

# Data Governance & Strategy

*La culturización del dato en las organizaciones data-driven*

Madrid, 14 de noviembre de 2019  
Tony Rodríguez

# Qué necesitamos para tener cultura del dato

## SDG Data Culture Framework

El proceso de culturización del dato puede enmarcarse en cinco áreas que deben abordarse exhaustivamente. Los datos aprovechan la tecnología para brindar capacidades analíticas, por lo que pueden ser beneficiosos para las personas dentro de la organización. El círculo finaliza cuando las personas de la organización poseen el hábito de tomar decisiones basadas en datos..

‘Asegurar que existan los mecanismos organizativos adecuados para institucionalizar las **decisiones basadas en datos**’

‘Buscando, motivando y reteniendo el **talento** con los **skills necesarios** para los retos de Data & Analytics’

‘Obtener **insights** de los datos, y hacer de éstos **acciones** que promuevan los objetivos de la organización’

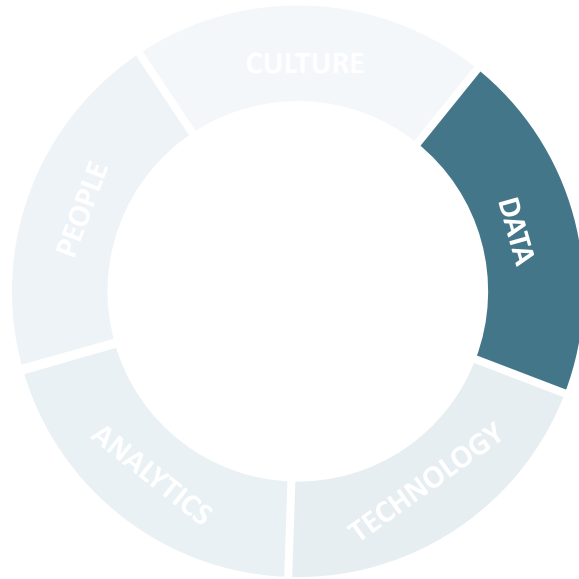


‘**LOS DATOS** a través de la organización deben representar la **mejor versión de la verdad**, y deben ser verídicos para la correcta toma de decisiones’

‘Entender la **tecnología** como un **acelerador** para lograr un mejor rendimiento de la analítica en los negocios’

# Qué necesitamos para tener cultura del dato

Data Analytics



*‘Los Datos a través de la organización deben representar la **mejor versión de la verdad**, y deben ser verídicos para la correcta toma de decisiones’*



# Qué necesitamos para tener cultura del dato

## Data Analytics

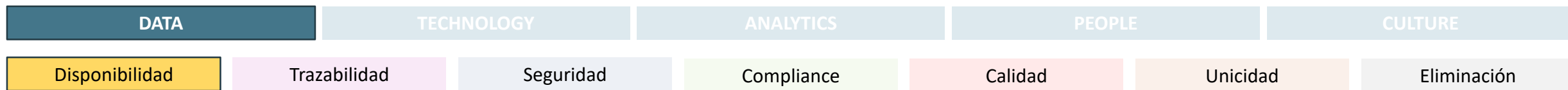


Para que una compañía **confíe plenamente en sus datos** y los use como un generador de valor, tanto el negocio como la TI deben centrarse en dominar los "ejes de datos".



# Qué necesitamos para tener cultura del dato

## Data Analytics



### ¿Qué es Disponibilidad de los Datos?

El término disponibilidad de los datos se refiere a la **habilidad de asegurar que los datos sean siempre accesibles** cuando y dónde se necesiten dentro de la infraestructura de la organización, y que estos **datos sean conocidos en toda la organización**.

### ¿Por qué Disponibilidad de los Datos?

La realidad es que **los datos que no son accesibles cuando se les necesita o que se desconoce su significado no tienen valor**. Por este motivo es crucial mantener un alto nivel de disponibilidad de los datos para garantizar el éxito continuado de la organización.

### Áreas de enfoque de la Disponibilidad de los Datos

#### Modelo centralizado de almacenamiento de los datos



Evitar **silos de información** por departamento y compartir toda la información en un repositorio único. Ofrece la ventaja de que todos los departamentos o unidades de negocio saben donde localizar los datos guardados.

*Ambas opciones...*

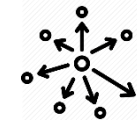


#### Glosario de Negocio

El Glosario de Negocio permite a las personas construir y gestionar un **vocabulario de negocio común** y ponerlo a la disponibilidad de toda la organización, **independientemente del modelo de almacenamiento**.

*Ambas opciones...*

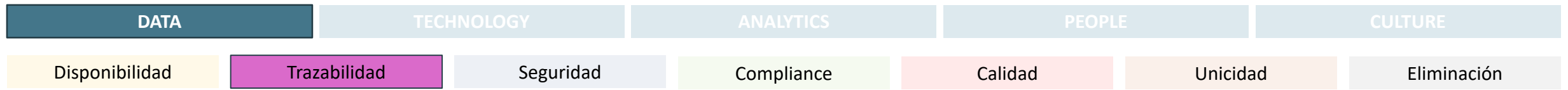
#### Modelo descentralizado de almacenamiento de los datos



Aprovechar las sinergias de una **cultura organizacional flexible**, en la que cada departamento o unidad de negocio puede gestionar o almacenar sus propios datos.

# Qué necesitamos para tener cultura del dato

## Data Analytics



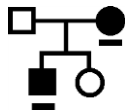
### ¿Qué es la Trazabilidad de los Datos?

La trazabilidad de los datos describe **orígenes de los datos, movimientos, características y calidad a lo largo del ciclo de vida de los datos**. Este término también describe qué pasa con los datos al pasar por los diferentes procesos.

### ¿Por qué la Trazabilidad de los Datos?

La trazabilidad de los datos puede ayudar a analizar como la información es usada y rastrear partes clave de la información que sirven un objetivo concreto. Por ejemplo, cómo se han recogido la información de ventas y qué rol podría jugar en nuevos procesos, nuevos o mejorados, que pongan los datos a través de diagramas de flujo adicionales dentro de un negocio.

### Áreas de enfoque de la Trazabilidad de los Datos



#### Trazabilidad de negocio

El quién, qué, dónde, por qué y como de los datos de negocio. Los informes muestran una visión simplificada del linaje que destaca la transformación y agregación de datos que se necesita por parte de negocio.



#### Análisis de Impacto

Identifica las **consecuencias potenciales derivadas de un cambio** y evalúa los riesgos asociados a este cambio.

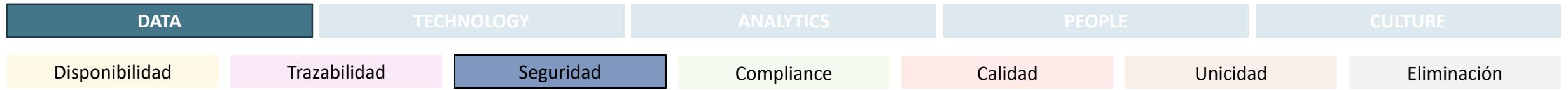


#### Trazabilidad Técnica

Muestra el **flujo físico de los datos** a través de aplicaciones subyacentes, servicios y almacenes de datos, enfocándose al desarrollo, actualización y mantenimiento de la Arquitectura de Datos.

# Qué necesitamos para tener cultura del dato

## Data Analytics



### ¿Qué es la Seguridad de los Datos?

La **seguridad de los datos** es un conjunto de estándares y tecnologías que protegen los datos de la destrucción, modificación o divulgación, ya sea **intencionada o accidental**. Puede ser aplicada utilizando un amplio rango de técnicas y tecnologías, incluyendo controles administrativos, seguridad física, controles lógicos, estándares organizacionales y otras técnicas de salvaguarda para limitar el acceso a usuarios o procesos no autorizado o maliciosos.

### ¿Por qué la Seguridad de los Datos?

Organizaciones de todo el mundo están invirtiendo fuertemente en capacidades de ciberdefensa de las tecnologías de la información (TI) **para proteger sus activos críticos** (proteger una marca, capital intelectual o información de clientes, o proporcionar controles para infraestructura crítica).

### Áreas de enfoque de la Trazabilidad de los Datos



#### Gestión del Control de Acceso

**Monitorizar, controlar y evaluar** los accesos y las autorizaciones en el sistema de los usuarios.



#### Enmascaramiento de los datos

El enmascaramiento de áreas específicas de datos puede **protegerlos de divulgación a fuentes maliciosas**, y personal interno que no debería ver o usar los datos.



#### Encriptación de los datos

Aplica un código a cada dato individual y **no permite el acceso a los datos sin una clave autorizada**.

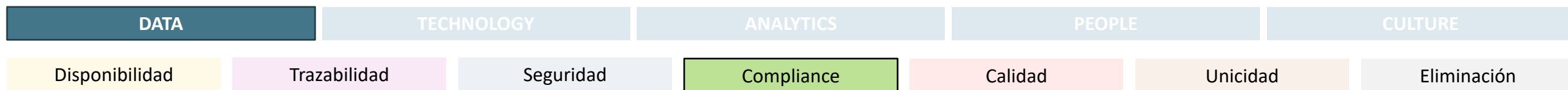


#### Resiliencia de los datos

Mediante la creación de una **copia de seguridad de los datos** las organizaciones pueden recuperarlos en caso de borrado o corrupción accidental, o de robo en una brecha de seguridad.

# Qué necesitamos para tener cultura del dato

## Data Analytics



### ¿Qué es la Conformidad de los Datos?

El cumplimiento representa todas las operaciones que se pueden realizar potencialmente con datos para **estar conforme con las pautas o especificaciones establecidas**. En este sentido, la explotación de los datos debe estar alineada con las pautas internas y externas a la organización.

### ¿Por qué la Conformidad de los Datos?

El término Conformidad de los Datos gana relevancia en **industrias altamente reguladas** como la farmacéutica. Por otra parte, los datos ofrecen constantes oportunidades de **evolucionar modelos de negocio existentes**, sin embargo, las regulaciones que gobiernan estas industrias tienen que tenerse en cuenta con el fin de **evitar costes económicos o mala reputación**.

### Áreas de enfoque de la Conformidad de los Datos

#### Compliance interno



Operaciones analíticas de Datos tienen que respetar los **principios, políticas y procedimientos internos del Gobierno de los Datos**. Hacer que la organización respete las normas internas agrega valor a un programa de gobernanza de datos (si existe) y ayuda a promocionar la cultura de los datos.

#### Compliance Externo



El **cumplimiento normativo con la GDPR** cuando los datos son usados en escenarios analíticos avanzados es imprescindible, y más aún si se las técnicas de “machine learning” se aplican a los datos.

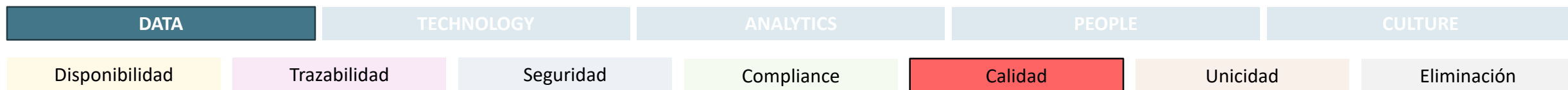


La **Conformidad con los estándares de la industria farmacéutica** y las guías regulatorias en relación con el tratamiento de los datos también es necesario.



# Qué necesitamos para tener cultura del dato

## Data Analytics



## ¿Qué es la Calidad de los Datos?

La calidad de los datos se refiere a **las condiciones bajo las cuales los datos se ajustan** a los usos previstos en las operaciones y decisiones de negocio. El nivel de calidad aceptado depende en gran medida de los requerimientos de negocio y de los umbrales establecidos.

## ¿Por qué la Calidad de los Datos?

Una de las principales **condiciones** para que una empresa **adopte una Cultura de Análisis de Datos** es el **nivel de confianza** que las personas tienen en los datos. Una vez que la calidad de los datos supera un umbral determinado, **las personas comienzan a confiar en los datos** y, en consecuencia, pueden realizar decisiones basadas en la información.

## Áreas de enfoque de la Calidad de los Datos



### Completitud:

Aplicar controles de calidad de los datos con el fin de asegurar la **completitud de los datos**. El contexto de negocio es relevante dado que permite saber cuando falta información.



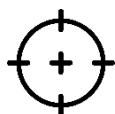
### Estandarización:

Aplicar controles de calidad de los datos con el fin de asegurar la **estandarización del formato de los datos** como una función de los requerimientos de negocio.



### Consistencia:

Aplicar reglas de calidad de los datos de negocio con el fin de asegurar consistencia entre los diferentes sets de datos. La **lógica de negocio** es esencial para controlar esta cuestión.



### Precisión:

Aplicar controles de calidad con el fin de asegurar que **los valores de los datos son correctos**, mejorando la obsolescencia o otros aspectos incorrectos de los datos.



### Duplicación:

Aplicar controles de calidad de los datos con el fin de **asegurar valores únicos para cada concepto** o objeto representado por los datos.

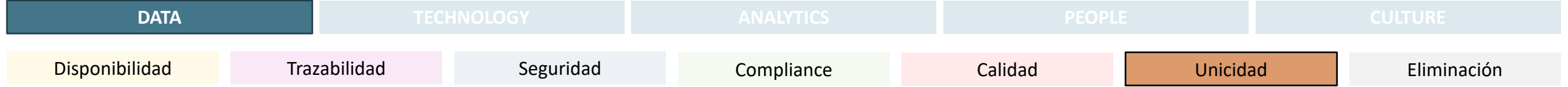


### Integridad:

Aplicar controles de calidad de los datos, dependiendo del set de datos, con el fin de asegurar que **todos los valores están relacionados con otros valores**, respetando la integridad de la información.

# Qué necesitamos para tener cultura del dato

## Data Analytics



### ¿Qué es la Unicidad de los Datos?

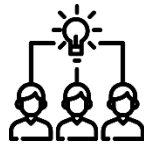
La Unicidad es un componente de la calidad de los datos. Define el aspecto más valioso de las iniciativas de integración de los datos en las compañías: Diferentes **valores de los datos** alrededor del mismo concepto deben estar guardados (y preparados para ser analizados) sólo una vez (datos maestros), y ésta debe representar la **mejor versión de la verdad**. La unicidad también aplica a los valores de los metadatos.

### ¿Por qué la Unicidad de los Datos?

En un contexto de negocio en el que hay una gran cantidad de datos, y diferentes fuentes de información tratando potencialmente con los mismos valores, es necesario controlar la unicidad con el fin de permitir la **confianza de la compañía en un único valor de la verdad**.

### Áreas de enfoque de la Unicidad de los Datos

#### Unicidad en el valor de los metadatos



Las personas de la compañía necesitan utilizar un **mismo lenguaje**, es decir, que un único valor de los metadatos haga referencia a un único concepto, evitando sinónimos que puedan llevar a malentendidos.

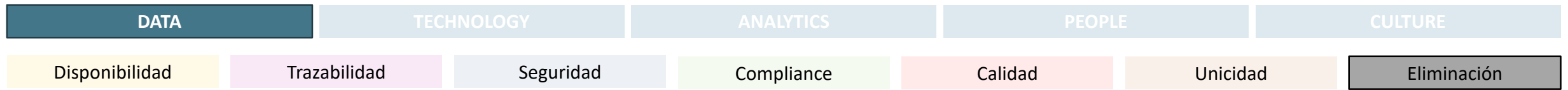
#### Unicidad en el valor de los datos



Las personas de la compañía necesitan tomar decisiones sobre un **único valor de los datos referenciando a un único concepto**, evitando diferentes versiones de un mismo concepto. Lograr esto aumentará la consistencia entre los departamentos, y el intercambio de datos será más fácil.

# Qué necesitamos para tener cultura del dato

## Data Analytics



### ¿Qué es la Eliminación de los Datos?

La Eliminación de los Datos hace referencia a la **congelación temporal** de los datos por la necesidad de prescindir de ellos para su análisis o la **eliminación permanente** de dichos datos cuando se vuelven obsoletos y no pueden proporcionar ningún valor en su explotación.

### ¿Por qué la Eliminación de los datos?

La Eliminación de los Datos se realiza en la **fase final del ciclo de vida de los datos**. Normalmente, cuando el consumo de los datos se convierte en obsoleta, esta información representa un coste en términos de almacenamiento. Con el fin de evitar la obsolescencia de los datos, es importante tener en cuenta que los datos deberían ser eliminados, ya sea temporalmente o permanentemente.

### Áreas de enfoque de la Eliminación de los Datos

#### Congelar los Datos Temporalmente



Es posible que haya la necesidad de dispensar de manera temporal determinados datos. Este hecho se llama **congelación de los datos** dado que no implica una eliminación completa.

#### Enterrar los Datos Permanentemente

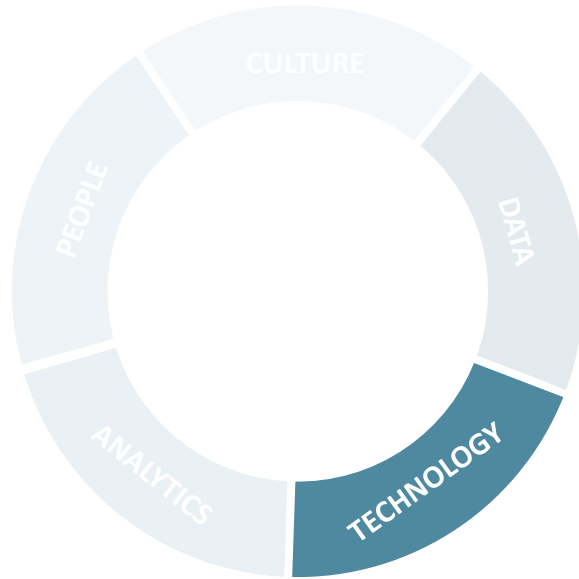


Cuando los datos paran de aportar valor a la organización y no han sido utilizados en un cierto periodo de tiempo, son **removidos permanentemente** del sistema.

# Qué necesitamos para tener cultura del dato

*Data Analytics*

---

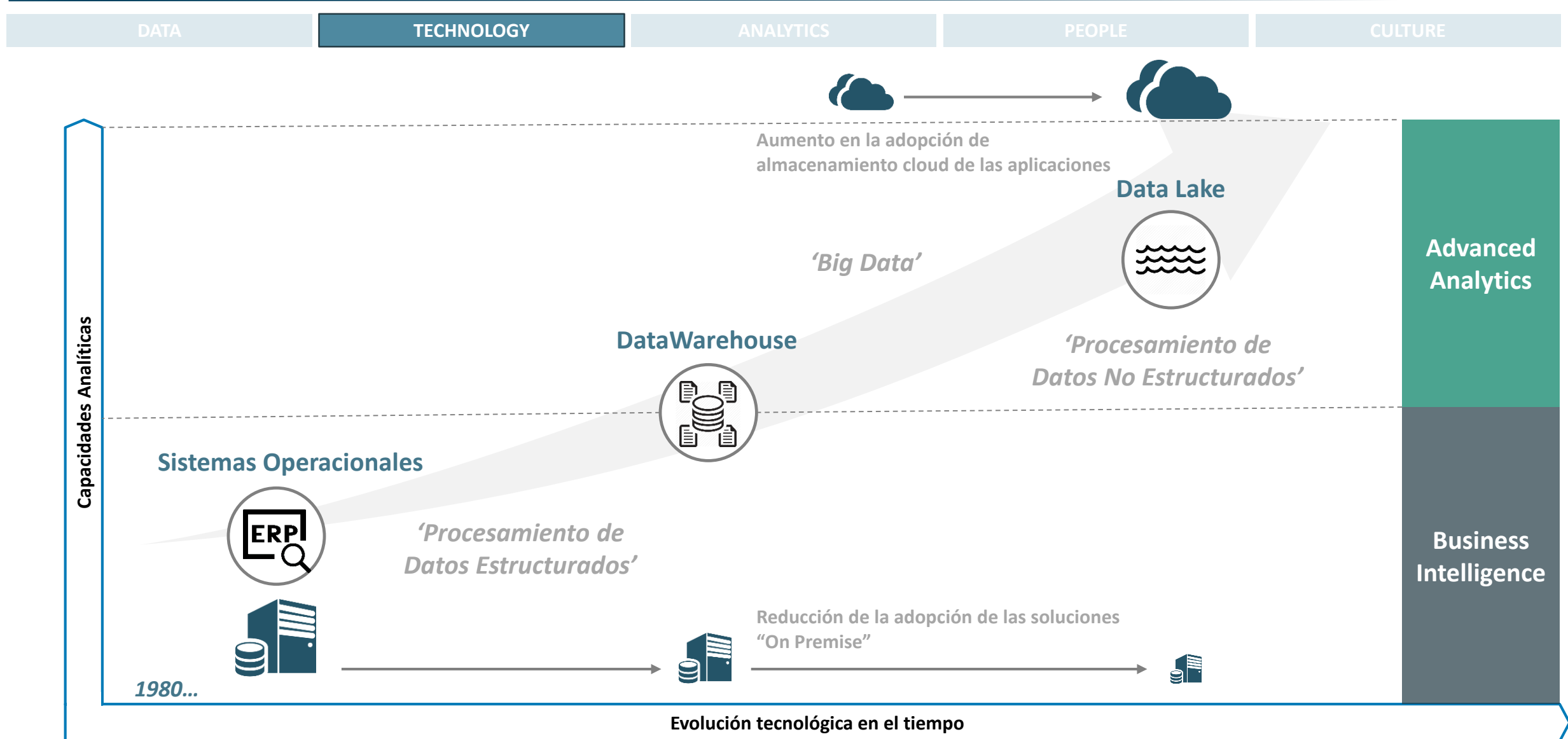


*‘Entender la **tecnología** como un **acelerador** para lograr un mejor rendimiento de la analítica en los negocios ’*



# Qué necesitamos para tener cultura del dato

## Data Analytics

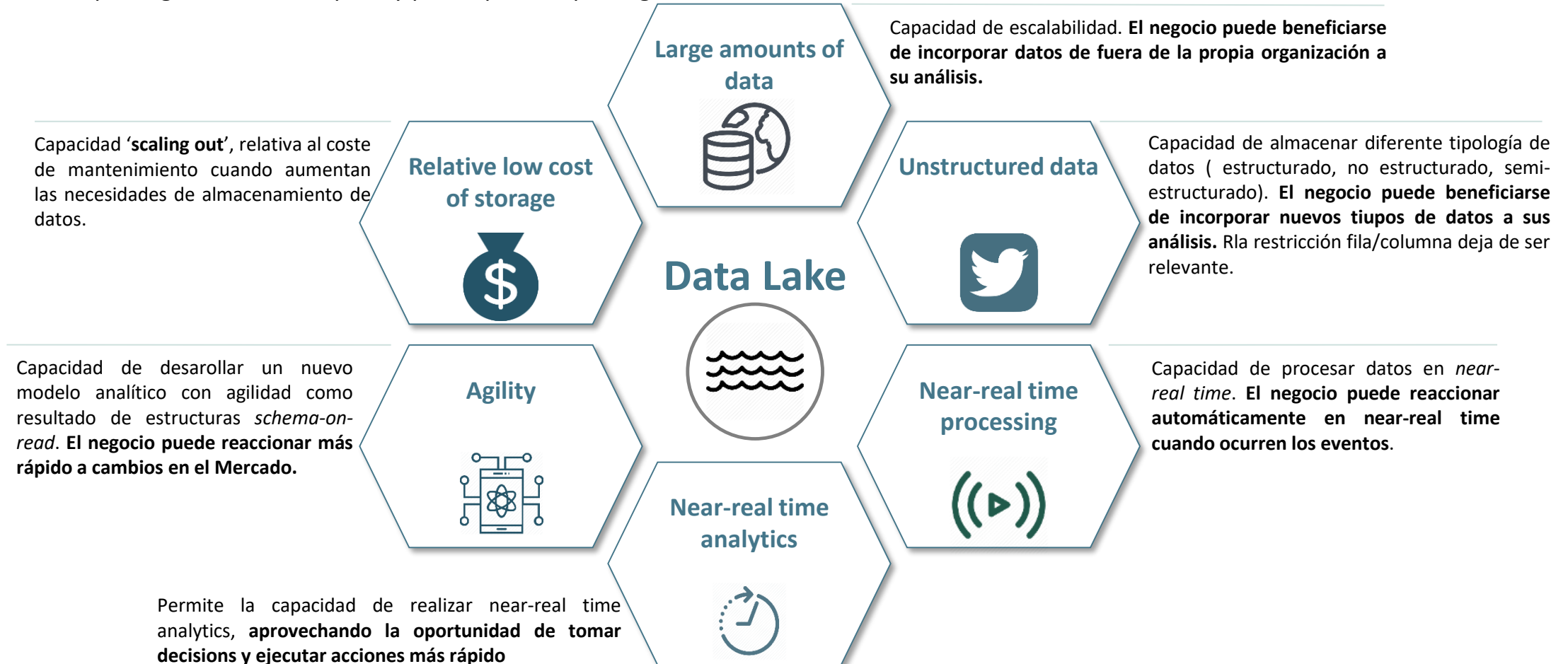


# Qué necesitamos para tener cultura del dato

## Data Analytics

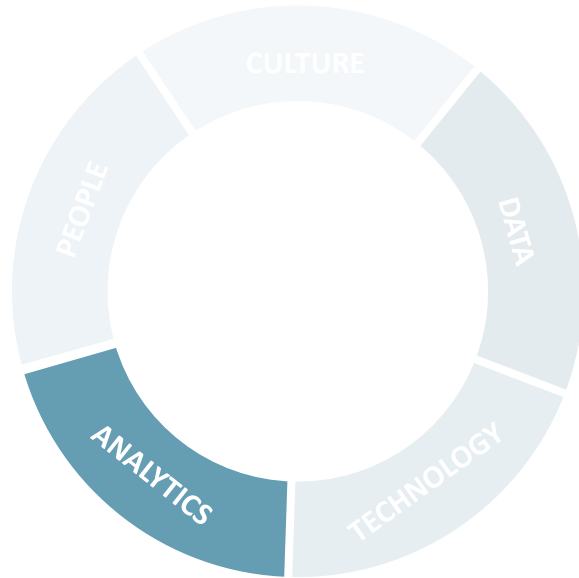


Un Data Lake es el centro de almacenamiento de todos los datos de la Compañía, está compuesto por una copia de los datos operacionales, un área en la que se procesan, transforman, se les aplica calidad, y otra área en la que se agrupan para que las áreas de negocio los puedan analizar de manera confiable para funciones de reporting, análisis descriptivo y prescriptivo, reporting, etc.



# Qué necesitamos para tener cultura del dato

Data Analytics



*‘Obtener **información** de los datos, y convertir esas ideas en **acciones** que promuevan los objetivos de la organización’*

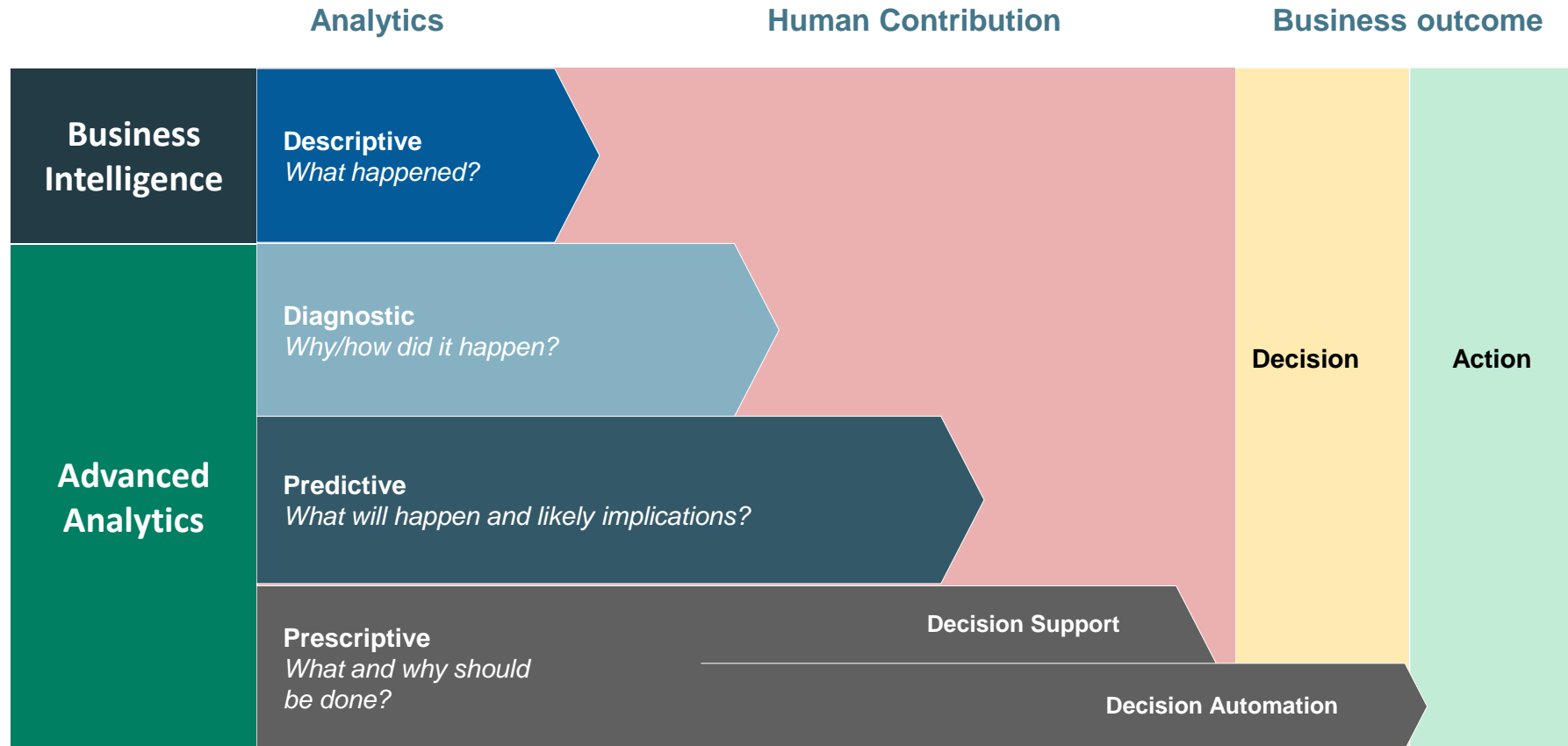


# Qué necesitamos para tener cultura del dato

## Data Analytics



Analytics es el **descubrimiento, interpretación y comunicación** del significado de los **patrones en los datos**, y el proceso de aplicar esos patrones para una **toma de decisiones efectiva**.

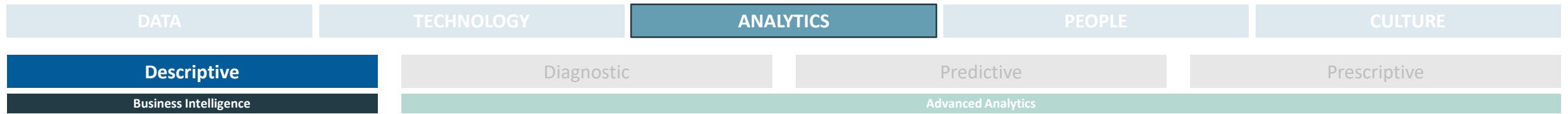


Source: Gartner



# Qué necesitamos para tener cultura del dato

## Data Analytics



La Analítica Descriptiva proporciona una imagen ajustada de lo que ha ocurrido en un negocio y cómo difiere esta situación respecto otros periodos. Esta analítica se basa en la interpretación humana como *input* para generar una decisión o una nueva información.

## Analítica Descriptiva

Qué ocurrió?



Analítica que ayuda a comprender cómo han ocurrido las cosas.



### Reporting

Provisión regular de información para los decisores dentro de la organización para apoyarlos en sus actividades.



### Consulta

Análisis básico de datos para responder preguntas puntuales apuntando a una table de base de datos o combinación de tablas.



### Visualización

Es la representación gráfica de la información mediante elementos como gráficos, mapas, infografías, etc.

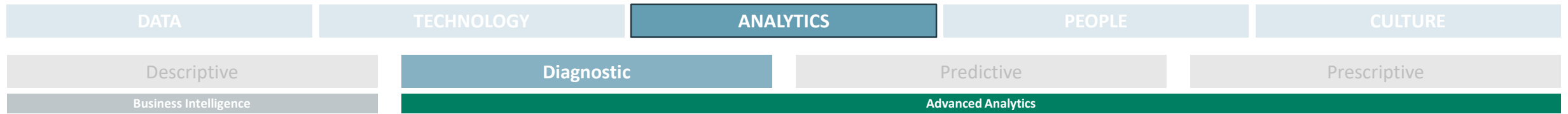


### Estadística Descriptiva

Resume el conjunto de datos obtenidos por un pequeño número de valores descriptivos, (promedio, mediana, varianza, la desviación típica, etc). Estas medidas descriptivas ayudan a conocer las principales propiedades de los datos observados.

# Qué necesitamos para tener cultura del dato

## Data Analytics



**Diagnostic Analytics se centra en por qué y cómo ocurrieron los eventos.** A este nivel, las herramientas analíticas y procesos, evolucionan hacia habilitar una mayor comprensión de los datos que **junto con la Analítica Descriptiva**, es lo que proporciona obtener una situación de diagnóstico.

Las funciones de *diagnostic analytics* se resumen en 3 categorías:

### Por qué/Cómo ocurrió ?



*Examinar las causas de resultados pasados para comprender el impacto en el negocio.*

### Identificar anomalías

Basado en los resultados del Análisis Descriptivo, los analistas identifican áreas que requieren mayor profundidad analítica que no pueden ser contestadas con un simple vistazo a los datos.

### Profundizar en la Analítica (discovery)

Los analistas deben identificar las fuentes de datos que les ayudarán a explicar las anomalías. A menudo, este paso puede requerir que los analistas necesiten incorporar datos de otras aplicaciones (internas / externas) para comprender los motivos de las situaciones.

### Determinar relaciones causales

Identificación de relaciones ocultas de impacto entre variables y eventos. Funcionalidades como probabilidad teórica, análisis de regresión, filtrado en series temporales, son necesarias para descubrir nuevos impactos en el negocio.

# Qué necesitamos para tener cultura del dato

## Data Analytics



**Predictive analytics** es una rama de Advanced Analytics que se utiliza para **realizar predicciones sobre eventos futuros desconocidos**. Se basa en datos históricos y diferentes técnicas analíticas.

### Predictive Analytics

*Qué puede ocurrir y sus implicaciones?*



*Analítica que ayuda a ajustarse al future.*

**Datos Estructurados**  
(principalmente datos internos e históricos.)

**Datos No Estructurados**  
(principalmente datos externos y Big Data.)

#### Modelo Predictivo

##### Estadística Multi-variante

Incluye observación y análisis de más de una variable concurrentemente.

##### Análisis de Regresiones

Tipo de análisis multi-variable. Permite examinar las relaciones entre variables.

##### Forecasting

Mide los datos recopilados a lo largo del tiempo para identificar tendencias.

##### Pattern Matching

Acto de verificar una secuencia dada para la presencia de los constituyentes de algún patrón.

##### Machine Learning

Aplicación de Inteligencia Artificial (IA) que Brinda la capacidad de aprender y mejorar automáticamente en base a la experiencia.

#### Business outcome

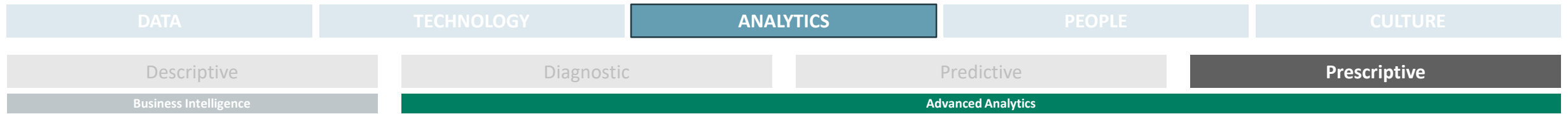
Qué ocurrirá ?



Decisión de Negocio

# What do we need to adopt Data Analytics

## Data Analytics



Prescriptive Analytics sugiere decisiones en base a los resultados obtenidos durante la Analítica Descriptiva y Predictiva.

### Prescriptive Analytics

*Qué se debería hacer?*



*“Si tenemos datos, miremos los datos. Si todo lo que tenemos son opiniones, vamos con la mía” Jim Barksdale, CEO of Netscape*

**Datos Estructurados**  
(principalmente datos internos e históricos.)

**Datos No Estructurados**  
(principalmente datos externos y Big Data.)

#### Descriptive analytics

Reporting Estándar

Estadística Descriptiva

#### Predictive analytics

Estadística Multi-variante

Análisis de Regresión

Forecasting

Pattern Matching

Machine Learning

#### Business outcome

**Motores de Sugerencias**

**Lógica Computacional**

Captura conocimiento del negocio de manera estructurada a menudo en forma de reglas.

**Técnicas de Optimización**

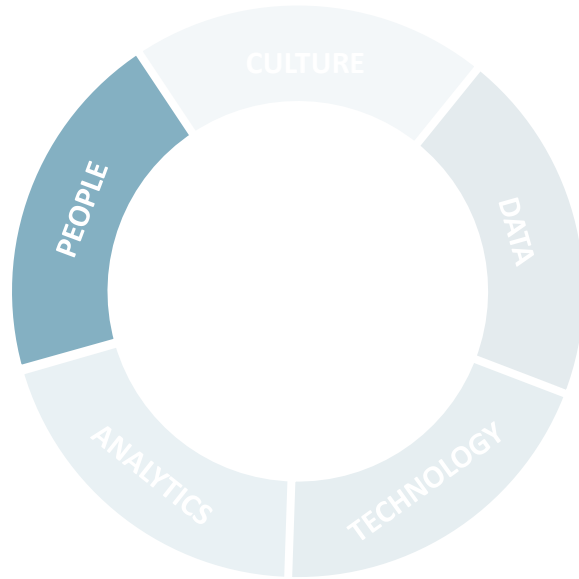
Maximizar los beneficios y gestionar compensaciones comerciales

**Qué se debería hacer?**

**Decisión de Negocio**

# Qué necesitamos para tener cultura del dato

Data Analytics



*‘Encontrar, motivar y retener el **talento** adecuado con las **habilidades** necesarias para promover las mejores decisiones.’*



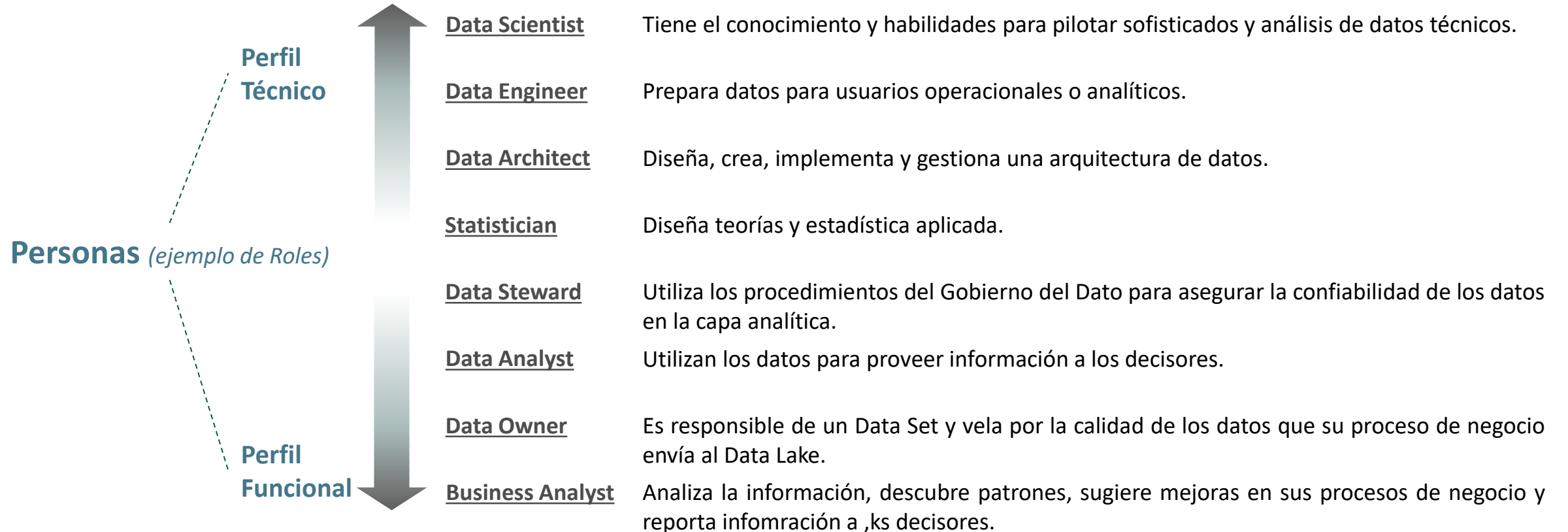
# Qué necesitamos para tener cultura del dato

## Data Analytics



Toda Organización que toma decisiones basadas en datos, necesita contar con personas con el adecuado rol analítico para extraer y generar información.

### Presente

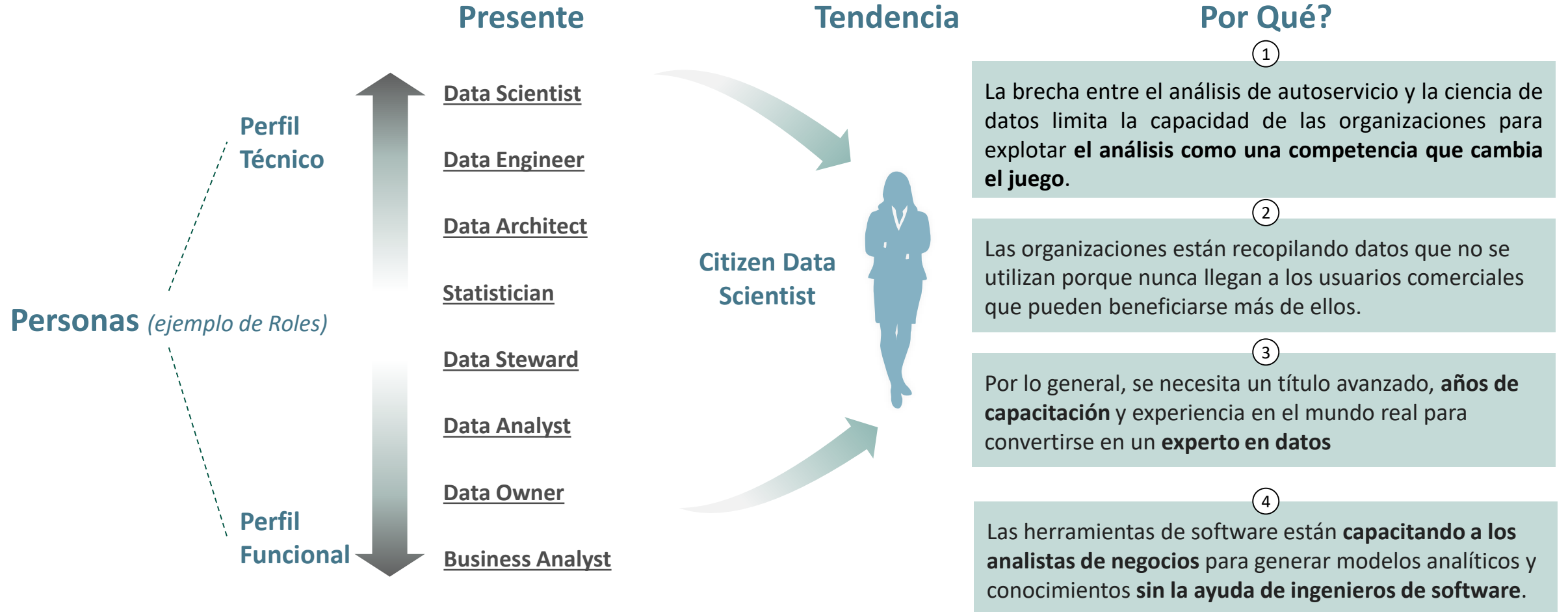


# Qué necesitamos para tener cultura del dato

## Data Analytics



Toda Organización que toma decisiones basadas en datos, necesita contar con personas con el adecuado rol analítico para extraer y generar información. La tendencia está creando un nuevo Rol denominado 'Citizen Data Scientist'.





# Qué necesitamos para tener cultura del dato

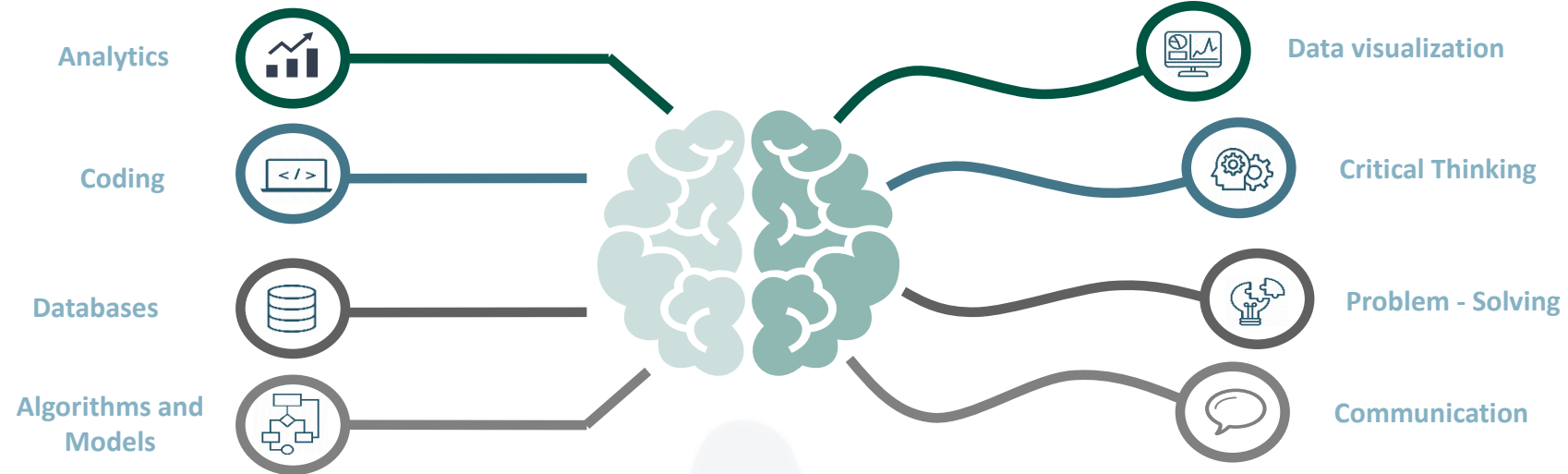
## Data Analytics



### Data Scientist

### Rasgos de un Citizen Data Scientist

### Business Analyst



Tanto los analistas de negocios como los científicos de datos **son expertos en el uso de datos para informar decisiones**, pero aplican sus habilidades de diferentes maneras, utilizando las mismas herramientas o herramientas similares ...

...El rol de **Citizen Data Scientist es una posición híbrida** que cierra la brecha entre el científico de datos técnicos de alto nivel y el analista de descubrimiento de datos más orientado a los negocios.





# Qué necesitamos para tener cultura del dato

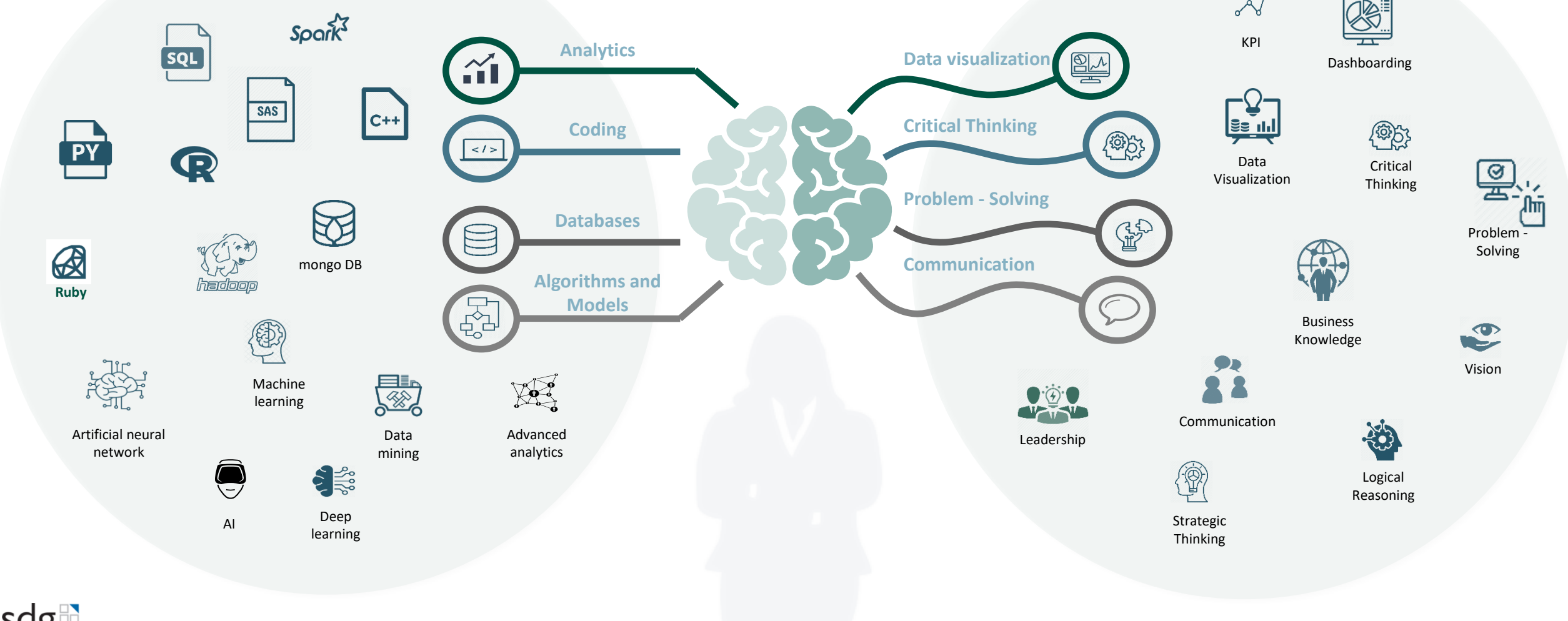
## Data Analytics



### Data Scientist

### Citizen Data Scientist

### Business Analyst



# Qué necesitamos para tener cultura del dato

Data Analytics

DATA

TECHNOLOGY


ANALYTICS

PEOPLE

CULTURE

Uno de los supuestos más transformadores de Gartner, es que para 2020...

El número de expertos en datos y análisis en unidades de negocios crecerá tres veces más que los expertos en departamentos de TI, lo que obligará a las empresas a repensar sus modelos organizacionales y conjuntos de habilidades

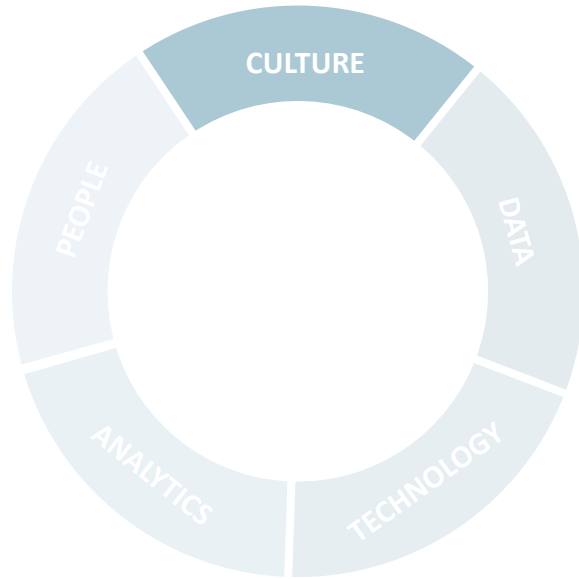


...el 55% de Citizen Data Scientist superará a Data Scientist en la cantidad de análisis avanzado producido

# Qué necesitamos para tener cultura del dato

*Data Analytics*

---



‘Asegurar que existan los mecanismos organizativos adecuados para institucionalizar las **decisiones basadas en datos.**’



# Qué necesitamos para tener cultura del dato

## Data Analytics



La Cultura del Dato se genera como combinación entre Datos, Tecnología, Analítica y Personas, y transforma los datos en un activo estratégico para el negocio. Todos estos factores juntos crean la cultura Data Driven.

*‘Entender la **tecnología** como un **acelerador** para lograr un mejor rendimiento de la analítica en los negocios’*

*‘Encontrar, motivar y retener el **talento** adecuado con las **habilidades** necesarias para promover las mejores decisiones’*



*‘**Los Datos** a través de la organización deben representar la **mejor versión de la verdad**, y deben ser verídicos para la correcta toma de decisiones’*

*‘Obtener **información** de los datos, y convertir esas ideas en **acciones** que promuevan los objetivos de la organización’*

## Data Analytics



Strategy. Decision. Governance.