



Felipe Maggi

CDO

DATARMONY



datarmony.com

Asegurar la calidad de los datos

Datokey frente a la espada de Damocles

La calidad de los datos es esencial para alcanzar el éxito en los negocios. Sin embargo, es el eslabón más débil de la cadena de explotación y activación de la enorme cantidad de información de la que ahora disponen las empresas. En este contexto, el objetivo de Datokey es reforzar dicho eslabón y asegurar así la toma de decisiones acertadas.

La Real Academia Española define la expresión *espada de Damocles* como “amenaza persistente de un peligro”. La frase proviene de una anécdota de la cultura clásica griega: Dionisio I de Siracusa se intercambia por un día con Damocles, un cortesano que envidiaba su riqueza y su poder. Damocles recibe entonces las atenciones de un rey, y disfruta de su suerte hasta que repara en la espada que colgaba sobre su cabeza, sostenida por una fina crin de caballo.

Esta es una realidad que han sufrido empresas de los sectores *mice* y *retail*, y a las que Datarmony ha ayudado a través de su solución de monitorización y alerta de calidad del dato: Datokey.

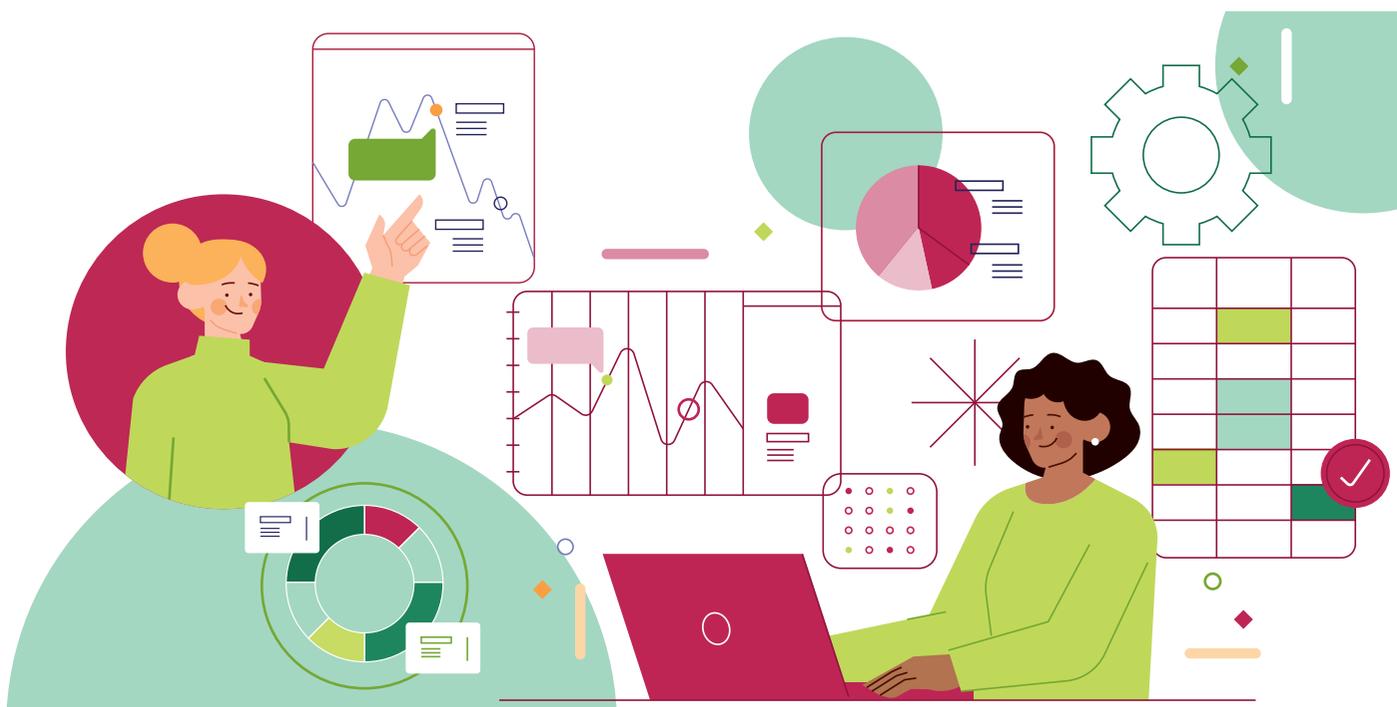
LA CALIDAD DE LOS DATOS

El Ministerio para la Transformación Digital, dependiente de la Secretaría de Estado de Digitalización e Inteligencia Artificial del Gobierno de España, dice: “Los modelos de IA se basan en grandes volúmenes de datos para obtener resultados precisos y confiables. Si los datos utilizados están contaminados o son de baja calidad, los resultados de los algoritmos de IA serán poco confiables o incluso erróneos. Por lo tanto, garantizar la calidad de los datos es esencial para lograr el máximo rendimiento de las aplicaciones de IA, reducir o eliminar sesgos y aprovechar su potencial al máximo”.

Pero no es necesario adentrarse en cuestiones tan sofisticadas como la inteligencia artificial para apreciar el peligro de tomar decisiones erróneas basadas en datos de baja calidad. Por poner un ejemplo, basta con perder la trazabilidad correcta de las acciones de los usuarios en nuestras webs o *apps* para visualizar embudos de conversión que no reflejen la realidad y nos lleven a programar cambios de diseño mal enfocados. Para garantizar la calidad de los datos necesitamos, en primer lugar, definir qué entendemos por tal. Así, según la norma UNE ISO/IEC 25012, las características que definen la calidad de los datos son exactitud, completitud, consistencia, credibilidad, actualidad, accesibilidad, conformidad, eficiencia, precisión, trazabilidad, comprensibilidad, disponibilidad, portabilidad y recuperabilidad.

Con Datokey los tiempos dedicados a la detección y solución de errores en la recopilación y procesado de datos digitales se reducen en un 50%

La historia simboliza aquellas situaciones en las que nos amenaza un peligro constante, a veces no evidente, que nos hace imposible disfrutar de una situación que parece envidiable. En la era de los datos, y en el contexto de su explotación, la calidad es precisamente nuestra espada de Damocles. La moraleja de la historia es perfectamente aplicable en este caso: contamos con una riqueza y un poder para tomar decisiones informadas desconocidos hasta la fecha, pero la calidad de la información que usamos para tomar esas decisiones es una espada que pende de un hilo muy delgado, siempre a punto de romperse.



De ellas, la accesibilidad, la disponibilidad, la portabilidad y la recuperabilidad están muy bien cubiertas con los servicios disponibles en la nube (Google Cloud, AWS, Azure), gracias a los cuales contamos con bases de datos de alto rendimiento y espacio de almacenamiento prácticamente ilimitado.

En cuanto a la comprensibilidad, el concepto está bastante bien resuelto por las herramientas de visualización de datos, cada vez más potentes y capaces no solo de conectarse a diversas fuentes de información, sino de procesarla para presentarla de forma clara, concisa y directa. Entre ellas destacan PowerBI, Tableau, Looker y Qlik.

UNA ESPADA DE DOBLE FILO

Cubiertos los aspectos anteriores, resta dar una solución satisfactoria a los puntos de exactitud, completitud y consistencia, ya que, si los datos son erróneos, la facilidad para acceder a ellos es más una trampa que una ventaja. Es precisamente en estos aspectos donde herramientas como Datokey adquieren un papel protagonista.

Las empresas deben adoptar políticas de gobernanza de datos, algo nada sencillo ni de plantear ni de ejecutar. Es más: incluso tras haber adoptado dichas políticas en el marco de un proyecto bien definido y estructurado, su mantenimiento suele ser un desafío que rara vez se resuelve bien.

Si tenemos en cuenta la cantidad de factores que hay en juego veremos que es prácticamente imposible garantizar el cumplimiento de las normas relacionadas con la calidad del dato, y ni establecer normas, patrones o directrices ni nombrar responsables para velar por dichas normas garantizará que se cumplan.

El resultado, en la práctica, es que los equipos internos de desarrollo y mantenimiento de aplicaciones y servicios están más ocupados corrigiendo errores de recogida y procesamiento de datos que desarrollando y mejorando las aplicaciones necesarias para la marcha del negocio.

MONITORIZACIÓN CONSTANTE

Ante esta situación, existen soluciones —completamente personalizables— que monitorizan de forma constante la cantidad, la calidad y la consistencia de los datos que las empresas utilizan para tomar decisiones.

Datokey compara constantemente la información de los sitios webs y las aplicaciones que reciben las herramientas de análisis con las reglas y patrones definidos por el cliente en materia de calidad de datos; después envía alertas cuando esas reglas no se cumplen o cuando las tendencias de las métricas esenciales se salen de los patrones normales.

La solución combina reglas de negocio, técnicas estadísticas avanzadas e inteligencia artificial. Se trata no solo de alertar de un problema tan pronto como este se produce, sino de, a partir

de cambios de tendencia no esperados, anticiparse a posibles conflictos.

Los usuarios pueden acceder a un cuadro de mando para observar directamente las tendencias de calidad de los datos, así como las alertas que se hayan generado. Además —y para ser fieles a la filosofía de que, en lugar de ir a buscar el dato, el dato debe ir a buscar a sus responsables—, se informa de cada incidencia inmediatamente a través de mensajes enviados por los canales que el cliente desee habilitar (con el correo electrónico como vía más común). Una vez recibida la alerta, los responsables pueden descargarse un fichero con toda la información técnica relacionada con la incidencia.

Un aspecto esencial de esta solución es que los datos que monitoriza nunca salen del entorno del cliente. Datokey se conecta con los datos, disponibles en la plataforma que se utilice, pero no los almacena en otro sitio ni los duplica. Como empresa dedicada al análisis de datos, Datarmony es plenamente consciente de que no puede favorecer que surjan silos de información.

¿El resultado?: Datarmony estima que los tiempos dedicados a la detección y solución de errores en la recopilación y procesamiento de datos digitales se reducen en un 50%. De este modo se libera a los equipos de TI de tareas repetitivas y de escaso retorno. ■