

DataPuebla

- *Ciencia de datos e inteligencia artificial para mejorar la toma de decisiones y apoyar la proactividad gubernamental.*
- *Agilizar trámites y servicios públicos.*
- *Formular mejores políticas públicas.*



Gobierno de Puebla
Hacer historia. Hacer futuro.



Secretaría
de Administración

Gobierno de Puebla



Los datos

¿Cómo podemos
conocer al
ciudadano?

¿Cómo podemos
mejorar nuestra
oferta?

¿Cómo podemos
ser mas
competitivos?

Los datos

¿Para qué?

Conocimiento

Eficiencia

Decisiones acertadas

Proactividad

¿Cómo?

Encuestas

Solicitudes

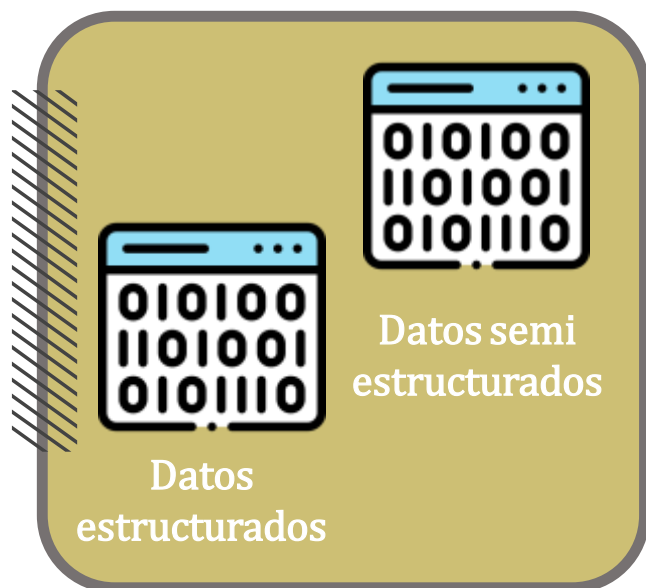
Censos

Bases de datos

Diversos mecanismos



El Problema



Bases de datos aisladas



Datos no estructurados

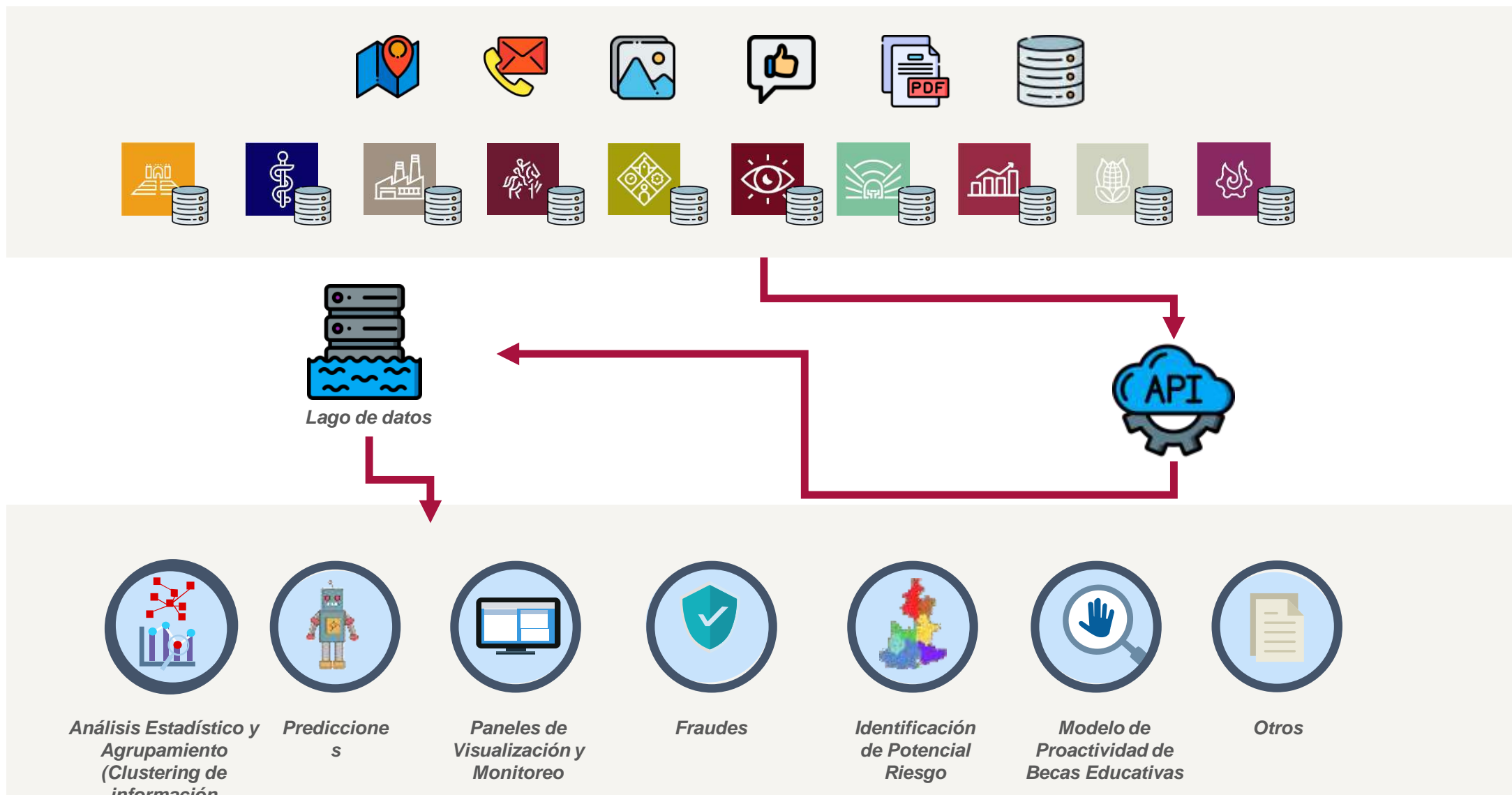


Bases de datos relacionadas
Datos estructurados

Múltiple información pero no conectada

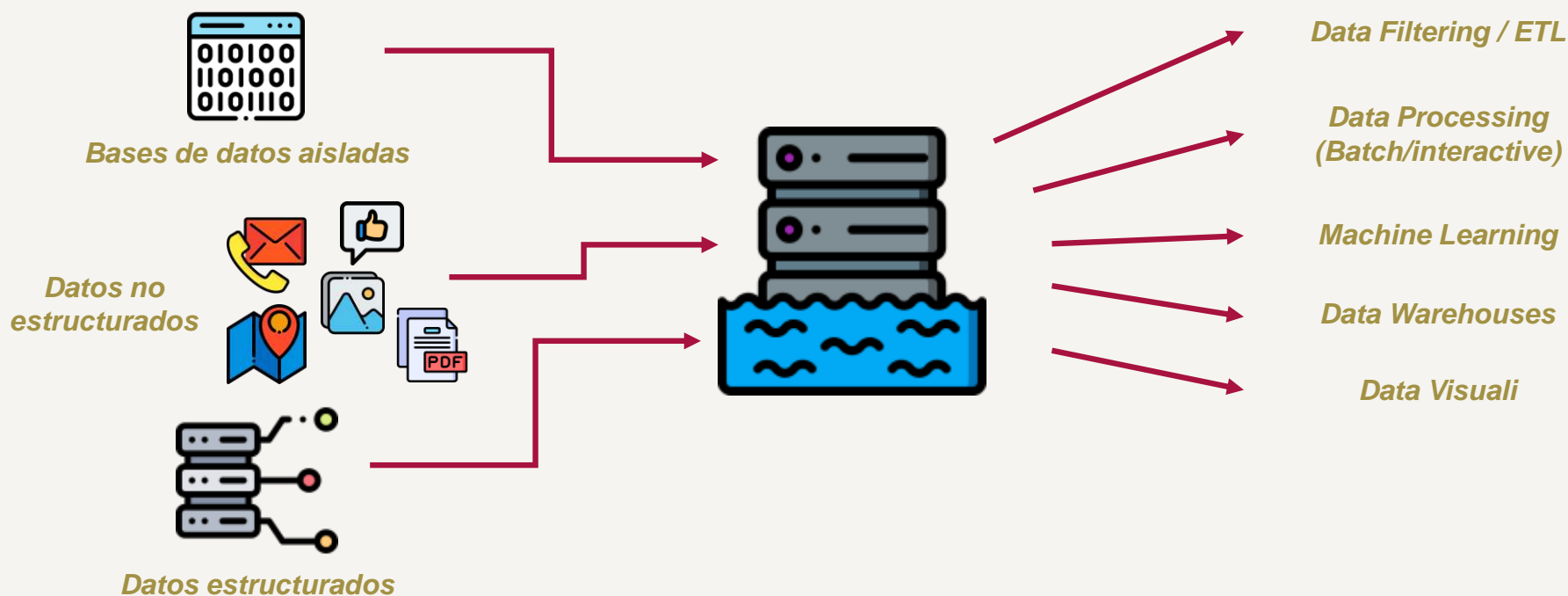


A partir de la información existente y de la construcción del lago de datos



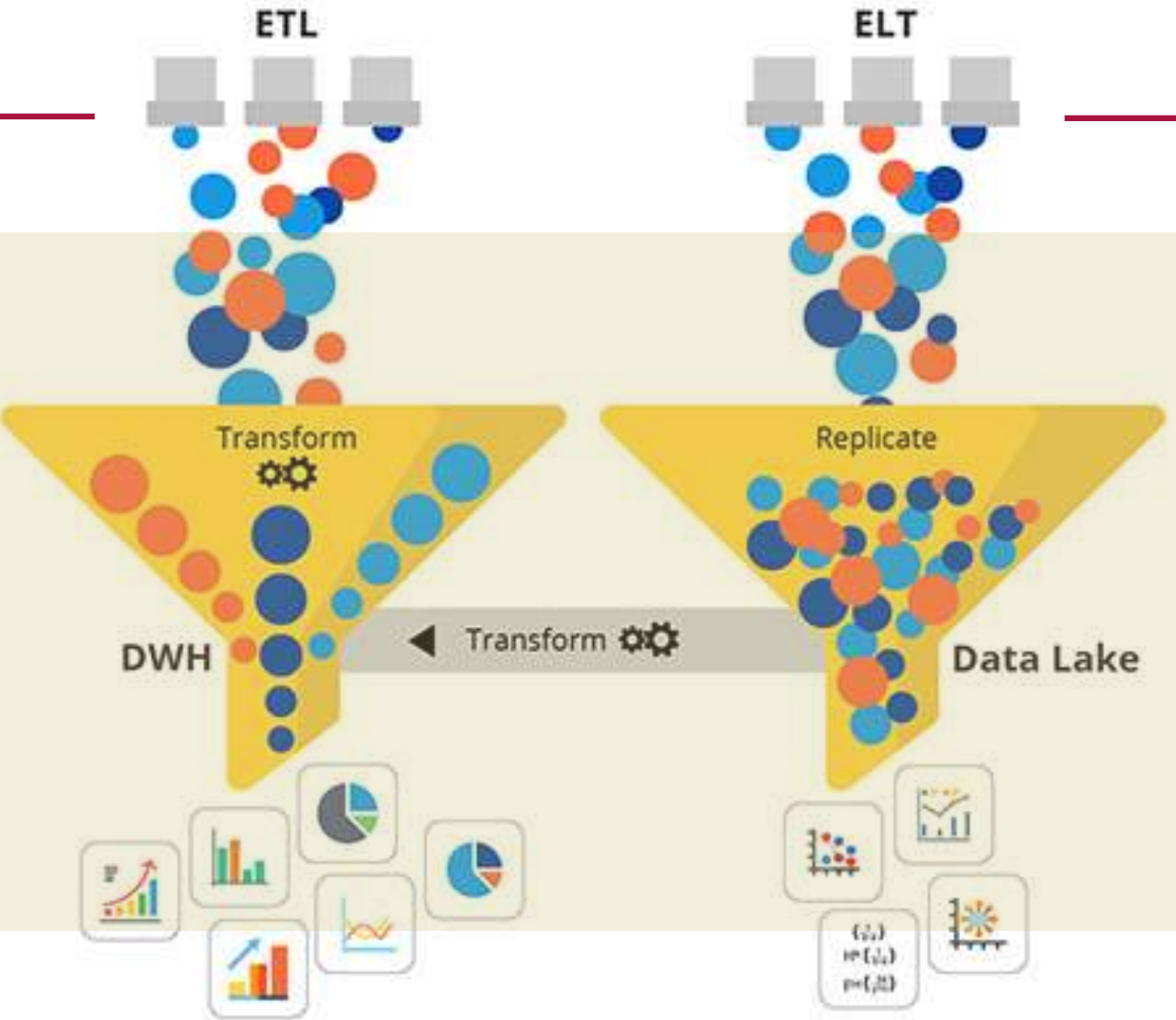
¿Qué es un lago de datos?

Un repositorio diseñado para almacenar grandes cantidades de información en su forma nativa, los datos pueden ser estructurados, semi estructurados o no estructurados y se incluyen tablas, textos, registros del sistema y más.



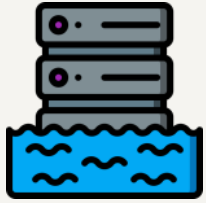
¿Porqué?

- Estadística Descriptiva
- Extraer
- Transformar
- Cargar



- Estadística Descriptiva y Predictiva
- Extraer
- Cargar
- Transformar

¿Cuál es la estrategia?



Beneficios del lago de datos

Democratizar los datos

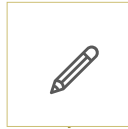
Calidad de los datos

Esquema flexible

Múltiples formatos de datos

Analítica avanzada

Escalabilidad



Definir las ventajas



Identificación de los datos



Identificación de las fuentes de datos



Uso de metodologías apropiadas para explotación de los datos



Optimización de la infraestructura



Indicadores iniciales



Cadena de valor del conocimiento



Componentes DataPuebla

1ra etapa

1.

Arquitectura
del lago de
datos

2.

Infraestructura
del lago de
datos

3.

Analítica
avanzada
(Machine
Learning)

4.

Indicadores
iniciales

5.

Capacitación

6.

Soporte
Técnico



Ejemplos

Proactividad en programas de gobierno

Becas educativas
Financiamientos a MiPymes

Escuchar que opinan los turistas del Estado

Observatorio turístico a partir de Twitter, Airbnb, Tripadvisor, etc.

Competencias laborales requeridas

Competencias demandadas para un mejor nivel socio-económico.

Diversificación económica inteligente

Identificar actividades y regiones óptimas para implementación de programas que fomenten la inclusión en cualquiera de sus formas

¿Cómo puedo participar?

*Generando datos y
compartiéndolos*

*Presentando
hipótesis*

*¿Cómo puede ayudar la ciencia
de datos a tu dependencia?*



¿Qué modelo de trabajo se va a utilizar?

El modelo de trabajo se conoce como **Laboratorios**

Los proyectos de **ciencia de datos** tienen una duración de 6 o 7 semanas dependiendo de la complejidad del despliegue.

Por cada área de negocio y/o servicio se deberán identificar los **indicadores** que serán resultado de los proyectos de ciencia de datos.

Se usarán **herramientas** para conocer las necesidades de las áreas, la definición de los sub proyectos y resultados esperados (como **Crisp DM**).

El equipo multidisciplinario de cada laboratorio se conformará así:

1 **Product Owner (Analytics Leader)**

1 **Científico de Datos (Business Analyst)**

1 **Científico de Datos (Data Scientist)**

1 **Ingeniero de Datos (Data Engineer)**



DataPuebla



Gobierno de Puebla
Hacer historia. Hacer futuro.



Secretaría
de Administración

Gobierno de Puebla