

CASE STUDY VI: DATA IN THE AUDIOVISUAL BUSINESS: TRENDS AND OPPORTUNITIES

prepared by Ms. Leticia Ange Pozza and Ms. Ana Paola Sifuentes

Pilot Project on Copyright and the Distribution of Content in the Digital Environment
Committee on Development and Intellectual Property
World Intellectual Property Organization
2021

Disclaimer

This study was commissioned as part of the Pilot Project on Copyright and the Distribution of Content in the Digital Environment¹ of the WIPO Committee on Development and Intellectual Property (CDIP). This document is not intended to reflect the opinions of Member States or the WIPO Secretariat.

1 INTRODUCTION: THE IMPORTANCE OF DATA IN THE PRESENT AND FUTURE

Everyone's relationship with data has changed in the past years. What started as bytes was quickly transformed into "the new oil" and "the sexiest jobs in the 21st century"², only to - maybe-not-so-quickly – realize that data raises issues related to power, organizational assets, privacy, and people.

Organizations of all sizes live nowadays in an ocean of data where people are its natural source. By the time children are born, especially, but not limited to high-income countries, they become drops of water in this ocean: there will be some trace of who they are in public and private datasets. At the same time, many organizations will be already predicting who they will become in life. Although this sounds very dystopian and far away from reality, numbers show that the current reality is that of big data, data-oriented, algorithmic and artificial intelligence (AI) ocean.

Technology employing data also lead to tailored products and services but it could also mean less privacy while possibly wrongfully judging millions of people with one misplaced intention embedded in a code. In this new world, being offline, unpredictable and untraceable is the real challenge. There are several studies linking people's personality traits to how they will behave when interacting with social platforms (likes, comments, or posts on Instagram, for example)³. These traits in association with algorithms that deliver content based on their social behavior are some of the ways social media has created content bubbles.

This reality is highly relevant as this is the digital environment that over-the-top services emerged and exist. Behind somewhat pervasive algorithms there are highly skilled programmers and decision-makers that understand how they work and apply those skills to their favor, as we have seen, for example, in video streaming platforms that base their recommendations on the preferences and likes of other users that have watched similar content. Data is now, centric and strategic for large and small corporations. Algorithms have become powerful tools, accelerating and changing the way we consume, communicate, perceive and access anything, from information, to goods and services, and even to ourselves.

It is predicted that by 2025 we will have more data than we can analyze. The estimate is that by the end of that year, only 15% of the data in the global datasphere will be tagged and only one-fifth of that will actually be analyzed⁴. Data, software and hardware are becoming valuable commodities. But how guaranteeing that the 3% of data analyzed are relevant to businesses and our society?

¹ Studies available at: https://dacatalogue.wipo.int/projects/DA_1_3_4_10_11_16_25_35_01

² Sexiest Job in the 21st Century <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>

³ Computer-based personality judgements are more accurate than those made by humans. <https://www.pnas.org/content/pnas/112/4/1036.full.pdf>

⁴ Data Age 2025 <https://www.import.io/wp-content/uploads/2017/04/Seagate-WP-DataAge2025-March-2017.pdf>

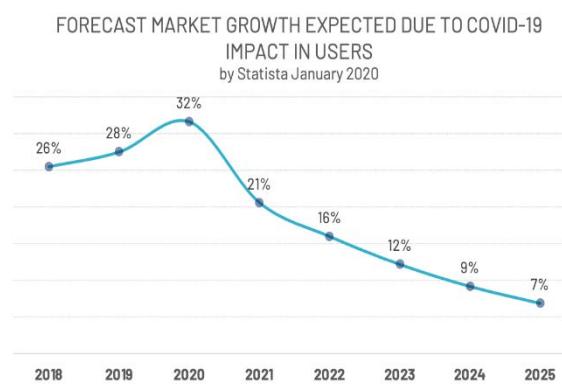
The technologies involved in the Fourth Industrial Revolution are all directly related to collecting, processing and analyzing data. These include major areas like mobile, the internet of things, artificial intelligence, augmented reality, advanced robotics, genomics, and blockchain⁵.

The creative industry is also a big part of this revolution. Major production companies can now base their decisions on data. For example, tests that were once conducted for a few hundred potential audience members are now done in milliseconds for millions of people, deciding which products are produced and which are not. This technology along with artificial intelligence and machine learning, has allowed content creators to directly cater to a specific consumer's palate and even anticipate their preferences and present them with highly desirable products.

2 DATA AND THE CREATIVE INDUSTRY

The creative industry has changed dramatically in the last few years, there are several contributing factors, the main one being the increase of internet bandwidth, even with prepaid internet access (3G and 4G) that allows users to watch content from their mobile devices regardless of their location. Video content in itself has exponentially increased, streaming platforms, also known as over the top (OTT), directly provide an extensive catalog of films and series to consumers. Some websites even allow users to be the content creators, the primary example of this is Youtube⁶, the largest and most popular video hosting platform worldwide. A more recent growing trend is stand-alone virtual reality gears (such as Oculus) that have paved the way to a new form of cultural consumption. These devices were so popular that they sold out during the pandemic⁷.

The COVID-19 pandemic has played its role in increasing the OTT activities and audiences. Subscription video on demand (SVOD), for example, reached its biggest growth in recent years, as shown in the graph below⁸.



Because OTT platforms provide services directly to consumers, they can collect, manage and measure data related to their catalogs and the specific users. Differently from previous distribution models, data collection offers an easier, centralized, and real-time possibility for feedback, allowing stakeholders to readjust their strategies and targeted audience. In OTT, while users stream the content, there is an opportunity to collect information regarding location, user profile, habits and preferences, equipment, and internet characteristics (also for technical purposes

of maintaining the experience at home, for example). The possibilities are endless, and it is only at the beginning of what can be accomplished with the data collected from and/or by OTTs⁹.

An example is that some OTT services combine the users' login from social media or email account to enrich customer's profile with information about social status, family composition, living

⁵ The Fourth Revolution <https://www2.deloitte.com/content/dam/Deloitte/global/Documents/Strategy/gx-strategy-ops-the-fourth-revolution-now.pdf>

⁶ 30 most popular Youtube stars in the world. <https://www.businessinsider.com/most-popular-youtubers-with-most-subscribers-2018-2>

⁷ Oculus surpasses \$100 million in Quest content sales https://techcrunch.com/2020/05/18/oculus-surpasses-100-million-in-quest-content-sales/?guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAFs-8JJyE6PFUHDSJWlkYEP9Hv5vd7a8UuBt-gBAjHAIKTEXJ8BPUHsHRWZyZ6XFGZbXqdWgyaxNTD6GU1Xn2uIRJGvMMNv0TmImVTNCZ-a9efZ2HjKsu5VHvhLSG7EFBI9mJoLZGOQhvOOBFEDbsjYked22ui8_lvqrXX79&guccounter=2

⁸ Statista Video Streaming <https://www.statista.com/outlook/206/100/video-streaming--svod-/worldwide>

⁹ 4 ways the entertainment industry is harnessing AI for growth. <https://www.analyticsinsight.net/4-ways-that-the-entertainment-industry-is-harnessing-ai-for-growth/>

conditions and other types of interests that could be available. What is a convenient option for the user, is an interesting opportunity for the OTT. This, combined and compared to other profiles, helps target and select what customers would like to watch. Below are a few ways companies, including OTTs, are using data to add value to their organization and their customers.

2.1 PERSONALIZATION ON A MASS SCALE WITH CONTENT MARKETING

A significant part of investment in series and films is allocated to marketing. Studios are increasingly using algorithms to help them creating promotional movie trailers, where systems analyze trailer footage frame by frame, labeling/classifying objects and events, and then comparing them to data generated for other trailers. The idea is that movies with similar sets of labels will attract similar types of people.

20th Century Fox has partnered with Google to use the company's servers and the open-source AI framework TensorFlow to create their experimental movie attendance prediction and recommendation system, called Merlin. First, Merlin scans the trailer, labeling objects like "forest," "car," and "animal". Then, analyses the frequency of these labels, how long these objects appear on-screen, and the order and timeframe in which they appear. 20th Century Fox's engineers say that this temporal, or time-based information is particularly rich because it correlates with a film's genre. "For example," they explain, "a trailer with a long close-up shot of a character is more likely in a drama movie than for an action movie, whereas a trailer with quick but frequent shots is more likely in an action movie"¹⁰. The AI is used to predict user attendance and segment indexes a year in advance.

An example of this is 20th Century Fox's 2017 dark superhero film Logan. Its trailer features lots of slow shots of the main character Hugh Jackman looking bloody and beaten. By comparing this information with analyses of other trailers, Merlin tries to predict what films might interest the people who saw Logan, thus, targeting and personalizing both the trailer and the audience on a massive scale.

The system gets quite a few of the films correct, including other superhero movies like X Men: Apocalypse, Doctor Strange, and Batman vs Superman: Dawn of Justice. It even correctly identifies John Wick: Chapter 2 as a pair for Logan. That's an impressive inference since John Wick is certainly not a superhero movie. All in all, Merlin identifies all of the top five picks, even if it does fail to put them in the same order of importance.

In any type of algorithm, there are always those recommendations that do not necessarily make sense. Merlin predicts that The Legend of Tarzan will be a big hit with Logan fans, for example. Neither Google nor 20th Century Fox explains this, but it could have something to do with the "forest," "tree," and "light" found in Logan, elements which also feature heavily in the Tarzan trailer. This shows how important a human decision is in these cases, training an algorithm requires business intelligence and experience, one of the hardest things to teach a machine¹¹.

Even with mistakes, the use of AI becomes an important asset to understand audiences. With fewer resources and time spent, they could substitute the lengthy process of targeted interviews and closed sessions with customers. These algorithms cross-reference numerous trailers and movies providing valuable input on how it should be promoted and who is more likely to be interested in the film.

¹⁰ Convolutional collaborative filter network or video-based recommendation systems <https://arxiv.org/pdf/1810.08189.pdf>

¹¹ Teaching machines to reason about what they see. <https://news.mit.edu/2019/teaching-machines-to-reason-about-what-they-see-0402>

2.2 CONTENT CREATION

Writing is a creative area where artificial intelligence and machine learning algorithms have also started to be employed. They have been used for simple factual text creation, with one-third of Bloomberg's news using some kind of automation to be written - something called "robot journalism" or "automated journalism"¹². In a similar use, Sverker Johansson¹³ created a computer program to write Wikipedia articles. The bot has written a total of 2.7 million articles, which is about 8.5 percent of the total articles on Wikipedia—more than any human has ever written. At some point, we have all read a partially automated article, and we will see that increase a lot in the next few years.

AI has also been applied to writing fictional stories. Sunspring¹⁴ is an experimental short film, based on an algorithm and produced previously in 2016. Americans Oscar Shapin and Ross Godwin programmed a computational system they named Benjamin and fed it with science-fiction movie scripts, including X-Files, Star Trek, and 2001, A Space Odyssey. Benjamin analyzed similarities and repetitions in the scripts, drew up scenarios, and simulated a dialogue. It even "composed" part of the soundtrack, based on 30.000 pop songs.

The credits of the film read: « WRITTEN BY Benjamin, a system-on-chip (SOC) computer with graphics processing unit (GPU), running a "long short-term memory (LSTM), Recurrent neural network (RNN)". The experiment seemed promising, but the feedback on the script was not good. Although they can assist in the development of texts, we are yet to see a full computer creation that humans approve of.

Another recent example is the British series Black Mirror from Netflix. Their film, Bandersnatch¹⁵, allowed the viewer to choose the upcoming scene from an array of choices. The viewer could choose between story options and control the fate of the protagonist - a game developer named Stefan. And this is where Netflix user data kicks in: Stefan's first dilemma was whether to eat Sugar Puffs cereal or Kellogg's Frosted corn flakes for breakfast. Sixty percent of the viewers opted for Frosted Flakes. A choice that does not seem relevant to the story, is data that can be combined with user profiles for branded content opportunities and marketing research. Paired with profiling data, interaction can be an infinite source of data collection in the creative process.

Although we are still in the early stages of this development for scripting with algorithms, other formats of data collection help determine the direction of the content development as it is on air. An example of using data to make strategic choices for Netflix's original productions is the development of the series House of Cards. The main advantage in using and benefitting from user data was to identify viewer profiles and viewer preferences. This enabled Netflix to make a connection between a large potential audience and content that would match the viewers' expectations¹⁶.

In 2011 Netflix invested nearly 100 million dollars in the first two seasons of House of Cards, which was an adaptation of a British series by the same name. (The original House of Cards was broadcast on BBC in 1990). The new version of House of Cards is produced by David Fincher

¹² Putting Europe's robots on the map <https://reutersinstitute.politics.ox.ac.uk/our-research/putting-europe-s-robots-map-automated-journalism-news-agencies>

¹³ Sverker Johansson https://es.wikipedia.org/wiki/Sverker_Johansson

¹⁴ Sunspring <https://www.youtube.com/watch?v=LY7x2Ihqjmc>

¹⁵ Bandersnatch. https://en.wikipedia.org/wiki/Black_Mirror:_Bandersnatch

¹⁶ Giving viewers what they want <https://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html>

who also directed the first episodes. The investment was a risky one and it was based on audience data Netflix had been gathering from the 33 million subscribers it had at the time.

The data showed that the original British version of House of Cards was very popular and that most viewers watched director David Fincher's feature film The Social Network from the beginning to the end; this showed interest and commitment to Fincher's work. Lastly, the data showed that films starring Kevin Spacey were very popular. By combining these findings, Netflix ended up producing an extremely successful American version of House of Cards.

Netflix also chose a data-driven marketing campaign for *House of Cards*. The main poster image was designed by a machine learning system and it replicated the style and colors of earlier successes, such as the film *Macbeth*. Promotional trailers were also produced with the help of data analysis, with season five of *House of Cards* having up to ten different trailers, each targeted at a specific audience segment. For fans of Kevin Spacey, the trailer showed him in action, whereas fans of *Thelma and Louise* were shown trailers with female characters, such as Claire Underwood played by actress Robin Wright.

Amazon's Prime Video had a different approach to using audience data¹⁷. Amazon produced pilot episodes of several series. Based on the audience reaction, Amazon decided which pilots were turned into longer series. Thereby, the decision rests on the audience's acceptance and not Amazon Studios itself. In addition to the audience's comments, each pilot is evaluated on the following criteria: Moments the viewer pauses viewing; Re-watched scenes; Skipping over scenes; Scenes that create a specific reaction (based on collecting social media posts and reactions); How well is the casting liked by viewers (also based on collecting social media posts and reactions).

By combining data from these points Amazon concluded that out of the pilot episodes, the audience wanted a situation comedy about four republican US senators, - and from that, the *Alpha House* series pilot went into production. The series concept relied on audience data and left creative freedom to its writers, as opposed to the development and writing of *House of Cards*.

This comparison indicates that while the audience data is important, the real difference lies in how it is used. Amazon decided to produce a series that was statistically tailored to match the tastes and expectations of the viewers, whereas Netflix provided the creator's data and conclusions which would help them in guiding creative choices, and, most importantly, it allowed more risk-taking.

2.3 SPEECH RECOGNITION

One of the early uses of AI in television has been transcribing spoken language into written text to generate automated subtitling. The usage of algorithms in this context can help to recognize speech (when are they speaking?), cluster speech (who speaks and when?), language (which language is being spoken?), speech (what is being said?), and synchronize video and script. Companies like Amazon, Hulu and HBO are now using AI to create flawless fully or semi auto-generated subtitles¹⁸. Usually, those systems use neural networks to analyze all of the vocabulary used in dialogues to ensure accurate subtitling. Speech-to-text technologies could become especially important also to broadcasters who often are required to provide audiovisual content in an accessible format. Subtitling is an important feature not only for the accessibility of the

¹⁷ Alpha house vs House of Cards. <https://medium.com/@cxdai99/alpha-house-vs-house-of-card-how-data-analysis-helped-netflix-but-not-amazon-92c71eb55b3e>

¹⁸ Subtitling Captioning Streaming Platforms. <https://www.betranslated.com/blog/subtitling-captioning-streaming-platforms/>

audiovisual content but also to reach audiences in different territories. In an OTT market that can distribute the content almost instantly around the globe, the availability of subtitles in different languages becomes an important feature to allow this larger distribution.

Speech recognition is also being used to improve content classification and flag potentially problematic or harmful content, often with the use of sentiment recognition to identify anything that might be described as being sensitive¹⁹. This way, the process of moderating content can be done with much greater speed and efficiency than previously possible. Eventually, it will be possible to apply these technologies in real-time, with live audiences and television, flagging cursing and behaviors deemed inappropriate for live television. We already have something similar with Youtube auto-generated subtitles²⁰, which are created by Google's machine learning algorithms. The user interface operation by speech (voice control) is an invaluable addition improving accessibility to all users.

2.4 HOW DATA IS PAVING THE FUTURE IN THE AUDIOVISUAL INDUSTRY

Big data has transformed the entertainment industry allowing digital experiences to be designed and catered for each user. Programmers, designers and marketers can predict a person's preferences based on their data and automatically make changes or adjustments based on the user's previous click. Graphics, animations and ads are laid out to capture our attention and encourage us to continue interacting with the material and ultimately to make a purchase.

These types of considerations are massively present in digital experiences, from shopping websites to mobile games. They can be as simple as changing the default language based on our location, recommending what clothes to buy based on our browsing of a website's apparel, or as complex as a chat-bot using natural language processing to assist us in asking for information and interacting with a website²¹. This is reinforced by the increasing speed and availability of the Internet, including beyond the computer in the current digital environment of multi-screen users.

The widespread processing of different sets of data on a large scale has already affected the audiovisual industries. Information obtained from data processing shaped the way movies and tv series are promoted, with specific movie trailers generated and customized to target a specific user's profile. With an almost infinite catalog available to users, data plays an important role in prioritizing what audiences are shown on streaming platforms. Algorithms employed by OTT platforms will present you with the most enticing content personalized based on your preferences and personal information. The choices made by users feed the platforms with their choices, an important element for the investments made in the OTT's catalog and in particular in the commissioning or financing of new content. The audience's decisions might impact core elements of the production such as topic, theme, genre, setting, and even actors and directors of a film.

3 DATA AND VIDEO GAMES

The industry of video games is one of the most important of the entertainment industries. Video games have always had a strong connection with data. The code behind video games is constantly calculating multiple variables to determine what will happen (called game mechanics and physics). For example, when the user jumps, do they stomp on an enemy or fall into a hole? Based on the user's input, the code would have to decide if the player has won or lost in the end.

¹⁹ Deep learning for detecting inappropriate content in text. <https://link.springer.com/article/10.1007/s41060-017-0088-4>

²⁰ Youtube auto-generated subtitles. <https://vocalmatic.com/blog/how-to-automatically-generate-subtitles-for-youtube/>

²¹ 5 examples of Artificial Intelligence in web apps <https://hub.packtpub.com/5-examples-of-artificial-intelligence-in-web-apps/>

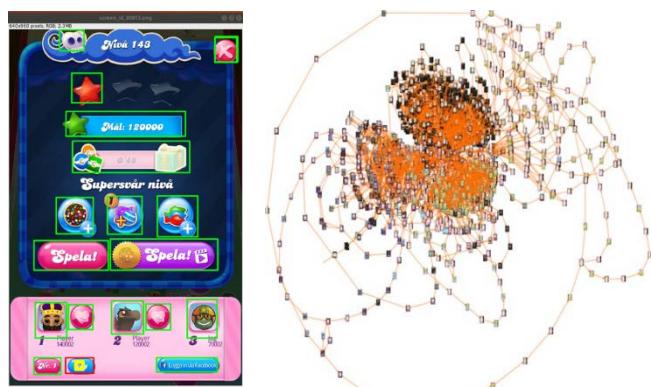
Through the last few decades, video game code has become more and more complex, with the programs having state of the art calculations and in some cases, using AI to respond to how the user is playing, adapting in real-time.

They also have become more widely accessible, not only being restricted to playing on a computer or a console but mobile devices, like smartphones. Thanks to this, video games are now a huge industry that continues to grow. While there were almost two billion video gamers across the world in 2015, this figure is expected to rise to over three billion gamers by 2023. According to estimates, there were over 2.9 billion video gamers in the world in 2020²². All of these users represent a huge amount of revenue with the global video game market being valued at USD 151.06 billion in 2019 and expected to grow at a Compound Annual Growth Rate (CAGR) of 12.9% from 2020 to 2027²³.

The design of video games has been impacted by data, especially in the case of online mobile games. More than 2 billion gamers produce 50 Tb of data per day, and the data collected through mobile gaming can be used to better cater ads to users and create a unique relationship between them and a brand. User's preferences and reactions can also be taken into consideration when designing the game²⁴.

AI can be used to give direct feedback to developers and to test their products. In the case of Candy Crush Saga, a puzzle game that has been on top of the charts for almost 8 years and that gained \$1 billion in sales of digital goods alone in 2019²⁵, the game developers, King Digital

Entertainment, use AI to test their levels²⁶. They have a system called BAIT, or Bot for AI Testing, that navigates each level and tests all the clickable parts of the screen to ensure that if a user were to do the same task, say click on the 'next level button', the game would correctly perform that action. BAIT will also find any graphical rendering issues or missing text on the levels, making it ideal for testing functionality and finding any performance issues. Candy Crush has over 4500 levels and it releases around 45 new levels each week. Testing is key to ensure players enjoy a smooth gaming experience and that the game stays on top of the charts.



AI testing levels in Candy Crush, detecting elements and a graph generated after traversing all the levels in-game

Another application of AI feedback is in the language learning app, Duolingo, which prides itself in making learning fun by gamifying the experience and backing its design with data. The app uses machine learning algorithms to tailor the lessons based on the learner's knowledge of the language and by being the most downloaded language learning app in the world, they have access to incredible amounts of data generated by their users. The developers elaborate this on their research page: "With more than 500 million learners, Duolingo has the world's largest

²² Number of video gamers worldwide 2015-2023 <https://www.statista.com/statistics/748044/number-video-gamers-world/>

²³ Video Game Market Size, Share & Trends Analysis Report By Device <https://www.grandviewresearch.com/industry-analysis/video-game-market#:~:text=The%20global%20video%20game%20market,12.9%25%20from%202020%20to%202027.&text=The%20rising%20inclination%20from%20physical,on%20hardware%20compatibility%20and%20efficiency>

²⁴ How big data is disrupting the gaming industry <https://www.cio.com/article/3251172/how-big-data-is-disrupting-the-gaming-industry.html>

²⁵ Why are people still obsessed with candy crush saga? <https://www.wsj.com/articles/why-people-are-still-obsessed-with-candy-crush-saga-11603288963#:~:text=Yet%20while%20there%20have%20been,analytics%20firm%20Sensor%20Tower%20Inc>

²⁶ How King uses AI to test Candy Crush Saga <https://www.infoq.com/articles/candy-crush-QA-AI-saga/>

collection of language-learning data at its fingertips. This allows us to build unique systems, uncover new insights about the nature of language and learning, and apply existing theories at scales never before seen.”²⁷

On the screen itself, players have been interacting with AI in video games since the early stages of the industry, most commonly in the form of non-playable characters and enemies, or computer-controlled opponents. With more advanced AI in these elements, games can be more exciting with smarter enemies, can have more realistic dialogues, and games responding in real-time to a player’s actions.²⁸

This can be seen in *Zelda: Breath of the Wild*, an action-adventure game where you play as Link trying to save the fictional land of Hyrule by solving a series of puzzles and defeating enemies in an open-world design. The game introduces a consistent physics engine, letting players approach problems in different ways rather than trying to find a single solution²⁹. It also introduces an AI component to battles; players have enemy encounters where the opponents use all available resources to try to defeat the player. This seemingly thinking AI means that players will have to take caution when engaging in combat since their enemies are capable of coming up with survival solutions on the go³⁰.

A friendlier version of AI comes in the form of companions, who assist the player in navigation, puzzle-solving, and combat. They alleviate loneliness, enable banter and teasing, and facilitate character development³¹. An example is the game *The Last of Us*, an action-adventure game where you accompany a teenage girl, Ellie, across a post-apocalyptic United States. The AI behind her character makes her anticipate the player’s movement and have different sets of reactions based on what the player is doing, whether that is moving, hiding, or fighting. Besides helping the player navigate the world, the AI also makes Ellie interesting as a character, giving her special animations and audio dialogue³².

Like any other type of AI application, its implications have to be considered, with tailored game design a disease recognized by the World Health Organization (WHO) becomes more frequent, video game addiction. In recent years, there have been games accused of addictive design and even predatory business practices.

Such is the case for *Fortnite*, a free-to-play Battle Royal game developed by Epic Games. The company has been in the news cycle repeatedly due to the multiple lawsuits surrounding it. In 2020, it launched lawsuits against Apple and Google because it was removed from their online stores as a response to Epic lowering the cost of in-game currency. The game was also sued in the past because of the popular dances it featured, none of them attributing their original creators nor compensating them in any way³³. Parents of children who made unauthorized purchases in-game, and children addicted to the game, have also filed lawsuits against Epic since they have a no money back policy³⁴. Regarding the design of the game itself, *Fortnite* incentivizes players to play constantly by creating daily challenges and seasonal campaigns. It also has a hostile design

²⁷ Duolingo <https://www.duolingo.com/approach>

²⁸ How Artificial Intelligence Will Revolutionize the Way Video Games are Developed <https://insidebigdata.com/2020/11/27/how-artificial-intelligence-will-revolutionize-the-way-video-games-are-developed/>

²⁹ *Zelda Breath of the Wild* <https://web.archive.org/web/20171214212039/https://www.eurogamer.net/articles/2017-01-18-zelda-breath-of-the-wild-boss-eiji-aonuma-on-the-large-extra-burden-of-moving-to-nintendo-switch>

³⁰ Dynamic AI in *Breath of the Wild* <https://www.zeldadungeon.net/dynamic-ai-in-breath-of-the-wild/>

³¹ *The Last of Us part II* <https://www.theringer.com/2020/6/30/21307939/last-of-us-part-ii-ai-sidekicks-non-player-characters>

³² Endure and Survive the AI of *The Last of Us*

https://www.gamasutra.com/blogs/TommyThompson/20200617/364927/Endure_and_Survive_the_AI_of_The_Last_of_Us.php

³³ Epic Fortnite dance lawsuits <https://www.polygon.com/fortnite/2018/12/18/18147009/epic-fortnite-dance-lawsuits>

³⁴ Epic games faces class action lawsuit in Canada over Fortnite addiction <https://www.gamesindustry.biz/articles/2019-10-08-epic-games-faces-class-action-lawsuit-in-canada-over-fortnite-addiction>

in its item store that repeatedly (and aggressively) shows players what they are missing by not paying for the ‘gold tier’ items³⁵.

While AI and data are making video game design easier for developers, and the experience more immersive for players, it is important to regulate games with these practices to avoid feeding into the existing problem of video game addiction.

4. DATA IN THE OTT: ENHANCING ADVERTISEMENT AND MARKETING OPPORTUNITIES THROUGH DATA

Combining data, audiovisual content, branding, advertisement, and digital distribution means lead to a new era of marketing that is based on AI product placement. The endless possibilities in films, series, and videos on open platforms make this strategy much easier to identify opportunities through analyzes and video breakdown, develop strategies and monitor success. BEN³⁶, a Bill Gates-owned product placement agency has developed an AI that has 40 years’ worth of industry data, both structured—such as likes, comments, and views—and unstructured, including images, audio, and text within the content itself. “Using this data, we create customized algorithms for each client that are guaranteed to predict the best content opportunities for generating results,” says Erin Schmidt, BEN’s Chief of Traditional Integration. “Since 2015 we’ve increased the BEN’s revenue five-fold and established the company as an industry leader with a valuation of more than \$1 billion.”

In practice, this means that the AI can surface shows, films, influencers, streams, and social media platforms—perhaps even audiences—that a brand might not have previously considered otherwise. “Ultimately, this means that the AI will make stronger recommendations with each use, which elevates our capabilities and allows us to truly target the best channels for every individual brand and campaign. Persol is an example of a brand that has spiked after placing it in specific targeted celebrity audiences, even making a model that was discontinued to come back in stock due to its demand. The idea behind the data model is not placing products everywhere but rather knowing where to place and when and where not to.

Product placement opportunities could even exist in older movies by placing branded coffee mugs in hands with a blank canvas or even bus posters. It is a strategy from Chinese giant Tencent, owner of the WeChat platform³⁷. They have partnered with UK company Mirraid to place products in spaces that are empty in older movies. This means that every Christmas, you could see a different ad in the same place in Home Alone with Macaulay Culkin, or even in Mariah Carey’s video clip “All I want for Christmas is you”, a song that hits Billboard charts seasonally³⁸.

4.1 PROGRAMMATIC ADVERTISEMENT

Real-time interaction made targeted advertisement a very interesting feature but, at the same time, highly challenging due to technical challenges to define, process, and deliver it. The solution was to use “programmatic advertising”, a process fundamentally based on data.

Programmatic advertising is a technology-enabled process of bidding on where to sell a company’s advertising in real-time. Millions of operations happen at the same time, where the

³⁵ Folding Ideas: Manufacturing discontent <https://www.youtube.com/watch?v=dPHPNglhR0&t=709s>

³⁶ How product placement works in 2020 <https://www.forbes.com/sites/lalondon/2020/06/24/how-product-placement-works-in-2020with-ai-deep-learning-and-more/>

³⁷ Tencent uses AI to put product placement into old movies. <https://lovelymobile.news/tencent-uses-ai-to-put-product-placement-ads-into-old-movies/>

³⁸ All I want for Christmas is you <https://time.com/5708874/all-i-want-for-christmas-is-you-mariah-carey/>

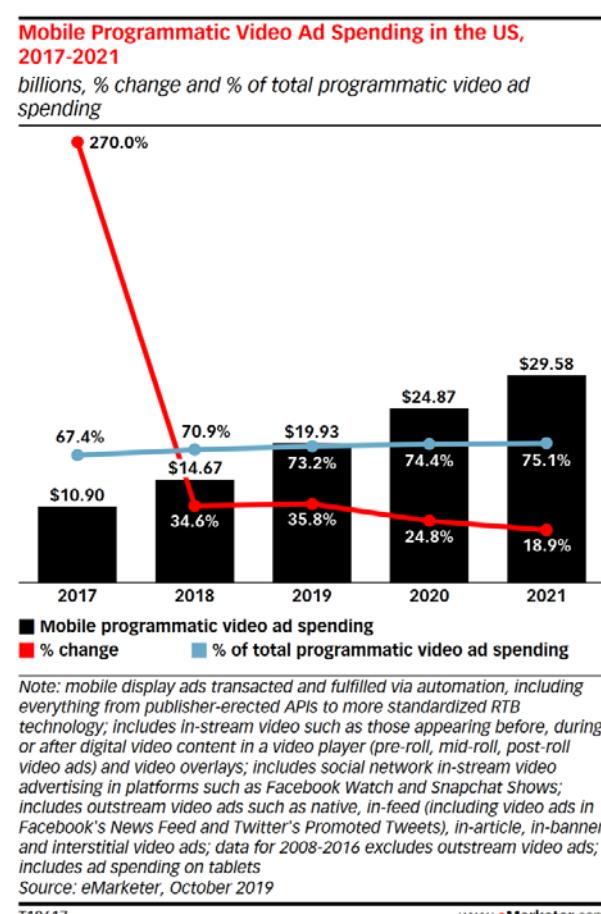
lowest and highest bidding values are set by the advertising company at any time of the day, similar to a stock price operation³⁹.

Programmatic is an informational tool that helps advertisers decide their campaign parameters, and whether they are interested in buying ad space or not. Real-Time Bidding (RTB) is one way of acquisition. Its technology allows buying and selling display impressions (the way the advertisement is valued) while a page loads. RTB could be understood as an online auction marketplace for purchasing “impressions.” The advertiser decides how much it is willing to pay, the targeted audience, and the message, but not the space in which it will be displayed.

One platform that has its business model based on programmatic advertising is Google, in particular through YouTube. In 2020, more than 80% of advertising was done programmatically. Overall, mobile video ads sold programmatically generated in the United States alone, \$19.93 billion in 2019 and \$24.87 billion in 2020⁴⁰.

"Even if the seller and the buyer know each other, they are using programmatic pipes for executing a transaction and for serving a campaign," said Kevin Schaum, vice president of the advanced solutions group at SpotX, an ad-serving and supply-side platform (SSP) for video publishers. "That shift has been one of the main things that we've seen."

The system offers many important benefits, such as:



- (1) Reaching a super-specific audience (Traditionally you had the option of buying ad space in a popular time slot on TV or a site your audience may frequent. Programmatic advertising uses behavioral data to find your audience and display the perfect video ad, making it much more effective);
- (2) Cross-channel reach (Because programmatic advertising relies on data, it allows us to reach across channels and serve targeted ads on mobile, desktop, and even TV sets);
- (3) Easy to readjust advertising campaigns (adjustments to the parameters could be done after the beginning of the campaign);
- (4) No need for real-life deal-making (Programmatic media buying lets the algorithms do all the work, with no need of building direct relationships with buyers and sales teams. All you need is to give the algorithms enough information to be able to make data-based decisions);
- (5) Quantitative reporting (All media ad space reports are based on data and communications

³⁹ What is programmatic video advertising <https://www.google.com/url?q=https://biteable.com/blog/what-is-programmatic-video-advertising&sa=D&source=editors&ust=1613087969856000&usg=AOvVawfEb2tzlYnsnM3wadMv9eQ>

⁴⁰ Programmatic advertising is driving rapid growth in mobile video <https://www.emarketer.com/content/programmatic-advertising-is-driving-rapid-growth-in-mobile-video>

between machines. Reporting is factual and data proves how successful the campaign was).

Programmatic Advertising is one of the technologies that has changed the most the advertising industry in the past years. It is certainly one of the key factors behind the successful business model behind AVOD and its growing attraction of “premium” audiovisual content⁴¹.

4.2 REDE GLOBO: DATA IN THE CREATION OF AN OTT BY A BRAZILIAN BROADCASTER

Rede Globo is not only the biggest broadcasting company in Latin America, but also one of the biggest exporters of soap opera, with hits like “Avenida Brasil” being seen in more than 106 countries. The Globo media conglomerate includes an array of companies⁴², from newspapers to videogames. The company has the biggest audience in Brazil, having reached an audience of 100 million per day⁴³, and was the only company that has remained its leadership during the COVID-19 pandemic. In 2020, streaming has become the second-biggest audience in the country⁴⁴, with 15% of TVs watching content through different OTT platforms instead of open or cable TV.

Following the market trend, Rede Globo created the company’s streaming service named Globoplay. Since 2015, the company has been heavily investing in the platform and its catalog as part of a digital strategy that began as early as 2010⁴⁵, which is five years before the main SVOD platform in Brazil, Netflix, arrived in the country.

Data was a crucial aspect of investing in streaming. Rede Globo started by integrating their data into what is called a *data lake*: an infrastructure that gathers all information that was once segregated or divided into different departments, into one place. Previously, most of the company’s data were kept in simple datasets and visualized in Microsoft Excels. The transformation allowed making data accessible and available through automated dashboards, enabling quicker assessment and decision making. This investment reduced time spent cleaning and organizing data, democratizing information, and allowing to rearrange human resources towards business insights.

Another benefit of shifting into a data lake was the possibility of cross-matching different audiences and analyzing a wider variety of users. Different content targeted different audiences and different age groups.

The company also invested in automation tools, such as targeting the automation of ads between commercials and recommended videos for the OTT platform through machine learning algorithms. These algorithms analyze comments from social media to place the most commented and liked videos on top of the user, crossing different sources and identifying possible interests of that particular user. After only 3 months of large-scale use of data, the company increased its audience by 35% compared to previous offers⁴⁶.

In 2019, the company announced a 3.3-million-dollar investment in Pixellot⁴⁷, an AI solution that uses unmanned multi-camera devices at arenas to track the flow of play, identify and clip highlights, update the scoreboard, as well manage ad inserts with zero human intervention. This

⁴¹ Programmatic. <https://www.thinkwithgoogle.com/marketing-strategies/programmatic/>

⁴² Grupo Globo. <https://brazil.mom-rsf.org/en/owners/companies/detail/company//grupo-globo-1/>

⁴³ Globo reaches 100 million. <https://redeglobo.globo.com/novidades/noticia/globo-celebra-alcance-de-mais-de-100-milhoes-de-pessoas-por-dia.ghtml>

⁴⁴ Streaming is the second biggest audience in the country. <https://olhardigital.com.br/2020/07/08/noticias/streaming-ja-e-a-2-maior-audiencia-do-pais-atras-apenas-da-globo/>

⁴⁵ Investment in Globoplay <https://teletime.com.br/23/04/2019/investimento-no-globoplay-e-prioridade-na-estrategia-da-globo/>

⁴⁶ Big Data project in Rede Globo <https://redeglobo.globo.com/novidades/noticia/projeto-em-big-data-da-globo-e-premiado-no-ibc.ghtml>

⁴⁷ Pixellot investment <https://www.sportspromedia.com/news/globo-pixellot-artificial-intelligence-streaming-technology-brazil>

big investment could certainly allow it to reap the benefits, especially for a company that holds the exclusivity of the main soccer events in Brazil.

With a long way to become the biggest streaming company in the country, Rede Globo has shown that making long term investments in digitalization and data is perhaps the only way to compete in the OTT market. Technology and data were an ally to the company during the COVID-19 pandemic and will continue to transform even the most traditional media companies in the years to come.

5 DATA AND AUDIOVISUAL: THE PERFECT MATCH

Product placement, programmatic advertising, branded content and many other advertising strategies have just started their way into the data world. Multiply their reach by new and diverse ways of accessing their audience (series, short videos, movies, documentaries, mockumentaries, podcasts, songs, and video games) and the different screen formats that allow the user to be connected anywhere, anytime (mobile, tablets, desktop, tv), as well as the services catered to people's interests (pay-per-view, streaming, social media), make infinite collaborations possible. AI is the growth tool that allows human creativity to take action on its own.

Neil Gaiman's book American God, written in 2001 and recently turned into a series by HBO tackles the war between the relationship of "old" and "new" gods. In one scene, the old god Argus Panoptes, a many-eyed giant from Greek mythology, meets the god Media⁴⁸, supposedly in decadence due to being too traditional for the new century. The encounter makes her absorb him and his all-seeing knowledge, transforming into a "New Media", and being able to watch over all her followers, having many more "worshippers", in the writer's analogy. She does so by using the tools given by new gods "Tech Boy" (technological screens) and "Mr. World" (world wide web).

Although farfetched and creative, to say the least, the writer depicts a story that is not far from the reality we live in. The combination of these "gods", old and new, make the audiovisual industry the perfect match for the birth of a New Media, one derived from data and in search of the perfect content for the perfect follower. This transformation is just the beginning of an era that will be determined by those that can properly analyze, collect and understand the data, to generate new formats, languages, and technology.

⁴⁸ American Gods [https://screenrant.com/american-gods-technical-boy-birth-death/2/](https://screenrant.com/american-gods-technical-boy-birth-death/)