



# GESTIÓN DE DATOS EN EL NEGOCIO AUDIOVISUAL: NETFLIX COMO ESTUDIO DE CASO

Data management in audiovisual business:  
*Netflix as a case study*



Eva-Patricia Fernández-Manzano, Elena Neira y Judith Clares-Gavilán

**Nota:** This article can be read in its original English version on:  
<http://www.elprofesionaldelainformacion.com/contenidos/2016/jul/06.pdf>



**Eva-Patricia Fernández-Manzano** es licenciada y doctora en ciencias de la información por la *Universidad Complutense de Madrid (UCM)*, *Master en dirección de empresa audiovisual* por el *IE* y *Master en big data y business intelligence* por la *Escuela de Organización Industrial (EOI)*. Profesora de la *Universidad Rey Juan Carlos (URJC)*. Participa en congresos nacionales e internacionales, y es autora de artículos y libros sobre sus líneas de investigación: *big data*, gestión y producción en medios de comunicación.

<http://orcid.org/0000-0001-7655-872X>

*Universidad Rey Juan Carlos*  
Edificio Departamental, Despacho 207  
Camino del Molino, s/n. 28943 Fuenlabrada (Madrid), España  
[evapatriciafer@gmail.com](mailto:evapatriciafer@gmail.com)



**Elena Neira** es licenciada en derecho y en comunicación audiovisual. Vinculada a los medios de comunicación desde el año 2000, primero en *Grupo Zeta*, en los departamentos de *Marketing* e *Investigación de mercado*. Con posterioridad trabajó como directora de marketing de la distribuidora *Wide Pictures*. En la actualidad asesora a empresas del sector audiovisual. Experta en marketing, redes sociales, tecnología y nuevos modelos de distribución audiovisual. Es docente en la *Universitat Oberta de Catalunya (UOC)*. Miembro de *Innovación Audiovisual*. Autora de *El espectador social* (2013) y *La otra pantalla* (2015).

<http://orcid.org/0000-0002-8565-2889>

*La otra pantalla*  
Rosselló, 421 - 4, 1. 08025 Barcelona, España  
[elena@laotrapantalla.com](mailto:elena@laotrapantalla.com)



**Judith Clares-Gavilán** es profesora de los *Estudios de Ciencias de la Información y la Comunicación* de la *Universitat Oberta de Catalunya (UOC)*. Licenciada en periodismo y doctora en comunicación por la *Universitat Ramon Llull (URL)*. Sus principales líneas de investigación son: la estructura y las políticas del sector audiovisual y muy especialmente los cambios producidos en el sector con la irrupción de la distribución audiovisual en internet, el vídeo bajo demanda (*video on demand*, VOD) y la televisión conectada. Miembro del Grupo *GAME*, es investigadora del proyecto de I+D financiado por el *Ministerio de Economía y Competitividad* "Cultura lúdica, competencia digital y aprendizajes".

<http://orcid.org/0000-0002-7462-9712>

*Universitat Oberta de Catalunya, Estudios de Ciencias de la Información y de la Comunicación*  
Rambla Poblenou, 156. 08018 Barcelona, España  
[jclares@uoc.edu](mailto:jclares@uoc.edu)

## Resumen

El negocio de la distribución de contenidos audiovisuales bajo demanda ha encontrado en los *big data* un aliado estratégico para la obtención de información y la toma de decisiones en torno al *core content* de su negocio: el contenido y sus suscriptores. A través de este artículo analizamos qué son los *big data* y cuál es su flujo de gestión de datos en una empresa audiovisual orientada a la distribución de contenido audiovisual bajo demanda. Para ello, seleccionamos como caso de estudio la compañía *Netflix*. A partir de aquí veremos por qué los *big data* son considerados eje estratégico del negocio y cuáles son los retos que se deben afrontar para la expansión internacional.

Artículo recibido el 31-03-2016  
Aceptación definitiva: 10-04-2016

## Palabras clave

Audiencias; Análisis de datos; *Big data*; Inteligencia de negocio; Distribución audiovisual bajo demanda; Minería de datos; Modelos de negocio; *Netflix*; Vídeo bajo demanda (VOD).

## Abstract

Big data has become an enormous asset for on-demand content distribution services, helping information supply and decision-making, regarding both the content of the database and subscribers to the database. In this article we describe and define big data and data management in a media company devoted to on-demand audiovisual content distribution: *Netflix*. This article suggests that big data is a prime strategy in media business and outlines the upcoming challenges that follow its global expansion.

## Keywords

Audience; Big data; Business intelligence; Business models; Data analysis; Data mining; *Netflix*; On demand audiovisual content distribution; Video on demand (VOD).

**Fernández-Manzano, Eva-Patricia; Neira, Elena; Clares-Gavilán, Judith (2016). "Data management in audiovisual business: *Netflix* as a case study". *El profesional de la información*, v. 25, n. 4, pp. 568-576.**

<http://dx.doi.org/10.3145/epi.2016.jul.06>

## 1. Introducción

La incorporación de las nuevas tecnologías en el ámbito audiovisual ha suscitado una serie de cambios en los modelos de negocio. Algunos sectores y empresas han comprendido rápidamente los nuevos hábitos y flujos de trabajo pero, para la gran mayoría, la transición no se ha completado. La hoja de ruta de la adaptación viene marcada por dos aspectos fundamentales:

- utilidad
- franja temporal

Esto es así porque las nuevas formas digitales de acceso y consumo de productos audiovisuales requieren nuevos modos de actuar por parte del usuario. En la mayoría de los casos, el consumidor no integra los cambios tecnológicos en su día a día hasta que les ve la utilidad: desde una mejora en la calidad de la imagen, la satisfacción de que el sistema de recomendación acierte con los gustos propios o la posibilidad de consumir contenidos bajo demanda frente al diferido. Como indica **Fernández-Manzano (2016)**, en el caso concreto de los *big data*, o la gestión de datos masivos, en entornos audiovisuales "los datos en sí carecen de valor, la importancia radica en el uso que se haga con ellos". Esto es, por tanto, la utilidad. Además, se suma un matiz temporal que afecta a la propia gestión de datos y a los usuarios, y que supone compatibilizar en el mismo momento la demanda tecnológica de las personas con lo que los algoritmos y las herramientas pueden ofrecer. Es decir, la cantidad de datos que pueden gestionar las nuevas tecnologías no siempre va a la velocidad que demanda la sociedad usuaria. Esto es la denominada franja temporal y conecta con áreas tan prolíficas como la inteligencia artificial de las máquinas.

Ambos términos, utilidad y temporalidad, se encuentran ligados al hecho de que consumir contenidos audiovisuales en la Red provoca un cambio de actitud en los usuarios, favoreciendo la aparición de los formatos transmedia (**Scolari, 2013**) y del *prosumer* (**Jenkins, 2008; Napoli, 2008**), un consumidor que interviene en el *storytelling* (desarrollo de la narración) adoptando una actitud activa (**Gubbins, 2012**).

El consumo de contenidos audiovisuales a través de dispositivos como la tablet o el smartphone, por tanto, convierte al espectador en generador de datos. Esta circunstancia se produce cada vez que se identifica como usuario, comparte contenidos o indica que algo le gusta. El espectador de hoy en día genera información en el ecosistema de internet.

El negocio de la distribución de contenidos online es consciente de que obtener información de sus propios usuarios se convierte en una ventaja competitiva sin precedentes. En esta línea, *Netflix*, paradigma de la gestión interna de datos y explotación de metadatos, demuestra un uso de la información flexible y adaptativa al entorno y una consiguiente toma de decisiones basadas en *big data*. En apenas dos décadas, ha establecido un modelo de negocio que sitúa al consumidor en el centro de sus decisiones. En la práctica, esto se traduce en una aportación decisiva en términos de innovación en el consumo televisivo, con la tecnología como principal aliada. Gracias en parte a esta filosofía, *Netflix* es una marca global con más de 81 millones de suscriptores en todo el mundo que reproducen más de 125 millones de horas de contenido al día, según datos de la compañía correspondientes al 1<sup>er</sup> trimestre de 2016.

La estrategia aplicada a nuevos modelos mediáticos como el vídeo bajo demanda (*video on demand*, VOD) es un tema que suscita mucho interés académico. Como apunta **Izquierdo-Castillo (2015)**, las claves que configuran la estrategia de *Netflix* se resumen en tres elementos interrelacionados:

"un catálogo amplio y variado que incluya contenido *premium* en su oferta base; una tarifa económica y sin permanencia que permita acceso ilimitado al catálogo; y un servicio enfocado a la calidad de la imagen, la facilidad en la navegación, la búsqueda de contenidos afines y la accesibilidad a través de múltiples dispositivos".

Con este artículo queremos contribuir a esta línea de investigación, mostrando por qué los *big data* son considerados eje estratégico de los nuevos negocios mediáticos como el VOD, entre cuyas aplicaciones prácticas se encuentra, por ejemplo, la propuesta al usuario de contenidos personaliza-

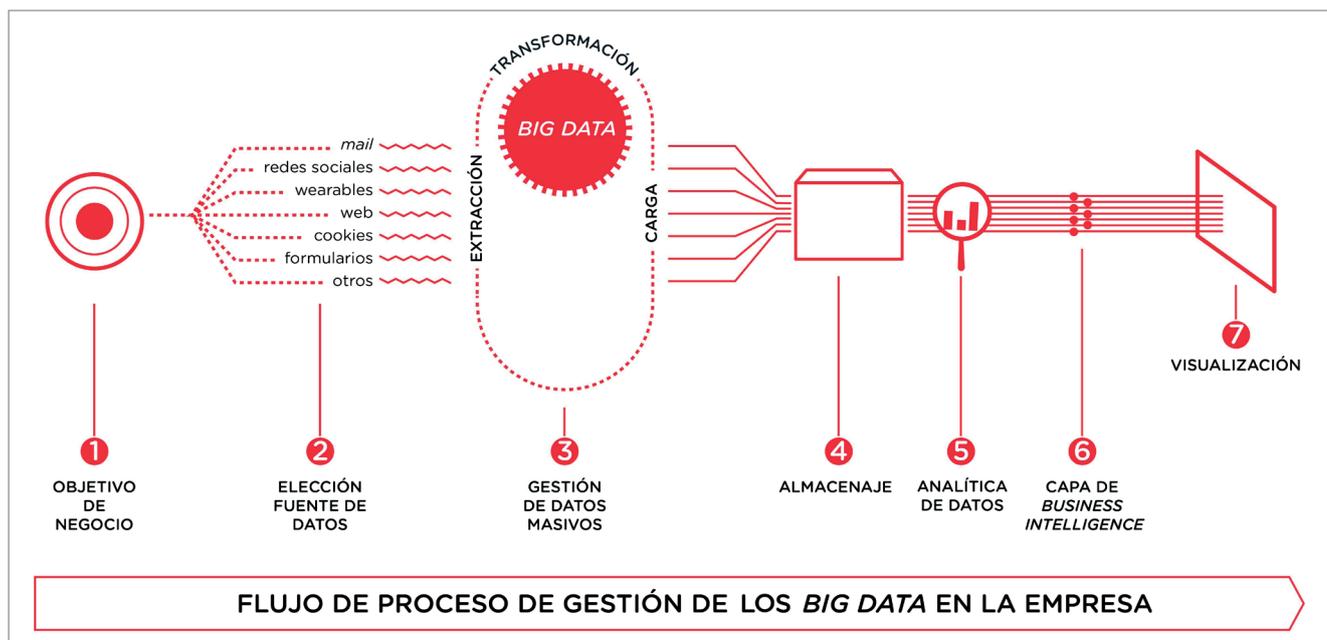


Imagen 1. Flujo de proceso de gestión de los *big data* en la empresab (Fernández-Manzano, 2016).

dos a través de una inteligente combinación de algoritmos de recomendación, la producción de contenidos propios o las nuevas adquisiciones de catálogo en sintonía con las demandas de sus clientes.

Como veremos a continuación, los *big data* son clave en la toma de decisiones concernientes tanto a la calidad técnica del servicio como al catálogo. Orientan sobre qué contenido ofrecer, cómo presentarlo y en qué cantidad. Ayudan, también, a tomar las mejores decisiones sobre las producciones propias (*Netflix Originals*).

El negocio de la distribución de contenidos online es consciente de que obtener información de sus propios usuarios se convierte en una ventaja competitiva sin precedentes

Esta generación de contenidos personalizada ha llevado a impulsar una estrategia de distribución diferente, adaptada a las recientes demandas de un nuevo tipo de espectador que reivindica libertad total a la hora de ver un contenido audiovisual, cómo, cuándo y dónde quiera (Neira, 2015). Su apuesta por el estreno simultáneo de series completas en su plataforma ha llevado al usuario al conocido como ‘atracción seriéfilo’ o *binge watching* (Jenner, 2014). Pioneros en este ámbito, se han convertido en referencia del cambio de modelo en el negocio de la televisión de pago.

## 2. Objetivos y metodología

Este artículo se plantea como objetivos comprender qué son los *big data* y conocer cuál es su flujo de gestión de datos en una empresa audiovisual. Así mismo, aporta nuevos argumentos al debate académico que permiten entender por qué esta gestión de datos masiva es considerada un eje estratégico en el negocio de la distribución audiovisual bajo demanda.

Para ello nos centramos en *Netflix* como estudio de caso. Se identifican sus objetivos de negocio, se analizan criterios de selección de fuentes, extracción de datos, transformación y carga, así como la monitorización y aplicación de sus algoritmos relacionados con la aplicación de la capa de toma de decisiones de negocio, también conocida como *Business intelligence*. Finalmente, procedemos a establecer las conclusiones del artículo atendiendo a los objetivos planteados.

Hacemos uso de la metodología del estudio de caso (Yin, 2009), trabajando con fuentes primarias y secundarias relativas a la gestión de datos de la compañía: revisión bibliográfica, análisis de los datos publicados por la propia empresa (*Netflix tech blog* y el área corporativa de *Netflix.com*), la información difundida por sus trabajadores en foros de debate y el análisis de datos facilitados por parte de la prensa especializada.

## 3. Gestión de datos en el negocio audiovisual

Las nuevas tecnologías permiten a las empresas almacenar y generar grandes volúmenes de información, incluso demandar acciones tecnológicas muy definidas a nivel de negocio. Tanto los usuarios como las empresas generan datos en su interacción con las nuevas tecnologías. Dicha producción de datos es tan prolífica que se ha denominado *big data* (grandes datos). Éste es el término empleado para referirse a toda aquella cantidad ingente de datos que, debido a sus características, no puede ser siempre procesada por los sistemas informáticos tradicionales. Esto puede deberse a una cuestión de volumen, lo que hará que se necesite trabajar con grandes almacenes de datos (*data centers*) físicos o en la nube. También podemos valorar características como la escala temporal o velocidad, lo que va a permitir establecer predicciones.

Partiendo del marco teórico expuesto por Fernández-Manzano (2016), relativo al flujo de gestión de los *big data* en la empresa, procedemos a incorporarlo en este artículo como guía que nos permita ubicar y analizar técnicamente las diferentes acciones que vamos a explicar en los próximos apartados.

La procedencia de los datos que se incluyen dentro de la terminología *big data* puede correlacionarse con fuentes tan dispares como la web y medios sociales, datos biométricos, transacciones bancarias o *machine to machine* (M2M), entre otros muchos. Como sugiere **Serrano-Cobos** (2014), hace años las empresas sólo tenían acceso a sus propios datos (propios o externos hechos por consultorías), pero en la actualidad, gracias a internet, estas empresas pueden contrastar su información con datos de mercado, opiniones de usuarios o incluso publicaciones científicas. El manejo de esta información genera una oportunidad crucial para plantearse correctas preguntas de negocio, así como tomar las mejores decisiones.

Podemos señalar como principales características de los *big data* los aspectos de volumen, variedad, veracidad y velocidad (**Martínez-Martínez; Lara-Navarra**, 2015):

- volumen: entendemos la gran cantidad de datos disponibles o accesibles;
- variedad: procede de la diversidad de fuentes de procedencia de la información imágenes, texto, música, cifras, tablas de *Excel*, diversos tipos de lenguaje, etc.;
- la calidad de los datos implica la característica de veracidad, dado que de ella depende que la información pueda ser útil o no;
- vuelve a incorporarse el valor de la velocidad: a mayor cantidad de datos, mayores posibilidades de tomar decisiones en menor tiempo; esto es una clara alusión al área de inteligencia de negocio.

#### 4. Estrategia de *Netflix* en el análisis de datos como estudio de caso

El flujo de procesos de una empresa dedicada a la distribución de contenido audiovisual bajo demanda, consciente de la utilidad de los datos, comienza con un planteamiento estratégico de negocio en el que se identifican los objetivos que se quieren conseguir; esto permitirá establecer sus métricas e indicadores de rendimiento, también conocidos como KPI (*key performance indicators*).

Aplicar estas lógicas a *Netflix* supone, como ya indicábamos, entender al usuario como generador de datos. Los 81 millones de suscriptores a lo largo y ancho del planeta, según datos del primer trimestre de 2016, que se traducen en 125 millones de horas de contenido reproducido al día, adquieren una magnitud dentro del negocio que va más allá de lo cuantitativo. En la práctica, cada una de esas horas de consumo en un dispositivo conectado a internet (*streaming*) lleva aparejados paquetes de información individualizada, cuya gestión y analítica ayuda a *Netflix* a comprender su situación en el mercado y también a identificar sus objetivos. Todo esto le permite diseñar un flujo de procesos de gestión de datos acorde con sus pretensiones.

#### 4.1. *Big data* en los objetivos de negocio

*Netflix*, autodenominada la “televisión por internet líder en el mundo”, basa su modelo de negocio en la suscripción y no en la inversión publicitaria propia de la televisión en abierto gratuita, pública o privada (**Clares-Gavilán; Ripoll-Vaquero; Tognazzi-Drake**, 2013). Utiliza los datos obtenidos de sus usuarios no para negociar con los anunciantes o las agencias de medios, sino para captar y retener suscriptores, minimizar las bajas, conseguir fidelidad con el servicio a largo plazo, y lograr buenos índices de satisfacción con su producto (**Gómez-Urbe; Hunt**, 2015).

“*Netflix* ha establecido un modelo de negocio que sitúa al consumidor en el centro de sus decisiones”

Sin descuidar este objetivo principal, es necesario enumerar todo un conjunto de subobjetivos, estrechamente relacionados con el primero, que se articulan para garantizar su cumplimiento. Así pues, los datos sirven a otros propósitos como los indicados a continuación:

- control de calidad de los contenidos cuyos derechos han adquirido (**Govind**, 2014);
- tamaño óptimo de su catálogo: como explica Jenny McCabe en *Netflix quick guide* (**McCabe**, 2013), *Netflix* no aspira a tenerlo todo, sino a garantizar que el contenido que tienen interés y se consume. De ahí la rotación de su catálogo y la no renovación de títulos con menor acogida entre sus usuarios (menos rentables en términos de volumen de audiencia y coste de su licencia). Las páginas que, mensualmente, informan de los títulos que están a punto de salir de la oferta de *Netflix* gozan de gran popularidad;
- posibilidad de ofrecer recomendaciones personalizadas de contenidos a sus suscriptores y ayudarles a elegir evitando que se sientan abrumados por el volumen del catálogo;
- producir contenidos teniendo en cuenta las preferencias de sus clientes.



Imagen 2. Usuario viendo *Netflix* mediante dispositivo conectado  
<https://media.netflix.com/es/company-assets>

La combinación de estos elementos ayuda a garantizar el cumplimiento del primer objetivo de negocio citado: captar y evitar la baja de suscriptores. De todos ellos, el sistema de recomendación se sitúa en el núcleo de la retención de usuarios y de la fidelización con el servicio a largo plazo. Las bajas en el servicio, como apuntan **Gómez-Uribe y Hunt (2015)** son más frecuentes en el mes gratuito de prueba ya que:

“el servicio de recomendación realiza un trabajo satisfactorio ayudando a elegir a los clientes con un largo historial de uso, pero no es tan bueno en el caso de nuevos suscriptores, de los que se sabe muy poco”.

El análisis de datos que les permite conocer a sus clientes, por tanto, es una de las claves estratégicas que garantizan el éxito de este servicio de VOD. Al no depender de la publicidad, el beneficio de *Netflix* es proporcional al número de suscriptores que pagan una cuota. Como indica **McCabe (2013)**, para que el pago de cada licencia de contenido sea rentable, deben asegurarse unos índices óptimos de visionado.

Los *big data* resultan clave en la toma de decisiones concernientes tanto a la calidad técnica del servicio como al catálogo. Orientan sobre qué contenido ofrecer, cómo presentarlo y en qué cantidad. Incluso ayudan a tomar mejores decisiones (*business intelligence*) en torno a las producciones propias que deciden poner en marcha (*Netflix Originals*)

Por ello, el conocimiento de sus usuarios es la piedra angular de su *modus operandi*: permite la consecución del objetivo ya descrito, gracias a la mejora de las experiencias de usuario y a la constante búsqueda de la eficiencia en el tipo de contenido que ofrecen.

#### 4.2. Elección de fuentes de datos

En la selección de fuentes de datos se encuentran implícitos los objetivos de negocio. En el caso de *Netflix* vendría definida por la información que proviene de su propio usuario. Hablamos de un servicio que demanda un dispositivo conectado a la Red permitiendo muchísimas posibilidades de obtención de datos, interacción y convergencia.

Un usuario medio destina 60-90 segundos a decidir qué ver. Transcurrido ese período las posibilidades de que termine consumiendo un contenido se reducen ostensiblemente (**Gómez-Uribe; Hunt, 2015**). Esta ventana temporal es clave para entender la monitorización que *Netflix* realiza sobre los hábitos de sus consumidores. Como se desprende de las declaraciones de Mohammad Sabah en *Hadoop summit* (**Sabah, 2012**) y de Todd Yellin en *How Netflix uses big data* (**Yellin, 2015**), se analizan aspectos como:

- tipo de contenido reproducido o descartado;
- características de la reproducción (si se realiza a ritmo normal, si se avanza o retrocede, se pausa o se abandona);
- intensidad de la reproducción (frecuencia y número de horas de *streaming* por sesión);
- valoración que el usuario le otorga al contenido (teniendo

en cuenta las valoraciones otorgadas a otros contenidos);

- tipo de dispositivo desde el que se accede (televisión conectada, tableta, móvil o descodificador).

También guarda registro de los caminos que sigue el usuario, búsquedas, ubicación geográfica, día y hora, ubicación del contenido dentro de la página de inicio cuando se decidió reproducir y hasta los contenidos recomendados descartados.

*Netflix* ha optado por centrar su modelo de negocio orientado a la demanda en la calidad técnica del visionado, así como en la capacidad de movilidad y accesibilidad de los contenidos (**Izquierdo-Castillo, 2015**). Por este motivo, evalúa la situación tecnológica de cada reproducción (**Govind, 2014**):

- cómo la tasa de almacenamiento de contenido (*rebuffer*) y la velocidad de su transmisión (*bit rate*) afectan el comportamiento del usuario (las horas que dedica al servicio);
- cómo se pueden crear experiencias personalizadas (según la tecnología que emplean o la calidad de su ancho de banda);
- cómo mejorar la entrega de contenidos (qué contenidos deberían ubicar en la cola de los servidores para que se reproduzcan más rápido, mejoras técnicas de audio, subtítulos, etc.).

Estos datos se cruzan con los metadatos (información de los propios datos) que *Netflix* asigna a todos los contenidos que incorpora a la plataforma (**Madrigal, 2014**): información técnica y artística, atributos de género o marcadores semánticos (etiquetas relacionadas con la narración de ese programa o película), entre otros. A partir de aquí se podrá constatar hasta qué punto el conocimiento de las preferencias de sus suscriptores está vinculado a una tasa elevada de utilización del servicio.

Los datos que el usuario proporciona de manera activa (cuando valora el contenido, rellena formularios en la fase de inscripción, inicia sesión como un usuario específico de los varios que existen en el hogar o decide registrarse con sus perfiles sociales) son testimoniales si se comparan con la información generada (no suministrada) a partir del uso, de manera involuntaria (qué se reproduce y qué se descarta, a qué hora se consume, desde qué aparato conectado, con qué frecuencia, etc.). Este modo de monitorización, realizado dentro del entorno del usuario mediante un dispositivo conectado, y no percibido en el consumo, genera valiosas respuestas no condicionadas que permiten extraer conocimientos sobre la motivación de los usuarios para ver un contenido u otro (**Brogan, 2015**). Este tipo de investigaciones, de corte etnográfico, se ha convertido en el motor de la reconfiguración de la estructura audiovisual en lo que a producción y distribución se refiere (**Izquierdo-Castillo, 2012**), promovida por parte de este nuevo agente que es *Netflix*.

#### 4.3. Proceso ETL (*extract, transform and load*) y sistemas de almacenaje de datos

El protocolo técnico de manejo de los datos se conoce como ETL (*extract, transform and load*, extracción, transformación y carga). Supone el proceso propiamente de gestión masiva de la información (*big data*) en el que se van a aplicar criterios referentes a aspectos técnicos como la calidad (característica de veracidad), pudiendo así disponer de información útil tanto por no contener errores (datos dupli-

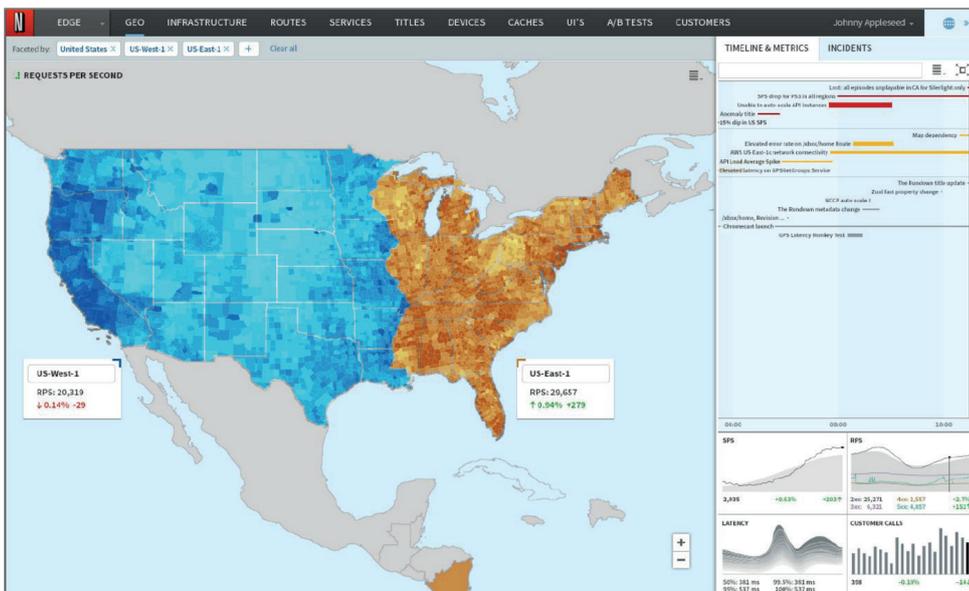


Imagen 3. Vista del front-end donde se presentan gráficos y datos dinámicos actualizados en tiempo real

cados, erróneos o inválidos) como por corresponder con los objetivos de negocio. El proceso de carga se enmarca dentro de la arquitectura de datos y va a definir aspectos como el modelo de integración de datos o dónde van a almacenarse. Las opciones habituales suelen ser dos:

- un gran repositorio (depósito o almacén) de datos, denominado genéricamente *data warehouse*;
- almacenaje en la nube.

En Netflix los datos se almacenan en la infraestructura de la nube de Amazon Web Service (AWS). Esta gestión a gran escala es la que permite alcanzar volúmenes de transmisión de contenidos de cerca de mil millones de horas al mes (Tse, 2015).

#### 4.4. Capa analítica

Una vez preparados los datos, esto es, una vez que se dispone de los *big data* acordes con las pretensiones del negocio, la siguiente acción a realizar es la denominada capa analítica que nos ofrece información sobre las métricas y los KPI (*key performance indicators*, o indicadores clave de rendimiento) establecidos previamente. La analítica de datos muestra información con idea de convertirse en próximas estrategias a realizar, pero también puede hablar de aspectos económicos como es el retorno de la inversión (denominado normalmente ROI, *return on investment*). Teniendo en cuenta que Netflix obtiene los ingresos por suscripción de abonados y no por publicidad (Clares-Gavilán; Ripoll-Vaquero;

Tognazzi-Drake, 2013), resulta un escenario ideal para lograr estrategias basadas en la información de perfiles segmentados. Lograr este retorno a través de los datos monitorizados supone comenzar un proceso de descubrimiento de valor en ellos. A tal efecto se realiza un proceso iterativo con grandes volúmenes de datos con idea de encontrar un modelo válido. Un ejemplo habitual en el consumo online puede ser la identificación de un algoritmo útil que detecte bien la información que queremos (¿en qué capítulo de una serie se fideliza la audiencia?), bien una conclusión (¿qué país tiene una mayor tasa de conversión del modelo de negocio?).

Entre los algoritmos más conocidos, Gómez-Urbe y Hunt (2015) destacan el *Personal video ranker*, una sofisticación de

El catálogo de Netflix descansa sobre una compleja red de algoritmos que, nutriéndose de la monitorización del consumo conectado de sus suscriptores, contribuye a formalizar decisiones de visionado que resulten satisfactorias. Como explican Gómez-Urbe y Hunt (2015), dependiendo de las características y las capacidades del dispositivo, Netflix puede llegar a mostrar 40 filas de selecciones (que acostumbran a corresponderse con un algoritmo específico). Dos de cada tres horas de contenido reproducido en Netflix, explican los autores, se “descubren” en este nivel. En la práctica, hay tantos modelos personalizados bajo la apariencia de selecciones recomendadas como usuarios suscritos a la plataforma.

Entre los algoritmos más conocidos, Gómez-Urbe y Hunt (2015) destacan el *Personal video ranker*, una sofisticación de

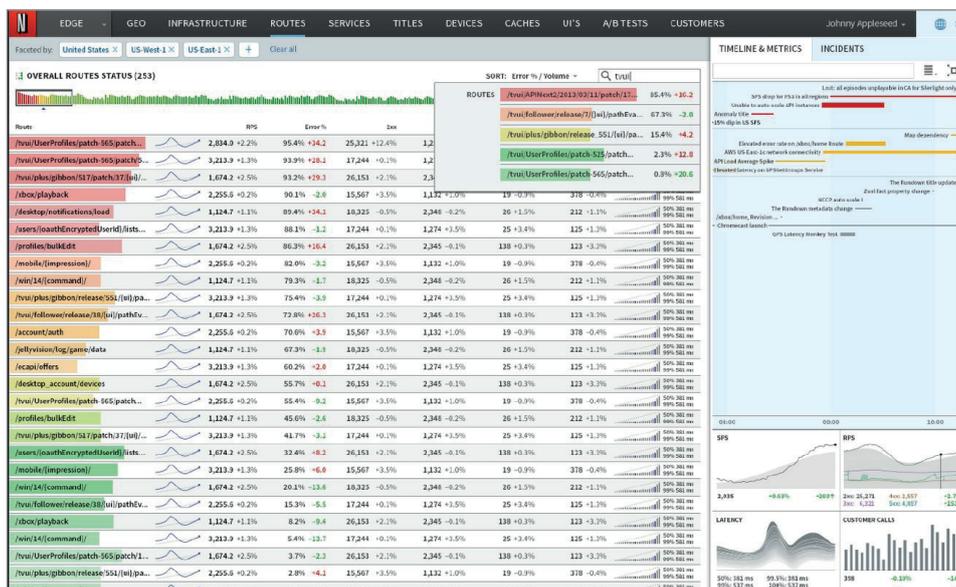


Imagen 4. En el front-end pueden verse las rutas de navegación seguidas por los usuarios, con tasas de petición, tasas de error y otros indicadores clave para cada ruta, que se actualiza casi en tiempo real.

la tradicional clasificación por géneros, hiperespecializada gracias al trabajo de asignación de metadatos a todo el contenido que se vuelca a la plataforma. *Netflix* cuenta con un equipo de personas (denominados *taggers*) cuyo trabajo consiste en visionar programas y asignarles etiquetas y categorías. Esta indexación de contenidos ha dado lugar a casi 80.000 microgéneros dentro de su ecosistema audiovisual (Madrigal, 2014). Les permite, además, encontrar conexiones entre películas y la creación de nuevas categorías (o subcategorías) en base a patrones como el origen de la producción, año, áreas temáticas, roles clave del equipo artístico o técnico, adjetivos ordenados por popularidad, etc.

Los autores también destacan *Top N video ranker*, que incluye todo lo que el usuario ve bajo el epígrafe *Top picks*, una selección de recomendaciones en las principales categorías generadas por el algoritmo anterior, filtradas según su popularidad pero personalizadas para cada cliente.

*Trending now* es el algoritmo encargado de generar categorías-tendencia, consecuencia lógica del entorno digital y social en el que está ubicado el nuevo espectador. Bajo este paraguas caen selecciones de contenido coyunturales basadas en picos de consumo que tienen un ciclo de vida corto como la Navidad, San Valentín o períodos vacacionales.

Como explica Yellin (2015), se monitoriza tanto la reproducción como el abandono de la misma. Dicha elección cobra todo el sentido cuando el algoritmo *Continue watching* se pone en acción. Genera una categoría que aglutina los títulos que previsiblemente el usuario querrá continuar viendo, tanto si se trata de contenido seriado (capítulos) o de un contenido específico cuya reproducción el usuario ha decidido ver en tandas o se ha visto forzado a interrumpir. Se trata de una categoría personalizada que tiene en cuenta todos aquellos datos que le permite saber qué contenido colocar en primer lugar (cuándo fue la última vez que lo vio, en qué punto de la reproducción se quedó, la frecuencia de visionado, dispositivo desde el que consume, tipo de conexión que emplea, etc.).

Otro de los algoritmos destacados por Gómez-Uribe y Hunt (2015) es el *Video-video similarity*, que genera la categoría *Because you watched*. Dentro de ella se encuentran los contenidos relacionados con un título que el usuario ya ha visto.

Aunque también existe un componente comercial que da mayor visibilidad a determinados productos (como los “Recién añadidos” o los “Estrenos”), la página de acceso a los contenidos de *Netflix* es un ecosistema dinámico que intenta adaptarse a la propia idiosincrasia de cada suscriptor. Por tanto, frente a una cultura audiovisual tradicionalmente orientada a ofrecer categorías estancas, en *Netflix* se observa una experiencia personalizada gracias a esta sensación

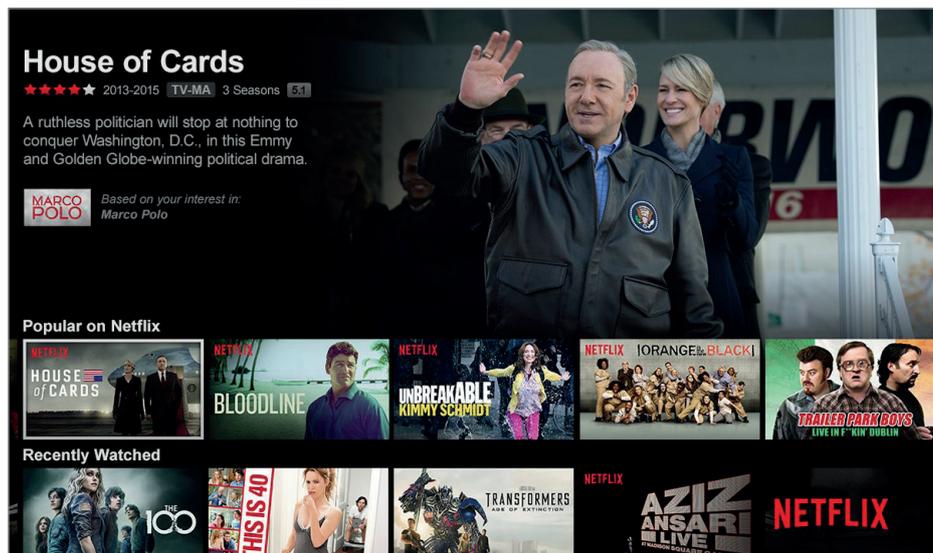


Imagen 5. Interfaz de *Netflix* <http://www.netflix.com>

de “producto nuevo”, fruto de la navegación por la plataforma. En la práctica, la selección algorítmica de contenidos representa el 80% de las horas de vídeo reproducidas por parte de sus usuarios (Gómez-Uribe; Hunt, 2015).

#### 4.5. Business intelligence

La información generada por los datos procedentes de la experiencia de navegación de los usuarios enlaza con la siguiente capa de negocio denominada *Business intelligence*, que es la destinada a tomar las mejores decisiones. En este sentido, tener una oferta de contenidos modelada en función de los gustos comprobados de sus usuarios presupone una toma de decisiones de negocio previa.

Es decir, estas lógicas de negocio parten de procesos previos, como la aplicación de algoritmos. En el caso de usuarios de una plataforma online es crucial poder establecer patrones de hábitos de visionado y consumo. A tal efecto, es útil el trabajo de aprendizaje automático, también llamado *Machine learning* (rama de la inteligencia artificial destinada a lograr que las computadoras aprendan), gracias al cual se evalúan modelos de datos para confirmar su validez y así aplicar lógicas de negocio. Esto es el *Business intelligence*.

La serie *House of cards* es el ejemplo recurrente cuando se habla de *big data* aplicados al diseño de contenidos. Como afirma Neira (2015):

“En 2012 el consejo de administración de *Netflix* tomó la decisión de crear una nueva unidad de negocio dentro de la compañía. Su objetivo era la creación de contenidos de producción propia para su plataforma (*Netflix originals*). Con varias propuestas sobre la mesa, concluyeron que la superposición de tres elementos básicos (el director David Fincher, la popularidad de la serie británica original y el actor Kevin Spacey) creaba un ‘círculo de éxito probado’ que garantizaba el triunfo del proyecto (Carr, 2013). Esa fue la conclusión, tras muchos meses de aplicación de minería de datos (...). *House of cards* era una apuesta segura por una confluencia de datos que ya tenían una relevancia positiva por separado,

pero al ser combinados dentro de un mismo contenido ofrecían una base incluso más sólida. El resultado fue una toma de decisiones de negocio basada en datos: la inversión de 3,8 millones de dólares por capítulo”.

El diseño de este tipo de productos “a medida” de los clientes es una manifestación clara de la capa de *Business intelligence*. Esto es posible gracias a que los *big data* perfilan un público potencial con unas dimensiones definidas cuya predisposición a visionar el contenido es el camino para conseguir lo que ellos quieren: maximizar la fidelización con el servicio.

Netflix utiliza los datos obtenidos de sus usuarios no para negociar con los anunciantes o las agencias de medios, sino para captar y retener suscriptores, minimizar las bajas, conseguir fidelidad con el servicio a largo plazo, y lograr buenos índices de satisfacción con su producto

## 5. Retos del análisis de datos de *Netflix* con su expansión internacional

Para una empresa de vocación global que a fecha de abril de 2016 dispone de servicio activo en más de 190 países, la personalización es la única vía para conseguir proveer de contenido audiovisual a todos los territorios en los que se encuentra presente sin descuidar la diversidad cultural de cada uno de ellos, atendiendo a los gustos audiovisuales locales.

Como explican **Raimond y Basilico** (2016), hasta 2015 las recomendaciones estaban principalmente basadas en el lugar de residencia. Cada vez que *Netflix* llegaba a un nuevo país, el servicio, además de un etiquetado local de contenidos, se reajustaba de manera manual. Esta aproximación a la recomendación pretendía evitar que la estandarización perjudicase la experiencia del usuario, y que se alejase de su cultura audiovisual. La expansión internacional afianzada en 2016 ha impuesto un cambio operativo: en lugar de escalar la aproximación del algoritmo a cada mercado, el sistema ha evolucionado aglutinando los países en regiones con catálogos similares y dejando que los algoritmos operen individualmente en cada región, pero procesando globalmente sus datos. Para evitar que los países pequeños perviertan el algoritmo, en los grandes, los datos de *engagement* o fidelización (identificados como reproducciones o intensidad de uso) se corrigen atendiendo al tamaño de los catálogos respectivos. No podemos olvidar que la selección que ofrece *Netflix* en los territorios en los que opera es distinta y varía en contenido y volumen. La gestión de derechos audiovisuales, a pesar de la falta de barreras territoriales en internet, sigue siendo territorial y delimitada a las distintas ventanas de distribución existentes: cine, *home-video*, VOD y televisión (**Clares-Gavilán**, 2014).

Por otro lado, su gestión de los *big data* ha crecido a medida que incorporaban más suscriptores al servicio (**Izrailevsky**, 2016) y se ha ido enriqueciendo con nuevas capas gracias a la expansión internacional. Los programas disponen ahora de nuevos territorios a modo de circuitos de pruebas, lo que ayuda a depurar el mapa de preferencias de contenidos

audiovisuales de los clientes y a identificar audiencias potenciales para sus productos.

La confluencia de la eficiencia en el diseño de programas, fruto de ese profundo conocimiento de sus clientes y la apertura a nuevos mercados, ha discurrido también en paralelo al incremento de *Netflix originals*. Este sería el caso de *Narcos* (2015), un drama seriado centrado en la figura de Pablo Escobar (narcotraficante colombiano, fundador y líder del Cartel de Medellín), con una importante proporción de diálogo en castellano. Esta producción llegó en un momento de fuerte consolidación y aumento de la audiencia hispana, tanto local, desde Estados Unidos, como internacional, y en paralelo al desembarco del servicio en varios países de Latinoamérica (**Betancourt**, 2015).

La expansión internacional les ha permitido también redimensionar a su público y, gracias a los *big data*, identificar los mismos segmentos de espectadores en distintos territorios. Sus nichos de audiencia son ahora globales. El contenido para audiencias generalistas, como *House of cards* (2013), *Jessica Jones* (2015) o *Fuller house* (2016), coexiste ahora con productos para los mencionados nichos globales. Las *sitcoms* (series cómicas televisivas situadas regularmente en los mismos escenarios) para *millennials*, como *Love* (2016) o *Masters of none* (2015) son un buen ejemplo de esta tendencia.

## 6. Conclusiones

En el caso analizado vemos cómo la extracción y manejo de datos, y su posterior análisis y toma de decisiones, se ha convertido en una ventaja competitiva para el negocio. Un negocio orientado a su oferta en función de la demanda y gustos del usuario.

El empleo de técnicas de minería de datos permite a la compañía *Netflix* la segmentación de sus clientes, y la identificación de clusters de usuarios a los que poder observar y ofrecer contenidos personalizados. En este sentido, estas técnicas han ayudado a crear su propia línea de negocio de producciones originales (*Netflix originals*) adaptadas en gran medida a los intereses y demandas de sus suscriptores.

Con el empleo de técnicas de minería de datos y de recomendación de contenidos hemos visto cómo logran también saber qué quieren o qué necesitan sus usuarios y con ello reducir las bajas de su servicio.

Netflix ha convertido el conocimiento de sus usuarios en la piedra angular de su *modus operandi*

Podemos concluir que la internacionalización del servicio ha sido, curiosamente, el paso decisivo para superar la fragmentación cultural por países. Gracias a su flujo de inteligencia de negocio, dinámico, circular y basado en modelos de *machine learning* (**Govind**, 2015) han podido evolucionar en la identificación de nichos más globales y en la producción de contenidos a medida de un público objetivo que ha crecido con la incorporación de nuevos mercados. *Netflix* ahora recomienda, adquiere y produce para el mundo.

De cara a futuras líneas de investigación consideramos que, de todos los aspectos señalados, sería pertinente ampliar el análisis a otros proyectos de VOD con el objetivo de comprobar si esta realidad es replicable a otros casos de estudio. Con ello, podríamos analizar, valorar e identificar tendencias que nos ayuden a seguir identificando por qué los *big data* son considerados eje estratégico en el negocio de la distribución audiovisual bajo demanda.

Asimismo, consideramos que el estudio de los *big data* aplicado al negocio de la distribución y consumo de contenido audiovisual bajo demanda sugiere comenzar nuevas líneas de investigación que entiendan la relación de las nuevas tecnologías con los usuarios y los creadores de contenidos.

## Bibliografía

**Betancourt, Manuel** (2015). "Netflix unequal marketing of 'Narcos' and 'Club de cuervos' ignores bilingual Latinos". *Remezcla*, 27 de octubre. <http://goo.gl/qcXmTf>

**Brogan, Jacob** (2015). "The case of the ornamental anthropologist. How Netflix puts a human face on big data". *Slate*, 31 May. <http://goo.gl/9Tu7Ja>

**Carr, David** (2013). "Giving viewers what they want". *The New York times*, 24 Febr. <http://goo.gl/lwuks8>

**Clares-Gavilán, Judith** (2014). *Estructura y políticas públicas ante los nuevos retos de la distribución y consumo digital de contenido audiovisual. Los proyectos de vídeo bajo demanda de cine Filmin y Universciné como estudio de caso*. Universitat Ramon Llull. Tesis doctoral. <http://www.tesisenred.net/handle/10803/247706>

**Clares-Gavilán, Judith; Ripoll-Vaquer, Jaume; Tognazzi-Drake, Alberto** (2013). *Distribución audiovisual en internet. VoD y nuevos modelos de negocio*. Barcelona: Editorial UOC. ISBN: 978 84 9029 911 1

**Fernández-Manzano, Eva-Patricia** (2016). *Big data. Eje estratégico en la industria audiovisual*. Barcelona: Editorial UOC. ISBN: 978 84 9116 380 0

**Gómez-Uribe, Carlos A.; Hunt, Neil** (2016). "The Netflix recommender system: Algorithms, business value, and innovation". *AMC transactions on management information systems (TMIS)*, v. 6, n. 4, article n. 13. <http://dx.doi.org/10.1145/2843948>

**Govind, Nirmal** (2014). "Optimizing the Netflix streaming experience with data science". *Netflix tech blog*, 11 June. <http://techblog.netflix.com/2014/06/optimizing-netflix-streaming-experience.html>

**Gubbins, Michael** (2012). "Digital revolution. Active audiences and fragmented consumption". En: Iordanova, Dina; Cunningham, Stuart (eds.). *Digital disruption: Cinema moves on-line*. St. Andrews: St. Andrews Film Studies. ISBN: 978 09563730 7 6

**Izquierdo-Castillo, Jessica** (2012). "Distribución online de contenidos audiovisuales: análisis de 3 modelos de negocio". *El profesional de la información*, v. 21, n. 4, pp. 385-390. <http://dx.doi.org/10.3145/epi.2012.jul.09>

**Izquierdo-Castillo, Jessica** (2015). "El nuevo negocio mediático liderado por Netflix: estudio del modelo y proyección en el mercado español". *El profesional de la información*, v. 24, n. 6, pp. 819-826. <http://dx.doi.org/10.3145/epi.2015.nov.14>

**Izrailevsky, Yury** (2016). "Completing the Netflix cloud migration". *Netflix media center*, 11 February. <https://goo.gl/T1rqW8>

**Jenner, Mareike** (2016). "Is this TVIV? On Netflix, TVIII and binge-watching". *New media & society*, v. 18, n. 2, pp. 257-273. <http://dx.doi.org/10.1177/1461444814541523>

**Jenkins, Henry** (2008). *Convergence culture: la cultura de la convergencia de los medios de comunicación*. Barcelona: Ediciones Paidós. ISBN: 978 84 493 2153 5 <http://dx.doi.org/10.1177/1461444814541523>

**Martínez-Martínez, Silvia; Lara-Navarra, Pablo** (2015). "El big data transforma la interpretación de los medios sociales". *El profesional de la información*, v. 23, n. 6, pp. 575-581. <http://dx.doi.org/10.3145/epi.2014.nov.03>

**Madrigal, Alexis** (2014). "How Netflix reverse engineered Hollywood". *The Atlantic*, 2 Jan. <http://goo.gl/7XCSWA>

**McCabe, Jenny** (2013). *Netflix quick guide: How does Netflix decide what's on Netflix*. Jenny McCabe, Director of global media relations, Netflix. <https://www.youtube.com/watch?v=VvpoUh9gx58>

**Napoli, Philip M.** (2008). "Toward the model of audience evolution: New technologies and the transformation of the media audiences". *McGannon Center working paper series*. Paper 15. [http://fordham.bepress.com/mcgannon\\_working\\_papers/15](http://fordham.bepress.com/mcgannon_working_papers/15)

**Neira, Elena** (2015). *La otra pantalla. Redes sociales, móviles y la nueva televisión*. Barcelona: Editorial UOC. ISBN: 978 8 4911 6116 5

**Raimond, Yves; Basilico, Justin** (2016) "Recommending for the world". *Netflix tech blog*, 17 Febr. <http://techblog.netflix.com/2016/02/recommending-for-world.html>

**Sabah, Mohammad** (2012). *Hadoop Summit*, 2012. Mohammad Sabah, Principal data scientist, Netflix. Silicon Angle. [https://www.youtube.com/watch?v=9h6wQXuke\\_E](https://www.youtube.com/watch?v=9h6wQXuke_E)

**Scolari, Carlos A.** (2013). *Narrativas transmedia: cuando todos los medios cuentan*. Barcelona: Deusto Ediciones. ISBN: 978 84 234133 6 2

**Serrano-Cobos, Jorge** (2014). "Big data y analítica web. Estudiar las corrientes y pescar en un océano de datos". *El profesional de la información*, v. 23, n. 6, pp. 561-565. <http://dx.doi.org/10.3145/epi.2014.nov.01>

**Tse, Eva** (2015). "Netflix case study". *Amazon web service*. <https://aws.amazon.com/es/solutions/case-studies/netflix>

**Yellin, Todd** (2015). *How Netflix uses big data*. Todd Yellin, VP of innovation, Netflix. <https://www.youtube.com/watch?v=Pu4myXu0ji0>

**Yin, Robert K.** (2009). *Case study research. Design and methods*. Thousand Oaks: Sage. ISBN: 978 1412960991

