



CONSEJO INTERNACIONAL
PARA LA LUCHA CONTRA LA CORRUPCIÓN

**MODELOS DE LENGUAJE DE GRAN TAMAÑO EN EL CONTROL DE LA
CONTRATACIÓN PÚBLICA: QUÉ DETECTAN Y QUÉ SE SABE SOBRE LOS
RESULTADOS. LA EXPERIENCIA DE BRASIL, COREA, CHILE, COLOMBIA Y LA
UNIÓN EUROPEA**

Dzianis Limarau, Ignat Krukovskiy

Consejo Internacional para la Lucha contra la Corrupción, asociación civil
Buenos Aires, 2026
Versión 1

DOI: 10.5281/zenodo.23241236
Publicado bajo licencia CC BY 4.0
<https://cilclegal.org/>

GRNTI 10.07.45, 06.56.31, 20.53.19; UNESCO 5605.99, 5909.01, 1203.04; OECD FOS 5.05 Derecho, 5.02
Economía y Empresa, 1.02 Ciencias de la Computación y de la Información

Cita recomendada: Limarau, D., Krukovskiy, I. (2026). Modelos de lenguaje de gran tamaño en el control de la contratación pública: qué detectan y qué se sabe sobre los resultados. Report. Consejo Internacional para la Lucha contra la Corrupción. DOI 10.5281/zenodo.23241236. <https://cilclegal.org/>

Índice

Índice	2
Objetivo del estudio	3
Destinatarios	3
Resumen ejecutivo	4
1. Introducción	5
2. Detección automatizada de riesgos en la contratación: de las banderas rojas a los modelos de lenguaje	6
3. Datos y método	6
3.1. Selección de las herramientas	6
3.2. Ficha de la herramienta	7
3.3. Cotejo con la tipología	7
3.4. Indicadores de resultados	7
3.5. Terminología	8
4. Herramientas basadas en modelos de lenguaje de gran tamaño en cinco jurisdicciones	8
4.1. Brasil: Alice 360 del Tribunal de Cuentas de la Unión	9
4.2. Brasil: VigIA del Tribunal de Cuentas del Estado de Santa Catarina y PRUMe AI del Tribunal de Cuentas del Estado de Amazonas	10
4.3. Corea: monitoreo de anuncios del Servicio de Contratación Pública	11
4.4. Corea: verificación de documentos contractuales en el distrito de Seocho	12
4.5. Chile: análisis de documentos de ChileCompra	12
4.6. Italia: monitoreo de anuncios de la Autoridad Nacional Anticorrupción	13
4.7. Unión Europea: la Oficina Europea de Lucha contra el Fraude y Arachne+	13
4.8. Colombia y otros casos sin modelo de lenguaje confirmado	13
5. Qué mecanismos cubren los modelos de lenguaje	14
5.1. Funciones y mecanismos	14
5.2. Matriz de cobertura	16
5.3. Coincidencia con el límite de visibilidad del participante externo	18
6. Qué se sabe sobre los resultados	18
7. Indicios fuera de la tipología	20
8. Limitaciones y una pregunta abierta sobre las banderas rojas	21
9. Conclusiones	21
Fuentes	22
Anexo. Herramientas sin modelos de lenguaje en las mismas jurisdicciones	25

Objetivo del estudio

Los organismos públicos de Brasil, la República de Corea, Chile, Colombia y la Unión Europea han comenzado a aplicar modelos de lenguaje de gran tamaño para verificar la contratación pública. Las preguntas principales del estudio son: qué herramientas basadas en modelos de lenguaje aplican exactamente esos organismos, qué verifican tales herramientas, qué mecanismos de distorsión del procedimiento competitivo son capaces de detectar y qué se ha publicado sobre los resultados de su funcionamiento.

En el estudio se denomina herramienta de control al sistema informático con cuya ayuda un organismo público, ya sea un órgano de control financiero, una agencia de contratación, un órgano antimonopolio, de supervisión o de aplicación de la ley, detecta de manera automática posibles infracciones en los procedimientos de contratación. Se denomina modelo de lenguaje de gran tamaño (en adelante también LLM, del inglés large language model) al modelo de inteligencia artificial que procesa texto en lenguaje natural, y herramienta LLM a la herramienta de control en la que al menos una función la ejecuta un modelo de ese tipo. Las herramientas de otra naturaleza, construidas sobre reglas, cruces de registros, umbrales estadísticos y aprendizaje automático clásico, se describen en el anexo a efectos de comparación.

Las tareas del estudio:

1. Elaborar un inventario de las herramientas LLM de control de la contratación en cinco jurisdicciones a octubre de 2026. Describir, para cada herramienta, sus funciones, sus datos, el tratamiento de la alerta y la etapa de implantación. La fuente son los documentos de los propios organismos que aplican las herramientas.
2. Cotejar las funciones de cada herramienta con los 30 mecanismos de la tipología del Consejo Internacional para la Lucha contra la Corrupción (en adelante, CILC; Limarau, 2026). Determinar qué etapas del ciclo de contratación y qué grupos de mecanismos cubren las herramientas y cuáles quedan fuera de su campo de visión.
3. Calcular, a partir de los datos publicados, cuatro indicadores de resultados: la cobertura, la proporción de alertas verificadas por una persona, la proporción confirmada y la proporción que dio lugar a medidas. Señalar para qué herramientas es imposible calcular esos indicadores.
4. Identificar los indicios que buscan las herramientas y que no figuran en la tipología, como candidatos a nuevos mecanismos.

Destinatarios

El estudio se dirige a los órganos de control y a las agencias de contratación que implantan o planean implantar modelos de lenguaje para verificar la contratación, a los investigadores de los riesgos de corrupción que trabajan con datos de los sistemas electrónicos de contratación, así como a las organizaciones que evalúan las herramientas de gobernanza de la inteligencia artificial, entre ellas el Grupo Científico Internacional Independiente sobre la IA de las Naciones Unidas, cuyo Informe preliminar de 1 de julio de 2026 señala que tales herramientas rara vez miden su eficacia real.

No se exige al lector conocer el derecho de contratación de cada país. Cada sistema se describe a partir de los documentos abiertos del organismo que lo aplica.

Resumen ejecutivo

Los modelos de lenguaje llegaron al control de la contratación más tarde que las reglas y la estadística. A octubre de 2026, en ninguna de las cinco jurisdicciones se han convertido en la herramienta principal de verificación. El inventario cuenta seis sistemas en los que la aplicación de un modelo de lenguaje está confirmada por documentos del organismo y dos proyectos anunciados para los que no está confirmada o bien la tecnología (Italia), o bien la función de contratación (Oficina Europea de Lucha contra el Fraude). De los seis sistemas confirmados, dos se encuentran en operación, y solo respecto de uno el organismo ha publicado una resolución con resultados.

Datos y método. El estudio se apoya en los documentos de los organismos que aplican las herramientas: resoluciones y artículos técnicos del Tribunal de Cuentas de la Unión de Brasil y de los tribunales de cuentas de los estados, comunicados de prensa y documentación de concurso del Servicio de Contratación Pública de Corea, publicaciones de la Dirección de Compras y Contratación Pública de Chile (ChileCompra) y el registro de adjudicación del contrato en el sistema chileno de contratación Mercado Público, la directiva de la Autoridad Nacional Anticorrupción de Italia y los informes de la Comisión Europea y de la Oficina Europea de Lucha contra el Fraude. Para cada herramienta se completó una ficha de seis campos. Las funciones de las herramientas se cotejaron con los 30 mecanismos de la tipología del CILC según tres valores: detecta directamente, detecta indirectamente, no detecta. Los resultados se calcularon a partir de las cifras publicadas, con indicación del denominador.

Resultados principales. La única herramienta LLM respecto de la cual el organismo ha publicado una resolución con resultados es Alice 360 del Tribunal de Cuentas de la Unión de Brasil. En ella el modelo de lenguaje no lee el pliego en busca de infracciones, sino que resuelve una tarea preparatoria: reconoce el objeto de la contratación a partir de descripciones imprecisas y depura los datos, tras lo cual los precios se comparan con métodos estadísticos. En el ciclo de monitoreo 2025-2026 el tribunal dirigió requerimientos a las entidades contratantes respecto de 49 contrataciones, de las cuales 15 se iniciaron por alertas de Alice 360; en dos casos la entidad contratante, tras el requerimiento, redujo el valor estimado de la contratación o la retiró, y el tribunal computa esa reducción como efecto financiero del control.

El Servicio de Contratación Pública de Corea construye un sistema en el que el modelo de lenguaje verifica los anuncios conforme a 23 puntos legales; hasta su puesta en marcha, los anuncios se verifican de forma manual y por muestreo. En el primer trimestre de 2026 se verificaron 15 261 anuncios, respecto de 797 se enviaron requerimientos de corrección, y los fundamentos más frecuentes entre ellos fueron el incumplimiento del plazo de publicación (437 casos) y los requisitos de experiencia excesivos junto con las restricciones regionales (82 casos).

ChileCompra celebró en enero de 2026 un contrato para el análisis de documentos mediante un modelo de lenguaje conforme a 20 reglas. El sistema se encuentra en la etapa de ejecución del contrato, con puesta en marcha anunciada para el cuarto trimestre de 2026. La Autoridad Nacional Anticorrupción de Italia y la Oficina Europea de Lucha contra el Fraude llevan a cabo experimentos sin resultados publicados. El sistema VigIA del Tribunal de Cuentas del Estado de Santa Catarina somete desde abril de 2024 las preguntas de los auditores sobre el texto de cada anuncio al modelo GPT-4o; en el primer semestre de 2024 el tribunal informó de 215 anuncios respecto de los cuales se adoptaron medidas. La plataforma PRUMe AI del Tribunal de Cuentas del Estado de Amazonas sigue siendo un prototipo de investigación.

Mecanismos cubiertos. Las dos herramientas en operación cubren cinco mecanismos de la tipología de 30: cuatro en las etapas de pliego y publicación, donde el modelo de lenguaje lee el anuncio (VigIA), y uno

en la etapa de justificación del precio, donde la inflación del precio inicial la detecta un algoritmo estadístico a partir de los datos preparados por el modelo (Alice 360). Si se cuentan los sistemas en construcción de Corea, Chile y el distrito de Seocho, la cobertura potencial alcanza 10 mecanismos en cinco de las diez etapas: justificación del precio, pliego, publicación, presentación de ofertas y ejecución. Ninguna herramienta en operación verifica la presentación de ofertas, su evaluación, la impugnación, la ejecución ni el pago; la presentación de ofertas solo la aborda la herramienta en construcción de ChileCompra, y la ejecución del contrato solo el sistema en construcción del distrito de Seocho. El grupo de mecanismos «neutralización del control», para el que en el conjunto de datos del CILC la proporción de reclamaciones confirmadas es mayor que para los demás grupos, queda fuera del campo de visión de las herramientas en operación; de las que están en construcción, solo lo roza la verificación de la información obligatoria del anuncio en el sistema coreano. El modelo de lenguaje lee los mismos documentos que el participante externo en la licitación, y en el mismo momento, antes de la presentación de ofertas, por lo que el límite de su visibilidad coincide con el límite de visibilidad del participante en la tipología.

Conclusiones. Ningún organismo publica la proporción de alertas fundadas por función del modelo de lenguaje. Las únicas proporciones publicadas corresponden a herramientas sin modelo de lenguaje y van del 3,4 al 23,5 %. La ganancia que aporta el modelo de lenguaje frente a las reglas solo se describe en la resolución del TCU: el modelo reconoce el objeto de la contratación a partir de descripciones imprecisas, cosa que las reglas no saben hacer. Los indicios que buscan las herramientas de las cinco jurisdicciones, incluidas las herramientas sin modelo de lenguaje del anexo, arrojan cuatro candidatos a nuevos mecanismos de la tipología. Para monitorear la calidad del modelo de lenguaje y de su aplicación, el organismo debe recopilar al menos cuatro magnitudes por cada función: el número de alertas emitidas, el número de alertas verificadas por una persona, el número de alertas confirmadas y el número de alertas que dieron lugar a medidas. Sin esos datos, la eficacia del modelo de lenguaje es indistinguible de la eficacia de las reglas a las que se ha añadido.

1. Introducción

La detección automatizada de riesgos en la contratación pública se desarrolló durante dos decenios como un sistema de banderas rojas. Una bandera roja es un rasgo observable del procedimiento que está estadísticamente asociado a una u otra infracción: ya sea un participante único, un plazo breve de presentación de ofertas o una reducción mínima respecto del precio inicial del contrato. Sobre esa base se construyeron los sistemas en operación: el robot ALICE de la Contraloría General de la Unión de Brasil desde 2015, el sistema BRIAS de análisis de indicios de colusión en licitaciones de la Comisión de Comercio Justo de Corea desde 2006 y la herramienta de evaluación de riesgos Arachne de la Comisión Europea desde 2013.

Los modelos de lenguaje de gran tamaño permitieron leer documentos no estructurados: anuncios, especificaciones técnicas, actas. En ellos el modelo puede encontrar lo que las reglas aplicadas a los campos de datos no ven, por ejemplo un requisito redactado a la medida de un solo proveedor. Entre 2024 y 2026 los organismos de cinco jurisdicciones comenzaron a implantar funciones de ese tipo. El estudio establece qué implantaron exactamente, qué verifican esas funciones y qué se sabe de sus resultados a partir de los documentos de los propios organismos.

La metodología del CILC examina los 30 mecanismos más típicos de distorsión del procedimiento de contratación, distribuidos en diez etapas del ciclo de contratación (Limarau, 2026). La tipología separa el

mecanismo, es decir, el modo de distorsión, del indicador, es decir, el rasgo observable del procedimiento. Para cada mecanismo señala la posición de observación: quién puede advertirlo, el participante externo en la licitación, el órgano de auditoría o la persona interna. En el apartado 7 de la tipología se plantea la tarea de cotejar el catálogo de mecanismos con los campos de los datos abiertos y determinar qué mecanismos admiten detección automática. El presente estudio resuelve esa tarea para las herramientas basadas en modelos de lenguaje.

La segunda fuente del estudio es el conjunto de datos abierto del CILC sobre las reclamaciones de reclamantes externos contra procedimientos de contratación y sobre los resultados de su examen (Limarau, 2026; Krukovskiy, 2026). Para cada reclamación se indica en él si fue examinada en cuanto al fondo, si fue declarada fundada y si dio lugar a medidas. Esas mismas cuatro magnitudes se aplican aquí a las alertas de las herramientas (apartado 3.4).

2. Detección automatizada de riesgos en la contratación: de las banderas rojas a los modelos de lenguaje

El enfoque de las banderas rojas está sistematizado en las guías de la OCDE y del Banco Mundial. Un estudio para la Comisión Europea lo contrastó sobre 192 casos en ocho Estados (PwC, Ecorys, 2013). Un índice compuesto de riesgo de corrupción se calculó sobre más de 2,8 millones de contratos de la base europea de anuncios TED (Fazekas, Toth, 2016). Una rama aparte, el cribado económico de la colusión en licitaciones, detecta las ofertas de cobertura por la estructura de los precios (Porter, Zona, 1993; Imhof, Karagök, Rutz, 2018). La revisión del Comité de Competencia de la OCDE de 2022 resume el balance: los filtros aislados no funcionan en la mayoría de los casos, un conjunto de filtros funciona mejor, y la bandera por sí sola no prueba la infracción (OECD, 2022).

El aprendizaje automático sobre esos mismos indicadores está descrito en trabajos académicos. Un bosque aleatorio sobre datos italianos se contrastó con investigaciones y sentencias condenatorias (Decarolis, Giorgiantonio, 2020). Clasificadores basados en filtros se contrastaron sobre licitaciones suizas y japonesas (Huber, Imhof, 2019). En los sistemas estatales, el aprendizaje automático funciona en la Comisión Nacional de los Mercados y la Competencia de España (algoritmo BRAVA) y en la Contraloría General de la República de Colombia (modelo predictivo de 2025).

La revisión de la OCDE «Governing with Artificial Intelligence» describe cerca de 200 casos de aplicación de la inteligencia artificial en la administración pública, entre ellos en la contratación, y destaca ALICE en Brasil y el Observatorio ChileCompra (OECD, 2025). La revisión de Open Contracting Partnership de 2026 señala que las afirmaciones sobre el efecto de tales sistemas no han sido verificadas de manera independiente y que muchos proyectos no salen de la etapa piloto (OCP, 2026). Ninguna de las revisiones separa las funciones que ejecuta el modelo de lenguaje de las funciones que ejecutan las reglas. Esa separación constituye precisamente el objeto del presente estudio.

3. Datos y método

3.1. Selección de las herramientas

En el inventario se incluyen los sistemas que aplica un organismo público de una de las cinco jurisdicciones (federal o nacional, un órgano de la Unión Europea o un órgano subnacional, si la herramienta está documentada), que están destinados al control de los procedimientos de contratación y en los que al menos

una función la ejecuta un modelo de lenguaje. El inventario se divide en dos grupos. En el primero entran seis sistemas para los que tanto el modelo de lenguaje como la función de contratación están confirmados por documentos del organismo: una resolución, un artículo en una publicación del organismo, documentación de concurso o una declaración del organismo transmitida por la prensa. En el segundo entran dos proyectos anunciados para los que una de las dos condiciones no está confirmada: la Autoridad Nacional Anticorrupción de Italia informa de experimentos con inteligencia artificial sin nombrar la tecnología, y la Oficina Europea de Lucha contra el Fraude informa de un piloto de inteligencia artificial generativa sin vincularlo a la verificación de la contratación. El segundo grupo se describe para completar el panorama y no entra en los cálculos de los apartados 5 y 6. Las herramientas en las que el modelo de lenguaje sirve solo de asistente del funcionario y carece de función de contratación (ChatTCU en Brasil, LIA en la Contraloría General de la Unión de Brasil, «Qu'en dit la Cour» del Tribunal de Cuentas de Francia) no se incluyen en el inventario. Se mencionan en el apartado 4 a efectos de delimitación.

La búsqueda se realizó en octubre de 2026. Se examinaron documentos de los organismos (resoluciones, informes anuales, artículos técnicos, documentación de concurso, comunicados de prensa), documentos de la OCDE, del Tribunal de Cuentas Europeo y de Open Contracting Partnership, así como la prensa especializada en portugués, coreano, español, italiano e inglés. Todas las afirmaciones sobre las funciones de las herramientas se apoyan en documentos de los organismos. Las afirmaciones procedentes de fuentes secundarias están señaladas.

El apartado 7 va más allá del inventario de modelos de lenguaje: los indicios no descritos en la tipología se recopilaban para todas las herramientas de las cinco jurisdicciones, incluidas las herramientas sin modelo de lenguaje del anexo. Es un resultado complementario del estudio y se apoya en un conjunto de sistemas más amplio que los apartados 5 y 6.

3.2. Ficha de la herramienta

Para cada herramienta se completaron seis campos: órgano y año; datos y técnica, con indicación de qué función ejecuta el modelo de lenguaje y cuál las reglas o la estadística; la lista de indicios que busca la herramienta; el tratamiento de la alerta; los indicadores de resultados publicados; las evaluaciones independientes. Las fichas figuran en el apartado 4 y, para las herramientas sin modelo de lenguaje, en el anexo.

3.3. Cotejo con la tipología

Cada función de la herramienta se cotejó con los mecanismos de la tipología del CILC. Una función detecta un mecanismo directamente si está dirigida a la huella de ese mecanismo en particular: por ejemplo, la verificación de los requisitos de calificación por exceso está dirigida al mecanismo SPC-02. Una función detecta un mecanismo indirectamente si el indicador que encuentra, según la descripción de la tipología, corresponde también a ese mecanismo y no solo al que apunta la función: por ejemplo, un plazo de presentación de ofertas acortado artificialmente (PUB-01) puede aparecer entre las respuestas a las preguntas de auditoría sobre el anuncio, aunque las preguntas se hayan redactado para verificar los requisitos a los participantes. No se asigna la relación indirecta si el indicador solo podría coincidir con la huella del mecanismo bajo supuestos adicionales: así, un precio superior al de mercado no se considera huella de la subvaloración del precio en previsión de modificaciones del contrato (BUD-02) ni de las ofertas de cobertura (BID-01), puesto que la tipología describe esos mecanismos mediante otros indicadores. El resultado se resume en la matriz «mecanismo x herramienta» (tabla 3).

3.4. Indicadores de resultados

Los resultados de la herramienta se evalúan mediante cuatro magnitudes: la cobertura, es decir, la proporción de procedimientos que pasan por la herramienta; la proporción verificada, es decir, la proporción de alertas que verificó una persona; la proporción confirmada, es decir, la proporción de alertas que el verificador declaró fundadas; y la proporción con consecuencia, es decir, la proporción de alertas confirmadas a las que siguió una corrección, una suspensión, una anulación o una sanción. Con esas mismas magnitudes el CILC describe los resultados del examen de las reclamaciones de reclamantes externos en su conjunto de datos. Para cada herramienta se indica cuáles de las cuatro magnitudes pueden calcularse a partir de los datos publicados. Los indicadores que no se ajustan a esas definiciones (el número de requerimientos a las entidades contratantes, el número de anuncios con medidas, la exactitud de extracción de datos F1) se presentan por separado y no se equiparan a las cuatro magnitudes. Todas las cifras son autodeclaradas. No se encontraron evaluaciones independientes de la exactitud de ninguna herramienta LLM.

3.5. Terminología

Los términos siguen la tipología del CILC: mecanismo, indicador de riesgo, posición de observación (participante externo, órgano de auditoría, persona interna). Se denomina indicio de restricción de la competencia al indicador que el participante externo puede ver en el pliego. Una alerta es el mensaje que la herramienta emite a una persona respecto de un procedimiento. La tasa de confirmación es la proporción de reclamaciones declaradas fundadas tras el examen en cuanto al fondo (Krukovskiy, 2026). Aquí se aplica a las alertas.

4. Herramientas basadas en modelos de lenguaje de gran tamaño en cinco jurisdicciones

La tabla 1 resume el inventario: seis sistemas con modelo de lenguaje confirmado y dos proyectos anunciados sin tal confirmación. A continuación se presenta la ficha de cada herramienta.

Tabla 1. Herramientas LLM de control de la contratación, octubre de 2026

Jurisdicción	Herramienta y órgano	Función del modelo de lenguaje	Confirmación	Etapas	Resultados publicados
Brasil	Alice 360, Tribunal de Cuentas de la Unión (TCU)	Reconocimiento del objeto de la contratación a partir de descripciones imprecisas, depuración de datos antes de la comparación de precios	Resolución del tribunal	Operación, ciclo 2025-2026	Acórdão 1145/2026: 15 de 49 requerimientos a entidades contratantes, 2 con reducción del valor de la contratación
Brasil	VigIA, Tribunal de Cuentas del Estado de Santa Catarina (TCE-SC)	Respuestas a las preguntas de los auditores sobre el texto del anuncio antes de la publicación (GPT-4o)	Artículo en la revista del tribunal	Operación desde abril de 2024	7 711 anuncios analizados, 215 con medidas entre abril y octubre de 2024

Jurisdicción	Herramienta y órgano	Función del modelo de lenguaje	Confirmación	Etapas	Resultados publicados
Brasil	PRUMe AI, Tribunal de Cuentas del Estado de Amazonas (TCE-AM)	Extracción de campos y cláusulas de contratos y anuncios	Artículo académico de funcionarios del tribunal	Prototipo	F1 de 0,86 sobre una submuestra etiquetada de 20 documentos
Corea	Monitoreo de anuncios, Servicio de Contratación Pública (PPS)	Verificación del anuncio conforme a 23 puntos legales y de su coherencia	Documentación de concurso, comunicado de prensa	Desarrollo, puesta en marcha anunciada para el segundo semestre de 2026	Antes de la puesta en marcha: verificación manual por muestreo, 797 requerimientos sobre 15 261 anuncios en el primer trimestre de 2026
Corea	Verificación de documentos contractuales, administración del distrito de Seocho (Seúl)	Verificación de la completitud, la coherencia y la conformidad con las normas	Declaración de la administración en la prensa	Desarrollo, plazo septiembre de 2026	no hay
Chile	Análisis de documentos, ChileCompra	20 reglas sobre documentos no estructurados; clasificación de denuncias confidenciales	Documentación de concurso, publicaciones de ChileCompra	Ejecución del contrato, puesta en marcha anunciada para el cuarto trimestre de 2026	no hay
Italia	Monitoreo de anuncios, Autoridad Nacional Anticorrupción (ANAC)	Detección de condiciones anticompetitivas en los anuncios	No confirmado: la tecnología no se nombra	Experimento (directiva de 2026)	no hay
UE	GETI con IA generativa, Oficina Europea de Lucha contra el Fraude (OLAF)	Integración de indicadores de riesgo en una herramienta de IA generativa	No confirmado: la función de contratación no se nombra	Piloto de 2025	no hay

4.1. Brasil: Alice 360 del Tribunal de Cuentas de la Unión

Órgano y año. Tribunal de Cuentas de la Unión de Brasil (Tribunal de Contas da União, en adelante TCU), unidades de auditoría digital y de auditoría de contrataciones. Alice 360 continúa la línea de la herramienta ALICE, cuyo código el TCU recibió de la Contraloría General de la Unión en 2016. La versión basada en un modelo de lenguaje se introdujo en el ciclo de monitoreo continuo de las contrataciones federales de junio de 2025 a marzo de 2026. Los resultados del ciclo fueron aprobados por la resolución del pleno del TCU Acórdão 1145/2026, de 6 de mayo de 2026 (expediente TC 009.250/2025-5).

Datos y técnica. La primera versión de la herramienta era una pantalla de consulta. El auditor introducía la descripción del objeto, la unidad de suministro y el método estadístico (rango intercuartílico, puntuación z y otros), el sistema buscaba en la base de contratos celebrados y devolvía las desviaciones respecto del precio de mercado (punto 78.1.1 de la resolución). Tras obtener acceso a plataformas de modelos de lenguaje, «en especial Claude», la herramienta, según la resolución, «comenzó a utilizar la IA de manera verdaderamente intensiva» para reconocer y analizar los objetos de contratación. Ahora genera a diario

informes sobre contrataciones y objetos con posible sobreprecio (punto 78.1.2). El modelo de lenguaje ayuda a «depurar y procesar datos e información incorrecta» y reconoce el objeto «incluso con descripciones imprecisas» (puntos 75.2 y 11 del voto del relator). La parte estadística: la versión actual utiliza la media recortada, y el panel Alice 360 utiliza la mediana tras eliminar los valores atípicos (puntos 78.1.3, 78.1.5). Los datos proceden del sistema federal de contratación Compras.gov.br y del Portal Nacional de Contrataciones Públicas (PNCP). La lectura de anuncios y especificaciones técnicas por el modelo de lenguaje no se describe en la resolución; en otros documentos del TCU, Alice 360 tampoco tiene esa función.

Indicios. Precio del objeto de la contratación que se desvía del de mercado. Está prevista la ampliación a los vínculos de los participantes con funcionarios y a las ofertas con indicios de inviabilidad (punto 78.3.1).

Tratamiento de la alerta. El panel Alice 360, incluido el panel de medicamentos, está disponible para los auditores desde octubre de 2025. La alerta se convierte en fundamento de un requerimiento a la entidad contratante (diligência) y, en caso de confirmación, de una orden vinculante de corrección. La Plataforma de Alertas, Riesgos y Tipologías (PARTS) debe entregar las alertas a las herramientas de trabajo de los auditores del TCU y de los tribunales de cuentas estatales asociados (punto 78.4).

Resultados. En el ciclo 2025-2026 el tribunal dirigió requerimientos a las entidades contratantes respecto de 49 contrataciones. En 15 de ellas el fundamento fue una alerta del panel Alice 360. En dos casos la entidad contratante, tras el requerimiento, redujo el valor estimado de la contratación o la retiró. El TCU computa esa reducción como efecto financiero del control (benefício financeiro). En otros dos casos se espera una reducción (puntos 78.1.4, 82). Para las 49 contrataciones en conjunto, el valor estimado se volvió a publicar tras los requerimientos por R\$ 105,6 millones menos que el inicial. Los importes de las dos contrataciones detectadas por Alice 360 no figuran en la parte legible de la resolución. El único ejemplo de retiro mencionado: el ítem 233 (lenalidomida) de la licitación electrónica 90044/2025 del Comando de la Marina (punto 64.1). El número de alertas y la proporción confirmada no figuran en la resolución. La resolución señala que el desarrollo y las pruebas de Alice 360 restaron al equipo un tiempo que de otro modo se habría dedicado al análisis de contrataciones (punto 78.2). La propuesta de la unidad de recomendar la continuación del proyecto no fue adoptada por el pleno, que se limitó a encargar la coordinación de las iniciativas tecnológicas (punto 9.2 de la parte dispositiva).

Calidad de los datos. La resolución registra 502 contrataciones federales con un valor estimado superior a R\$ 100 millones cuyo valor real resultó al menos cien veces menor. La incoherencia de los datos, según la resolución, «llevaba regularmente al equipo a falsos positivos» (puntos 73, 74).

Evaluaciones independientes. No hay evaluaciones independientes. El artículo de auditores del TCU sobre la plataforma PARTS (Santos et al., 2026) describe el paso del aprendizaje automático a la inteligencia artificial generativa como una forma de reducir la necesidad de especialistas en entrenamiento de modelos, pero no contiene indicadores de exactitud.

4.2. Brasil: VigIA del Tribunal de Cuentas del Estado de Santa Catarina y PRUMe AI del Tribunal de Cuentas del Estado de Amazonas

El sistema VigIA. Órgano: Tribunal de Cuentas del Estado de Santa Catarina (Tribunal de Contas do Estado de Santa Catarina, en adelante TCE-SC). Funciona desde abril de 2024 con el apoyo científico de la Universidad Federal de Santa Catarina. La técnica se describe en un artículo de funcionarios del tribunal en su propia revista (Albuquerque, Wensing, Joppi, 2024, p. 156): las preguntas de auditoría se ejecutan de

manera automática, sin intervención humana, sobre cada anuncio recibido mediante la integración con el modelo generativo GPT-4o. El sistema recibe los anuncios antes de su publicación a través del sistema e-Sfinge, unos 250 al día; las preguntas las registran los auditores. En los dos primeros meses, según el mismo artículo, se procesaron más de 2 000 procedimientos y se ejecutaron más de 10 000 preguntas. Entre el 18 de abril y el 8 de octubre de 2024 se recibieron 33 507 anuncios, se analizaron 7 711 y se generaron 63 445 respuestas del modelo. Respecto de 215 anuncios se adoptaron medidas: corrección, retiro, anulación o suspensión, por un valor cercano a R\$ 2 000 millones (TCE-SC, 2024). Cobertura: el 23 % de los anuncios recibidos. Proporción de anuncios con medidas entre los analizados: 2,8 %; no es la proporción de alertas confirmadas, puesto que no se publican el número de alertas ni el número de respuestas verificadas. Los autores del artículo señalan que la construcción de métricas de exactitud sigue siendo una tarea pendiente.

La plataforma PRUMe AI. Órgano: Tribunal de Cuentas del Estado de Amazonas (Tribunal de Contas do Estado do Amazonas, en adelante TCE-AM). La plataforma se describe en un artículo académico (Bezerra, Lopes, 2025) como un sistema de control externo basado en modelos de lenguaje. El modelo extrae datos de los documentos por el método RAG (generación con recuperación de fuentes) y registra de dónde procede cada valor. Muestra del estudio: 150 documentos reales (anuncios, contratos y modificaciones, informes y dictámenes), de los cuales el 55 % son PDF digitales y el 45 % documentos escaneados. El tiempo de clasificación se redujo de 21,4 a 7,9 minutos por documento, y el tiempo de análisis completo de 39,2 a 17,8 minutos. La exactitud de extracción (F1) se calculó sobre una submuestra etiquetada por expertos de 20 documentos: 0,86 para los campos del contrato y 0,82 para las cláusulas. El artículo no nombra el modelo concreto. La operación productiva no está confirmada. Es la única herramienta del inventario con métricas de exactitud publicadas, y estas se refieren a la extracción de datos sobre 20 documentos, no a la detección de infracciones.

4.3. Corea: monitoreo de anuncios del Servicio de Contratación Pública

Órgano y año. Servicio de Contratación Pública (Public Procurement Service, en adelante PPS). El sistema de monitoreo de las contrataciones autónomas basado en inteligencia artificial se desarrolla en el marco de un programa del Ministerio del Interior y Seguridad. La puesta en marcha productiva está anunciada para el segundo semestre de 2026 (ETNews, 2026). Base legal: la reforma de la Ley de Contratación Electrónica, que entró en vigor el 23 de diciembre de 2025 y otorgó al PPS la facultad de exigir a las entidades contratantes la corrección de los anuncios en caso de infracción manifiesta de la ley.

Datos y técnica. Como entrada se reciben los anuncios, las especificaciones y los términos de referencia de 16 sistemas electrónicos de contratación, unos 2 500 al día. El modelo de lenguaje verifica el texto en busca de probables infracciones de la Ley de Contratos del Estado y de la Ley de Contratos de las Entidades Locales y asigna una clase de riesgo según el tipo de infracción. Un funcionario revisa la alerta, el PPS envía a la entidad contratante un requerimiento de corrección y el resultado se devuelve a los datos de entrenamiento. El contenido de la verificación se reveló en el concurso abierto para la creación del modelo (DACON, agosto-octubre de 2026, organizado por el Ministerio del Interior y Seguridad y la Agencia Nacional de la Sociedad de la Información, con el PPS como entidad solicitante). Los participantes del concurso deben determinar las infracciones conforme a 24 puntos, de los cuales 23 se refieren a las normas sobre contratos estatales y locales, sobre el apoyo a los productos de la pequeña empresa y sobre el software, y uno verifica la coherencia del texto del anuncio con los valores registrados en el sistema. El modelo es fijo para todos los participantes (Gemma 4 26B con embeddings bge-m3, sin ajuste fino), los datos constan de 20 000 anuncios sin etiquetar y 200 etiquetados, con un conjunto de prueba de 1 853 anuncios, y la

calidad se evalúa por el F1 medio sobre los 24 puntos (DACON, 2026). El modelo de la versión productiva no se nombra.

Indicios. Errores en las condiciones de admisión, requisitos de experiencia excesivos, reconocimiento de la experiencia solo de determinadas organizaciones, restricciones regionales aplicadas de forma incorrecta, incumplimiento del plazo de publicación, superación de los límites de los criterios de evaluación de las propuestas, ausencia de información obligatoria.

Tratamiento de la alerta. Clasificación por riesgo, verificación única por un funcionario, requerimiento de corrección a la entidad contratante.

Resultados. Hasta la puesta en marcha del sistema, los anuncios se verifican de forma manual y por muestreo. En el primer trimestre de 2026 se verificaron 15 261 anuncios y respecto de 797 se enviaron requerimientos de corrección, es decir, en el 5,2 %. Fundamentos: 437 por incumplimiento del plazo de publicación, 113 por errores en la aplicación de la ley de apoyo a la pequeña empresa, 82 por requisitos de experiencia excesivos y restricciones regionales, 65 por superación de los límites de los criterios de evaluación (PPS, 2026). Entre enero y mayo de 2026 se verificaron unos 30 000 anuncios y se enviaron unos 1 250 requerimientos, de los cuales 1 207 fueron aceptados por las entidades contratantes (96,6 %). Esas cifras describen la verificación manual que el modelo de lenguaje debe sustituir. Para él sirven de línea de base: la distribución de las infracciones por tipos y la proporción de requerimientos aceptados.

4.4. Corea: verificación de documentos contractuales en el distrito de Seocho

La administración del distrito de Seocho de la ciudad de Seúl desarrolla desde abril de 2026 un sistema de verificación de documentos contractuales. Un modelo de lenguaje con recuperación de fuentes está alojado en la infraestructura propia del distrito. El sistema verifica la coherencia de los datos, la presencia de las cláusulas obligatorias, la conformidad con las leyes y las normas internas y la fundamentación de los cálculos. Para cada punto emite un dictamen de «conforme», «no conforme» o «advertencia», con referencia a la fuente. El plazo de finalización del desarrollo es septiembre de 2026 (Herald Economy, 2026). La puesta en operación a octubre de 2026 no está confirmada. La herramienta corresponde a la etapa de ejecución del contrato, que las demás herramientas LLM no cubren.

4.5. Chile: análisis de documentos de ChileCompra

Órgano y año. Dirección de Compras y Contratación Pública de Chile (ChileCompra), departamento Observatorio. Base legal: la Ley 19.886 en la redacción de la Ley 21.634 (2023). Esta otorgó a ChileCompra la facultad de realizar cruces automatizados y aplicar inteligencia artificial, y la obligó a remitir los antecedentes sobre infracciones a la Contraloría General de la República, al Ministerio Público y al órgano de defensa de la competencia.

Contrato. La licitación 869591-2-LP25 «Análisis de Documentos mediante modelos de IA» se publicó el 14 de agosto de 2025; de diez participantes, el contrato fue adjudicado el 9 de enero de 2026 a la empresa Sociedad de Profesionales X-Consulting Limitada, por un monto de CLP 160 millones (registro de Mercado Público). Objeto del contrato: un procedimiento automatizado de detección de anomalías conforme a 20 reglas de incumplimiento de las normas de contratación. El procedimiento se aplica a los documentos de texto no estructurados de Mercado Público, históricos y actuales, unos 200 000 procesos al año. El contrato prevé la transferencia de competencias al equipo del Observatorio. En diciembre de 2025 el proyecto pasó una evaluación de impacto algorítmico según la metodología del GobLab de la Universidad Adolfo Ibáñez y fue clasificado en el nivel de impacto alto (ChileCompra, 2025).

Funciones del modelo de lenguaje según las publicaciones de ChileCompra. En publicaciones de distintos meses, ChileCompra describe cuatro funciones: el análisis de las bases de licitación, los anexos y los actos administrativos en busca de incoherencias, restricciones indebidas e infracciones de la lógica de las normas (noviembre de 2025); el análisis de los documentos PDF adjuntos, con reconocimiento de texto, en busca de posibles indicios de colusión y sobreprecios (julio de 2026); el análisis de las bases de licitación, las especificaciones técnicas y los contratos para detectar 20 nuevos tipos de infracciones (septiembre de 2026); la clasificación de denuncias confidenciales (noviembre de 2025).

Etapas. A octubre de 2026 la herramienta se encuentra en la etapa de ejecución del contrato. ChileCompra la describe como un piloto y espera su implantación en el cuarto trimestre de 2026 (ChileCompra, julio de 2026). No hay resultados. A efectos de comparación: el mecanismo en operación del Observatorio basado en reglas aplica 15 reglas a todos los procedimientos de Mercado Público. Entre agosto y noviembre de 2025 emitió más de 6 199 alertas, y respecto del 54 % de ellas los organismos adoptaron medidas (anexo, tabla A1).

4.6. Italia: monitoreo de anuncios de la Autoridad Nacional Anticorrupción

La Autoridad Nacional Anticorrupción de Italia (Autorità Nazionale Anticorruzione, en adelante ANAC) adoptó el 26 de mayo de 2026 la directiva programática sobre la actividad de supervisión para 2026. La directiva prevé la creación, «entre otros medios, mediante experimentos con inteligencia artificial», de un sistema estructurado de monitoreo de anuncios. Finalidad del monitoreo: detectar condiciones anticompetitivas e infracciones graves de las normas sectoriales para que la ANAC pueda impugnar los anuncios conforme al artículo 220 del Código de Contratos Públicos (ANAC, 2026). La tecnología, el ejecutor, los plazos y los resultados no se indican en la directiva. El informe anual de la ANAC correspondiente a 2025 no contiene indicadores. La aplicación de un modelo de lenguaje no está confirmada por documentos de la ANAC, por lo que el proyecto se asigna al segundo grupo del inventario y no entra en los cálculos de los apartados 5 y 6.

4.7. Unión Europea: la Oficina Europea de Lucha contra el Fraude y Arachne+

La Oficina Europea de Lucha contra el Fraude (en adelante, OLAF). La información procede del informe de la Comisión Europea sobre la ejecución de la Estrategia de Lucha contra el Fraude (SWD(2026) 246). La herramienta analítica de la OLAF denominada GETI busca simultáneamente en datos estructurados y no estructurados en distintos idiomas. Se ha complementado con un conjunto de indicadores de riesgo para la detección proactiva de infracciones. A mayo de 2026 se habían elaborado seis informes analíticos. Durante 2025 se trabajó en un piloto de integración de los indicadores de riesgo en una herramienta de inteligencia artificial generativa. Los resultados del piloto no se han publicado. El informe de la OLAF correspondiente a 2025 informa de nuevas herramientas de inteligencia artificial para tres tareas: la detección de comunicaciones relacionadas con fraude en grandes volúmenes de datos, la anonimización de comunicaciones y la gestión de los documentos de los expedientes. La técnica no se indica y la relación con la contratación no se revela. Ninguno de los documentos vincula el piloto generativo con la verificación de los procedimientos de contratación, por lo que el proyecto se asigna al segundo grupo del inventario y no entra en los cálculos de los apartados 5 y 6.

La herramienta Arachne+. Es la herramienta de la Comisión Europea para la evaluación de riesgos en los proyectos financiados con fondos de la UE. Se puso en funcionamiento el 27 de abril de 2026 y está construida sobre 78 indicadores predefinidos. Según la descripción de la Comisión, la herramienta «no

adopta decisiones automáticas». La aplicación de la inteligencia artificial se asigna a la tercera etapa de desarrollo y a la agenda del grupo de expertos para 2026 (European Commission, 2026). Arachne+ no utiliza un modelo de lenguaje, por lo que la herramienta se describe en el anexo.

4.8. Colombia y otros casos sin modelo de lenguaje confirmado

En Colombia ninguna herramienta de control de la contratación utiliza un modelo de lenguaje, a juzgar por los documentos de los propios organismos. El modelo predictivo de la Contraloría General de la República (2025) está construido sobre árboles de decisión y bosque aleatorio. El sistema HELIOS del Ministerio de Defensa Nacional (agosto de 2026) se describe como aprendizaje automático sobre los datos estructurados del sistema de contratación SECOP. El sistema de alertas tempranas de la agencia de contratación Colombia Compra Eficiente (2025) consta de cuatro reglas. El centro de analítica avanzada de la Procuraduría General de la Nación (mayo de 2026) se creó para la vigilancia de las elecciones. La Circular 023 de la Contraloría General de la República, de 4 de diciembre de 2025, permite a los funcionarios utilizar la inteligencia artificial solo como medio auxiliar y prohíbe fundar las conclusiones exclusivamente en ella.

En Brasil el robot ALICE de la Contraloría General de la Unión funciona con expresiones regulares, cruces SQL y clasificadores de aprendizaje automático. No utiliza un modelo de lenguaje. La plataforma generativa de la misma contraloría, LIA (2025), no está vinculada a la verificación de la contratación. En Corea el sistema BRIAS de la Comisión de Comercio Justo sigue siendo un sistema de reglas ponderadas. Su modernización de 2026 añade la búsqueda vectorial de similitud por grupos de productos, pero no un modelo de lenguaje (KFTC, 2026). La herramienta BRAVA de la Comisión Nacional de los Mercados y la Competencia de España, el modelo del Servicio de Investigaciones Especiales de Lituania (2026) y la herramienta de la Autoridad de Integridad de Hungría están contruidos sobre aprendizaje automático con datos estructurados.

5. Qué mecanismos cubren los modelos de lenguaje

5.1. Funciones y mecanismos

La tabla 2 descompone el trabajo de cada herramienta LLM en funciones separadas y vincula cada función con los mecanismos de la tipología.

Tabla 2. Funciones de los modelos de lenguaje en las herramientas de control y mecanismos de la tipología del CILC

Herramienta	Estado	Función del modelo de lenguaje	Quién detecta la infracción	Mecanismo, directamente	Mecanismo, indirectamente	Resultados de la función
Alice 360, TCU	en operación	Reconocimiento del objeto de la contratación y depuración de datos para la comparación de precios	algoritmo estadístico sobre datos preparados por el modelo	BUD-01 (inflación del precio inicial)		No se publican por función; 15 de 49 requerimientos a entidades contratantes, 2 con reducción del valor

Herramienta	Estado	Función del modelo de lenguaje	Quién detecta la infracción	Mecanismo, directamente	Mecanismo, indirectamente	Resultados de la función
Alice 360, TCU	plan	Vínculos de los participantes con funcionarios; ofertas inviables	no se revela	BID-02, EVL-04		no implantado
VigIA, TCE-SC	en operación	Respuestas a las preguntas de los auditores sobre el anuncio antes de la publicación	modelo de lenguaje	SPC-01, SPC-02, SPC-05 (según el carácter de las preguntas)	PUB-01	215 anuncios con medidas de 7 711 (2,8 %); las alertas y las confirmaciones no se publican
PPS, Corea	en construcción	Condiciones de admisión y requisitos de experiencia	modelo de lenguaje	SPC-02	SPC-01	Línea de base de la verificación manual: 82 de 797 requerimientos en el trimestre
PPS, Corea	en construcción	Restricciones regionales	modelo de lenguaje	SPC-02		incluidas en los mismos 82
PPS, Corea	en construcción	Plazo de publicación	modelo de lenguaje	PUB-01		437 de 797
PPS, Corea	en construcción	Límites de los criterios de evaluación de las propuestas	modelo de lenguaje	SPC-03		65 de 797
PPS, Corea	en construcción	Aplicación de la ley de apoyo a la pequeña empresa	modelo de lenguaje	SPC-02		113 de 797
PPS, Corea	en construcción	Coherencia del texto con los valores registrados	modelo de lenguaje	PUB-02	SPC-05	no se publican
ChileCompra	en construcción	Incoherencia de las bases de licitación y los actos, restricciones indebidas	modelo de lenguaje	SPC-05, SPC-01, SPC-02		no implantado
ChileCompra	en construcción	Indicios de colusión y sobrepregios en PDF	modelo de lenguaje	BID-01, BUD-01		no implantado

Herramienta	Estado	Función del modelo de lenguaje	Quién detecta la infracción	Mecanismo, directamente	Mecanismo, indirectamente	Resultados de la función
ChileCompra	en construcción	Clasificación de denuncias confidenciales	modelo de lenguaje	ninguno (tratamiento de denuncias, no detección)		no implantado
Seocho, Seúl	en construcción	Compleitud y coherencia de los documentos contractuales, fundamentación de los cálculos	modelo de lenguaje	IMP-01	IMP-02	no implantado
ANAC	segundo grupo	Condiciones anticompetitivas de los anuncios	no se revela	SPC-01, SPC-02	PUB-01	experimento; no entra en los cálculos
OLAF	segundo grupo	Indicadores de riesgo en una herramienta generativa	no se revela	no se revela		piloto; no entra en los cálculos

5.2. Matriz de cobertura

La tabla 3 muestra cuáles de los 30 mecanismos están cubiertos por las herramientas del primer grupo con composición de funciones conocida: Alice 360 del TCU, VigIA, PPS, ChileCompra y el sistema del distrito de Seocho. Las notaciones distinguen quién detecta la infracción y en qué estado se encuentra la herramienta: las letras D e I significan detección directa e indirecta por el modelo de lenguaje en una herramienta en operación; E, detección por un algoritmo estadístico sobre datos preparados por el modelo de lenguaje; D* e I*, detección directa e indirecta por el modelo de lenguaje en una herramienta en construcción; y la marca «plan», una función anunciada pero no desarrollada.

Tabla 3. Cobertura de los mecanismos de la tipología del CILC por las herramientas LLM

Código	Mecanismo	TCU	VigIA	PPS	ChileCompra	Seocho
PLN-01	Necesidad ficticia o inflada					
PLN-02	Acuerdo de los requisitos con el futuro adjudicatario					
BUD-01	Inflación del precio inicial	E			D*	
BUD-02	Subvaloración del precio en previsión de modificaciones del contrato					
PRC-01	Contratación injustificada con un solo proveedor					
PRC-02	Fraccionamiento del volumen					
PRC-03	Abuso de la urgencia					
SPC-01	Requisitos a la medida de un participante predeterminado		D	I*	D*	
SPC-02	Requisitos de calificación excesivos		D	D*	D*	

Código	Mecanismo	TCU	VigIA	PPS	ChileCompra	Seocho
SPC-03	Criterios de alto peso y valoración indeterminada			D*		
SPC-04	Garantías onerosas					
SPC-05	Requisitos indeterminados sobre el contenido de la oferta		D	I*	D*	
PUB-01	Plazo de presentación acortado artificialmente		I	D*		
PUB-02	Publicación sin los identificadores obligatorios			D*		
PUB-03	Modificación del pliego al final del plazo					
BID-01	Ofertas de cobertura				D*	
BID-02	Vinculación oculta	plan				
EVL-01	Rechazo discriminatorio de ofertas					
EVL-02	Rechazo sin motivación					
EVL-03	Asignación de puntos con infracción del procedimiento					
EVL-04	Conflicto de intereses en la comisión	plan				
RVW-01	Cierre del plazo de impugnación antes de los resultados					
RVW-02	Devolución de la reclamación por motivos formales					
RVW-03	Negativa de la segunda instancia					
RVW-04	Contrato antes de concluir el examen					
IMP-01	Modificaciones del contrato con aumento del precio					D*
IMP-02	Recepción de obras no ejecutadas					I*
IMP-03	No aplicación de sanciones al proveedor					
PAY-01	Pago contra documentos ficticios					
PAY-02	Venta del orden de los pagos					

La cobertura efectiva la dan las dos herramientas en operación: cinco mecanismos de 30. El modelo de lenguaje detecta directamente tres de ellos (SPC-01, SPC-02 y SPC-05 en VigIA) e indirectamente uno (PUB-01). El quinto, la inflación del precio inicial (BUD-01), lo detecta en Alice 360 un algoritmo estadístico de comparación de precios; el modelo de lenguaje solo le prepara los datos, y su contribución a la detección no se ha medido por separado. La desviación del precio al alza respecto del de mercado no se asigna a las huellas de la subvaloración del precio (BUD-02) ni de las ofertas de cobertura (BID-01); la tipología describe esos mecanismos mediante otros indicadores, el aumento del precio por modificaciones del contrato y la estructura de las ofertas, que Alice 360 no analiza. La cobertura potencial, contando los

sistemas en construcción de Corea, Chile y el distrito de Seocho, es de 10 mecanismos en cinco etapas, de ellos 9 directamente. Es una estimación de las funciones anunciadas, no un resultado alcanzado.

Por grupos, la cobertura potencial es desigual. La restricción del acceso (grupo C2) está cubierta casi por completo: SPC-01, SPC-02 y PUB-01 están cubiertos, SPC-04 y PUB-03 no. Del grupo de distorsión de la elección (C4) las herramientas solo ven el pliego, SPC-03 y SPC-05; la etapa de evaluación de las ofertas queda fuera de su campo. La simulación de la competencia (C3) no la cubren las herramientas en operación; los indicios de colusión en los documentos solo están anunciados en la herramienta en construcción de ChileCompra. La extracción del beneficio fuera de la fase competitiva (C6) se toca en dos puntos: el precio inicial (BUD-01, Alice 360 en operación) y la ejecución del contrato (IMP-01 e IMP-02, solo el sistema en construcción del distrito de Seocho). La evitación de la competencia (C1) no está cubierta por nada. Del grupo de neutralización del control (C5), el sistema en construcción del PPS cubre solo la publicación sin la información obligatoria (PUB-02); la etapa de impugnación no está cubierta por ninguna herramienta.

5.3. Coincidencia con el límite de visibilidad del participante externo

La tipología divide el ciclo de contratación en tres segmentos según la posición de observación. El participante externo en la licitación ve el segmento central: del pliego a la impugnación. Las herramientas LLM, en operación y en construcción, leen los mismos documentos que el participante, y en el mismo momento, antes de la presentación de ofertas. Por eso su cobertura recae en ese mismo segmento central, y además en su primera mitad. Las etapas de evaluación de las ofertas y de impugnación también las ve el participante, pero las herramientas no las leen: las actas de las comisiones y las resoluciones de los órganos de impugnación no forman parte de sus datos de entrada. Los segmentos inicial y final, visibles para el órgano de auditoría y para la persona interna, las herramientas LLM casi no los cubren. En el segmento inicial hay un punto, la justificación del precio inicial (BUD-01), que la tipología asigna a la auditoría y en parte al participante externo; lo cubre Alice 360, y lo cubre con estadística, no con la lectura de documentos. En el segmento final solo el sistema en construcción del distrito de Seocho lee los documentos contractuales y entra con ello en el campo de la auditoría.

El modelo de lenguaje no amplió el conjunto de mecanismos accesibles a la observación en comparación con el participante externo. Cambió la escala de la observación: el participante lee procedimientos aislados, el modelo lee todos los anuncios del flujo. El grupo «neutralización del control» (C5), para el que en el conjunto de datos del CILC la proporción de reclamaciones confirmadas es mayor que para los demás grupos, queda fuera del campo de visión de las herramientas en operación, y de las que están en construcción solo lo roza la verificación de la información obligatoria del anuncio (PUB-02) en el sistema coreano. Ninguna herramienta verifica los plazos de impugnación, las devoluciones de reclamaciones ni las actuaciones de la segunda instancia.

6. Qué se sabe sobre los resultados

La tabla 4 muestra qué indicadores de resultados pueden calcularse a partir de los datos publicados.

Tabla 4. Indicadores de resultados de las herramientas LLM según los datos publicados

Herramienta	Cobertura	Proporción verificada	Proporción confirmada	Proporción con consecuencia	Otros indicadores publicados
Alice 360, TCU	Contrataciones desde R\$ 10 millones, sin obras ni TI compleja; la proporción del flujo no se publica	no se publica	no se publica	no se publica	15 de los 49 requerimientos a entidades contratantes del ciclo se iniciaron por alertas; 2 de los 15 requerimientos con reducción del valor o retiro de la contratación
VigIA, TCE-SC	7 711 de 33 507 anuncios recibidos (23 %)	no se publica	no se publica	no se publica	215 anuncios con medidas de 7 711 analizados (2,8 %)
PPS, Corea (verificación manual antes de la puesta en marcha)	15 261 anuncios en el trimestre; la proporción del flujo no se publica	no aplicable: no hay alertas, la verificación es manual	no aplicable	no aplicable	797 requerimientos de corrección (5,2 % de los verificados); 96,6 % de los requerimientos aceptados por las entidades contratantes (enero-mayo)
ChileCompra	anunciado el 100 % de los procedimientos	no implantado	no implantado	no implantado	evaluación de impacto algorítmico (GobLab UAI) realizada, resultados no revelados
Seocho, Seúl		no implantado			no hay

Las métricas de exactitud de PRUMe AI (F1 de 0,86 y 0,82) no se incluyen en la tabla: miden la calidad de la extracción de datos sobre una submuestra etiquetada de 20 documentos y no corresponden a ninguna de las cuatro magnitudes del apartado 3.4. Las dos herramientas del segundo grupo, ANAC y OLAF, no publican indicadores.

Ninguna de las cuatro magnitudes del apartado 3.4 puede calcularse a partir de los datos publicados para ninguna herramienta LLM; la cobertura solo se conoce para VigIA. La proporción de alertas confirmadas es el indicador clave, y no lo ha publicado ningún organismo que aplique una herramienta LLM. Las cifras publicadas por el TCU y el TCE-SC caracterizan otra cosa: la proporción de requerimientos a las entidades contratantes iniciados por alertas y la proporción de anuncios analizados respecto de los cuales se adoptaron medidas. Las únicas proporciones publicadas corresponden a una herramienta sin modelo de lenguaje, ALICE de la Contraloría General de la Unión de Brasil: 23,5 % de alertas fundadas sobre una muestra de 17 (Dantas, Martins, 2022) y 3,4 % sobre una muestra de 29 (Ministerio de Defensa de Brasil, 2026). En el PPS la proporción de requerimientos de corrección aceptados, 96,6 %, corresponde a la verificación manual. Muestra la conformidad de la entidad contratante, no la confirmación por un órgano independiente. Tras la puesta en marcha del modelo de lenguaje, ese mismo indicador se convertirá en medida de su exactitud solo con una condición: que el PPS publique el número de alertas antes de la verificación por el funcionario.

La ganancia que aporta el modelo de lenguaje frente a las reglas solo se describe en la resolución del TCU, y se describe de forma cualitativa: el modelo reconoce el objeto de la contratación «incluso con descripciones imprecisas», cosa que las reglas aplicadas a los campos no saben hacer, y la incoherencia de los datos antes de ello «llevaba regularmente a falsos positivos». No hay una estimación numérica de la ganancia. El pleno del TCU no incluyó en la parte dispositiva la propuesta de recomendar la continuación del proyecto y se limitó a encargar la coordinación de las iniciativas tecnológicas. Es el único caso del

inventario en que el organismo nombró el costo de la implantación: el tiempo del equipo restado al análisis de contrataciones.

7. Indicios fuera de la tipología

Los indicios que buscan las herramientas de las cinco jurisdicciones encajan en su mayoría en los 30 mecanismos de la tipología. Como se advirtió en el apartado 3.1, aquí se consideran todas las herramientas, incluidas las herramientas sin modelo de lenguaje del anexo; los cuatro candidatos a nuevos mecanismos proceden precisamente de ellas. Otros indicios resultan ser indicadores de mecanismos ya descritos.

Tabla 5. Indicios buscados por las herramientas que no están descritos en la tipología del CILC

Indicio	Herramienta	Relación con la tipología	Propuesta
Participación de un proveedor inscrito en un registro de sancionados (registros CEIS y CNEP en Brasil; análogos en HELIOS y en el Ministerio Público de Minas Gerais)	ALICE CGU, ALICE TCU, HELIOS, MPMG	no descrito	mecanismo de la etapa VII: admisión de un proveedor inhabilitado a través de una persona vinculada o por inacción de la comisión
Donaciones de los participantes a campañas electorales, afiliación partidaria de los socios, servicios a campañas	MPMG, trails T14-T16	no descrito	mecanismo de la etapa I o VII: la contratación como recompensa por el apoyo político
Doble financiación de un mismo proyecto con cargo a distintos fondos	Arachne	no descrito	mecanismo de la etapa X: pago de lo mismo dos veces
Pólizas de seguro falsificadas durante la ejecución	DIARI, Colombia	IMP en parte	mecanismo de la etapa IX: garantía de cumplimiento ficticia
Empresa sin indicios de actividad (sin conexión a la red eléctrica, sin empleados, registro en vísperas de la licitación)	MPMG, ALICE TCU, PPS	indicador de BID-02	indicador
Una misma dirección IP en varios participantes, textos de ofertas idénticos	CADE Cérebro	indicador de BID-01	indicador
Precio de adjudicación del 97-99 % del estimado	Sherlock, BRIAS	indicador de BID-01, BUD-01	indicador
Incumplimiento del plazo de publicación del anuncio	PPS	PUB-01 directamente	ya descrito

De los cuatro candidatos, tres corresponden a segmentos que el participante externo no ve (etapas I, IX y X) y uno a la etapa de evaluación de las ofertas. La tipología predice exactamente eso: las herramientas de los órganos de auditoría y del ministerio público ven lo que el participante no ve, y los indicios que buscan enriquecen precisamente esos segmentos. La inclusión de los candidatos en el catálogo exige su contrastación con materiales de investigaciones y se remite a la siguiente versión de la tipología.

8. Limitaciones y una pregunta abierta sobre las banderas rojas

El inventario se construyó a partir de documentos abiertos y está limitado por lo que los organismos comunican sobre sus herramientas. Un organismo puede utilizar un modelo de lenguaje sin nombrarlo en sus comunicados de prensa; así, el modelo de VigIA se nombra solo en el artículo de funcionarios del

tribunal, no en sus comunicaciones oficiales. También es cierto lo contrario: las palabras «inteligencia artificial» en un comunicado de prensa no siempre significan un modelo de lenguaje. Tales casos (ALICE de la Contraloría General de la Unión, el modelo de la Contraloría General de la República de Colombia, Arachne+) se excluyeron del núcleo. La situación a octubre de 2026 cambiará: ChileCompra y el PPS anunciaron la puesta en marcha para fin de año, la ANAC y la OLAF llevan a cabo experimentos.

Los resultados se calcularon sobre cifras autodeclaradas, y sus denominadores son distintos. El TCU cuenta verificaciones, el TCE-SC anuncios con medidas, el PPS requerimientos de corrección. No es posible comparar las herramientas entre sí con esas cifras, y en el estudio no se hace.

El cotejo de las funciones con los mecanismos lo realizó un solo investigador a partir de las descripciones de las funciones. No se evaluó la concordancia entre codificadores independientes. Donde la descripción de la función es incompleta (OLAF, ANAC), no se asignaron mecanismos.

Pregunta abierta. Todas las herramientas del inventario, con modelo de lenguaje y sin él, emiten la alerta con lógica binaria: la regla se activó o no, el umbral se superó o no. La puntuación de colusión de BRIAS y la clase de riesgo del PPS son escalas ordinales, pero a la salida vuelven a convertirse en la decisión «verificar o no». Los trabajos académicos construyen en cambio estimaciones probabilísticas (Decarolis, Giorgiantonio, 2020; Huber, Imhof, 2019). La tasa de confirmación del CILC ofrece un modo de calibrar esas estimaciones: la proporción de alertas declaradas fundadas sirve de estimación empírica de la probabilidad de infracción dado un indicio determinado. La estimación es condicional: se refiere solo a las alertas seleccionadas para verificación y depende de los criterios con los que el verificador declara fundada una alerta. Sin una selección aleatoria de las alertas para verificación y sin criterios fijos de confirmación, no puede trasladarse a todos los procedimientos con el mismo indicio. ¿No es más exacto describir el riesgo de un procedimiento mediante una probabilidad calibrada con intervalo de confianza que mediante un conjunto de banderas rojas? Esa pregunta excede el marco del presente estudio y se plantea aquí como tarea de un trabajo aparte.

9. Conclusiones

1. A octubre de 2026, los modelos de lenguaje en el control de la contratación de las cinco jurisdicciones se encuentran en la etapa de implantación. El inventario cuenta seis herramientas con modelo de lenguaje confirmado y dos proyectos anunciados sin tal confirmación. Dos herramientas funcionan: respecto de Alice 360 el tribunal publicó una resolución con resultados, y respecto de VigIA el tribunal publicó el número de anuncios con medidas y en el artículo de sus funcionarios nombró el modelo. Las otras cuatro están en desarrollo, en ejecución del contrato o en prototipo.
2. La única función del modelo de lenguaje sobre la que un organismo ha rendido cuenta de resultados mediante una resolución no es la lectura del pliego en busca de infracciones, sino el reconocimiento del objeto de la contratación a partir de descripciones imprecisas; la infracción misma la detecta un algoritmo estadístico. La lectura de anuncios en busca de requisitos excesivos y condiciones a la medida de un solo proveedor funciona en un estado brasileño (VigIA) y está en construcción en Corea y Chile.
3. Las herramientas LLM en operación cubren cinco mecanismos de la tipología de 30, de ellos tres directamente por el modelo de lenguaje. Contando los sistemas en construcción, la cobertura potencial es de 10 mecanismos en cinco etapas. Con excepción de la justificación del precio inicial (BUD-01, que la tipología asigna a la auditoría y en parte al participante externo) y de la ejecución del contrato (IMP-01 e IMP-02 en el sistema en construcción del distrito de Seocho), todos ellos se sitúan en el

segmento visible para el participante externo, y principalmente en la etapa de pliego. Las etapas de planificación, elección del procedimiento de contratación, evaluación de las ofertas, impugnación y pago no están cubiertas por ninguna herramienta. Del grupo «neutralización del control», el sistema en construcción del PPS cubre un mecanismo de cinco. El modelo de lenguaje amplió la escala de la observación, no sus límites.

4. Ningún organismo publica la proporción de alertas fundadas por función del modelo de lenguaje. Sin ese indicador no puede medirse la ganancia que aporta el modelo frente a las reglas a las que se ha añadido. Para monitorear la calidad del modelo de lenguaje y de su aplicación, el organismo debe recopilar al menos cuatro magnitudes por cada función: el número de alertas emitidas, el número de alertas verificadas por una persona, el número de alertas confirmadas y el número de alertas que dieron lugar a medidas. Los organismos disponen de esos datos. El PPS ya publica la distribución de los requerimientos de corrección por tipos de infracción para la verificación manual.
5. Los indicios que buscan las herramientas de las cinco jurisdicciones, incluidas las que carecen de modelo de lenguaje, y que no figuran en la tipología arrojan cuatro candidatos a nuevos mecanismos. Tres de ellos se sitúan en segmentos invisibles para el participante externo, lo que concuerda con la estructura de la tipología por posiciones de observación y define el contenido de su siguiente versión.
6. La lógica binaria de las banderas rojas es común a todas las herramientas del inventario. Deja abierta la pregunta sobre el paso a probabilidades calibradas. La tasa de confirmación se propone como medio de esa calibración, a condición de una selección aleatoria de las alertas para verificación y de criterios fijos de confirmación.

Fuentes

Las fuentes se agrupan por apartados del estudio.

Publicaciones del CILC (apartados 1, 3, 5, 6, 8)

Limarau, D. Mechanisms of Corruption in Public Procurement: A Typology by Stage of the Procurement Cycle. Version 1.3. CILC, 2026. DOI 10.5281/zenodo.21939393.

Krukovskiy, I. Cost Assessment of Procurement with Confirmed Violations at the Notice Publication and Bid Submission Stage. CILC, 2026. DOI 10.5281/zenodo.22237028.

Limarau, D. A Dead-End Branch of Appeal: How the Supervisory Authority Ignores the Signals of an External Reporter. CILC, 2026. DOI 10.5281/zenodo.22819817.

Revisiones y trabajos académicos (apartado 2)

OECD. Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions. OECD Publishing, 2025. DOI 10.1787/795de142-en.

OECD. Data Screening Tools for Competition Investigations. DAF/COMP/WP3(2022)5, 2022.

Open Contracting Partnership. AI in public procurement: between experimentation and impact. 9 July 2026.

PwC, Ecorys. Identifying and reducing corruption in public procurement in the EU. European Commission, 2013.

Fazekas, M., Toth, I. J. From corruption to state capture: a new analytical framework. Political Research Quarterly, 2016.

Porter, R., Zona, J. D. Detection of bid rigging in procurement auctions. *Journal of Political Economy*, 1993.

Imhof, D., Karagök, Y., Rutz, S. Screening for bid rigging: does it work? *Journal of Competition Law and Economics*, 2018.

Huber, M., Imhof, D. Machine learning with screens for detecting bid-rigging cartels. *International Journal of Industrial Organization*, 2019.

Decarolis, F., Giorgiantonio, C. Corruption red flags in public procurement: new evidence from Italian calls for tenders. *Banca d'Italia, Questioni di Economia e Finanza* 544, 2020.

Brasil (apartados 4.1, 4.2, 4.8)

Tribunal de Contas da União. Acórdão 1145/2026-Plenário de 6 de maio de 2026, TC 009.250/2025-5, relator Ministro Jorge Oliveira. Acompanhamento. Edição 2025/2026 da fiscalização contínua de aquisições públicas federais. <https://portal.tcu.gov.br/imprensa/noticias/fiscalizacao-do-tcu-acompanha-compras-publicas-federais-entre-2025-e-2026>

Santos, R. F., Santos, P. G., Marins, F., Hans, E., Bezerra, S., Terêncio, F. A., Feitoza, A. Integrando a Análise de Dados ao Dia a dia do Auditor com a Plataforma de Alertas, Riscos e Tipologias do TCU. Seminário Internacional O Futuro da Auditoria Pública, Revista TCU, 2026.

Albuquerque, A. M. de, Wensing, I. M., Joppi Filho, N. L. Inteligência Artificial Generativa (IAG): perspectivas inéditas no controle público. *Revista do Tribunal de Contas do Estado de Santa Catarina*, ano 2, n. 3, p. 150-164, maio/out. 2024. DOI 10.52028/tce-sc.v02.i03.ART08.SC.

Tribunal de Contas do Estado de Santa Catarina. Inteligência artificial criada pelo TCE/SC possibilita retificação em 215 editais de licitação, com previsão de investimentos de R\$ 2 bilhões. 9 de outubro de 2024. <https://www.tcsc.tc.br/inteligencia-artificial-criada-pelo-tcsc-possibilita-retificacao-em-215-editais-de-licitacao-com>

Bezerra, A. de S., Lopes, L. C. Controle externo baseado em LLM: avaliação empírica do PRUMe AI. *Revista Aracê*, v. 7, n. 9, e8409, 2025. DOI 10.56238/arev7n9-271.

Borges, A. O., Belga, J. L., Gazzola, P. O. L., Morais, A. B. ALICE: Análise automatizada de contratações públicas para prevenção de irregularidades e combate à corrupção. *Cadernos Técnicos CGU*, 2025. https://revista.cgu.gov.br/Cadernos_CGU/article/download/867/552/5714

Dantas, F., Martins, R. Efetividade dos alertas da ferramenta ALICE. Universidade Católica de Brasília, 2022.

Ministério da Defesa, Secretaria de Controle Interno. Relatório final de consultoria sobre a ferramenta Alice. 22 de janeiro de 2026. <https://www.gov.br/defesa/pt-br/aceso-a-informacao/auditorias-1/acoes-de-consultoria-1/arquivos/relatorios/2025/2-relatorio-final-alice-8615445.pdf>

Controladoria-Geral da União. Alice: Analisador de Licitações, Contratos e Editais. <https://www.gov.br/cgu/pt-br/assuntos/auditoria-e-fiscalizacao/alice>

Portal Tela. IA da CGU emitiu 35 mil alertas de risco em licitações federais. 30 de março de 2026. <https://www.portaltela.com/noticias/politica/2026/03/30/ia-da-cgu-emitiu-35-mil-alertas-de-risco-em-licitacoes-federais/>

Conselho Administrativo de Defesa Econômica (CADE). Data screening tools for competition investigations, Note by Brazil. OECD, DAF/COMP/WP3/WD(2022)36, 2022. [https://one.oecd.org/document/DAF/COMP/WP3/WD\(2022\)36/en/pdf](https://one.oecd.org/document/DAF/COMP/WP3/WD(2022)36/en/pdf)

Ministério Público de Minas Gerais, Universidade Federal de Minas Gerais. Painel Trilhas de Licitações. Anais do Simpósio Brasileiro de Banco de Dados (SBBBD), 2023. https://sol.sbc.org.br/index.php/sbbbd_estendido/article/download/44196/43959/

Corea (apartados 4.3, 4.4, 4.8)

ETNews. 조달청, AI로 자체입찰 전수 점검, 공공조달 공정성 높인다. 6 de julio de 2026.

DACON. 나라장터 자체입찰 공고 법령 위반사항 모니터링 AI 경진대회. Ministerio del Interior y Seguridad, Agencia Nacional de la Sociedad de la Información, Servicio de Contratación Pública, 2026.

Servicio de Contratación Pública (PPS). Resultados de la verificación de las contrataciones autónomas en busca de infracciones de la legislación en el primer trimestre de 2026 (según informaron 뉴스서울 el 27 de abril de 2026 y 정보통신신문 el 1 de junio de 2026).

Herald Economy. 서초구, AI 기반 계약서류 검토시스템 구축. 7 de mayo de 2026.

Korea Fair Trade Commission. Competition and Corruption in Public Procurement, Note by Korea. OECD, DAF/COMP/WD(2026)77, 9 June 2026. [https://one.oecd.org/document/DAF/COMP/WD\(2026\)77/en/pdf](https://one.oecd.org/document/DAF/COMP/WD(2026)77/en/pdf)

Korea Fair Trade Commission. Data screening tools for competition investigations, Note by Korea. OECD, DAF/COMP/WP3/WD(2022)31, 2022. [https://one.oecd.org/document/DAF/COMP/WP3/WD\(2022\)31/en/pdf](https://one.oecd.org/document/DAF/COMP/WP3/WD(2022)31/en/pdf)

Chile (apartado 4.5)

Mercado Público. Licitación 869591-2-LP25 «Análisis de Documentos mediante modelos de IA»: registro de adjudicación (Resolución 16-B de 9 de enero de 2026).

ChileCompra. ChileCompra lanza licitación inédita para aplicar inteligencia artificial en compras públicas. 20 de agosto de 2025; ChileCompra expone avances en uso de datos e inteligencia artificial. 28 de noviembre de 2025; ChileCompra aplicó herramienta para el uso ético de IA en su proyecto de detección de anomalías. 31 de diciembre de 2025; Cuenta Pública Participativa 2026. 27 de julio de 2026; ChileCompra presenta en instancia internacional impulsada por el Banco Mundial. 13 de julio de 2026; ChileCompra comparte avances en inteligencia artificial en encuentro G7/MMGP. Septiembre de 2026.

ChileCompra. ChileCompra incrementa en 207% la detección de irregularidades con cruce masivo de datos y tecnologías de inteligencia analítica. Diciembre de 2025. <https://www.chilecompra.cl/2025/12/chilecompra-incrementa-en-207-la-deteccion-de-irregularidades-con-cruce-masivo-de-datos-y-tecnologias-de-inteligencia-analitica/>

ChileCompra. ChileCompra detecta potenciales conflictos de interés por \$3.452 millones tras cruce de datos masivo. Enero de 2026. <https://www.chilecompra.cl/2026/01/chilecompra-detecta-potenciales-conflictos-de-interes-por-3-452-millones-tras-cruce-de-datos-masivo/>

Italia y Unión Europea (apartados 4.6, 4.7)

ANAC. Direttiva programmatica sull'attività di vigilanza, anno 2026. 26 de mayo de 2026.

European Commission. Commission Staff Working Document SWD(2026) 246: monitoring of the Commission Anti-Fraud Strategy action plan, state of play 31 May 2026.

European Commission, Anti-Fraud Knowledge Centre. What is Arachne? (Arachne+), 2026.

OLAF. The OLAF Report 2025. 20 April 2026.

European Court of Auditors. Special report 06/2026. 2026. https://www.eca.europa.eu/ECAPublications/SR-2026-06/SR-2026-06_EN.pdf

Comisión Nacional de los Mercados y la Competencia. Memoria 2024. https://www.cnmc.es/sites/default/files/editor_contenidos/CNMC/Memorias/Memoria_CNMC_2024_.pdf

Specialių tyrimų tarnyba (STT). Piloto de aprendizaje automático explicable para la detección de riesgos en la contratación. Febrero de 2026. <https://stt.lt/en/news/7481/>

Colombia (apartado 4.8, tabla A1)

Contraloría General de la República de Colombia. Modelo Analítico Predictivo en Contratación, 19 de mayo de 2025; Circular 023 de 4 de diciembre de 2025.

Contraloría General de la República de Colombia. Comunicado de prensa sobre la DIARI. Septiembre de 2024. <https://ambitojuridico.com/sites/default/files/2024-09/COMUNICADO-LaDiariha-2024%20%282%29.pdf>

Cambio Colombia. El cerebro digital de la Contraloría que protege los recursos públicos. Octubre de 2025. <https://cambiocolombia.com/contenido-especial/articulo/2025/10/el-cerebro-digital-de-la-contraloria-que-protege-los-recursos-publicos>

Superintendencia de Industria y Comercio. Herramientas de investigación: Sherlock. Presentación para la UNCTAD. https://unctad.org/system/files/non-official-document/tdrbpconf10_unrc_present_investigative_tech_rusinke_sp.pdf

Ministerio de Defensa Nacional de Colombia. Sistema HELIOS (según informaron Vanguardia y Cambio el 6 de agosto de 2026).

Anexo. Herramientas sin modelos de lenguaje en las mismas jurisdicciones

Tabla A1. Herramientas de control de la contratación basadas en reglas, cruces y aprendizaje automático clásico

Herramienta, órgano, año	Técnica	Indicios	Grupos de mecanismos de la tipología	Resultados publicados
ALICE, Contraloría General de la Unión (CGU), Brasil, 2015	expresiones regulares sobre los textos de los anuncios, cruces SQL con registros, rangos de precios, clasificadores	requisitos restrictivos, especificaciones a la medida de un proveedor, incoherencia de los documentos, plazos, sobreprecios, proveedores sancionados	C2, C4, C6	2024: 161 000 procesos, 212 auditorías, R\$ 1 250 millones de efecto; 01.2023-02.2026: 35 000 alertas, 729 auditorías; proporción de alertas fundadas del 23,5 % (17 alertas, 2021) y del 3,4 % (29 alertas, 2025)
Projeto Cérebro, órgano de defensa de la competencia CADE, Brasil, 2014	pruebas de Newcomb-Benford, entropía de rotación, una misma dirección IP, similitud de textos	ofertas de cobertura, rotación	C3	un caso documentado (2018-2021)
Painel Trilhas de Licitações, Ministerio Público del Estado de Minas Gerais (MPMG), 2023	grafos, 20 trails	participante único, datos de contacto compartidos, vínculos con funcionarios, donaciones	C3, C4, C6	MRR de 0,210 frente a 0,035 con selección aleatoria
BRIAS, Comisión de Comercio Justo (KFTC), Corea, 2006	reglas ponderadas, puntuación de colusión	proporción del precio, número de participantes, presentación simultánea, consorcios repetidos	C3	2018-2021: 26 casos, 9 sanciones; 2007-2017: 10 036 banderas, 7 investigaciones

Herramienta, órgano, año	Técnica	Indicios	Grupos de mecanismos de la tipología	Resultados publicados
Observatorio ChileCompra (reglas), 2024-2025	15 reglas, cruces con el registro de funcionarios públicos SIAPER	fraccionamiento, funcionario con participación en un proveedor, conflicto de intereses, defectos formales	C1, C3, C4, C5	2025: más de 13 000 notificaciones, 54 % de medidas adoptadas, 206 remisiones a los organismos
DIARI, Contraloría General de la República, Colombia, 2025	bosque aleatorio	retrasos, incumplimiento, pólizas falsificadas, contratos directos	C1, C6	946 alertas (2025), proporción confirmada no publicada
Sherlock, Superintendencia de Industria y Comercio (SIC), Colombia	tres reglas sobre el SECOP	precio del 97 %, adjudicaciones repetidas, participante único	C3	no hay casos procedentes de la herramienta
HELIOS, Ministerio de Defensa Nacional de Colombia, 2026	aprendizaje automático sobre datos estructurados	participante único, adjudicaciones rápidas, prórrogas, sanciones	C2, C3, C6	no hay
Arachne y Arachne+, Comisión Europea, 2013 y 2026	107 y 78 indicadores, cruces con la base de datos Orbis	concentración, vínculos, modificaciones del contrato, doble financiación	C3, C4, C6	no hay estadísticas de resultados; el Tribunal de Cuentas Europeo (informe 06/2026) no evaluó la eficacia
BRAVA, Comisión Nacional de los Mercados y la Competencia (CNMC), España	clasificador de aprendizaje automático	patrones de precios	C3	un caso; la exactitud no se publica
Modelo del Servicio de Investigaciones Especiales (STT), Lituania, 2026	aprendizaje automático explicable	no se revela	no determinado	piloto