

Big Data Applications in Media and Entertainment Industry

Jiaan Wang

Indiana University Bloomington

3209 E 10th St

Bloomington, Indiana 47408

jervwang@indiana.edu

ABSTRACT

This paper demonstrates the growth of big data and its various applications in media and entertainment industry. We showcase the rapid surge of big data and the increasing need for big data technologies. We also 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 in the advancement of media and entertainment industry.

KEYWORDS

Big Data, Media, Entertainment Industry, Technology

1 INTRODUCTION

"2013 is the first year known as the beginning of big data, the world officially enter the era of big data. But big data is not clearly defined, until now, except for large enterprise data also have different definitions, such as Wanda defines the big data as DIKW hierarchical model, that is, Data, Knowledge and wisdom" [3].

"The era of big data is not coming; it is here. The birth and growth of big data was the defining characteristic of the 2000s. As obvious and ordinary as this might sound to us today, we are still unraveling the practical and inspirational potential of this new era. Google processes over 20 petabytes of data a day (a little less than half the entire written works of mankind from the beginning of recorded history in all languages). In addition to collecting and searching for more information, the technologies that allow us to capture and interpret that data are improving every time we blink. Something as simple as a snapshot has become a data collection event" [1].

"Big Data is about the growing challenge that organizations face as they deal with large and fast-growing sources of data or information that also present a complex range of analysis and use problems. Big Data technologies describe a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery, and/or analysis" [2].

"IDC, 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. This will be particularly true in industries experiencing high rates of business change and aggressive consolidation" [2].

"New data sources for Big Data include industries that just recently began to digitize their content. In virtually all of these cases, data growth rates in the past five years have been near infinite, since in most cases it started from zero. 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" [2].

"The problem with the massive data collection and distribution system we have created is: big data is a big mess. Most of the data we capture in our daily lives just sits around, cluttering up storage space on our devices and slowing down our connections" [1].

"Under the era of big data, the traditional TV media are facing opportunities and challenges, how to deal with challenges and to seize the opportunity is the traditional TV media should concern. Comparison to the Traditional TV media, network TV and new media, the biggest advantage is that the traditional TV media have high-quality TV content, and the strong support of national policy. Traditional TV media itself has a lot of data, but traditional media did not make good use of these data that has been the impact of new media" [3].

2 APPLICATIONS IN MEDIA AND ENTERTAINMENT INDUSTRY

"Social media solutions such as Facebook, Foursquare, and Twitter are the newest new data sources. A number of new businesses are now building Big Data environments, based on scale-out clusters using power-efficient multicore processors like the AMD Opteron 4000 and 6000 Series platforms, that leverage consumers' (conscious or unconscious) nearly continuous streams of data about themselves (e.g., likes, locations, opinions). Thanks to the "network effect" of successful sites, the total data generated can expand at an exponential rate. One company IDC spoke with collected and analyzed over 4 billion data points (Web site cut-and-paste operations) in its first year of operation and is approaching 20 billion data points less than a year later" [2].

"Some of the most interesting, but also most challenged, industries when it comes to Big Data adoption will be utilities and content service providers (e.g., cable TV, mobile carriers). These communities (with assists from related companies such as video gaming system and appliance manufacturers) are building out Big Data generating fabrics. Their opportunity now is to figure out how to handle and then do something with all that data, despite the fact that from a cultural standpoint data guardianship and use were much less in the past" [2].

"An additional hurdle for these industries is that it isn't enough to just get the "answers" from Big Data platforms. They also need to implement automated response systems (e.g., automated power management or "in game" ad placement) that will ultimately be the foundation of their business models" [2].

"We would like to offer a set of rules for the new data world: 1) big is not enough, and 2) it is neither necessary nor practical to fix every piece of data we have collected as a species into some particular

order” [1].

“We are already capturing massive quantities of data about our entertainment. Take, for example, *Supernatural*, an American horror series, created by Eric Kripke in 2005.¹ Now in its seventh season, it has generated roughly 112 hours of footage. So we have a lot of pixels, yes, but we also have much more. We 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” [1]. “Today, all of that footage and all of that information is locked away in old style data collections: fixed and unwieldy. But if we can store all that information in a system, modeled more on biology than books, and apply our significant and increasing processing power to analyze and respond to the world, rather than just move it around mechanically, then we have the possibility of generating and interacting with the world and the characters of *Supernatural* (or possibly even a story you like). This requires computational intelligence, not a Google search. It is not the ability to hunt down a single piece of data in the massive haystack of global information but rather the ability to make something new and interesting emerge out of that data” [1].

“In the era of big data, mass user behavior data is used to model predictions. Where big data are the personal recommendation system in a typical application of radio and television, The traditional approach is based on the user’s clicking behavior, to analyze the user’s preferences, then recommend related programs. But now in order to recommend more accurate, use not just the set-top box data for statistical analysis, but also dig out the sharing behavior on the user network along with the comment feature behavior and other behaviors, in order to better characterize user portrait. In the era of big data, television media should be the depth of excavation and analysis of user information on the user’s viewing behavior, the initiative to understand what users really want to see, in order to provide better services for television users. In other countries, the television media successful application of large data typical case is the “house of cards”, which analyzes the form of selection and decision-making with actors play using the big data” [3].

“Technically, the first to take in consideration is television media are capable of producing large amounts of data every day, how to integrate their data, define combining their data assets to create a connection between the television media and their users, effective analysis of audience preferences to realize customization. secondly the traditional TV media have with respect to network operators, the biggest advantage is that they have high-quality TV content, but how to use these high quality content effectively disseminated to users. in addition to drawing telecommunications powerful content communication technologies and outside network framework, also taking into account the characteristics of the television media itself” [3].

“The most important point that TV media can use big data technology is that television is the media itself has data, the data are the main source of set-top boxes, network management systems. To collect the data more widely, some companies such as Nielsen TV media can also take the technology to provide brain waves, using 32 sensors, acquisition frequency of 500 times / Sec, measurable

indicators are mainly about emotional investment, triggering memories and attention. Therefore, data collection is more mature as it showed” [3].

“But the TV media data from multiple data sources and scattered, besides the internal data such as set-top box data, network management systems data, BOSS system, etc., as well as external data, such as online user behavior data, data integration is the primary challenge in television media big data applications. How TV media make internal data and external data streams to achieve mutual exchange, how to create their own big data, sort out their own data assets, which need the support by big data technology. And television media use Big Data technologies to meet the individual needs of the “precise communication”, which can improve service quality, protection of cultural rights and interests of the public, the media TV plays disseminating information, building culture, guide public opinion, the responsibility to resist foreign cultural erosion at the same time, therefore more need to focus on high-tech applications” [3].

3 CONCLUSION

Put here an conclusion. Conclusions and abstracts must not have any citations in the section.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Gregor von Laszewski for his support and suggestions to write this paper as well as Lee Yang for her proofreading on this paper.

REFERENCES

- [1] 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.
- [2] 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.
- [3] 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.