

25 DE SEPTIEMBRE DEL 2020

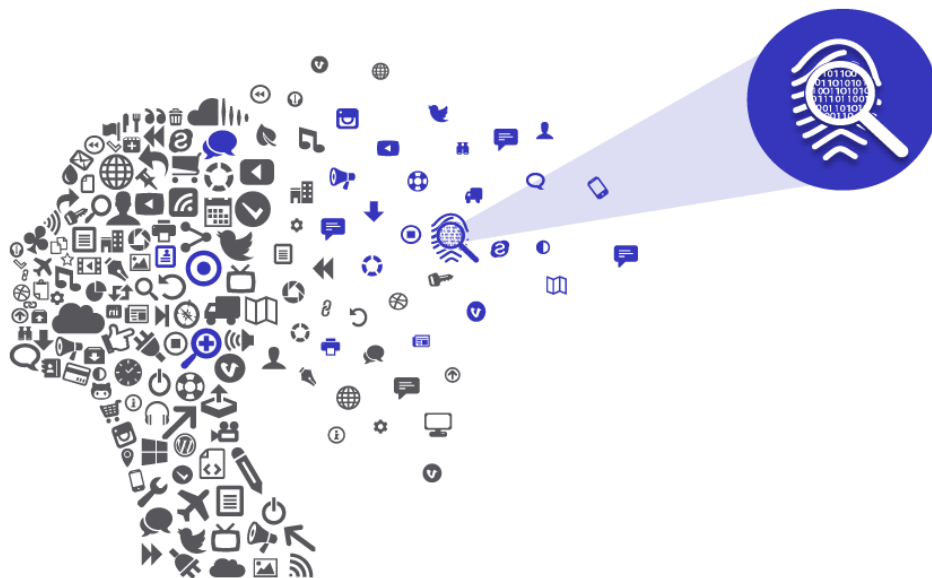
**El uso de información proveniente de fuentes no tradicionales como mecanismo de control y prevención de riesgos. (Bases de datos externas, redes sociales, fuentes abiertas, etc.)**

**Luis Fernando Cadena Barrera**



# Ciencia de Datos aplicadas a los Negocios

Identificar, diseñar e implementar estrategias para la recolección de información, a través de una metodología de control inteligente, mediante la segmentación y selección de potenciales clientes con un mayor potencial de negocio.



**Inteligencia**

*intelligentia - intellegere*

*Intel - capacidad    legere - escoger*

***Quien sabe escoger***

# Toma de Decisiones basadas en Datos

La **Ciencia de Datos** aplicada a los negocios establece que es posible sistematizar la extracción de conocimiento a partir de grandes cantidades de información para resolver problemas del negocio.

Existen dos principios que buscan encontrar patrones a partir del análisis de los datos:

## 1. Similitud:

Determinar la probabilidad de que dos entidades descritas por un conjunto de datos o atributos sean iguales.

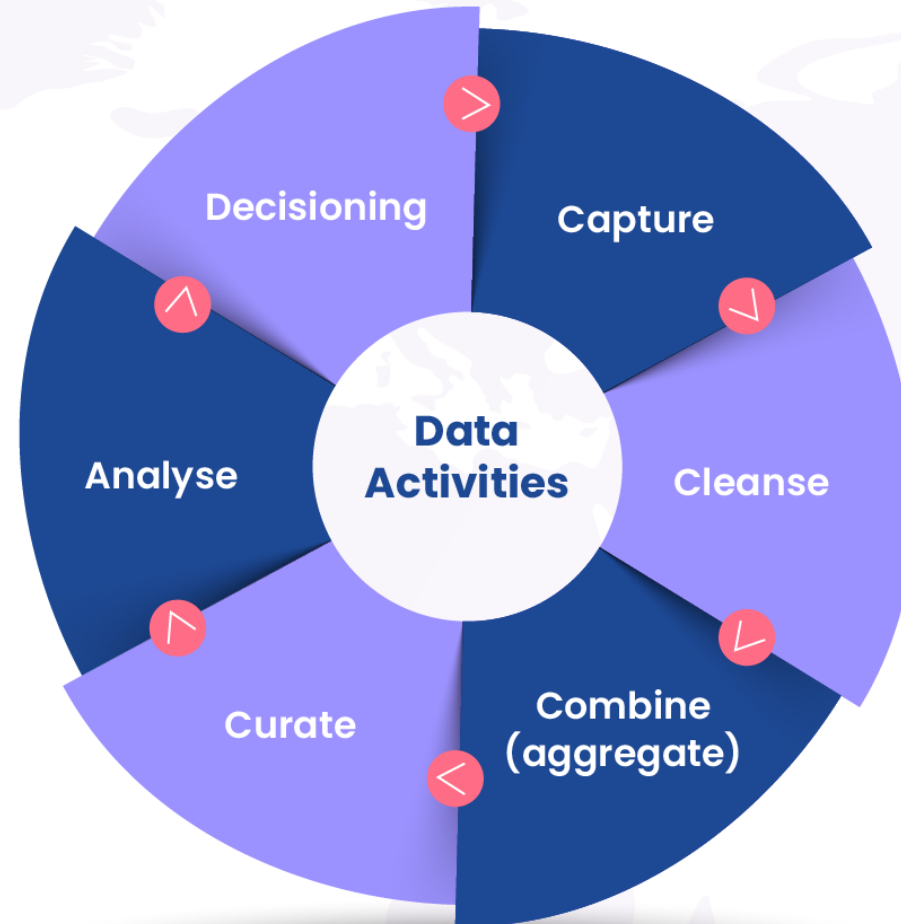


## 2. Regularidad:

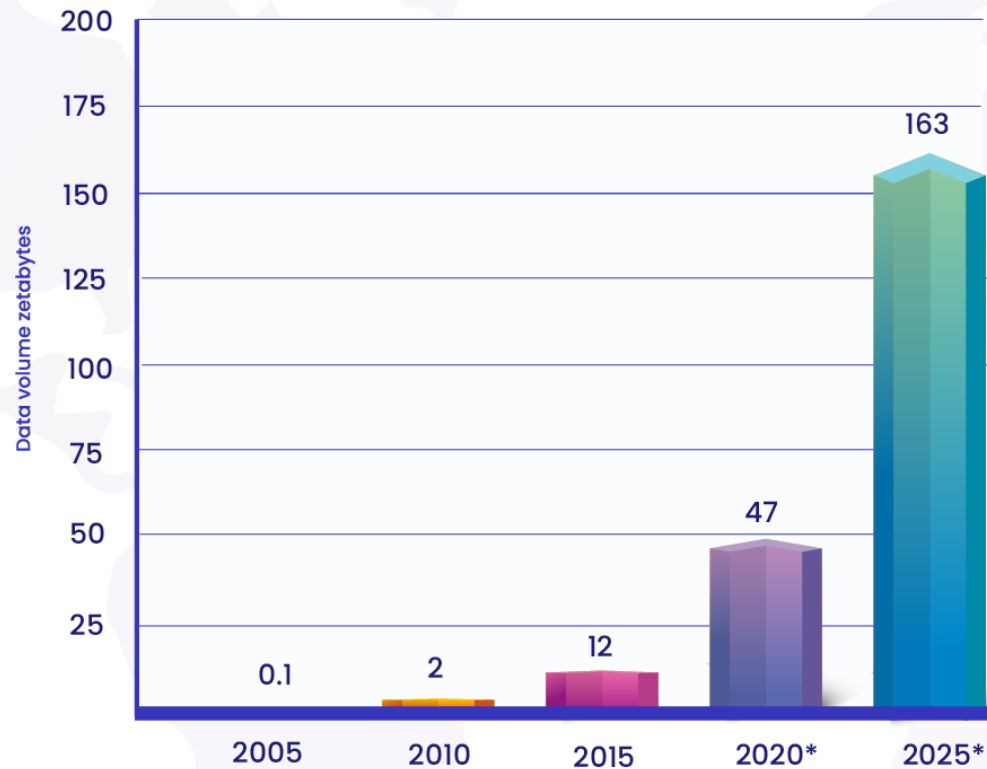
Determinar la probabilidad de que un fenómeno se repita a partir de la información que existe sobre su comportamiento en el pasado.



## Ciclo de vida de los datos



## Creación diaria de datos 2005 – 2025 \* (zettabytes, \* esperada)



**90%**  
de los datos disponibles  
hoy, fueron creados en los  
últimos 2 años.

**0.5%**  
de los datos que hoy  
producimos son  
analizados.

**2.3**  
Trillones de megabytes  
se producen en 1 día.

# Las fuentes de información tradicionales

El entorno digital ha permitido el desarrollo de fuentes de información que podemos denominar tradicional y que se construye alrededor de tres ejes:



Datos de identidad



Datos censales, estadísticos  
y administrativos



Datos económicos



# Las fuentes de información NO tradicionales



# Fuentes agregadas – Productos de información

Los productos de información se pueden clasificar en tres categorías generales:




1. Los datos como servicio (**as a service**) se ofrecen como plataformas que generan un producto específico: noticias, clima, comparadores de viaje, etc.
2. Los productos mejorados (**data-enhance**) se ofrecen para mejorar bienes o servicios (tangibles e intangibles): asistentes digitales, streaming o vehículos autónomos.
3. Los productos de conocimiento (**data-insight**) son aquellos en los que se hacen procesos de análisis con el objetivo de identificar e informar hechos o tendencias: financieras, fiscales, legales, de mercado, preferencias, etc.



# Fuentes transaccionales automatizadas

## La factura electrónica

 Campos en cada transacción (hasta) **161**

**Desde la identidad fiscal hasta la clasificación de los productos y la ubicación.**

Sector, producto, características, precio, hora, Código postal, método de pago, etc.

**1,297**



**Campos adicionales**  
En 26 anexos para sectores específicos

### Detalle sobre transacciones de sectores específicos

Nómina, líneas aéreas, vehículos, gasolinas, divisas, comercio exterior, notarios, colegiaturas, etc.

# Uso de datos para cumplimiento y prevención de riesgos

La información digital proveniente de fuentes de datos tradicionales y no tradicionales, puede ayudar a las organizaciones a mitigar los riesgos y facilitar el cumplimiento regulatorio a partir de soluciones analíticas, un ejemplo es:



Análisis de  
Congruencia de  
ISR e IVA



Redes confiables  
de proveeduría



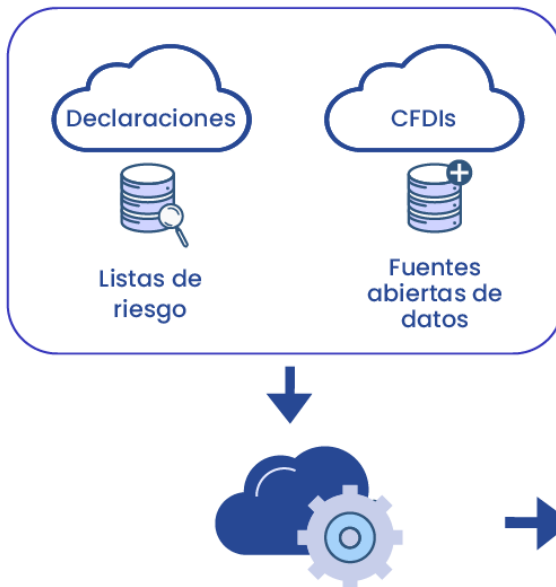
Análisis de Controles  
Volumétricos

Powered by  Intelitax

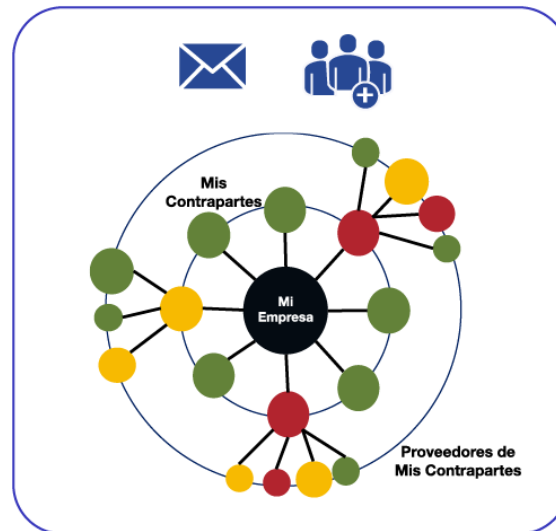
# Redes Confiables de Proveduría

Plataformas que te permitan monitorear al comportamiento de tu cadena de proveduría en varios niveles de forma **SEGURA** y **NEUTRAL** sin poner en riesgo su confidencialidad y garantizando su veracidad.

## Automatización



## Red de Proveedores



## Seguridad



***“Puedes tener datos sin tener información, pero no puedes tener información sin tener datos”.***

*Daniel Keys Moran  
Programador y escritor*

