

Las 5 Vs de Big Data En Netflix

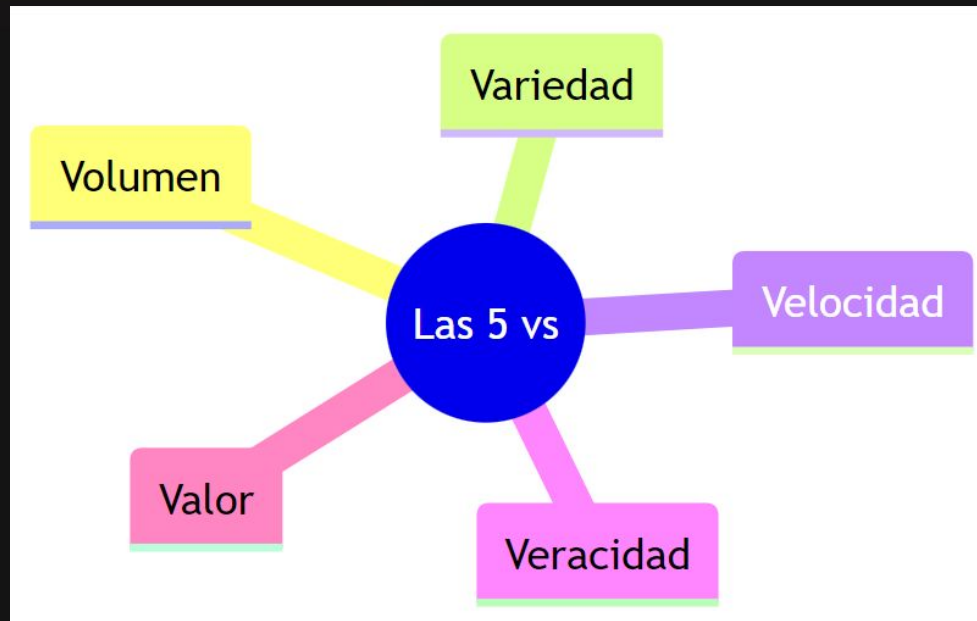


Katherine Jimenez Vivanco
Mauricio Carazas Segovia
Oscar Ramos



Los servicio de streaming son cada vez más utilizados por las personas

- Al tener más usuarios estos servicios tienen más datos que analizar para poder brindar un servicio de calidad a sus usuarios.
- En este trabajo analizaremos cómo se aplican las 5v's de Big Data en la empresa Netflix



¿Dónde se ejecutan todos los microservicios de netflix?



Netflix tenía bajo su poder una amplia red de servidores, sin embargo se dieron cuenta que la velocidad de su crecimiento haría que fuese imposible mantener estos a largo plazo y significa tiempo valioso que podrían invertir en otras áreas de la empresa.

Por lo cual decidieron trasladar todo su material a la nube, para poder ejecutarlo en servidores de otras empresas en este caso **Amazon Web Services**.

Para lo cual los programadores de Netflix escribieron cientos de programas y los instalaron rápidamente en los servidores de **Amazon Web Services**.

¿Dónde se ejecutan todos los microservicios de netflix?



Netflix tenía bajo su poder una amplia red de servidores, sin embargo se dieron cuenta que la velocidad de su crecimiento haría que fuese imposible mantener estos a largo plazo y significa tiempo valioso que podrían invertir en otras áreas de la empresa.

Por lo cual decidieron trasladar todo su material a la nube, para poder ejecutarlo en servidores de otras empresas en este caso **Amazon Web Services**.

Para lo cual los programadores de Netflix escribieron cientos de programas y los instalaron rápidamente en los servidores de **Amazon Web Services**.

¿Dónde se ejecutan todos los microservicios de netflix?



Netflix tenía bajo su poder una amplia red de servidores, sin embargo se dieron cuenta que la velocidad de su crecimiento haría que fuese imposible mantener estos a largo plazo y significa tiempo valioso que podrían invertir en otras áreas de la empresa.

Por lo cual decidieron trasladar todo su material a la nube, para poder ejecutarlo en servidores de otras empresas en este caso **Amazon Web Services**.

Para lo cual los programadores de Netflix escribieron cientos de programas y los instalaron rápidamente en los servidores de **Amazon Web Services**.

¿Dónde se ejecutan todos los microservicios de netflix?



Netflix tenía bajo su poder una amplia red de servidores, sin embargo se dieron cuenta que la velocidad de su crecimiento haría que fuese imposible mantener estos a largo plazo y significa tiempo valioso que podrían invertir en otras áreas de la empresa.

Por lo cual decidieron trasladar todo su material a la nube, para poder ejecutarlo en servidores de otras empresas en este caso **Amazon Web Services**.

Para lo cual los programadores de Netflix escribieron cientos de programas y los instalaron rápidamente en los servidores de **Amazon Web Services**.

1. Volumen

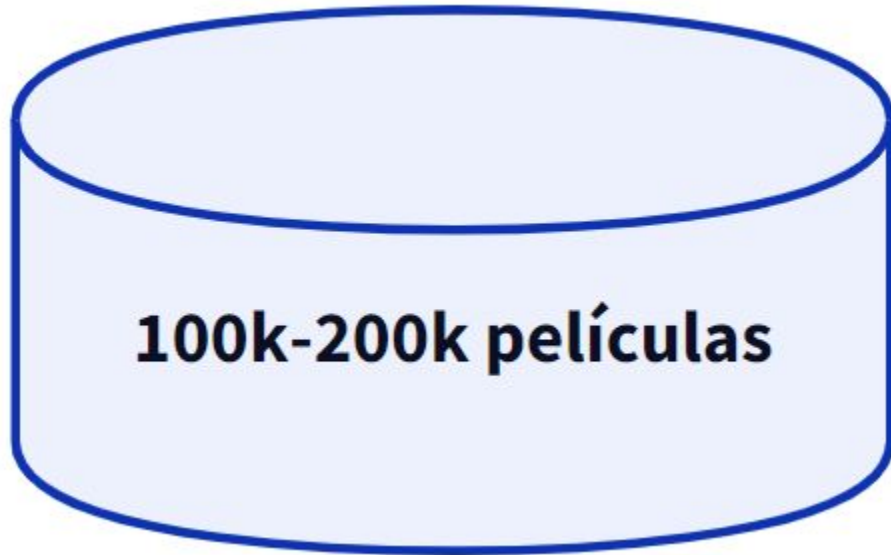
¿Qué aspecto tienen los servidores de almacenamiento que albergan el contenido de Netflix y lo transmiten?

Una mezcla de discos duros empacados juntos en una computadora se compone de ellos. Utilizan **36 unidades** que pueden transportar aproximadamente **100 TB de datos**. Estos servidores son capaces de almacenar y cargar continuamente entre **10,000 y 20,000 películas**.

1. Volumen

¿Qué as
conteni

Una me
compon
aproxim
almace

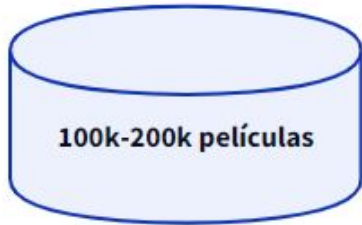
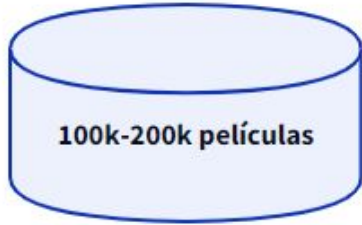
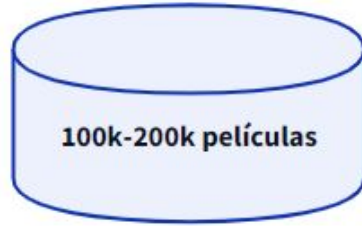
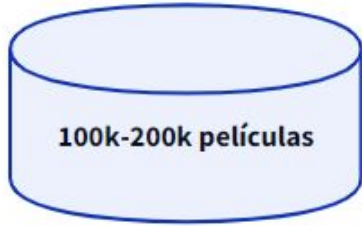


el

se
tar
de

¿Qué as
contenid

Una me
compon
aproxim
almacen

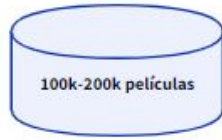
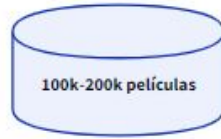
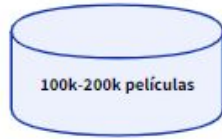
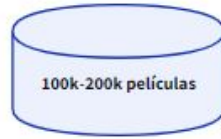
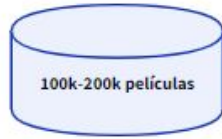
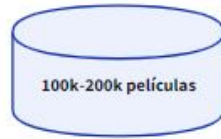
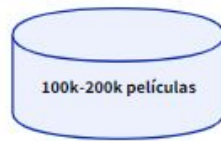
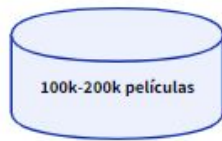


ergan el

dora se
sportar
ces de

¿Qué aspecto tiene
el contenido de Netflix?

Una mezcla de dispositivos
que componen de ello
aproximadamente
almacenar y cargar

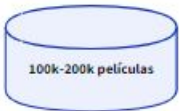
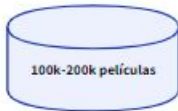
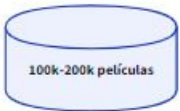
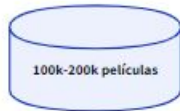
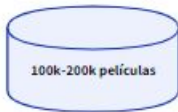
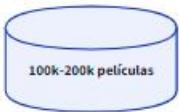
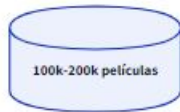
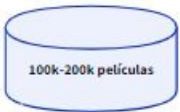
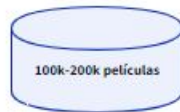
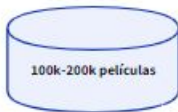
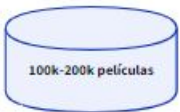
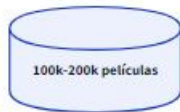


to que albergan el

la computadora se
ueden transportar
s son capaces de
0 películas.

¿Qué aspecto
contiene

Una mezcla
compone
aproximada
almacena

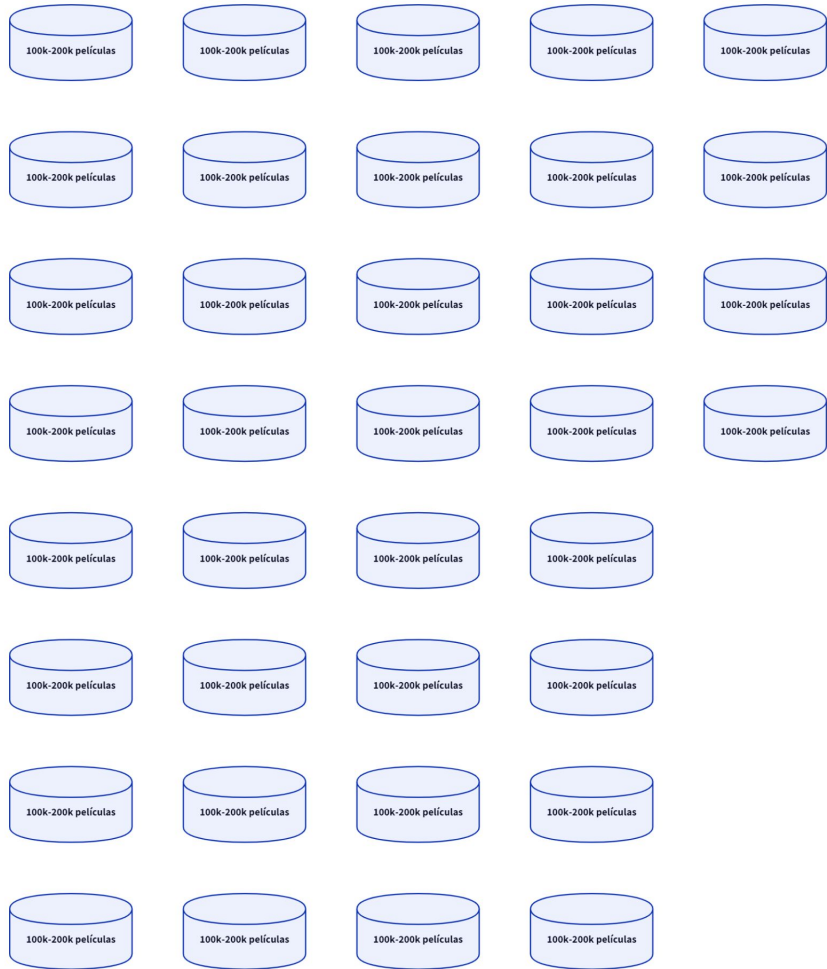


bergan el

adora se
nsportar
aces de

¿Qué aspecto
contenido de








Una mezcla
compone de
aproximadam
almacenar y c










bergan el

adora se
nsportar
aces de
.








2. Variedad

- **Netflix** brinda una diversidad de contenidos que incluyen:
 - Series
 - Películas
 - Programas en múltiples idiomas
- **Tipos de datos** que Netflix maneja:
 -  **Formatos de video:** Para series, películas y documentales.
 -  **Imágenes:** Portadas, fondos y avatares de perfil.
 -  **Textos:** Títulos, descripciones, subtítulos y metadatos.
 -  **Audio:** Pistas de sonido en diferentes idiomas y efectos sonoros.
 -  **Datos de interacción:** Hora de visualización, duración, nivel de popularidad e interacciones sociales.
 -  **Información de búsqueda:** Palabras clave, historial y recomendaciones.
 -  **Datos de transmisión:** Calidad del streaming, interrupciones y velocidad de carga.

2. Variedad

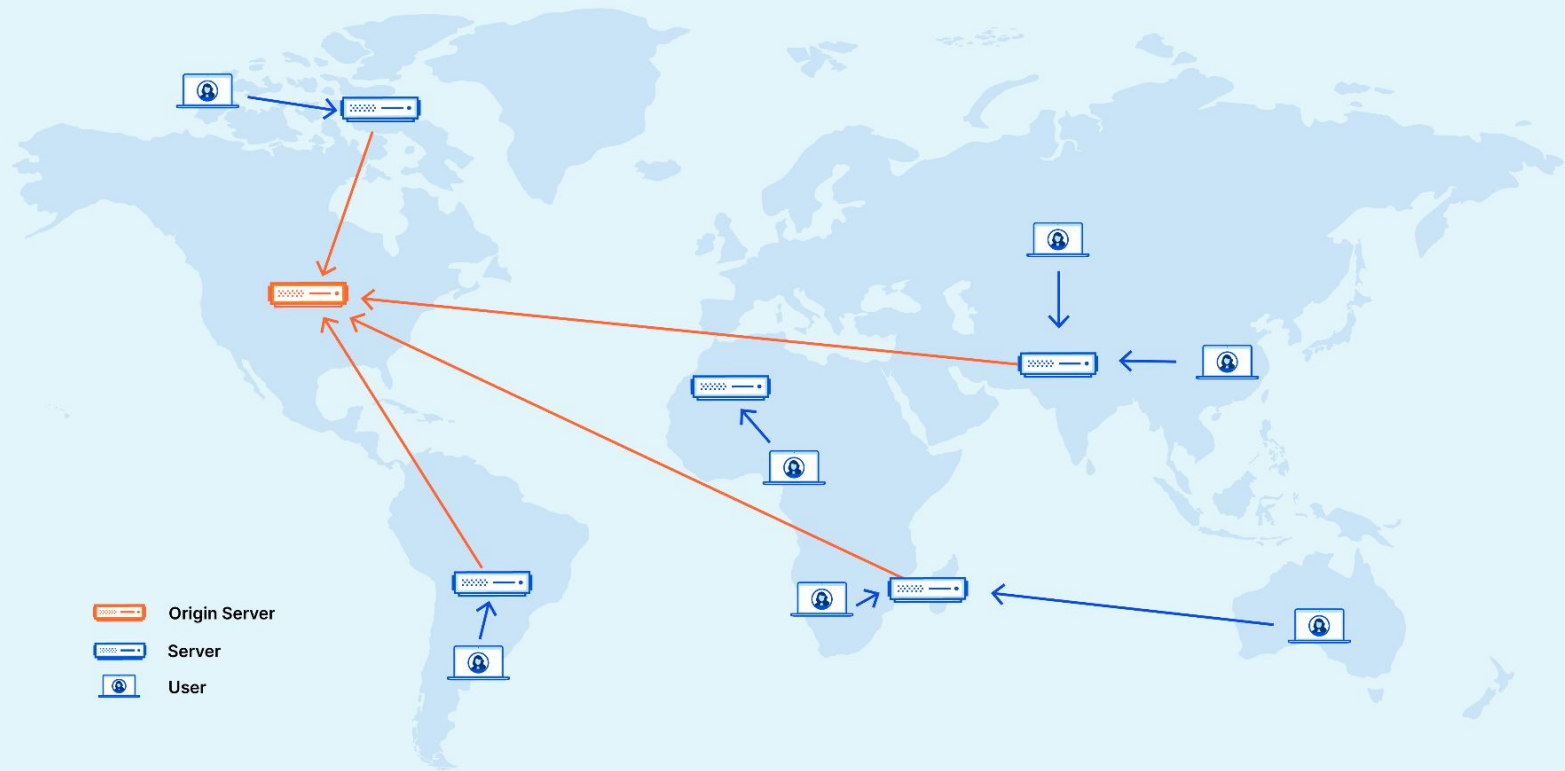
- **Netflix** brinda una diversidad de contenidos que incluyen:
 - Series
 - Películas
 - Programas en múltiples idiomas
- **Tipos de datos** que Netflix maneja:
 -  **Formatos de video:** Para series, películas y documentales.
 -  **Imágenes:** Portadas, fondos y avatares de perfil.
 -  **Textos:** Títulos, descripciones, subtítulos y metadatos.
 -  **Audio:** Pistas de sonido en diferentes idiomas y efectos sonoros.
 -  **Datos de interacción:** Hora de visualización, duración, nivel de popularidad e interacciones sociales.
 -  **Información de búsqueda:** Palabras clave, historial y recomendaciones.
 -  **Datos de transmisión:** Calidad del streaming, interrupciones y velocidad de carga.

2. Variedad

- **Netflix** brinda una diversidad de contenidos que incluyen:
 - Series
 - Películas
 - Programas en múltiples idiomas
- **Tipos de datos** que Netflix maneja:
 -  **Formatos de video:** Para series, películas y documentales.
 -  **Imágenes:** Portadas, fondos y avatares de perfil.
 -  **Textos:** Títulos, descripciones, subtítulos y metadatos.
 -  **Audio:** Pistas de sonido en diferentes idiomas y efectos sonoros.
 -  **Datos de interacción:** Hora de visualización, duración, nivel de popularidad e interacciones sociales.
 -  **Información de búsqueda:** Palabras clave, historial y recomendaciones.
 -  **Datos de transmisión:** Calidad del streaming, interrupciones y velocidad de carga.

3. Velocidad

- Según Netflix, cuando transmites un programa de TV o una película en calidad estándar, utilizas alrededor de **1 GB de datos por hora**. Si optas por videos en alta definición, esta cantidad aumenta a aproximadamente **3 GB por hora**.
- Para manejar esta gran cantidad de datos y asegurar un acceso rápido, Netflix emplea un **sistema de administración de bases de datos distribuidas** (DDBMS). Esto permite que los datos se almacenen en ubicaciones cercanas a donde se necesitan, lo que agiliza el tiempo de acceso.
- Cuando seleccionas una película en Netflix, la consulta para acceder a los datos se procesa en múltiples ubicaciones dentro de la base de datos. Esto aprovecha la capacidad de procesamiento paralelo de la CPU que ofrece el **DDBMS**, lo que no solo mejora la velocidad, sino que también reduce las necesidades de hardware y optimiza la eficiencia del sistema.



3. Velocidad

- Según Netflix, cuando transmites un programa de TV o una película en calidad estándar, utilizas alrededor de 1 GB de datos por hora. Si optas por videos en alta definición, esta cantidad aumenta a aproximadamente 3 GB por hora.
- Para manejar esta gran cantidad de datos y asegurar un acceso rápido, Netflix emplea un **sistema de administración de bases de datos distribuidas** (DDBMS). Esto permite que los datos se almacenen en ubicaciones cercanas a donde se necesitan, lo que agiliza el tiempo de acceso.
- Cuando seleccionas una película en Netflix, la consulta para acceder a los datos se procesa en múltiples ubicaciones dentro de la base de datos. Esto aprovecha la capacidad de procesamiento paralelo de la CPU que ofrece el **DDBMS**, lo que no solo mejora la velocidad, sino que también reduce las necesidades de hardware y optimiza la eficiencia del sistema.

3. Velocidad

- Según Netflix, cuando transmites un programa de TV o una película en calidad estándar, utilizas alrededor de 1 GB de datos por hora. Si optas por videos en alta definición, esta cantidad aumenta a aproximadamente 3 GB por hora.
- Para manejar esta gran cantidad de datos y asegurar un acceso rápido, Netflix emplea un **sistema de administración de bases de datos distribuidas** (DDBMS). Esto permite que los datos se almacenen en ubicaciones cercanas a donde se necesitan, lo que agiliza el tiempo de acceso.
- Cuando seleccionas una película en Netflix, la consulta para acceder a los datos se procesa en múltiples ubicaciones dentro de la base de datos. Esto aprovecha la capacidad de procesamiento paralelo de la CPU que ofrece el **DDBMS**, lo que no solo mejora la velocidad, sino que también reduce las necesidades de hardware y optimiza la eficiencia del sistema.

4. Veracidad

Netflix, como plataforma líder de streaming, maneja una cantidad inmensa de datos diariamente, y con ello, enfrenta desafíos en cuanto a la veracidad de la información.

Se ha observado una notable variación en los datos de calificaciones:

- No todas las películas reciben la misma cantidad o tipo de calificación por parte de los usuarios.
- Mientras que algunas películas pueden tener solo 3 calificaciones, hay usuarios que han calificado más de 17,000 películas.
- Esta variabilidad en las calificaciones y la interacción del usuario puede introducir sesgos, ruido y anomalías en los datos.

Es esencial para Netflix filtrar y validar esta información utilizando tecnologías de **Big Data** para identificar y descartar datos no fiables o atípicos.

4. Veracidad

Netflix, como plataforma líder de streaming, maneja una cantidad inmensa de datos diariamente, y con ello, enfrenta desafíos en cuanto a la veracidad de la información.

Se ha observado una notable variación en los datos de calificaciones:

- No todas las películas reciben la misma cantidad o tipo de calificación por parte de los usuarios.
- Mientras que algunas películas pueden tener solo 3 calificaciones, hay usuarios que han calificado más de 17,000 películas.
- Esta variabilidad en las calificaciones y la interacción del usuario puede introducir sesgos, ruido y anomalías en los datos.

Es esencial para Netflix filtrar y validar esta información utilizando tecnologías de **Big Data** para identificar y descartar datos no fiables o atípicos.

4. Veracidad

Netflix, como plataforma líder de streaming, maneja una cantidad inmensa de datos diariamente, y con ello, enfrenta desafíos en cuanto a la veracidad de la información.

Se ha observado una notable variación en los datos de calificaciones:

- No todas las películas reciben la misma cantidad o tipo de calificación por parte de los usuarios.
- Mientras que algunas películas pueden tener solo 3 calificaciones, hay usuarios que han calificado más de 17,000 películas.
- Esta variabilidad en las calificaciones y la interacción del usuario puede introducir sesgos, ruido y anomalías en los datos.

Es esencial para Netflix filtrar y validar esta información utilizando tecnologías de **Big Data** para identificar y descartar datos no fiables o atípicos.

4. Veracidad

Netflix, como plataforma líder de streaming, maneja una cantidad inmensa de datos diariamente, y con ello, enfrenta desafíos en cuanto a la veracidad de la información.

Se ha observado una notable variación en los datos de calificaciones:

- No todas las películas reciben la misma cantidad o tipo de calificación por parte de los usuarios.
- Mientras que algunas películas pueden tener solo 3 calificaciones, hay usuarios que han calificado más de 17,000 películas.
- Esta variabilidad en las calificaciones y la interacción del usuario puede introducir sesgos, ruido y anomalías en los datos.

Es esencial para Netflix filtrar y validar esta información utilizando tecnologías de **Big Data** para identificar y descartar datos no fiables o atípicos.

5. Valor

- Netflix destaca por ofrecer una amplia selección de contenido de calidad a precios razonables, desde películas de estreno hasta series originales.
- La plataforma utiliza los datos de sus usuarios para ofrecer recomendaciones personalizadas y perfiles de usuario, mejorando la experiencia de visualización según los gustos y preferencias de cada uno.
- Estas características, derivadas del análisis de datos, no solo enriquecen la experiencia del usuario, sino que también guían decisiones estratégicas de la empresa, como la producción de contenido y la adquisición de licencias.
- **Conclusión:** Netflix transforma grandes volúmenes de datos en valor tangible para sus usuarios y para su negocio.

5. Valor

- Netflix destaca por ofrecer una amplia selección de contenido de calidad a precios razonables, desde películas de estreno hasta series originales.
- La plataforma utiliza los datos de sus usuarios para ofrecer recomendaciones personalizadas y perfiles de usuario, mejorando la experiencia de visualización según los gustos y preferencias de cada uno.
- Estas características, derivadas del análisis de datos, no solo enriquecen la experiencia del usuario, sino que también guían decisiones estratégicas de la empresa, como la producción de contenido y la adquisición de licencias.
- **Conclusión:** Netflix transforma grandes volúmenes de datos en valor tangible para sus usuarios y para su negocio.

5. Valor

- Netflix destaca por ofrecer una amplia selección de contenido de calidad a precios razonables, desde películas de estreno hasta series originales.
- La plataforma utiliza los datos de sus usuarios para ofrecer recomendaciones personalizadas y perfiles de usuario, mejorando la experiencia de visualización según los gustos y preferencias de cada uno.
- Estas características, derivadas del análisis de datos, no solo enriquecen la experiencia del usuario, sino que también guían decisiones estratégicas de la empresa, como la producción de contenido y la adquisición de licencias.
- **Conclusión:** Netflix transforma grandes volúmenes de datos en valor tangible para sus usuarios y para su negocio.

5. Valor

- Netflix destaca por ofrecer una amplia selección de contenido de calidad a precios razonables, desde películas de estreno hasta series originales.
- La plataforma utiliza los datos de sus usuarios para ofrecer recomendaciones personalizadas y perfiles de usuario, mejorando la experiencia de visualización según los gustos y preferencias de cada uno.
- Estas características, derivadas del análisis de datos, no solo enriquecen la experiencia del usuario, sino que también guían decisiones estratégicas de la empresa, como la producción de contenido y la adquisición de licencias.
- **Conclusión:** Netflix transforma grandes volúmenes de datos en valor tangible para sus usuarios y para su negocio.

5. Valor

- Netflix destaca por ofrecer una amplia selección de contenido de calidad a precios razonables, desde películas de estreno hasta series originales.
- La plataforma utiliza los datos de sus usuarios para ofrecer recomendaciones personalizadas y perfiles de usuario, mejorando la experiencia de visualización según los gustos y preferencias de cada uno.
- Estas características, derivadas del análisis de datos, no solo enriquecen la experiencia del usuario, sino que también guían decisiones estratégicas de la empresa, como la producción de contenido y la adquisición de licencias.
- **Conclusión:** Netflix transforma grandes volúmenes de datos en valor tangible para sus usuarios y para su negocio.

Referencias

- AMAN SINGH, CLOUD STORAGE: HOW DOES NETFLIX STORE THEIR CONTENT?, Enero, 2021
disponible en:
<https://www.globaltechcouncil.org/big-data/cloud-storage-how-does-netflix-store-their-content/>
- Eric Walz, How Netflix Uses a Distributed Database Management System to Deliver Your Movies, 6 de julio de 2016 disponible en :
<https://www.linkedin.com/pulse/how-netflix-uses-distributed-database-management-system-eric-walz/>
- Chaithanya Pramodh Kasula, Netflix Recommender System — A Big Data Case Study, Junio 2020
Disponible en:
<https://towardsdatascience.com/netflix-recommender-system-a-big-data-case-study-19cfa6d56ff5>