



Big Data: ¿Big Governance? El Gobierno de Datos en entornos Hadoop

WEBINAR

Marzo de 2017

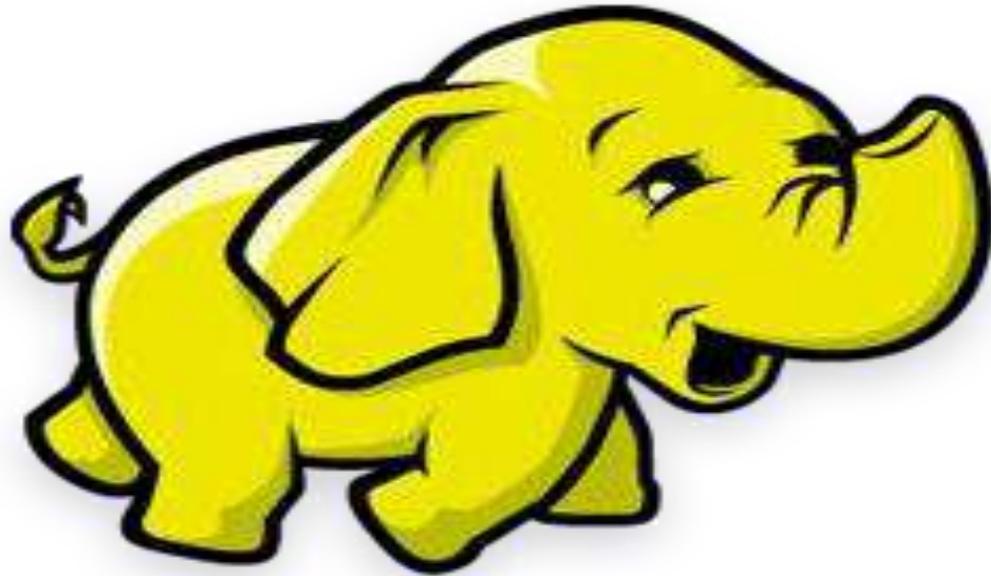


PowerData

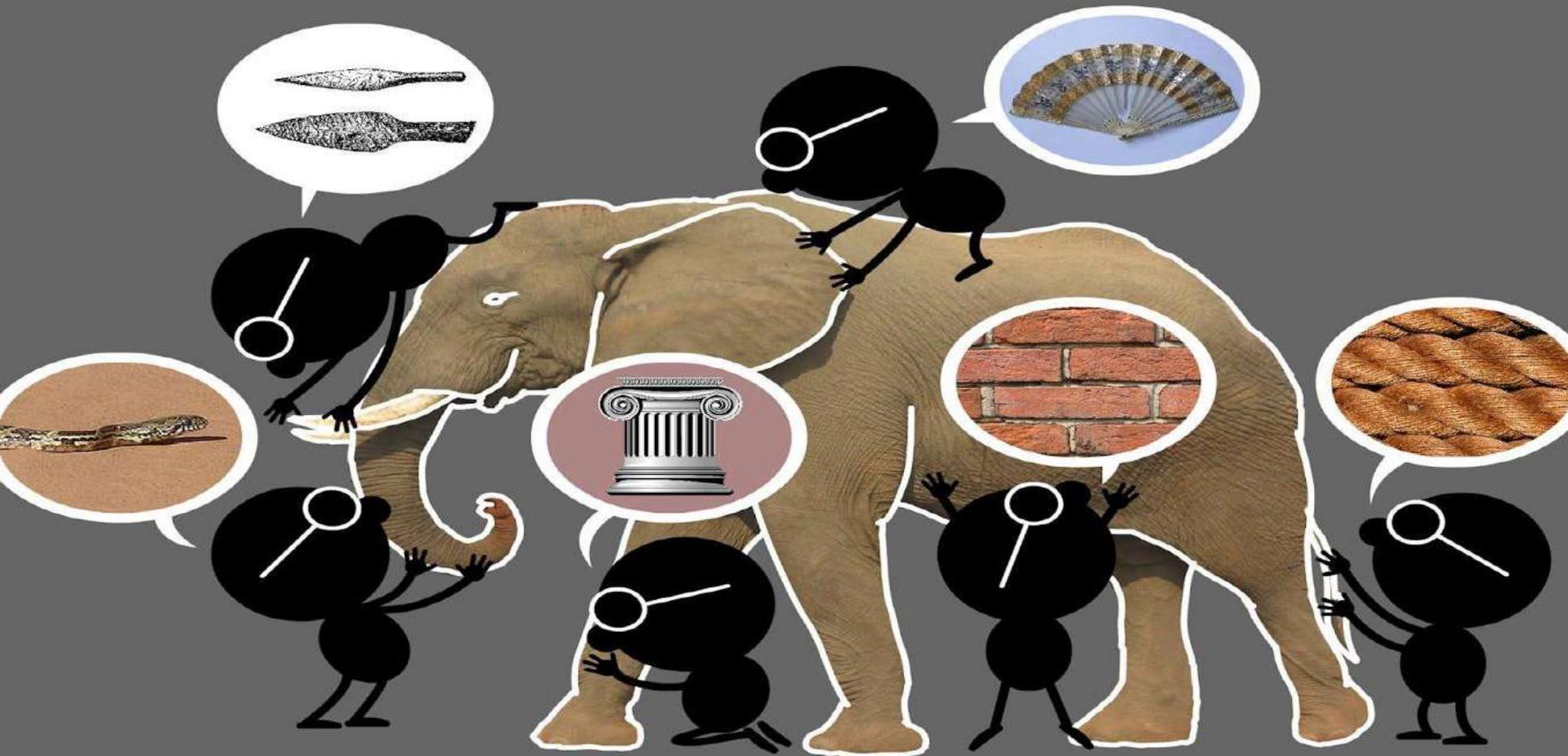
Especialistas en Gestión de Datos

DEFINICIONES Y CONTEXTO

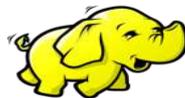
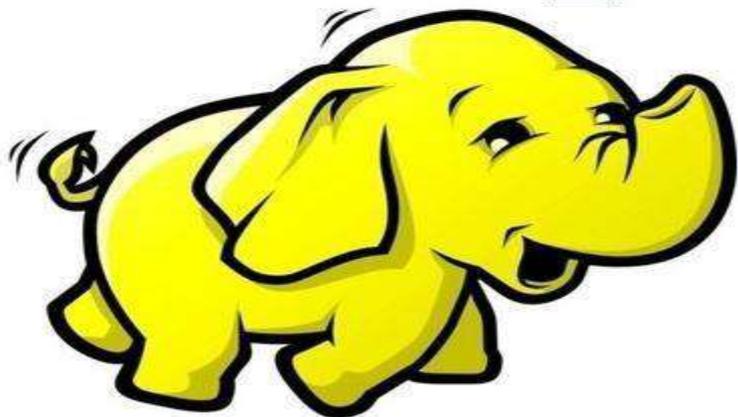




LOS 6 CIEGOS Y EL ELEFANTE



hadoop



hadoop <==> analytics

**GOBIERNO BIG
DATA**

**SEGURIDAD -
MONITOREO -
CONTROL DE
ACCESO**

CAPA DE PRESENTACIÓN

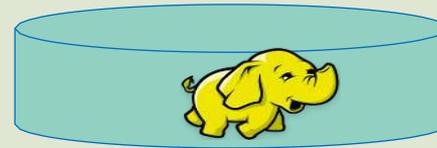
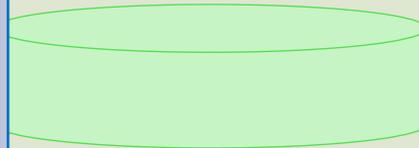
CAPA ANALÍTICA

CALIDAD - GESTIÓN DE DATOS

INTEGRACIÓN - METADATA

INGESTA - ETL

ALMACENAMIENTO



AUTOSERV

Conjunto de procedimientos para asegurar la correcta gestión de los datos en una organización

Cumplimiento Regulatorio

Business & IT

Datos sólidos y creíbles

Seguridad y Privacidad

Responsabilidad ¿quién?

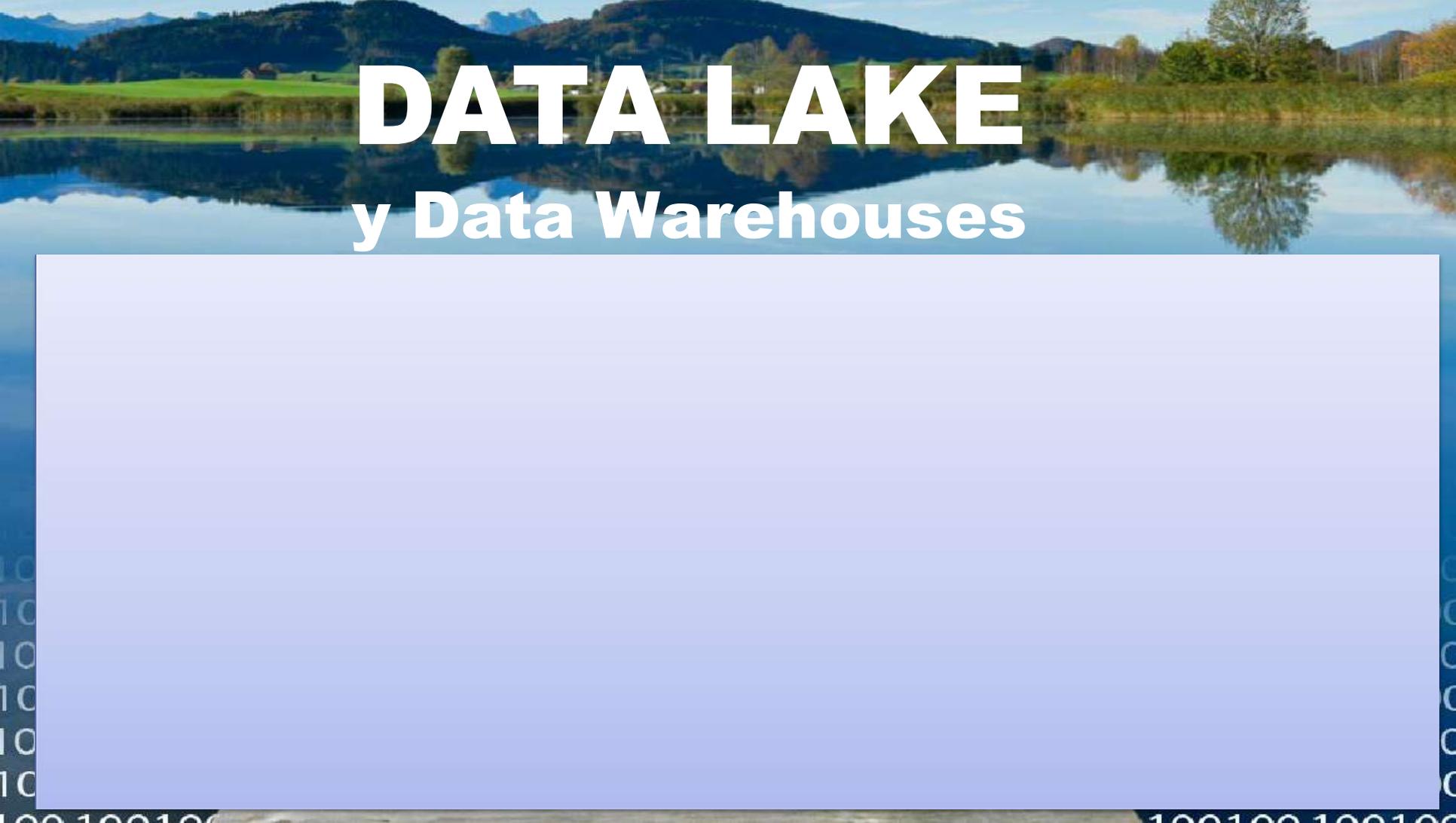
Calidad de Datos

Datos como activos

Predictibilidad

GOBIERNO DE HADOOP





DATA LAKE y Data Warehouses

0
10
10
10
10
10
100100100

100100 10010

“Raw”

- Datos desde orígenes
- Sin transformación
- Sin control de calidad
- Restringido con privilegios
- Datos con su metadata
- Almacenamiento en hadoop

Procesados

- Formatos estandarizados
- Controles de calidad simples
- Restringido con privilegios
- Con “keys”
- Almacenamiento en hadoop

Validados

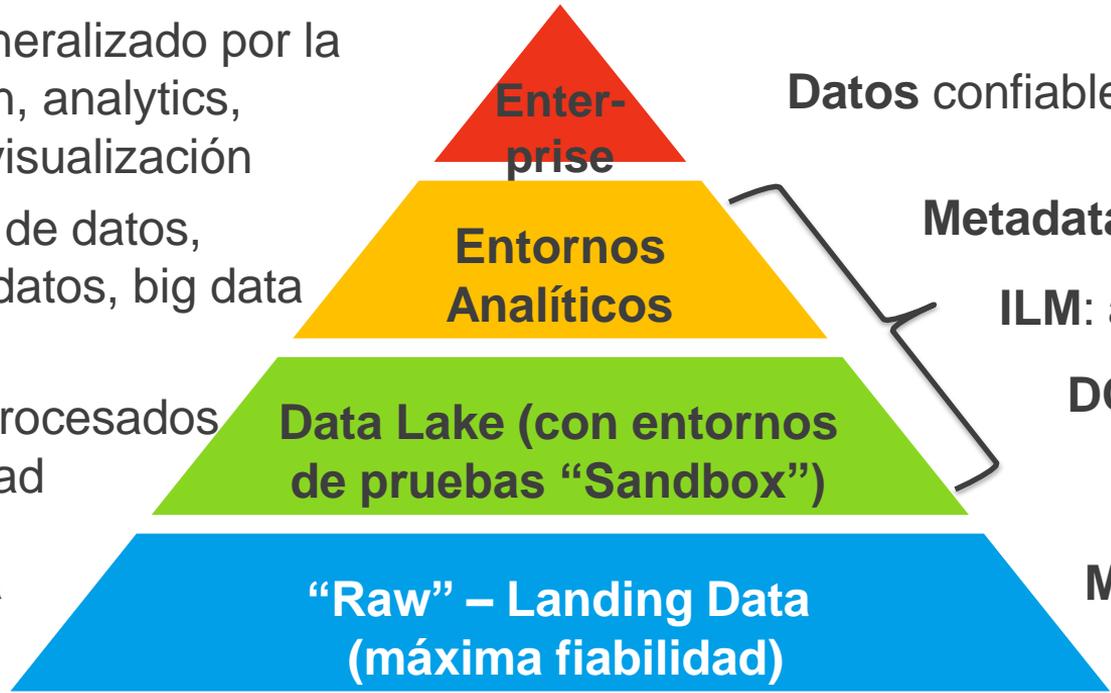
- Uso masivo
- Calidad aplicada
 - Masking aplicado
- Datos pueden ser extraídos de hadoop
- Seguridad y Protección de Datos
- Datos “raw” y procesados pueden eliminarse

Refinados

- Se aplica analítica
- Enriquecimiento de datos
- Calidad de Datos
- Merge & Parse
- Datos pueden ser extraídos de hadoop

PIRÁMIDE DEL GOBIERNO DE DATOS

- 4.- Uso generalizado por la corporación, analytics, reporting, visualización
- 3.- Modelo de datos, completar datos, big data analytics
- 2.- Datos procesados y con calidad
- 1.- Captura de datos



Datos confiables (gobernados)

Metadata: catálogo y glosario

ILM: acceso, retención

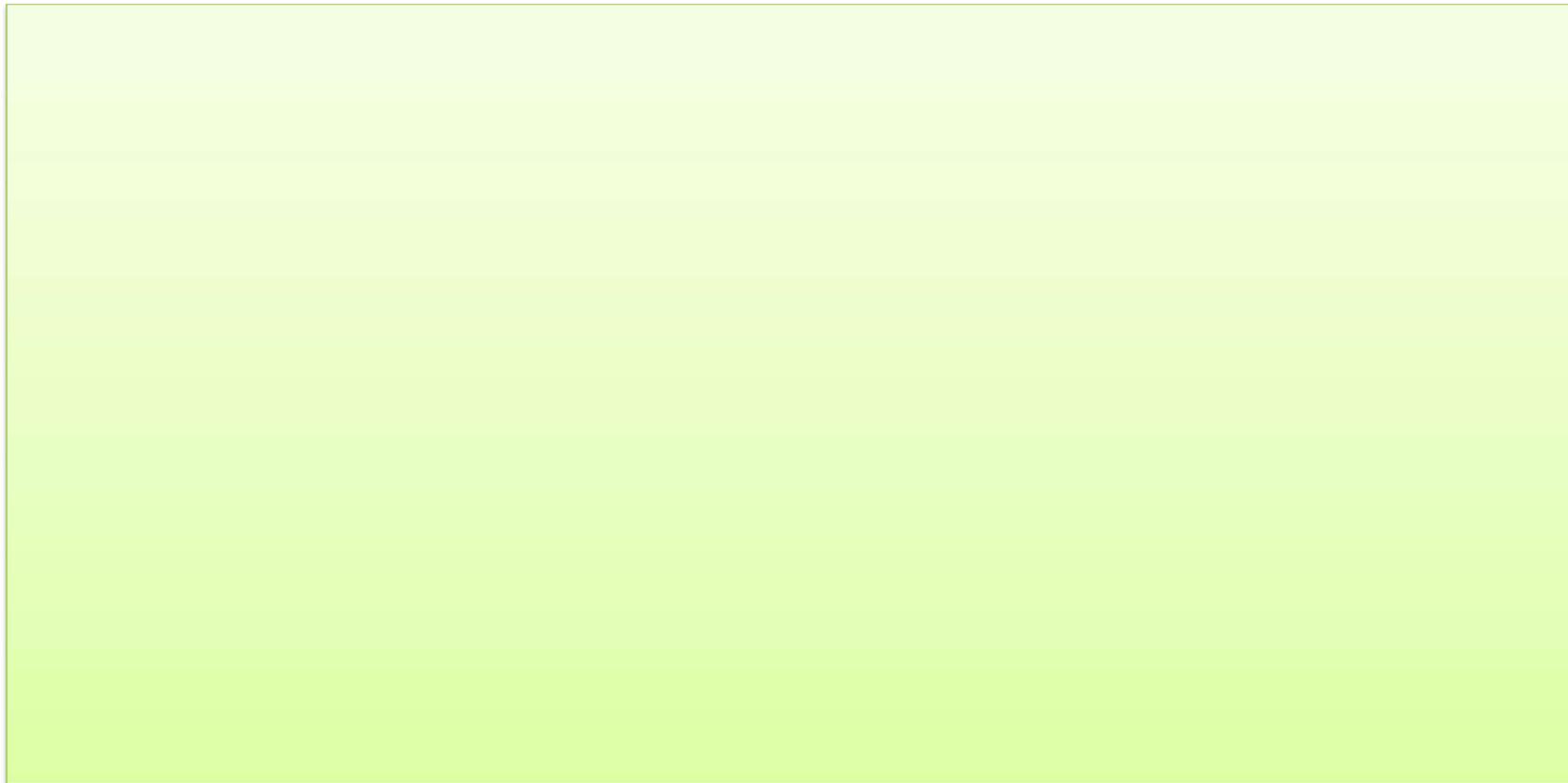
DQ: Validación y reglas

Metadata y linaje

ILM: acceso, retención



**COMPONENTES DEL GOBIERNO
BIG DATA**



MÍNIMOS COMPONENTES





Qué supone una arquitectura hadoop.

Autoservicio de Datos

Carga primero, refina después
¿Existe un modelo de gobierno de datos?



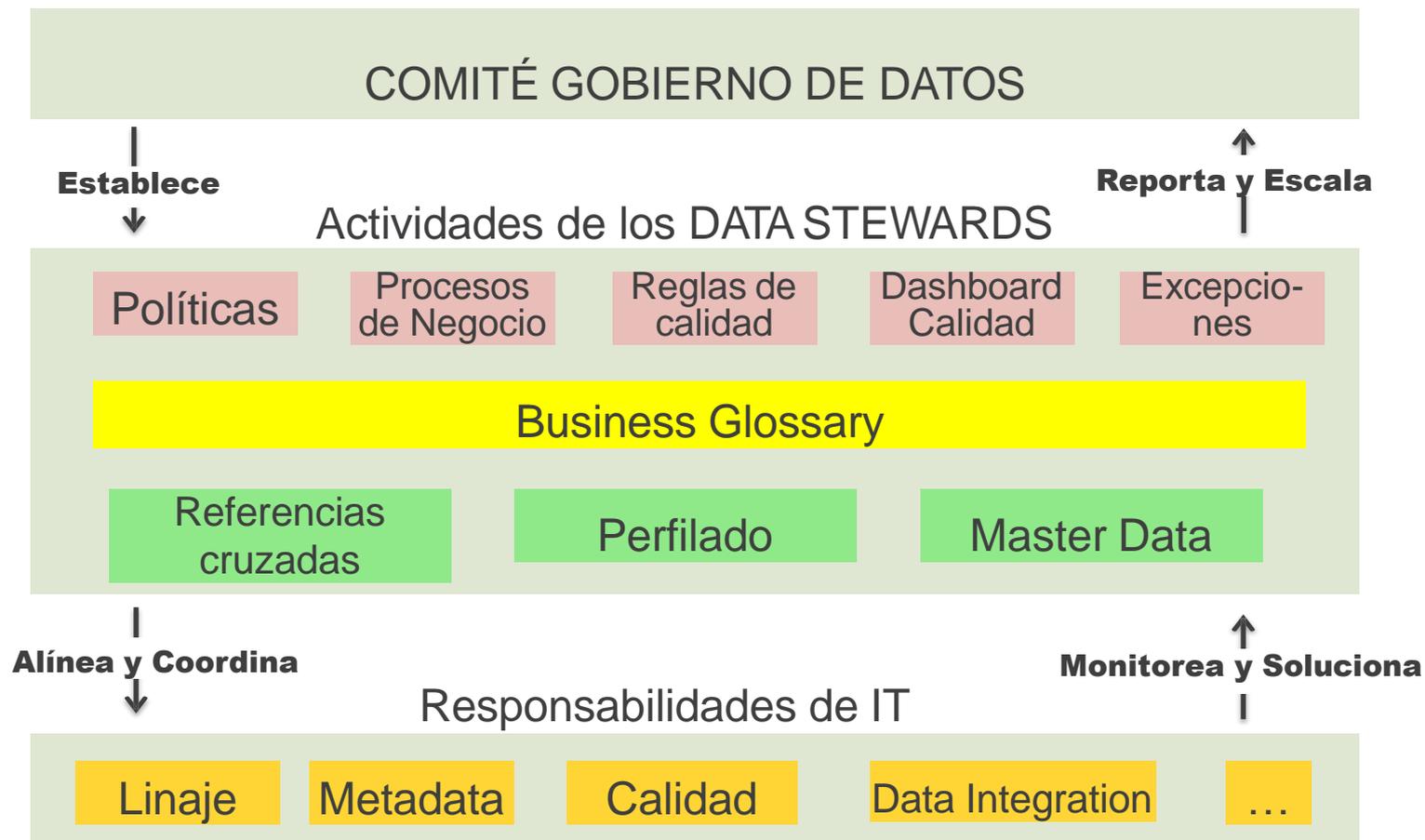
Reglas básicas e irrenunciables para el gobierno: políticas, standards, requisitos regulatorios, etc.



Definiciones, responsabilidades, gestión de excepciones, reglas de calidad, etc.



Identificar herramientas críticas para el proceso.
Fasear la implantación
No duplicar herramientas



Dudas, preguntas



CONSULTA A UN
EXPERTO

•Síguenos en:



•Y contacta con nosotros a través:

webinar@powerdataam.com
marketing@powerdata.es

• ARGENTINA • CHILE • COLOMBIA • ECUADOR • ESPAÑA • MÉXICO • PERÚ • URUGUAY •