



Conoce las 5 tendencias en Data Quality

que dominarán las organizaciones hasta 2027



Material elaborado por Merlin Data Quality, en el marco del 8° Congreso Tecnología y Negocios América Digital México 2023.



Índice

Data Quality en las organizaciones	3
5 tendencias en Data Quality a 2027	4
Aumento del uso de tecnologías de automatización y su sustento en la calidad de datos.....	5
Enfoque en la gestión de datos en tiempo real.....	6
Integración y calidad de datos.....	7
Mayor énfasis en la privacidad y la seguridad de los datos.....	8
Uso de tecnologías de inteligencia artificial y aprendizaje automático.....	9
Prioridades de acción en materia de calidad de datos	10

Data Quality

en las organizaciones

La calidad de los datos es un factor clave en el éxito de cualquier empresa, especialmente en un mundo cada vez más impulsado por los datos. Data Quality es la dimensión de los datos que refiere a la precisión, integridad, relevancia, coherencia y oportunidad de la información, y es esencial para la toma de decisiones informada y estratégica, así como también para el crecimiento del negocio. Por esta razón, es importante que las empresas establezcan una cultura de calidad de datos en la que todos los miembros de la organización valoren y prioricen la excelencia de los mismos.

Un informe de Gartner ya estimaba que para el año 2022, **las iniciativas de calidad de datos serían de prioridad crítica para el 60% de las organizaciones**. La misma consultora destacó, en otra investigación posterior, que el **47% de las organizaciones consideran que la calidad de los datos sería su principal desafío en la gestión de datos**, seguido de la gobernanza de datos y la seguridad de los datos.

Otro reporte, de Forrester Research, encontró que **el 22% de las empresas pierde más de \$20 millones al año debido a datos inexactos o incompletos**.

Las tendencias en data quality están evolucionando a medida que la tecnología y las demandas del mercado cambian. Hacia 2027 se espera que la calidad de los datos siga siendo un tema crítico para las organizaciones y que los responsables de datos continúen en la búsqueda de nuevas soluciones para mejorarla.

Son varias las predicciones en torno al data quality para 2027. Resumimos en este documento las cinco más relevantes, que ninguna empresa puede dejar de monitorear.

5 tendencias

en Data Quality a 2027

1.

Aumento del uso de **tecnologías de automatización** y su sustento en la calidad de datos

El número de organizaciones que utilizan tecnologías de automatización inteligente en su modelo de negocio no para de crecer. El estudio anual de Deloitte “Automation with intelligence”, ya mostraba en 2020 que tres de cada cuatro empresas habían incorporado tecnologías como la RPA (automatización de procesos mediante robots), el machine learning o aprendizaje automático, el procesamiento de lenguaje natural o el BPM (Gestión de Procesos de Negocio), entre otras.

En nuestro campo, la automatización de procesos de gestión de datos se ha vuelto cada vez más importante, y se espera que esto continúe en los próximos años. En la era digital actual, las empresas dependen cada vez más de la tecnología para procesar y analizar grandes volúmenes de datos. Sin embargo, estos registros sólo son útiles si son precisos, completos y confiables.

La adopción de herramientas y soluciones de automatización para mejorar la calidad de los datos puede ayudar a reducir la cantidad de errores humanos. Por ejemplo, en lugar de ingresar manualmente los datos en una base de datos, se pueden utilizar herramientas online de normalización de datos durante su carga. Esto minimiza la probabilidad de errores tipográficos y otras fallas que pueden afectar la calidad de la información.

Otra forma en que la automatización puede mejorar la calidad de los datos es a través de la validación automática de los mismos. Las herramientas de validación pueden detectar al instante datos que no cumplen con ciertos criterios, como direcciones de correo electrónico inválidas o números de teléfono incorrectos.

Además, la automatización puede ayudar a garantizar la consistencia de los datos. Si se utiliza una herramienta de automatización para importar datos de diferentes fuentes, se pueden aplicar las mismas reglas de limpieza de datos para garantizar que los registros sean coherentes y estén estandarizados.

2.

Enfoque en la gestión de **datos en tiempo real**

Otro fenómeno de los últimos años es la exponencial relevancia que las organizaciones están asignando al uso de los **datos en tiempo real**. A nivel corporativo es masiva la adopción de soluciones para recopilar y analizar datos en tiempo real, impulsadas por la necesidad de gestionar información online en distintas operaciones y transacciones de clientes.

Sabemos que son múltiples los beneficios de trabajar con datos en tiempo real:

- Ayuda a minimizar los costos en diferentes áreas de la organización.
- Permite ofrecer un tiempo de respuesta rápido y soluciones más acertadas.
- Facilita la observación de cambios en el mercado.
- Mejora la toma de decisiones y hace que esta sea más rápida a medida que aumenta el ritmo del negocio.
- Reduce la probabilidad de errores tanto en procesos operativos como en los de comunicación.

Pero, obtener datos precisos y actualizados al instante es otra cosa, ya que implica la detección y corrección de errores de datos en tiempo real, antes de que ingresen a las bases de datos y afecten la calidad de los datos en su conjunto. Vale la pena explorar, además de las soluciones que permiten el procesamiento y análisis de datos de forma inmediata a medida que se generan, cómo complementarlo con herramientas que aseguren la calidad de los datos generados.



3.

Integración y calidad de datos

Según una encuesta realizada por Ventana Research, **el 56% de las empresas cree que la integración de datos es uno de los mayores desafíos que enfrentan en torno a la gestión de datos.** Es que, si hablamos de retos, uno que se ubica en los primeros puestos, es sin dudas la integración de diferentes fuentes de datos en un solo sistema. La combinación de datos heterogéneos provenientes de múltiples fuentes en una sola aplicación, es conocida como proceso de integración de datos.

Este proceso permite la fusión de diversos tipos de información, como documentos, tablas y matrices de datos, para que sean utilizados en procesos de negocio o en funciones, tanto por usuarios individuales como por organizaciones y aplicaciones. La integración de diversas fuentes de datos en una única fuente es importante porque permite tener una visión más completa y precisa de la información que se está manejando. Al combinar diferentes fuentes de datos, se pueden descubrir patrones y relaciones que no serían visibles si se analizara cada fuente de forma independiente.

Además, la integración de datos también ayuda a reducir la redundancia y la inconsistencia de los datos, ya que se pueden eliminar datos duplicados y armonizar la información que proviene de diferentes sistemas y formatos. IDC, la firma líder mundial en investigación de mercado que se enfoca en tecnología de la información, telecomunicaciones y electrónica de consumo, encontró que **la falta de datos de calidad puede llevar a una disminución del 20-30% en la eficacia de los proyectos de análisis de datos.**

La calidad de los datos es crucial en el proceso de integración de datos. Si los datos no son precisos, completos o consistentes, la integración no producirá resultados útiles y puede incluso ser perjudicial para la organización. Es importante asegurar que los datos sean de alta calidad antes de integrarlos en una única fuente.

La integración de datos también puede mejorar la calidad de los datos en sí misma, ya que, al combinar diferentes fuentes de datos, se pueden detectar errores y discrepancias que pueden ser corregidos antes de integrarlos. De esta manera, la integración de datos no solo mejora la precisión y la integridad de la información, sino que también ayuda a mantener la calidad de los datos en el futuro.

4.

Mayor énfasis en la **privacidad** y la **seguridad** de los datos

La privacidad y la seguridad de los datos se han vuelto una preocupación constante para las organizaciones, y se espera que esto continúe en los próximos años. Las empresas están implementando soluciones para proteger los datos confidenciales de sus clientes y mejorar la transparencia en el uso de la información.

Un estudio de la empresa de software de gestión de datos Talend encontró que el **84% de los responsables de TI creen que los datos son inexactos, incompletos o duplicados en su organización**. La falta de calidad de los datos puede tener un impacto significativo en la toma de decisiones y la eficacia de la organización, pero, por sobre todo, la información falsa o errónea puede ocasionar problemas de seguridad y falta de confianza en las organizaciones.

La falta de control y transparencia en la recopilación de datos puede llevar a la inclusión de datos no precisos o incompletos, lo que afecta la calidad de los datos. A su vez, la ausencia de políticas de calidad de datos puede ocasionar la pérdida de datos, la inclusión de datos incorrectos o incompletos y la disminución de la veracidad y confianza en los mismos.



5.

Uso de tecnologías de inteligencia artificial y aprendizaje automático

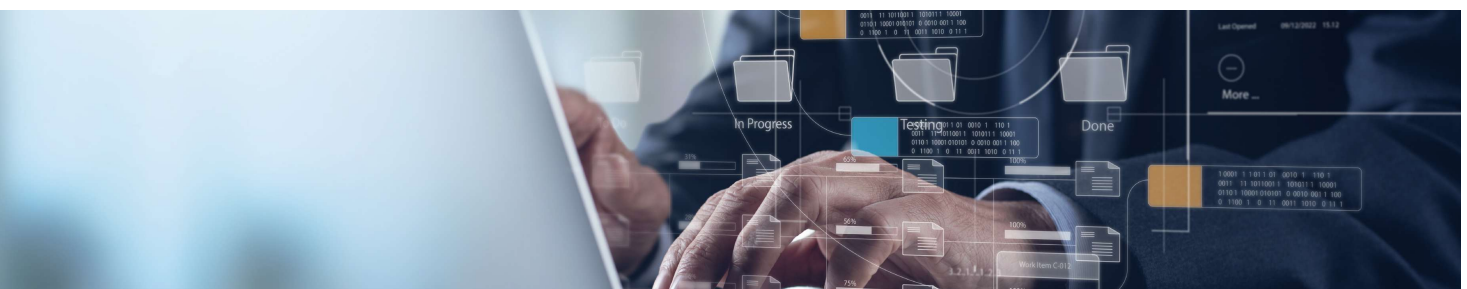
El uso de tecnologías de inteligencia artificial (IA) y aprendizaje automático (AA) ha revolucionado la forma en que las empresas trabajan con sus datos. Estas tecnologías permiten procesar grandes cantidades de información de manera rápida y eficiente, lo que a su vez puede mejorar significativamente la calidad de los datos.

Gartner predice en uno de sus reportes que **para el año 2025, el 75% de las empresas utilizarán IA y ML para automatizar sus procesos de negocios.**

La IA y el AA pueden usarse para detectar y corregir errores en los datos, así como para identificar patrones y tendencias que podrían pasar desapercibidos para los analistas humanos. Además, pueden ayudar a identificar y resolver problemas de calidad de los datos en tiempo real.

Sin embargo, es importante tener en cuenta que la calidad de los datos utilizados para entrenar los modelos de IA y AA es crucial para garantizar que los resultados sean precisos y confiables. Los datos deben ser limpios, completos y consistentes para que los modelos puedan generar predicciones y recomendaciones precisas.

Por lo tanto, es esencial que las empresas se centren en mejorar la calidad de sus datos antes de implementar tecnologías de IA y AA. Esto puede lograrse mediante la implementación de políticas y prácticas efectivas de gestión de datos y la utilización de herramientas de calidad de datos. Al hacerlo, las empresas pueden aprovechar al máximo estas tecnologías para mejorar la calidad de sus datos y tomar decisiones basadas en información precisa y confiable.





Prioridades de acción en materia de calidad de datos

La calidad de los datos se ha convertido en una prioridad crítica para muchas empresas, ya que los datos precisos y confiables son esenciales para la toma de decisiones empresariales y para el éxito en el mercado. Para mejorar la calidad de sus datos, las empresas están tomando diversas iniciativas que incluyen:

Evaluación y auditoría de la calidad de los datos, para identificar problemas de calidad y definir planes de acción para abordarlos.

Implementación de procesos de entrada de datos más rigurosos, para garantizar que los datos sean capturados de manera precisa y consistente desde el principio.

Establecimiento de políticas de gestión de datos, incluyendo la definición de estándares de calidad de datos, roles y responsabilidades, y procesos de resolución de problemas.

Uso de herramientas de calidad de datos, para identificar y corregir problemas de calidad de datos de manera más eficiente.

Capacitación del personal, para mejorar la conciencia y la comprensión de la importancia de la calidad de los datos y para asegurar que el personal tenga las habilidades y conocimientos necesarios para gestionar los datos de manera efectiva.

En conclusión, las iniciativas para mejorar la calidad de los datos son cruciales para el éxito empresarial. Al establecer políticas claras, mejorar los procesos de entrada de datos, implementar herramientas y capacitar al personal, las empresas pueden garantizar que los datos sean precisos y confiables, lo que se traduce en una mejor toma de decisiones y en una ventaja competitiva en el mercado.

Merlin Data Quality

Es una plataforma integral de calidad de datos cuyo objetivo es colaborar en minimizar las problemáticas expuestas y maximizar los resultados del negocio mediante la utilización de datos confiables, precisos y veraces. Somos una compañía argentina con expansión regional dedicada desde hace más de dos décadas al tratamiento de la calidad de datos.

Nuestro propósito ha sido siempre promover la importancia de la calidad de los datos para la toma de decisiones exitosa.

**Te invitamos a explorar nuestras soluciones para
la gestión de calidad de datos.**

merlindataquality.com



hola@merlindataquality.com

