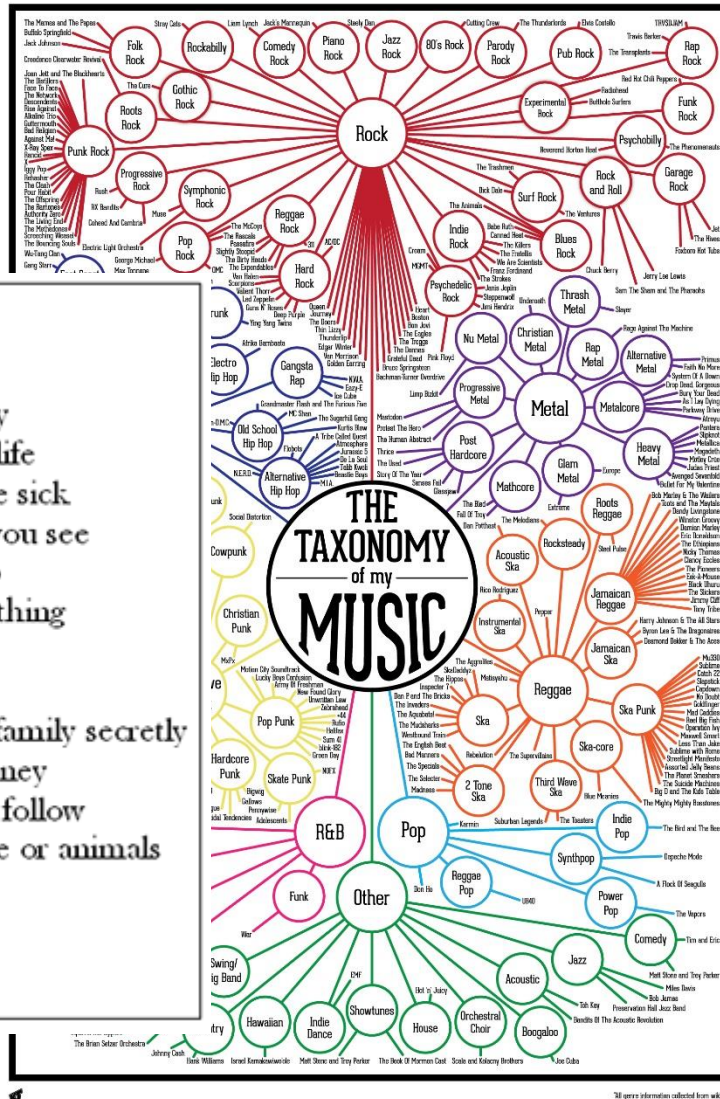


# Modelos y Persistencia de Datos

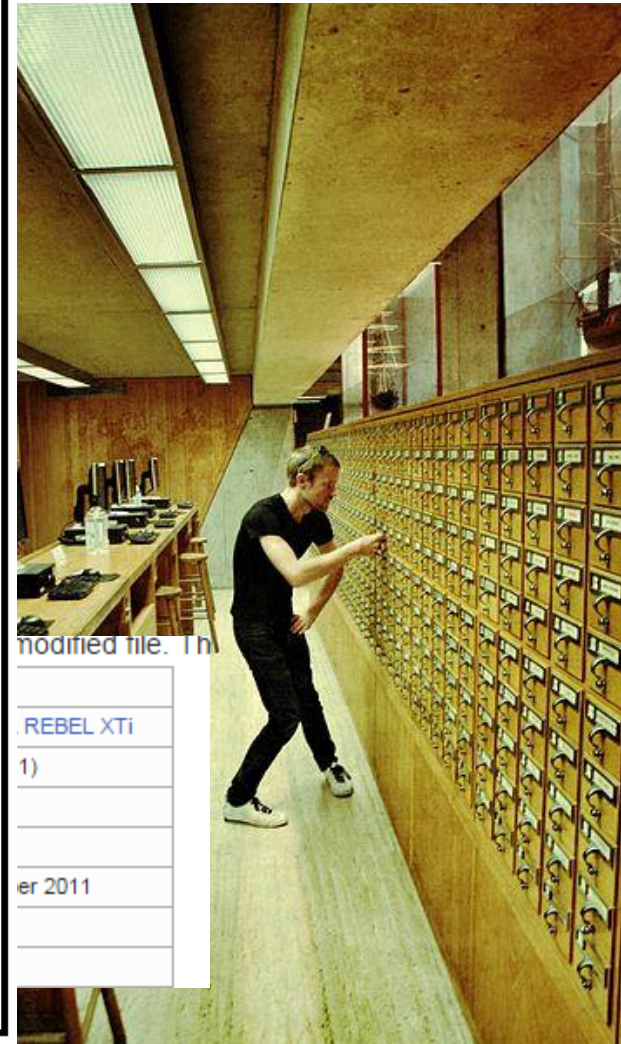
Alejandro Sierra

# Ejemplos de Meta Datos



## Glossary

<b>Begin</b>	Start
<b>Cry</b>	Speak loudly
<b>Death</b>	The end of life
<b>Doctor</b>	He helps the sick
<b>Dream</b>	Something you see when asleep
<b>Fire</b>	Hot orange thing
<b>Pearl</b>	Lulu
<b>Return</b>	Go back
<b>Run Away</b>	Leave your family secretly
<b>Sell</b>	Give for money
<b>Trackers</b>	People who follow other people or animals
<b>Trouble</b>	Problems
<b>Try</b>	Have a go



# Datos no estructurados

- Los meta datos sirven de puente entre los datos no estructurados y los datos estructurados

Camera manufacturer	Canon
Camera model	Canon EOS DIGITAL REBEL XTi
Exposure time	1/10 sec (0.1)
F-number	f/5.6
ISO speed rating	400
Date and time of data generation	11:43, 1 September 2011
Lens focal length	18 mm
<a href="#">Show extended details</a>	

# Administración de Meta Datos

- Modelo de Madurez

## 5. Meta datos para habilitar la Innovación en el Negocio

- Predecir, analizar nuevas ideas

## 4. Meta Datos para Administrar el Negocio

- Optimización del Negocio

## 3. Meta Datos de Negocio y Técnicos alineados

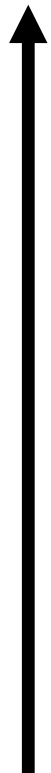
- Relacionar términos de Negocio con objetos de TI

## 2. Describir el Negocio con Meta datos (Meta Datos de Negocio)

- Glosario, reglas de negocio

## 1. Describir la tecnología con Meta Datos (Meta Datos Técnicos)

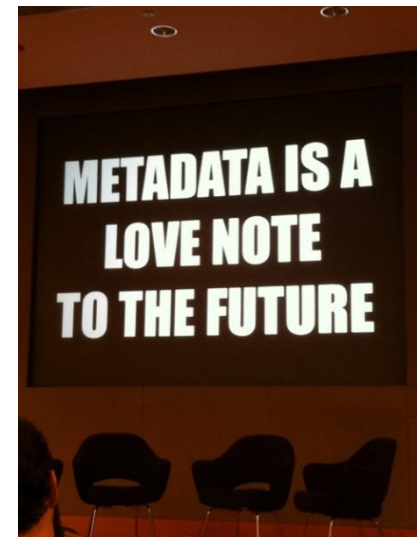
- Catálogos de DBMS



# Fuentes de Meta Datos

- Sistemas Operacionales
- Middleware (Por ejemplo EII)
- Sistemas de BD
- Aplicaciones
- DW
- Herramientas de Reportes
- Herramientas de Desarrollo
- Documentación de SW
- Políticas y procedimientos
- Procesos
- GUIs
- Personas

Los metadatos no son creados con el fin de ser consumidos



# Preguntas básicas que debemos responder

1. ¿Qué datos residen en nuestros archivos/BDs?
2. ¿Qué significa cada elemento de datos en términos de negocio y en términos técnicos?
3. ¿Dónde está ubicado?
  - ¿Está replicado en otras ubicaciones?
4. ¿Cómo llegó a la ubicación en la que está?
5. ¿Cómo se puede acceder?

Otras preguntas:

- Dueño, calidad, usuarios, reglas, otros nombres,...

# Componentes de una estrategia de Admon de Meta Datos

- Repositorio en el que almacenamos los meta datos.
  - Meta meta datos. Diseño del repositorio.
- Administrador de meta datos.
  - Persona. Cualidades técnicas. Mantenimiento del repositorio como un servicio.
- Políticas
  - Obligan a otras entidades a ciertas reglas sobre los meta datos. Nombres de objetos, formatos de meta datos, ...
- Procedimientos
  - Procedimientos para cumplir con las políticas

# Categorías

Negocio

Técnica

Uso

Procesos



# Meta datos de Negocio

- Proviene de los usuarios o de las personas de negocio.
- Capa semántica
- Términos de negocio
  - Cliente Preferencial vs ~~cliente~~Preferencial
  - Clases (Tipos)
    - Cantidad, fecha, porcentaje, texto
- Glosario de Negocio

# Glosario de Negocio

- Comunicación clara a través de la organización
- Mejorar la colaboración Negocio-IT
- Bases fuertes para cumplir con:
  - Regulaciones
  - Gobierno de Información
  - Responsable de Información

# Glosario de Negocio

The screenshot shows a web application titled "Sales Glossary" with tabs for "Glossary", "Glossary Editor", and "Attribute Manager". The breadcrumb trail is "BISAMPLE (Relational Model) > BISAMPLE > SAMP\_REVENUE\_F >". A search bar contains "Search SAMP\_REVE". The main content area displays the details for "SAMP\_REVENUE\_F", which is a "Custom Attribute". The details are as follows:

Name	SAMP_REVENUE_F
Definition	Revenue Fact Sample
Steward	Scott
Type	Entity
Abbreviation	
Alternative Abbreviation	
Data Type	

Below the details is a "Documentation" section with a red box around the "Relationship" link. The text "No documentation. Click Edit button to edit." is displayed. The "Edit" button is visible. Below the "Relationship" link, there are sections for "Synonyms:", "See Also: SAMP\_ADDRESSES\_D, SAMP\_TIME\_MTH\_D", "More General:", "More Specific:", and "Contains: COST\_FIXED, COST\_VARIABLE, DISCNT\_RATE, DISCNT\_VALUE, ORDER\_PAY\_DT, ORDER\_NUMBER".

**SAMP\_REVENUE\_F** Custom Attribute

Name: SAMP\_REVENUE\_F  
Definition: Revenue Fact Sample  
Steward: Scott  
Type: Entity

Documentation

No documentation. Click Edit button to edit.

Relationship

+ Synonyms:  
+ See Also: SAMP\_ADDRESSES\_D, SAMP\_TIME\_MTH\_D  
+ More General:  
+ More Specific:  
+ Contains: COST\_FIXED, COST\_VARIABLE, DISCNT\_RATE, DISCNT\_VALUE, ORDER\_PAY\_DT, ORDER\_NUMBER

# Glosario de Negocio

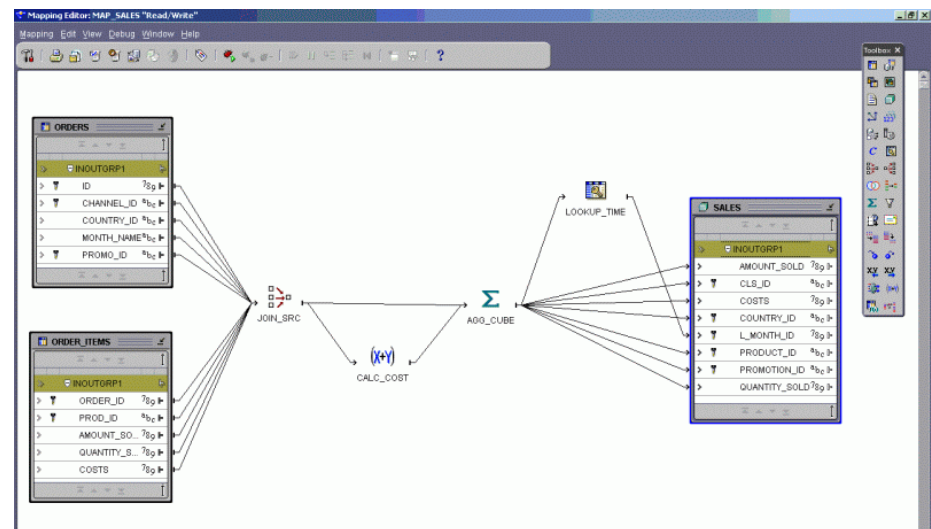
- Usuarios
  - De negocio: Para entender terminología y validar fuentes de datos
  - Custodios (Data Stewards): Apoyar el ciclo de vida de los datos
  - Técnicos: Tomar decisiones de diseño/arquitectura

# MD Técnicos

- Documentación en términos técnicos.
- Catálogos de BD, herramientas CASE, IDEs.
- Abreviaciones, guiones, ....
  - Cliente preferencial -> CLI\_PREF.
- En un escenario ideal los nombres se rigen por una política clara.

# MD de Procesos

- Otro tipo de md técnica.
- Describe acciones tomadas por programas.
  - ETL, procedimientos de BD, software.
- Describe transformaciones.
  - Por ejemplo formato de origen y formato de destino.



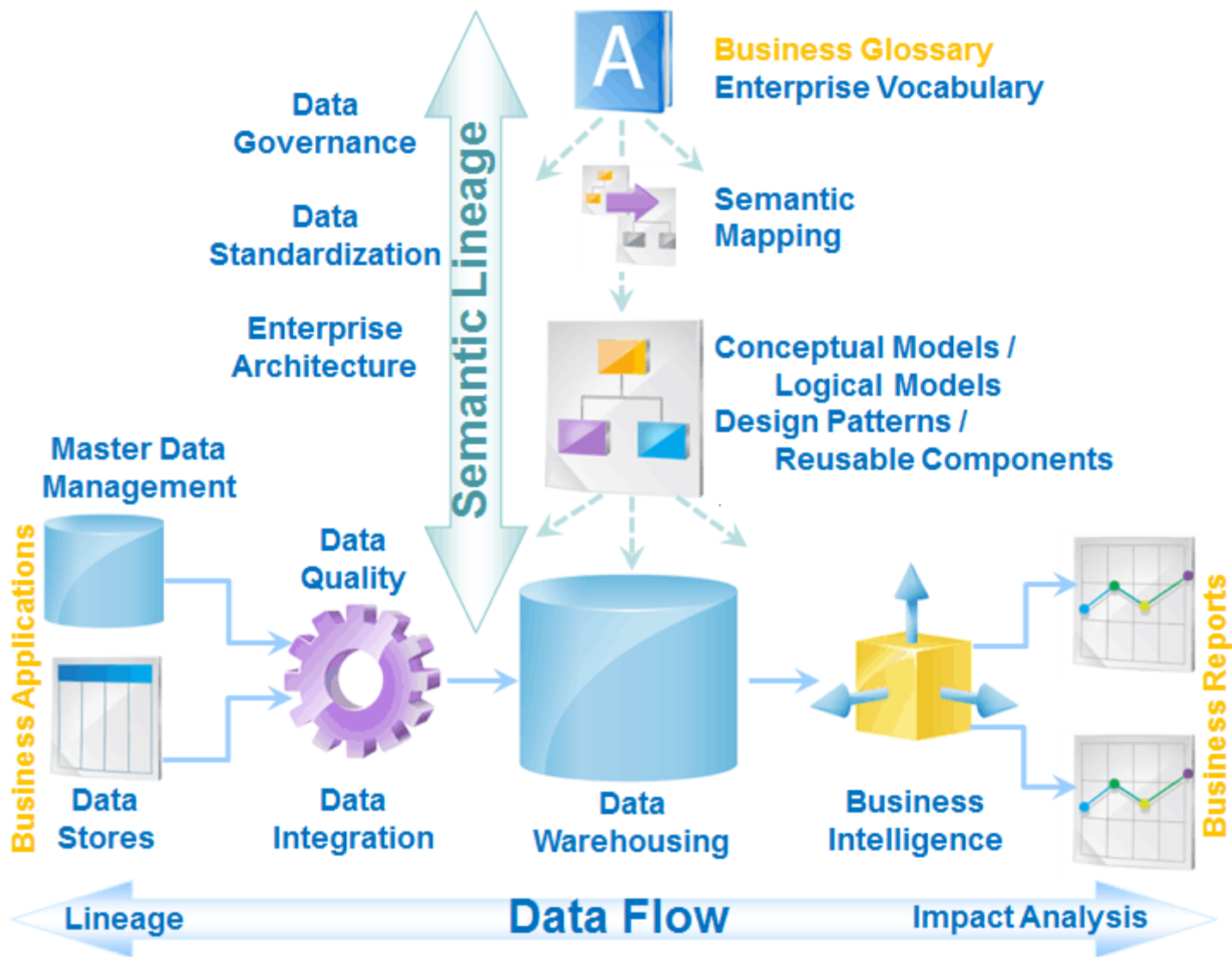
# MD de Uso

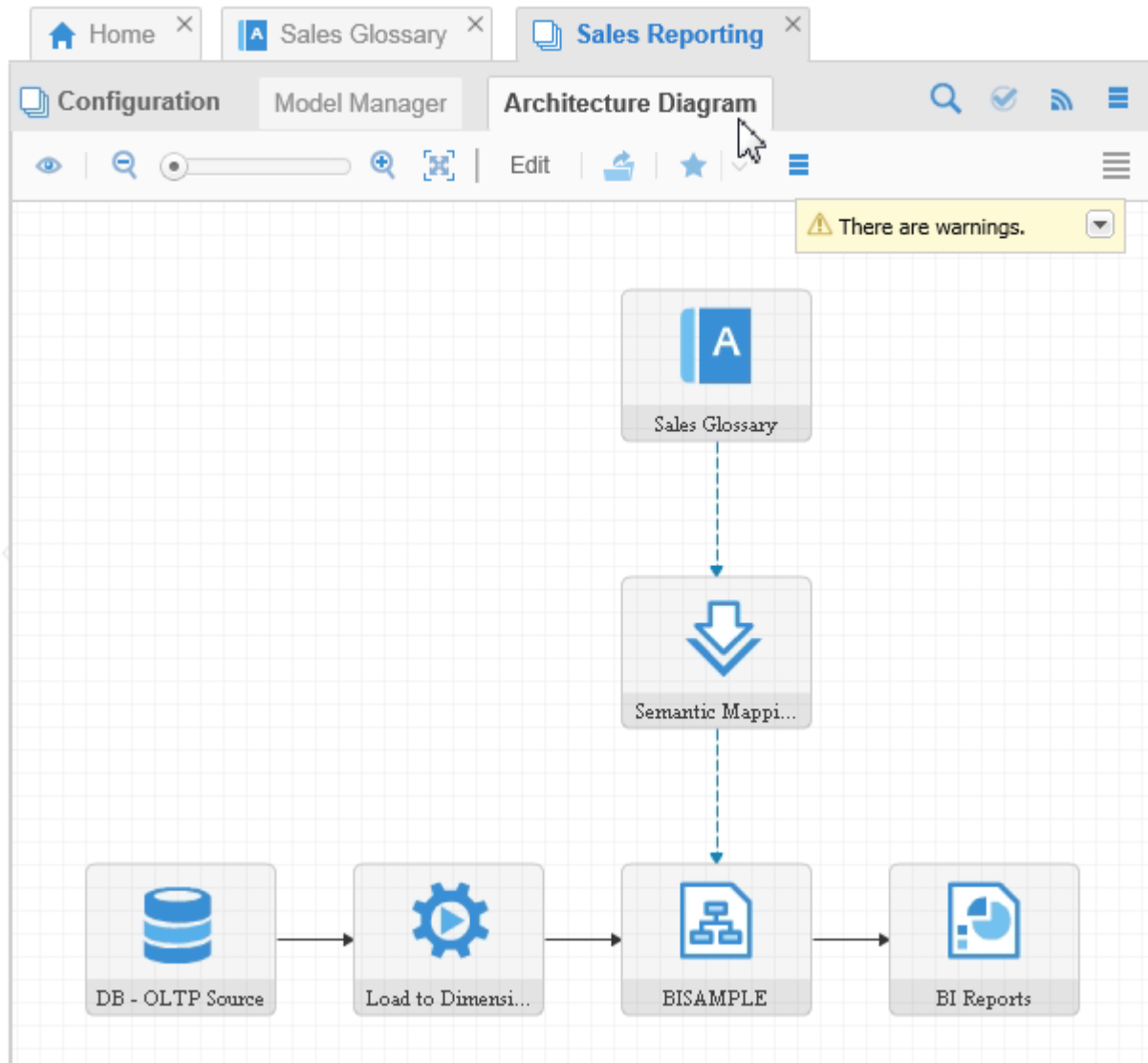
- Patrones y frecuencias de acceso
  - Diario, mensual, durante el día.
- Monitoreo de uso.
- ¿Con qué herramientas se accede?
- Queries vs Tablas
- Joins

# Linaje de punta a punta

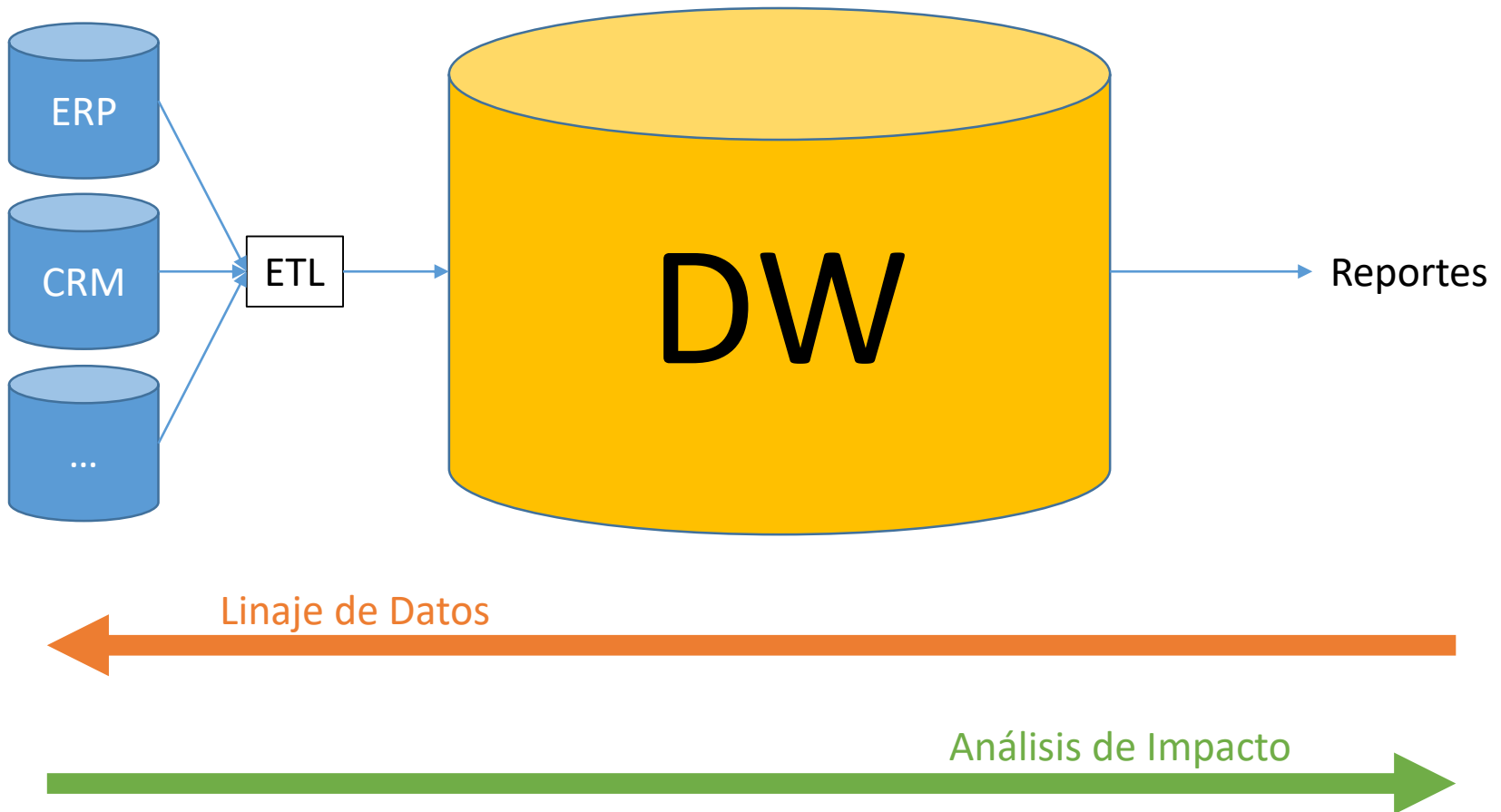
- Linaje Semántico
  - ¿Cómo está representado un objeto de negocio en uno o más objetos físicos?
  - Glosario <-> Objetos Físicos
- Movimiento de Datos
  - Conocer desde el origen hasta las diferentes copias o productos de un Dato.
  - ETL, Replicación.
  - Análisis de Impacto
    - Origen -> Destino
    - Si modifico esta tabla, ¿Qué procesos se afectan?
  - Linaje de Datos
    - Destino -> Orígenes
    - ¿De dónde viene este dato que me muestra un reporte?







# Análisis de Impacto vs Linaje de Datos



# Repositorio de MD

## Centralizado

- Sin redundancia
- Limpieza previa a consulta
- Mayor esfuerzo

## Federado

- No hay movimiento de datos.
- Menos esfuerzo
- Depende de la facilidad de consultar meta datos en tiempo de consulta.
- Limitado a lectura

Requiere procesos que mantengan actualizados los metadatos

# Análisis y Diseño

- ¿Cuáles preguntas vamos a responder?
- Meta-meta datos
  - Cardinalidad
  - Tipos de meta datos
  - Relaciones entre Tipos de MD

**Figure 4.3. Example of meta metadata**

Name: Entity  
Definition: A fully normalized business object.  
Relationship: Related to one or many tables  
Read by all, create/update/delete by the DA  
Ownership: Data Administrator (DA)  
Origin: ERwin  
Physical Location: MDRSYSENT table  
Cleanliness: 2 percent missing data  
Applicability: Since November 1, 2003

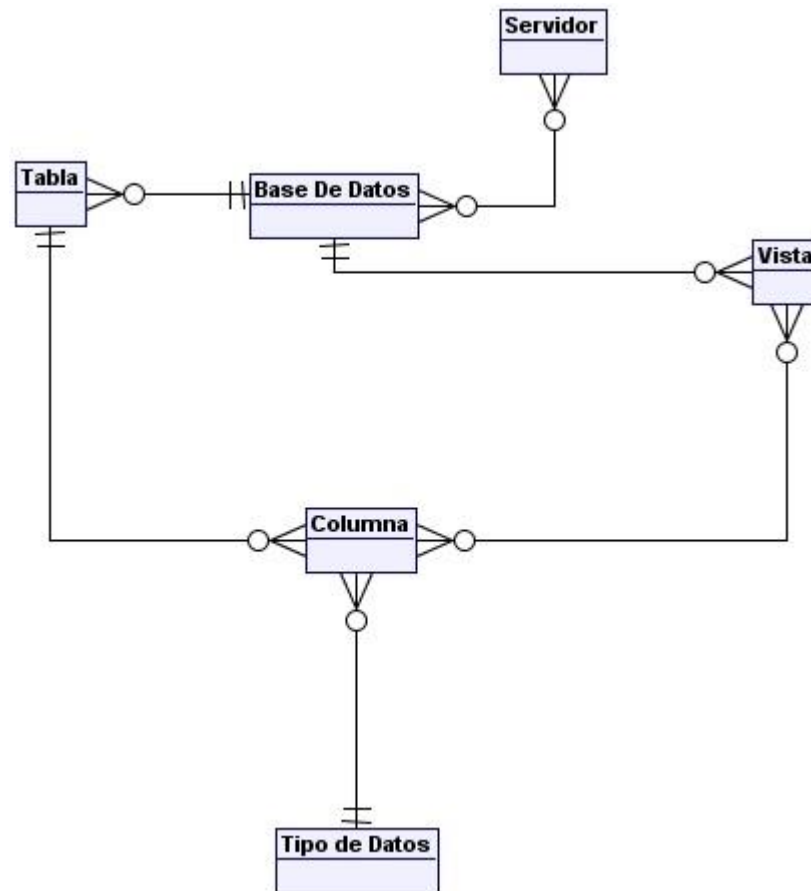
# Requerimientos no funcionales de un Repo. De MD

- Alta prioridad
  - Facilidad de adaptarse a nuevas fuentes
  - Seguridad
    - Describe casi todo el negocio
- Media
  - Desempeño
  - Disponibilidad continua
  - Mantenibilidad
- Bajos
  - Capacidad de volumen
  - Escalabilidad

# Ejemplo en tablero para meta metadatos técnicos

- Diagrama ER Conceptual con
  - Servidor
  - Base de Datos
  - Tabla
  - Columna
  - Tipo de Datos
  - Vista

# Boceto de Clase





# Ejercicio

- Ampliar el modelo para que incluya:
  - Metadatos de negocio
  - ETLs
    - Caja negra
    - Extracción de columnas
    - Carga en columnas

# Referencias

- The Art of Enterprise Information Architecture
  - A Systems-Based Approach for Unlocking Business Insight
  - Mario Godinez, Eberhard Hechler, Klaus Koenig, Steve Lockwood, Martin Oberhofer, Michael Schroeck
  - IBM Press
  - 2010
- Business metadata capturing enterprise knowledge
  - Inmon, William H.
  - Elsevier/Morgan Kaufmann
  - 2008
- DAMA-DMBOK: Data Management Body of Knowledge
  - 2nd Edition.
  - dama.org
  - 2017
- Oracle Enterprise Metadata Management
  - <https://www.oracle.com/middleware/technologies/enterprise-metadata-management.html>