



SKILLING
CENTER

TECMILENIO



Fundamentos del Big Data

Sectores estratégicos de Big Data y open data





Karla es coordinadora dentro de una empresa farmacéutica que crece de forma acelerada. Últimamente ha leído sobre el Big Data, lo que la ha llevado a hacer una propuesta para la empresa de cómo el Big Data puede ayudarlos en algunos proyectos. Al presentarle la idea a su jefa, ella le hizo varias preguntas a Karla que implican que investigue más a fondo.

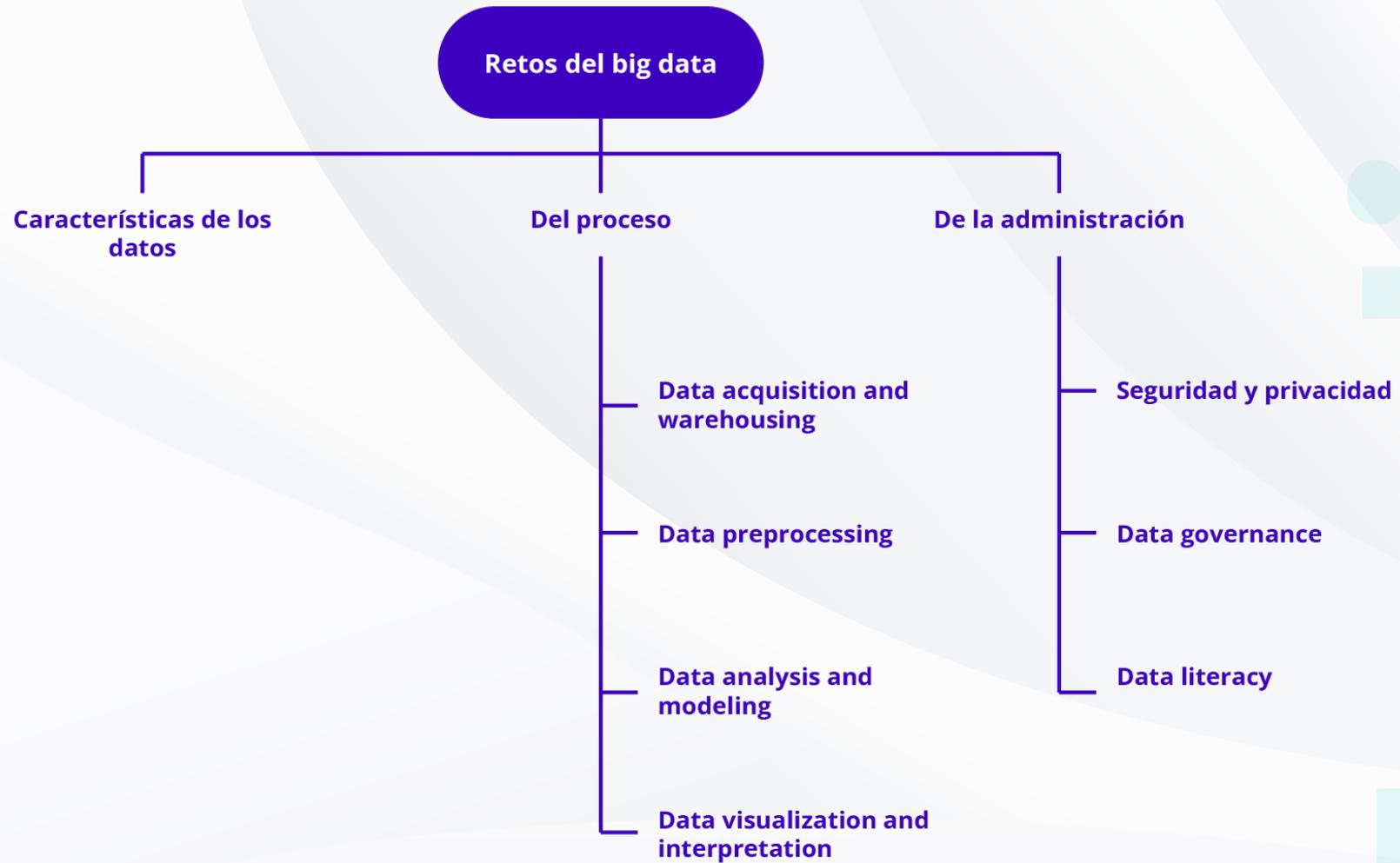
En este tema acompaña a Karla a determinar los beneficios que pudiera tener el sector farmacéutico con el potencial del Big Data y también a aprender sobre cuáles son sus principales retos.





El uso del Big Data tiene un gran potencial para los negocios, sin embargo, se requiere tener una estrategia definida antes de emprender el viaje.

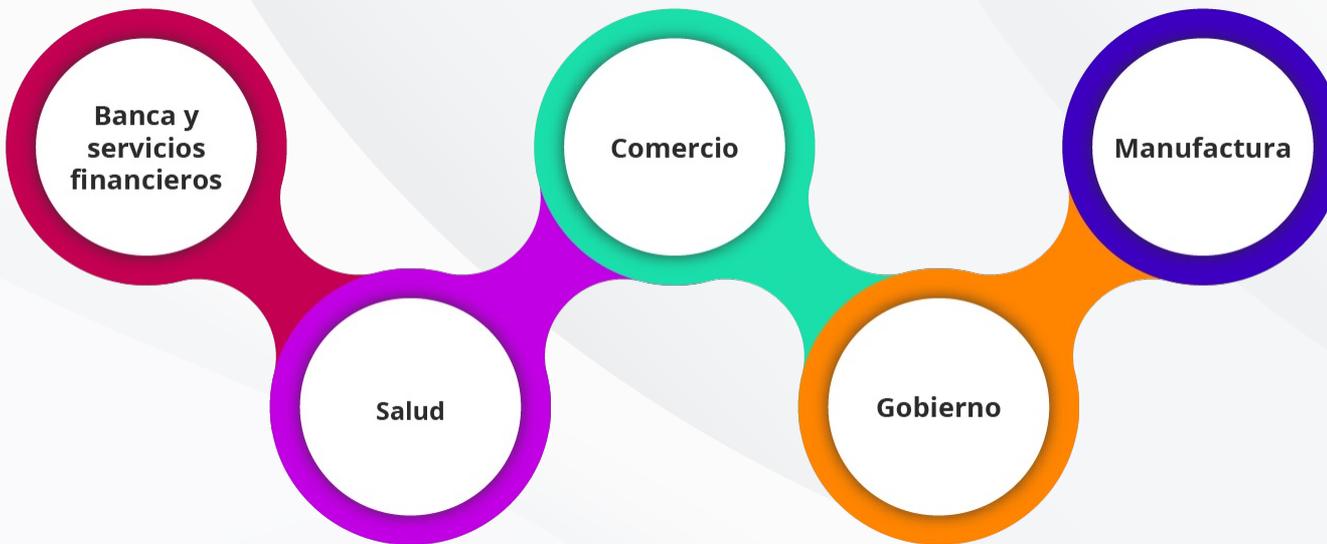






Muchas industrias ya han incorporado el Big Data como parte de su estrategia y modelo de negocio.

Las más sonadas son:





Open data es una iniciativa que tiene como objetivo la existencia de datos que sean de libre acceso, uso y distribución por cualquier persona para promover la utilización de datos y la innovación a partir de ellos.

Características:

- **Disponibilidad:** acceso a todo el público, descargable, sin comprometer la privacidad o seguridad de la fuente y de los usuarios.
- **Reúso y redistribución:** formato que permita su reutilización y redistribución sin restricciones.



Los sitios web que ofrecen servicios de *open data* son conocidos como interfaz de programación de aplicaciones (API, por sus siglas en inglés) y se utilizan mediante lenguajes de programación que permiten el intercambio de información entre dos agentes, cliente y servidor.

Una fuente de información sobre las API se puede consultar en el sitio de SYDLE.



El siguiente enlace es externo a la Universidad Tecmilenio, al acceder a él considera que debes apegarte a sus términos y condiciones.

Sydle. (2023). *¿Qué es API? Ejemplos, ventajas y tipos*. Recuperado de <https://www.sydle.com/es/blog/api-6214f68876950e47761c40e7>



En el fomento del acceso a datos libres participan tanto entidades públicas como privadas para poder crear dicho servicio, por ejemplo, el sitio web de Yahoo Finance. Además, implica la existencia de una norma que aumenta y legisla para mayor transparencia y acceso a la información. Algunas instituciones a partir de las cuales puedes encontrar bases de datos públicas nacionales son:

- Instituto Nacional de Estadística y Geografía
- Portal de Datos del Gobierno Federal
- Banco de México

Para bases de datos provenientes del sector privado y otro tipo de organizaciones:

- Kaggle
- Google Dataset Search
- DataHub
- Github



Para reforzar tu comprensión del tema, responde las siguientes preguntas:

1. Menciona el potencial del Big Data para mejorar los procesos de negocio y los apoyos del gobierno en la agricultura.
2. Comenta la manera en que resolverías los retos del Big Data para esta industria, los cuales incluyen las características de los datos, del proceso y de la administración.
3. Comenta las características y beneficios del ecosistema de los datos abiertos para la organización en donde laboras.
4. Investiga y describe la norma internacional para implementar y mantener la seguridad de la información en las organizaciones.



Karla ahora puede completar su propuesta poniendo los principales retos a enfrentar de acuerdo con el contexto de la empresa en la que trabaja. Por ejemplo, tienen muy claro qué datos necesitan y de dónde provienen, por lo que la *data acquisition and warehousing* no son un gran reto para la adopción del Big Data.

Su propuesta también se puede ver enriquecida argumentando cómo el Big Data ha transformado a la industria médica y cómo ellos podrían ser líderes en innovación próximamente. Incluso si la industria farmacéutica no se encontrara dentro de las más beneficiadas por el Big Data, Karla sabe que puede acceder a datos de libre acceso y que a partir de ellos puede innovar y generar valor para la empresa.





- Sydle. (2023). *¿Qué es API? Ejemplos, ventajas y tipos*. Recuperado de <https://www.sydle.com/es/blog/api-6214f68876950e47761c40e7>

Fundamentos del Big Data

Big Data en la empresa





Luis es un analista de finanzas dentro de una empresa de ventas minoristas que está haciendo sus primeros esfuerzos en Big Data, por lo que se han abierto nuevas vacantes para profesionistas en datos. Luis tiene muchas dudas sobre esta iniciativa: quiere saber cómo el Big Data impactaría a todas las áreas de la empresa, si su empleo está en riesgo ahora o si él podría ocupar una de las vacantes y desempeñarse como un científico de datos.

Luis nunca se queda con dudas, por lo que ha decidido aventurarse en aprender cómo funciona el Big Data dentro de la empresa en donde trabaja para lograr sus beneficios.

Acompaña a Luis en esta experiencia de aprendizaje y conoce cómo es que el Big Data funciona en el interior de una empresa, y cómo es que tú puedes hacer uso de la analítica para agregar valor al negocio, incluso si no eres un especialista en la ciencia de datos.





Existe una estrecha relación entre el Big Data y los ingresos de una empresa. Por un lado, la ciencia de datos siempre aporta procedimientos para encontrar soluciones a problemas como reducción de costos o para incrementar los ingresos.





Sedkaoui (2018) señala que las ventajas de la analítica de Big Data son:





La utilización del Big Data por sí sola no resuelve los problemas, se requiere también modificar la cultura de la organización para adoptar la toma de decisiones basada en datos, alineación entre áreas y establecer políticas de seguridad efectivas.

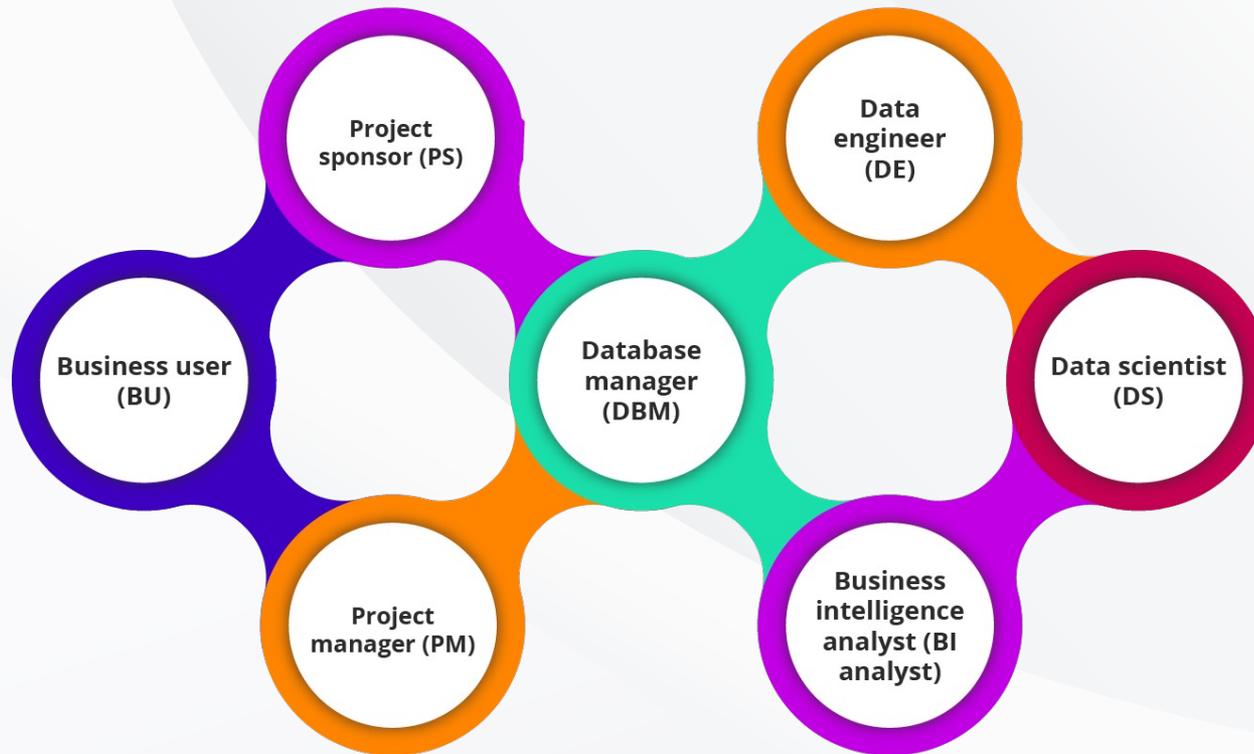
Schmelzer (2021) aborda algunos casos de usos comunes donde los hallazgos sobre los clientes son los protagonistas con los que las empresas transforman su día a día con base en Big Data:

Tener visión 360° del cliente.

Mejorar la adquisición y retención de clientes.

Personalización y recomendación.

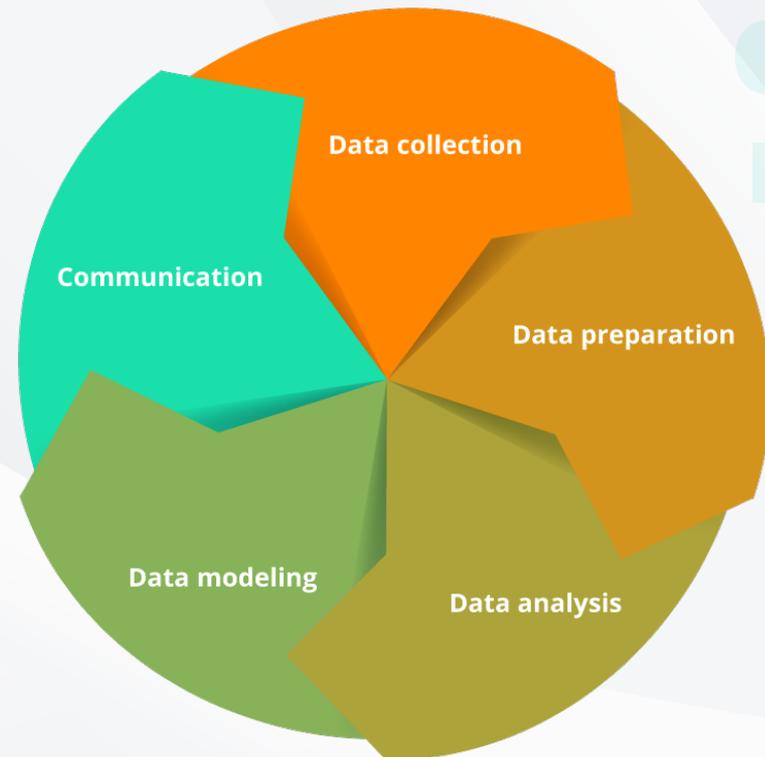
Roles involucrados en proyectos de ciencia de datos:



Fuente: EMC Education Services. (2015). *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Estados Unidos: Wiley.



Normalmente la ciencia de datos sigue un proceso que contempla las siguientes etapas, en las cuales están involucrados en mayor o menor medida todos los roles, con base en los criterios de cada organización:





Davenport y Patil (2022) reconocen que conforme el tiempo pase habrá empleos más especializados alrededor de la ciencia de datos, pero que, a pesar de ello, los científicos de datos “deben contribuir a la recopilación adecuada de datos, el análisis responsable, los modelos completamente implementados y los resultados comerciales exitosos”.



El *Business Intelligence* (BI) determina estrategias, herramientas y personas que permiten a la empresa tomar decisiones más fundamentadas en datos, y con el tiempo, desarrolla la cultura de tomar decisiones con base en información, o *data-driven enterprise*, convirtiendo el proceso en una estrategia para lograr sus objetivos, desarrollando un valioso activo no tangible.

Para esto, es necesario que ocurran tres cosas:

Los datos estén actualizados y sean confiables.

Los datos sean fáciles de interpretar y sean presentados en formatos fáciles de explorar.

Los datos estén disponibles en cualquier momento y su acceso sea sencillo.



La estrategia BI está fuertemente apoyada por la visualización, que incluye técnicas para comunicar historias (*storytelling*) para compartir y explicar fácilmente la situación y la solución propuesta.



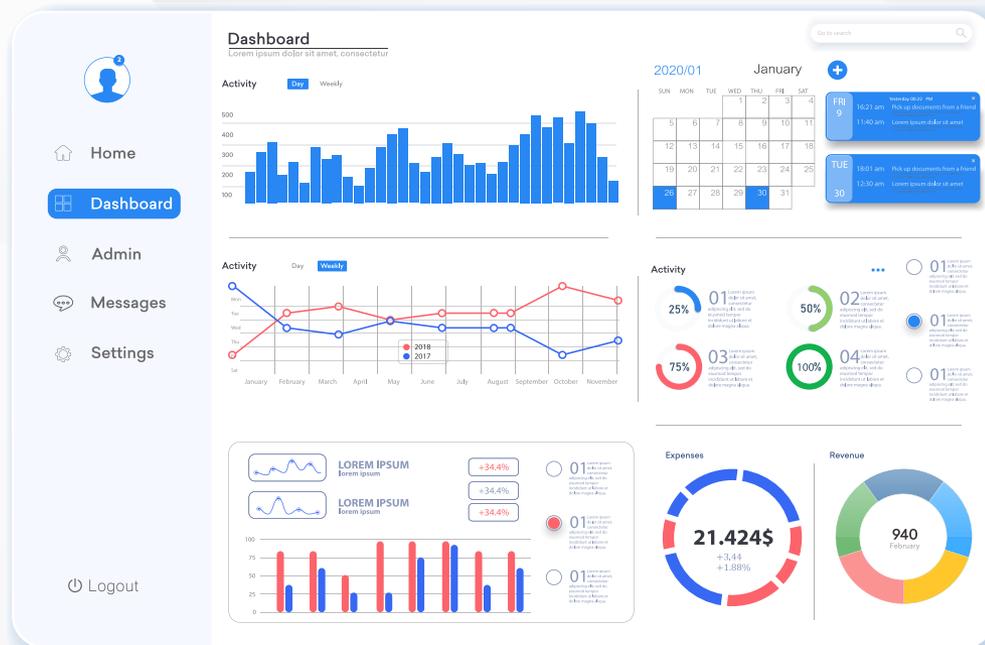


Para reforzar tu comprensión del tema, considera el caso hipotético de una empresa que ofrece servicios de telecomunicaciones de telefonía celular a millones de usuarios y responde las siguientes preguntas:

1. Menciona las ventajas de utilizar Big Data para obtener reducción de costos, satisfacer las necesidades de los clientes, soportar la toma de decisiones y la creación de nuevos servicios. Considera cada uno de los propósitos por separado.
2. Menciona dos funciones específicas que pueden asignarse a cada uno de los roles del equipo de trabajo en proyectos de la ciencia de datos, que aplican en la empresa hipotética.
3. Investiga y menciona una herramienta que se puede utilizar para cada una de las etapas de un proyecto de ciencia de datos en la empresa hipotética.
4. Menciona un plan general para transformar la empresa hipotética en *data-driven enterprise*.



Con el aprendizaje obtenido en este tema Luis sabe que su trabajo no está en riesgo, pues el Big Data y los roles alrededor de él están destinados a existir al mismo tiempo y a colaborar continuamente a través del proceso de ciencia de datos. También sabe que puede obtener información a través de *dashboards* sin tener que ser un especialista en datos y está tan emocionado de formar parte de una empresa en camino de ser *data-driven*.





- Davenport, T., y Patil, D. (2022). *Is Data Scientist Still the Sexiest Job of the 21st Century?* Recuperado de <https://hbr.org/2022/07/is-data-scientist-still-the-sexiest-job-of-the-21st-century>
- EMC Education Services. (2015). *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Estados Unidos: Wiley.
- Schmelzer, R. (2021). *8 big data use cases for businesses and industry examples*. Recuperado de <https://www.techtarget.com/searchbusinessanalytics/feature/8-big-data-use-cases-for-businesses-and-industry-examples>.
- Sedkaoui, S. (2018). *Data Analytics and Big Data*. Wiley

Tecmilenio no guarda relación alguna con las marcas mencionadas como ejemplo. Las marcas son propiedad de sus titulares conforme a la legislación aplicable, estas se utilizan con fines académicos y didácticos, por lo que no existen fines de lucro, relación publicitaria o de patrocinio.

Todos los derechos reservados @ Universidad Tecmilenio

La obra presentada es propiedad de ENSEÑANZA E INVESTIGACIÓN SUPERIOR A.C. (UNIVERSIDAD TECMILENIO), protegida por la Ley Federal de Derecho de Autor; la alteración o deformación de una obra, así como su reproducción, exhibición o ejecución pública sin el consentimiento de su autor y titular de los derechos correspondientes es constitutivo de un delito tipificado en la Ley Federal de Derechos de Autor, así como en las Leyes Internacionales de Derecho de Autor. El uso de imágenes, fragmentos de videos, fragmentos de eventos culturales, programas y demás material que sea objeto de protección de los derechos de autor, es exclusivamente para fines educativos e informativos, y cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por UNIVERSIDAD TECMILENIO. Queda prohibido copiar, reproducir, distribuir, publicar, transmitir, difundir, o en cualquier modo explotar cualquier parte de esta obra sin la autorización previa por escrito de UNIVERSIDAD TECMILENIO. Sin embargo, usted podrá bajar material a su computadora personal para uso exclusivamente personal o educacional y no comercial limitado a una copia por página. No se podrá remover o alterar de la copia ninguna leyenda de Derechos de Autor o la que manifieste la autoría del material.