

DATA MANAGEMENT IN AUDIOVISUAL BUSINESS: NETFLIX AS A CASE STUDY

Gestión de datos en el negocio audiovisual: *Netflix* como caso de estudio

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

Nota: Este artículo puede leerse traducido al español en:
http://www.elprofesionalde lainformacion.com/contenidos/2016/jul/06_esp.pdf



Eva-Patricia Fernández-Manzano graduated in audiovisual communication and obtained her PhD degree (2005) from *Complutense University of Madrid*. She has a Master's in audiovisual business management from *Instituto de Empresa*, and a Master's in big data and business intelligence management from *Escuela de Organización Industrial (EOI)*, both in Madrid. She is a professor at the *Rey Juan Carlos University* on media management. Her research focuses on how media companies work with technology. She has published several papers and books about media and big data, the last one being *Big data, eje estratégico de la industria audiovisual* (UOC, 2016).
<http://orcid.org/0000-0001-7655-872X>

Universidad Rey Juan Carlos
Edificio Departamental, Despacho 207
Camino del Molino, s/n. 28943 Fuenlabrada (Madrid), Spain
evapatriciafer@gmail.com



Elena Neira has a bachelor of law and media studies (audiovisuals). She has worked as a market researcher in *Grupo Zeta*, and marketing manager in the movie distribution business (*Wide Pictures*). Currently, she is working as a consultant for several media companies. She specializes in online marketing, social media, VoD, and new audiovisual distribution models. She is a teacher and lecturer. Member of *Innovación Audiovisual*. Author of *El espectador social* (2013) and *La otra pantalla* (2015).
<http://orcid.org/0000-0002-8565-2889>

La otra pantalla
Rosselló, 421 - 4, 1. 08025 Barcelona, Spain
elena@laotrapantalla.com



Judith Clares-Gavilán is a lecturer and researcher at the *Faculty of Information Sciences and Communication* at the *Open University of Catalonia (UOC)*. She holds a degree in journalism and a PhD in communication from the *Ramon Llull University*. Her main research interests are: structure and policies of the audiovisual sector and especially the changes in the industry with the advent of audiovisual distribution through internet, video on demand (VoD), and online TV. Member of the *GAME* group, she is a researcher of the R+D project "Cultura lúdica, competencia digital y aprendizajes" funded by the *Ministry of Economy and Competitiveness* of Spain.
<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, Spain
jclares@uoc.edu

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.

Manuscript received on 31-03-2016

Approved on: 10-04-2016

Keywords

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

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.

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).

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. Introduction

The application of new technologies in the audiovisual field has led to a series of systemic changes in business models. Certain sectors and companies have been quick to adapt to the new trends and workflows, but for the great majority the transition has not been completed. The adaptation roadmap is marked by two fundamental aspects:

- utility
- time slot.

This is because the new digital methods for accessing and using audiovisual products require new modes of action on the user's part. We notice that, in most cases, consumers do not integrate these technological changes in their day-to-day use until they realize their usefulness, which can range from improved image quality, satisfaction that the recommendation system understands the individual consumer's tastes, or the possibility of viewing content on demand rather than pre-recorded. As **Fernández-Manzano** (2016) points out, in the case of big data or mass data management in audiovisual fields, "the data is of no value in itself, its importance stems from the use that is made of it". This is, therefore, the utility. Furthermore, a temporary nuance appears which affects both data management and users, it entails the immediate alignment of people's technological demands with what the tools and algorithms can supply. That is, the amount of data new technologies can handle does not always offer the speed that users demand. This is the so-called time slot and is linked to such prolific fields as artificial intelligence and machine learning.

Both terms, utility and temporality, are intrinsically linked to the fact that the viewing of online audiovisual content leads to a change in attitude amongst the users, favoring the emergence of transmedia formats (**Scolari**, 2013) and the *prosumer* (**Jenkins**, 2008; **Napoli**, 2008), a user who takes part in the *storytelling*, adopting an active role (**Gubbins**, 2012).

Therefore, viewing audiovisual content on devices such as *tablets* or *smartphones* makes the viewer a data generator. This happens whenever a user is identified, content is shared, or something is "liked". Today's spectator generates information in the internet ecosystem.

The audiovisual, on demand, distribution industry is aware that its competitive advantage comes from obtaining information from its own users. Thus, *Netflix* has become the exemplar of internal data management and use of metadata, demonstrating a use of information that is flexible and adaptable to its environment and results in decision making based on big data. In under two decades, *Netflix* has established a business model that places the user at the core of its decisions. In practice, this translates into a major contribution, in terms of TV viewing innovation, with technology as its main ally. Owing, to a certain degree, to this philosophy, *Netflix* is currently a global brand with over 81 million subscribers, reproducing over 125 million hours of content per day, according to official company figures, corresponding to the first quarter of 2016.

The strategies applied to new media models, such as video on demand (VOD), are of significant academic interest. As **Izquierdo-Castillo** (2015) points out, the key factors that make up *Netflix's* strategy can be summed up in three inter-linked elements:

"a wide and varied catalogue that includes *premium* content in its basic package; a permanence-free economical rate that provides unlimited access to the listings; and a service geared towards image quality, simple browsing, similar content tips and accessibility through multiple devices."

In this paper we aim to contribute to this research path, outlining why big data are considered a strategic axis of the most recent media enterprises such as Video on Demand (VoD), which provides the user personalized content by

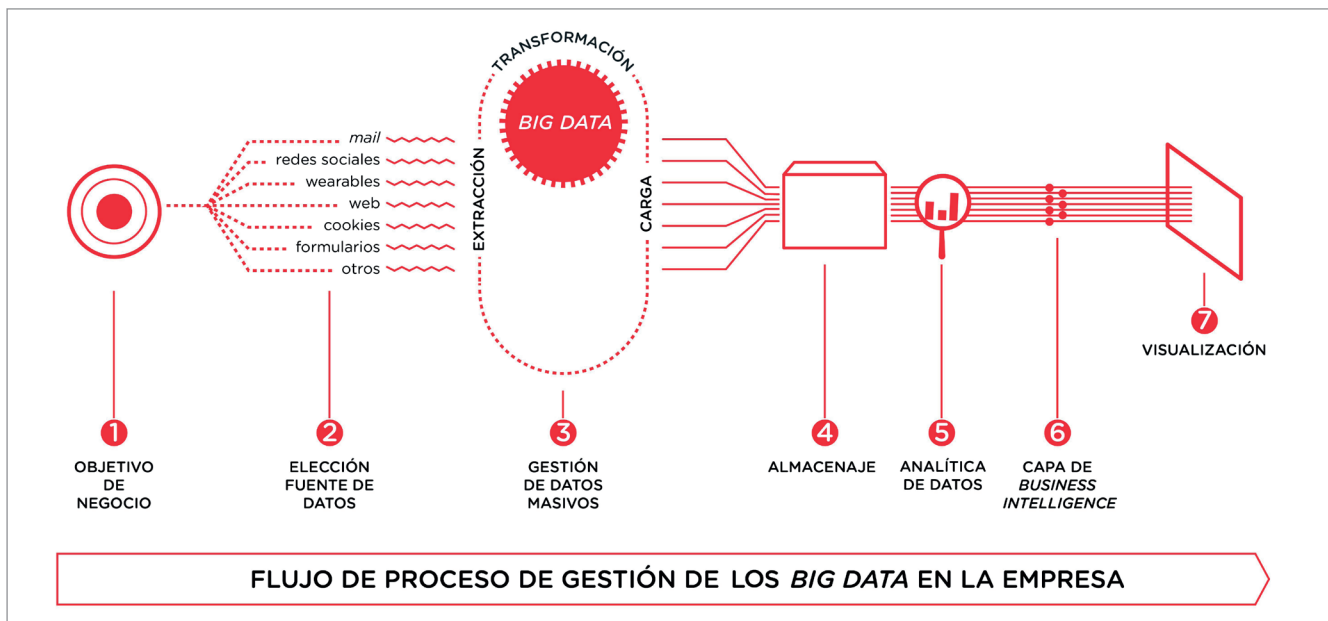


Figure 1. Process flow of big data management in the company (Fernández-Manzano, 2016)

means of an intelligent combination of suggested viewing algorithms, own content production, or new catalog acquisitions in line with client demand.

Big data are key to decision making for both the service’s technical quality as well as those pertinent to the catalog. It provides guidance on what content to offer, how to present it, and in what quantity. It also helps to make better decisions related to in-house productions (*Netflix originals*). This generation of personalized content has also led to promoting an alternative distribution strategy, adapted to recent demands by a new type of spectator who calls for total freedom when it comes to viewing audiovisual content, where, how, and whenever he wishes (Neira, 2015). Netflix took a gamble on releasing the entire season of original television shows at once, which ultimately led to what is now known as *binge watching* (Jenner, 2014). As pioneers in this field, Netflix has become a reference point for the changing model in the pay-TV industry.

“The audiovisual, on demand, distribution industry is aware that its competitive advantage comes from obtaining information from its own users”

2. Objectives and methodology

In this article we describe and define big data and data management. We will provide new arguments to the academic debate, and explain why mass data management is thought to be a strategic axis for the on demand audiovisual content distribution industry.

To do this we focus on *Netflix* as our case study. Its business targets will be identified, source selection, data gathering, transformation, and loading criteria will be

analyzed, as well as the monitoring and application of their own algorithms related to the application of the business decision-making layer, also known as *Business intelligence*. Finally, we will draw conclusions, focusing on the targets set.

To carry out this paper, we follow the methodology of the Yin (2009) study case, working with primary and secondary sources related to the firm’s own data management: bibliographic reviews, analysis of the data published by the firm itself (*Netflix Tech Blog* and the corporate area of *Netflix.com*), information provided by its staff in various discussion forums, and analysis of data provided by the specialized press.

3. Data management in the audiovisual business

New technologies are allowing companies to store and generate large volumes of information, and even demanding highly defined technological action, at a business level. In fact, users and businesses generate data in their interaction with new technology. The created data is so prolific that it has been termed big data. This is the term used to refer to the vast amount of data, which owing to its characteristics cannot always be processed by existing computing systems. This could be due to a volume issue, which will make it necessary to work with physical mass data storage (datacenters) or in the cloud. We can also take into account details such as time scales and velocity, enabling prediction making.

Based on the theoretical framework presented in Fernández-Manzano (2016), concerning big data management flow in business, we incorporate it into this paper as a guide to place and analyze the various technical aspects

Currently, the origin of the data included in big data terminology includes such dispirit sources as the web and the social media, biometrics, banking transactions, and machine to machine (M2M). As Serrano-Cobos (2014) suggests,

years ago firms only had access to their own data (in-house or through consultancies), but today, owing to the Internet, these businesses can contrast their information with market data, consumer feedback, or even scientific publications. Handling this information provides a crucial opportunity to come up with clever business questions, in addition to making the best decisions.

The main features of big data include volume, variety, veracity, and velocity (**Martínez-Martínez; Lara-Navarra, 2015**):

- volume, an enormous amount of available or accessible data;
- variety characteristically comes from diverse media sources, images, texts, music, figures, spreadsheets, different language types, and so on;
- data quality implies the need for veracity, given that its usefulness depends upon it;
- once again, the importance of velocity comes into play: the larger the data load, the greater the chance of speeding up decision making. This is a clear allusion to the business intelligence sector.

4. The *Netflix* strategy in data analysis as a case study

The process flow of a company devoted to on demand audiovisual content distribution, aware of the usefulness of data, begins with a strategic business approach in which objectives are set; this will establish their parameters and performance indicators, also known as KPIs (key performance indicators).

Applying this logic to *Netflix* involves, as previously mentioned, understanding the user as a data generator. The 81 million subscribers worldwide, according to 2016 1T data, represent 125 million viewing hours per day. They acquire such magnitude within the company that it goes beyond the merely quantitative. In practice, each of these hours consumed by means of a device connected to the internet (streaming) is linked to individual information batches which, when controlled and analyzed, helps *Netflix* to understand its niche and identify its goals. All this allows them to design a flow of data management processes in accordance with their aspirations.

4.1. Big data in *Netflix*'s business goals

Netflix, self-proclaimed "world's leading internet TV service", bases its business model on subscription rather than on the advertising typical of open television, whether it be public or commercial (**Clares-Gavilán; Ripoll-Vaquer; Tognazzi-Drake, 2013**). It uses data collected from its users, not to negotiate with advertisers or media agen-

cies, in order to improve members' retention, reduce cancellations, achieve long term fidelity, and obtain positive satisfaction ratings for their product (**Gómez-Urbe; Hunt, 2015**).

Without neglecting the main objective, it is paramount to list a complete set of sub-objectives, closely linked to this main one, in order to guarantee its compliance. Thus, data serves other functions, such as those listed below:

- Quality control of acquired rights (**Govind, 2014**).
- Optimal size of their catalog. As Jenny McCabe explains in *the Netflix quick guide for programming* (**McCabe, 2013**), *Netflix* does not aspire to have it all, but guarantees that their content is interesting and consumable. Hence, their catalog rotation and the non-renewal titles less popular amongst its users (not as profitable in terms of audience volume and licensing costs). Monthly pages containing information about upcoming titles are highly popular.
- The chance to offer subscribers personalized recommendations about content and help them choose without being overwhelmed by the size of the catalog.
- Produce content with their clients' preferences in mind.

“*Netflix* has established a business model that places the user at the core of its decisions”

All these elements combined help to guarantee compliance with the first business objective mentioned: attract and avoid subscriber cancellation. Above all, the recommender system is at the core for keeping subscribers and long-term customer fidelity. Service cancellations, as **Gómez-Urbe; Hunt (2015)** point out, are more frequent during the free trial month because

“the recommendation service does a satisfactory job helping long-standing clients, but is not as effective with new subscribers, of whom little is known.”

Data analysis which allows them to know their clients is,



Figure 2. User watching *Netflix* on an internet connected device.
<https://media.netflix.com/es/company-assets>

therefore, one of the strategic keys that guarantee the success of this VOD service. By not being advertising dependent, *Netflix's* profit is directly proportional to the number of paying subscribers. As McCabe points out in *the Netflix quick guide for programming* (McCabe, 2013), for the content license payment to be profitable, optimum viewing rates must be ensured.

Therefore, knowledge of their clients is the cornerstone of their *modus operandi*; allowing them to meet their aforementioned targets thanks to improved viewer experience and the constant search for efficiency in the type of content they offer.

Big data are key to decision making for both the service's technical quality as well as those pertinent to the catalog. It provides guidance on what content to offer, how to present it, and in what quantity. They also helps to make better decisions related to in-house productions (*Netflix originals*)

4.2. Netflix's choice of data sources

In data source selection lies the implicit business goals. In the case of *Netflix*, it would be defined by the information provided by the users themselves. We are talking about a service requiring an internet-connected device, which allows for a significant number of data gathering, interaction, and convergence possibilities.

The average user spends 60-90 seconds deciding what to watch. After this time the chances of viewing any content are greatly reduced (Gómez-Uribe; Hunt, 2015). This window is key to understanding the monitoring that *Netflix* carries out on its consumers' habits. As can be understood from Mohammed Sabah's statements in *Hadoop Summit* (Sabah, 2012) and Yellin's speech on *How Netflix uses big data* (Yellin, 2015), aspects such as:

- viewed or discarded content type;
- playback characteristics (if it plays at normal speed, if forwarded or rewind, paused or abandoned);
- playback intensity (frequency and number of *streaming* hours per session);
- rating the user gives the content (keeping in mind the ratings given to other content);
- the sort of device used to gain access (connected TV, tablet, mobile phone, or decoder).

It also keeps track of the paths taken by the user: searches, geographic location, time and day, where on the page the content was found when selected and even the discarded recommendations).

Netflix has opted to focus its business model on the demand for technical viewing quality, as well as on the mobile and accessibility capacity of content (Izquierdo-Castillo, 2015). This is why it evaluates the technological state of each playback (Govind, 2014):

- how the content storage rate and transmission speed (*re-buffer* and *bitrate*) affect user behavior (hours spent on the service);
- how personalized experiences can be created (depending on the technology used or the quality of their bandwidth);
- how to improve content delivery (how content should be placed on the servers so they can be played more quickly, technical audio improvements, subtitles, and so forth).

This data is crossed with the metadata (information in the data itself) that *Netflix* assigns to all the content it adds to the platform (Madrigal, 2014): technical and artistic information, genre attributes or semantic markers (tags related to the narrative of the program or film), amongst others. From here, it is possible to clarify to what extent knowing the subscribers' preferences is linked to a service's high rate of use.

The data that the users actively provide (when rating content, filling out forms during the enrollment phase, logging on as specific users from a household with several members, or deciding to register with their social profiles) are testimonial in comparison with information generated (not supplied) by use, in an involuntarily manner (what is played and what is discarded, at what time it is viewed, on what connected device, how often, and so on). This monitoring mode, carried out within the user's own environment and not perceived by any connected device, produces valuable unconditioned answers that enable the gathering of knowledge about the users' motivation to watch some content or other (Brogan, 2015). This type of research, of an ethnographic nature, has become the engine of audiovisual structure reshaping as far as production and distribution is concerned (Izquierdo-Castillo, 2012), promoted in part by this new agent that is *Netflix*.

4.3. The extract, transform and load process and Netflix's data storage systems

The data handling technical layer is known as ETL (extract, transform, and load). It involves the mass data (big data) management process in which criteria relevant to technical characteristics, such as quality (veracity characteristic), are applied. This allows the use of information that is useful, both because it does not contain errors (duplicate, erroneous, or invalid data) and because it corresponds with business goals. The loading process is framed within the data architecture and defines such aspects as data integration models or where it will be stored. There are usually two options:

- a huge data repository (warehouse or store), known generically as data warehouse, or
- storage in the cloud.

Netflix stores data on the *Amazon Web Service* (AWS) cloud infrastructure. This large scale management is what allows monthly streaming volumes of close to one billion hours per month (Tse, 2015).

4.4. The analytical layer at Netflix

Once the data has been prepared, that is, once the big data required for the business' needs are available, the next step to be taken is known as the analytical layer. It offers informa-

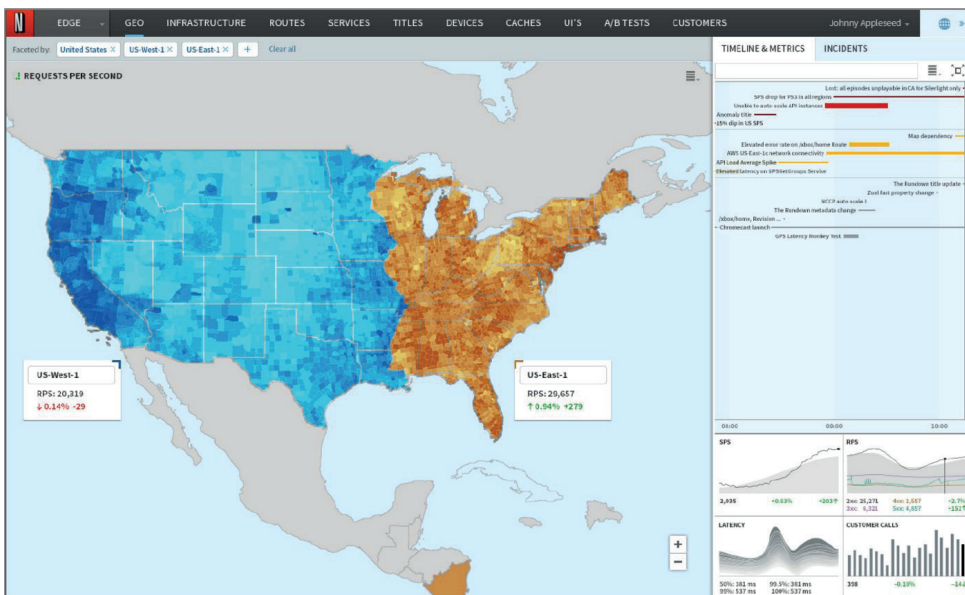


Figure 3. View of the front-end where graphics and dynamic data are presented and updated in real time

recommended viewing as users subscribing to the platform.

Amongst the better-known algorithms, **Gómez-Uribe and Hunt (2015)** highlight the *Personal video ranker*, a sophisticated upgrade of traditional genre classification, hyperspecialized due to the task of assigning metadata to all the content uploaded onto the platform. *Netflix* counts on a team of people (known as *taggers*) whose job it is to watch programs and assign tags and categories. This content indexing has given rise to almost 80,000 microgenres within its audiovisual ecosystem (**Madri-**

gal, 2014). It allows them, furthermore, to find similarities among films and breathe new categories (or subcategories) based on such patterns as production details, year, thematic fields, key roles of artistic or technical staff, adjectives ordered by popularity, and so on. The authors also highlight *Top N video ranker*, which includes everything the viewer watches under the label *Top Picks*, a set of recommendations for the main categories generated by the previous algorithm, filtered on the basis of their popularity, but personalized for each client. *Trending now* is the algorithm responsible for generating category trends, the logical result of the digital and social environment in which the new spectator is placed. Under this umbrella fall content choices relevant to consumption peaks with a limited life cycle, such as Christmas, Saint Valentine's

tion on metrics and previously established key performance indicators (KPIs). Data analysis shows information intended to become future strategies, but can also deal with economic aspects such as *return on investment* (ROI). Bearing in mind that *Netflix* has a ROI based on member subscription rather than on advertising (**Clares-Gavilán; Ripoll-Vaquer; Tognazzi-Drake, 2013**), it is the ideal setting to achieve strategies based on profiling information through segmentation. Achieving this return by means of monitored data involves starting a process to find out their value. To this end, an iterative process with large volumes of data is carried out in order to find a valid model. A common example in *online* consumption could be the identification of a useful algorithm that can detect either the desired information (in which episode of a series do the viewers become loyal followers?), or a conclusion (which country has the highest business model conversion rate?).

Netflix's catalog relies upon a complex web of algorithms that, nourished by monitoring its subscribers online use, aims to contribute towards formalizing satisfactory viewing decisions. As **Gómez-Uribe and Hunt (2015)** point out, depending on the characteristics of the device, *Netflix* can show up to forty selection rows (which tend to correspond to a specific algorithm). Two out of three hours of played *Netflix* content, explain the authors, are "discovered" at this stage. In practice, there are as many personalized models appearing under

the front-end where graphics and dynamic data are presented and updated in real time

recommended viewing as users subscribing to the platform. Amongst the better-known algorithms, **Gómez-Uribe and Hunt (2015)** highlight the *Personal video ranker*, a sophisticated upgrade of traditional genre classification, hyperspecialized due to the task of assigning metadata to all the content uploaded onto the platform. *Netflix* counts on a team of people (known as *taggers*) whose job it is to watch programs and assign tags and categories. This content indexing has given rise to almost 80,000 microgenres within its audiovisual ecosystem (**Madri-**

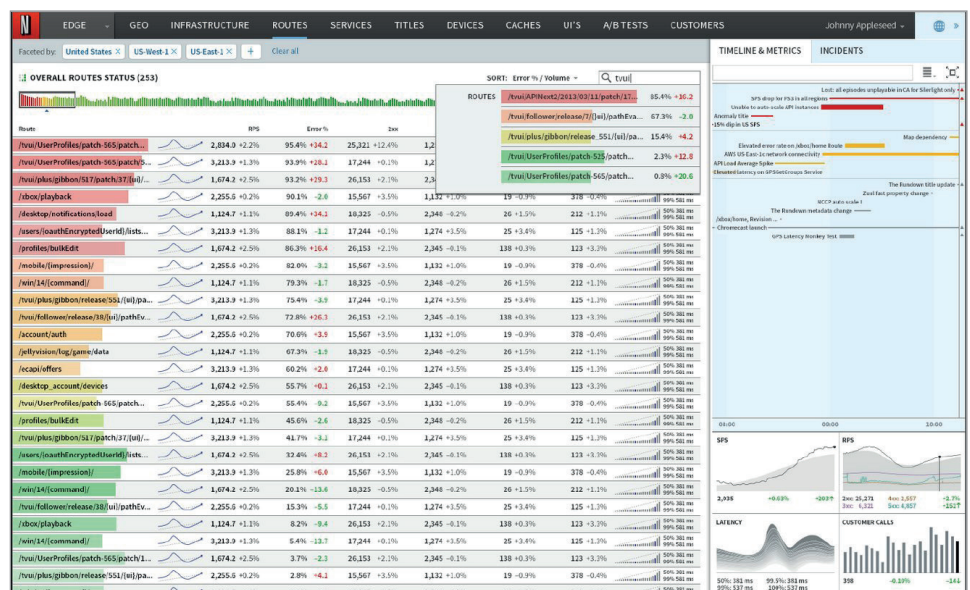


Figure 4. On the front-end can be seen the navigation routes followed by users, request rates, error rates and other key indicators for each route, which are updated almost in real time.

day, or the holiday periods.

As Yellin explains in *How Netflix uses big data* (2015), monitoring is performed on any playback as well as its abandonment. Such choice makes sense when the *Continue watching* kicks in. It creates a category that holds those titles that the user will most likely want to continue watching, whether it is serialized (episodes) or of specific content that the user has opted to watch in separate stages or has had to stop. It is a customized category that takes into account all the data that allows it to know what content to place first (last viewing time, where the reproduction was stopped, viewing frequency, device through which it was viewed, means of connection used, and so forth).

Another of the algorithms highlighted by Gómez-Urbe and Hunt (2015) is *Video-video similarity*, creates the category *Because you watched*. In this category we can find content related to a title that the user has already watched.

Even though there is also a commercial component that gives viewing priority to certain products (like “Recently added” or “Premieres”), *Netflix’s* access to the content page is a dynamic ecosystem that attempts to adapt to each subscriber’s idiosyncrasies. Therefore, faced with an audiovisual culture that is traditionally geared towards offering non-flexible categories, in *Netflix* we observe a highly customized experience due to this feeling of “novel product”, the result of browsing through the platform. In practice, the algorithmic content selection represents 80% of video hours played by its users (Gómez-Urbe; Hunt, 2015).

4.5. Business intelligence

The information generated by data from user browsing experience is linked to the next business layer, known as *Business intelligence*. It has the mission of making better decisions. In this sense, having a content range modelled in accordance with the users’ proven tastes presupposes previous business decision making.

This business logic stems from previous processes, such as the application of algorithms. For users of an online platform it is crucial to be able to establish viewing and consumption patterns. To this end, *Machine learning* (a branch of artificial intelligence aiming to get computers to learn) performs a useful task, thanks to which data models are assessed in order to confirm their validity and, thus, apply business logics. This is *Business intelligence*.

The series *House of cards* is a recurring example when discussing big data applied to content design. As Neira (2015) states,

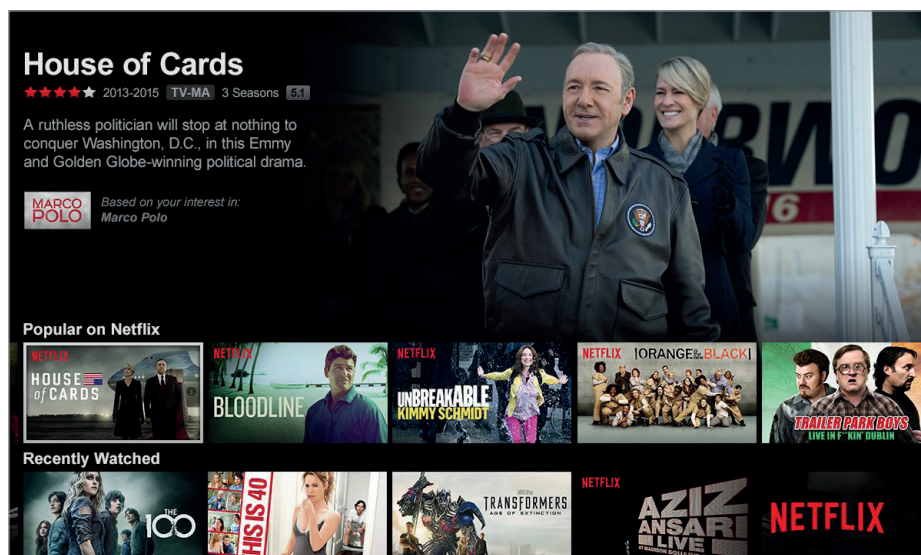


Figure 5. *Netflix* interface
<http://www.netflix.com>

“in 2012 *Netflix’s* management board decided to create a new business unit within the company, whose purpose was to make in-house productions for its platform (*Netflix Originals*). With several proposals on the table, they concluded that overlapping three key elements (director David Fincher, The popularity of the original British series, and actor Kevin Spacey) created “proven success circle”, guaranteeing the project’s success (Carr, 2013). That was the conclusion after many months of data mining (...). *House of Cards* was a safe bet due to a confluence of data that already stood out due to individual merit, but formed a much more solid base when combined in one single product. The result was a data based business decision: investment came to 3.8 million dollars per episode”.

The design of this type of “tailor-made” product is a clear example of business intelligence. It is made possible because big data profiles a potential audience with a set of defined dimensions whose readiness to view the content is the path towards achieving their goals; maximizing client fidelity.

5. Data analysis challenges for *Netflix’s* international expansion

For a global-minded company that, as from April 2016, has had an active service in over 190 countries, customizing is the only way to supply audiovisual content to all those territories in which it operates, without neglecting cultural diversity and catering to local audiovisual preferences.

As Raimon and Basilisco (2016) explain, until 2001, recommendations were mainly based on the place of residence. Every time *Netflix* went to a new country the service, as well as local content tagging, was reset manually. This recommendation approach intended to avoid standardization from damaging the viewing experience, and, thus, distancing itself from audiovisual cultures. Global expansion, strengthened in 2016, has imposed an operative change. Rather than sca-

ling the algorithmic approach to each market, the system has evolved by placing countries into regions with similar catalogs and letting the algorithms operate individually in each region, but processing their data globally. To prevent smaller countries from spoiling the algorithm, in larger countries, *engagement* data (identified as playbacks or intensity of use) are corrected based on each respective catalog size. We should not forget that the choice offered by *Netflix* in its various operating territories is different and varies in content and volume. The management of audiovisual rights, despite the lack of territorial boundaries on the internet, is still territorial and tied to the different existing distribution windows: home-video, VOD, and television (Clares-Gavilán, 2014).

“*Netflix* uses data collected from its users, not to negotiate with advertisers or media agencies, in order to improve members’ retention, reduce cancellations, achieve long term fidelity, and obtain positive satisfaction ratings for their product”

On the other hand, its big data management has been growing as more subscribers join the service (Izrailevsky, 2016) and it has been enriched with new layers due to global expansion. Programs can now use new territories as testing ground and it helps them to refine their clients’ audiovisual content preference maps and identify potential audiences for their products.

The confluence of efficiency in program design, the result of profound client awareness and new market openings, has also run parallel to the rise in *Netflix Originals*. This would be the case of *Narcos* (2015), a drama series based on the figure of Pablo Escobar, with a significant proportion of dialogue in Spanish. This production came at a time of strong consolidation and increasing Hispanic audiences, both local, from the United States, as well as international, and parallel to the service’s landing in a number of Latin American countries (Betancourt, 2015).

Global expansion has allowed them to resize their audience and, thanks to big data, identify the same viewer segments in the different territories. Their viewing niches are now global. General viewing content, such as *House of Cards* (2013), *Jessica Jones* (2015), or *Fuller House* (2016), now co-exist with products for the afore-mentioned global niches. *Sitcoms for millennials*, such as *Love* (2016) or *Masters of none* (2015) are a good example of this trend.

6. Conclusions

Having developed the paper, we can draw the following conclusions:

In the analyzed case we see how extraction and data management, and subsequent analysis and decision-making, has become a competitive business advantage. A business gearing its supply in relation to the demands and preferences of the user.

Using data mining techniques enables *Netflix* to segment its clients and identify user clusters in order to observe and offer personalized content. In this sense, these techniques have helped to set up its own production brand (*Netflix Originals*) adapted, largely, to subscriber interests and demands.

Through the use of data mining and content recommendation techniques we have been able to see how they manage to also know what their users want or need, and, as a result, lower the number of service cancellations.

We can also conclude that the service’s internationalization has been, curiously, the decisive step towards overcoming cultural fragmentation by countries. Owing to their business intelligence flow, dynamic, circular, and based on *machine learning* models (Govind, 2015), they have been able to evolve in the identification of more global niches and in tailor-made production content for an objective audience that has grown with the addition of new markets. *Netflix* now recommends, acquires, and produces worldwide.

Looking ahead to future research paths, we believe that of all the mentioned aspects it would be pertinent to widen analysis to other VOD areas in order to check if this reality is replicable to other case studies. In this way, we could analyze, evaluate, and identify trends that will help to continue identifying why big data are considered to be a strategic axis in the on demand audiovisual content distribution industry.

“*Netflix* knowledge of their clients is the cornerstone of their *modus operandi*”

Likewise, we consider that research on big data applied to the audiovisual content on demand distribution and consumption business calls for new research paths that can understand the relationship of new technologies with users and content generators.

7. References

- Betancourt, Manuel (2015). “*Netflix* unequal marketing of ‘Narcos’ and ‘Club de cuervos’ ignores bilingual latinos”. *Re-mezcla*, 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 video 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://media.netflix.com/en/company-blog/completing-the-netflix-cloud-migration>

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>

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>

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://www.theatlantic.com/technology/archive/2014/01/how-netflix-reverse-engineered-hollywood/282679>

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

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

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



El profesional de la
información


Bienvenido a EPI Indexada por ISI y Scopus
ISSN 1386-6710 / ISSN-e 1699-2407
Revista internacional, científica y profesional

<http://www.elprofesionaldelainformacion.com>

Revista internacional de
Información y Comunicación
Indexada por ISI Social Sciences Citation Index (Q3),
Scopus (Q2) y otras bases de datos

Factor de impacto JCR:
IF 2015 = 0,710

Scopus/SCImago Journal Rank:
SJR 2015 = 0,422



Presentación del Director

<http://www.elprofesionaldelainformacion.com>

Colección de libros de bolsillo

El profesional de la información (Editorial UOC)



NOVEDAD



El profesional de la
información

EDITORIAL UOC

Más información:

<http://www.elprofesionaldelainformacion.com/libros.html>