

# APLICACIÓN DE AUDITORÍA DE SISTEMAS A LOS SERVICIOS DE INTEGRACIÓN DE DATOS CON PROCESOS ETL

Oscar Orlando Galvis Cifuentes<sup>1</sup>

## Resumen

El auditor debe considerar como parte de su labor, la aplicación del conocimiento sobre las características de uso, diseño y automatización de los procesos ETL (Extract, Transform and Load), las cuales se resumen en extracción, transformación y cargue de datos.

Adicionalmente, los escenarios de riesgo establecidos conforme a la Norma Técnica Colombiana NTC-ISO 31000, sirven de guía, no solo al diseñador del ETL, sino también al auditor para identificar los posibles eventos generadores de riesgo y por ende la aplicación de controles para su mitigación. A partir del marco de referencia COBIT, se establecen aspectos y criterios importantes para la evaluación de los procesos ETL y sus controles. También se presenta una lista de buenas prácticas que un auditor de sistemas debe considerar al realizar su evaluación a los servicios de integración de datos.

## Palabras clave

ETL, Servicio de integración de datos, Auditoría de sistemas, Interoperabilidad, Riesgos y controles.

## Abstract

---

<sup>1</sup> Ingeniero de sistemas de la Universidad Antonio Nariño en Bogotá Colombia. Estudiante de último semestre de la Especialización en auditoría de sistemas en la misma universidad. [ogalvis35@uan.edu.co](mailto:ogalvis35@uan.edu.co)

The auditor should consider as part of his work, the application of knowledge on the characteristics of use, design and automation of ETL processes, which are summarized in extraction, transformation and data loading. Additionally, the risk scenarios established in accordance with the Colombian Technical Standard NTC-ISO 31000, serve as a guide, not only for the ETL designer, but also for the auditor to identify possible risk generating events and therefore the application of controls for their mitigation. Based on the COBIT reference framework, important aspects and criteria are established for the evaluation of ETL processes and their controls. It also presents a list of good practices that a systems auditor should consider when evaluating data integration services.

### **Key words**

ETL, Data integration service, Systems Audit, Interoperability, Risks and controls

### **Introducción**

Los servicios de integración de datos hacen referencia a la funcionalidad de realizar trasposos de información entre diferentes fuentes y realizar transformaciones, operaciones y agrupaciones de datos. Su uso se ha generalizado en migraciones de datos o tareas repetitivas que impliquen transportar información entre diferentes fuentes, que pueden estar dentro de la misma entidad o fuera de ella, e inclusive distantes geográficamente.

La auditoría de sistemas a este tipo de servicios de integración implica que los miembros del equipo auditor tengan conocimiento de la forma en que estos procesos funcionan, los riesgos que tiene su procesamiento y los controles que comúnmente se utilizan para mitigarlos. Sin embargo, el rol del auditor debe ir más allá de aplicar este conocimiento: debe tener la capacidad de realizar una buena evaluación al proceso, sobre

todo a los controles que mitigan los riesgos. Debe usar su experiencia, criterios, imparcialidad e independencia para generar el diagnóstico correcto y brindar las sugerencias del caso.

Por lo tanto, este artículo pretende responder la pregunta, ¿Cuáles serían las mejores prácticas de auditoría de sistemas que el auditor debe considerar en su revisión de los procesos que usan servicios de integración de datos?

Para cumplir con este propósito, se utilizarán fuentes de información relacionadas con procesos ETL (Extract, Transform and Load), procesos de interoperabilidad y de servicios de integración de datos.

## **Objetivos**

### **Objetivo general**

Determinar las mejores prácticas de una auditoría de sistemas en la evaluación de los controles implementados para los servicios de integración de datos.

### **Objetivos específicos**

Describir las características de uso, diseño y automatización de los servicios de integración de datos que usan procesos ETL.

Identificar escenarios de riesgos y controles presentes en los servicios de integración de datos que usan procesos ETL.

Establecer los aspectos, criterios y buenas prácticas de la auditoría de sistemas para la evaluación de los controles relacionados con el diseño, pruebas y automatización de servicios de integración de datos que usan procesos ETL.

## **Metodología**

Se usará el método analítico, debido a que es necesario descomponer los elementos del proceso de auditoría para confrontarlos con las particularidades que tiene la funcionalidad de los servicios de integración y así tener bases para establecer las mejores prácticas en la realización de una auditoría. Respecto al tipo de estudio, será explicativo puesto que se busca exponer las buenas prácticas en una auditoría a los servicios de integración de datos.

## **Marco contextual**

El concepto de *servicio de integración de datos* es utilizado en los ámbitos informáticos y hace referencia a aquellos aplicativos o procesos técnicos en los que se mueve información de una fuente a otra.

Los servicios de integración de datos también son conocidos con la sigla ETL, que significa Extract, Transform and Load. Hace referencia a cada uno de los pasos que compone un proceso de integración de datos: extraer información de las fuentes, hacer transformación a los datos y cargar el resultado sobre una bodega final de datos.

### **Aplicación de los servicios de integración de datos**

Los servicios de integración de datos tienen como propósito reunir en un solo sitio la información necesaria, para luego poder realizar nuevos procesos para la aplicación de estos nuevos datos como, por ejemplo:

*Minería de datos*: de acuerdo con Perez Lopez & Santín González (2007), la disponibilidad de grandes volúmenes de información da origen a la posibilidad de realizar análisis de dichos datos a lo cual se le conoce como minería de datos o datamining.

*Migración a nuevos procesos*: se describe como el paso de información de un sistema a otro debido a factores como actualización de software y/o hardware, reubicación de los medios de almacenamiento y otros que obliguen al movimiento de la información.

*Análisis de comportamiento y tendencias*: es quizá una de las funcionalidades más empleadas al usar técnicas de minería de datos. “Permiten extraer patrones, tendencias y regularidades para describir y comprender mejor los datos y para predecir comportamientos futuros” (Perez Lopez & Santín González, 2007, pág. 2)

### **Aplicativos comunes para diseñar e implementar servicios de integración de datos**

En el mercado del software, existe gran variedad de aplicativos orientados a ETL, cada uno con características, limitaciones y costos, además de incluir otras funcionalidades para complementar los procesos de extracción, transformación y carga de datos. A continuación, a modo de ejemplo, se enuncian tres aplicativos de uso común:

*SSIS (SQL Server Integration Services)*: Creada por Microsoft y definida como:

Microsoft Integration Services es una plataforma para construir soluciones de integración y transformación de datos a nivel empresarial. Use Integration Services para resolver problemas complejos del negocio copiando o descargando archivos, cargando bodegas de datos, limpieza y minería de datos y gestionando los datos y objetos de SQL Server. (Microsoft, 2018)

*Cognos Analytics*: es una herramienta de IBM. “Cognos Analytics es una plataforma de inteligencia empresarial alimentada por inteligencia artificial que da soporte a todo el ciclo de analítica, desde el descubrimiento hasta la operacionalización.” (IBM, s.f.)

*Oracle Warehouse Builder*: es la herramienta exclusiva de Oracle para el uso de los servicios de integración de datos. En Oracle la presentan como:

Oracle Data Integrator (ODI) es la contribución de Oracle para la integración de datos estratégicos a largo plazo. Warehouse Builder Enterprise ETL 11gR2 (OWB-EE) actualiza más ampliamente el diseño del almacén de datos, ETL y contribución a la calidad de los datos para la base de datos Oracle con infraestructura de integración de datos basada en ODI (Oracle, 2010)

## **Procesos ETL y la interoperabilidad**

La interoperabilidad, en el ámbito de la tecnología, se puede definir como:

La capacidad de un sistema de información de comunicarse y compartir datos, información, documentos y objetos digitales de forma efectiva (con una mínima o nula pérdida de su valor y funcionalidad), con uno o varios sistemas de información (siendo generalmente estos sistemas completamente heterogéneos, distribuidos y geográficamente distantes), mediante una interconexión libre, automática y transparente, sin dejar de utilizar en ningún momento la interfaz del sistema propio. (Gomez, 2007, pág. 28)

Esto sugiere, que los procesos ETL son interoperables, porque cumplen con el objetivo de intercambiar datos con mínima pérdida de información y entre diferentes fuentes, generando a su vez nuevos conocimientos.

## **La auditoría de sistemas y los procesos ETL**

Como en todo proceso tecnológico, los servicios de integración de datos también contienen riesgos que pueden generar diferentes impactos sobre los procesos, la información y la bodega de datos. Para mitigar los efectos de esos riesgos se debe complementar el diseño y/o implementación del ETL con controles que tengan la capacidad de manejar los eventos que generan los riesgos y disminuir de esta forma el impacto sobre los procesos.

Quienes diseñan e implementan procesos ETL, deben conocer los requerimientos, las entradas de información, las transformaciones a realizar y el resultado esperado. Por lo tanto, la gestión y análisis de esos riesgos y controles son de su responsabilidad.

Sobre el tema de controles, vale la pena complementar que estos pueden ser de tres tipos, tal como se describe en la tabla 1.

**Tabla 1** *Tipología de Controles Según su Clase*

Tipo	Descripción
<b>Preventivo</b>	Son aquellos controles incluidos como parte del proceso y que tienen por objetivo evitar la frecuencia y ocurrencia de las desviaciones
<b>Detectivo</b>	Son aquellos que se activan cuando se materializa un riesgo, registran los eventos en logs y los notifican para que un funcionario los pueda revisar y determinar las correcciones del caso.
<b>Correctivo</b>	Son aquellos que se implementan para dar solución a los incidentes que se detectan en el proceso.

Fuente: elaboración propia, basado en (Tamayo Alzate, 2001, págs. 18-19)

### Modelos y normativas

En el desarrollo se emplearán los modelos y normas descritos en la tabla 2, las cuales determinan el marco para la auditoría a los servicios de integración de datos.

**Tabla 2** *Modelos y Normas*

Modelo	Aplicación
<b>COBIT 2019</b>	Se determinan los criterios para tener en cuenta en la auditoría a procesos ETL. COBIT es definido como “un marco de referencia para el gobierno y la gestión de la información y la tecnología, dirigido a toda la empresa” (ISACA, 2019, pág. 9)
<b>NTC-ISO 31000</b>	Se emplea como guía para el manejo y gestión de los escenarios de riesgo presentes en los servicios de integración de datos. Esta norma técnica “brinda los principios y directrices genéricas sobre la gestión del riesgo” (ICONTEC, 2011, pág. 3)
<b>ISO 27001</b>	Brinda las pautas y directrices relacionadas con la seguridad de la información. Esta norma la definen como: La norma ISO 27001 es una solución de mejora continua en base a la cual puede desarrollarse un Sistema de Gestión de Seguridad de la Información (SGSI) que permita evaluar todo tipo de riesgos o amenazas susceptibles de poner en peligro la información de una organización tanto propia como datos de terceros. (ISOTools, pág. 4)

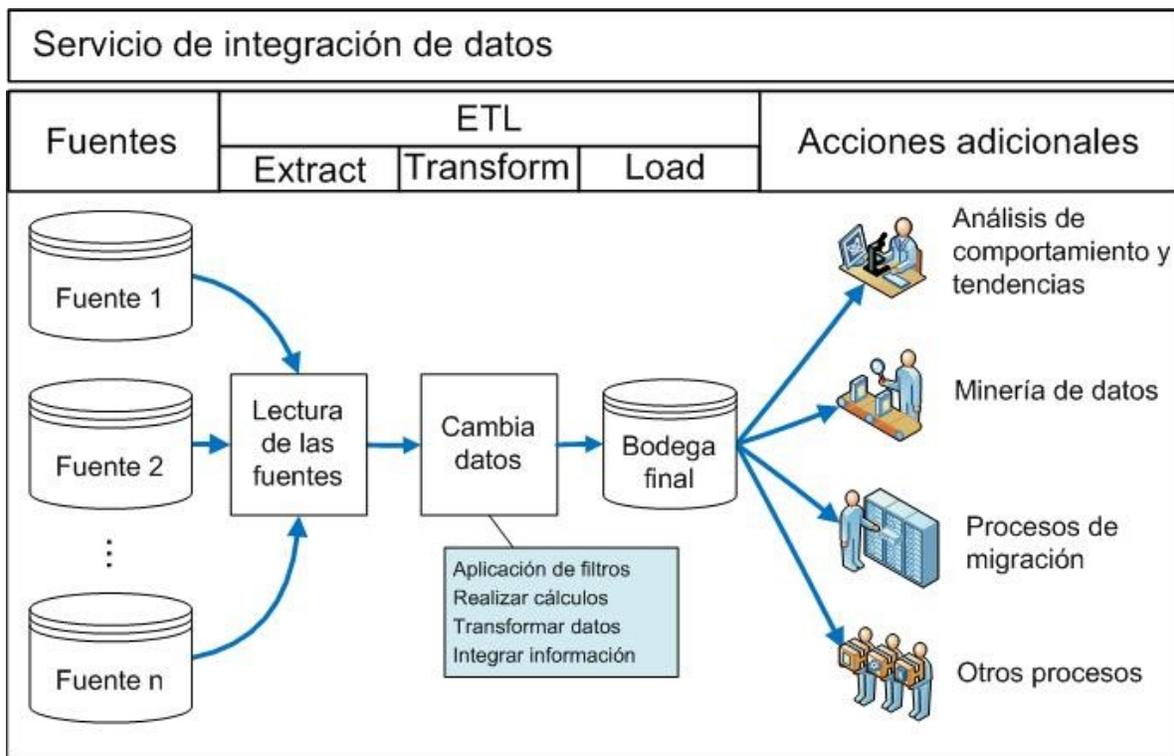
Fuente: Elaboración propia.

## Resultados

### Características de uso, diseño, desarrollo y automatización de los servicios de integración de datos que usan procesos ETL

El auditor de sistemas deberá estar en capacidad de comprender, analizar y diagnosticar cada proceso que compone el ETL implementado, sobre todo en lo referente a los controles asociados para mitigar sus riesgos. Por ello, es necesario que además del conocimiento y experiencia en auditoría, el auditor posea conocimientos sobre los servicios de integración de datos y tenga los suficientes criterios para emitir un buen diagnóstico basado en una medición confiable siguiendo modelos y normas de aceptación general.

En la figura 1 se describe el flujo que tiene un servicio de integración de datos.



**Figura 1** Modelo del Servicio de Integración de Datos

Fuente: Elaboración propia (Ilustraciones de Microsoft Visio)

De acuerdo con la figura 1, el proceso ETL requiere como entradas una serie de fuentes, que pueden ser archivos o bases de datos, para posteriormente transformarlos y cargarlos en la bodega final. Esta bodega sirve a su vez de nueva fuente para otros procesos como minería de datos, migraciones y/o análisis de comportamientos y tendencias, entre otros.

Por consiguiente, el proceso ETL implica la ejecución de tres acciones: extraer, transformar y cargar y a su vez cada acción involucra la realización de unas tareas. Estas acciones y tareas se describen en la tabla 3.

**Tabla 3** *Relación de las Actividades a Realizar en los Procesos ETL*

Acción	Descripción	Tareas
<b>Extract</b>	Extracción de los datos desde las fuentes	<i>Tarea 1:</i> identificar las fuentes de datos de las cuales se realizará la extracción de la información <i>Tarea 2:</i> Realizar la lectura de la información contenida en las fuentes.
<b>Transform</b>	Transformación de los datos que fueron extraídos	<i>Tarea 3:</i> teniendo los datos leídos, se pueden transformar generando nuevos datos a través de procesos como filtros, cálculos, entre otros.
<b>Load</b>	Carga de los datos transformados a la bodega de almacenamiento final	<i>Tarea 4:</i> unificar la información. <i>Tarea 5:</i> selección del destino final de los datos. <i>Tarea 6:</i> unificar los atributos de la información con los atributos de la bodega de datos final. <i>Tarea 7:</i> comprende el paso de la información a la bodega de datos final.

Fuente: Elaboración propia basada en la explicación de (Duque Méndez, Hernández Leal, Pérez Zapata, Arroyabe Tabares, & Espinosa, 2016)

Por otro lado, los procesos ETL tienen generalmente una actividad final en donde se realiza la automatización de su ejecución de acuerdo con criterios de tiempo o eventos del sistema.

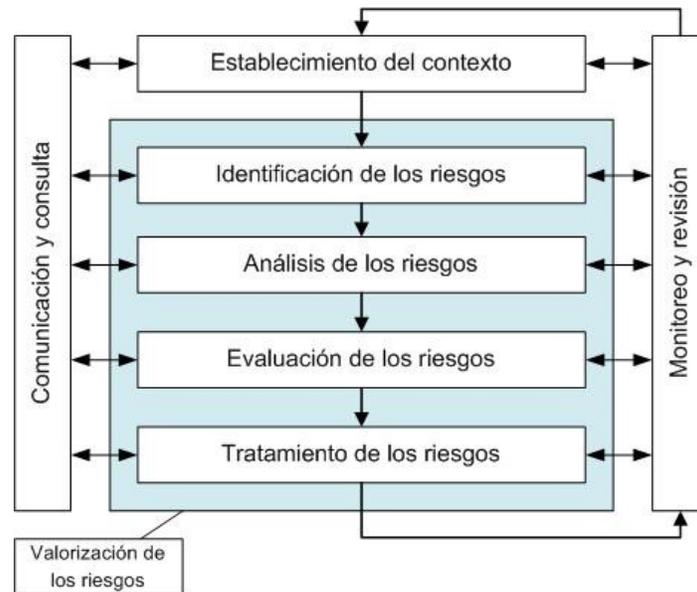
Los aplicativos de software sobre servicios de integración de datos tienen rutinas que permiten esa automatización. Como ejemplo, en la figura 2 se muestran las opciones de programación automática de un proceso ETL en SSIS:

**Figura 2** *Opciones de Programación en SQL para Procesos ETL*

Fuente: Captura de pantalla en Microsoft SQL Server Management Studio

## **Escenarios de riesgo y controles presentes en los servicios de integración de datos que usan procesos ETL**

Esta sección del artículo se desarrolla con la metodología de la Norma Técnica Colombia NTC-ISO 31000. En la figura 3, se muestra el esquema de la gestión de riesgo sugerida por esta norma.



**Figura 3** Proceso para la Gestión de los Riesgos

Fuente: Elaboración propia basada en los procesos para la gestión de riesgos documentados en la Norma Técnica Colombiana NTC-ISO 31000 (ICONTEC, 2011, pág. 17)

Cada una de las etapas del proceso para la gestión de riesgos, es explicada en la tabla 4 dándoles la orientación respectiva hacia la aplicación en procesos ETL. La responsabilidad de la ejecución de estas etapas es del dueño del proceso ETL y no del auditor.

**Tabla 4** Descripción de Etapas de la Gestión de Riesgos sobre ETL

Actividad	Definición	Explicación sobre ETLs
<b>Establecer el contexto</b>	“La organización articula sus objetivos, define los parámetros externos e internos que se van a considerar al gestionar el riesgo y establece el alcance y los criterios del riesgo para el resto del proceso” (ICONTEC, 2011, pág. 18)	Establecer aquellos factores o escenarios internos y externos que pueden ser generadores de riesgo para el proceso ETL.
<b>Identificar los riesgos</b>	“El objeto de esta fase es generar una lista exhaustiva de riesgos con base en aquellos eventos que podrían crear, aumentar, prevenir, degradar, acelerar o retrasar el logro de los objetivos.” (ICONTEC, 2011, pág. 21)	A partir de los escenarios de riesgo especificados y aquellos propios del proceso ETL, se debe realizar la identificación de los riesgos.

Actividad	Definición	Explicación sobre ETLs
<b>Análisis de riesgos</b>	“El análisis del riesgo involucra la consideración de las causas y las fuentes de riesgo, sus consecuencias positivas y negativas, y la probabilidad de que tales consecuencias puedan ocurrir.” (ICONTEC, 2011, pág. 21)	Se deben determinar las consecuencias sobre el proceso ETL en caso que se materialicen los riesgos y la probabilidad de su ocurrencia.
<b>Evaluación de riesgos</b>	“Implica la comparación del nivel de riesgo observado durante el proceso de análisis y de los criterios del riesgo establecidos al considerar el contexto. Con base en esta comparación, se puede considerar la necesidad de tratamiento.” (ICONTEC, 2011, pág. 22)	Evaluar la severidad del riesgo y considerar los posibles tratamientos para mitigar el impacto sobre el proceso ETL.
<b>Tratamiento de riesgos</b>	“Involucra la selección de una o más opciones para modificar los riesgos y la implementación de tales opciones. Una vez implementado, el tratamiento suministra controles o los modifica.” (ICONTEC, 2011, pág. 22)	Determinar los tratamientos que se deben dar a los riesgos considerando la disminución del impacto sobre el proceso ETL.
<b>Comunicar y consultar</b>	“La comunicación y la consulta con las partes involucradas externas e internas deberían tener lugar durante todas las etapas del proceso para la gestión del riesgo.” (ICONTEC, 2011, pág. 17)	Conciérne a mantener informado al equipo encargado de la gestión de riesgos sobre los procesos ETL.
<b>Monitoreo y revisión</b>	De acuerdo con la NTC-ISO 31000 (2011), es una parte esencial del proceso para la gestión del riesgo que pretende mantenerlo en vigilancia y mejorar la valorización del riesgo.	El equipo encargado de la gestión del riesgo sobre procesos ETL, debe realizar el proceso de monitoreo y revisión.

Fuente: elaboración propia basada en las definiciones de la NTC-ISO 31000 para cada una de las etapas del proceso de gestión de riesgos.

Los escenarios de riesgo hacen referencia a aquellos ámbitos internos o externos de la organización que pueden ser generadores de riesgo. Se identifican durante la primera etapa del proceso de gestión de riesgos (establecimiento del contexto) en la NTC-ISO 31000.

Particularmente para lo referente a los servicios de integración de datos con ETLs, se sugieren a continuación un conjunto universal de escenarios de riesgo, los cuales se describen en la tabla 5; sin embargo, el auditor y/o gestor de los riesgos debe considerar si los aplica o no y adicionar los que considere oportunos de acuerdo con el alcance y funcionalidad de su ETL.

**Tabla 5** *Escenarios de Riesgo en los Procesos ETL*

Escenario	Descripción
<b>Fuentes de la información</b>	<p>Corresponde a los archivos y/o bases de datos desde donde se extrae la información que será usada en el proceso ETL. Se deben considerar situaciones como:</p> <ul style="list-style-type: none"> <li>• Tipos de archivos y bases de datos compatibles con el proceso</li> <li>• Tamaño de los archivos y/o bases de datos a leer</li> <li>• Complejidad en la lectura de los archivos.</li> </ul>
<b>Datos contenidos en las fuentes de información</b>	<p>Hace referencia a la forma, calidad e integridad de los datos que están contenidos en las fuentes. Se deben considerar situaciones como:</p> <ul style="list-style-type: none"> <li>• Formatos de fecha, ¿las fechas tienen siempre el mismo formato?</li> <li>• Separadores de miles y de decimales en los datos numéricos.</li> <li>• Integridad de la información.</li> </ul>
<b>Formas de relación entre diferentes fuentes de información</b>	<p>Implica la forma de poder relacionar tablas o datos entre diferentes archivos y/o bases de datos. Sucede en casos donde se debe complementar la información de una fuente con la existente en otra. Se deben considerar factores como:</p> <ul style="list-style-type: none"> <li>• Integridad de las llaves que se usan para la relación.</li> <li>• Información repetida entre las fuentes.</li> </ul>
<b>Cálculos sobre los datos leídos</b>	<p>Son las operaciones matemáticas, estadísticas, trigonométricas, etc. que se realizan sobre la información leída y recopilada de las fuentes. Se deben tener en cuenta situaciones como:</p> <ul style="list-style-type: none"> <li>• Validez de los datos para los cálculos.</li> <li>• Validaciones previas a los cálculos (Ejemplo, en una división validar que el divisor sea diferente a cero, de lo contrario generará error)</li> <li>• Uso adecuado de las funciones matemáticas que tiene la herramienta ETL con la que se está trabajando.</li> </ul>
<b>Filtros a la información</b>	<p>Es la manera en que se descarta información porque no tiene validez para el proceso, pero se respeta aquella que si tiene relevancia para el objeto de estudio. Se deben considerar elementos como:</p> <ul style="list-style-type: none"> <li>• Variables empleadas para la definición de los filtros.</li> </ul>

Escenario	Descripción
	<ul style="list-style-type: none"> <li>• Uso adecuado de las funciones y/o instrucciones que tiene la herramienta ETL para la aplicación de filtros.</li> </ul>
<b>Agrupamiento de datos</b>	<p>Es la generación de nuevos datos a partir de la consolidación de resultados. Por ejemplo, generar los totales de venta por mes de cada regional a partir un archivo de Excel que contiene cada una de las facturas emitidas en un año contable de todas la regionales. Se deben tener en cuenta elementos como:</p> <ul style="list-style-type: none"> <li>• Criterios para la agrupación de los datos y la generación de nueva información</li> <li>• Validez de los datos que son clave para su agrupación.</li> <li>• Uso apropiado de las funciones y/o instrucciones que la herramienta ETL ofrece para permitir la agrupación de datos.</li> </ul>
<b>Bodega de almacenamiento de datos</b>	<p>Corresponde al archivo o base de datos donde se hace la carga de la información resultante del proceso de lectura y transformación. Se deben considerar elementos como:</p> <ul style="list-style-type: none"> <li>• Compatibilidad del tipo de archivo o base de datos con el proceso ETL.</li> <li>• Distribución de los datos resultantes en el archivo o base de datos.</li> </ul>
<b>Seguridad de la información</b>	<p>Comprende todos los aspectos relacionados con garantizar la seguridad de la información durante el flujo que tiene el proceso ETL. De acuerdo con la normativa ISO 27001, se deben considerar aspectos como: “La elaboración de un plan de contingencia frente a los incidentes, el análisis de riesgos, las competencias, el grado de involucración de la Dirección, las inversiones en seguridad y el grado de implementación de controles.” (ISOTools, pág. 3)</p>

Fuente: elaboración propia.

### **Aspectos, criterios y buenas prácticas de la auditoría de sistemas para la evaluación de los controles aplicados a los procesos ETL**

En la etapa del tratamiento de riesgo, que tiene el proceso de gestión de la NTC-ISO 31000, se debe considerar la aplicación de controles con el fin de disminuir el impacto que pueda llegar a presentar la materialización del riesgo. Estos controles, como se explicaron anteriormente, pueden ser preventivos, detectivos y/o correctivos.

En la tabla 6 se presenta un conjunto de controles de tipo preventivo que, desde la experiencia del autor, son de utilidad para quien se encuentra diseñando e implementado un proceso ETL. Se pueden considerar otros controles adicionales a criterio del dueño del proceso. El rol del auditor es validar el buen uso de los controles y confirmar su eficiencia.

**Tabla 6** *Controles Preventivos para los Procesos ETL*

CONTROL	EXPLICACIÓN
<b>Unificación de formatos de fecha</b>	Dada la posibilidad de existir fechas en formatos diferentes, se sugiere implementar procesos que permitan tomar cada fecha y formatearla a un único formato para su uso.
<b>Unificar formatos numéricos</b>	Pueden existir diferencias entre la forma de separación de miles y decimales en los datos numéricos. Por ello, se sugiere implementar un proceso de unificación tomando cada dato y dándole el formato que se usará en el resto del proceso ETL.
<b>Ajustar el tamaño de las cadenas de texto</b>	Cuando el destino de un dato tipo texto es un campo de una tabla, se recomienda controlar que el tamaño del texto no sobrepase la longitud definida en el campo de la tabla.
<b>Validación de datos</b>	Antes de realizar transformaciones o cálculos con los datos, se sugiere efectuar procesos de validación a los mismos. Aquellos datos que no superen las validaciones deben ser notificadas. Para ello, puede emplear las funcionalidades de notificación que pueda tener la herramienta ETL.

Fuente: elaboración propia.

Respecto a los controles detectivos, estos deben ser implementados sobre el proceso ETL de tal forma que pueda generar información adicional sobre el riesgo materializado; por ejemplo, un log de eventos o un archivo de seguimiento (trace). Se debe tener en cuenta que los controles correctivos se aplican cuando se presenten incidentes.

Para finalizar el tema de controles, se presenta en la tabla 7 información de ejemplo sobre algunos riesgos de los servicios de integración de datos, controles que pueden incluirse y el monitoreo que debe efectuar el gestor y/o auditor para garantizar su funcionamiento.

**Tabla 7:** Ejemplo de Riesgos, Controles y Monitoreo

RIESGO	CONTROLES	MONITOREO	FRECUENCIA
Formatos de fecha diferentes entre las fuentes de información.	Implementar funciones dentro del diseño o código fuente para unificar los formatos de fecha	Mediante la técnica de inyección de datos, ejecutar pruebas sobre la función enviando como parámetros fechas en diferentes formatos y corroborar que las retorne en el formato esperado.	Si el proceso ETL se ejecuta frecuentemente, se sugiere realizar el monitoreo cada 2 meses.
Cambios repentinos en la estructura de los archivos o bases de datos que son fuente de información para el proceso	Validaciones previas sobre la estructura esperada del archivo. Si presenta modificaciones debe notificar y suspender la ejecución del proceso	Mediante la técnica de simulación, hacer cambios intencionales a la estructura de los archivos o bases de datos para comprobar la efectividad de las validaciones y verificar que haya efectuado la notificación.	Si el proceso ETL se ejecuta frecuentemente, se sugiere realizar el monitoreo cada 2 meses
La información resultante del proceso de integración fue calculada de forma errónea.	Pruebas integrales al proceso completo.	Usar herramientas de software especializadas en análisis de datos, para confirmar que los algoritmos implementados estén realizando correctamente los cálculos a los datos.	Se sugiere realizar el monitoreo durante la implementación y pruebas del proceso ETL.

Fuente: elaboración propia

Para complementar la información sobre la herramienta especializada en análisis de datos que se menciona en la tabla 7, se puede tomar como ejemplo el software IDEA, el cual es descrito como:

IDEA es una poderosa herramienta para Análisis, Extracción y Auditoría de Datos, que provee funcionalidades para aplicar numerosas “buenas y mejores prácticas de Auditoría Asistida con Computador, universalmente aceptadas para Auditorías Basadas en Datos o Transacciones”. Es fácil de utilizar (amigable), con funciones especialmente diseñadas para verificar la calidad e integridad de la información de

bases de datos y archivos de computador, analizar y clasificar los datos aplicando criterios de acuerdo con las reglas del negocio, automatizar técnicas de auditoría asistidas con el computador (CAATs), generación de reportes y gráficos, exportar archivos y enviar correos electrónicos desde el software IDEA. (Audisis)

Por otra parte, el auditor de sistemas debe considerar en su labor, aspectos y criterios propios de un desarrollo de software. Para ello, el marco de referencia COBIT en su versión 2019 brinda los lineamientos necesarios para auditar este tipo de servicios tecnológicos. En la tabla 8 se muestran los objetivos de gestión que aplican para la auditoría a procesos ETL.

**Tabla 8** *Objetivos de Gestión y su Aplicabilidad en Procesos ETL*

OBJETIVO DE GESTIÓN	DESCRIPCIÓN
<b>BAI01 – Gestionar los programas</b>	Antes del proceso técnico que implica un ETL, debió presentarse una fase de análisis de requerimientos y gerencia de proyectos. Este objetivo de gestión contiene lineamientos específicos para poder auditar el proceso previo a la parte técnica del ETL.
<b>BAI02 – Gestionar la definición de requisitos</b>	Contiene los aspectos necesarios para evaluar la definición del alcance del proceso ETL.
<b>BAI03 – Gestionar la identificación y construcción de soluciones</b>	Brinda las herramientas de evaluación necesarias para garantizar la efectividad de las soluciones tecnológicas.

Fuente: Elaboración propia, basado en el dominio *Construir, Adquirir e Implementar* (BAI) (ISACA, 2019, págs. 153-180)

Para concluir esta sección y recopilando los aspectos relevantes de este documento, se presenta en la tabla 9 una serie de buenas prácticas que un auditor de sistemas debe considerar para realizar su labor en la evaluación a los procesos ETL.

**Tabla 9** Buenas Prácticas de Auditoría sobre Procesos ETL

GUIA	EXPLICACIÓN
<b>Conocer sobre procesos ETL</b>	El auditor de sistemas debe tener conocimiento sobre los aspectos técnicos que involucra un servicio de integración de datos; así mismo debe dominar y entender los términos asociados a este tipo de procesos tecnológicos.
<b>Tener acceso a la documentación sobre alcance y funcionalidad del ETL que se pretende auditar</b>	Este acceso es necesario para que el auditor pueda tener de primera mano información sobre el proceso que se va a evaluar, su alcance, los requerimientos del cliente y todos los demás aspectos sobre lo que se espera del proceso ETL.
<b>Tener acceso a la documentación relacionada con la gestión de riesgos del proceso ETL que se pretende auditar</b>	La gestión de riesgos se convierte en fuente de información para que el auditor pueda identificar aquellos eventos de riesgo a los que se debe realizar seguimiento y evaluación. Si no existe esta documentación sobre la gestión de riesgo, se sugiere que el auditor la realice como parte de su labor.
<b>Identificar los controles que fueron implementados en el proceso ETL</b>	Ya sea mediante la documentación que el diseñador tenga disponible o identificarlos directamente sobre el proceso (código fuente o herramienta ETL empleada). Lo importante es que el auditor pueda tener acceso a dichos controles para realizar las evaluaciones que considere pertinentes con las técnicas y herramientas respectivas.
<b>Habilidad para detectar falencias en el proceso de integración de datos</b>	El auditor debe poseer perspicacia para poder identificar sobre el código fuente o sobre la herramienta ETL aquellas falencias que puedan incurrir en materialización de riesgos y por ende en generación de impactos negativos sobre el proceso.
<b>Basar su labor en los marcos y normas de conocimiento universal</b>	Se sugiere emplear en la auditoría, los lineamientos que ofrece el marco de referencia COBIT. Sin embargo, otras normas y métodos son válidos conforme sean necesarios para la evaluación de los procesos y controles empleados en el diseño de ETLs.

Fuente: elaboración propia

De esta forma, el auditor se encuentra preparado para asumir una auditoría a procesos ETL, considerando el conocimiento del proceso, empleando la metodología acorde con la evaluación, detectando las falencias en los controles y generando el diagnóstico y sugerencias para propiciar el mejoramiento continuo de estos procesos.

## Conclusiones

La ejecución de una auditoría de sistemas a servicios de integración de datos mediante procesos ETL (Extract, Transform and Load) implica tres aspectos fundamentales: primero, conocer las características de uso, diseño y automatización de este tipo de procesos tecnológicos para que el auditor tenga los suficientes criterios técnicos y logre emitir el diagnóstico correcto; segundo, considerar los escenarios de riesgo y los controles implementados por el dueño del proceso ETL, para que el auditor pueda emplear herramientas de evaluación sobre dichos controles y determinar su efectividad; tercero, aplicar un conjunto de buenas prácticas de auditoría relacionadas con la evaluación a los servicios de integración de datos.

Utilizando la Norma Técnica Colombiana NTC-ISO 31000 se identifican los escenarios de riesgo que pueden impactar negativamente el proceso ETL y la información que allí se utiliza. Estos escenarios son: las fuentes de información, los datos contenidos en esas fuentes, la relación entre los datos, los cálculos, filtros y agrupaciones que se realicen sobre los datos, bodega de almacenamiento y la seguridad de la información.

Para un adecuado control preventivo es recomendable que el proceso ETL incluya como parte de sus funciones la unificación de los formatos de fecha y numéricos, considerar las longitudes máximas de las cadenas de texto y, sobre todo, la realización de validaciones a los datos, además de otros controles que, de acuerdo con el requerimiento del ETL, deban implementarse para garantizar su correcto funcionamiento.

## Referencias

- Audisis. (s.f.). *Audisis*. Obtenido de IDEA:  
[https://www.softwareaudisis.com/Documentos/Folleto\\_IDEA.pdf](https://www.softwareaudisis.com/Documentos/Folleto_IDEA.pdf)
- Duque Méndez, N. D., Hernández Leal, E. J., Pérez Zapata, A. M., Arroyabe Tabares, A. F., & Espinosa, D. A. (2016). Modelo para el proceso de extracción, transformación y carga en bodega de datos. Una aplicación con datos ambientales. *CIENCIA E INGENIERÍA NEOGRANADINA*, 95-109.
- Gomez, L. F. (2007). Interoperabilidad en los Sistemas de Información Documental (SID): la información debe fluir. *Revista Códice Vol. 3 N° 1*, 23-39.
- IBM. (s.f.). *Cognos Analytics*. Obtenido de IBM Cognos Analytics:  
<https://www.ibm.com/co-es/products/cognos-analytics>
- ICONTEC. (2011). *Norma Técnica Colombiana NTC-ISO 31000. Gestión del riesgo. Principios y directrices*. Bogotá: Instituto Colombiano de Normas Técnicas y Certificación (ICONTEC).
- ISACA. (2019). *COBIT 2019. Marco de referencia. Objetivos de gobierno y gestión*. Schaumburg, IL, USA: ISACA.
- ISOTools. (s.f.). *La norma ISO 27001. Aspectos clave de su diseño e implementación*. [www.isotools.org](http://www.isotools.org). Obtenido de <https://www.isotools.org/pdfs-pro/iso-27001-sistema-gestion-seguridad-informacion.pdf>
- Microsoft. (7 de Junio de 2018). *SQL Server Integration Services*. Obtenido de <https://docs.microsoft.com/en-us/sql/integration-services/sql-server-integration-services?view=sql-server-ver15>
- Oracle. (Mayo de 2010). *Oracle Warehouse Builder*. Obtenido de <https://www.oracle.com/technetwork/developer-tools/warehouse/owb-11gr2-new-features-summary-129693.pdf>
- Perez Lopez, C., & Santín González, D. (2007). *Minería de datos. Técnicas y herramientas*. Madrid España: Thomson Ediciones.
- Tamayo Alzate, A. (2001). *Auditoría de sistemas. Una visión práctica*. Manizales: Universidad Nacional de Colombia.