

Data Science Basics

Fundamentos de Ciencia de datos

El impacto de los datos



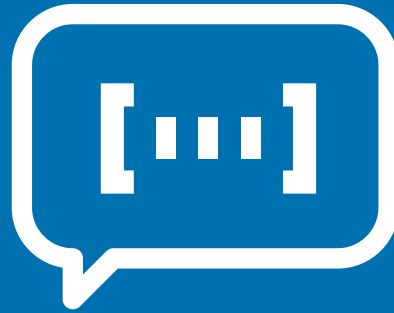
Source: ivyproschoool.com

Según SnapLogic, las compañías están **perdiendo \$140 mil millones cada año** en tiempo y recursos desperdiciados, esfuerzos duplicados y oportunidades perdidas debido a la **desconexión de datos**.

Las **empresas altamente automatizadas** tienen seis veces más probabilidades de lograr un **crecimiento de ingresos** superior al 15%.

Con una **sólida ingeniería de datos**, las empresas pueden acelerar la prestación de servicios a los clientes, mejorar la calidad, flexibilidad y escalabilidad.

Los **productos y servicios** se vuelven más innovadores y ágiles; y se logra mayor productividad y ahorro, generando un **"círculo virtuoso"**.



DEBATE

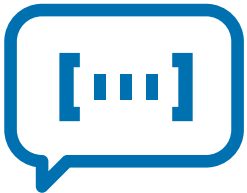
¿Cómo podemos procesar datos a escala para obtener valor?

OBJETIVO

Obtener visión sobre el procesado e integración de datos a escala.

INSTRUCCIONES

1. Piensa en el proceso que sigues en Amazon al comprar un producto.
2. **Únete** a tu compañero para responder estas **preguntas**:
 - ¿Qué datos se capturan? ¿En qué canales?
 - ¿Qué valor accionable les saca Amazon?
 - ¿Cómo se procesan los datos para sacar ese valor?
 - ¿Qué etapas/pasos debemos seguir?
3. Usa post-its para **capturar** vuestras ideas. Sintetiza 1 post-it para cada pregunta.
4. **Prepárate** para compartir vuestra visión con el resto de la clase.



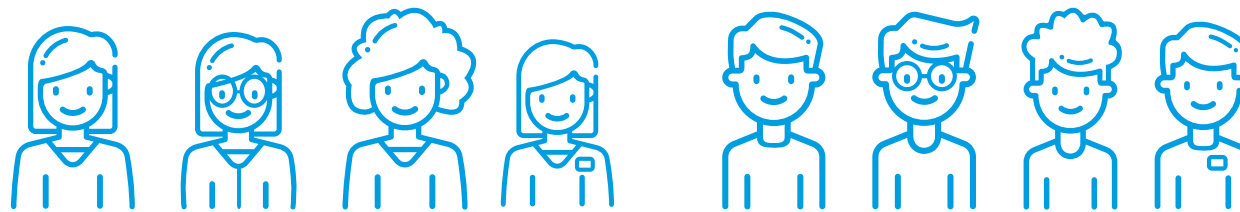
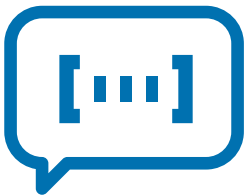
20 min

OBJETIVO

Comparte tu visión

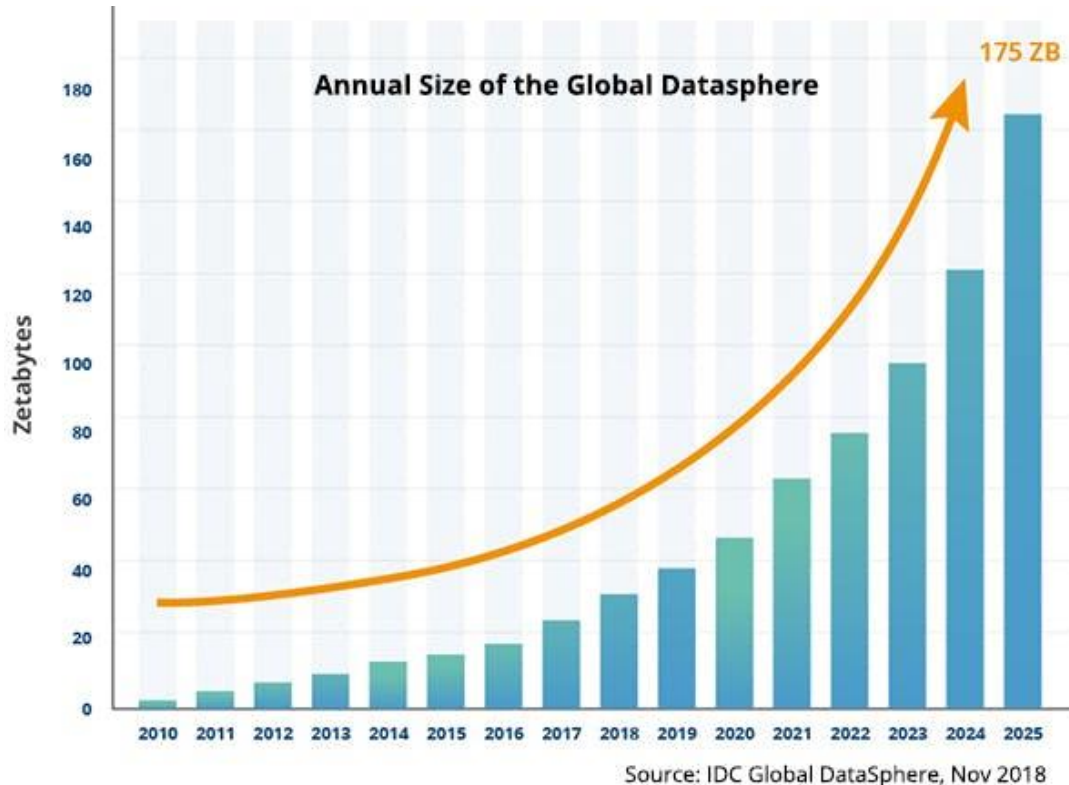
INSTRUCCIONES

1. **Comparte** tu visión con el resto de la clase.
2. Genera unas **conclusiones comunes**.



10 min

La era de los datos...



El universo digital de los datos para el 2025 alcanzará un volumen de **175ZB**.

2,5 quintillones de bytes de datos se crean cada día.

El **Análisis de datos** es una necesidad para obtener un conocimiento útil de los datos en bruto.

Surge una (no tan) nueva disciplina: **La Ciencia de datos**

y una nueva profesión: el **Científico de datos** y su demanda cada día va en aumento.

<https://www.statista.com/statistics/871513/worldwide-data-created/>

Ciencia de datos (DS: Data Science)

"La extracción de conocimiento útil de los datos para resolver problemas empresariales mediante un proceso sistemático con etapas bien definidas"
(Provost & Fawcett, 2013).

Aplicación de técnicas analíticas avanzadas y principios científicos para extraer información valiosa de los datos para la toma de decisiones comerciales, la planificación estratégica y otros usos.

La ciencia de los datos nace del método científico y de la evolución del análisis de datos.

"Disciplina que crea sistemas y algoritmos para descubrir conocimiento, detectar patrones, generar información útil y realizar predicciones a partir de datos a gran escala"
(Molina-Solana, Ros, Dolores Ruiz, Gomez Romero, & Martin-Bautista, 2017).

"Extracción de conocimiento accionable directamente de los datos a través de un proceso de descubrimiento o formulación y prueba de hipótesis"
(NIST, 2015).

Ciencia de datos (DS: Data Science)

La ciencia de datos trata de utilizar datos para **encontrar respuestas o resolver problemas**.

Implica **recopilar, analizar e interpretar** grandes cantidades de datos para obtener información y tomar decisiones fundamentadas.

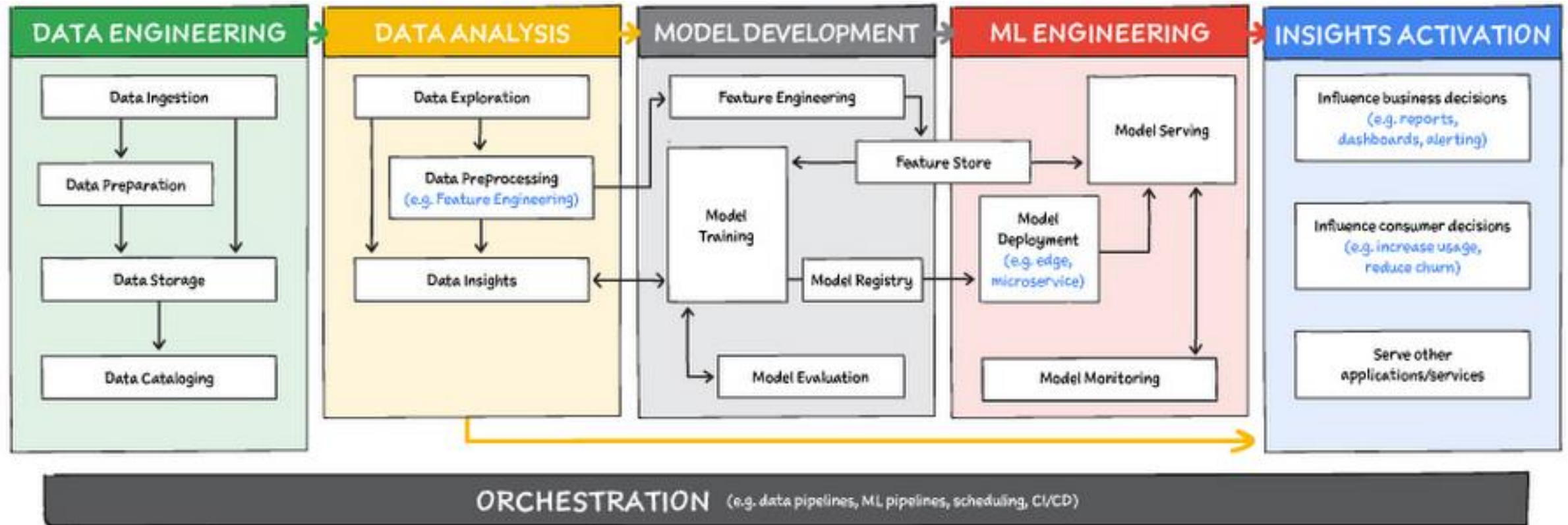
En términos prácticos, la ciencia de datos **combina técnicas de diversos campos**, como la estadística y la informática , con el conocimiento del dominio para dar sentido a los datos.

Implica tareas como **limpiar y preparar datos, explorar patrones y relaciones** dentro de los datos y crear modelos predictivos para pronosticar resultados futuros o identificar tendencias.

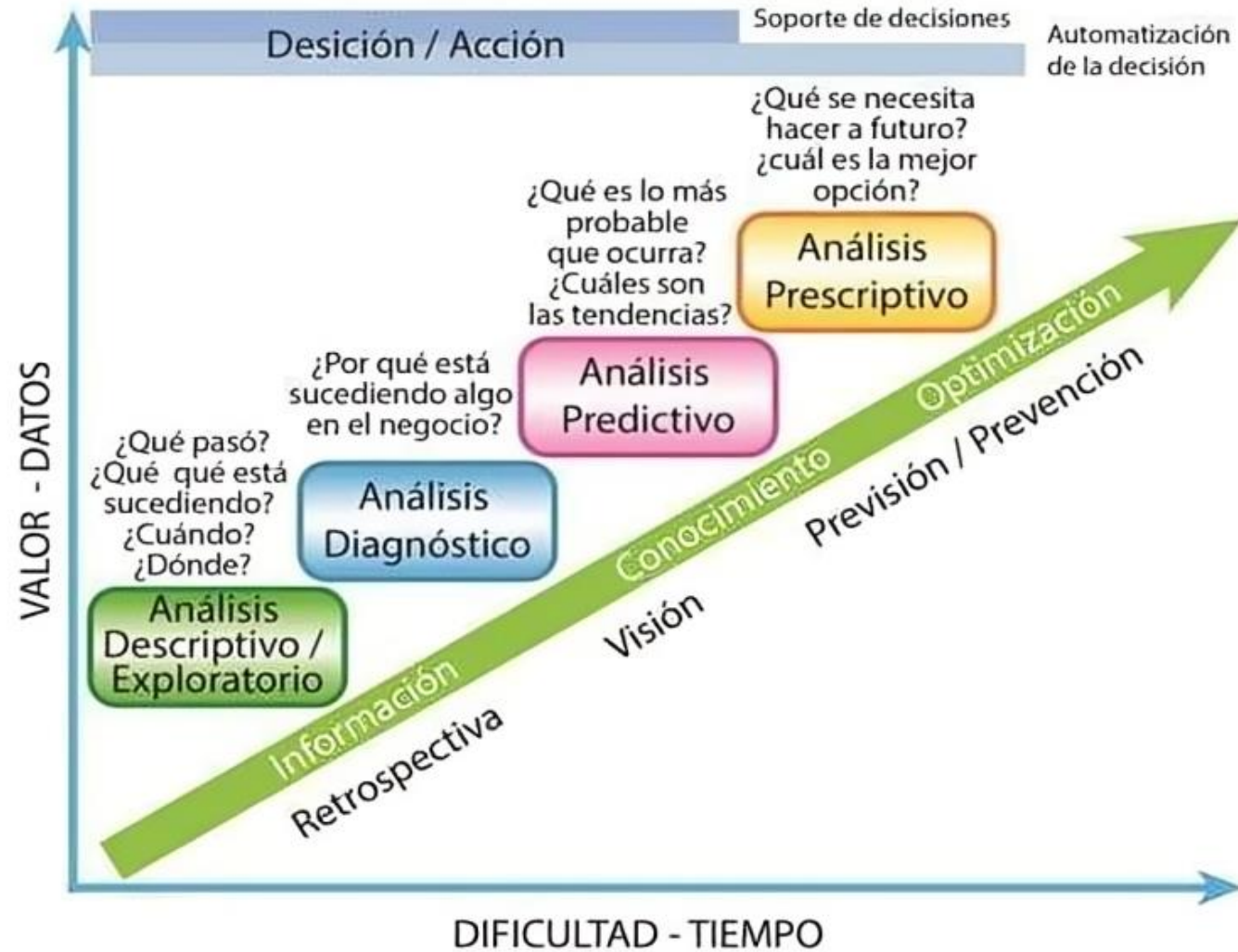
Al aprovechar el poder de los datos, los científicos de datos pueden **descubrir información valiosa** que ayude a las empresas a mejorar sus operaciones, desarrollar nuevos productos o comprender mejor a sus clientes.

En última instancia, la ciencia de datos nos **permite aprovechar la gran cantidad de información disponible** en el mundo moderno para impulsar la innovación y tomar decisiones más inteligentes.

Etapas del DS

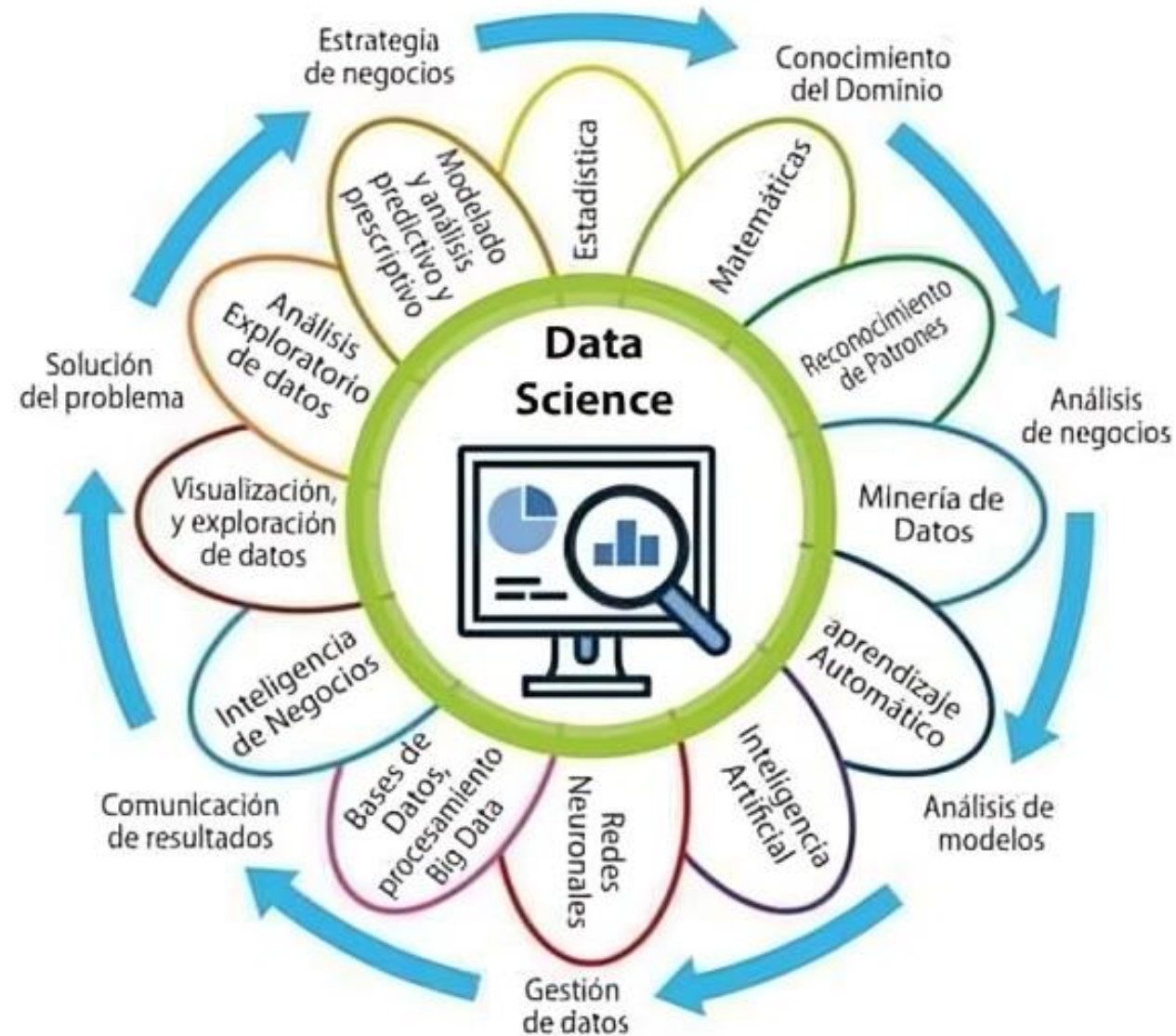


Tipos de análisis de datos

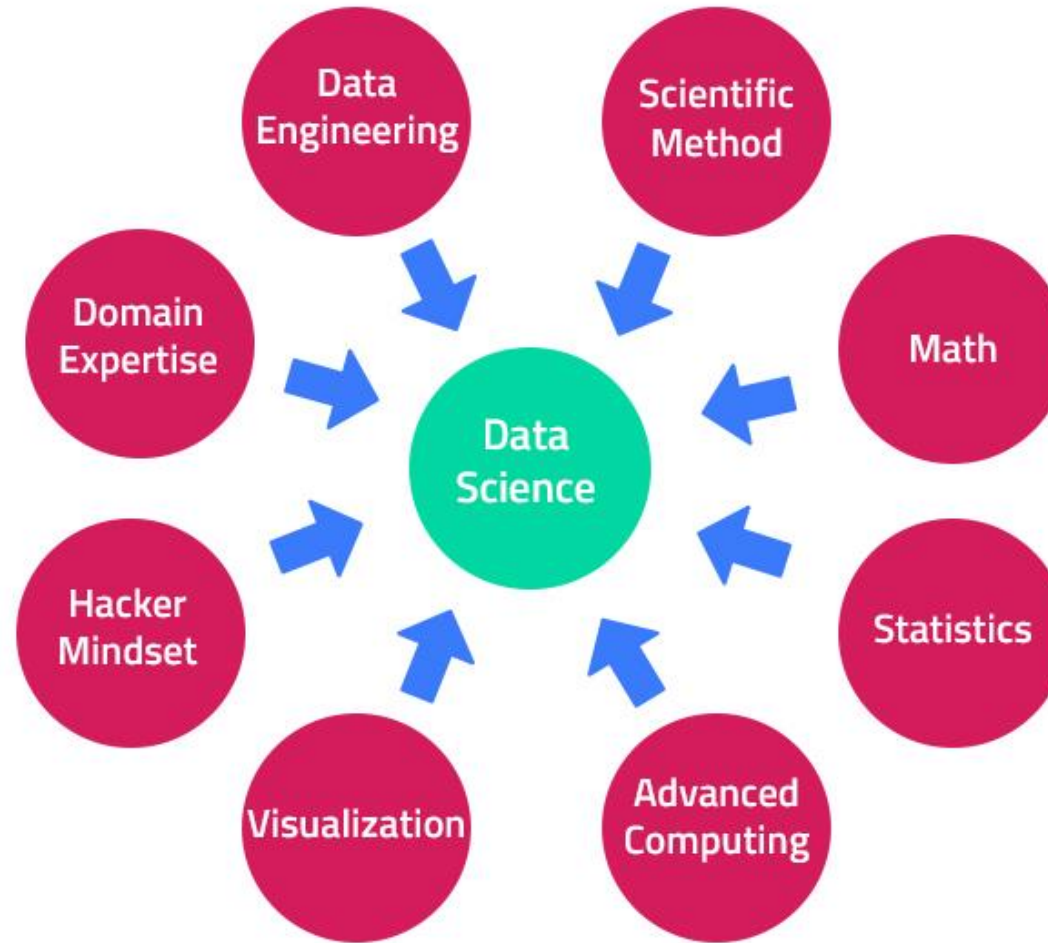


(Gartner, 2012; Loury, 2014; National Academi of Science, 2017; Sivarajah, 20t7)

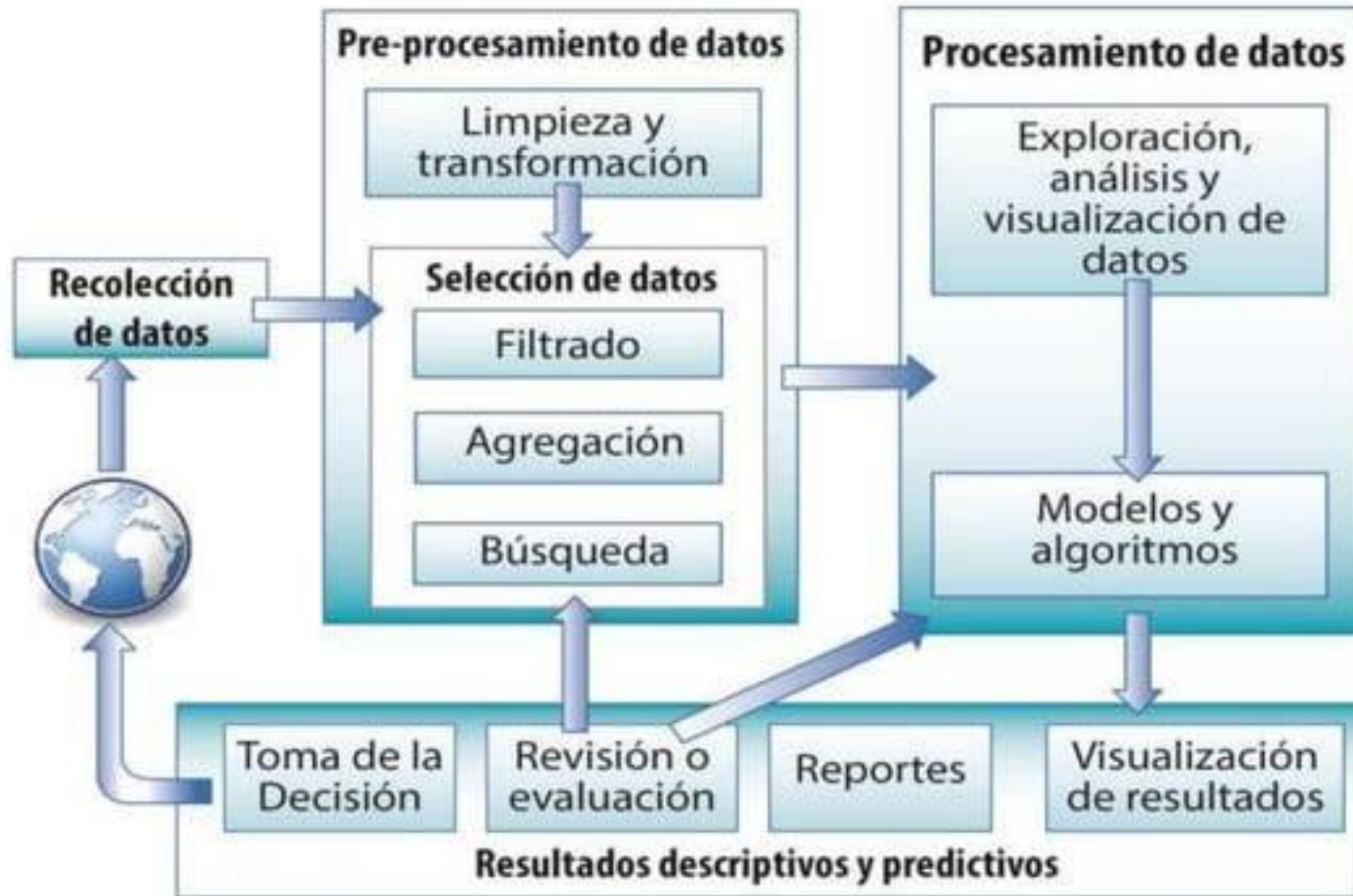
Disciplinas relacionadas con DS



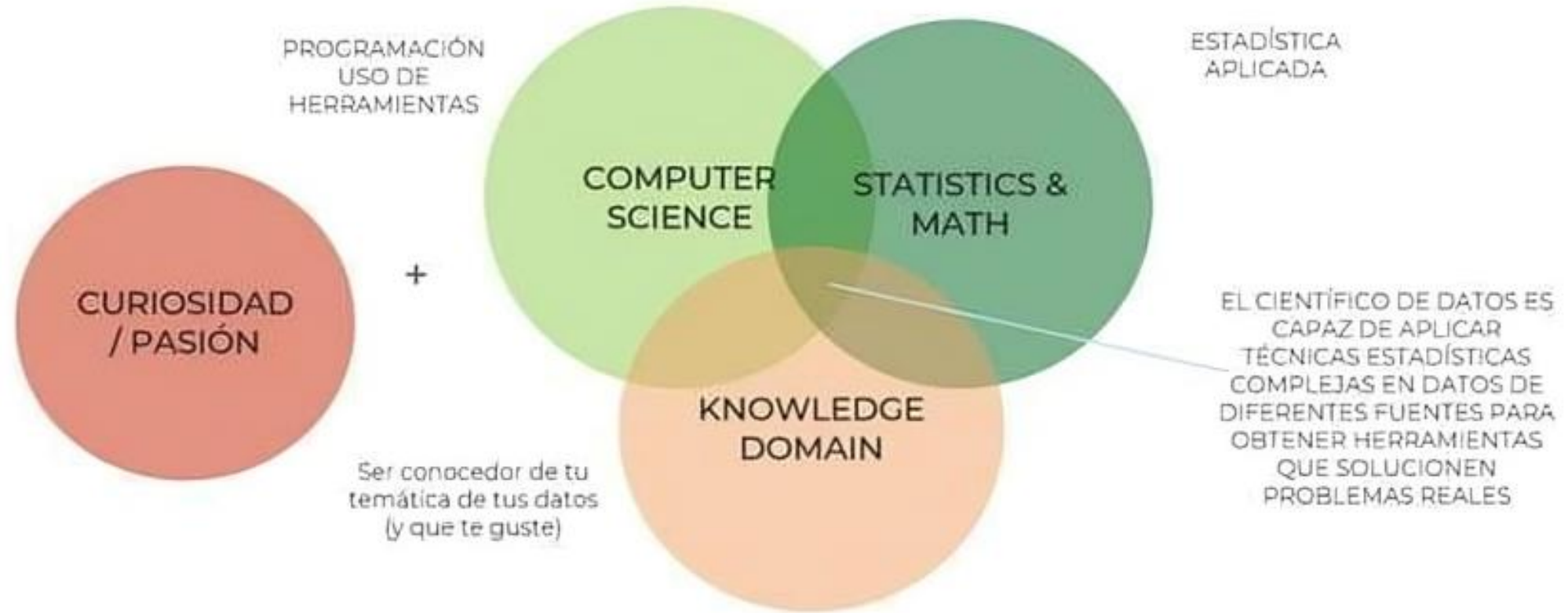
Disciplinas relacionadas más destacadas



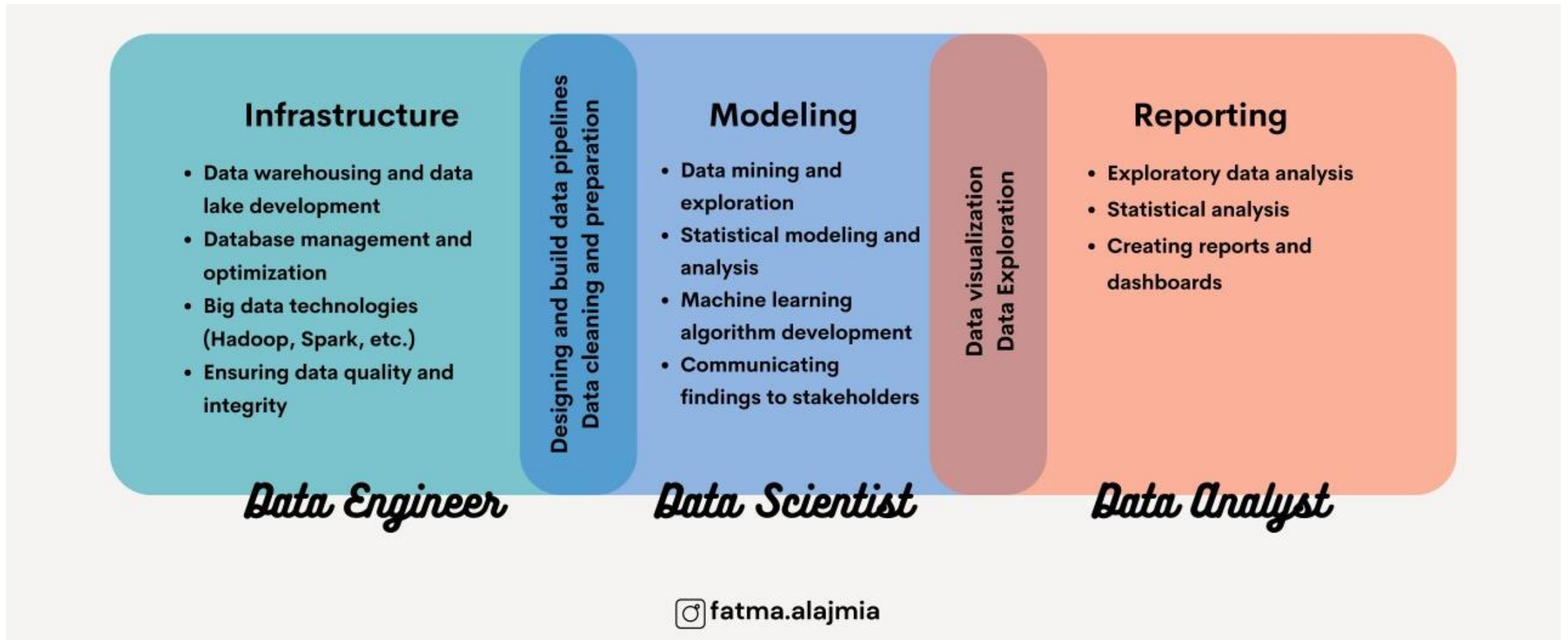
Ciclo de vida de DS



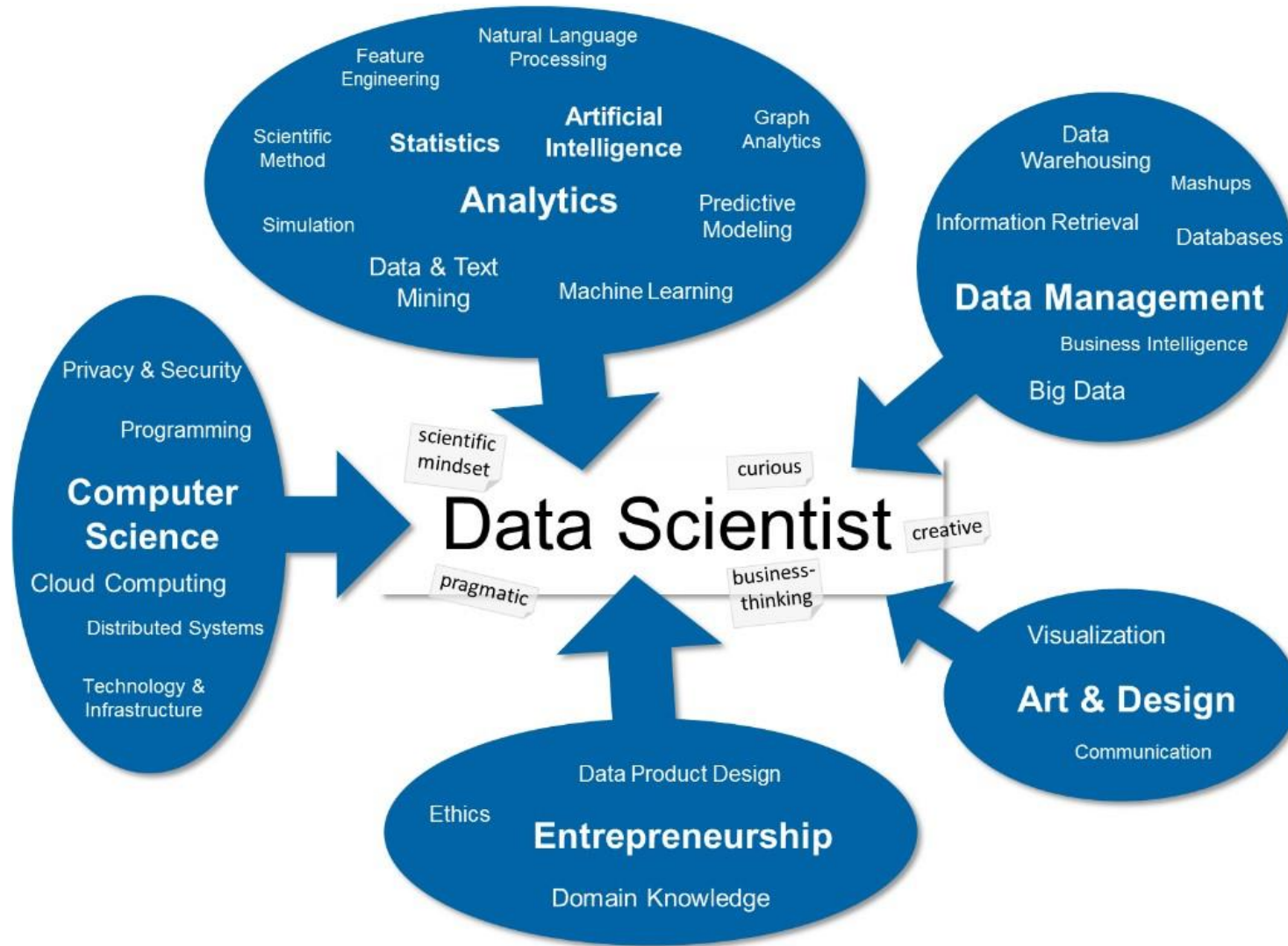
Científico de datos: ramas de conocimiento



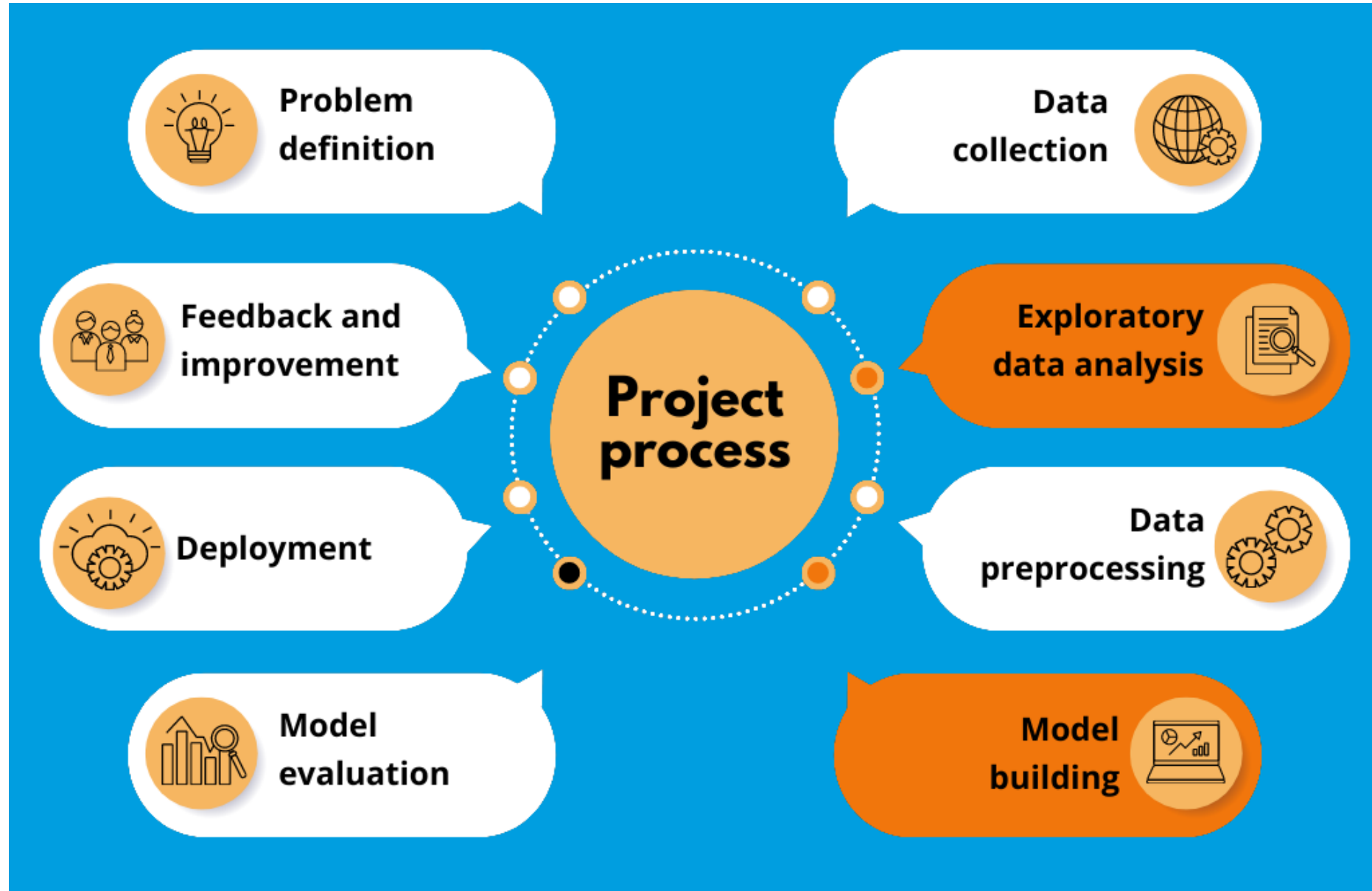
Científico de datos vs otros roles



Científico de datos



Proyecto DS



Herramientas DS



Herramientas DS





Next steps



We would like to know your opinion!

Please, let us know what you think about the content.
From Netmind we want to say thank you, we appreciate time
and effort you have taking in answering all of that is
important in order to improve our training plans so that you
will always be satisfied with having chosen us
quality@netmind.es

Thanks!

Follow us:

