utilizan técnicas computacionales.



Gráfica 4 - Estudios primarios por año.

4. Conclusiones y Trabajo Futuro

Los resultados obtenidos en esta revisión sistemática demuestran que el uso de técnicas estadísticas para solventar el problema de la identificación temprana de la corrupción es una idea que en los últimos años ha tenido una buena acogida, así como la combinación de estas técnicas con tecnologías de la información. Lo anterior sumado al incremento en los casos de corrupción global reportados, llevan a la necesidad de seguir en el desarrollo de modelos que permitan solventar de manera más eficaz y eficiente el problema de la detección de la corrupción en procesos de contratación. Los estudios muestran que, dada la complejidad del problema, se hace necesario contar con un conjunto (batería) de herramientas especializadas para atacar los diferentes delitos que ocurren en contratación.

Respecto al trabajo futuro, el grupo de investigación pretende trabajar en lo siguiente: 1) El desarrollo de un modelo que integre diversas técnicas estadísticas y de computación, que permita la identificación temprana de la corrupción en un dominio específico de aplicación y 2) La elaboración de uno o más índices que permitan evaluar el modelo desarrollado.

Agradecimientos

A la Universidad del Cauca y al Clúster CreaTIC por financiar parcialmente el desarrollo de la presente investigación.

Referencias

- Bailey, J., Budgen, D., Turner, M., Kitchenham, B., Brereton, P., & Linkman, S. (2007). Evidence relating to Object-Oriented software design: A survey. Paper presented at the First International Symposium on Empirical Software Engineering and Measurement (ESEM 2007).
- Bajari, P. (2001). Comparing competition and collusion: a numerical approach. *Economic Theory*, 18(1), 187-205. doi: <https://doi.org/10.1007/PL00004128>
- Bajari, P., & Ye, L. (2003). Deciding between competition and collusion. *Review of Economics and statistics*, 85(4), 971-989. doi: <https://doi.org/10.1162/003465303772815871>
- Belokrylov, K. A., & Runova, L. P. (2017). Localization of Subjective Assessment Risks in the Public Procurement System Based on Fuzzy Logic. *Journal of Advanced Research in Law and Economics*, 8(2 (24)), 411-417. doi: [https://dx.doi.org/10.14505/jarle.v8.2\(24\).09](https://dx.doi.org/10.14505/jarle.v8.2(24).09)
- Chotibhongs, R., & Arditi, D. (2012). Detection of collusive behavior. *Journal of Construction Engineering and Management*, 138(11), 1251-1258. doi: [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000542](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000542)
- Dorofeeva, J. A., & Tokmakov, M. (2020). The Use of Identical IP Address by Tender Participants as an Indication of Cartel Scheme Digital Transformation of the Economy: Challenges, Trends and New Opportunities (pp. 244-252): Springer.

- Ehlermann-Cache, N. (2007). Bribery in public procurement: methods, actors and counter-measures: OECD.
- Fazekas, M., Tóth, I. J., & King, L. P. (2016). An objective corruption risk index using public procurement data. *European Journal on Criminal Policy and Research*, 22(3), 369-397. doi: <https://doi.org/10.1007/s10610-016-9308-z>
- Huber, M., & Imhof, D. (2019). Machine learning with screens for detecting bid-rigging cartels. *International Journal of Industrial Organization*, 65, 277-301. doi: <https://doi.org/10.1016/j.ijindorg.2019.04.002>
- Imhof, D., Karagök, Y., & Rutz, S. (2018). Screening for Bid Rigging—Does It Work? *Journal of Competition Law & Economics*, 14(2), 235-261. doi: <https://doi.org/10.1093/joclec/nhy006>
- Islam, M. T., & Yousuf, M. A. (2018). Development of a Corruption Detection Algorithm using K-means Clustering. Paper presented at the 2018 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE).
- Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering.
- López-Iturriaga, F. J., & Sanz, I. P. (2018). Predicting public corruption with neural networks: An analysis of spanish provinces. *Social Indicators Research*, 140(3), 975-998. doi: <https://doi.org/10.1007/s11205-017-1802-2>
- Morozov, I., & Podkolzina, E. A. (2013). Collusion detection in procurement auctions. Higher School of Economics Research Paper No. WP BPR, 25. doi: <http://dx.doi.org/10.2139/ssrn.2221809>
- Owusu, E. K., Chan, A. P., & Ameyaw, E. (2019). Toward a cleaner project procurement: Evaluation of construction projects' vulnerability to corruption in developing countries. *Journal of cleaner production*, 216, 394-407. doi: <https://doi.org/10.1016/j.jclepro.2019.01.124>
- Parente, S. T., Schulte, B., Jost, A., Sullivan, T., & Klindworth, A. (2012). Assessment of predictive modeling for identifying fraud within the Medicare program. *Health Manag Policy Innov*, 1(2), 8-37.
- Peltier-Rivest, D. (2017). The prevention and detection of corruption in pharmaceutical companies. *Pharmaceuticals Policy and Law*, 19(1-2), 17-31. doi: <https://dx.doi.org/10.3233/PPL-170451>
- Peltier-Rivest, D. (2018). A model for preventing corruption. *Journal of Financial Crime*, 25(2), 545-561.
- Porter, R. H., & Zona, J. D. (1993). Detection of bid rigging in procurement auctions. *Journal of political economy*, 101(3), 518-538. doi: <https://doi.org/10.1086/261885>
- Ralha, C. G., & Silva, C. V. S. (2012). A multi-agent data mining system for cartel detection in Brazilian government procurement. *Expert Systems with Applications*, 39(14), 11642-11656. doi: <https://doi.org/10.1016/j.eswa.2012.04.037>
- Shan, M., Chan, A. P., Le, Y., Xia, B., & Hu, Y. (2015). Measuring corruption in public construction projects in China. *Journal of Professional Issues in Engineering Education and Practice*, 141(4), 05015001. doi: [https://doi.org/10.1061/\(ASCE\)EI.1943-5541.0000241](https://doi.org/10.1061/(ASCE)EI.1943-5541.0000241)
- Wachs, J., & Kertész, J. (2019). A network approach to cartel detection in public auction markets. arXiv preprint arXiv:1906.08667. doi: <https://doi.org/10.1038/s41598-019-47198-1>