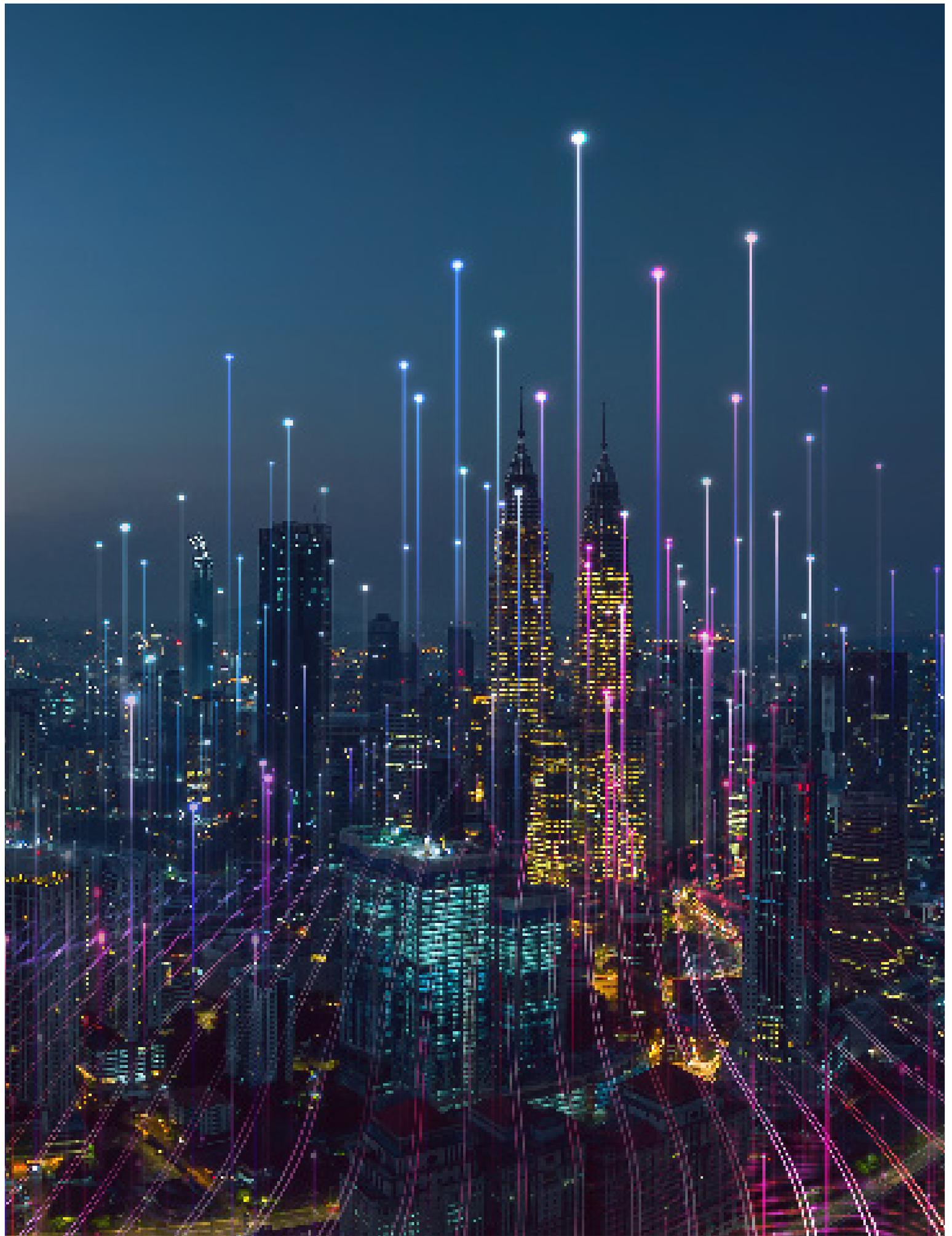




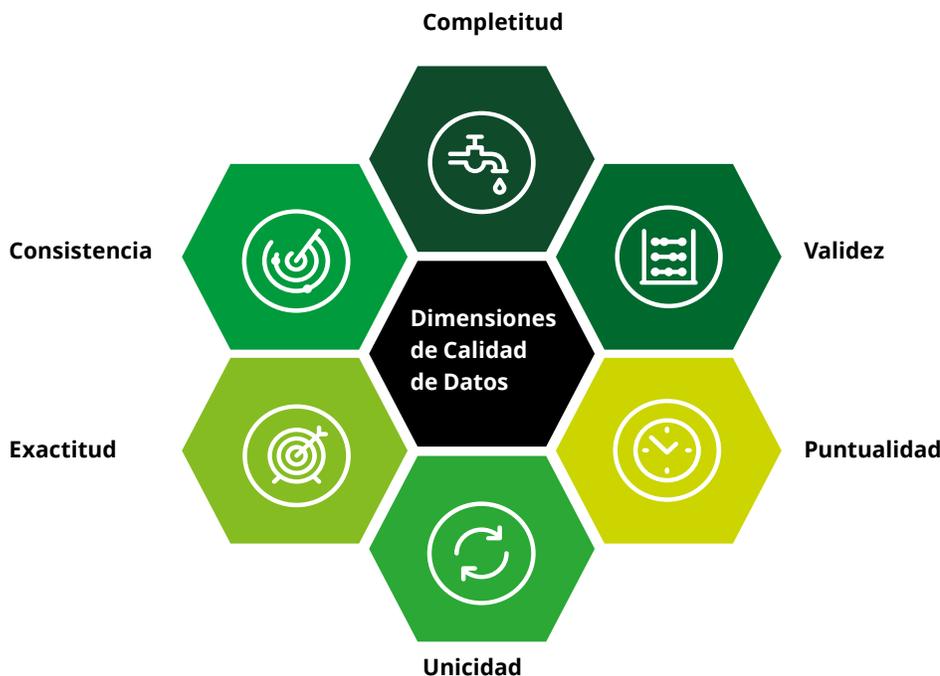
Calidad de Datos en
la era del *Big Data*
Automatización y
priorización como
habilitadores clave





¿Qué es la calidad de los datos y por qué es importante?

Es una comprensión cuantificada de la idoneidad de los datos. Mide la calidad en términos de qué tan bien habilita los procesos de la organización. Esto incluye la planificación, implementación, y control de actividades que aplican técnicas de gestión de calidad a los datos para asegurar el cumplimiento de las necesidades de sus consumidores.



- Cumplimiento normativo
- Toma de decisiones más informadas
- Habilitación de procesos de negocio



¿Por qué puede ser difícil y costoso?

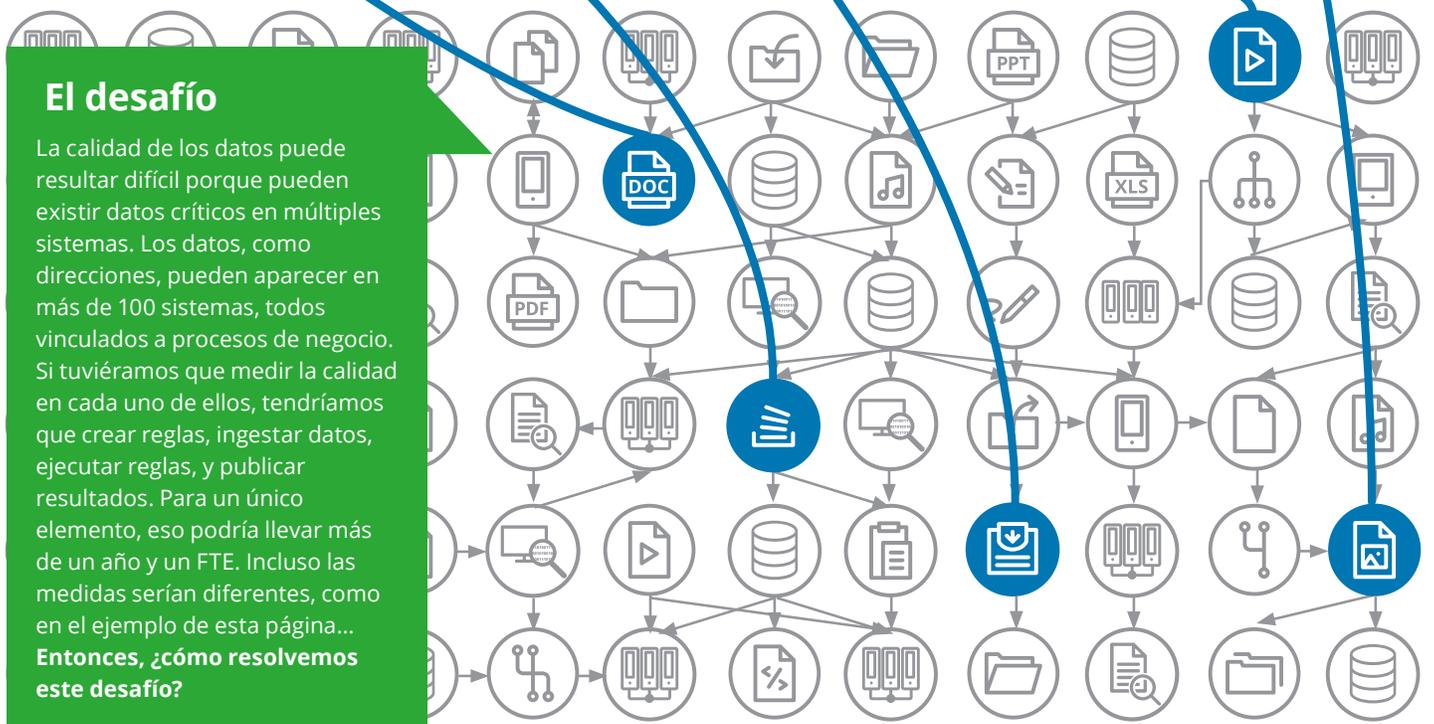
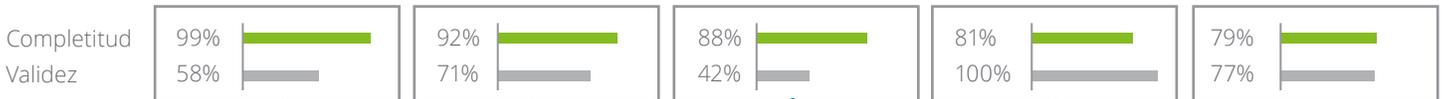
Sergio, Data Steward,
Dominio de Clientes

Objetivo: verificar la calidad de los datos para el elemento de datos direcciones



¿Tenemos un problema de calidad de datos o no? ¿Cómo definir si los datos tienen calidad? ¿Por dónde empezar?

Mediciones de calidad de datos para **direcciones** en las mismas 2 dimensiones, para diferentes sistemas





Dos formas principales que resolverán el desafío de la calidad de los datos



Automatización

Utilizar tecnología disponible para automatizar las tareas de calidad de datos tanto como sea posible:

- Medir la calidad de los datos durante la ingesta y la transformación.
- Utilizar *machine Learning*, módulos relacionado con enriquecimiento de datos, *semantic matching* y descubrimiento de datos.
- Deducir el linaje de los datos escaneando bases y sus metadatos.



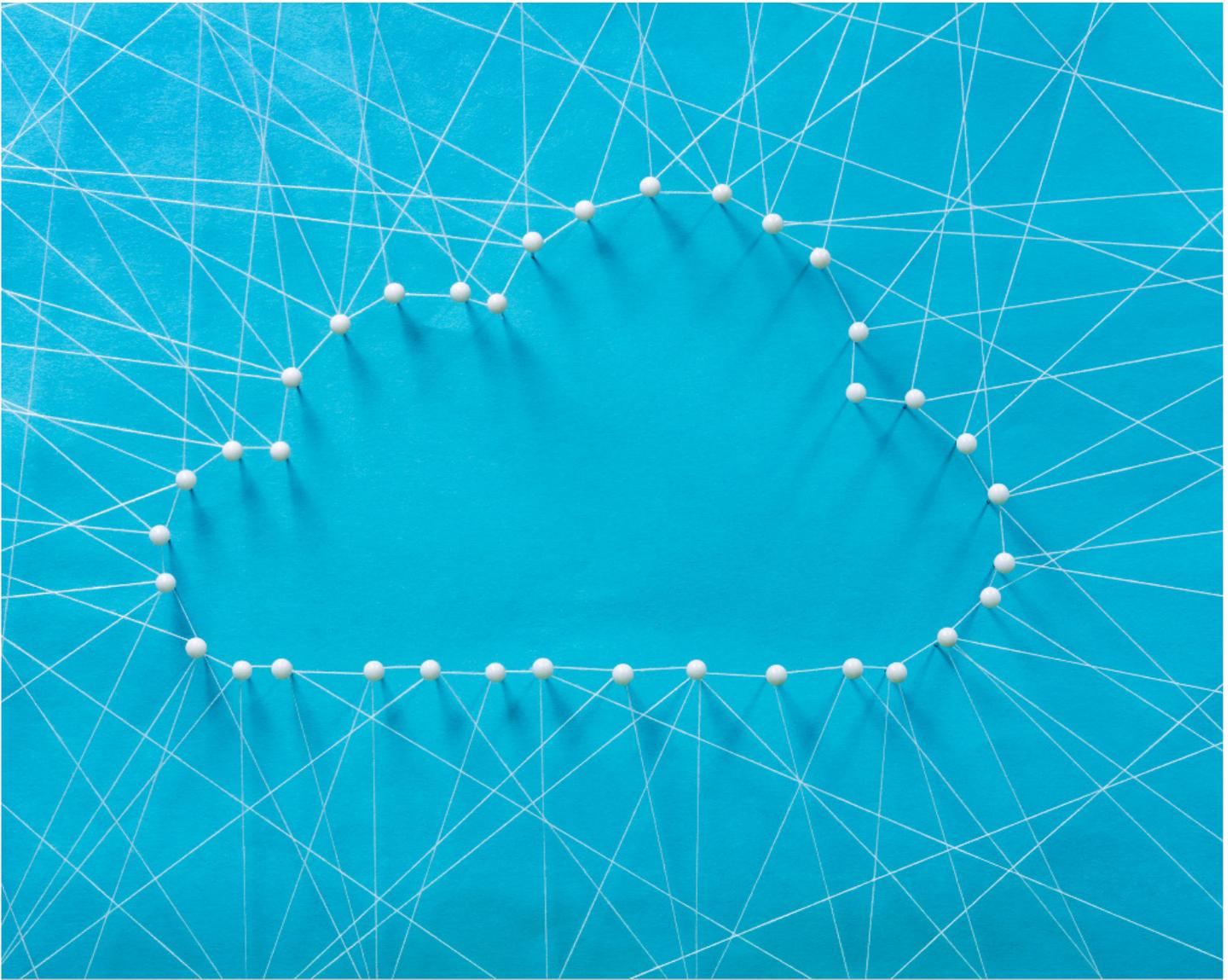
Priorización

Priorizar la ejecución de la calidad de los datos en dos dimensiones principales:

- Solo medir la calidad de los datos para aquellos que sean críticos. Solo priorizar los dominios y los elementos que sean los más críticos para su organización.
- Medir inteligentemente, no todo y todas las fuentes. Adoptar fuentes de oro ("*Golden Sources*") y confiables ("*Trusted Sources*"), y concentrar los esfuerzos de calidad de datos ahí.

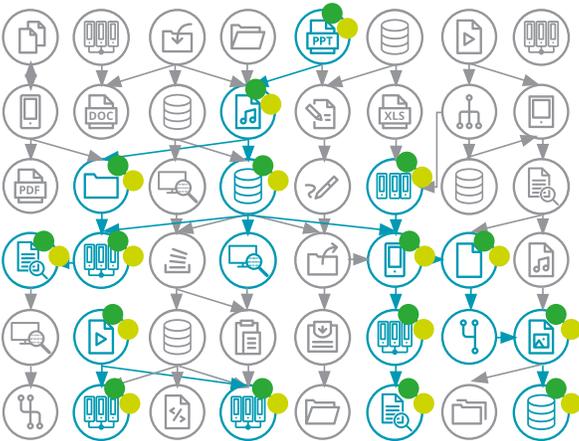
Diseño inteligente

Un diseño inteligente puede reducir el número de reglas de calidad de los datos y, al mismo tiempo, impulsar la coherencia.





Diseño extenso y repetido



Controles / medidas:

- Completitud (16x)
- Validez (16x)

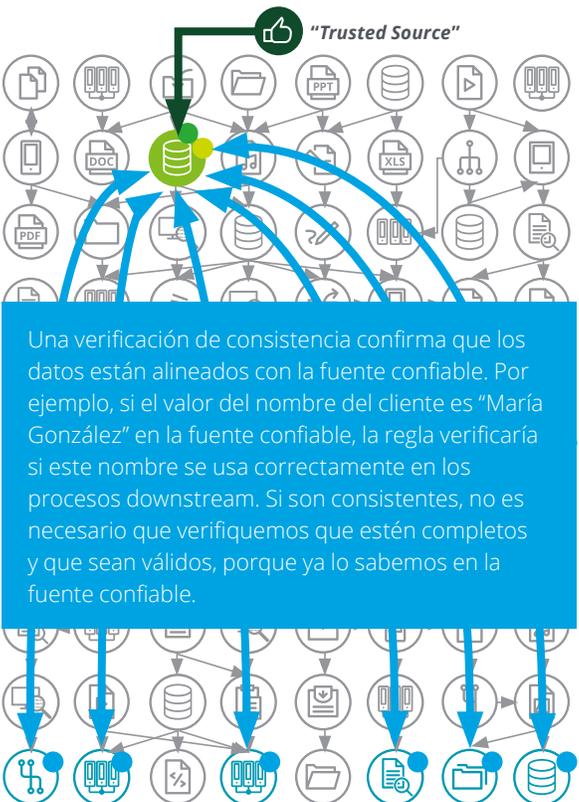
En este ejemplo, tenemos un único elemento de datos existente en 16 sistemas. Esta es una situación bastante común: los datos de clientes y productos, por ejemplo, se pueden utilizar en más de 100 sistemas en muchas organizaciones.

En este escenario, se han implementado 32 controles de integridad y validez para medir la calidad de los datos. Hay dos conclusiones:

- Es extremadamente caro. Tuvimos que trabajar con datos de 16 sistemas y escribir 32 reglas de calidad de datos.
- No hay garantía de coherencia. Idealmente, la calidad de los datos es perfecta en todo el flujo, pero si no, tendremos diferentes medidas de integridad y validez, ¿y cuál vamos a tomar como "verdadera"?



Diseño inteligente



Una verificación de consistencia confirma que los datos están alineados con la fuente confiable. Por ejemplo, si el valor del nombre del cliente es "María González" en la fuente confiable, la regla verificará si este nombre se usa correctamente en los procesos downstream. Si son consistentes, no es necesario que verifiquemos que estén completos y que sean válidos, porque ya lo sabemos en la fuente confiable.

Controles / medidas:

- Completitud (1x)
- Validez (1x)
- Consistencia (6x)

Para el elemento en cuestión, 1 fuente se designa como "Fuente Confiable". Aquí es donde se miden la completitud y la validez. Esta fuente contiene la "verdad" de los datos; si hay alguna pregunta o duda, esta fuente proporcionará la respuesta. No necesitamos medir todos los demás sistemas para verificar su integridad y validez – simplemente tenemos que verificar que los sistemas / procesos finales estén ingesting los datos correctos. Basta una simple comprobación de consistencia. El resultado:

- Solo tenemos 6 reglas de consistencia, y 8 reglas en total (~75% menos). Solo necesitamos ingestar datos de 7 sistemas (~56% menos).
- Tenemos claridad sobre cuál es el verdadero estado de la calidad de los datos.
- Solo si tenemos un problema *downstream*, seguimos los datos *upstream*.

Factores de éxito



Fuentes confiables

Combine esfuerzos de calidad de datos con fuentes confiables y MDM para un mayor impacto. Si puede probar la calidad de los datos más críticos en una base centralizada, será mucho más fácil convencer para que se utilice este punto central. Esto impulsará la calidad de los datos en toda la organización.



La alineación del negocio

Como siempre, asegúrese de que el negocio y el liderazgo estén alineados desde el principio. Esto es necesario para determinar qué datos son realmente críticos. También impulsará que las medidas de calidad puedan convertirse posteriormente en mejoras de los datos.



Políticas y estándares

No inicie la calidad de los datos como un ejercicio separado. Específicamente, asegúrese de que la calidad de estos se apoye en la política de datos y los estándares subyacentes. Esto permitirá el cumplimiento e impulsará la adopción de mejores prácticas.



Enfoque replicable

Cree un enfoque estandarizado que se pueda aplicar a toda la organización. Escriba reglas de forma coherente, guarde los metadatos en un inventario y, cuando sea posible, utilice herramientas compartidas. Esto ayudará a crear una versión compartida de la verdad.

Contactos

Para saber cómo su compañía puede anticipar retos y escoger las mejores prácticas, por favor contactar a:

Carlos Labanda

Socio

Analytics & Cognitive
Deloitte Spanish Latin America
Tel: +57 317 5590866
clabanda@deloitte.com

Willem Koenders

Gerente Senior

Analytics & Cognitive
Deloitte Spanish Latin America
Tel: +505 8395 9756
wkoenders@deloitte.com

www.deloitte.com

Deloitte.

Deloitte se refiere a Deloitte Touche Tohmatsu Limited, sociedad privada de responsabilidad limitada en el Reino Unido, a su red de firmas miembro y sus entidades relacionadas, cada una de ellas como una entidad legal única e independiente. Consulte www.deloitte.com para obtener más información sobre nuestra red global de firmas miembro.

Deloitte presta servicios profesionales de auditoría y assurance, consultoría, asesoría financiera, asesoría en riesgos, impuestos y servicios legales, relacionados con nuestros clientes públicos y privados de diversas industrias. Con una red global de firmas miembro en más de 150 países, Deloitte brinda capacidades de clase mundial y servicio de alta calidad a sus clientes, aportando la experiencia necesaria para hacer frente a los retos más complejos de los negocios. Los más de 330,000 profesionales de Deloitte están comprometidos a lograr impactos significativos.

Tal y como se usa en este documento, "Deloitte S-LATAM, S.C." es la firma miembro de Deloitte y comprende tres Marketplaces: México-Centroamérica, Cono Sur y Región Andina. Involucra varias entidades relacionadas, las cuales tienen el derecho legal exclusivo de involucrarse en, y limitan sus negocios a, la prestación de servicios de auditoría, consultoría fiscal, asesoría legal, en riesgos y financiera respectivamente, así como otros servicios profesionales bajo el nombre de "Deloitte".

Esta presentación contiene solamente información general y Deloitte no está, por medio de este documento, prestando asesoramiento o servicios contables, comerciales, financieros, de inversión, legales, fiscales u otros.

Esta presentación no sustituye dichos consejos o servicios profesionales, ni debe usarse como base para cualquier decisión o acción que pueda afectar su negocio. Antes de tomar cualquier decisión o tomar cualquier medida que pueda afectar su negocio, debe consultar a un asesor profesional calificado. No se proporciona ninguna representación, garantía o promesa (ni explícito ni implícito) sobre la veracidad ni la integridad de la información en esta comunicación y Deloitte no será responsable de ninguna pérdida sufrida por cualquier persona que confíe en esta presentación.