



BIG DATA



1. MUNDO DIGITAL

2. NECESIDADES

3. PROFESIONES

4. PRIVACIDAD

5. ESTILOS DE VIDA

6. EXTENSIÓN



Para Pensar

¿QUÉ SON LOS DATOS?

¿QUÉ SE ESPERA CONOCER?

ÁREAS DE DESEMPEÑO

TECNOLOGÍA



DESAFÍO

NATIVOS DIGITALES INMIGRANTES DIGITALES

Marc Prensky

En *On the Horizon* (MCB University Press, Vol. 9 No. 6,
December 2001)

Big Data

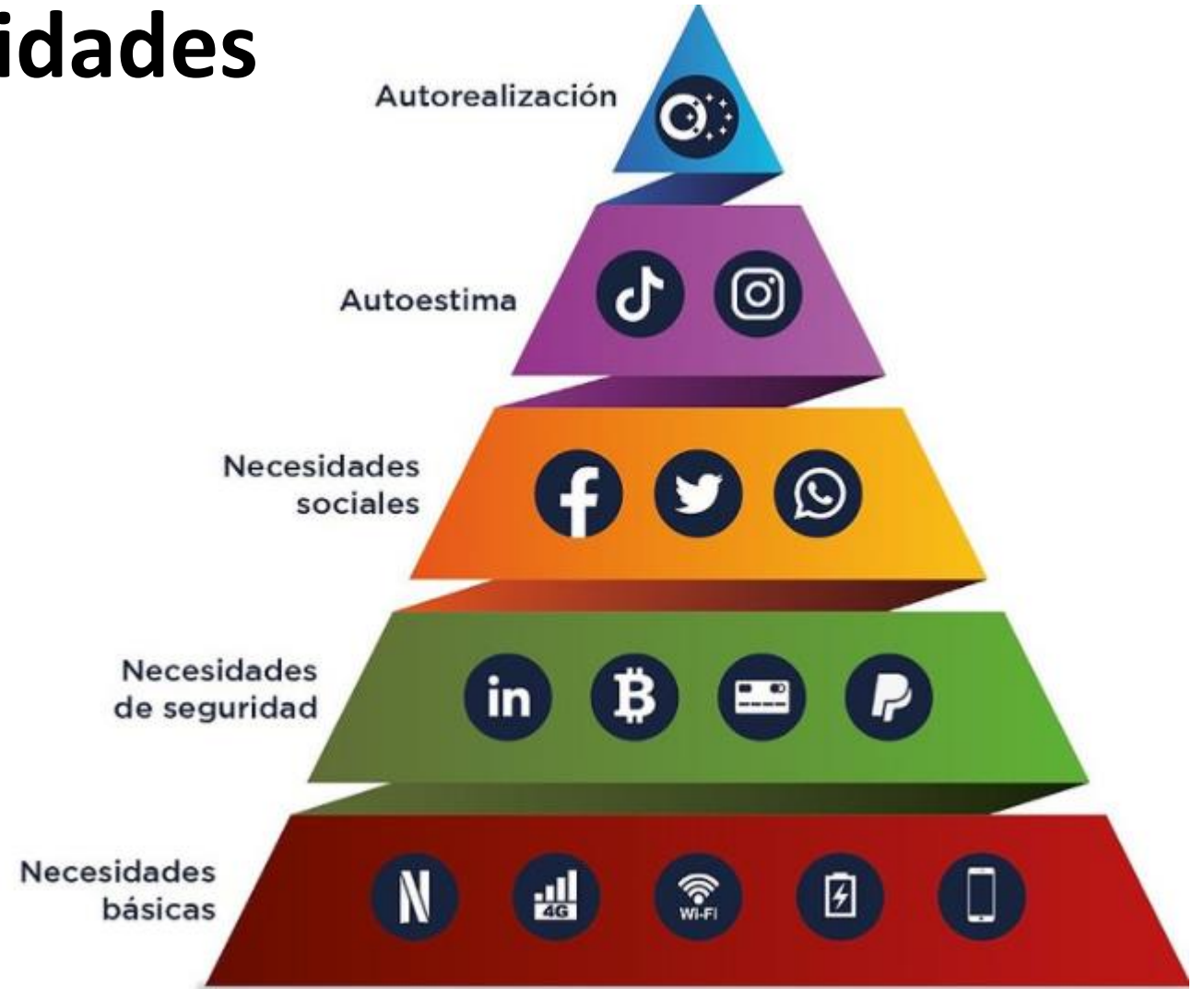


1. mundo digital

¿Qué datos nuestros tienen
Google, Whatsapp, ,
Instagram, Home Banking?

Políticas de Seguridad y
Privacidad?

2. necesidades



Big Data

3. profesiones

Contador
Médico
Ingeniero
Abogado
Maestro

Carpintero
Plomero
Modista
Dibujante

Desarrollador de aplicaciones de voz
Profesional de Big Data
Experto en bockchain
Experto en Biometría
Experto en marketing y automatización

Influencer
Instagramer
Gamer
Experto en social media
Community Manager

<https://www.universia.net/ar/actualidad/orientacion-academica/profesiones-tecnologicas-mas-demandadas-en-2021.html>

Big Data

4. privacidad

RAE: Ámbito de la vida privada que se tiene derecho a proteger de cualquier intromisión

**Protección por constituciones provinciales, constitución nacional, DDHH, organismos internacionales, etc.
Se agregan Datos personales y más.**



**Ley de protección de datos personales
Derecho al Olvido
Aceptación de políticas
Anonimización
Disociación**

Big Data

5. Estilo de vida

- Gran Hermano
- Celular
- Situación actual
- Redes sociales
- Internet de las cosas
- Localización.
- Dietas Gustos
- Intereses
- Política
- Opiniones
- Gastos

Estilo de vida se refiere a cómo se orientan los intereses, las opiniones, y los comportamientos y conductas de un individuo, grupo o cultura. ...

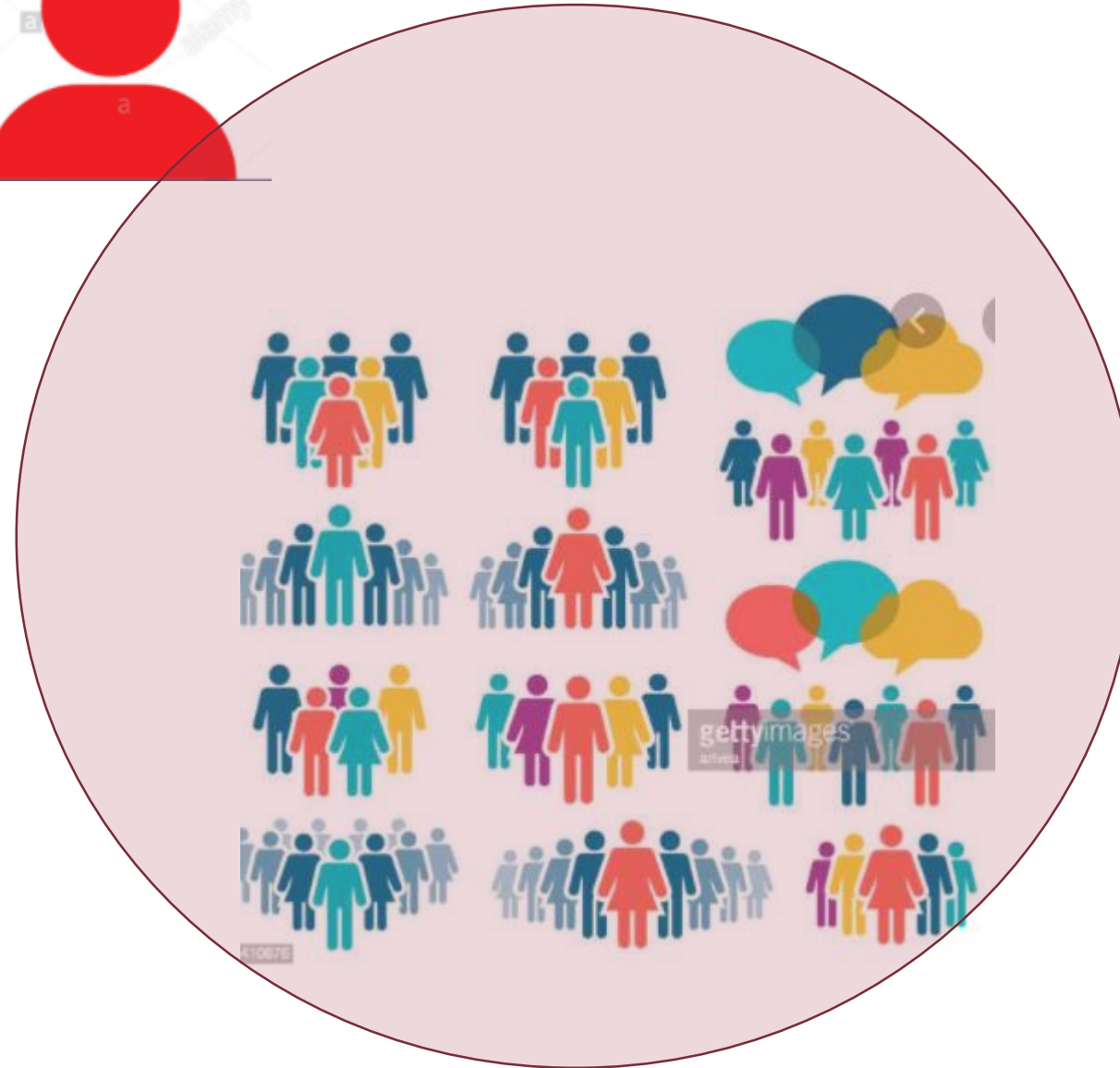
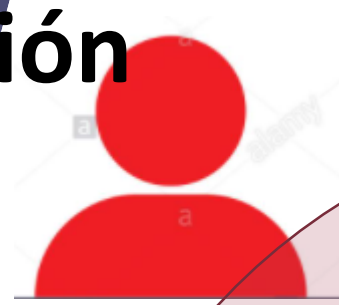
Tener un **estilo de vida** específico implica una opción consciente o inconsciente entre un sistema de comportamientos y otros



Big Data

6. Extensión

- Redes Sociales
- Redes de trabajo
- Grupos
- Intereses
- Contactos
- Imágenes compartidas
- Deportes
- Marcas
- Aplicaciones que utilizamos



Big Data



BIG DATA: ORIGEN Y DEFINICIÓN

¿QUÉ ES BIG DATA Y QUÉ REPRESENTA EN REALIDAD?

- Definición
- Lo que no es (MITOS)
- ¿En qué consiste?
- Origen del Big Data



DEFINICIÓN DE BIG DATA

- “Sets de datos extremadamente grandes que deben ser analizados computacionalmente para revelar patrones, tendencias y asociaciones, especialmente relacionadas al comportamiento humano y a las interacciones”

LO QUE NO ES BIG DATA (MITOS)

- **Mito 1: Todo el mundo lo está haciendo**
- **Mito 2: Big Data es todo acerca del tamaño**
- **Mito 3: Big Data dirá lo que pasará después**
- **Mito 4: Trabajar con Big Data requiere grandes presupuestos**
- **Mito 5: Big Data es algo importante solo para las áreas de tecnología**

¿EN QUÉ CONSISTE BIG DATA?

- Volumen
- Velocidad
- Variedad

ORIGEN DE BIG DATA

Big Data Desde los inicios hasta hoy

Memoria Virtual (Fritz-Rudolf Güntsch)

Concepto desarrollado por este físico alemán como una idea que trataba el almacenamiento finito como infinito y permitía procesar datos sin las limitaciones de memoria del hardware.



Reconocimiento de voz (William C. Dersch)

Presenta "Shoebox" la primera máquina en comprender 16 palabras y 10 dígitos en inglés hablado mediante el uso de los datos disponibles en ese momento, y la capacidad de procesarlo de manera eficiente.



El auge de la comunicación bidireccional

El Censo de flujo de información de Japón, comenzó a rastrear el volumen de información. Con el número de palabras utilizadas como unidad de medida en los medios, concluyó que la demanda de comunicación unidireccional se había estancado. Sin embargo, aumentaba la demanda de comunicación bidireccional y más personalizada.



La fundación de la World Wide Web (Tim Berners-Lee)



La World Wide Web explota

La década de 1990 fue un momento de crecimiento explosivo para la tecnología y los datos de Business Intelligence comenzaron a acumularse en forma de documentos de Microsoft Excel.



Las tres V de Big Data de Gartner (Doug Laney)

Laney publicó un trabajo de investigación titulado *3D Data Management: Controlling Data Volume, Velocity, and Variety de donde se salieron las "3V" aceptadas del Big Data.*



1956



Ley bibliométrica del aumento exponencial (Derek Price)

Observa el aumento exponencial (x2 cada 15 años y x10 cada 50 años) del número de publicaciones científicas y artículos. Con esta ley explica que "cada avance [científico] genera una nueva serie de progresos a una tasa de natalidad razonablemente constante, de modo que el número de nacimientos es estrictamente proporcional al tamaño de la población de descubrimientos en cualquier momento".



1961



1962



La Era de la Automatización

Debido a la afluencia de información en los años 60, las organizaciones comienzan a diseñar, desarrollar e implementar sistemas de computación centralizados que les permiten automatizar sus sistemas de inventario.

1966



1975



Ley de Parkinson (I. A. Tjomsland)

En su charla, "Where Do We Go From Here?" dijo que la primera Ley de Parkinson se puede parafrasear para describir la industria: "Los datos se expanden hasta llenar el espacio disponible para el almacenamiento".

1980



1989



El primer informe de Bases de Datos

Crystal Reports crea la primera base de datos simple con Windows, lo que facilita a las empresas el trabajo. De esta forma, comprar más memoria incentiva el uso de técnicas de programación que usan la memoria de forma más intensiva.

1992



1995



El problema del Big Data (Michael Cox & David Ellsworth)

El término "Big Data" fue utilizado por primera vez en un artículo de estos investigadores de la NASA donde afirman que "el aumento de los datos se estaba convirtiendo en un problema para los sistemas informáticos actuales". Esto también se conoce como el "problema del Big Data".

1997



2001



Google publica GFS y MAPREDUCE

Google publica en 2003 y 2004 las publicaciones de GFS (Google File System) y MapReduce que son los dos pilares fundamentales de Hadoop y de las tecnologías Big Data, que en 2006 incluyó Yahoo! en Hadoop.

2003



ORIGEN DE BIG DATA

Hadoop: código abierto para Big Data

Sistema de código 100% abierto para almacenar, procesar y analizar grandes volúmenes de datos.



Lanzamiento de Hortonworks

Competidor de Cloudera, líder de plataformas de datos abiertas y conectadas permite acumular, analizar y actuar sobre la información derivada de los datos, y es 100% de código abierto.



250 libros sobre Big Data en Amazon

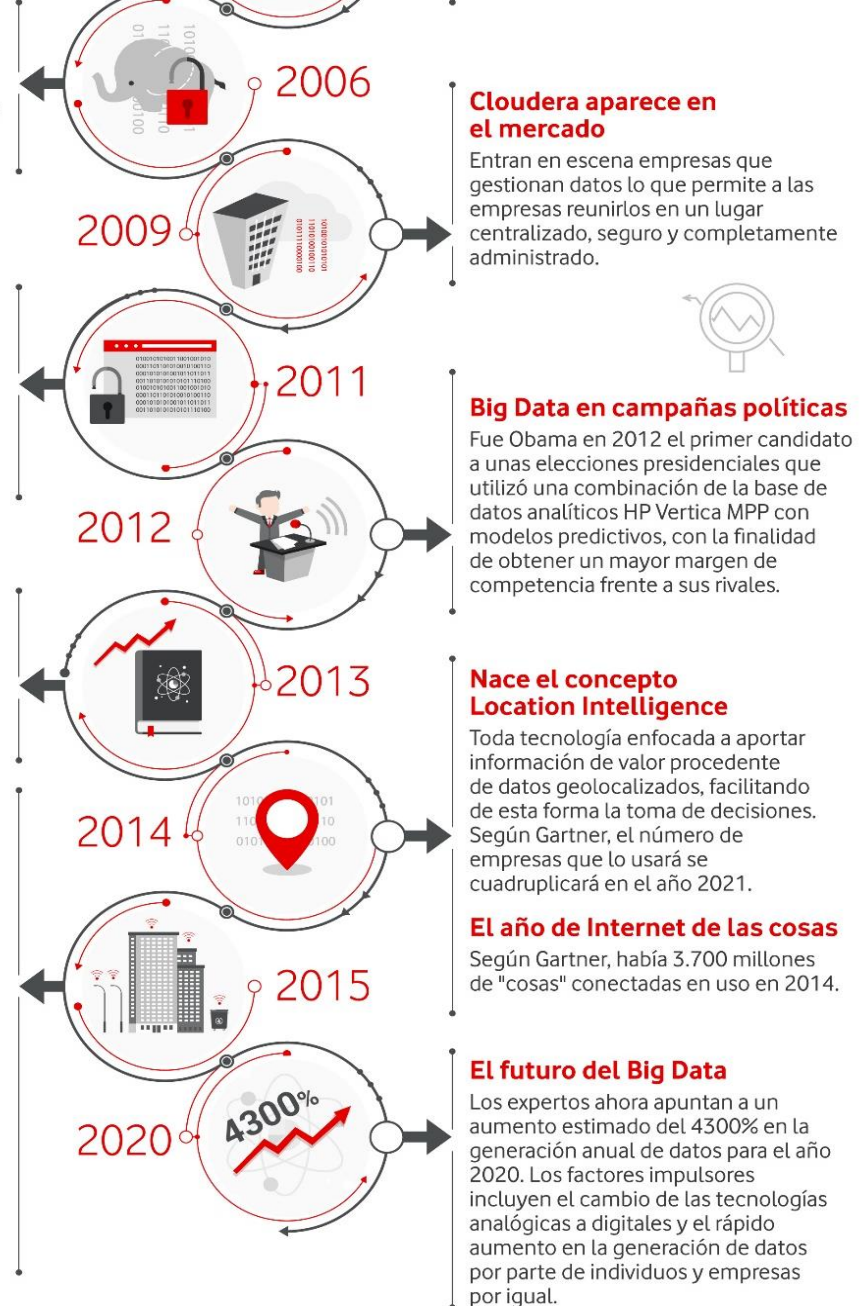
Aumenta el interés de este tema y por ende el número de libros sobre Big Data.

La eliminación de Big Data del Gartner Hype Cycle

El Big Data deja de aparecer en lo más alto del Hype Cycle de Gartner ya que deja de considerarse una tecnología emergente.

Smart Cities y los datos

Gartner estima que más de 1.100 millones de "cosas" conectadas serán utilizadas por ciudades inteligentes en 2015, incluida la iluminación LED inteligente, monitoreo de atención médica, cerraduras inteligentes y redes de sensores para "cosas" como detección de movimiento o control de la calidad del aire.



Fuentes:

forbes.com | winshuttle.com | wikipedia.org | Gartner.com | www-03.ibm.com

EVOLUCIÓN DE BIG DATA

1956-2011

El problema de Big Data
Cox & Ellsworth: utilizan por primera vez el término en un artículo de la NASA, donde afirman que "el aumento de los datos se estaba convirtiendo en un problema para los sistemas informáticos actuales"

1997

1956

1956

1961

1962

1966

1975

1995

1992

1989

1980

2001

2003

2006

2009

2011

Memoria Virtual

Fritz- Rudolf Guntch: Idea que trataba el almacenamiento finito como infinito y permitía procesar datos sin las limitaciones de memoria del hardware

Reconocimiento de Voz

William Dersh: Presenta 'Shoebbox' la primera máquina en comprender 16 palabras y los 10 dígitos en inglés.

El auge de la comunicación bidireccional

El censo de flujo de información en Japón comenzó a rastrear el volumen de información. Estancamiento de la comunicación unidireccional. Aumenta la demanda de comunicación bidireccional y mas personalizada

Bibliométrica del aumento exponencial

Derek Price explica que cada avance científico genera una nueva serie de progresos a una tasa constante de modo que el nro de nacimientos es estrictamente proporcional al tamaño de la población de descubrimiento en cualquier momento

World Wide Web explota

La década del 90 fue un momento explosivo para la tecnología y los datos de Business Intelligence comenzaron a acumularse en documentos de Excel

La era de la automatización

Las organizaciones comienzan a desarrollar sistemas de computación centralizados que permiten automatizar su sistema de inventarios

Fundación de la World Wide Web

Tim Berners-Lee

Lanzamiento de Hortonworks

Competencia de Cloudera, líder de plataformas de datos abiertas y conectadas permite acumular, analizar y actuar sobre la información derivada de los datos. 100%abierto

Ley de Parkinson

Tjomsland: dijo que esa ley se puede parafrasear para describir la industria: "Los datos se expanden hasta llenar el espacio disponible para el almacenamiento"

Primer informe de Bases de Datos

Crystal Reports crea la primer base de datos simple con Windows, facilitando el trabajo de las empresas, así comprar mas memoria incentiva el uso de técnicas de programación que usan la memoria más intensivamente

Las 3 V de Big Data de Gartner

Laney publicó: 3D Data Management: Controlling Data volumen, Velocity and Variety de donde surgieron las 3v de big data.

Hadoop: código abierto para Big Data

Sistema de código 100% abierto para almacenar, procesar y analizar grandes volúmenes de datos.

Google publica GFS y MAPREDUCE

Son los 2 pilares fundamentales de Hadoop y de las tecnologías big data que en 2006 incluyeron Yahoo! En Hadoop

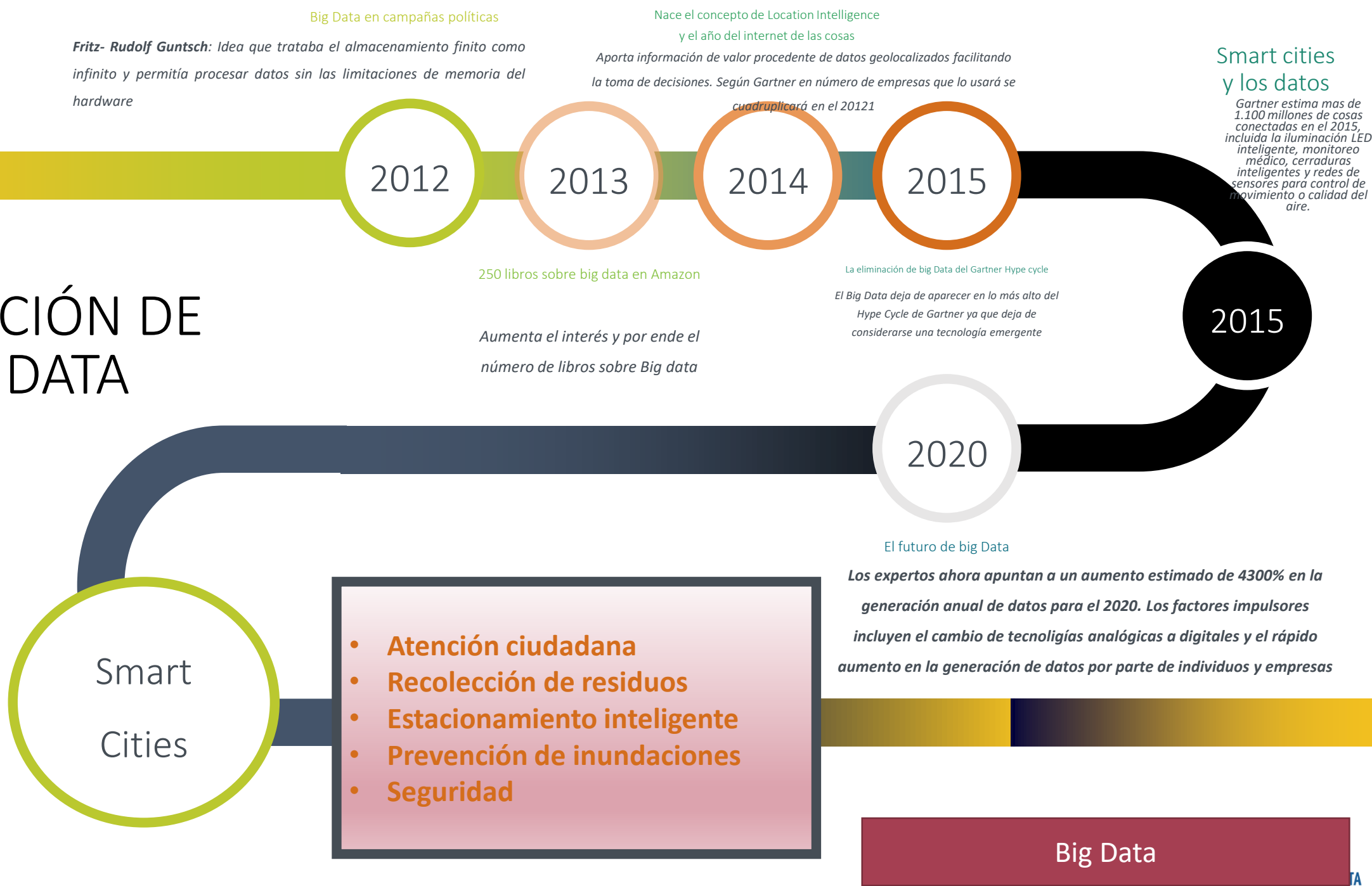
Cloudera aparece en el mercado

Entran en escena empresas que gestionan datos lo que permite a las organizaciones reunirlos en un lugar centralizado, seguro y completamente administrado.

Big Data

EVOLUCIÓN DE BIG DATA

2012-2020





CARACTERÍSTICAS Y ESCALABILIDAD

Big Data

CARACTERÍSTICAS DE BIG DATA Y COMO HACERLO ESCALABLE ¿EN QUÉ CONSISTE?

- Características de los equipos de Big Data (Veracidad)
- La escalabilidad en Big Data
- Críticas al Modelo de las V

CARACTERÍSTICAS PRINCIPALES DE BIG DATA

Equipo

- Almacenamiento: recursos hardware y software que permite el almacenamiento de los datos.
- Procesamiento: herramientas de procesamiento de los datos.
- Análisis: metodología seguida para realizar el análisis de los datos que derivan en información de valor.

Usos comunes y realistas

- Gobierno
- Desarrollo internacional
- Salud
- Marketing digital
- Educación
- Industria
- Automatización



LA ESCALABILIDAD EN BIG DATA

- Procesamiento
- Memoria
- Almacenamiento



ESCALABILIDAD: ALTERNATIVAS

- Aumentar capacidad de equipos propios
- Implementar espacio y procesamiento en la nube pagando solo por lo utilizado.



CRÍTICAS AL MODELO DE LAS V

- Completitud de los datos
- Correlación de los datos
- Decisiones relegadas a la automatización

PRINCIPALES CAMINOS PARA DEDICARSE AL BIG DATA

Infraestructura y
arquitectura de soluciones
tecnológicas.

- Conocimiento del Negocio
- Habilidades técnicas y de programación
- Habilidades estadísticas y matemáticas

Analytics y data science.

- Entendimiento de tendencias, patrones, segmentación
- Perfilado de usuarios
- Análisis de comportamiento
- Conocimientos matemáticos y estadísticos
- Capacidad de análisis
- Conocer el negocio

Big Data

EJEMPLOS

El Real Madrid y el Big data

<https://www.youtube.com/watch?v=yoSqojO2-CQ&t=53s>

CÓMO LA CIENCIA CREÓ EL INCREÍBLE
LIVERPOOL DE KLOPP

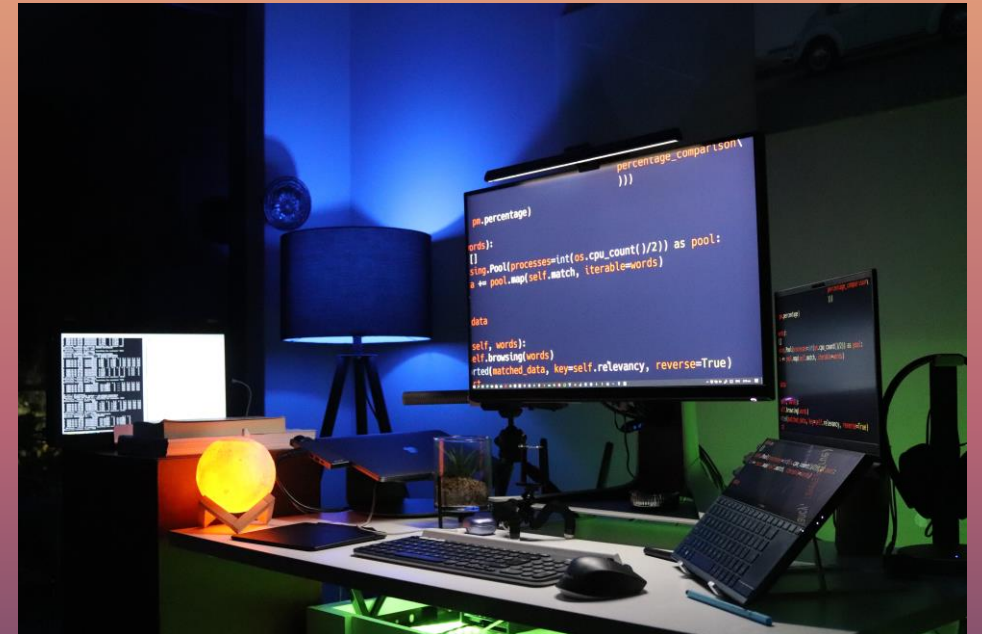
https://www.youtube.com/watch?v=1Kx_JpZlXB0

¿Qué es big data?

<https://www.youtube.com/watch?v=w4vsFKMO7XA>

Big data de las redes sociales para predecir el
comportamiento ciudadano

<https://www.youtube.com/watch?v=yoSqojO2-CQ&t=53s>



Big Data