

## INTELIGENCIA ARTIFICIAL Y AUDITORÍAS BASADAS EN DATOS NO ESTRUCTURADOS

### *Artificial intelligence and audits based on unstructured data*

Hayat Massoud<sup>1</sup>

<https://orcid.org/0000-0001-5207-857X>

Recibido: 15/03/2024

Aceptado: 15/06/2024

Publicado: 30/06/2024

Cómo citar este artículo: Massoud, H. (2024). Inteligencia artificial y auditorías basadas en datos no estructurados. *Entrelíneas*, 3(1), 5-18. <https://doi.org/10.56368/Entrelíneas312>

### RESUMEN

En contabilidad, auditoría y finanzas, se manejan numerosos documentos para los cuales no existe una metodología que permita extraer los datos verdaderamente relevantes. Actualmente, el uso de la inteligencia artificial permite tratar datos no estructurados, cuya complejidad para procesarlos y organizarlos requiere una considerable cantidad de horas-hombre, ralentizando la labor del auditor. En los últimos años, términos como Big Data, análisis de datos e inteligencia artificial han emergido como conceptos clave en la contabilidad. Con el objetivo de describir la aplicación de la inteligencia artificial en el análisis y procesamiento de datos no estructurados para su uso en auditorías operacionales, se realizó un estudio descriptivo y documental, con un diseño no experimental y una perspectiva transversal. Los resultados demostraron el tratamiento de datos no estructurados en la auditoría, especificando su clasificación y manejo, destacando cómo el Big Data influye en la auditoría y cómo los datos estructurados y no estructurados están transformando los métodos tradicionales del proceso auditor. Se concluye que manejar grandes volúmenes de datos no estructurados siempre plantea riesgos de seguridad y privacidad, lo que justifica la necesidad de proteger estos datos contra accesos no autorizados. Al mismo tiempo, se enfatiza la necesidad de que los contadores y auditores adquieran nuevas habilidades y conocimientos técnicos para trabajar con herramientas avanzadas de inteligencia artificial y análisis de datos.

**Palabras clave:** análisis de datos, auditoría financiera, Big Data, contabilidad, Inteligencia Artificial.

---

<sup>1</sup> Universidad de Panamá, Facultad de Administración de Empresas y Contabilidad, Centro Regional Universitario de San Miguelito. Doctora en Gerencia. [hayat.massoud@gmail.com](mailto:hayat.massoud@gmail.com)

## ABSTRACT

In accounting, auditing and finance, numerous documents are handled for which there is no methodology that allows the truly relevant data to be extracted. Currently, the use of artificial intelligence makes it possible to process unstructured data, the complexity of which to process and organize requires a considerable number of man-hours, slowing down the auditor's work. In recent years, terms such as Big Data, data analytics and artificial intelligence have emerged as key concepts in accounting. With the objective of describing the application of artificial intelligence in the analysis and processing of unstructured data for use in operational audits, a descriptive and documentary study was carried out, with a non-experimental design and a transversal perspective. The results demonstrated the treatment of unstructured data in the audit, specifying its classification and management, highlighting how Big Data influences the audit and how structured and unstructured data are transforming the traditional methods of the audit process. It is concluded that handling large volumes of unstructured data always poses security and privacy risks, which justifies the need to protect this data against unauthorized access. At the same time, the need for accountants and auditors to acquire new skills and technical knowledge to work with advanced artificial intelligence and data analysis tools is emphasized.

**Keywords:** data analysis, financial audit, Big Data, accounting, Artificial Intelligence.

## Introducción

Los sistemas que se están implementando en la actualidad para la toma de decisiones basados en las TIC están causando impresión en el mundo empresarial moderno, sobre todo porque existe una presión creciente sobre los auditores para que desempeñen un papel más efectivo en el control de las entidades corporativas. El uso de sistemas de inteligencia artificial (IA, indistintamente) por parte de los auditores contaba hasta la década pasada con beneficios y limitaciones identificados, sobre todo cuando se pretendía evaluar el impacto de la inteligencia artificial en el diseño y monitoreo de los sistemas de control interno, y las implicaciones de usarlos para las operaciones y la supervivencia de firmas de auditoría pequeñas y medianas que viven constantemente tratando de cerrar la brecha entre las expectativas y el desempeño de la auditoría (Omoteso, 2012).

La inteligencia artificial ya no se considera un elemento del futuro lejano, porque la tecnología está cada vez más presente en la mayoría de los procesos industriales y económicos, haciendo que la IA sea un factor transformador en las organizaciones, mejorando significativamente la toma de decisiones. En el área de auditoría, particularmente, Montoya Hernández & Valencia Duque (2020) han examinado el uso de la inteligencia artificial desde perspectivas teóricas y prácticas en distintos sectores, por su aplicabilidad para impulsar el aumento de la eficiencia de los procesos, la reducción de los costos a largo plazo y ayudar con la disminución de errores que se alcanzan mediante la automatización. Además, la IA permite el acceso a sistemas capaces de realizar inferencias, retroalimentarse y generar conocimiento a partir de actividades previas, lo que la convierte en un área de gran interés actualmente, tal como comentan estos autores.

Para contextualizar el tema, en el área de contabilidad los datos no estructurados son cualquier tipo de información contable que no esté organizada, que no se encuentre en un formato predefinido y que no se pueda procesar fácilmente usando ninguno de los métodos tradicionales propios de la contabilidad (Nielsen, 2022). Es común que en una empresa se registre un alto movimiento de correos, facturas, recibos, estados financieros y otros que no están en forma tabular (datos que se organizan en filas y columnas, similares a una tabla, como las hojas de cálculo), requiriendo de técnicas especializadas como el procesamiento del lenguaje natural o la visión por computadora, para extraerles la información significativa.

El uso de la inteligencia artificial aquí, supone la aplicación de técnicas como el procesamiento de lenguaje natural (*Natural Language Processing*), la visualización por computadora y el aprendizaje automático. De esta forma se pueden procesar los datos no estructurados para analizarlos debidamente de una manera más rápida y eficiente.

Al ser parte de la IA, el procesamiento de lenguaje natural se centra en la comunicación entre el hombre y las computadoras, haciendo que la comunicación ambigua e imprecisa que transmite el primero, es un factor que aborda la automatización de manera lógica para eliminar la ambigüedad y facilitar la comprensión. En contabilidad, auditoría y finanzas, existe una gran cantidad de documentos que se manejan (desempeño financiero empresarial, evaluación presente y futura de la administración de la empresa, evaluaciones de analistas, estándares de dominio, regulaciones, evidencia del cumplimiento de normas y regulaciones, etc.), para los cuales no existe una metodología para extraer los datos realmente relevantes (Fisher *et al.*, 2016).

En los últimos años, el término Big Data, junto con otros temas como el análisis de datos y la inteligencia artificial, han surgido como un término más que manejar en la contabilidad. Las habilidades de los profesionales contables han progresado junto con los avances tecnológicos, evolucionando desde las técnicas tradicionales del uso de lápiz y papel, hacia máquinas de escribir y calculadoras, y finalmente adoptando herramientas avanzadas como hojas de cálculo y *softwares* especializados. Aquí se incursiona en el tema del análisis de datos en la contabilidad que, aunque relativamente nuevo, es un tema que está creciendo de manera significativa en todas las áreas de esta profesión. Los contadores que logren comprender el manejo de patrones y tendencias en grandes conjuntos de datos y lleguen a convertir esos hallazgos en narrativas estratégicas persuasivas, son los que ocuparán una posición concluyente en el mundo empresarial del siglo XXI (Bose *et al.*, 2023).

Esto se menciona porque el principal problema para la contabilidad en las empresas es cuando los profesionales se encuentran ante datos no estructurados, dado que la dificultad de procesarlos y organizarlos requiere de un cúmulo de horas-hombre que ralentiza la labor, y no siempre se puede obtener toda la data de manera sistematizada y a tiempo para analizar la información resultante de manera eficiente, con la precisión que amerita.

Los datos no estructurados en una organización se encuentran en los correos electrónicos, documentos escaneados y archivos multimedia, pero se denominan así porque no siguen un formato estandarizado para que se pueda organizar en filas y columnas. Para que puedan ser integrados en los sistemas contables tradicionales, la información y el procesamiento dependen de la estructuración de los datos, porque un formato inconsistente dificulta significativamente el trabajo de los contadores y auditores.

Para ilustrar esta problemática, es fácil imaginar la gran cantidad de datos de ese tipo que se recoge en los sistemas tradicionales de procesamiento diariamente. La gestión para extraer información considerada verdaderamente relevante de este ingente conjunto de datos requiere el uso de recursos y capacidades humanas y tecnológicas, con las que no siempre se cuenta (Adnan & Akbar, 2019).

Si bien los sistemas identifican y extraen información de los datos no estructurados, hasta la actualidad el proceso se presenta complicado. Por ejemplo, los algoritmos de Procesamiento del Lenguaje Natural (*Natural Language Processing*) y la búsqueda visual que realizan por medio de las computadoras tienen que ser capaces de interpretar correctamente el contexto y el contenido, pero esto nuevamente entra en conflicto con la ambigüedad y la complejidad del lenguaje natural que utiliza el individuo para comunicar y registrar sus ideas (Salinas, 2023).

Con respecto a la integración de este tipo de datos con los sistemas contables ya existentes (ahora denominados como tradicionales), hay que tomar en cuenta que, al estar diseñados para manejar datos estructurados, si se pretende integrar en ellos datos no estructurados, se necesitan

desarrollar nuevas interfaces y procesos que manejen lo complejo del lenguaje natural y lo inserten en el contexto adecuado.

A esta situación se le suma el hecho de que los datos no estructurados también pueden contener errores de los que nadie se encuentra exento, y no están limpios de información irrelevante ni redundante, que va afectando la calidad del análisis contable (Yang *et al.*, 2018). Como garantizar la precisión e integridad de estos datos es fundamental para una contabilidad fiable, vale la pena preguntarse hasta qué punto la inteligencia artificial puede mejorar la eficiencia y efectividad de las auditorías mediante el análisis de datos no estructurados (De La Hoz Suárez *et al.*, 2024).

Para tratar de responder a esa pregunta, se ha planteado como objetivo de este breve estudio describir la aplicación de la inteligencia artificial en el análisis y procesamiento de los datos no estructurados, para utilizarla en las auditorías operacionales. La finalidad es tratar de conocer si analizar y procesar datos no estructurados mediante el uso de la IA ha arrojado datos convincentes que hayan sido documentados, destacando los beneficios y oportunidades que conlleva esta integración en el área contable.

### **Revisión de la literatura**

Lo reciente del tema (y en general de la nueva etapa que atraviesa la inteligencia artificial) en diversos campos que trabajan con datos precisos, ha permitido realizar una revisión de la literatura para explorar brevemente el uso que la inteligencia artificial y las técnicas de aprendizaje automático tienen en la auditoría contable. Esta revisión se centra en cómo estas tecnologías emergentes pueden mejorar la precisión, eficiencia y eficacia de los procedimientos de auditoría, permitiendo a los auditores analizar una gran cantidad de datos (volúmenes que el ojo y la capacidad humana están lejos de alcanzar) con mayor profundidad, y detectar patrones y anomalías que podrían pasar desapercibidos con los métodos tradicionales (De la Hoz Suárez *et al.*, 2024). Diversos autores examinan la integración de herramientas avanzadas como redes neuronales, procesamiento del lenguaje natural y sistemas de análisis predictivo, y su impacto potencial en la reducción de riesgos y en la toma de decisiones más informada en el ámbito de la auditoría financiera.

La Fundación Latinoamericana de Auditores Internos (2022) explica que, cuando se menciona el término 'datos', la mayoría de las personas imagina cadenas de números en la pantalla de una computadora, probablemente en una hoja de cálculo; sin embargo, los datos abarcan mucho más, porque son cualquier tipo de información que pueda recopilarse y analizarse (Bridwater, 2018). En la actualidad pueden venir en forma de una red social como X, informes de tendencias de clientes, de ingresos, o datos de tiempo recopilados por máquinas en una fábrica, lo que hace que la variedad de datos sea casi infinita.

Hou *et al.* (2020) realizaron una investigación sobre la tecnología de procesamiento de datos no estructurados en la ejecución de auditorías, enfocándose particularmente en el contexto del presupuesto de Big Data. Ese estudio examinó cómo las herramientas y técnicas avanzadas de procesamiento de datos no estructurados pueden ser utilizadas para mejorar la eficiencia y precisión de las auditorías. En el estudio se destacó la importancia de integrar tecnologías de Big Data para manejar la creciente cantidad de datos no estructurados que las empresas generan hoy en día, tales como correos electrónicos, registros de transacciones, publicaciones en redes sociales y documentos internos.

Montoya Hernández & Valencia Duque (2020) presentaron una revisión sistemática de la literatura sobre el uso de la inteligencia artificial en el ámbito de la auditoría. En su estudio, los autores examinan diversas investigaciones y desarrollos tecnológicos que demuestran cómo la IA está siendo integrada en las prácticas de auditoría. La revisión abarca diferentes aplicaciones, desde el análisis de grandes volúmenes de datos hasta la automatización de tareas repetitivas y la mejora en la detección de fraudes y riesgos.

Fukas *et al.* (2021) desarrollaron un modelo de madurez de inteligencia artificial específicamente diseñado para la auditoría, que se centra en evaluar y guiar a las firmas de auditoría en su camino hacia la integración y utilización efectiva de tecnologías de inteligencia artificial en sus procesos. El modelo propuesto identifica varias etapas de madurez, desde el nivel inicial de adopción y experimentación con tecnologías básicas, hasta niveles más avanzados donde las soluciones de inteligencia artificial están completamente integradas y optimizadas para la automatización y mejora de la calidad de las auditorías.

Chowdhury (2021) analiza en su estudio las perspectivas y desafíos que plantea el uso de la inteligencia artificial en el proceso de auditoría. El estudio destaca cómo la inteligencia artificial puede transformar significativamente la auditoría al automatizar tareas rutinarias, mejorar la precisión de los análisis y permitir la identificación temprana de riesgos y fraudes. Chowdhury también señala que la implementación de IA en auditoría enfrenta varios desafíos, incluyendo la necesidad de un marco regulatorio adecuado, la capacitación y adaptación del personal, y la integración de estas tecnologías en los sistemas existentes de las firmas auditoras.

En estos cuatro casos se observa que los estudios más recientes sobre la incorporación de la inteligencia artificial en la auditoría contable representan un avance significativo en la búsqueda de precisión y eficiencia del proceso contable. A pesar de los desafíos que permanecen asociados con la integración y la necesidad de un marco regulatorio adecuado, la capacidad de analizar grandes volúmenes de datos y detectar patrones con mayor rapidez y precisión ofrece un panorama de mejoras a la calidad de las auditorías, siendo una herramienta conveniente para el ejercicio profesional contable (Hasan, 2021).

## **Metodología**

Por tratarse de un tema que representa una novedad en Panamá, este estudio es descriptivo, mostrando un acercamiento general desde la perspectiva documental, por lo que se ha llevado a cabo una revisión de la literatura seleccionada, abarcando estudios considerados relevantes por la autora para proporcionar una visión integral del Big Data, el análisis de datos y la inteligencia artificial en la profesión contable.

Con un diseño no experimental y ubicando la data de manera transversal, el procedimiento para el análisis comenzó con la recopilación de la información de fuentes académicas y profesionales sobre dos palabras clave: la inteligencia artificial y su aplicación en la auditoría. Se utilizaron fuentes como los artículos de investigación de Fisher, Garnsey y Hughes (2016), así como los estudios de Fukas, Rebstadt, Remark y Thomas (2021), que proporcionaron una síntesis de la literatura y modelos de madurez en gestión de las tecnologías de información.

Cuando se procedió a la clasificación y categorización de la información, se agruparon por temas, según fueron identificados en la literatura revisada, donde se incluyó la perspectiva en la aplicación de técnicas avanzadas de procesamiento de datos no estructurados, como las discutidas por Hou *et al.* (2020) y Sun & Vasarhelyi (2018), quienes exploraron el análisis de datos textuales y el aprendizaje profundo en la auditoría. A partir de esta pre-estructura, se llevó a cabo una revisión comparativa para identificar los puntos que se desarrollaron en los resultados y las áreas de convergencia entre los estudios registrados.

Esta comprobación se realizó siguiendo el enfoque metodológico propuesto por Montoya Hernández & Valencia Duque (2020), quienes realizaron una revisión sistemática de literatura sobre inteligencia artificial en la auditoría, tomando de estos autores el proceso de “resumir, agrupar y criticar la información y material bibliográfico existente sobre un tema o situación de interés, aplicando un proceso de búsqueda, clasificación, análisis, crítica y síntesis” (p. 215), aunque no de un modo tan sistemático.

La estructura de los resultados, para darle coherencia al estudio, partieron de una clasificación mucho más actual de los datos que realizó Bridwater (2018) para Forbes, donde se encontraba la distinción entre datos estructurados y no estructurados, una especie de contexto necesario para abordar el hecho de la auditoría como proceso sistemático, y a partir de allí desglosar los puntos necesarios para cumplir con el objetivo de la investigación y dar respuesta a la pregunta formulada.

## Resultados

### Clasificación de los datos según Bridwater (2018)

Para ayudar a organizar la cantidad de información que circula hoy en día, se utiliza una lista de 13 categorías de datos provistas por Bridwater (2018), de acuerdo con la definición aportada para facilitar su clasificación. Para los denominados datos grandes, por ejemplo, se encuentra el Big Data, porque son datos que exceden la capacidad de las bases de datos tradicionales para capturarlos, gestionarlos y procesarlos con baja latencia. Estos conjuntos de datos están estructurados, semiestructurados y no estructurados, provenientes de diversas fuentes y en volúmenes que van desde terabytes hasta *zettabytes*. Big Data ayuda con el aprendizaje automático, que es imprescindible para la inteligencia artificial. Con el análisis de grandes volúmenes de datos, es posible identificar patrones y entender por qué ocurren ciertos eventos, predecir futuros comportamientos y tomar decisiones estratégicas basadas en estos análisis.

Los datos pueden clasificarse según su nivel de estructura. Los estructurados tienen un modelo predefinido y están organizados de manera consistente, generalmente en formatos tabulares, mientras que los no estructurados, como los *feeds* de las redes sociales no siguen una estructura predefinida. El avance de las aplicaciones modernas y la programación orientada a objetos ha desafiado el paradigma de los datos estrictamente estructurados, y hoy en día los datos pueden variar en su estructura y cambiar rápidamente, lo que lleva a los desarrolladores a buscar alternativas más flexibles a las bases de datos relacionales para gestionar estos datos dinámicos.

Los datos con marca de tiempo son aquellos que incluyen una secuencia temporal que define cuándo se capturó o procesó cada punto de los datos. Son valiosos para analizar comportamientos a lo largo del tiempo, como las acciones de los usuarios en un sitio *web* y sirven a los sistemas que predicen comportamientos futuros o realizan análisis de recorrido, ya que es un rastro que permite reproducir y aprender de los pasos de los usuarios.

Los datos de la máquina son generados por sistemas y tecnologías que impulsan las operaciones empresariales modernas. Incluyen información de API, puntos finales de seguridad, colas de mensajes, aplicaciones en la nube y sensores industriales. Aunque suelen ser ignorados, contienen un registro en tiempo real de la actividad y el comportamiento de los sistemas y de los usuarios, haciéndolos valiosos para resolver problemas, identificar amenazas y predecir futuros inconvenientes mediante el aprendizaje automático.

Los datos espaciotemporales combinan información de ubicación y tiempo para un mismo evento, mostrando cómo cambian los fenómenos en una ubicación física a lo largo del tiempo. Con ellos se pueden rastrear vehículos en movimiento, describir cambios poblacionales y detectar anomalías en las redes de telecomunicaciones, permitiendo a los analistas usar estos datos para realizar cálculos y resumir estadísticas basadas en ubicaciones específicas.

Los datos abiertos son aquellos disponibles libremente para cualquier persona, sin restricciones de derechos de autor o patentes. Tecnologías emergentes como la computación multipartita y la privacidad diferencial permiten compartir datos privados sin comprometer la identidad del propietario y, aunque aún son técnicas académicas, se cree que en el futuro los datos abiertos se entenderán de nuevas maneras, promoviendo la innovación y la creación de riqueza. Por otro lado, los datos oscuros son información digital recopilada y almacenada pero no utilizada, una especie de activos de información que una organización tiene pero que generalmente no se utilizan

para otros fines, sin embargo, representan un potencial desperdiciado si no se aprovechan adecuadamente para generar valor.

Los datos en tiempo real se refieren a información que se procesa y actúa casi instantáneamente. Con la computación de borde y el aumento de 5G, los datos en tiempo real permitirán desplegar recursos de emergencia más rápidamente y mejorar la fluidez del tráfico en eventos masivos, fortaleciendo la relación entre consumidores y marcas mediante ofertas relevantes entregadas en momentos precisos.

Los datos genómicos analizan el ADN para identificar nuevos fármacos y personalizar tratamientos, un volumen que está creciendo rápidamente y superarán a los generados por plataformas como X y YouTube, con el desafío de procesar y analizar estos datos masivos para convertirlos en conocimiento útil, empleando herramientas especializadas y flujos de trabajo unificados.

Los datos operativos convierten la información de los sistemas de tecnología de la información en recursos utilizables por empleados de todas las áreas, llevando a analizar nuevas aplicaciones y objetivos comerciales para desarrollar métricas operativas efectivas lo que hace que, adoptar una mentalidad de datos operativos, las organizaciones pueden asegurarse de que la información correcta llegue a las personas adecuadas en el momento adecuado, mejorando la toma de decisiones.

Los datos de alta dimensión se están popularizando en tecnologías de reconocimiento facial, por la complejidad que tiene el reconocimiento de los contornos faciales humanos; estos son datos multifacéticos que permiten manejar cálculos que describen las fisonomías faciales, facilitando el procesamiento y reconocimiento de rostros de personas.

Los datos obsoletos y no verificados se han recopilado sin confirmar su relevancia, precisión o tipo correcto, y esta característica hace que confiar en estos datos sea peligroso, ya que las decisiones basadas en información no verificada o desactualizada pueden ser erróneas, debiéndose verificar la validez de los datos antes de utilizarlos para cualquier propósito comercial.

Los datos translíticos combinan transacciones y análisis, permitiendo el procesamiento y la generación de informes en tiempo real. Las empresas pueden realizar análisis de 'ventanas de transacciones', mejorando la eficiencia operativa y proporcionando acceso rápido a información relevante en tiempo real, respaldando tareas como la detección instantánea de fraudes y la orientación inteligente de marketing.

Es decir, los datos pueden adoptar muchas formas y se pueden clasificar de diversas maneras para facilitar su comprensión y uso. Desde grandes volúmenes de información hasta datos que se generan y procesan en tiempo real, pasando por datos que permanecen inactivos, cada tipo de dato ofrece diferentes oportunidades y desafíos para el análisis y la toma de decisiones (Fundación Latinoamericana de Auditores Internos, 2022). En este contexto, la auditoría de datos se vuelve esencial para garantizar la calidad y la integridad de los datos, permitiendo que el análisis y la toma de decisiones sean efectivos.

### **La auditoría como proceso sistemático**

Chowdhury (2021) define la auditoría como un proceso sistemático que implica obtener y evaluar objetivamente evidencias sobre afirmaciones relacionadas con acciones y eventos económicos, con el fin de determinar su grado de conformidad con criterios establecidos y comunicar los resultados a las partes interesadas. Ser objetivo en este campo significa que el auditor es independiente, trabaja con objetivos específicos y emite una opinión experta después de examinar y evaluar los documentos pertinentes.

Tras varios altibajos y dos períodos de estancamiento en la investigación de la inteligencia artificial, los estudios actuales en este campo de la auditoría se están volviendo cada vez más relevantes para la disciplina de los sistemas de información. Los conceptos para la gestión de

aplicaciones de IA y la integración de los sistemas de inteligencia artificial en los procesos operativos son todavía un desafío importante para el discurso científico actual. Con el desarrollo continuo de *hardware* nuevo y cada vez más potente, junto con nuevas arquitecturas de IA y marcos de programación, son muchas las industrias que están desarrollando soluciones prácticas para sus tareas específicas. Precisamente por esto es que la IA se beneficia del crecimiento de volúmenes de datos, gracias a otras tendencias emergentes como la computación en la nube y el Internet de las Cosas (Fukas *et al.*, 2021).

Este potencial de creación de valor en los próximos años está en la auditoría. Los autores mencionan la predicción de una amplia automatización de procesos mediante el uso de inteligencia artificial para el año 2035, y que las principales firmas de auditoría ya han estado invirtiendo en esta tecnología en los últimos años. Sin embargo, mientras que la IA se utiliza ampliamente en otros dominios, en la auditoría su aplicación práctica y productiva todavía se presenta de manera incipiente. Una razón importante para esto es que -explican-, en comparación con otras industrias, la auditoría está sujeta a una fuerte regulación y estandarización que aún no considera explícitamente el uso de la inteligencia artificial. Por lo tanto, el uso de estos sistemas no se debe limitar a satisfacer los criterios de efectividad y eficiencia del auditor, sino también a cumplir con todos los requisitos de otras partes interesadas en el ecosistema de la auditoría.

Dado que la decisión de otorgar la continuidad de una empresa tiene un gran impacto económico y puede ser una base para disputas legales, la explicabilidad y transparencia en la toma de decisiones del auditor también son criterios que deben considerarse al desarrollar sistemas de IA para la auditoría. Con la tecnología, este término (explicabilidad) es un neologismo utilizado para definir la capacidad que tienen los modelos de aprendizaje automático para ser entendidos y que también sean confiables, de acuerdo con Ortiz de Zárate Alcarazo (2022).

Para ayudar a las firmas de auditoría a desarrollar y utilizar soluciones de *software* basadas en IA, a pesar de los numerosos requisitos regulatorios, el uso de los Modelos de Madurez como herramienta de gestión estratégica son una opción adecuada. Para diferentes tecnologías y campos de aplicación han sido desarrollados y discutidos en la investigación científica durante años. Sin embargo, aquellos modelos que abordan el uso de IA son mayormente muy genéricos y no tienen en cuenta las especificidades de dominios de aplicación decisivos, y dado que, incluso un solo estándar de auditoría puede impedir la implementación de un caso de uso técnicamente viable que cumpla con todas las demás leyes, los modelos de madurez existentes para IA son de significado limitado en el uso en auditoría para la transformación de los datos brutos (Fukas *et al.*, 2021).

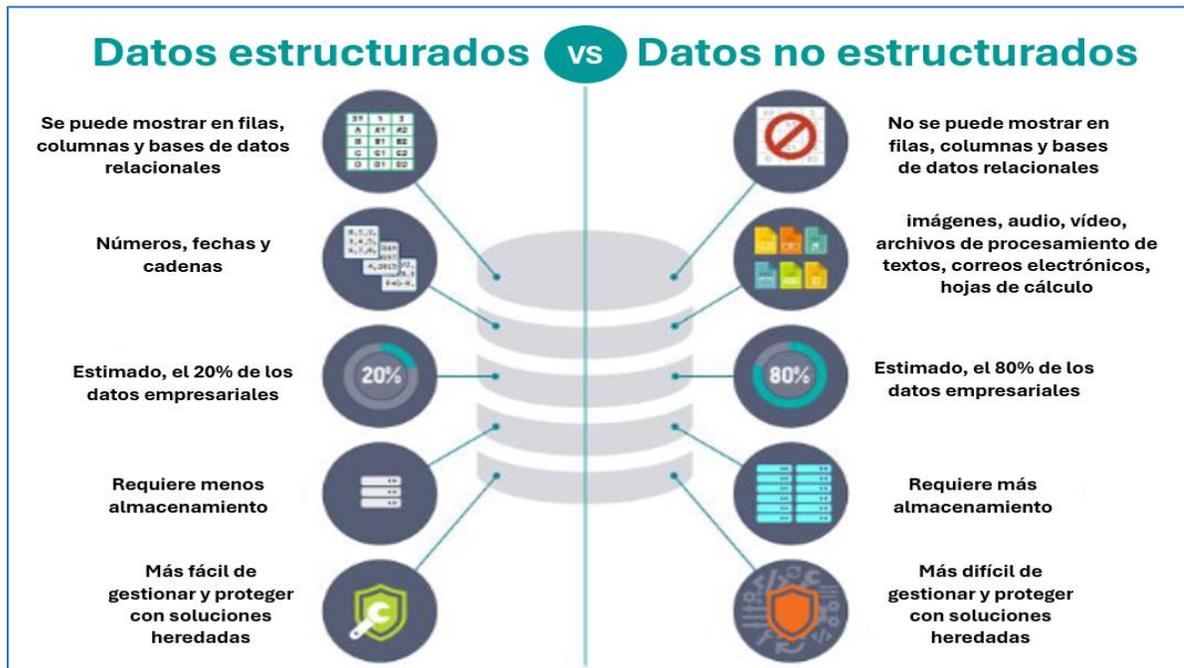
### **Datos de auditoría y datos no estructurados**

La conversión de los datos brutos que son recolectados en evidencias de auditoría sirven para alcanzar los objetivos de este proceso. Desde la recolección hasta la obtención de evidencia, los auditores deben centrarse en si pueden extraer datos valiosos para el análisis, lo que influye directamente en la calidad de los proyectos de auditoría y en la presentación de los resultados, y esta es la razón para que el análisis de los datos se considere el paso más crítico en los trabajos de auditoría (Hou *et al.*, 2022).

Estos autores explican que, actualmente, los métodos y herramientas asistidos por computadora más utilizados en auditoría incluyen los análisis de cuentas, de ratios económicos, de tendencias, los análisis estadísticos, de datos en Excel, Access, SQL, y *softwares* de auditoría como AO. Es normal que las firmas empleen Excel y *softwares* de auditoría para estados contables para recolectar y analizar datos de auditoría. Con la llegada de la era del Big Data, la gran cantidad de datos heterogéneos ha ampliado el alcance de las auditorías de Big Data, pues los métodos y herramientas tradicionales han quedado obsoletos para manejar y analizar datos no estructurados.

Señalan que los datos no estructurados son complejos y presentan características estructurales diversas. Las bases de datos relacionales tradicionales no pueden gestionar adecuadamente estos datos en términos de descripción y escala de gestión, por ello, es necesario desarrollar nuevas tecnologías y plataformas de gestión específicas para los datos no estructurados. El lenguaje de consulta y la visualización de datos son las herramientas a través de las cuales los usuarios interactúan y analizan Big Data. Este lenguaje debe ser capaz de expresar de manera precisa las intenciones del usuario al acceder a los datos y debe soportar una variedad de operaciones de datos flexibles y complejas. La Figura 1 presenta las diferencias entre los datos estructurados y los no estructurados.

**Figura 1**  
*Datos estructurados versus datos no estructurados*



*Nota. Hou et al. (2020, p. 2).*

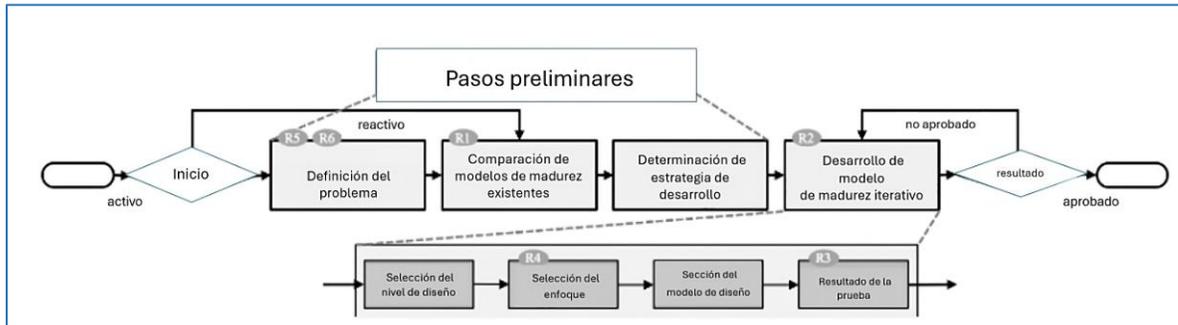
### Sistemas expertos en auditoría

Los sistemas expertos son programas de *software* que integran la experiencia de varios expertos para identificar problemas potenciales que podrían afectar la calidad del servicio. Muchas empresas utilizan *software* especializado para registrar transacciones, gestionar inventarios, preparar estados financieros y analizar resultados, con proveedores que personalizan sus productos según los requisitos de los clientes (Chowdhury, 2021). En el marco de las auditorías, los sistemas expertos son particularmente efectivos para abordar el problema de los datos no estructurados, convirtiéndose en una ventaja para la auditoría.

Encontrar patrones y anomalías que indiquen la existencia de irregularidades financieras, y la integración de datos estructurados con datos no estructurados, ayudan a actualizar con una visión más amplia la información financiera y operativa de una organización, porque está manejando datos por volumen cada vez más grandes. En el caso de los sistemas expertos en auditoría, todo depende del desarrollo y madurez de estos sistemas. Los Modelos de Madurez (MM) son herramientas que ayudan a evaluar y mejorar la efectividad de estos sistemas a lo largo del tiempo, combinando varias metodologías:

**Figura 1**

*Desarrollo de Modelos de Madurez (MM) en el contexto de los sistemas expertos en auditoría*



*Nota.* Fukas *et al.* (2021, p. 4).

Los sistemas expertos en auditoría se explican en Fukas *et al.* basándose en las ideas de Becker *et al.* (2009), donde se crean varios procesos de diseño de los Modelos de Madurez que son refinados a través de la investigación y las directrices de la Ciencia del Diseño propuestas por Hevner *et al.* (2004), para asegurar que el Modelo sea robusto y aplicable a diferentes contextos de la auditoría.

Un modelo de este tipo (señalan los autores) se prueba en escenarios reales para asegurar la eficacia y eficiencia, realizándose ajustes según sea necesario para mejorar su aplicabilidad, lo que confirma que los sistemas expertos sean efectivos en el manejo de datos no estructurados, mientras evolucionan continuamente en beneficio del campo de la auditoría.

### El manejo de datos no estructurados

El Big Data influye en muchos aspectos de la toma de decisiones empresariales, utilizando diversas fuentes de datos, que los contadores para mejorar la calidad de las auditorías y obtener conocimientos comerciales más relevantes. En el área contable se ha reconocido el impacto que tienen el análisis textual para mejorar la calidad de la auditoría, al poder leer miles de contratos y cualquier tipo de documentos complejos, mientras va extrayendo y estructurando la información textual que respalde la prácticas de auditoría (Sun & Vasarhelyi, 2018).

Los datos financieros, así como los no financieros, requieren ser verificados por un auditor independiente, que es una expansión en las funciones y responsabilidades de los auditores, aumentando la satisfacción de las partes interesadas y alineándose más con sus expectativas, algo que podría disminuir la brecha de expectativas en auditoría. De acuerdo con esto y para cumplir con esa nueva función de manera eficiente, con el tiempo, los auditores deberán apoyarse más en el análisis de Big Data y otras técnicas tecnológicas (Michael & Dixon, 2019).

Los procedimientos analíticos de evaluación de riesgos, por ejemplo, permiten a los auditores identificar relaciones entre datos, comparar, detectar patrones inesperados y evaluar resultados, similar a la revisión analítica en auditoría. Este proceso ofrece una visión general de los riesgos en la entidad, incluyendo la posibilidad de fraude financiero y no financiero. Con el uso de Big Data, la variedad de herramientas que existen proporcionan una evidencia más completa y profunda, ya que requiere un análisis más exhaustivo similar al realizado por auditores investigadores (Yudowati & Alamsyah, 2018).

Si bien es cierto que la auditoría ha avanzado menos que otras áreas en la implementación de técnicas de Big Data, Gepp *et al.* (2018) señalan que una razón podría ser que los auditores son cautelosos al usar tecnologías que sus clientes aún no han adoptado ampliamente, pero este argumento es discutible por los autores, porque los auditores tienen la capacidad y el conocimiento para utilizar estas técnicas de manera efectiva, independientemente de la adopción de los clientes,

especificando que la incorporación de las técnicas de Big Data puede proporcionar un valor significativo para mejorar la calidad de la auditoría y la capacidad para identificar riesgos y fraudes, lo que destaca la importancia de su adopción en la auditoría.

Revisando en la literatura se han encontrado algunas de las herramientas de aprendizaje profundo para auditorías contables, que sirven para ejemplificar de qué forma se están realizando las auditorías basadas en datos no estructurados. En el caso de *TensorFlow*, fue creada y mantenida por el equipo de Inteligencia Artificial de *Google Brain*, es una biblioteca de código abierto basada en un sistema matemático de programación de flujo, utilizada en la implementación de diversos algoritmos de aprendizaje automático que permite la construcción y entrenamiento de modelos de redes neuronales. Esta herramienta es aplicada en la predicción de índices bursátiles y en sistemas de gestión de la información, así como en el análisis de grandes volúmenes de datos financieros para detectar anomalías en los registros contables (Török, 2022).

Otra herramienta que vale la pena mencionar la describen Sun & Vasarhelyi (2018). El *IBM Watson Natural Language Understanding*, es un servicio en la nube que ofrece funciones avanzadas de procesamiento del lenguaje natural, para extraer metadatos esenciales de los textos (entidades, relaciones, conceptos, sentimientos y emociones). Se basa en una herramienta anterior, *Alchemy Language*, y utiliza algoritmos de aprendizaje profundo entrenados con texto de miles de millones de páginas *web* sobre una amplia variedad de temas.

Los auditores pueden aprovechar este servicio para automatizar el análisis de datos textuales, evaluando sentimientos y emociones en documentos y transcripciones de conferencias telefónicas. Con esta herramienta se mide el sentimiento dirigido hacia palabras clave específicas en el texto, proporcionando una visión detallada de las respuestas en contextos como las conferencias telefónicas. Por ejemplo, se pueden generar puntuaciones de sentimiento y alegría en las transcripciones de conferencias telefónicas trimestrales, demostrando que estas características mejoran en gran medida la capacidad del modelo para predecir debilidades en el control interno, en comparación con otros modelos que solo utilizan determinantes tradicionales (Sun & Vasarhelyi, 2018).

### **Argumentos finales**

Antes de la adopción generalizada de la inteligencia artificial, el proceso de planificación de auditoría solía depender en gran medida de sesiones de *brainstorming* con el cliente para evaluar los riesgos empresariales. Durante estas sesiones, el equipo de auditoría interactuaba con la dirección y el personal, revisaba las actas de la reunión de la Junta Directiva y utilizaba listas de verificación estructuradas para guiar la evaluación del riesgo. Esta metodología tenía limitaciones estructurales porque las listas de verificación restringían la toma de decisiones del auditor y podían inducir sesgos al limitar la evaluación de riesgos a los criterios predefinidos en la lista. Las notas tomadas durante estas reuniones también tenían limitaciones en términos de la amplitud de los elementos considerados (Sheils, 2020).

Con la incorporación de la inteligencia artificial en el proceso de auditoría (propone Sheils), se ha abierto la posibilidad de evaluar los riesgos de manera más integral, superando las limitaciones de las listas de verificación mediante el uso de tecnologías de asistentes cognitivos. Con la inteligencia artificial, las sesiones de *brainstorming* durante la planificación de auditoría, se pueden realizar con técnicas de procesamiento del lenguaje natural y aprendizaje automático. Esta herramienta es capaz de analizar los datos recopilados durante las sesiones de intercambio de ideas y presentarlos de manera estructurada y legible por máquina, facilitando una evaluación más completa y menos limitada por criterios predefinidos, haciendo que el auditor considere una gama más amplia de posibilidades y se mejore el proceso de toma de decisiones durante la planificación de auditorías.

Como indican Žager *et al.* (2020), la aplicación de la inteligencia artificial en contabilidad y auditoría, que empezó a producirse hace aproximadamente tres décadas y desde entonces ha experimentado una evolución significativa, ha transitado desde la investigación teórica hasta la práctica, demostrando que hay un amplio potencial para integrarla en los procesos contables. A pesar de los avances, aún persisten muchas incógnitas y desafíos por abordar, donde los autores se plantean preguntas para implementar efectivamente la IA en los negocios, y sobre el desarrollo de marcos regulatorios que guíen su uso, donde surgen preocupaciones legales y éticas sobre la responsabilidad por los errores que resulten de la contabilidad y la auditoría, por la necesidad de un enfoque cuidadoso y reflexivo en la integración de la IA en el campo contable, asegurando al mismo tiempo que se maximicen los beneficios potenciales mientras se mitigan los riesgos asociados.

## Conclusiones

Los resultados del estudio han planteado brevemente aspectos descriptivos sobre el tratamiento de los datos no estructurados en la auditoría, especificando la clasificación y el manejo de datos, destacando el modo en que influye el Big Data en la auditoría, planteando cómo los datos estructurados y no estructurados están transformando los métodos tradicionales de llevar el proceso auditor.

Para una mejor comprensión, se revisaron las categorías de datos para diferenciar aquellos de máquina, de los datos espaciotemporales, y datos genómicos, por cuanto cada uno representa un uso y oportunidad valiosos para los auditores. Siendo la inteligencia artificial una herramienta que emerge para suplir la capacidad del individuo ante ciertas tareas y actividades, la exposición de los sistemas expertos en el análisis de datos no estructurados destacan su capacidad para mejorar la calidad de la auditoría y detectar irregularidades financieras de manera más efectiva.

Manejar grandes volúmenes de datos no estructurados siempre va a plantear riesgos de seguridad y privacidad, especialmente cuando se trata de información sensible o confidencial. Esta es una razón más que válida para proteger estos datos de accesos no autorizados, y un factor importante que está haciendo que los contadores y auditores tengan que plantearse la adquisición de nuevas habilidades y conocimientos técnicos para trabajar con herramientas avanzadas de inteligencia artificial y análisis de datos.

Tal como se mostró en la revisión de la literatura, la capacitación en estas áreas es necesaria para aprovechar al máximo las tecnologías emergentes y mejorar las competencias del auditor pues, aunque los datos no estructurados pueden contener información valiosa para la contabilidad, su procesamiento y análisis presentan diversos inconvenientes que deben ser tratados mediante el uso de tecnologías avanzadas, lo que supone también la capacitación adecuada del personal.

## Referencias

- Adnan, K., & Akbar, R. (2019). An analytical study of information extraction from unstructured and multidimensional big data. *Journal of Big Data*, 6(1), 1-38. <https://doi.org/10.1186/s40537-019-0254-8>
- Becker, J., Knackstedt, R. & Pöppelbuß, J. (2009). Developing Maturity Models for IT Management. *Business & Information Systems Engineering*, 1(3), 213–222. <https://doi.org/10.1007/s12599-009-0044-5>
- Bose, S., Dey, S. K., & Bhattacharjee, S. (2023). Big data, data analytics and artificial intelligence in accounting: An overview. *Handbook of big data research methods*, 32-51. <https://acortar.link/yuKfTn>
- Bridwater, A. (2018). The 13 Types Of Data. In *Forbes*. <https://www.forbes.com/sites/adrianbridwater/2018/07/05/the-13-types-of-data/>

- Chowdhury, E. K. (2021). Prospects and challenges of using artificial intelligence in the audit process. *The Essentials of Machine Learning in Finance and Accounting*, 139-156. <http://dx.doi.org/10.4324/9781003037903-8>
- De La Hoz Suárez, B. A., Manjarres Tete, A. E., De La Hoz Suárez, A. I. & Luna Morán, I. F. (2024). Inteligencia artificial como estrategia para gestionar los procesos de auditoría financiera. *Revista Estrategia Organizacional*, 13(1), 57-72. <https://ideas.repec.org/a/col/000577/021148.html>
- Fisher, I. E., Garnsey, M. R., & Hughes, M. E. (2016). Natural language processing in accounting, auditing and finance: A synthesis of the literature with a roadmap for future research. *Intelligent Systems in Accounting, Finance and Management*, 23(3), 157-214. <https://doi.org/10.1002/isaf.1386>
- Fukas, P., Rebstadt, J., Remark, F., & Thomas, O. (2021). Developing an Artificial Intelligence Maturity Model for Auditing. In *ECIS, A Virtual AIS Conference*, 1-18. <https://acortar.link/1gNZKA>
- Fundación Latinoamericana de Auditores Internos. (2022). *Perspectivas y percepciones globales*. The Institute of Internal Auditors, Inc.
- Gepp, A., Linnenluecke, M. K., O'Neill, T. J., & Smith, T. (2018). Big data techniques in auditing research and practice: Current trends and future opportunities. *Journal of Accounting Literature*, 4(1), 102-115. <https://doi.org/10.1016/j.acclit.2017.05.003>
- Hasan, A. R. (2021). Artificial Intelligence (AI) in accounting & auditing: A Literature review. *Open Journal of Business and Management*, 1(1), 440-465. <https://doi.org/10.4236/ojbm.2022.101026>
- Hevner, A.R., Ram, S. & March, S.T. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1)75-105. <https://doi.org/10.2307/25148625>
- Hou, B., Zhang, Y., Shang, Y., Liang, X., Liu, T., & Su, J. (2020, October). Research on unstructured data processing technology in executing audit based on big data budget. In *Journal of Physics: Conference Series*, 1650(3), 032100. IOP Publishing. <https://doi.org/10.1088/1742-6596/1650/3/032100>
- Michael, A., & Dixon, R. (2019). Audit data analytics of unregulated voluntary disclosures and auditing expectations gap. *International Journal of Disclosure and Governance*, 16(4), 188-205. <https://doi.org/10.1057/s41310-019-00065-x>
- Montoya Hernández, A. Y. & Valencia Duque, F. J. (2020). Inteligencia artificial al servicio de la auditoría: una revisión sistemática de literatura. *Revista Ibérica de Sistemas e Tecnologías de Informação*, (E27), 213-226. <https://www.proquest.com/openview/8a2868ccf43245be9a642a31d5454ca4/1?pq-origsite=gscholar&cbl=1006393>
- Nielsen, S. (2022). Management accounting and the concepts of exploratory data analysis and unsupervised machine learning: a literature study and future directions. *Journal of Accounting & Organizational Change*, 18(5), 811-853. <https://doi.org/10.1108/JAOC-08-2020-0107>
- Omotoso, K. (2012). The application of artificial intelligence in auditing: Looking back to the future. *Expert Systems with Applications*, 39(9), 8490-8495. <https://doi.org/10.1016/j.eswa.2012.01.098>
- Ortiz de Zárate Alcarazo, L. (2022). Explicabilidad (de la inteligencia artificial). *Eunomía, Revista en Cultura de la Legalidad*, 22, 328-344.
- Salinas Rodríguez, S. I. (2023). Textarium: Un sistema de visualización de textos utilizando conocimiento libre y procesamiento de lenguaje natural. (Tesis de Ingeniería). Universidad de Chile.
- Sheils, B. P. (2020). *Investigación en la aplicación de técnicas de Inteligencia Artificial (IA) en la industria de la auditoría*. Universidad Pontificia Comillas.

- Sun, T., & Vasarhelyi, M. A. (2018). Embracing textual data analytics in auditing with deep learning. *International Journal of Digital Accounting Research*, 18, 49-67. [https://doi.org/10.4192/1577-8517-v18\\_3](https://doi.org/10.4192/1577-8517-v18_3)
- Török, R. M. (2022). Artificial intelligence algorithms applied in business and accounting. *Timisoara Journal of Economics and Business*, 15(1), 73-90. <https://doi.org/10.2478/tjeb-2022-0005>
- Yang, M., Adomavicius, G., Burtch, G., & Ren, Y. (2018). Mind the gap: Accounting for measurement error and misclassification in variables generated via data mining. *Information Systems Research*, 29(1), 4-24.
- Yudowati, S., & Alamsyah, A. (2018). Big data framework for auditing process. *International Journal of Engineering & Technology*, 7(4.38), 908-911. <https://doi.org/10.1287/isre.2017.0727>
- Žager, K., Dečman, N., & Rep, A. (2020). The impact of artificial intelligence on the accounting process. In *9th International Scientific Symposium Region, Entrepreneurship, Development 2020 (RED 2020)* (1430-1444). [https://www.researchgate.net/publication/342171038\\_The\\_Impact\\_of\\_Artificial\\_Intelligence\\_on\\_the\\_Accounting\\_Process](https://www.researchgate.net/publication/342171038_The_Impact_of_Artificial_Intelligence_on_the_Accounting_Process)