AUDITORÍA INTERNA

LA FÁBRICA DE PENSAMIENTO
INSTITUTO DE AUDITORES INTERNOS DE ESPAÑA

Auditoría Interna del Gobierno del Dato

The Institute of Internal Auditors

Instituto de Auditores Internos de España

El INSTITUTO DE AUDITORES INTERNOS DE ESPAÑA es una asociación profesional fundada en 1983, cuya misión es contribuir al éxito de las organizaciones impulsando la Auditoría Interna como función clave del buen gobierno. En España cuenta con más de 3.200 socios, auditores internos en las principales empresas e instituciones de todos los sectores económicos del país.

LA FÁBRICA DE PENSAMIENTO es el laboratorio de ideas del Instituto de Auditores Internos de España sobre gobierno corporativo, gestión de riesgos y Auditoría Interna, donde participan más de 150 socios y profesionales técnicos expertos.

AUDITORÍA INTERNA

BUENAS PRÁCTICAS EN GESTIÓN DE RIESGOS

OBSERVATORIO SECTORIAL

PRÁCTICAS DE BUEN GOBIERNO

El laboratorio trabaja con un enfoque práctico en la producción de documentos de buenas prácticas que contribuyan a la mejora del buen gobierno y de los sistemas de gestión de riesgos en organizaciones de habla hispana. Además de desarrollar contenido, fomenta el intercambio de conocimientos entre los socios.

ENCUENTRA TODOS LOS DOCUMENTOS DE LA FÁBRICA EN www.auditoresinternos.es

AUDITORÍA INTERNA

LA FÁBRICA DE PENSAMIENTO
INSTITUTO DE AUDITORES INTERNOS DE ESPAÑA

Auditoría Interna del Gobierno del Dato

Febrero 2020

MIEMBROS DE LA COMISIÓN TÉCNICA

COORDINACIÓN:

Eduardo Villalobos Fernández, CIA, CISA. GRUPO COOPERATIVO CAJAMAR.

María Aísa Sánchez-Horneros, COSO. TELEFÓNICA.

Isabel Arias Pozas, CISA. GRUPO BBVA.

Cristina Bausá Rosa, CIA, CISA, CRMA, CRISC, CISM, CGEIT. SAREB.

Ana Cendón. PwC.

Silvia Díaz Rodríguez, CISA, CISM. DELOITTE.

Fe Fernández Martín, GREA. PELAYO MUTUA DE SEGUROS.

Laura Lama Outeiriño. ALPHABET ESPAÑA - BMW GROUP.

Manuel Mendiola Antona, CIA, CISA, CRMA, CRISC, CISM, CGEIT. PKF ATTEST.

José Luis Picazo Martínez, CISA, CISSP, CEH. ING DIRECT.

Daniel Ponz Lillo, CIA, CISA, CRMA, CFE, CISM. IBERDROLA.

José Ignacio Sánchez Libreros. LIBERBANK

Jorge Sánchez López. MAPFRE, S.A.

Juan Santonja Lillo, CISA. SABIS - BANC SABADELL.

José Manuel Vidal Formoso, CISA, CRIS, CGEIT. GRUPO BBVA.

The Institute of Internal Auditors

Instituto de Auditores Internos de España

Contar con un buen gobierno del dato es una cuestión crítica para las organizaciones. El primer paso para realizar un gobierno del dato adecuado es definir y aprobar una política por parte del Consejo de Administración. El segundo, implica desarrollar e implantar procedimientos operativos para identificar y clasificar los datos, así como para medir su calidad y trazabilidad, entre otros aspectos. También habrá que adaptar procedimientos ya implantados para mejorar la seguridad y protección del dato. Auditoría Interna debe tener una implicación directa, proactiva y preventiva.

Este documento analiza al detalle los problemas a los que se enfrentan las compañías a la hora de tratar los datos con la tecnología y métodos tradicionales y facilita pautas sobre cómo se deben abordar estos problemas desde Auditoría Interna.

Desde una perspectiva teórico-práctica, este documento identifica y analiza las mejores prácticas en gobierno del dato y los nuevos roles que están surgiendo en las compañías para mejorarlo. También se explica el concepto de trazabilidad del dato y las diferentes fases de su ciclo de vida. Se apuntan, además, algunas reflexiones sobre la seguridad del dato y la infraestructura tecnológica necesaria para que las organizaciones puedan gobernar los datos adecuadamente. Y finalmente, se proponen métodos prácticos para medir un aspecto muy relevante: la calidad de los datos.

Confiamos en que sea un documento útil para la profesión y agradecemos a los miembros de la Comisión Técnica su colaboración para llevarlo a cabo.

Instituto de Auditores Internos de España

Índice

Resumen ejecutivo

Este documento aborda la problemática del gobierno del dato, especialmente desde la perspectiva de Auditoría Interna. No obstante, hay que destacar que el gobierno del dato es un aspecto de una enorme complejidad y extensión. Se analizará un conjunto limitado de cuestiones relacionadas con el gobierno del dato, ya que no se puede abarcar en toda su extensión sin pecar de falta de rigor. Así pues, es necesario comenzar con un *disclaimer* sobre algunas cuestiones:

- Este documento no versa sobre Big Data, un conjunto de datos que no se pueden "explotar" mediante técnicas tradicionales de tratamiento de datos. Normalmente, la dificultad en la explotación del Big Data se debe a que se maneja un gran volumen de información o porque los datos a tratar tienen una estructura compleja.
- Tampoco se plantean las implicaciones del nuevo Reglamento General de Protección de Datos (RGPD[1]) y cómo se les debe tratar para cumplir con el mismo, ya que esta cuestión es tan amplia y compleja que merece un documento aparte.
- La seguridad no es el tema central de este documento. Es evidente que es una parte relevante del gobierno del dato, pero ya existen múltiples fuentes, documentos y *know-how* de cómo abordar los temas de seguridad.

¿Qué se analiza entonces en este documento? Precisamente, los problemas a los que se enfrentan las organizaciones a la hora de tratar los datos con la tecnología y métodos tradicionales, y de cómo se deben abordar estos problemas desde Auditoría Interna. Se desarrollan los siguientes aspectos:

- Se identifican las mejores prácticas en relación con el gobierno del dato y cuáles son los nuevos roles que están surgiendo en las compañías para mejorarlo.
- Se analiza cuál puede ser el ciclo de vida del dato.
- Se proponen métodos prácticos para medir la calidad de los datos.
- Se estudia un concepto que genera cierta confusión: la trazabilidad del dato.
- Se analiza la tendencia que hay en algunas organizaciones de incluir "los informes" como un elemento más dentro de los procesos de calidad y gobierno del dato (especialmente, los que se remiten a los órganos de gobierno de las sociedades).

1. Los requisitos que establece el RGPD han sido recogidos en el ordenamiento español en la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y Garantía de los Derechos Digitales.

- Por último, se incluyen algunas pinceladas relacionadas con la seguridad del dato y con la infraestructura tecnológica necesaria para que las compañías puedan gobernar los datos adecuadamente.
- Se analiza el *Estado del arte.* Es decir, las normativas o metodologías que tratan sobre la calidad y el gobierno del dato.

LA NECESIDAD DE GOBERNAR EL DATO Y CÓMO SE ABORDA EN LAS ORGANIZACIONES

Las organizaciones cada vez gestionan mayor volumen de datos y, en consecuencia, cada vez hay más dificultades para realizar una gestión adecuada de estos datos. Por tanto, y derivado de todas las actividades que están apareciendo para gobernar el dato, ha surgido una **nueva disciplina** relacionada con la gestión de las TI: el gobierno del dato.

El primer paso para realizar un buen gobierno del dato es definir y aprobar una política por parte del Consejo de Administración. El segundo paso consiste en desarrollar e implantar procedimientos operativos para implantar dicha política.

Pero no todos los procedimientos a desarrollar serán nuevos, por lo que habrá que modificar algunos de los ya existentes. Por ejemplo, si las compañías ya han aprobado sus políticas de seguridad de la información y han desarrollado procedimientos sobre esta materia, es posible que sea necesario modificarlos dado que el gobierno del dato también incorpora un componente de seguridad del dato.

Los procedimientos más relevantes que una organización tendría que desarrollar en esta materia son:

- Procedimiento para la identificación y clasificación del dato.
- Procedimiento para medir la calidad del dato.
- Procedimiento para garantizar la trazabilidad del dato.

Algunas compañías ya han comenzado a medir la calidad de sus datos. Por ejemplo, son capaces de presentar un cuadro de mando al Consejo de Administración que indique el nivel de calidad de cada tipología de datos[2].

Si bien esas mediciones son una buena práctica, implican un nuevo riesgo: ofrecer una falsa sensación de seguridad. Este riesgo aparece desde el momento en que las mediciones de calidad no dejan de ser un mero artificio para tratar de medir la calidad real de los datos. Pero, como se verá, esa calidad se está midiendo con respecto a una serie de controles establecidos por el responsable del dato. Y, de partida, estas mediciones tienen una serie de limitaciones.

El primer paso para realizar un buen gobierno del dato es definir y aprobar una política por parte del Consejo de Administración.

2. En estos casos, se suele presentar con un elevado nivel de agregación. Por ejemplo, en esos cuadros de mando se podrían observar "afirmaciones" (de forma gráfica) de este tipo: i) la calidad de los datos de clientes es del 98,75%; ii) la calidad de los datos de operaciones es del 99,3%, etc.

Un ejemplo real: si el dispositivo que se usa para medir no es bueno, ofrecerá lecturas imprecisas. De poco sirve un velocímetro que marque una velocidad de 49km/h si un vehículo está viajando, en realidad, a 81km/h. Y, en el mundo de la calidad del dato, construir un velocímetro entraña un elevado grado de dificultad.

IMPACTO EN EL TRABAJO DE AUDITORÍA INTERNA

También está surgiendo una nueva disciplina dentro de Auditoría Interna: la auditoría del gobierno del dato. En principio, es una disciplina que nace en el ámbito de la auditoría de riesgos tecnológicos y que demanda un perfil de auditores de TI especializados en el análisis y tratamiento de datos.

La auditoría del gobierno del dato como disciplina nace en el ámbito de la auditoría de riesgos tecnológicos y demanda perfiles TI especializados en analítica de datos.

Por tanto, es necesario adaptar el plan de Auditoría Interna para introducir trabajos de auditoría que analicen cómo se gobierna el dato en las organizaciones. En la práctica, esto se podría hacer de dos formas compatibles entre sí:

- Añadiendo nuevos trabajos de auditoría.
- Modificando el alcance de trabajos ya existentes (ya sea de otros trabajos de auditoría de TI o de procesos de negocio).

En el caso de la auditoría de procesos de negocio, se trataría de añadir –como objetivo del trabajo– la opinión sobre el gobierno de los datos que se utilizan en el proceso que se va a auditar. Hay que ser cuidadoso con este enfoque, ya que verificar este objetivo consume un elevado número de recursos. En este sentido, es importante valorar los riesgos del proceso de negocio y la importancia de los datos. Hay miles de datos distintos en las compañías y no es posible opinar sobre la calidad de todos ellos, ya que no se estarían empleando los recursos de Auditoría Interna de forma eficiente, al no centrarse en los datos con mayor nivel de criticidad.

No hay una respuesta única a cuántos trabajos habría que incluir en el plan de Auditoría Interna. Depende de las características de cada organización y de la estrategia que quiera seguir su Dirección de Auditoría Interna:

- En algunas organizaciones puede que solo se introduzca un único trabajo de auditoría, donde se ofrezca una opinión integral de todo el proceso de gobierno del dato.
- Otras unidades de Auditoría Interna preferirán incluir varios trabajos en el plan, pero con alcances más limitados y reducidos en cada uno de ellos.

En cualquier caso, las revisiones de Auditoría Interna deben incluir una serie de puntos críticos, independientemente de si se analizan en un único trabajo o en varios trabajos a lo largo de varios años:

- La existencia y suficiencia de una política de gobierno del dato.
- Opinar sobre la calidad de los datos. En aquellas compañías en que ya se esté midiendo su calidad, Auditoría Interna también debería opinar sobre la suficiencia y calidad de esas mediciones.
- La trazabilidad de los datos.

Modelo de gobierno del dato

En primer lugar, es imprescindible definir qué se entiende por gobierno del dato. Existen distintas definiciones según la fuente analizada. A grandes rasgos –y realizando una definición propia– se puede decir que es el conjunto de personas, procesos y tecnología que permiten a una compañía capturar, mantener y gestionar los datos de su actividad de forma fiable y eficiente, para asegurar la veracidad, integridad, confidencialidad y disponibilidad de los datos; y conseguir que éstos ayuden a la compañía a tomar las decisiones oportunas en el momento adecuado.

El gobierno del dato es el conjunto de personas, procesos y tecnologías que permiten capturar, mantener y gestionar los datos de forma fiable, preservando la integridad y confidencialidad.

ESTÁNDARES PARA EL GOBIERNO DEL DATO

¿Existen normativas o estándares que ayuden a gobernar los datos? En realidad, apenas hay referencias internacionales en esta materia. Solo se pueden citar dos fuentes de cierta relevancia:

- Los *Principios para una eficaz agregación de datos sobre riesgos y presentación de informes de riesgos* (comúnmente conocidos como RDA[3])
- El DAMA-DMBOK2 *framework*[4].

RDA

Para entender su significado, hay que hacer varias precisiones:

- RDA fue publicado en enero de 2013 por el Comité de Supervisión Bancaria de Basilea[5] del BIS[6]. El BIS es un organismo constituido por los bancos centrales de 60 países cuya misión es ayudarles en su búsqueda de la estabilidad monetaria y financiera, así como fomentar la cooperación internacional en esas áreas y actuar como un banco para los bancos centrales.
- Una de las principales actividades de este Comité es emitir documentos que sirvan como guía tanto a las entidades financieras como a los bancos centrales.
- Uno de esos documentos es el ya mencionado *Principios para una eficaz agregación de datos sobre riesgos y presentación de informes de riesgos (RDA)".*

3. Comité de Supervisión Bancaria de Basilea. *Principios para una eficaz agregación de datos sobre riesgos y presentación de informes de riesgos.* Enero de 2013. El acrónimo RDA deriva del nombre de la norma en inglés: *Principles for effective Risk Data Aggregation and risk reporting.*

4. DAMA International. *The DAMA Guide to the Data Management Body of Knowledge* 2nd editon. Junio 2017. Lugar de publicación: DAMA International https://dama.org.

5. Comité de Supervisión Bancaria de Basilea: organización mundial que reúne a las autoridades de supervisión bancaria, cuya función es fortalecer la solidez de los sistemas financieros.

6. Acrónimo de *Bank for International Settlementes* (en castellano: *Banco de Pagos Internacionales).*

- Aunque en cuanto a su "forma" no es una norma, en el "fondo" –y mediante una serie de complejos mecanismos normativos– se ha convertido *de facto* en una norma internacional para bancos: los bancos sistémicos deben cumplir ya con RDA y se espera que los no sistémicos también cumplan en el futuro.

¿EN QUÉ CONSISTE?

El nacimiento de RDA está perfectamente razonado en el punto 1 de su introducción, que se reproduce a continuación por su interés:

> *Una de las principales lecciones de la crisis financiera mundial iniciada en 2008 fue que la inadecuación de las tecnologías de la información (TI) y las arquitecturas de datos de los bancos impidió realizar una gestión integral de los riesgos financieros. Muchos bancos fueron incapaces de agregar sus exposiciones al riesgo e identificar con prontitud y precisión sus concentraciones a nivel de grupo bancario, así como entre líneas de negocio y entre personas jurídicas. En algunos bancos, la incapacidad para gestionar adecuadamente los riesgos respondía a carencias en la agregación de datos sobre riesgos y en las prácticas de presentación de los correspondientes informes. Esto tuvo consecuencias graves para los propios bancos y para la estabilidad del sistema financiero en su conjunto*[7].

Por tanto, hay que tener en cuenta que dicha norma está muy orientada hacia los bancos, y a mejorar la capacidad que tienen de generar información correcta y relevante sobre sus principales riesgos.

RDA está construido en base a **14 principios:**

- **11 son los que deben cumplir los bancos:**
  - Gobernanza.
  - Arquitectura de Datos e Infraestructura TI.
  - Exactitud e Integridad.
  - Completitud.
  - Prontitud.
  - Adaptabilidad.
  - Exactitud.
  - Exhaustividad.
  - Claridad y utilidad.
  - Frecuencia.
  - Distribución.
- **3 principios que están dirigidos a los organismos supervisores:**
  - Examen.
  - Acciones correctivas y medidas de supervisión.
  - Cooperación origen/acogida.

DMBoK

Es un marco para la gestión de los datos publicado por DAMA *(Data Management Association International),* una asociación sin ánimo de lucro e independiente de los proveedores, que está compuesta por profesionales independientes especialistas en gestión de datos, y cuya misión principal es proporcionar ayuda y asistencia a los profesionales de los datos.

El marco publicado por DAMA es de pago y está organizado en **11 áreas de conocimiento:**

- Gobierno del dato.
- Arquitectura de datos.
- Modelado y diseño de datos.
- Almacenamiento y operación de los datos.
- Seguridad de los datos.
- Integración e interoperabilidad del dato.
- Gestión documental y de contenidos.
- Datos maestros y de referencia.
- *Data Warehousing* y *Business Intelligence.*
- Metadatos.
- Calidad de datos.

7. Comité de Supervisión Bancaria de Basilea. *Principios para una eficaz agregación de datos sobre riesgos y presentación de informes de riesgos.* Página 8.

MEJORES PRÁCTICAS EN MATERIA DE GOBIERNO DEL DATO

Lograr una buena gobernabilidad y gestión de los datos implica a diversas áreas, funciones y recursos dentro de una compañía, desde su máximo responsable (el Consejo de Administración) hasta el personal que actúa en su nombre; y desde el diseño de arquitectura de sistemas al minucioso control del registro o transformación de los datos, considerando posibilidades tanto de error como de omisión.

Es necesario un documento que regule las responsabilidades de ejecución, control y supervisión en todas las fases del proceso de gestión del dato, y en todos los niveles de la compañía, desde su captura hasta su apropiada utilización para cualquiera de los fines autorizados.

Por tanto, **la mejor práctica comienza con la aprobación por el Consejo de Administración de una política de gobierno y calidad del dato que regule las responsabilidades en cada proceso que implique gestión de datos y que otorgue facultades en la toma de decisiones.**

Para desarrollar la política, la principal premisa es que se debe cuidar su redacción y evitar la inclusión de demasiados detalles operativos. Sería ineficaz que –para implantar un cambio menor en cualquier procedimiento– fuera necesario modificar la política y, por tanto, ser aprobada de nuevo por el Consejo de Administración.

Una política de gobierno y calidad del dato debería incluir:

- Definición de lo que se entiende por dato.
- Definición de las funciones y responsabilidades en el proceso. En esta parte de la política se asignan responsabilidades concretas dentro de la compañía. Más adelante se mencionan las principales funciones que están surgiendo en relación al gobierno del dato.
- Enumeración de los principales textos normativos[8] que deben existir en una compañía. En este punto, el Consejo dará ciertas directrices a la dirección de la entidad, pero sin entrar en cuestiones eminentemente operativas. Por tanto, parece razonable que el Consejo exija que la Alta Dirección elabore, al menos, los siguientes procedimientos:
  - Para la identificación y clasificación del dato. Tiene una importante relación con los demás, ya que –por lo general– a los datos de mayor criticidad se les deben aplicar controles más intensos. Es decir, sería aconsejable que en los distintos procedimientos que se desarrollen se establezcan diferencias en función de la criticidad de los datos.
  - Para medir la calidad del dato.
  - Para garantizar la trazabilidad del dato.

Adicionalmente, el Consejo de Administración podría exigir a la Dirección la creación de otros procedimientos, en función del tamaño, estructura y complejidad de la organización. Algunos procedimientos son:

- Certificación de un conjunto de datos mediante un mecanismo que garantice una re-

Lograr un buen gobierno del dato implica a diversas áreas, funciones y recursos; desde el diseño de arquitectura de sistemas al registro y transformación de los datos.

8. Cada organización puede tener su propia forma de desarrollar sus políticas en diferentes textos normativos. No obstante, con objeto de facilitar la lectura, en adelante, se va a utilizar el término "procedimiento" para hacer referencia a estos textos normativos. De hecho, incluso, en algunas organizaciones, la palabra "política" tampoco sería el término exacto que se utiliza para definir los documentos de primer rango o nivel que aprueban las compañías.

visión exhaustiva de la calidad de un conjunto de datos.

- Certificar la calidad de un informe con la existencia de un mecanismo que la garantice.
- Garantizar la calidad de las transformaciones y cálculos que se realicen sobre los datos. Este punto tiene una fuerte relación con las metodologías de desarrollo de *software*, ya que se trata de garantizar que se realizan las pruebas adecuadas para asegurar que no se producen errores en los tratamientos de datos.
- Crear y mantener un diccionario único de datos[9].
- Gestionar el acceso a los datos de forma segura.
- Disociar los datos bajo determinadas circunstancias. Es decir, realizar una copia de una base de datos para que no sea posible identificar a qué persona física pertenece cada dato.

Es necesario determinar las responsabilidades de ejecución, control y supervisión en todas las fases del proceso de gestión del dato y en todos los niveles de la compañía.

- Almacenar, transmitir y eliminar los datos de manera segura.
- Adquirir datos de un proveedor externo y/o compartir/ceder datos con un tercero.

Dos consejos a la hora de elaborar una política de esta naturaleza:

- La política debe servir para trasladar directrices (no es un lugar para definir procedimientos), por lo que se recomienda evitar detalles excesivamente operativos.
- Si es la primera vez que la compañía va a crear una política de gobierno del dato, hay que evitar definir una política demasiado ambiciosa al principio por si no se consigue implantarla en un tiempo razonable al ser demasiado exigente. Hay que tener en cuenta que es costoso implantar la política y todos los procedimientos asociados.

  Una vez implantada la política aprobada por el Consejo, se podrá ir modificando para hacerla más exigente.

RELACIÓN ENTRE EL GOBIERNO Y CALIDAD DEL DATO Y OTRAS DISCIPLINAS DE SISTEMAS

El "gobierno y calidad del dato" se configura como una nueva disciplina en el mundo de la gestión de la tecnología.

Si se hiciera una división (muy simplista) de la gestión de sistemas, se podrían configurar tres disciplinas, que tienen interrelaciones entre sí.

9. El diccionario de datos es, en esencia, un repositorio donde se describe lo que significa cada dato. Aunque no existe un único criterio para documentar este significado, sí existe un consenso de que el diccionario de datos debe de ser entendible por el usuario final. Es decir, si un usuario quiere conocer qué significa un campo de una tabla, debería poder buscar su definición en el diccionario y entenderla.

Con la aparición del "gobierno y calidad del dato", se podría visualizar una cuarta disciplina, que presenta interacciones con las tres anteriores.

Esta interrelación provoca que no sea sencillo elaborar una política sobre gobierno y calidad del dato, ni desarrollar sus procedimientos asociados. Por ejemplo, esta política podría entrar en conflicto con la "política de seguridad de la información", existente en la mayoría de las organizaciones.

Si examinamos los procedimientos que se han recomendado en el punto anterior, se observa que existen varios relacionados con la disciplina de la seguridad. Y en otra tabla se muestra la relación entre la perspectiva RDA y los procedimientos.

PROCEDIMIENTOS RELACIONADOS CON LA DISCIPLINA DE LA SEGURIDAD
Gestionar el acceso a los datos de forma segura.
Disociar los datos bajo determinadas circunstancias.
Almacenar y transmitir los datos de manera segura.
Eliminar los datos de manera segura.
Adquirir datos de un proveedor externo y/o compartir/ceder datos con un tercero.

PROCEDIMIENTOS RELACIONADOS CON LA DISCIPLINA DE GOBIERNO Y CALIDAD DEL DATO
Procedimiento para la identificación y clasificación del dato.
Procedimiento para medir la calidad del dato.
Procedimiento para garantizar la trazabilidad del dato[10].
Garantizar la calidad de las transformaciones y cálculos que se realicen sobre los datos.
Certificación de un conjunto de datos.
Crear y mantener un diccionario único de datos.
Certificar la calidad de un informe.

NUEVOS ACTORES EN EL GOBIERNO DEL DATO

A medida que va creciendo la preocupación y la exigencia por un adecuado gobierno del dato, surge una serie de figuras organizativas que ayudan a mantener ese adecuado buen gobierno. Destacan:

Chief Data Officer (CDO)

Es el máximo responsable del gobierno del dato en las compañías.

Este rol empieza a surgir a medida que el tamaño de los datos va creciendo exponencialmente, a la vez que su complejidad. Se tiene constancia de que la primera CDO fue Cathryne Clay Doss, de la empresa Capital One, en 2002.

Existen muchas razones por las que este rol es necesario (especialmente, en compañías

10. Aunque también se podría considerar que este procedimiento tendría más relación con la disciplina de desarrollo de software.

El Chief Data Officer (CDO) es el máximo responsable del gobierno del dato en las compañías.

grandes). Por ejemplo, CAP GEMINI[11] menciona hasta 18 motivos por los que resulta aconsejable establecer este rol. Algunos son:

- Las operaciones suelen ser complejas, dispares y, a menudo, ineficientes en sus enfoques para la gestión de la información. La mayoría de las grandes compañías de servicios financieros de hoy son multinacionales. Sus operaciones abarcan varios países. Esto implica la intervención de personal con diferentes idiomas. Y, con frecuencia, son operados y gobernados de forma local y descentralizada.
- La información crítica suele estar en silos. Y eso perjudica los informes a nivel empresarial, la toma de decisiones y la optimización del rendimiento.
- Ciertas funciones comerciales requieren la información agregada, pero no está disponible fácilmente.
- Los departamentos de negocio y de sistemas no hablan el mismo idioma, ni tienen un entendimiento común sobre la gestión de la información. Esto provoca una brecha de conocimiento considerable con respecto a los elementos de datos críticos para la empresa.

Entre las principales funciones que suele tener el CDO conviene resaltar:

- Asegurar la implantación de las políticas aprobadas por el Consejo de Administración.
- Coordinar la actuación del resto de figuras que intervienen en la gestión de los datos.
- Racionalizar y estandarizar el uso de herramientas y mejoras prácticas para la gestión de los datos.
- Realizar un reporte periódico al Consejo de Administración sobre el gobierno y calidad del dato.

Responsable funcional del dato *(Data owner)*

En general, es una persona del negocio encargada de definir conceptualmente los datos dentro de su ámbito de responsabilidad. Sus principales funciones suelen ser:

- Identificar cuáles son los datos necesarios para la gestión del negocio, definiéndolos de forma clara en el diccionario único de datos.
- Establecer la calidad mínima necesaria que deban tener los datos.
- La toma de decisiones en lo referente a la forma de capturar los datos o de corregirlos en caso de incidencias.

Lo normal es que en una compañía existan varios *data owners,* normalmente en función de las diferentes áreas de negocio.

Responsable técnico *(Data architect)*

Diseña y/o mantiene los sistemas de información que dan soporte a la operativa, a los procesos o al reporting de la entidad.

Al igual que con los *data owners,* lo normal es que existan varios *data architect,* que ten-

11. CAP GEMINI. *The Role of the Chief Data Officer in Financial Services.* Junio de 2012. https://www.capgemini.com/gb-en/resources/the-role-of-the-chief-data-officer-in-financial-services/

drán una estrecha relación con los *data owners* e irán tomando con ellos decisiones consensuadas para mejorar la gestión de los datos.

Usuario de los datos *(Data user)*

Tendrá sus responsabilidades y obligaciones en el proceso.

Responsable de informes *(Report owner)*

Figura que está surgiendo en algunas organizaciones para:

- Garantizar la calidad de los informes.
- Decidir los criterios que se deben utilizar a la hora de elaborar un informe.
- Controlar el acceso y distribución de los informes.

EL ROL DE AUDITORÍA INTERNA

En relación con el gobierno del dato, es importante que exista en el plan de Auditoría Interna un trabajo para analizar el proceso de gobierno del dato. ¿Cuál debería ser el contenido y alcance de este trabajo de auditoría? y ¿cuáles son las verificaciones más comunes que se deberían llevar a cabo?

Antes, hay que enumerar algunas cuestiones que podrían formar parte de la revisión de Auditoría Interna (que cobrarán más sentido en los siguientes capítulos):

1. Analizar la suficiencia de la política aprobada por el Consejo de Administración.
2. Asegurar que ha sido comunicada adecuadamente dentro de la organización.
3. Verificar la existencia de los órganos de gobierno de la calidad del dato.
4. Verificar que se han desplegado los roles y funciones en la compañía.
5. Analizar los umbrales de calidad, cómo se vigilan y el reporte que se realiza.
6. Analizar los procesos relacionados con el diccionario de datos y la trazabilidad (funcional y técnica).
7. Verificar la existencia de planes de acción para remediar las incidencias de calidad del dato.
8. Analizar los procesos relacionados con la seguridad del dato.

Los cuatro primeros aspectos para revisar no conllevan una excesiva complejidad. Se pueden realizar con un único trabajo de Auditoría Interna (un único entregable), sin consumir demasiados recursos y en un tiempo de ejecución corto; las pruebas necesarias para verificar dichos puntos tampoco son excesivamente complejas.

En cambio, a partir del quinto punto, la verificación se vuelve más compleja y consumidora de recursos. Hay que tener en cuenta que cualquier compañía puede superar con facilidad el millar de tipologías de datos distintos. ¿Cómo verificar el establecimiento de umbrales sin entrar en las peculiaridades de cada tipo de dato? ¿Cómo verificar la trazabilidad sin entrar en cada uno de los datos que necesitan ser trazados? ¿Cómo compaginar la disciplina de seguridad con la del gobierno del dato? Y, aunque se pueda (y se deba) trabajar

Es clave que en Auditoría Interna exista un plan de trabajo para analizar el proceso de gobierno del dato.

con muestras, ¿cómo elegir esas muestras para poder llegar a conclusiones válidas para toda la compañía? Por ejemplo, suele ser normal que las políticas se hayan implantado con mayor intensidad en ciertos conjuntos de datos más críticos y que, en cambio, en los conjuntos de datos de menor criticidad apenas se hayan implantado (o sean más relajadas).

Auditoría Interna necesita contar con varios perfiles profesionales y colaborar con otras áreas para velar por el gobierno del dato.

Por tanto, parece razonable que la verificación de los puntos más complejos (del quinto en adelante) se realice a través de otros trabajos de auditoría independientes, y se limiten a la revisión de conjuntos acotados de datos (en lugar de a todos los datos de la compañía).

Lo que se plantea como una posible estrategia es incluir en el plan de Auditoría Interna:

- Un trabajo para verificar los puntos más "troncales" del gobierno del dato (por ejemplo, las cuatro primeras tareas expuestas al principio).
- Varios trabajos de auditoría para verificar el resto de puntos. Por ejemplo:
  - Dividiendo todos los datos de la organización en varios conjuntos o ejes[12]. Y poniendo el foco en aquellos ejes que tengan un mayor nivel de riesgo (desde el punto de vista del riesgo del proceso en que se usan esos datos).
  - Separando las "disciplinas" que se han mencionado en el apartado anterior *Relación entre el gobierno y calidad del dato y otras disciplinas del sistemas.* Es decir, se puede planificar una serie de trabajos para revisar los procedimientos relacionados con la seguridad, separando estos trabajos de los que están más relacionados con la disciplina del gobierno del dato (calidad, trazabilidad, etc).

No obstante, algunos departamentos de Auditoría Interna preferirán abordar el cumplimiento de la política con una menor cantidad de trabajos/informes, pero con un alcance mayor. Ambos enfoques son válidos y dependen de la estrategia que cada departamento de Auditoría Interna considere más eficaz en su compañía.

Para que los trabajos de Auditoría Interna sobre esta materia se realicen con solvencia y calidad, es necesario contar con varios perfiles de auditoría. En la mayoría de las compañías, el peso y el liderazgo de este tipo de auditorías recaerá dentro del área de Auditoría de TI. Se podrían distinguir varios perfiles para realizar estos trabajos:

- Ya se ha mencionado que el gobierno y la calidad del dato podría ser una nueva disciplina de la gestión de TI. También se podría decir que surge una nueva disciplina de auditoría de TI: Auditoría del gobierno y calidad del dato. Para auditar esta disciplina, es necesario un perfil de auditores de TI especializados en el análisis y tratamiento de datos.
- Para el éxito de estos trabajos de auditoría es fundamental que se establezcan colaboraciones robustas con el resto de las áreas de Auditoría Interna, ya que en ella reside un mayor conocimiento del negocio (y de la importancia y criticidad de los datos que maneja la organización). Por tanto, es im-

12. Se usa el término de ejes para hacer referencia a tipologías de datos de diferente naturaleza. Por ejemplo: los datos de clientes serían un eje y los datos de empleados podrían ser otro eje distinto (ya que sus atributos son diferentes, se capturan de forma distinta y, probablemente, tengan distinta criticidad).

prescindible que muchos de los trabajos de auditoría que se van a exponer en este documento se realicen de forma conjunta entre el área de Auditoría de TI y otras áreas de Auditoría Interna especializadas en los procesos de negocio cuyos datos se vayan a auditar. Habría que prescindir de estas colaboraciones en aquellos casos en los que el auditor de TI posea un gran conocimiento del negocio cuyo dato está auditando.

- Finalmente, como esta nueva disciplina tiene interrelación con otras disciplinas más tradicionales (la auditoría de sistemas, la auditoría de seguridad y la auditoría de desarrollo software) también sería necesaria la implicación de esos tres perfiles.

Gestión del dato

CICLO DE VIDA DEL DATO

La información es uno de los activos más importantes de las compañías, siendo el dato el elemento mínimo en la generación de conocimiento. Como cualquier bien de la empresa, la información tiene un ciclo desde su creación o captura y almacenamiento; pasando por su uso en distintas actividades de procesamiento; hasta que éstos se vuelven obsoletos y se eliminan.

La gestión del ciclo de vida de los datos de principio a fin es clave para que se extraiga la máxima utilidad de los mismos.

Las empresas que trabajan con datos como materia prima deberían adoptar el enfoque de "ciclo de vida del dato", es decir, identificar cómo se produce el flujo de información en sus procesos y actividades tanto internos como externos. Simplificando, el ciclo de vida de los datos se podría agrupar en tres grandes bloques.

Fases del ciclo de vida del dato

Cada uno de estos tres bloques se podría dividir en varias fases secuenciales[13].

13. Elaboración propia, a partir de SEARCHSECURITY. *Data-Lifecycle-Management-Model-Shows-Risks-and-Integrated-Data-Flow.* Web. Julio 2018

Obtención. Creación y captura de datos que no existían en la compañía. Hay diversas formas para adquirir datos, como la obtención de datos ya existentes que han sido creados por entes ajenos al negocio o la creación de datos por parte del factor humano y/o dispositivos del propio negocio.

Relevancia. Valoración del interés y pertinencia de la información para el negocio. Se debe establecer tanto la utilidad de la información, como los criterios de idoneidad de dicha información. Esta fase también se puede ubicar dentro del proceso de clasificación.

Clasificación. La clasificación de la información y su correcta identificación es uno de los pasos fundamentales del ciclo de vida de los datos. Cualquier tipo de información nueva para la empresa tendrá que ser clasificada antes de su incorporación a los flujos de datos de los sistemas.

Hay innumerable información sobre los diferentes criterios de clasificación. Por ejemplo, el estándar ISO 27001:2005 indica que la información debe ser clasificada en términos de su valor, requerimientos legales, confidencialidad y grado crítico para la compañía.

Otros criterios de clasificación, basados en riesgos e impactos asociados a la información pueden ser:

- Integridad. La información es actual, coherente, completa y solo se realizan cambios autorizados sobre ella.
- Disponibilidad. Solo los usuarios autorizados tienen acceso y pueden usar la información cuando sea necesario.
- Confidencialidad. A la información se accede solo por usuarios autorizados, entidades o procesos. El valor de la información para la empresa determina de forma directa los niveles de confidencialidad.

Almacenamiento. Los datos deben de ser almacenados en unos repositorios adecuados y seguros. Dependiendo del tipo de dato o información, la ubicación será distinta. Las bases de datos y sistemas de ficheros son los más frecuentemente utilizados. Esta característica es clave puesto que la organización, acceso y control de datos es indispensable para el correcto funcionamiento de las compañías.

Transmisión y transporte. Los datos se mueven según las necesidades de la empresa, tanto internamente como externamente. En general, existen dos grandes medios de transmisión y transporte: i) las redes de telecomunicaciones y ii) los dispositivos de almacenamiento portátiles (discos duros, *pendrives,* etc.).

En esta fase habría que prestar una especial atención a la transmisión del dato a otras compañías (ya sea como encargados del tratamiento, o como cesionarios del dato), tanto por los requerimientos legales como por el posible impacto sobre la privacidad de los usuarios

Tratamiento, conversión y alteración. Los escenarios en los que una compañía necesita tratar los datos son muy heterogéneos. Los tipos de tratamientos más usuales son:

- Tratamiento de los datos para elaborar información de gestión y contable.
- Combinación de los datos existentes entre sí para enriquecer los mismos y aportar mejor información a la compañía.

- Gestión operativa de los datos. Aquí se podría incluir desde el movimiento de los datos de un entorno a otro hasta la concesión o revocación de permisos a los usuarios.

Backup. Los sistemas de copia de seguridad son una pieza fundamental en los sistemas de gestión de la información, al ser elementos necesarios para garantizar la disponibilidad de los datos. Pero es necesario que los niveles de seguridad entre los datos en entornos productivos y sistemas de *backup* estén alineados.

Retención. Los datos pueden ser utilizados por la empresa durante el tiempo que legalmente este establecido para el tratamiento o esté autorizado por el propietario de la información. Se deben regular los criterios para retener los datos. Esta fase también se puede ubicar dentro del proceso de destrucción.

Destrucción. Particularmente importante con información sensible, pues será necesario el asegurar que la misma no recaerá en manos de terceras personas, así como garantizar que se cumplen los requerimientos legales relacionados con la retención.

El rol de Auditoría Interna

Auditar el ciclo de vida del dato es muy complejo. No tanto por las dificultades técnicas de las tareas de Auditoría Interna, sino por la enorme cantidad y heterogeneidad de datos que hay en una organización. Por ejemplo, no tienen nada que ver cómo se obtienen los datos de los clientes de una compañía a cómo se obtienen los datos de los empleados.

Como estrategia de Auditoría Interna, es muy probable que no tenga sentido auditar todas las fases descritas anteriormente en un solo trabajo de auditoría interna.

Lo primero que debería hacer Auditoría Interna es plantearse dos grandes preguntas acerca de **cómo abordar la auditoría del ciclo de vida del dato:**

- **¿Cómo voy a dividirlo?** En este capítulo, por ejemplo, se han realizado dos clasificaciones: una más general, en tre bloques (recolección, uso y conservación); y otra más detallada, en nueve fases (desde la obtención hasta la eliminación).
- Una vez realizada la división anterior, **¿cómo voy a auditar cada fase?** ¿Voy a hacer un único trabajo por cada fase o voy a hacer varios trabajos para auditar cada fase? ¿En cada trabajo voy a revisar todos los datos de la compañía o de cada fase voy a hacer varios trabajos desglosados por tipo de dato? ¿Voy a realizar una auditoría interna exclusivamente desde el ámbito de TI o es mejor integrar esta auditoría interna dentro de otros trabajos del proceso de negocio?

Auditar el ciclo de vida del dato es muy complejo por la enorme cantidad de datos y su heterogeneidad.

La respuesta depende de las características de cada organización. No obstante, sí que se pueden indicar dos factores que influyen a la hora de responder a estas preguntas:

- **Tamaño de la organización.** No parecería práctico que la auditoría interna de una gran compañía auditara todas las actividades del ciclo de vida (de todos los datos) en un único trabajo.
- **Tipo de actividad.** Existen determinadas fases del ciclo de vida que son más heterogéneas que otras. Y, por tanto, tiene un mayor sentido que la auditoría interna de esas fases se divida en varios trabajos distintos.

Tiene sentido que la auditoría interna de las fases del ciclo de vida del dato se divida en trabajos distintos.

Las fases de obtención y tratamiento suelen ser diferentes dependiendo del dato. Y, en cambio, el resto de las fases no suele presentar muchas diferencias en función del tipo de dato.

- Un par de ejemplos prácticos:
  - La fase de *backup* suele ser un proceso bastante estándar en las compañías que no depende demasiado del tipo de dato. Por lo que podría ser normal que este proceso se revisara con un único trabajo de auditoría interna, y con una versión más cercana a los procesos de TI.
  - La fase de obtención de datos es muy heterogénea. Es probable que la obtención de datos de clientes sea muy diferente a la obtención de datos de empleados; serán sistemas diferentes, con riesgos diferentes y controles diferentes. De ahí que probablemente tenga más sentido separar esos procesos en dos trabajos de auditoría interna distintos. Y tendría más sentido realizar un trabajo que se acerque más al proceso de negocio. Por ejemplo, se podría realizar un trabajo que se denomine "Auditoría del proceso de alta de los clientes" en el que, además de revisar los riesgos desde el punto de vista de gobierno del dato, se auditen los riesgos desde el punto de vista legal y de negocio.

Algunas tareas a realizar en las auditorías internas del ciclo de vida del dato.

Son muy diversas, y dependerán de las decisiones que haya tomado Auditoría Interna a raíz de las preguntas anteriores. Y también de los distintos riesgos que afecten a cada fase, al tipo de dato y a la compañía en concreto. Por ello, no es posible inventariar un conjunto de actividades que se deban realizar siempre en todas las auditorías internas de este tipo.

Algunas actividades a modo de ejemplo:

Recolección

- Grado de automatización de los procesos de carga y entrada de datos.
- Controles definidos sobre el proceso de entrada de información en las aplicaciones origen.
- Conciliaciones establecidas para validar la calidad de la información. Gestión de las discrepancias.
- Verificaciones existentes de la calidad de la información de entrada al proceso en cuanto a completitud, validez, consistencia, exactitud.
- En el caso de captura de datos de terceros, verificación de la existencia de un SLA con el proveedor que garantice la calidad y disponibilidad del dato.

Uso

- Identificación de los tratamientos realizados.
- Identificación de ajustes manuales realizados sobre los datos y recálculos.
- Conciliaciones de datos con otros repositorios de información. Gestión de las diferencias identificadas.
- Transformaciones realizadas sobre los datos y grado de dispersión de la información.
- Metodología de desarrollo y mantenimiento de las aplicaciones.

- Identificación de los responsables y usuarios de los datos.
- Identificación de los niveles de control sobre acceso al dato.
- Control sobre las cesiones de datos a terceros.

Eliminación

- Existencia de un proceso de identificación de datos que hayan dejado de ser útiles.
- Identificación de procedimientos de eliminación de la información.
- Identificación de los procedimientos de gestión de los sistemas de *backup* y restauración.

CALIDAD DEL DATO

Tal y como dijo el físico Sir William Thomson en el siglo XIX, *"Lo que no se define no se puede medir. Lo que no se mide, no se puede mejorar y lo que no se mejora, se degrada siempre"*.

Si se traslada este axioma al mundo de los datos, surge la siguiente pregunta: ¿es posible medir la calidad de los datos de una organización? La respuesta es que sí es posible. Y, además, dado que la fiabilidad de los datos es clave para el buen desarrollo del negocio, cada vez más organizaciones están llevando a cabo proyectos cuyo objetivo es medir la calidad de los datos que maneja. De esta forma, realizan un seguimiento periódico de la calidad de los datos, analizan su evolución y toman las medidas oportunas para remediar las deficiencias identificadas.

¿Cómo se puede medir la calidad de los datos y cuál es el papel de Auditoría Interna en este tipo de procesos? Antes de exponer estas cuestiones, hay que reflexionar sobre el gobierno del dato y preguntar por las causas que provocan que un informe utilizado para la toma de decisiones contenga errores, o que un fichero que una compañía utiliza para analizar un proceso de negocio contenga errores.

Simplificando, se pueden mencionar cuatro causas:

- **Mala calidad de los datos cuando se capturan** en los sistemas de las compañías. Una de las causas más frecuentes son los errores en la grabación de los datos por parte de los usuarios (cuando los datos se capturan de forma manual).
- **Errores en los desarrollos informáticos realizados por Tecnología** (cuando los informes o ficheros son elaborados por Tecnología).
- **Errores en los procedimientos que han aplicado los usuarios** (cuando los informes o ficheros son elaborados por ellos).
- **Errores colaterales causados por desarrollos informáticos.** Es un hecho que ocurre con cierta frecuencia en las compañías. Y consiste en que se descubre que hay un error en un informe o en un fichero, desconociéndose la causa del error, ya que los procesos de generación de esos informes o ficheros no se han modificado. Finalmente, se descubre que la causa está en un efecto no deseado de un desarrollo informático que, aunque no debería haber afectado a este informe (o fichero), sí tenía ciertas de-

Cada vez más compañías realizan un seguimiento periódico de la calidad y fiabilidad de sus datos.

pendencias que no se conocían. Suele estar causado por problemas en la trazabilidad.

¿Cómo se están afrontando en muchas compañías los procedimientos de medición de calidad del dato? Hay que tener en cuenta que la utilidad fundamental de estos procesos de medición es minimizar las causas de los errores expuestas anteriormente. Son especialmente útiles para mitigar la primera y cuarta causas y, en menor medida, para mitigar las causas segunda y tercera.

El objetivo de los procedimientos de medición de la calidad del dato es minimizar las causas de los errores.

Mejores prácticas para medir la calidad del dato

A continuación se proponen varios pasos teóricos a realizar en un proyecto para medir la calidad del dato y su aplicación, mediante un ejemplo concreto (se utilizará el mismo ejemplo en todo el capítulo).

Las fases que se siguen para la implantación de un proyecto para medir la calidad de los datos suelen ser:

- Definir el alcance (qué vamos a medir).
- Definir cómo se medirá la calidad.
- Realizar las mediciones.
- Analizar los resultados y establecer planes de remediación.

❶ DEFINIR EL ALCANCE

El primer paso es seleccionar un conjunto de datos acotado sobre el que se va a medir la calidad. Los proyectos de medición de la calidad son complejos, y el conjunto de datos de una compañía suele ser inmenso. Por ello, es importante acotar el alcance del proyecto para que sea manejable.

Seleccionar un conjunto de datos demasiado amplio suele ser una garantía para el fracaso en este tipo de proyectos. Es más recomendable ir poco a poco, pero consiguiendo resultados visibles. Y mejor, comenzar por aquellos datos que tengan una mayor criticidad para el negocio de la compañía.

Ejemplo de implantación

Una compañía comienza el proyecto de medición y, para acotar el alcance, toma dos decisiones:

1. Comenzar a medir la calidad de los datos personales de sus clientes.
2. De todo el conjunto de datos de los clientes, selecciona un subconjunto de los que considera más relevantes. Tras analizarlos, observa que está compuesto por 90 campos diferentes. Pero, tras un análisis experto entre el *data owner* y el *data architect,* decide que los realmente relevantes son 50 de esos 90 campos. Por tanto, acota el alcance del proyecto a esos 50 campos.

❷ DEFINIR CÓMO SE MEDIRÁ LA CALIDAD

Una vez elegidos qué datos concretos se van a medir, hay que decidir cómo se va a medir la calidad. Es decir, hay que definir los requisitos que debe cumplir la información incluida en el perímetro de actuación para poder certificarla en base a unos criterios homogéneos. Para ello será imprescindible la colaboración de los dueños de los datos, ya que son los que pueden determinar en base a qué requisitos considerar que un dato tiene la calidad suficiente como para ser explotado de forma fiable.

Para ayudar a los propietarios de los datos a realizar esta definición, hay que facilitarles los criterios más comunes para medir si un dato es de calidad.

CRITERIOS MÁS COMUNES Y LA TIPOLOGÍA DE CONTROLES TÍPICOS QUE SE PUEDEN REALIZAR PARA MEDIR EL DATO

Criterio	Descripción	Tipología de controles
Completitud	El dato debe estar informado para todo el perímetro aplicable.	Controles de blancos, de valores inexistentes y de valores por defecto.
Unicidad	El dato es único y no hay más de una forma de registrar la misma información.	Controles de duplicados.
Validez	El dato tiene un valor válido de acuerdo a los requisitos establecidos (formato, rango de valores, datos tipificados, etc).	Controles de formato, integridad referencial, contraste contra catálogo para campos tipificados y/o rango de valores de acuerdo a negocio.
Consistencia	El dato es coherente y consistente con la información almacenada en el mismo u otro sistema de información.	Controles de conciliación contable (saldo sistemas vs contabilidad), coherencia entre datos, validaciones de negocio, réplica de cálculos complejos y/o anclaje de información entre distintos sistemas de información.
Disponibilidad	El dato está disponible para su explotación en el momento requerido.	Controles de monitorización del proceso (rechazos, totales y tendencias, accesibilidad y disponibilidad, etc).
Exactitud	El dato es exacto y refleja fielmente la realidad.	Controles de revisión documental contra expedientes físicos o anclaje contra fuentes externas reputadas. Control estadístico de variaciones con respecto a periodos anteriores.
Efectividad	El dato debe ser relevante y pertinente para los procesos de negocio.	Grado de conformidad de los propietarios / usuarios de los datos con la información recibida, revisión periódica de las solicitudes de cambio y/o evolutivos referentes al perímetro de actuación considerado, etc.

Fuente: Elaboración propia

Como entregables de esta fase, se tendrían:

- Un conjunto de campos a medir.
- Un conjunto de controles a ejecutar (medir) por cada campo que se ha seleccionado.
- El umbral de tolerancia al error de ese control en ese dato. Para la definición de estos umbrales será necesaria la colaboración de los dueños de los datos, que poseen el juicio experto para determinar a partir de qué umbral se requerirá el reproceso o remediación del dato de forma previa a su explotación. En este tipo de proyectos novedosos también resulta razonable fijar este umbral en fases posteriores.

Para medir la calidad hay que definir los requisitos que debe cumplir la información incluida en el perímetro de actuación.

Ejemplo de implantación

Siguiendo con el ejemplo anterior, en la fase previa se han seleccionado 50 campos de clientes que se quieren medir. Ahora hay que tomar decisiones "campo a campo". Por ejemplo, vamos a suponer que uno de los 50 campos es el campo PROFESIÓN. En una aproximación simplicista[14], se puede decidir que se va a medir en base a dos criterios:

1. Completitud: Se medirá si hay algún cliente que tenga la PROFESIÓN sin rellenar (que esté vacío). Y se establecerá a partir de cuántos clientes con el campo sin informar supone un problema y se debe desencadenar alguna alerta o aviso. Es decir, el umbral de tolerancia al error. Por ejemplo, en este caso, el *data owner* ha decidido que el control será conforme si el porcentaje de clientes con el campo PROFESIÓN vacío está por debajo del 0,1% del total de clientes que se están midiendo.
2. Validez: Se medirá que todos los valores informados para cada cliente están dentro de una lista concreta de profesiones permitidas. Por ejemplo, contra una tabla referencial del sistema que ha definido el *data owner*. Y también habrá que definir un umbral.

En esta fase, los análisis que hay que realizar, así como las decisiones, se toman individualmente por cada uno de los atributos (o campos) que forman parte del cliente. De ahí que es importante limitar el alcance de estos proyectos, dado el elevado nivel de detalle al que hay que llegar.

❸ REALIZAR LA MEDICIÓN

Una vez definidos cuáles son los controles de calidad a realizar, es necesario comenzar a medir, ejecutando esos controles de calidad.

El ciclo típico de este proceso es:

1. **Se ejecutan uno a uno todos los controles de calidad definidos.**

   Para que el proceso sea eficiente, es aconsejable que estas ejecuciones se realicen con herramientas automáticas[15]. Hay aplicaciones de mercado que están pensadas para realizar este tipo de mediciones.

   Otra decisión a tomar sería la periodicidad con la que se van a tomar las mediciones. En condiciones normales, se realizarán mensualmente.

2. **Se muestra el resultado al usuario de forma amigable.** Se debería mostrar en un cuadro de mando que permita la visualización y navegación desde la visión general de calidad (informes ejecutivos) hasta el detalle de la incidencia de calidad identificada.

Ejemplo de implantación

A la hora de presentar los resultados, hay que ser capaces de mostrar lo siguiente:

1. En relación al campo PROFESIÓN:
   a. Número de clientes que tienen el campo vacío (y porcentaje sobre el total). Y, probablemente, un campo "semafórico" para indicar si se ha superado el umbral de tolerancia al error.
   b. Idem para el segundo control definido en este ejemplo (total de clientes que tienen una PROFESIÓN que no coincide con los valores de profesión permitidos).
2. Los usuarios de estas mediciones deberían poder ver con facilidad cuáles son los clientes

14. Por simplista debe entenderse que, a modo de ejemplo, solo se describen dos controles. En la práctica, lo normal es que se definan más controles por cada campo.

15. Aunque lo normal es que los controles de exactitud se tengan que ejecutar manualmente, ya que requerirán una revisión muestral contra una documentación física.

concretos que incumplen las mediciones, para que puedan empezar a analizar el problema.

3. Los usuarios deben poder ver cierta profundidad histórica de los análisis (recordemos que los controles se ejecutan periódicamente). De esta forma, podrán visualizar fácilmente si la calidad ha mejorado o ha empeorado con respecto a las mediciones anteriores.

4. El usuario puede ver "controles aislados" sobre un campo. Pero, ¿cuál es la calidad del campo PROFESIÓN? Una buena práctica sería realizar una agregación de todos los controles de cada campo para mostrar al usuario un nivel de calidad de ese campo. Para ello, es necesario recurrir al criterio experto.

   Supongamos, siguiendo este ejemplo, que las mediciones dan un 98% para el primer control y un 96% para el segundo. ¿Cuál es la calidad del campo PROFESIÓN? En este caso, el *data owner* decide que ambos controles son igual de importantes, por lo que se realiza una media aritmética. Por tanto, en el informe resultante se puede agregar, de forma que se indica que la calidad del campo PROFESIÓN es de un 97%.

5. Pero se puede seguir ampliando la pregunta: ¿Y cuál es la calidad global de los datos de clientes? ¿Es posible dar ese dato a un consejero de la compañía? De la misma forma que se ha "agregado" la calidad de un campo, se puede componer una fórmula agregada que, bajo criterio experto, ofrezca una visión de la calidad global. Tan sólo habría que hallar la media de la calidad de todos los campos y ponderar por la importancia del campo, según criterio experto.

   En este caso (y de cara a seguir con este ejemplo en los siguientes puntos de este documento), vamos a suponer que, en la última medición, el dato de CLIENTES tiene una calidad del 95%[16].

❹ ANALIZAR LOS RESULTADOS Y ESTABLECER PLANES DE REMEDIACIÓN

Una vez que se tiene el resultado de las mediciones, los responsables de los datos pueden comenzar a analizar los resultados y las desviaciones respecto a los umbrales de tolerancia establecidos. De tal forma que, si la medición no es satisfactoria, el *data owner* debe establecer los planes de remediación necesarios para mejorar la calidad del dato.

La repetición periódica de estas mediciones y análisis deben permitir que la compañía observe si las remediaciones aplicadas son realmente eficaces y si la calidad de los datos va mejorando con el tiempo.

En este tipo de proyectos suele ser frecuente que, en las primeras mediciones, se ajusten los controles y umbrales de calidad mediante juicio experto hasta que el resultado de la medición represente lo que realmente queremos medir desde el punto de vista funcional.

Ejemplo de implantación

Supongamos que se han empezado a realizar las mediciones del campo PROFESIÓN. Y que el responsable del dato observa que el sistema de calidad indica que hay un 20% de clientes con el campo vacío. Un escenario plausible y lógico que se podría desencadenar a raíz de ese análisis sería:

1. El propietario del dato comienza a analizar casos concretos de clientes, seleccionando una muestra.
2. Después de analizar diez clientes, observa que se trata de personas jurídicas. Y, obviamente, no tienen una profesión.

Tras realizar las mediciones, los responsables de los datos analizan los resultados y las desviaciones respecto a los umbrales de tolerancia.

16. Como se puede apreciar, una forma de presentar los resultados es mediante porcentajes, representando el 100% la máxima calidad posible.

3. Llega a la conclusión de que:
   a. Hay que modificar la medición para excluir a las personas jurídicas de ese control.
   b. Hay que incluir un nuevo control en las personas jurídicas que sea el equivalente a la profesión de las personas físicas. Por ejemplo, el CNAE.

Con el ejemplo anterior se observa que este enfoque se apoya en un sistema de mejora continua que presenta una peculiaridad: no es necesario que esté perfecto para ponerlo en producción. Incluso puede ser razonable renunciar a la perfección para poder obtener resultados visibles lo antes posible.

La "fuente de datos reputada" es la "tabla" concreta a la que los usuarios deben acudir para obtener un dato con el mayor nivel de fiabilidad.

La "fuente de datos reputada" o *Golden Source*

En entornos de datos complejos, los datos suelen estar repetidos en diferentes sistemas de información. Esto ocurre, en ocasiones, por necesidades técnicas y, en otras, por ineficiencias heredadas de los sistemas. Así, podemos tener un mismo dato en el entorno transaccional, en el entorno informacional e, incluso, en el entorno final del usuario.

Por tanto, ¿cómo se mide la calidad de los datos en todos estos sistemas?, ¿se deben repetir las mediciones en todas las "tablas"[17] dónde se repite un dato?, ¿existen formas más eficientes de medir la calidad del dato?

En este contexto ha surgido un nuevo concepto que se denomina "fuente de datos reputada" o *Golden Source*. Se puede decir que es la "tabla" concreta a la que los usuarios deben acudir para obtener un dato concreto con el mayor nivel de fiabilidad.

Constituye una buena práctica que las compañías determinen qué "tablas" tienen la consideración de "fuente de datos reputada". Este concepto tiene varias implicaciones en los procesos de gobierno del dato:

- El nivel de control sobre esas tablas debe ser mayor. Esto implica que sobre estas fuentes se ejecutará el mayor número de controles de calidad.
- Sobre otras fuentes de información parece razonable aplicar una tipología de controles diferente, orientada a asegurar que la información mantiene la integridad con la fuente reputada. Es decir, en las fuentes que no sean una *Golden Source* debe predominar la ejecución de controles de "Consistencia"[18]. Siendo especialmente útiles los controles de anclaje: la verificación de que los datos son los mismos que los de la fuente de datos reputada.
- Los usuarios saben que, en caso de duda con respecto a algún dato, pueden acudir a la fuente reputada.

Reflexión final sobre las mediciones de calidad

Es conveniente reflexionar sobre el riesgo de ofrecer una falsa sensación de seguridad en las mediciones de calidad.

Ejemplo

¿Podemos estar tranquilos si las mediciones indican que la calidad de los datos de clientes es del 99,8%? Pero pudiera ser que la calidad real no fuera tan buena…

17. Se ha utilizado el término "tabla" para facilitar el entendimiento de este concepto. Aunque, en realidad, sería más preciso hablar de unidades lógicas de información.

18. Ver pagina 23 de esta guía en la que se mencionaban varios criterios de medición.

Este riesgo aparece desde el momento en que las mediciones de calidad no dejan de ser un mero artificio para tratar de medir la calidad real de los datos. Hay que recordar que esa calidad se está midiendo con respecto a una serie de controles establecidos por el responsable del dato y que estas mediciones tienen limitaciones:

- Solo se miden los datos que el responsable considera más relevantes. Por tanto, hay un conjunto de datos sobre los que no se realizan mediciones de calidad.
- Solo se miden con respecto a los controles definidos por el responsable. Por tanto, las mediciones pecarán de optimismo si se han definido pocos controles (o si esos controles son poco relevantes).
- Normalmente, solo se suelen incluir controles automatizables y estos controles no son suficientemente buenos como para detectar errores en la introducción de datos por parte de los usuarios.

Para mitigar este riesgo, se proponen varias actuaciones:

- Realizar pruebas de muestreo físico para validar la calidad de los datos introducidos[19]. Los resultados de estos muestreos deberían introducirse como un control adicional en las mediciones. Estos muestreos tienen un coste muy elevado con respecto a los controles automatizados, por lo que podría ser admisible el que se realicen con una periodicidad distinta a los controles automáticos.
- Ser transparentes a la hora de presentar los resultados de calidad. Y, especialmente, de cara al Consejo de Administración, que debe conocer la existencia de este riesgo de falsa sensación de seguridad.
- Medir un conjunto de datos relevante (en relación al riesgo del uso de cada dato). Y que los controles sean adecuados y exigentes. Tal y cómo ya se ha comentado, Auditoría Interna puede aportar gran valor a la hora de garantizar esta cuestión, aplicando su sentido crítico al valorar la suficiencia de este apartado.

El rol de Auditoría Interna

Ofrecer aseguramiento en la calidad del dato supone un reto considerable para los auditores internos. El enfoque con el que realizar las revisiones dependerá mucho del grado de madurez de la organización con respecto a la medición de los datos. En este sentido, se pueden distinguir dos tipos de compañías:

- Compañías maduras: las que tengan implantado un sistema de medición de calidad de los datos descrito en este capítulo.
- Compañías poco maduras: aquellas que no disponen de un mecanismo para medir la calidad de los datos.

AUDITORÍA INTERNA DE LA CALIDAD DEL DATO EN COMPAÑÍAS POCO MADURAS

Resulta razonable que la labor de Auditoría Interna se centre en analizar si el nivel de calidad del dato es adecuado en compañías que no pueden cuantificarlo. Y, por tanto, la labor de Auditoría Interna se centra más en opinar sobre esa calidad.

Es conveniente reflexionar sobre el riesgo de ofrecer una falsa sensación de seguridad en las mediciones de calidad.

19. Estos muestreos no tienen relación con los muestreos que puede realizar Auditoría Interna. Las mediciones por muestreo que se sugieren en este punto deberían ser realizadas por la Segunda Línea de Control.

Auditoría Interna debe escoger un alcance "razonable" y para ello es importante acotar el perímetro a auditar.

En estos casos, uno de los enfoques más habituales consiste en auditar y supervisar la exactitud, validez, completitud y existencia de los datos que residen en las bases de datos de una determinada aplicación o proceso de negocio. Es decir, Auditoría Interna se encargará de seleccionar un conjunto de datos, definir y ejecutar controles de calidad, y analizar los resultados.

Para implantar esta estrategia, es conveniente seguir algunos consejos que facilitarán la tarea:

- Como se ha recomendado en el apartado *Mejores prácticas para medir la calidad del dato*, Auditoría Interna debe escoger un alcance "razonable". Es importante acotar el perímetro a auditar. Probablemente, sea más adecuado realizar diez trabajos de auditoría interna diferentes (de diez aplicaciones o procesos de negocio) antes que intentar abarcar en un único trabajo muchas aplicaciones o procesos de negocio. Y se deben seleccionar aquellos atributos que se consideren más relevantes.
- Se ha de intentar llegar a la causa raíz que ha provocado los hechos observados. En ocasiones esto no será posible por la complejidad técnica y tiempo que podría suponer evaluar las incidencias detectadas; pero, al menos, se debe asegurar que el dato realmente presenta debilidades antes de incluir el hecho en el informe.
- Probablemente, el auditor interno no aporte mucho valor si simplemente se limita a emitir el informe con el resultado de los controles de calidad que ha ejecutado. De igual forma que el propietario de los datos debe analizar una "cata" o muestra de las incidencias para tener una percepción de lo que está ocurriendo, el auditor interno debería hacer una "cata" de las incidencias, y no únicamente para tratar de identificar la causa raíz, sino también para tener la seguridad de que realmente se trata de incidencias reales en los datos. En caso contrario, un informe de Auditoría Interna puede alarmar a la compañía innecesariamente.

Ejemplo de enfoque

Supongamos que nos encontramos ante una compañía que no está realizando mediciones de calidad de sus datos y, por tanto, desconoce su grado de calidad. Aquí la labor de Auditoría Interna será muy similar a la que se ha descrito en los ejemplos del capítulo anterior:

1. **Definir el alcance del trabajo.** El departamento de Auditoría Interna toma dos decisiones en esta fase del trabajo:
   a. Decide centrar la revisión en analizar la calidad de los datos de clientes porque no parece eficiente ni razonable opinar sobre todos los datos de la organización en un solo trabajo de auditoría interna.
   b. Además, se observa que un cliente está caracterizado por 90 campos. Y Auditoría Interna decide que el alcance del trabajo se va a centrar en 40 campos, por considerarse los más relevantes.
2. **Definir cómo se medirá la calidad.** Estas decisiones hay que tomarlas individualmente por cada uno de los 40 campos. Por ejemplo, con respecto el campo "PROFESIÓN", Auditoría Interna decide realizar tres mediciones (controles) distintos.
3. **Medir la calidad de los datos elegidos en el alcance.** Se ejecutan los controles de calidad definidos en el punto 2 para cada uno de los 40 campos seleccionados en el alcance. A diferencia del proceso que debería ejecutar la compañía, no se espera que se definan agregaciones de calidad de las diferentes mediciones.
4. **Realizar recomendaciones de mejora,** en función de los resultados obtenidos y del análisis de las incidencias detectadas en el paso anterior.

AUDITORÍA INTERNA DE LA CALIDAD DEL DATO EN COMPAÑÍAS MADURAS QUE ESTÁN MIDIENDO LA CALIDAD DE SUS DATOS

En este caso, el principal trabajo de Auditoría Interna será revisar el programa de medición de calidad que tenga la compañía. Es decir, si una compañía ya está midiendo cuál es la calidad de sus datos, el objetivo principal de Auditoría Interna sería opinar acerca de si está de acuerdo con esas mediciones.

El trabajo de Auditoría Interna se podría dividir en dos grandes bloques:

- Analizar si la forma en que se miden los datos es adecuada.
- Analizar el grado de cobertura de la medición de datos.

Analizar si la forma en qué se miden los datos es adecuada. Se podría, por ejemplo, coger una muestra de los diferentes cuadros de mando que se han descrito en este capítulo y analizar la razonabilidad de los procesos que se aplican hasta llegar a ese cuadro de mando.

En esta revisión, el auditor interno debe cuestionarse el funcionamiento del sistema ¿con qué periodicidad se mide? ¿cómo se van agrupando los cálculos de calidad? ¿son las reglas de medición adecuadas? ¿se miden todos los campos relevantes de ese conjunto de datos? ¿se definen planes de remediación para los datos que presentan baja calidad?

Ejemplo de enfoque de Auditoría Interna

Siguiendo el ejemplo, nos encontramos ante una compañía que está midiendo la calidad de los datos de sus clientes, indicando que el nivel de calidad es del 95%. Ese dato agregado se descompone, a su vez, en varios componentes. Uno de ellos es el campo PROFESIÓN, cuya calidad es del 97%.

En este caso, el trabajo de Auditoría Interna se podría centrar en responder a estas preguntas:

- ¿Estoy de acuerdo con que la calidad de los datos de CLIENTES es del 95%? ¿cómo se ha calculado este dato? ¿cómo se realizan las agregaciones?
- ¿Estoy de acuerdo con que la calidad de los datos del campo PROFESIÓN es del 97%? ¿cómo se ha calculado este dato? ¿cómo se realizan las agregaciones? ¿son adecuados los controles que se realizan sobre ese campo? ¿sobran o faltan algunos controles?
- ¿Se están llevando a cabo planes de remediación? ¿Dichos planes ofrecen buenos resultados y la calidad de los datos va mejorando con el tiempo? ¿Los umbrales marcados para fijar los "semáforos" son adecuados?
- ¿Los cuadros de mando con los resultados de las mediciones son adecuados para que los responsables del dato puedan realizar su función (usabilidad)? ¿Cada cuánto tiempo se mide? ¿Se puede hacer *drill down* hasta el dato último, con el objeto de analizar las incidencias?

Incluso en compañías con controles maduros, también se debe reflexionar acerca de la importancia que tienen para la compañía los datos que se están midiendo, de forma que se podrían plantear pruebas de auditoría interna *adhoc* en datos especialmente críticos.

Analizar el grado de cobertura de la medición de datos. Se trata de analizar si se está aplicando el proceso de medición de datos a los conjuntos de datos que se consideran más relevantes en la entidad.

Ejemplo de enfoque de Auditoría Interna

En este caso, el trabajo de Auditoría Interna se podría centrar en analizar si, además de medir los datos de clientes, se están midiendo otras tipologías de datos en la compañía. Y si existen planes creíbles de extender estas mediciones al resto de conjunto de datos.

Si la compañía ya mide la calidad de sus datos, Auditoría Interna analizará la forma y el grado de cobertura de esa medición.

Conclusiones

En las compañías poco maduras el esfuerzo de Auditoría Interna para realizar pruebas analíticas de datos es mayor. Es Auditoría Interna la que define en muchas pruebas las reglas que quiere testar para opinar sobre si un conjunto de datos tiene la suficiente calidad o no.

En cambio, en las compañías maduras, Auditoría Interna se centra más en analizar la razonabilidad del procedimiento de medición, y no tanto en realizar pruebas analíticas.

La trazabilidad del dato se refiere al registro de evidencias que se pueden consultar para averiguar cuándo y quién ha modificado un dato concreto.

TRAZABILIDAD DEL DATO

Es un término ambiguo que se utiliza para muchas cuestiones relacionadas con los sistemas de información. Uno de los usos más habituales de este término es el de referirse al **registro de evidencias que se pueden consultar para averiguar cuándo y quién ha modificado un dato concreto.** A esta trazabilidad también se la conoce como *log* o *Audit Trail*[20].

En cambio, cuando se habla de calidad del dato, el término trazabilidad se utiliza para definir otro concepto: **la capacidad de una compañía para conocer el ciclo de vida de sus datos.** En general, la trazabilidad debe permitir a una empresa responder a las siguientes preguntas:

- En relación a un dato concreto e individual: ¿cómo nace? ¿cómo se transforma? ¿cómo y dónde se utiliza?
- En relación a un informe: ¿qué datos se utilizan para construir el informe? ¿cómo se utilizan esos datos para construir el informe?

Se puede decir que una organización tendrá una buena trazabilidad si es capaz de contestar a esas preguntas en un periodo corto de tiempo y sin emplear demasiados recursos.

Sin embargo, disponer de esta trazabilidad es costoso para las compañías. ¿Qué ventajas aporta? Entre otras, que la trazabilidad ayuda a mejorar y optimizar los procesos, facilita la identificación de inconsistencias en los datos, la detección rápida del origen de las incidencias, así como el impacto que tiene en los procesos cualquier cambio que se realice en el ciclo de vida del dato.

Mejores prácticas para medir la trazabilidad del dato

En el contexto de trazabilidad de datos, aún se está en un estadio inicial de definición. En nuestra opinión, aún no existen unos términos claros y unas expectativas claras sobre qué es la trazabilidad y cómo debe construirse.

La trazabilidad completa implica que, para cada uno de los datos, debe conocerse –en cualquier momento y de manera rápida– el uso que tiene, las transformaciones que ha sufrido, las dependencias con otros datos y el origen de los mismos. Esta información debe

20. *Audit Trail:* versión anglosajona para el término "traza de auditoría", referido al registro de la secuencia de actividades que han afectado a una operación o evento.

disponerse a nivel de sistema, tabla/fichero y campo físico. Además, es importante que cada dato y transformación tenga asignado un responsable que garantice la calidad del mismo.

Esta implementación puede llegar a ser muy costosa dependiendo de la complejidad de los sistemas, procesos y datos de la compañía. Por ello, éstas deben decidir la profundidad y el nivel de detalle al que quieren llegar.

En cualquier caso, empieza a existir cierto consenso en hablar de dos tipos distintos de trazabilidad:

TRAZABILIDAD FUNCIONAL

Es la representación, a alto nivel, del proceso de generación de un dato desde su origen. Debe definir y documentar claramente, para cada dato considerado relevante por la compañía, los siguientes aspectos:

- Qué otros datos son requeridos y qué fórmulas se aplican para obtener ese dato relevante.
- Si el dato es simple o calculado a partir de otros datos, si se ajusta manualmente, y el sistema donde reside.
- El uso que la compañía hace del dato. Es decir, qué áreas lo utilizan en sus procesos y para qué; y si está incluido en algún informe que se reporte a la Dirección o al Consejo de Administración.

En esta trazabilidad funcional se muestra el sistema/aplicación dónde se almacena el dato, pero no se indica el nombre de la tabla y del campo donde se almacena físicamente. Esta información es necesaria para poder replicar el dato y esto es lo que aportaría la trazabilidad técnica.

TRAZABILIDAD TÉCNICA

Llega a un mayor grado de detalle ya que para cada dato (simple o calculado) se debe documentar la base de datos, tabla y campo donde se almacena físicamente. Además, debe incluir el formato y los valores permitidos en cada campo. Esta trazabilidad es más completa y costosa.

Dentro de los diferentes grados de madurez de las empresas, suele elaborarse en un primer momento la trazabilidad funcional y, posteriormente, la trazabilidad técnica, ya que requiere un mayor esfuerzo de documentación. Para una correcta implantación es necesario disponer de una herramienta robusta que facilite la incorporación de información y su mantenimiento a lo largo del ciclo de vida del dato. Esta herramienta debe permitir visualizar la trazabilidad a los usuarios autorizados y analizar dependencias a la hora de cambiar parte del proceso.

La trazabilidad implica conocer, en cualquier momento y de manera rápida, el uso que tiene cada dato, su origen y dependencia de otros datos y posibles transformaciones.

El rol de Auditoría Interna

Es habitual que la verificación de la trazabilidad de los datos de una organización no se aborde dentro de una única revisión de Auditoría Interna, sino en varias (normalmente por grupo o tipología de datos). También es habitual que se verifique sólo un subconjunto de datos (aquellos que se puedan considerar de mayor criticidad).

Las tareas habituales en este tipo de revisión son:

- Verificar si se ha definido un modelo de trazabilidad, y si está alineado con la estrategia de la organización.

- Verificar la suficiencia e idoneidad de la trazabilidad implantada en la compañía. Para ello, se suele actuar de la siguiente forma:
  - Seleccionar una muestra de datos o informes.
  - Para esa muestra, comprobar si, en un tiempo razonable, se puede contestar a las preguntas que se mencionaban en la introducción:
    - Dado un dato concreto e individual, ¿cómo nace? ¿cómo se transforma? ¿cómo y dónde se utiliza?
    - Dado un informe, ¿qué datos se utilizan para construir el informe? ¿cómo se utiliza?
- Verificar si existe una herramienta robusta que soporte el modelo de trazabilidad establecido.
- Validar los controles establecidos para garantizar la integridad y trazabilidad a lo largo del ciclo de vida del dato.

Muchas compañías integran en su marco de gestión de calidad del dato aspectos relacionados con la calidad de los informes.

CALIDAD DE LOS INFORMES

Los órganos de gobierno de las empresas necesitan informes de gestión adecuados y confiables para la toma de decisiones. Estos informes son clave para que las compañías alcancen sus objetivos.

Según los *Principios para una eficaz agregación de datos sobre riesgos y presentación de informes de riesgos"*del Comité de Supervisión Bancaria de Basilea, el proceso de *reporting* de información de riesgos está relacionado directamente con los principios de Exactitud, Exhaustividad, Claridad y utilidad, Frecuencia y Distribución (principios 7 al 11).

Impulsadas especialmente por estos principios, muchas organizaciones integran en su marco de gestión de calidad del dato aspectos relacionados con la calidad de los informes.

Por ejemplo, como se ha visto en el apartado de este documento sobre nuevos actores en el gobierno del dato, algunas compañías han introducido la figura del **Responsable de informes** cuyas principales funciones son:

- Garantizar la calidad de los informes.
- Decidir los criterios que se deben utilizar a la hora de construir un informe.
- Controlar el acceso y distribución de los informes.
- Elaborar un registro de informes para mejorar el grado de control sobre los mismo.

El rol de Auditoría Interna

Para evaluar la calidad de los informes hay que plantearse, fundamentalmente, dos cuestiones:

¿CÓMO ESTRUCTURAR LOS TRABAJOS EN EL PLAN DE AUDITORÍA INTERNA?

Existen dos posibles enfoques:

- **Realizar un único trabajo** en el que se seleccionen los informes más relevantes que se reportan a la Dirección y al Consejo de Administración (o a sus comisiones delegadas). De esta forma se revisaría una muestra de todos los informes.
- **Encajar estas revisiones dentro de las auditorías internas de los procesos de negocio** que, de forma habitual, son parte de la actividad de Auditoría Interna. Una prueba

adicional que se realizaría sería valorar la calidad de los informes más relevantes que se generan en ese proceso de negocio. Probablemente, este segundo enfoque aporte un mayor valor a la compañía.

¿CÓMO ABORDAR LOS TRABAJOS Y QUÉ ASPECTOS EVALUAR DE LOS INFORMES?

Las revisiones de Auditoría Interna pueden ir orientadas a verificar si en los informes se están siguiendo los siguientes principios:

- **Prontitud.** Identificar los niveles de servicio, frecuencias y plazos para la agregación, generación y distribución de los datos que se requieren para la elaboración del informe. Y comprobar que están formalmente establecidos, que se monitorean y que cumplen los criterios del negocio.
- **Exhaustividad.** Analizar que los informes de riesgos cubren todas las áreas de riesgo significativas, y su profundidad y alcance están en consonancia con el tamaño y complejidad de las actividades de la compañía. Evaluar si los informes dirigidos a la Alta Dirección ofrecen una evaluación prospectiva del riesgo y no se basan únicamente en datos pasados y presentes.
- **Claridad y Utilidad.** Evaluar si los informes son fáciles de entender, y si existe un glosario donde se definan los conceptos que se utilizan. Valorar si la información presentada a la Alta Dirección es pertinente y adecuada para el proceso de gobierno y la toma de decisiones.
- **Frecuencia.** Revisar si la periodicidad establecida para los informes es acertada, está alineada con los criterios de negocio y se cumple.
- **Distribución.** Evaluar el circuito de distribución establecido, comprobando que garantiza tanto la confidencialidad de la información como su envío a los destinatarios pertinentes.
- **Exactitud.** Evaluar si el informe final es correcto, en función de los datos disponibles. Para esta evaluación, habitualmente se aplican dos tipos de enfoque:
  - Partir del reporte final e identificar y revisar los diferentes procesos y entradas hasta llegar al origen de la información.
  - Evaluar el proceso final a partir de un sistema intermedio sin necesidad de llegar al sistema fuente.

  Es importante destacar que la revisión de este último punto entraña un alto grado de dificultad, y consume muchos recursos de Auditoría Interna.

Auditoría Interna verificará si los informes están siguiendo los criterios de prontitud, exhaustividad, claridad, frecuencia y exactitud, entre otros.

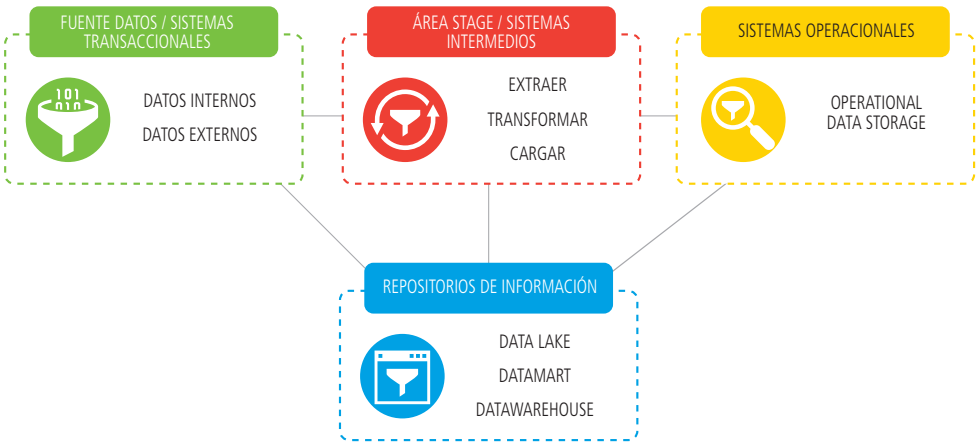
Arquitectura técnica: facilitador del buen gobierno del dato

La arquitectura técnica soporta el ciclo de vida completo del dato. Engloba a los sistemas fuente, sistemas intermedios de transformación, sistemas transaccionales, así como otros de tipo informacional que permiten la explotación de los datos. Es decir, debe permitir que se puedan tratar los datos para tomar las decisiones de negocio.

Además, una adecuada arquitectura permitirá la aplicación de los atributos a gobernar mencionados anteriormente (confidencialidad, trazabilidad, calidad, etc.).

Por estos motivos, la arquitectura desempeña un papel importante a la hora de determinar el éxito o el fracaso de un modelo de gobierno del dato robusto.

En términos de sistemas, y considerando el ciclo de vida del dato, el siguiente gráfico refleja el modelo típico de arquitectura técnica.



Fuente: Elaboración propia

ETAPAS

Fuentes de datos

Son aquellos sistemas origen del dato. Pueden ser tan heterogéneos como las tipologías de información y procesos de negocio, desde la capa de presentación *(front)* de cualquier sistema con acceso por parte de empleados y/o clientes hasta dispositivos del internet de las cosas *(IoT)*[21], una aplicación instalada en el *smartphone* o el último dispositivo de uso diario (relojes, pulseras, e incluso las últimas deportivas), pero también los servidores del regulador o una contraparte, la información de los mercados o los propios tornos de acceso a un edificio. Puesto que los orígenes pueden ser tan variados el diseño de la arquitectura debería considerar las amenazas desde este punto.

21. IoT: Acrónimo *Internet of Things.*

Sistemas intermedios

Área temporal *(middleware)* en la que se recogen los datos que se necesitan de los sistemas fuente. En esta área, y de forma escalonada, se realiza el proceso ETL[22] completo a través de interfaces de comunicaciones (SOAP[23], microservicios, etc.) y se envían datos desde múltiples fuentes. Los datos deben adecuarse al formato de los sistemas transaccionales, mantener la integridad y calidad en el proceso y asegurarse su confidencialidad cuando se requiera. Según el proceso de negocio, en ocasiones estos sistemas solo actúan de puente hacia el ODS[24], lugar en el que puede concluirse el proceso de ETL.

Sistemas operacionales (ODS)

Los datos recibidos de múltiples fuentes, si no se han producido errores en el proceso, serán explotados por el sistema operacional, ya sea para la realización de nuevas operaciones o como base para herramientas de *reporting*. Los sistemas trabajan en línea, no almacenan datos históricos y muestran solo la imagen del momento actual; por tanto, es imprescindible que la calidad y disponibilidad de los datos sea adecuada a los requisitos del proceso.

Repositorios de Información

El almacén de datos corporativos o *Data Lake,* sirve de repositorio de datos en bruto. Por el contrario, el *DataMart* o *DataWarehouse* son sistemas orientados al análisis de datos (OLAP[25]) y contienen datos históricos optimizados para su explotación y la inteligencia de negocio. En él se almacenan datos que pueden provenir tanto de sistemas intermedios, de sistemas operacionales o sistemas fuente. Para una adecuada explotación de los datos es necesario que los mismos estén organizados, sean coherentes entre distintas fuentes y mantengan un formato común.

Es importante disponer de un diccionario de datos a nivel funcional (procesos y actividades de negocio) y físico (sistemas TI). Este diccionario debe estar reflejado en la arquitectura técnica que soporta el ciclo de vida. Ésta deberá implementar un **MARCO DE CONTROL** robusto que ayude a identificar:

- **Errores en los datos:** de formato, datos vacíos, desactualizados, falta de integridad.
- **Fallos en los procesos** que impidan disponer de los datos.
- **Impacto en la calidad** de los datos ante cambios, procesos nuevos o manualidad del proceso.

Los datos deben adecuarse al formato de los sistemas transaccionales y mantener la integridad y calidad.

EL ROL DE AUDITORÍA INTERNA

Los equipos de Auditoría Interna, dentro de sus planes, deben analizar si la infraestructura tecnológica es adecuada para soportar el gobierno del dato. Entre los principales aspectos

22. ETL: concepto sin referencia concreta, de uso común en el marco de la arquitectura. Es el acrónimo de *Extract, Transform* and *Load* o "Extraer, Transformar y Cargar".
23. SOAP o *Simple Object Access Protocol,* terminología habitual para referirse a servicios web.
24. ODS u *Operational Data Storage*. Sistemas finales donde los datos son operados y almacenados.
25. Referencia para *On-Line Analytical Processing,* termino con el que se conoce a muchos sistemas de inteligencia de negocio.

que deben ser evaluados durante la revisión estarían:

- Instalación y mantenimiento de la infraestructura tecnológica.
- Escalabilidad de la infraestructura.
- Disponibilidad, monitorización y tiempos de respuesta para el usuario.
- Gestión de incidencias.
- Respaldo y Recuperación.
- Metodología de Desarrollo y mantenimiento de las aplicaciones

Auditoría Interna debe analizar en sus planes si la infraestructura tecnológica es adecuada para soportar el gobierno del dato.

En este apartado, el equipo de Auditoría Interna debe decidir si realizar una revisión específica centrada en este aspecto o si, por el contrario, estas tareas se incluyen dentro de otros trabajos de Auditoría Interna de TI.

En el primer caso, el equipo de Auditoría Interna podría realizar un trabajo específico que bien podría denominarse *"Adecuación de la plataforma tecnológica al gobierno del dato"*.

En el segundo enfoque, igualmente válido, pasaría por integrar las tareas de revisión de otros trabajos de auditoría interna más amplios.

Aquí se podría citar como ejemplo la gestión de incidencias. Si utilizamos el primer enfoque, se analizaría, entre otras cuestiones, la gestión de las incidencias de los sistemas que estén relacionados con el gobierno del dato.

Pero esto implicaría dejar de ver la gestión de incidencias en otros sistemas que pueden ser relevantes. Lo que a su vez implica que, al final, un proceso como el de gestión de incidencias, que es relativamente uniforme, se revise en varios trabajos diferentes de auditoría interna. Por esto, otras unidades prefieren tener en sus planes de Auditoría Interna algunos trabajos más orientados a procesos de TI. Es decir, tendrán, por ejemplo, un trabajo de *"Gestión de incidencias y problemas"* dónde verán ese proceso de forma general.

Mejores prácticas en la seguridad en acceso a los datos

La información se ve sujeta diariamente a múltiples riesgos o vulnerabilidades: el robo de propiedad intelectual, *ramsonware*, ciberterrorismo, disrupción del servicio, multas por infracción de la regulación de protección de datos, impacto en la reputación, etc.; lo que hace necesario proteger la información clave de la compañía de estos riesgos.

Se necesitaría elaborar varios documentos técnicos para reflejar todos los aspectos que hay que tener en cuenta para garantizar la seguridad de los datos. Existen múltiples fuentes y publicaciones para profundizar en el mundo de la seguridad de los datos. Y, en especial, cabe mencionar a la ISACA[26] como referencia obligada para los auditores de TI.

26. ISACA es el acrónimo de *Information Systems Audit and Control Association* (Asociación de Auditoría y Control de Sistemas de Información). Es una asociación internacional que apoya y patrocina el desarrollo de metodologías y certificaciones para la realización de actividades de auditoría y control en sistemas de información.

La seguridad de la información es el conjunto de prácticas para prevenir el acceso, uso, revelación, disrupción, modificación, inspección, registro o destrucción de la información no autorizados. Dichas prácticas suelen regirse sobre los principios de Confidencialidad, Integridad y Disponibilidad que –según la ISO 27001[27]– se extenderían a los sistemas implicados en su tratamiento, dentro de una compañía.

El marco de gobierno de acceso al dato que establezcan las compañías debe asegurar que las personas adecuadas acceden a la información adecuada, en el momento adecuado y por las razones adecuadas. Este marco debería incluir la implementación de los siguientes componentes clave:

- Un registro completo de todos los sistemas, propietarios, y la criticidad asociada a la compañía (valor aportado).
- En base a dichos niveles de criticidad, determinar el grado de seguridad a aplicar y el conjunto de medidas de seguridad que cumplan con los requisitos definidos.
- Clarificación de la responsabilidad para el cumplimiento de las normas de seguridad relacionadas con la gestión de accesos.
- Un conjunto estandarizado de procesos y procedimientos para la gestión de accesos a los sistemas de información que se encuentre documentado e implantado. Este debería incluir una definición de los roles aceptados y no aceptados, y las combinaciones de roles para cada sistema.
- Actividades de formación y concienciación en una cultura de seguridad del dato.
- Un régimen de aseguramiento aplicado en base a la criticidad de los sistemas.

Es de vital importancia que se mantenga un adecuado equilibrio entre la protección de la Confidencialidad, Integridad y Disponibilidad de la información y una implementación eficiente de la política, sin afectar la productividad de la compañía. El coste de la implantación de los controles debe ser proporcional al riesgo que se quiere mitigar.

Seguridad de la información implica prevenir el acceso, uso, revelación, disrupción, modificación, inspección, registro o destrucción no autorizado de la información.

EL ROL DE AUDITORÍA INTERNA

El conjunto de actividades que se deben desarrollar en este ámbito es muy variado y excede del objetivo de este documento. No obstante, a continuación se indican algunas de las actividades susceptibles de llevar a cabo por Auditoría Interna:

Medidas Técnicas

- Evaluar la efectividad de las soluciones de prevención de pérdida de datos o DLP *(Data Loss Prevention)* orientadas a la monitorización y control.
- Asegurar la eficacia de las medidas de seguridad en fuentes extraíbles, enfocados a garantizar la seguridad de la información que se transmite entre puntos físicos.
- Asegurar el correcto diseño y efectividad de los controles establecidos en relación con el funcionamiento de medidas de seguridad concretas (por ejemplo, cifrado, *firewalls*, gestión de actualizaciones).

27. ISO 27001: Norma internacional emitida por la Organización Internacional de Normalización (ISO) que describe cómo gestionar la seguridad de la información en una empresa. https://www.iso.org/isoiec-27001-information-security.html

- Revisión de las medidas de protección de dispositivos y aplicaciones a través de los cuales se puede acceder a grupos de datos, teniendo en cuenta la criticidad de estos.
- Revisión del proceso de gestión de cambios a programas.

El coste de la implantación de los controles debe ser proporcional al riesgo que se quiere mitigar.

Medidas Organizativas

- Asegurar la adecuación, aprobación y revisión periódica de políticas, procedimientos y normativa interna. Verificar que se comunica correctamente al conjunto de profesionales implicados (incluyendo terceros, si aplica).
- Revisar la adecuada asignación de responsabilidades sobre los datos, asegurando la claridad en cuanto a la propiedad de estos.
- Revisión de la correcta seguridad en los accesos (revisión de usuarios, perfiles y roles, y parámetros de seguridad).
- Asegurar que los procesos de almacenamiento garantizan la confidencialidad de los datos almacenados.

Medidas Jurídicas

- Revisión de contratos realizados con proveedores que puedan impactar en la seguridad y confidencialidad de los datos, asegurando que la compañía establece requerimientos adecuados en materia de seguridad de los datos tratados de la compañía y de terceros.
- Correcta definición de acuerdos de nivel de servicio con los proveedores o acuerdos de confidencialidad o no divulgación con los empleados, regulando los aspectos clave e incluyendo sanciones en caso de incumplimiento.

Instituto de Auditores Internos de España

Santa Cruz de Marcenado, 33 · 28015 Madrid · Tel.: 91 593 23 45 · Fax: 91 593 29 32 · www.auditoresinternos.es

Depósito Legal: M-5490-2020
ISBN: 978-84-120500-3-5

Diseño y maquetación: desdecero, estudio gráfico

Impresión: Impresión Artes Gráficas, IAG

OTRAS PRODUCCIONES DE LA FÁBRICA DE PENSAMIENTO

AUDITORÍA INTERNA DE LA INFORMACIÓN EXTERNALIZADA

Este documento aborda los principales aspectos normativos que deben considerarse en relación con la gestión y control de la información externalizada; y recomendaciones relativas al rol que Auditoría Interna debería desempeñaren las etapas del proceso de externalización, desde la fase de pre-contratación hasta la finalización de la prestación del servicio.

SUPERVISIÓN DE LA ÉTICA EMPRESARIAL

La ética empresarial es el resultado de las acciones impulsadas por la Alta Dirección de una organización, siguiendo las directrices de su Consejo de Administración, para fomentar que empleados y colaboradores actúen conforme a unos principios éticos que respondan a las expectativas de los stakeholders.

EVALUACIONES INTERNAS DE CALIDAD DE LA DIRECCIÓN DE AUDITORÍA INTERNA

La Guía Técnica 3/2017 de la CNMV sobre Comisiones de Auditoría de Entidades de Interés Público señala que la Comisión de Auditoría debe supervisar la actividad de Auditoría Interna, para lo que puede tomar como referencia el Marco Internacional para la Práctica Profesional de la Auditoría Interna del Instituto de Auditores Internos.

EVALUACIONES INTERNAS DE CALIDAD DE LA DIRECCIÓN DE AUDITORÍA INTERNA

La Norma 1300 - Programa de Aseguramiento y Mejora de la Calidad, del Marco Internacional para la Práctica Profesional de la Auditoría Interna asigna al Director de Auditoría Interna el desarrollo y mantenimiento de un completo Programa de Aseguramiento y Mejora de la Calidad (PAMC).

LA FÁBRICA DE PENSAMIENTO
INSTITUTO DE AUDITORES INTERNOS DE ESPAÑA

A medida que crece el caudal de datos que gestionan las organizaciones, se vuelve crítico contar con una gestión adecuada de los datos. Y para ello es clave contar con un marco sólido de gobierno del dato.

Este documento aborda tanto los problemas como las mejores prácticas para definir un buen gobierno del dato. Se analizan a fondo varios aspectos, desde el ciclo de vida del dato –incluyendo la trazabilidad y calidad de este– hasta metodologías y normativas aplicables en el proceso de gobierno del dato. Todo ello desde la perspectiva de Auditoría Interna.

Instituto de Auditores Internos de España

The Institute of Internal Auditors