# Big Data Applications in Media and Entertainment Industry

Jiaan Wang
Indiana University Bloomington
3209 E 10th St
Bloomington, Indiana 47408
jervwang@indiana.edu

## ABSTRACT

The growth of big data and its various applications in media and entertainment industry has been swift in recent years as well as the rapid surge of big data and the increasing need for big data technologies. We describe the problems that come with big data and its challenges in the industry. We then present various utilization of big data and why big data is important to the advancement of media and entertainment industry.

## KEYWORDS

i523, hid233, Big data, Media, Entertainment industry, Technology, Recommendation

## 1 INTRODUCTION

The amount of data being generated is increasing exponentially every year. Currently, we don't have the resources to process or analyze all the data. For example, giant tech companies like Google process over 20 petabytes of data daily [6]. "The rate at which we are generating data is rapidly outpacing our ability to analyze it and the trick here is to turn these massive data streams from a liability into a strength" [2]. Despite that, the technologies used to collect, analyze and interpret data are continuously improving [6].

IDC, the International Data Corporation, believes that "organizations that are best able to make real-time business decisions using big data streams will thrive, while those that are unable to embrace and make use of this shift will increasingly find themselves at a competitive disadvantage in the market and face potential failure", which will be especially accurate for companies who face high rates of changes in business [7].

But what exactly is big data? Wanda Group, a multinational conglomerate company based in China, defines big data as a DIKW hierarchical model, which stands for Data, Information, Knowledge and Wisdom [9]. Big data is about the rising challenge that companies face as they handle vast and rapidly-increasing sources of data and knowledge that further introduce a complicated field of inquiries and use issues. Big data technologies define a new era of technologies and frameworks which aim to efficiently gather useful information from huge array of data by using various data science techniques [7].

Emerging sources for big data include industries that are preparing to digitize their content. Particularly, "the media and entertainment industry moved to digital recording, production, and delivery in the past five years and is now collecting large amounts of rich content and user viewing behaviors" [7]. This will essentially prepare them to adapt for the upcoming big data era and make good use of these so called big data.

## 2 CHALLENGES IN MEDIA AND ENTERTAINMENT INDUSTRY

The issue with the immense data gathering and distributing structure we have created is: "big data is a big mess" [6]. All the information and data captured in our everyday lives simply have nowhere to be processed and it just gets put into storage forever [6].

Media and entertainment industry has always been embracing new technologies because companies believe that big data technologies are crucial to solving their business problems such as reducing operating expenses in an increasingly competitive market and generating enough revenue from producing data and content from various platforms and products [4].

Traditional TV media is facing challenges as its data is scattered. "It has internal data from set-top boxes, network management systems, BOSS systems, etc. as well as external data from online user behaviors. Data integration is the primary challenge in big data applications of traditional TV media" [9]. In China, the overall economy of traditional TV media does not look promising. The amount of time user spent on traditional TV has declined while more time is spent on Internet TV. "Studies have shown that, in 2012, Internet TV user base has reached 26.1 million while traditional TV user base is only 600 million. In addition, traditional TV operation rate has decreased from 70 percent in 2009 to 30 percent in 2012" [9].

These are the main challenges the media and entertainment industry need to deal with in order to better utilize big data to make a difference:

- "Making sense of data streams, whether text, image, video, sensors, and so on. Sophisticated products and services can be developed by extracting value from heterogeneous sources" [4].
- "Exploiting big data step changes in the ability to ingest and process raw data, so as to minimize risks in bringing new data-driven offerings to market" [4].
- "Curating quality information out of vast data streams, using algorithmic scalable approaches and blending them with human knowledge through curation platforms" [4].
- Businesses need to quickly get used to big data. Consumers nowadays are getting more familiar with the idea of big data. In addition, the cost of storage and analytics tools has been greatly reduced. Hence, it is extremely critical for businesses to understand the efficient use of big data to meet their needs and properly set up non-technical aspects such as management of personnel and staff in advance [4].

## 3 APPLICATIONS IN MEDIA AND ENTERTAINMENT INDUSTRY

Enormous amount of information is already being obtained about the entertainment industry [6]. "For example, *Supernatural*, an American horror series, created by Eric Kripke in 2005. Now in its seventh season, it has generated roughly 112 hours of footage. We have a lot of pixels and we also have every action of every character; every line of dialogue; a history of when, where, and how often everyone dies. Because all of that information is data, what we actually have, in and around those 112 hours of pixels, is a map to the world of *Supernatural*, and the characters inside it. Today, all of that footage and all of that information is locked away in old style data collections: fixed and unwieldy" [6]. However, suppose we take all these information and store it in a system. Later, we apply analytics tools using our powerful and growing processing power. Then it is possible that we can create and interact with the world and the characters of *Supernatural* [6].

Big data applications are also widely used in film industries as well. "IBM worked with a media company and ran its predictive models on social media buzz for the movie Ram Leela. According to the reports, IBM predicted a 73 percent success for the movie based on right selection of cities. Such rich analysis of social media data was conducted for Barfi and Ek Tha Tiger. All these movies had a runaway success at the box office. Shah Rukh Khan's Chennai Express, one of the biggest box office grosses in 2013, used big data and analytic solutions to drive social media and digital marketing campaigns. IT services company *Persistent Systems* helped Chennai Express team with the right strategic inputs. Chennai Express related tweets generated over 1 billion cumulative impressions and the total number of tweets across all hash tags was more than 750 thousand over the 90-day campaign period" [3]. Crayon, a big data analytic firm that is based in Singapore, teamed up with producers from leading Hindi film industries to select and release the right kind of music for movies so that they create the perfect hype. Furthermore, Lady Gaga and her associates also used optimization techniques to create the best impact at her live events by going through listening preferences [3].

Another field where big data applications are making influence is sports. Germany, FIFA 2014 champion, has been utilizing SAP's Match Insights software to analyze team performance which made a big difference for the team. It analyzes data such as *touch maps* of player positions, passing abilities, ball retention and so on. In addition, Match Insights was also used by an Indian Premier League team, Kolkata Knight Riders to test the consistency of its players which helped in both auction and ongoing training as well [3].

In order for entertainment companies to create better products and advertising strategies to appeal to more clients, they should figure out why their customers subscribe or unsubscribe using big data analytics [5]. "Unstructured sources best handled by big data apps like email, call detail records and social media sentiment reveal factors that are often overlooked for driving customer interest. Big data makes possible the understanding of consumption of digital media and entertainment and behavior that could be used together with traditional data demographic for personalized advertising in the right context at the right time, in the right place" [5].

One of the most impressive and powerful personalization tools created is recommendation engine. It gathers information from people's past records and suggests or predicts items that they might like. Companies like Amazon has profited by successfully combining its recommendation engine with its online shopping experience from browsing goods to checkout [1].

Artificial intelligence and deep learning technology are just two of the areas that Amazon is spending a vast amount of funds and resources in order to improve its recommendation engine to serve customers more effectively [1]. "In May 2016, Amazon opened up its sophisticated artificial intelligence technology as a cloud platform. The company unveiled DSSTNE, an open source artificial intelligence framework that Amazon developed to power its own product recommendation system" [1].

In addition, Amazon claims that customers frequently watch and review pilots created and released at Amazon Studios. Executives from Amazon then choose which pilots will develop into a full series based on customer feedback. *Transparent*, a comedy show based on a transgender patriarch whose family lives in Los Angles, was one result from this system. After its initial release in 2014, it received positive reviews from the general public and raised awareness about transgender problems [8]. "It was rewarded the following year with the Golden Globe for best TV series, musical or comedy" [8].

Another big media company who uses recommendation engines big time is Netflix who believes that content discovery is the number one priority. This is not surprising because its on-demand video and media streaming possibly dominates the world's market for digital content consumption. Just like Amazon, Netflix has spent immense amount of expenses and resources to make sure that its recommendation engine is one of the best to display its content library as much as possible [1]. "In December 2015, Netflix revamped the technology behind its content recommendation engine, deciding to do away with region based preferences in light of their ongoing global expansion" [1].

Netflix utilizes sophisticated statistical equations to advertise series that customers might like, for example, *House of Cards* and *Orange Is the New Black*. Those equations usually contains predictor variables such as popularity of the series, customers' previously watched content or their demographics [8].

"A typical Netflix user may lose interest unless something interesting is found within 60 seconds, two employees of the Los Gatos, California-based company wrote in a paper published in a scholarly journal last year" [8]. The personalization systems that Netflix has created for its customers has helped to lower the number of subscription cancellations, saving more than 1 billion dollar a year [8].

## 4 CONCLUSION

The rapid growth of big data has given media and entertainment industry an unique opportunity to utilize resources in order to benefit from big data applications and technologies. However, there are still some key challenges media companies are facing such as how to quickly adapt to the big data era, how to deal with and analyze immense amount of data pouring in every minute and how to make cost-effective products and consumer experiences.

Examples of current effective big data applications and technologies such as Match Insights from SAP and personalized recommendation engines from Amazon and Netflix are provided. In summary, big data applications and technologies are crucial in the success of media and entertainment companies.

## ACKNOWLEDGMENTS

## REFERENCES

[1] Shabana Arora. 2016. Recommendation Engines: How Amazon and Netflix Are Winning the Personalization Battle. Web Page. (June 2016). https://www.martechadvisor.com/articles/customer-experience/recommendation-engines-how-amazon-and-netflix-are-winning-the-personalization-battle/ HID: 233, Accessed: 2017-10-06.

[2] Lauren Browning. 2015. We sent men to the moon in 1969 on a tiny fraction of the data that's in the average laptop. Web Page. (June 2015). http://www.businessinsider.com/mind-blowing-growth-and-power-of-big-data-2015-6 HID: 233, Accessed: 2017-10-07.

[3] Ashok Karania. 2014. How Big Data Is Changing The Entertainment Industry! Web Page. (July 2014). https://www.linkedin.com/pulse/20140730194648-8949539-how-big-data-is-changing-the-entertainment-industry HID: 233, Accessed: 2017-10-03.

[4] Helen Lippell. 2016. *Big Data in the Media and Entertainment Sectors* (1 ed.). Springer International Publishing, Gewerbestrasse 11 CH-6330 Cham (ZG) Switzerland, Chapter 14, 245–259. https://doi.org/10.1007/978-3-319-21569-3_14 HID: 233, Accessed: 2017-10-03.

[5] Ritesh Mehta. 2017. Big Data in the Field of Entertainment. Web Page. (Aug. 2017). https://insidebigdata.com/2017/08/20/big-data-field-entertainment/ HID: 233, Accessed: 2017-10-03.

[6] Tawny Schlieski and Brian David Johnson. 2012. Entertainment in the Age of Big Data. *Proc. IEEE* 100, Special Centennial Issue (May 2012), 1404–1408. https://doi.org/10.1109/JPROC.2012.2189918 HID: 233, Accessed: 2017-09-20.

[7] Richard L. Villars, Carl W. Olofson, and Matthew Eastwood. 2011. Big data: What it is and why you should care. *White Paper, IDC* 14 (June 2011). www.tracemyflows.com/uploads/big_data/idc_amd_big_data_whitepaper.pdf HID: 233, Accessed: 2017-09-20.

[8] Angus Whitley. 2016. How Entertainment Companies Use Big Data. Web Page. (July 2016). https://www.comstocksmag.com/bloomberg/how-entertainment-companies-use-big-data HID: 233, Accessed: 2017-10-03.

[9] Chunjie Zhang, Wenqian Shang, Weiguo Lin, Yongan Li, and Rui Tan. 2017. Opportunities and challenges of TV media in the big data era. In *2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*. IEEE, Wuhan, China, 551–553. https://doi.org/10.1109/ICIS.2017.7960053 HID: 233, Accessed: 2017-09-20.

## 5 BIBTEX ISSUES

## 6 ISSUES

DONE:

Example of done item: Once you fix an item, change TODO to DONE

### 6.1 Formatting

in key words - replace Recommendation with Recommendation Systems

### 6.2 Writing Errors

Errors in title, e.g. capitalization - in the Media and Entertainment Industry

Are you using *a* and *the* properly?

### 6.3 Citation Issues and Plagiarism

It is your responsibility to make sure no plagiarism occurs. The instructions and resources were given in the class

Need to paraphrase long quotations (whole sentences or longer)