

# Testing de migración de base de datos

Ing. Elena Martirena  
Junio 2023



# Contenido

- Objetivo y contexto
- ¿Qué buscamos con el testing de datos en las migraciones?
- Actividades previas
- Metodología para el testing de migración de datos
- Ejemplos de aplicación
- Ejercicio
- Herramientas de Calidad de datos
- Conclusiones

# Objetivo y Contexto

- Compartir experiencias en testing aplicada a migraciones
- Presentar una metodología para el testing de la migración de datos
- Contexto: Bases de datos estructurados

# ¿Qué buscamos con el testing de datos en las migraciones?

- Que el universo de datos definido llegue completo al destino
- Mejorar la calidad de los datos que se requiere migrar
- Que los datos de origen lleguen bien a destino y sean correctamente impactados en la nueva base
- Proporcionar a la organización una base de datos limpia para incrementar su valor
- Minimizar el impacto negativo en el cliente

# Porque se migran las bases de datos?

- Fusión o adquisición de empresas, cuyo impacto a nivel de los datos será una *fusión e integración de las bases de datos* con un cambio total del modelo de datos.
- Cambiar el software de aplicación cuyo impacto a nivel de los datos será *una evolución o cambio total del modelo de datos existente*.
- Cambio de infraestructura o actualización de la versión del motor de base de datos, quedaría incluido en el segundo caso debido a la similitud en el impacto a nivel de los datos. También puede darse una combinación de escenarios.

# Actividades previas - Causas de los problemas

## Técnicos

- Sistemas Legados
- Evolución tecnológica/ Cambios de arquitectura / Software de base discontinuado
- Existencia de Información redundante. Varias fuentes para la misma entidad.
- Problemas en el diseño de la base de datos
- Errores u omisiones en el ingreso de información (ingreso manual o por IO)
- Modificaciones realizadas sobre los sistemas origen
- Integración de múltiples bases (fusión)

## ● Negocio

- Cambios en reglas de negocio
- Expansión de empresas (campañas comerciales)
- Reglamentaciones de organismos regulatorios (Bancos centrales, impositiva, seguridad social, etc.)
- Falta de capacitación o conocimiento por parte de los que ingresan los datos

# Actividades previas

- Definir universo de datos a migrar
  - Definir el universo de datos a migrar. Ej. Clientes activos, operaciones activas
  - Decidir acciones sobre el el histórico
- Conocer datos de origen a migrar
  - Modelo de datos
  - Inventario de entidades/Tablas a migrar y diccionarios de datos de origen
  - Obtener un diagnóstico de la calidad de los datos fuente

# Actividades previas

- Conocer base de datos destino
  - Modelo de datos
  - Inventario de entidades/Tablas destino, reglas de negocio (Ej. Datos no nulos, codificadores, etc.)
  - Mapeo y transformaciones
- Enfoque técnico: definir herramientas y ambientes donde se ejecutará la limpieza de datos (limpieza previa, aplicación de reglas, generación de informes de resultados)
- Definir un plan de acción que permita mejorar la calidad de la información y que sirva como entrada al proyecto de migración
- Identificar actores



# Principales actores

Área TI

Áreas de negocio

Áreas  
comerciales

Áreas  
Financieras

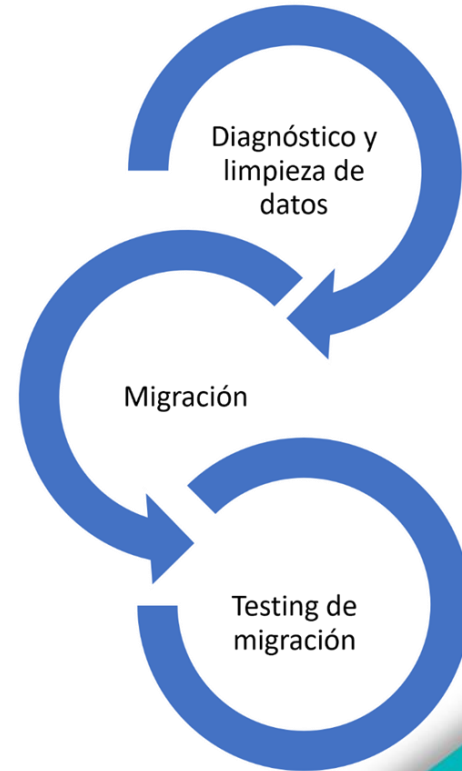
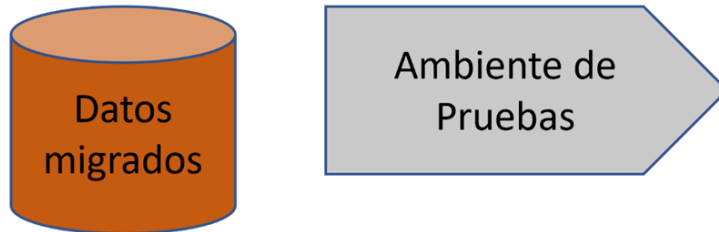
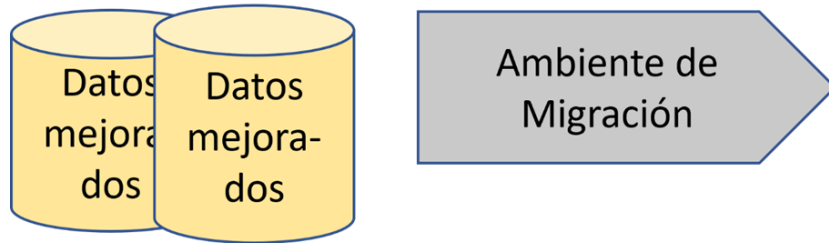
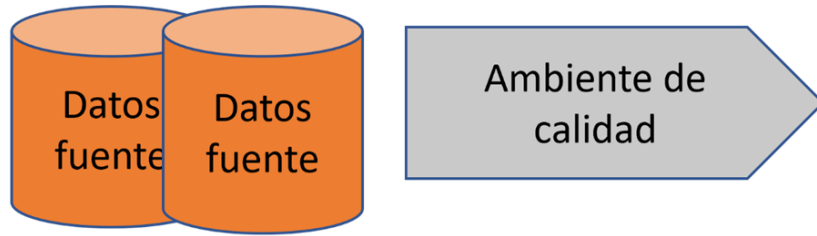
Auditorías

Oficial de  
cumplimiento

Gerencias

Custodios de  
datos

# Metodología para el testing de migración de datos



Plantear 3 proyectos en paralelo

# Diagnóstico de la calidad

Datos inexactos con  
la realidad

Inconsistencia en  
datos cruzados

Datos incompletos

Datos inexactos  
sintácticamente

Datos antiguos

Entidades duplicadas

# Diagnóstico de la calidad-Beneficios

- Contar con el conocimiento del estado actual de la calidad de datos
- Contar, ante cada problema de calidad de datos, con una evaluación del impacto y alternativas de solución
- Reducción de los riesgos asociados a problemas de calidad de datos durante el proceso de migración
- Entregar la nueva base de datos limpia, enfocado a la mejora continua de los datos
- Permitir que la operativa de la empresa sea normal, luego de la migración y que el impacto en el cliente sea el menor posible
- Evitar indisponibilidad del sistema por problemas en la migración de los datos. Ejemplo en sistema de *ebanca* que debe Funcionar 7x24

# Limpieza de datos

- Definir las dimensiones a controlar: (unicidad, exactitud, integridad...)
- Definir acciones correctivas dependiendo del origen de los errores
- Priorizar según la gravedad y la necesidad del negocio.  
Ejemplo: unicidad de la identificación del cliente en una fusión de empresas. Datos frescos y actualizados para *ecommerce*.

# Limpieza de datos

- Planificar donde y cuando se corrigen los datos
  - En las bases de datos origen y los aplicativos para evitar que se continúen produciendo errores
  - En el proceso de migración
  - En el destino
  - A realizar con baja prioridad o incluso desestimar

# Testing de migración -Calidades de los datos migrados



# Cualidades de los datos migrados - dimensiones de calidad

<i>Cualidad de datos migrados</i>	<i>Preguntas</i>	<i>Dimensión de calidad</i>
Unicidad de las entidades migradas	¿Se migraron entidades repetidas en el sistema de información? ¿Hay datos que vinieron de distintas fuentes y quedaron contradictorios para una misma entidad?	Unicidad
Universo de la migración	¿Se migraron todas las entidades esperadas para la migración? ¿Llegaron a destino todos los registros del universo definido?	Compleitud
Trazabilidad de los datos migrados	¿Cuáles son los registros migrados? ¿Qué usuario realizó el alta del registro? ¿De qué fuente provino el dato de un registro? ¿Es auditable la base de datos migrada?	Trazabilidad
Compleitud del registro	¿El modelo de datos destino contiene todos los datos esperados?  ¿Se cumplen todos los mapeos y transformaciones definidas?	Compleitud



# Cualidades de los datos migrados - dimensiones de calidad

<i>Cualidad de datos migrados</i>	<i>Preguntas</i>	<i>Dimensión de calidad</i>
Correctitud sintáctica de los datos	¿Tienen los datos el formato esperado? ¿Tienen la precisión definida? ¿Contienen los datos esperados? ¿Hay valores fuera de rango?	Correctitud
Usabilidad de los datos migrados	¿El nuevo sistema funciona con los datos migrados? ¿Los usuarios se sienten seguros con los datos migrados?	Usabilidad
Integridad referencial de las entidades migradas	¿El valor de una columna está incluido en la clave primaria de la entidad referenciada?	Integridad

# Cualidades de los datos migrados - dimensiones de calidad

<i>Cualidad de datos migrados</i>	<i>Preguntas</i>	<i>Dimensión de calidad</i>
Compleitud de los datos	¿Están las columnas/atributos poblados de acuerdo a las reglas de la migración? ¿Tenemos muchos valores nulos?	Compleitud
Consistencia interna de entidades migradas	¿Hay vacíos en los datos? ¿Hay partes faltantes en los numeradores? ¿Hay inconsistencia entre los atributos relacionados de una misma entidad?	Consistencia

# Qué podemos corregir y cómo

- Dimensiones para pensar cuando considerarlas
  - Exactitud semántica: Que los datos reflejen la realidad
  - Frescura: que los datos sean recientes/actualizados

# Conceptos fundamentales de la metodología

## Migración

- Modelo de datos origen
- Modelo de datos destino
- Reglas, mapeo y transformación

**Conocer**

Diagnóstico y limpieza de datos

**Diseñar casos de prueba**

- Elegir las cualidades
- Casos de prueba
  - Por totales
  - Detallados

- Manuales
  - Muestreo Lineal
  - Muestreo estadístico
- Automatizadas

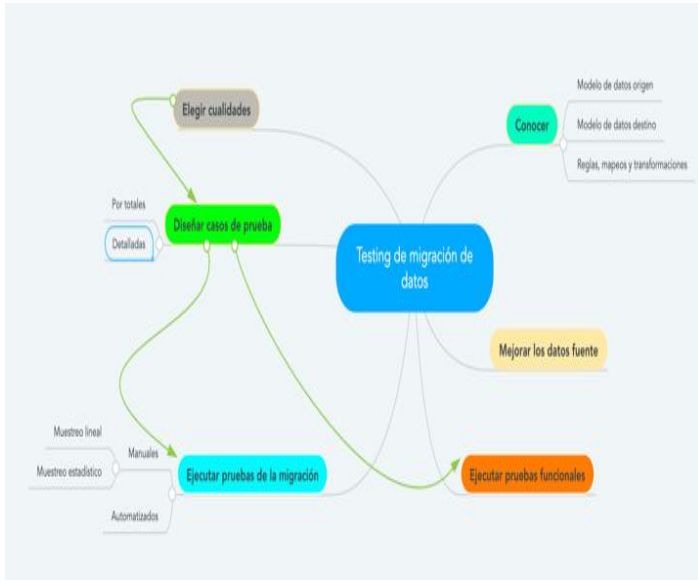
**Ejecutar pruebas de datos migrados**

Testing de Migración

**Ejecutar pruebas funcionales**

- Sistema transaccional
- Reportes
- Procesos por lote
- Interoperabilidad

# Mapa mental-Conceptos que guían a la metodología propuesta



- **Conocer** todos los aspectos relevantes para la migración, sea de los datos fuentes, del proceso de migración y de los datos de destino
- Invertir tiempo en **mejorar los datos fuente**;
- Elegir las **cualidades** a ser verificadas las cuales conducen al **diseño de los casos de prueba**;
- Ejecutar las **pruebas de los datos** migrados y **pruebas funcionales** del software de aplicación en la nueva instalación, considerándola como un todo para la verificación del buen funcionamiento luego de la migración de datos.

# Ejemplo de aplicación 1

- Empresa de microcrédito - Cambio de sistema informático

Cualidad	Casos de prueba
Compleitud de los datos	11
Compleitud de registro	13
Consistencia interna de entidades migradas	8
Correctitud sintáctica de los datos	5
Integridad referencial de las entidades migradas	10
Trazabilidad de los datos migrados	6
Unicidad de entidades migradas	1
Universo de migración	5
<b>Total</b>	<b>59</b>

Cantidad de casos de prueba por cualidades

# Ejemplo de aplicación 2

- Empresa financiera- Fusión de empresas

- Corrección de datos

- Antes de la migración: 43%

Se utilizaron los procedimientos de mantenimientos habituales de información:  
Muy costosa y lenta. Requiere de llamados a los clientes para confirmar y corregir los Datos, análisis exhaustiva de la información

Dimensiones: unicidad, consistencia y correctitud semántica

- Durante la migración: 57 %

Corrección directa sobre los datos con criterios establecidos con los usuarios.  
Se utiliza base de datos de referencia: ejemplo Identificación Civil para las cédulas  
Corrección en ETL (Extract-Transform-Load)

Dimensiones: Completitud de los datos, correctitud sintáctica, integridad

*Causa: Restricciones de nulos, cambios en el modelo  
o nuevas reglas de negocio*

# Ejemplo de aplicación 2

- Después de la migración
  - Se continuó corrigiendo los datos que por priorización, restricciones de tiempo/recursos no se logró corregir
  - Focalización en procesos y datos críticos:
    - Generación de reportes regulatorios
    - Análisis y corrección de los valores introducidos por omisión
    - Información para la toma de decisiones (Tableros de control, Business Analytics, etc.)
- ¿Qué dejamos?
  - Reporte de limpieza de datos
  - Trazabilidad: datos que permitan a una auditoría conocer los datos migrados (fuente, fecha de actualización)
  - Base de datos limpia para la migración y como entrada al nuevo aplicativo
- ¿Cómo verificamos el resultado?
  - Testeamos los datos en la base de datos migrada
  - Verificamos el nuevo aplicativo funcionando con los datos migrados
  - Volvemos a medir la calidad de los datos: mejora continua



# Pensemos... el rol de la calidad de datos en:

- Medicina
- Inteligencia artificial
- Seguridad Social
- Logística: Despacho de valijas
- Ejercicio: Nombrar otros ejemplos

Índice de  
Precios al  
consumo

Chatbot

Instituciones  
financieras

Indicadores  
económicos

Factura  
electrónica

# Ejercicio

- Se migrará la base de Datos de un Sistema de Oracle 11c a Oracle 19c (también se migrará el servidor JBoss). La base contiene datos sensibles de una importante entidad bancaria siendo de vital importancia la migración de los datos. Datos personales de clientes y cuentas asociadas.
- Suponga que piden que planifique las pruebas de la migración de la base de datos de Clientes y Cuentas.
- Qué cualidades de los datos probaría. Justifique
- Considerar y comentar la cobertura de las pruebas planificadas.

# Herramientas Calidad de datos



Algunas herramientas libre:  
OpenRefine  
<https://openrefine.org/>

Fuente: Magic Quadrant for Data Quality Solutions. Gartner. Ankush Jain, Melody Chien. 1 November 2022

# Conclusiones

- La importancia del testing de datos en migraciones
- Tres proyectos en paralelo que requieren (Diagnóstico y limpieza de datos, Migración, Testing de migración):
  - planificación, articulación y sincronización
  - equipo multidisciplinario
- Cualidades de datos migrados y dimensiones de calidad de datos
- Probar el proceso de migración y el funcionamiento del sistema con los datos migrados
- La calidad de datos en migraciones disminuye los riesgos asociados a problemas de calidad de datos durante y posterior al proceso de migración
- Mejora uno de los principales activos de la organización que son los datos
- La calidad de datos es un proceso continuo: Seguir trabajando en calidad de datos luego de las migraciones

# Referencias

- Metodología de testing para migración de datos. E. Martirena. M. Wodzislowski. Conisoft 2021.
- Batini, C.; Scannapieco, M.: "Data and Information Quality: Dimensions, Principles and Techniques", Springer, 2016. ISBN 978-3-319-24104-3.
- [DAMA](#) International: "Data Management Body of Knowledge", 2nd Edition (DAMA, DMBOK2). Technics Publications, Basking Ridge, New Jersey, 2017. ISBN 978-1-634-62883-9.
- Magic Quadrant for Data Quality Solutions. Gartner. Ankush Jain, Melody Chien. 1 November 2022

# Testing de migración de base de datos

Muchas gracias