

# Webinars

Solex  
2020

1. BA

2. Pricing Analytics

3. BigData

4. Computación Cloud

5. BI

6. Tableau Prep

7. Análisis de Sentimiento

8. Dilo. Dashboard

TEMAS

2020



BA

# Inteligencia de Negocios (BI)

Inteligencia de Negocios **es la habilidad para transformar** los datos en información, y la información en **conocimiento**, de forma que permita tomar de decisiones generando **Valor**.

Podemos definir la Inteligencia de Negocio como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos, para su **explotación directa** o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre el negocio.

# Analítica de Negocios (BA)

Impulsa **escenarios avanzados** de análisis de la información permitiendo que la toma de decisiones se base en modelos estadísticos que **no son necesariamente descriptivos** como lo hace en muchos escenarios el BI.

# Evolución de Arquitecturas



# Desafíos Actuales del Mundo Analítico

La **velocidad** actual para tomar **decisiones de gran valor e impacto** exige un desafío en la forma de **consumir los datos** incursionando en nuevas técnicas, arquitecturas y escenarios de análisis



# Cambio de Mentalidad

## Mentalidad Ideal

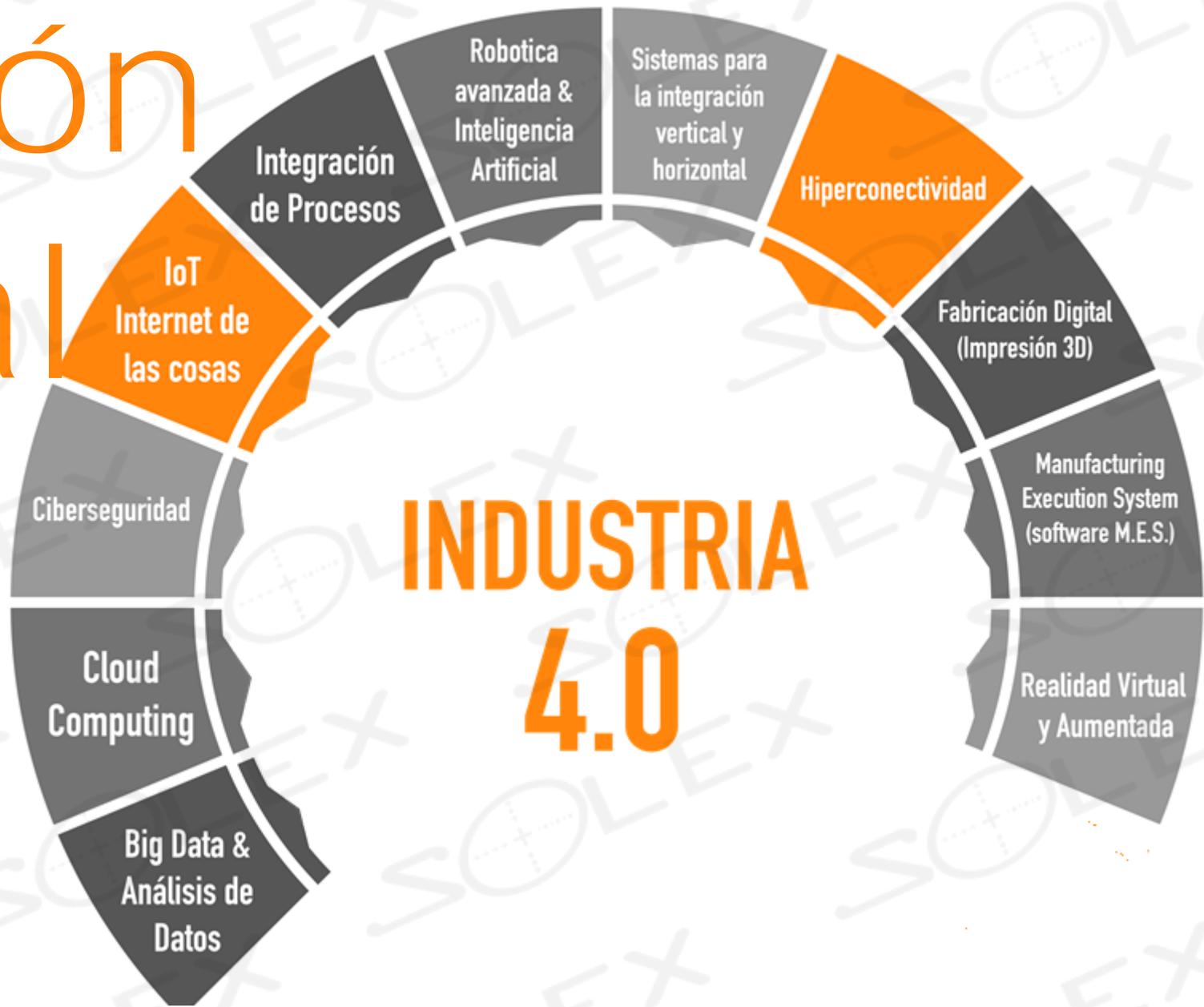
- + Cambio del clima del dato.
- + Especificar tendencias y patrones.
- + Escenarios diferenciadores, **generadores de valor**.
- + Información Clara y apoyada en resultados.
- + Impacto Fácil.
- + Habilitar y **preparar equipo**.

## Mentalidad Estándar

- + **Suposiciones**.
- + Creencias en volúmenes datos.
- + Puntos de vista para consumo de los datos.
- + Dificultad en un modelo de la verdad.
- + Definición de **Roles no íntegros**.
- + Hacer Promesas sobre datos que no son viables.



# Revolución Industrial



# Generación De Valor

**NETFLIX**

**NO MATÓ A BLOCKBUSTER**

Los ridículos cargos por pagos atrasados lo hicieron

**UBER**

**NO MATÓ A LOS TAXIS**

El acceso limitado a taxis, el mal servicio y el control de tarifas lo hicieron

**Apple**

**NO MATÓ A LA INDUSTRIA DE LA MÚSICA**

El obligar a comprar álbumes completos lo hizo

**amazon**

**NO MATÓ A OTROS MINORISTAS**

El mal servicio y experiencia del cliente lo hicieron

**airbnb**

**NO ESTÁ MATANDO A LA INDUSTRIA HOTELERA**

La disponibilidad limitada y las opciones de precios lo hicieron

La tecnología por sí misma no es el verdadero disruptor

**NO CENTRARSE EN EL CLIENTE ES LA MAYOR AMENAZA PARA CUALQUIER NEGOCIO**

*#Viral*

# Minería De Datos



Proceso analítico diseñado para explorar datos (usualmente grandes cantidades) en busca de patrones, tendencias y comportamientos consistentes entre variables.

# Escenarios NO Descriptivos

¿Cómo puedo predecir la deserción de mis clientes?

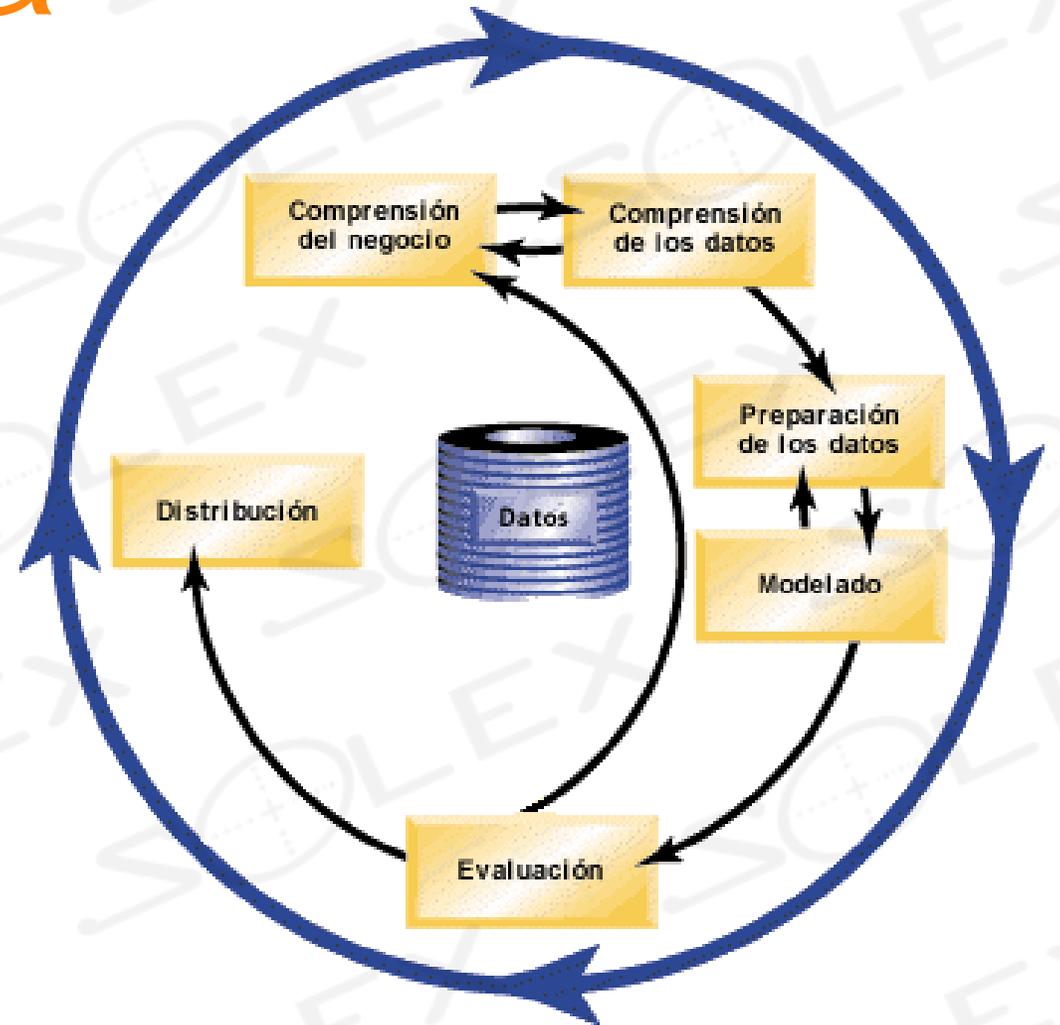
¿Cómo puedo dirigir mis campañas a los clientes más propensos a comprar?

¿Cómo puedo adelantarme a problemas potenciales en vez de tomar decisiones cuando ya es demasiado tarde?

# + Ejemplos

- Predicción de estudiantes que pueden desertar
- Predicción de clientes que pueden retirarse
- Perfil de clientes que incumplen los pagos
- Segmentación de clientes
- Clientes más propensos a aceptar ofertas
- Predicción de ingresos, gastos.
- Predicción de enfermedades en humanos, plantas
- Características de empleados “fieles”
- Características de los mejores vendedores
- Predicción de fallos de maquinaria
- Factores que influyen en emprendimientos exitosos
- Productos que impulsan otros (combos)

# Metodología Crisp – DM: Preparación de los datos



# Metodología Crisp – DM: Modelado

- Descubrir grupos demográficos con comportamientos distintos (Ej. Clientes, proveedores, almacenes)
- Descubrir grupos objetivo
- Detectar desviaciones
- Identificar prospectos

Clustering

- Descubrir Afinidades (eventos, salidas, ítems, etc.)
- Identificar el evento más probable a ocurrir, etc.
- Detectar desviaciones

Asociaciones y secuencias

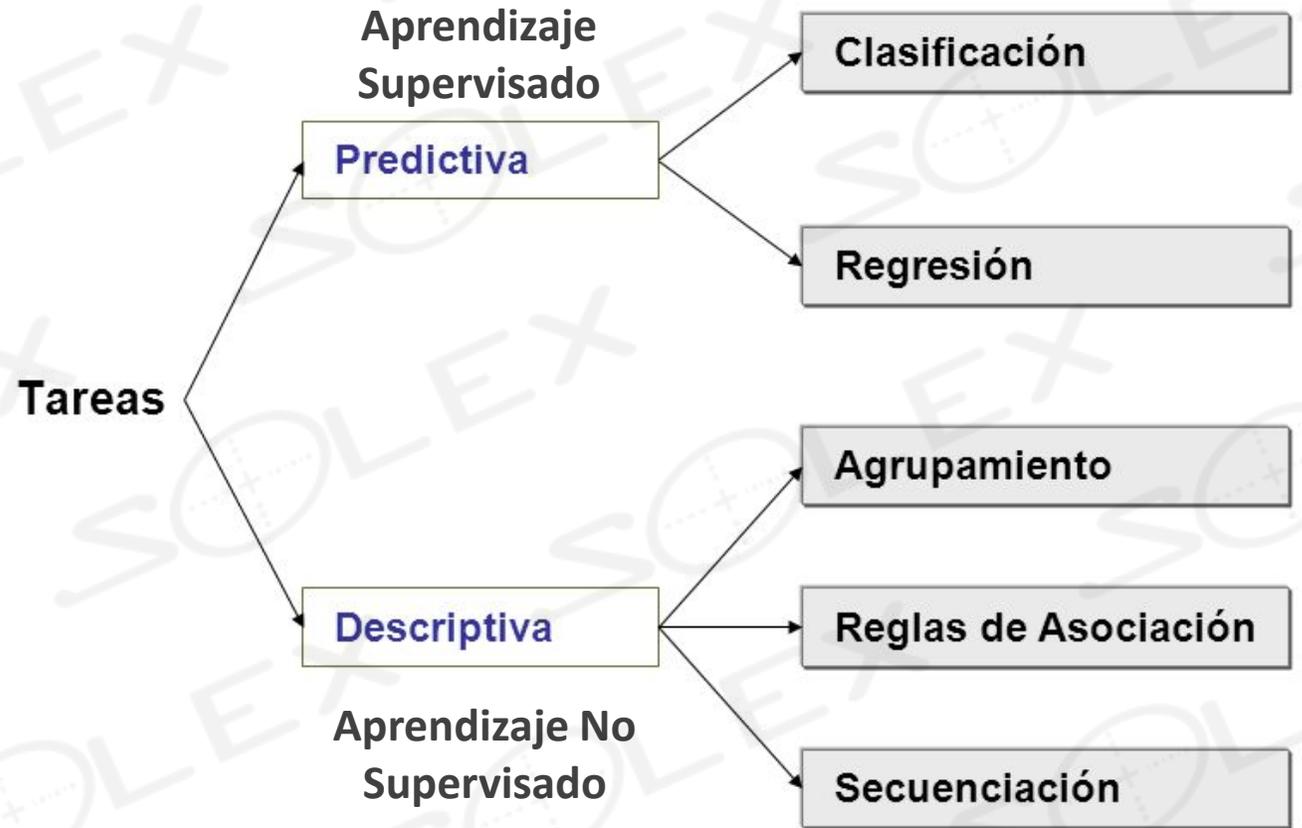
- Predecir respuestas (SI/NO)
- Clasificar Resultados (Deserción, desgaste, fraude, tipos de fallas, etc.)

Clasificación

- Predecir valores numéricos
- Predecir probabilidad de resultados (0 – 1)
- Predecir gastos esperados, costos, utilización, etc.

Predicción

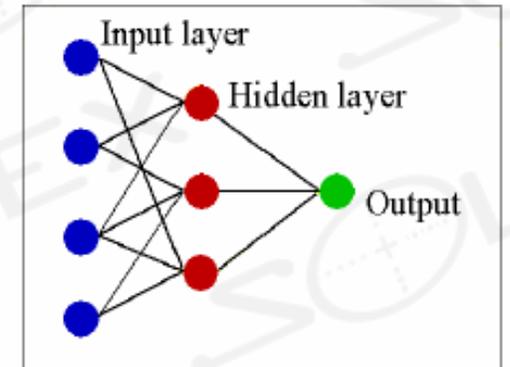
# Tareas de Minería de Datos



# Aprendizaje Supervisado

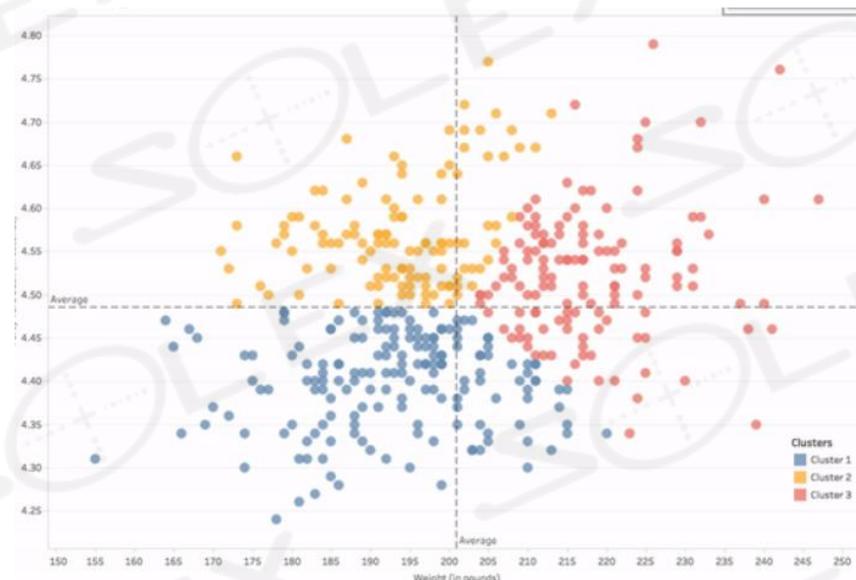
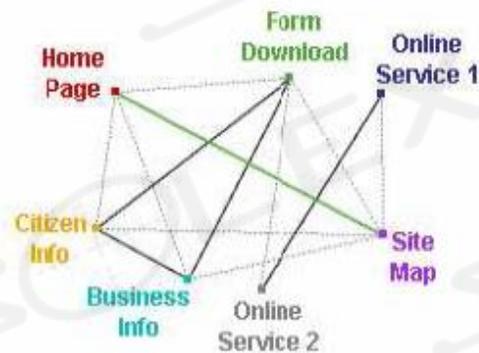
Construir un modelo predictivo, con base en un conjunto de datos histórico cuyo resultado es conocido.

Usar el modelo para predecir casos futuros.



# Aprendizaje NO Supervisado

Encuentra patrones emergentes y casos inusuales.



# Técnicas

PROBLEMA	DESCRIPCIÓN	EJEMPLOS	TÉCNICAS	
EXPLORACION RESUMEN DATOS	Descripción concisa de datos, elemental y agregada. Análisis exploratorio	Casi cualquier problema incluye elementos de descripción	ERPs, Stats, OLBI, cuadro de mando	Técnicas de Aprendizaje Supervisado
SEGMENTACION	Separación de datos en subgrupos significativos	Market Segmentation,	SOM K-MEANS	Técnicas de Aprendizaje NO Supervisado
DESCRIPCION CONCEPTUAL	Descripción inteligible y útil de conceptos / clases / grupos. <b>Prima conocimiento</b> sobre la precisión. Ligado a clasif / segmentación	Descripción de grupos de clientes según lealtad. Perfilado de segmentos if SEX=male and age>45 then CUST=loyal	Rule Induction, Decision Trees	Técnicas de Aprendizaje Supervisado
CLASIFICACION	Es asumido que diferentes observaciones, definidas por variables, pertenecen a clases dadas ( <b>supervisado</b> )	Bancarrota, Credit Scoring	Neural Nets Rule induction Decision Trees KNNs	Técnicas de Aprendizaje Supervisado
PREDICCIÓN (REGRESIÓN, FORECASTING)	Variable dependiente continua. Datos valores de variables predictivas, encontrar valor de predicción ( <b>supervisado</b> )	Bolsa, beneficios empresa, cuota de mercado	Regression Analysis, Regression Trees, Time Series, NN	Técnicas de Aprendizaje Supervisado
ANALISIS DE ASOCIACIONES	Búsqueda de dependencias entre variables A menudo con segmentación	Basket Analysis 30% de los que compraron cacahuets compraron también cerveza	Correlation Analysis, Association Rules, Bayesian Networks,	Técnicas de Aprendizaje NO Supervisado

# MODELO DE MADUREZ

GESTIÓN DE DATOS



OPTIMIZACIÓN DEL CANAL



CONOCIMIENTO DEL NEGOCIO



MODELOS ESTADÍSTICOS



TÉCNICAS DE ANÁLISIS

TÉCNICAS DE CIENCIA DE DATOS

# Científico de Datos

## MODERN DATA SCIENTIST

Data Scientist, the sexiest job of 21st century requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.



### MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

### DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative

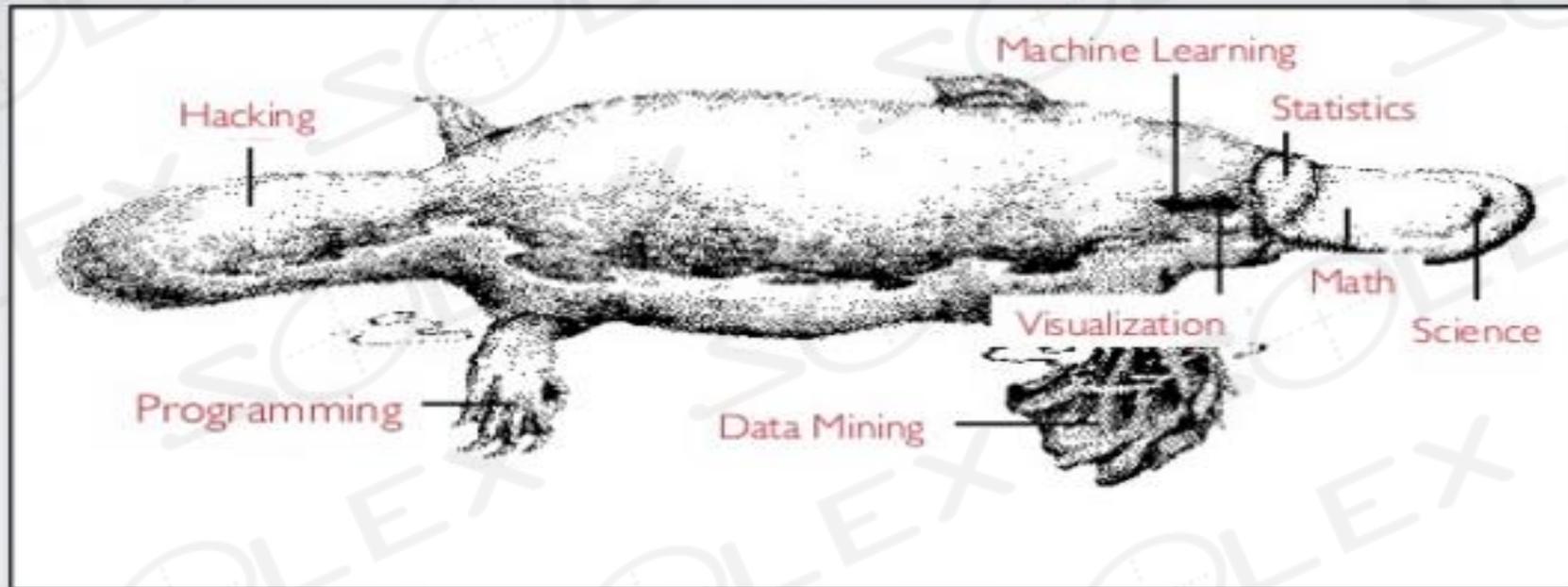
### PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing package e.g. R
- ☆ Databases SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

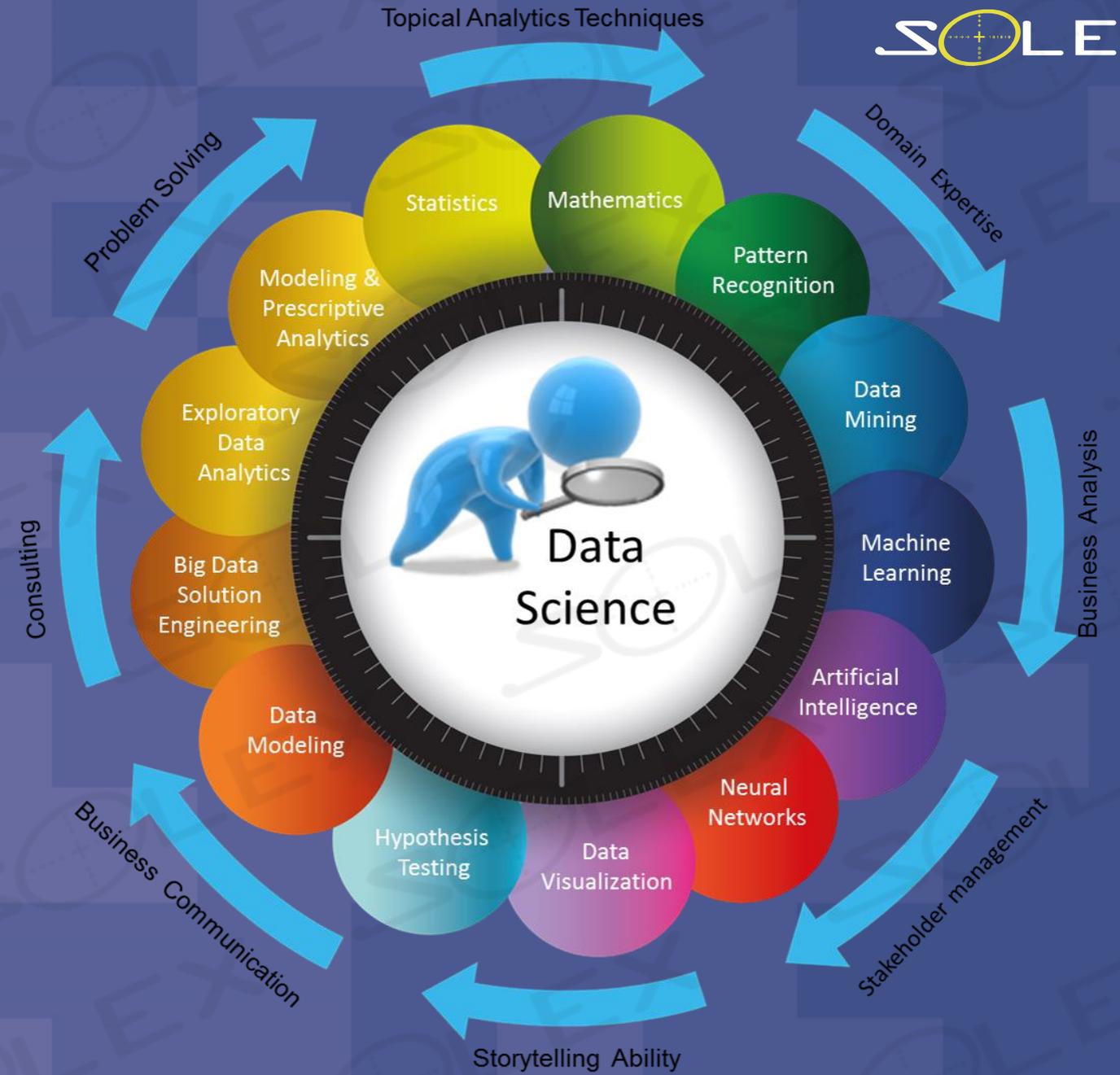
### COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

# The Platypus – Billed Data Scientist



# Áreas del Saber



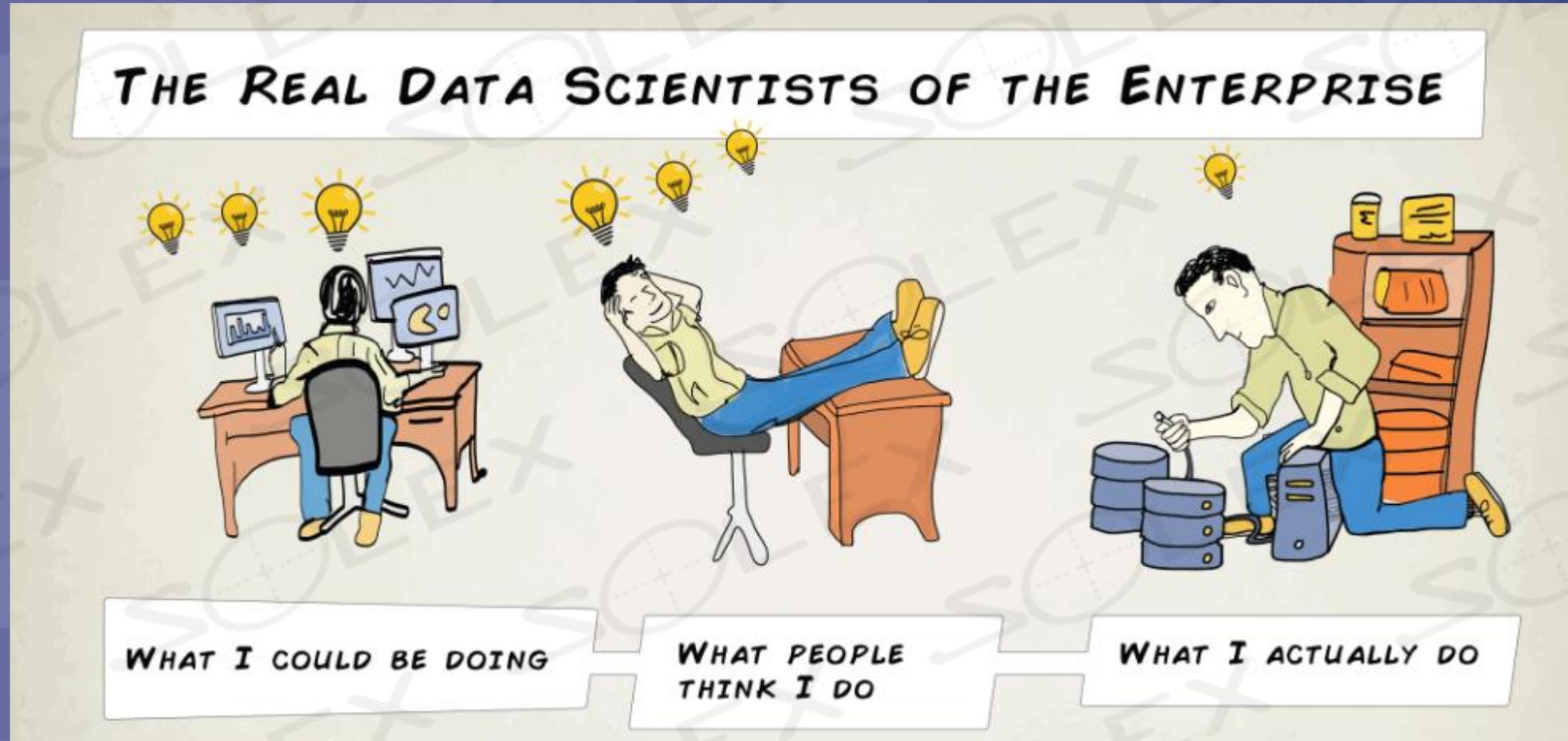
# Tareas I

- Descubrir con los expertos de negocio qué pretenden hacer, con qué datos se cuenta, qué es lo que no conocen.
- Proponer alternativas para resolver los problemas y ofrecer posibles usos de los datos con los que se cuenta.
- Llevar a cabo experimentos, definir hipótesis y realizar pruebas.
- Sumergirse en los datos de negocio, buscar los más importantes y pedir los que puedan proporcionar valor.
- Asear los datos, pelearse con ellos y dejarlos preparados para comenzar el análisis (70-80% del tiempo).
- Explorar los datos y jugar con ellos para dar respuesta a las preguntas planteadas y descubrir otros aspectos de interés que puedan ser modelizados.

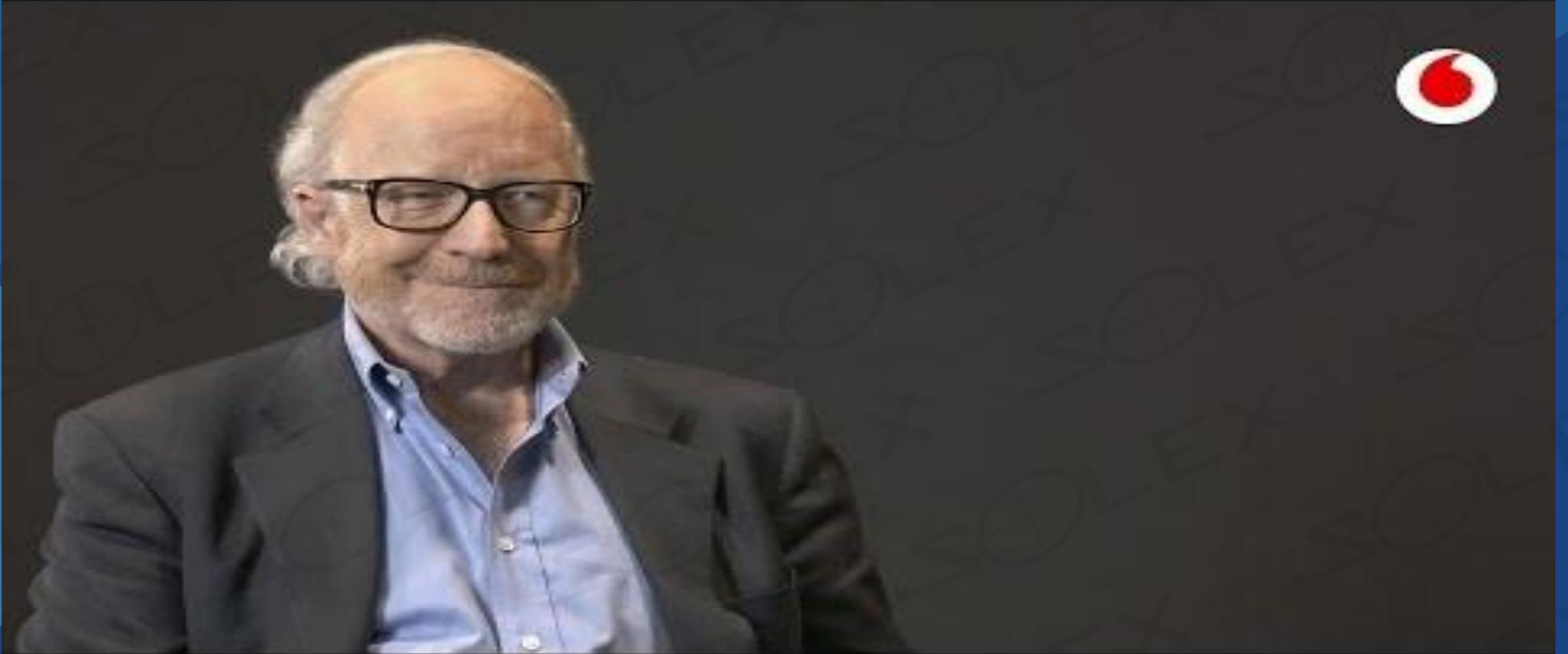
# Tareas II

- Probar diferentes modelos y algoritmos. Utilizar herramientas de Machine Learning y validar los modelos seleccionados.
- Analizar los resultados y evaluar las relaciones entre variables.
- Generar un esquema de aprendizaje que permita realimentar a los algoritmos elegidos.
- Llevar a producción los modelos y resultados basados en los datos, de este modo se finaliza el producto basado en datos y se ofrece la visualización de los resultados más destacados.
- Elaborar informes, que pueden ser interactivo.
- Resumir los principales resultados de valor conseguidos para la empresa y que permitan tomar las decisiones basadas en datos.

# Realidad — Calidad Datos



# Silicon Valley – Oportunidad Negocio



The background consists of a dark blue field filled with various-sized circles in shades of blue and teal. A faint, repeating watermark of the word 'SOLEX' is visible across the entire background. The word 'SOLEX' is written in a light blue, sans-serif font, with a small crosshair symbol positioned above the letter 'O'.

Gracias