

Guía de Calidad de Datos

El 80% de las decisiones empresariales dependen de datos confiables.

En un mundo cada vez más impulsado por los datos, la **calidad de los datos** es un factor crucial que influye directamente en la capacidad de una organización para tomar decisiones precisas, oportunas y efectivas. La fiabilidad de los datos que utilizamos determina la eficacia operativa, la competitividad y la capacidad de innovar. **Sin datos confiables, incluso las mejores estrategias pueden fracasar.**

Realizado por: Keisa Avila

keisaavila.com





¿Qué es la Calidad de Datos?

La calidad de los datos es el ADN de las decisiones inteligentes. Sin datos precisos, completos y consistentes, las organizaciones están navegando a ciegas. En un mundo impulsado por datos, la calidad no es negociable. Pero, ¿qué significa realmente tener datos de alta calidad? Vamos a desglosarlo:

Precisión

Datos que reflejan con exactitud la realidad que representan. Sin errores ni distorsiones.

Integridad

Información completa sin omisiones críticas que afecten su interpretación o uso.

Consistencia

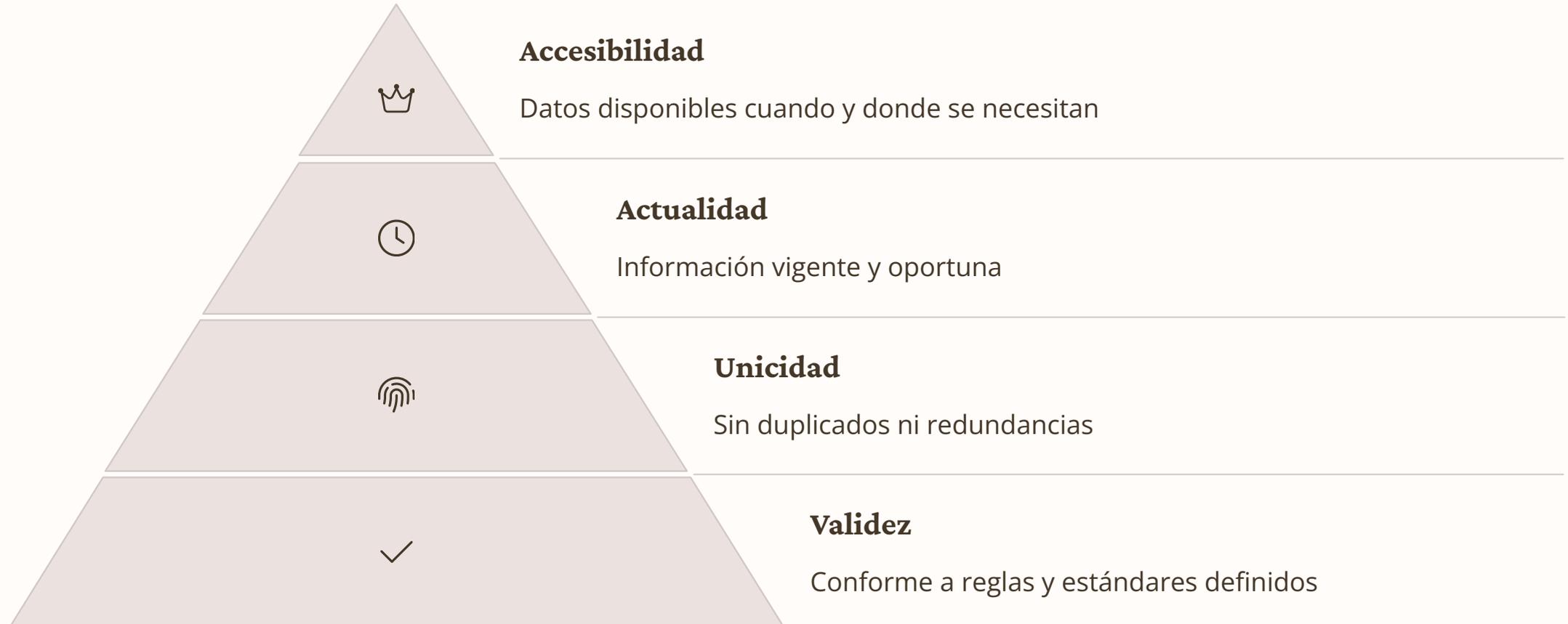
Datos uniformes entre diferentes sistemas, reportes y periodos de tiempo.

Estándares y Dimensiones de Calidad

La Calidad de los Datos: El Pilar de las Decisiones Inteligentes

No basta con tener datos; **necesitas los datos correctos**, en el momento adecuado y en la forma que mejor te permita tomar decisiones estratégicas. Los datos de calidad no solo se miden por su exactitud, sino por su **accesibilidad, actualidad, unicidad** y **validez**. Cada uno de estos pilares es crucial para garantizar que las decisiones basadas en datos se tomen de manera informada.

Si los datos no son confiables, la estrategia nunca será efectiva. La pirámide que verás a continuación muestra las dimensiones esenciales que debes dominar para construir una infraestructura de datos sólida y capaz de transformar la información en ventajas competitivas.



Proceso para Garantizar Calidad

La calidad de los datos es un **proceso constante y estructurado** que abarca desde su recolección hasta su uso en decisiones empresariales claves. Este proceso involucra la captura de datos correctos y requiere una gestión eficiente en cada etapa para asegurar que los datos sean confiables y útiles.

El ciclo presentado a continuación muestra las **etapas esenciales**, para mantener la **calidad de los datos** en todo momento, garantizando su precisión, integridad y validez a lo largo del tiempo:



Empowering Data, Securing Futures

Roles y Responsabilidades

La calidad de los datos no es tarea de una sola persona o equipo, sino un esfuerzo colaborativo que involucra a diversas partes interesadas en la organización. **Cada rol** tiene una responsabilidad clave para garantizar que los datos sean gestionados adecuadamente, desde su creación hasta su uso en la toma de decisiones. A continuación, se presentan los **roles principales** involucrados en la gestión de datos y sus respectivas responsabilidades para asegurar la calidad en cada etapa del proceso:



Chief Data Officer

Responsable de la estrategia global de datos.
Define políticas y estándares organizacionales.



Data Steward

Custodio de datos por dominio.
Implementa reglas de calidad y resuelve problemas.



Equipo de TI

Infraestructura técnica y automatización.
Gestiona herramientas y procesos de datos.



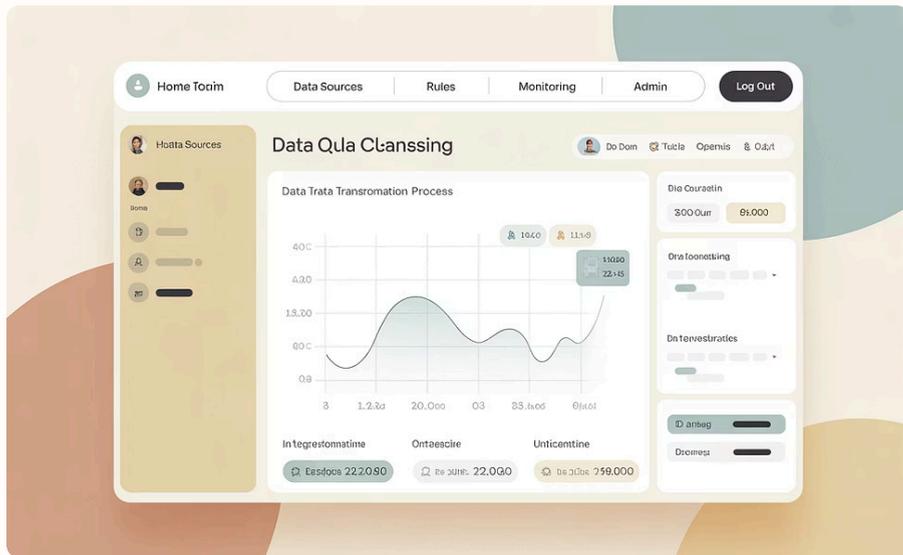
Usuarios de Negocio

Reportan problemas de calidad.
Validan resultados y aplicaciones prácticas.

Herramientas y Mejores Prácticas

Software Especializado

En el mercado hay una serie de herramientas que podemos utilizar, que automatizan la detección de problemas.



Limpieza Continua

Políticas de mantenimiento regulares previenen la acumulación de problemas de calidad.



Monitoreo de KPIs

Tableros de control permiten seguimiento en tiempo real de métricas críticas de calidad.



Ventajas de utilizar herramientas especializadas:

Mejora de la calidad de los datos:

Al automatizar la detección y corrección de errores, se mejora la calidad de los datos y se reduce el riesgo de resultados incorrectos en el análisis.

Ahorro de tiempo y recursos:

Al automatizar tareas de limpieza y corrección de datos, se ahorra tiempo y recursos que podrían ser utilizados en otras tareas.

Mayor eficiencia en el proceso ETL:

Las herramientas especializadas ayudan a que el proceso ETL sea más eficiente y rápido.

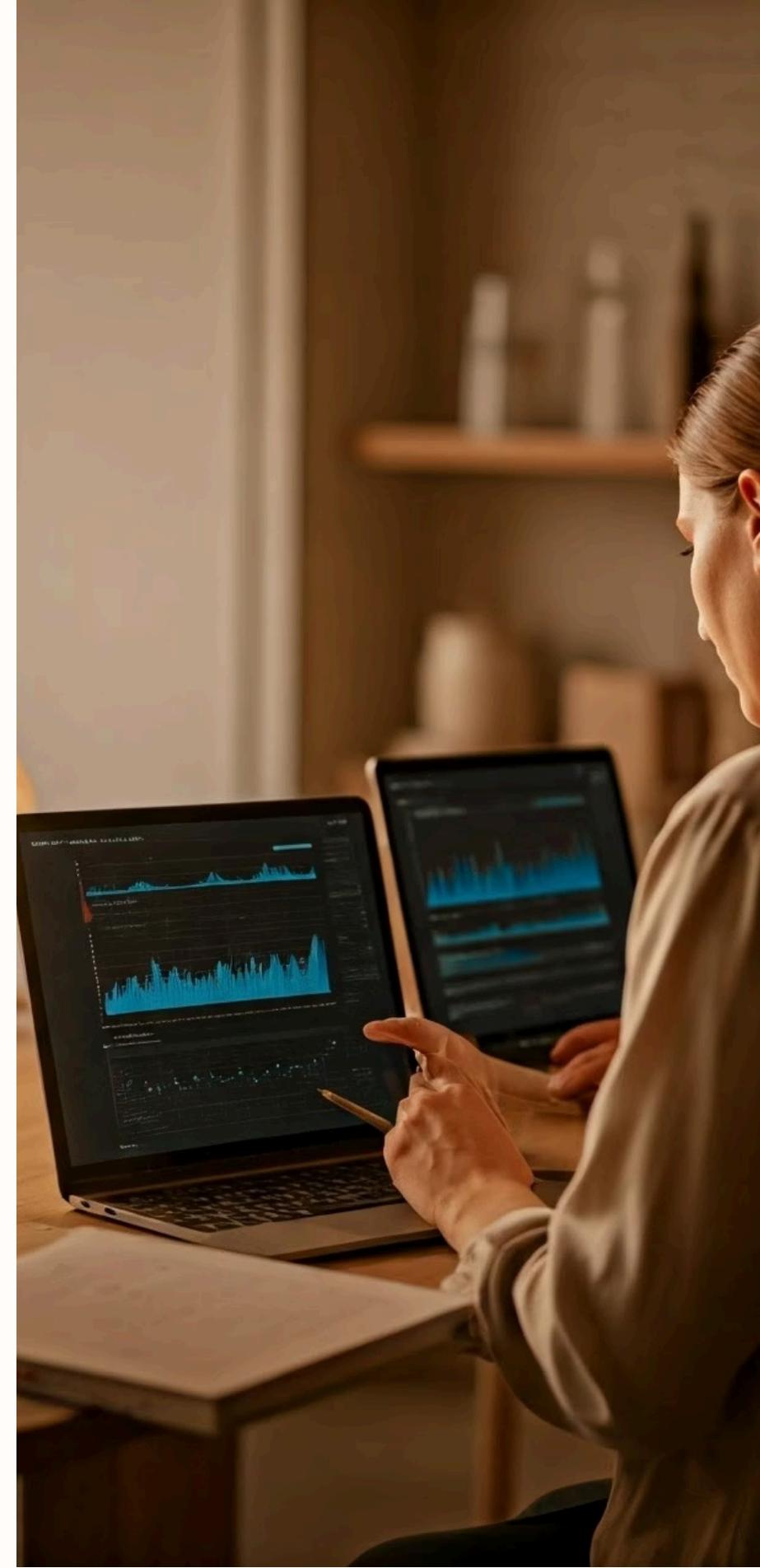
Ejemplos de herramientas especializadas:

Talend Open Studio:

Plataforma de integración de datos que ofrece funciones de limpieza y transformación de datos.

Atacama Uno:

Herramienta de integración de datos con capacidades avanzadas de limpieza y transformación de datos. Cuenta con una opción gratuita de perfilado de datos.





Recomendaciones

Evaluación Inicial

Realice un diagnóstico detallado del estado actual de sus datos.
Identifique áreas críticas y establezca una línea base de calidad.

Implementación Progresiva

Comience con proyectos piloto en áreas de alto impacto.
Expanda gradualmente a toda la organización.

Mejora Continua

Establezca ciclos de revisión regulares.
Actualice procesos según evolucionen las necesidades del negocio.