Communication and Media Education in an Era of Big Data

Rama Venkatasawmy
University of Newcastle, Australia

ABSTRACT

This theoretical paper argues for the need to review communication and media education curricula in the light of how roles and jobs are undergoing transformation in the communication and media industries wherein work carried out daily is influenced by Big Data analytics and derivations. Educators have to ensure that communication and media students comprehensively understand how and why large-scale data-sets are collected, scrutinized and interpreted for meaning and value to be ultimately applied when generating content. Big Data analytics and derivations need to be more present in contemporary communication and media programs considering how they are impacting on ways of doing and of knowing as well as on the negotiation of value within the communication and media industries.

This theoretical paper proposes an argument for the necessity to review communication and media education curricula in the light of how roles and jobs are undergoing transformation at all levels in the communication and media industries. Big Data are rapidly re-inventing how content is produced in the communication and media industries, as illustrated for instance by the extent to which Big Data-centric phenomena are impacting on elementary aspects of work in television, journalism, advertising and public relations (that are specifically addressed here).

It seems reasonable to view the gathering and use of big data in the media sector as the latest step in the historical process of the 'rationalization of audience understanding', that dates back at least to the 1930s when communication researchers such as George Gallup and Paul Lazarsfeld began working with media industry stakeholders on more data-driven approaches to understanding and predicting audience tastes and preferences (Napoli, 2016, p.2).

Continuously expanding Big Data as well as the computing methods associated with their collection and processing bear explicit resonance to how the generation of content by the contemporary communication and media industries reflects major current social and professional changes as influenced by technological advances. And this resonance needs to be taken into account in a better way when communication and media programs are taught in higher education contexts.

It becomes indispensable for various aspects pertaining to Big Data analytics and insights to be more effectively integrated into how communication and media programs are taught, so as to nurture in future communication and media professionals the ability to understand, evaluate and apply Big Data insights when investing resources, making creative decisions and producing content. Communication and media students ought to learn how large quantities of information – generated about and by netizens, social media users, online service providers, subscribers, shoppers, and retailers – constitute Big Data which are then categorized, dissected, processed and analyzed. As a result, significant extrapolations, indications and prognostications are derived and subsequently shape funding and the production of content in the communication and media industries.

Many of the ideas and points raised in this paper to develop the proposed argument come from various informal discussions and conversations with senior practitioners in the contemporary communication and media industries. It is acknowledged that, without the application of legitimate scientific methodology, such informal material can hardly be considered as serious validation of the argument proposed here. But the author's understanding of the contemporary state of the communication and media industries has undoubtedly benefited from the above-mentioned informal professional insights – some of which have been incorporated in the discussion – into contemporary industrial practices.

What is 'Big Data'?

In 2017, messaging applications – such as WhatsApp, Facebook Messenger, WeChat, QQ, Viber, Line, Telegram and Kakaotalk – had more than five billion active users monthly (HubSpot Research, 2017). It was not uncommon for the average individual anywhere to have five social media accounts in 2014 (Bennett, 2014) and to be engaging with social media for at least two hours every day by 2017 (Asano, 2017). And in many parts of the world, most people now obtain daily news almost exclusively from social

media platforms like Facebook, YouTube and Twitter (Wagner, 2017). Massive trails of data – that are exchanged, merged and stored (through the use of Apache Hadoop, Microsoft HDInsight, NoSQL, Hive, Sqoop or PolyBase for instance) – are generated by the intense interaction of people with social media networks as well as by their constant Internet activity.

The escalating coupling between social media network sites – for example to maintain a list of favorite TV shows and songs or to share photos – enables large quantities of data generated on the Internet to be exchanged and merged. The pervasive diffusion of apps, social media networks and service subscription sites, accessible 'anywhere-anytime' from increasingly 'smart' mobile appliances, has indeed led to vast amounts of personal data to become available, from: every search and social media reference made, every photograph uploaded or downloaded, every news article read, every positive or negative rating given to any form of creative content, every geographical location, every log in, subscription, payment, song list, textual entry, bookmark, and so on.

As people browse, search, communicate, share and buy throughout any day in a digitized world, each of the devices they use are essentially talking to each other and ultimately creating enormous amounts of structured, semi-structured and unstructured data. According to Richterich (2018, p.7), while

> structured data (such as demographic information or usage frequencies) can be easily standardized and, for example, numerically or alphabetically defined according to a respective data model, unstructured and semi-structured data are more difficult to classify. Unstructured data refer to visual material such as photos or videos, as well as to text documents which are/were too complex to systematically translate into structured data. Semi-structured data refer to those types of material which combine visual or textual material with metadata that serve as annotated, structured classifiers of the unstructured content.

In 2012, there was a thousand exabytes of data around the planet according to IBM Research's Michael Karasick: that figure was then projected to grow to approximately 9,000 to 10,000 exabytes of data by 2015 – attributable predominantly to the generation of data by mobile devices and social media networks, and to the growth of enterprise data (King, 2012).

Xu, Frankwick & Ramirez (2016, p.1562) explicate how 'Big Data' has been used as "a term that primarily describes datasets that are so large (terabytes to exabytes), unstructured, and complex (from genome-analysis, political science, sensor, social media, or smart phone apps, to Internet-based gadgets data) that require advanced and unique technologies to store, manage, analyze, and visualize." In practical terms, Big Data fundamentally constitute "the overwhelming volume and variety of digital information produced by and about human activity" (Lewis, 2015, p.322) that can be mined for trends, patterns and associations pertaining to the behavior of people online. Big Data

can hence be defined as "high-*volume*, high-*velocity* and/or high-*variety* information assets that demand cost-effective, innovative forms of information processing for enhancing insight, decision-making, and process automation" (Gartner, Big Data, 2012-2018). The most prominent features that distinguish Big Data from traditional large-scale datasets include:

- volume – the scale or size of data;
- velocity – the speed and flow of data as well as their processing; this has to do with the "timeliness of Big Data, specifically, data collection and analysis must be rapidly and timely conducted, so as to maximize the use of the commercial value of Big Data" (Rodríguez-Mazahua *et al.*, 2016, p.3074);
- veracity – the degree of uncertainty or unreliability of data; this has to do with "the need to check the accuracy of the data by eliminating noise through methodologies such as data pedigree and sanitization. This is to ensure data quality so that decisions that are made from the collected data are accurate and effective" (Rodríguez-Mazahua *et al.*, 2016, p.3075);
- value – the worth of data based on how they can be potentially exploited for commercial gain; data can constitute "a 'commodity' that can be sold to third parties for revenue" and "understanding the cost or value of the data can aid in budget decision-making and estimating the storage cost of the data" (Rodríguez-Mazahua *et al.*, 2016, p.3075);
- variety – the different types, forms and sources of structured, unstructured and semi-structured data.

And "variety is becoming the single biggest driver of big-data investments, as seen in the results of a recent survey by New Vantage Partners. This trend will continue to grow as firms seek to integrate more sources and focus on the 'long tail' of big data," according to Tableau Software (2017).

Information is filtered, selected and evaluated by means of statistical methods associated with data mining, language processing, and various Big Data tools. The latter are categorized according to the kind of analysis being carried out, such as: "(1) batch analysis, where data are firstly stored and then analyzed; (2) stream analysis, which is the extraction of information and knowledge discovery from continuous, rapidly generated streams of data, and (3) interactive analysis, which process the data in an interactive environment, allowing users to undertake their own analysis of information" (Rodríguez-Mazahua *et al.*, 2016, p.3078). Ultimately, Big Data are not so much about data quantity – as articulated in terms of Gigabytes, Terabytes, or Petabytes – but more about the insights gained from analyzing the data in question by employing particular techniques and technologies. Edwards & Fenwick (2016, p.216) confirm how

> digital analytics have become a key means of performing knowledge through tagging, classification, standardization, calculation, circulation and visualization, and these require codes and algorithms. […]. In effect, digital analytics and the standardization they require

to function are integral to much of our knowledge, communication and decision-making, and are part of the enactment of new ways of working and governing work.

Big Data thus help corporate and public organizations to extensively comprehend online behavior more than most people actually realize: not just what they are doing, listening to or watching or but also where, when and with what technological device. *Tweeter* likes, *Facebook* pokes, *Google* searches or *YahooNews* preferences, for example, enable service providers and advertisers to collect a lot of raw information from appliances connected to the Internet and from smart phones about people's habits and preferences. Major advancements with regards to data creation, access and storage at constantly smaller costs have made it easier to navigate and analyze the shadow layer, in the form of Big Data, of public lives online. Social media analytics, for instance, have emerged to essentially evaluate "informatics tools and frameworks to collect, monitor, analyze, summarize, and visualize social media data to facilitate conversations and interactions to extract useful patterns and intelligence" (Fan & Gordon, 2014, p.74).

As a consequence of processing colossal quantities of eclectic data generated by always-on invisible sensors and tracking tools implanted in the technological appliances and mobile devices that people interact with daily, a surfeit of algorithms now generate all sorts of predictions and models of people's preferences. Those unseen algorithms subsequently influence what shows people watch, what news they read, or even who they might vote for (considering how topics that end up trending on social media are mostly determined by algorithms). About such a state-of-affairs, *The Guardian* (2018) has commented that "the conceit of data mining firms – and the politicians who use them – is that they could win elections by molding electorates based on new forms of identity and new value systems. This process is accelerated by the echo chamber of social media, which allows citizens to close themselves off from wider debate and become infatuated with their own truths." Through increasingly complex channels of communication, artificial intelligence, data-crunching algorithms and prediction engines are managing how people view their own society and the world beyond.

From the consumer's perspective, the prevalence of social media transforms how people obtain information, connect with others, endorse their favorite brands, and purchase products. Social data analysis grows out of these activities and combines disciplines such as social network analysis, multimedia management, social media analytics, trend discovery, and opinion mining (Xu, Frankwick & Ramirez, 2016, p.1563).

The gathering, processing and scrutiny of Big Data impact tremendously on how the contemporary communication and media industries operate: by influencing decisions about investment and funding, and by re-shaping the ways in which communication and media producers not only create content but also engage with and manage business-critical audiences. "From both the production and the consumption standpoint, big data-fuelled algorithms are

increasingly dictating how media consumers navigate their media environment, while also increasingly dictating content production decisions" (Napoli, 2016, p.3). Resorting to Big Data insights and inferences ultimately permits a more precise customization and personalization of content, product and services that more effectively fulfill the specific needs of each type of consumer and audience.

Big Data and Television

  An ever-expanding audience of viewers who want to watch whatever they want wherever and whenever has become a fast growing trend in the light of the exponential rate of smart phone penetration on all continents. To fulfil the needs and expectations of billions of viewers/consumers globally, the online streaming of original TV shows and the delivery of Video-On-Demand services have become some of the most lucrative markets on the Internet. This accentuates how the audience remains a dominant fixation in the hyper-competitive communication and media industries. All business is fundamentally risky, but business in the communication and media industries can be "particularly risky because they are centered on the production of texts intended to be bought and sold to audiences that subsequently use these texts in highly volatile and unpredictable ways" (Redvall, 2016, p.139). Because of high sunk expenses that frequently arise in the elaboration and production of communication and media content (such as for a new TV commercial or serial) and because of the absence of absolute certitude about audience interest in or demand for any particular content, in the communication and media industries "the need to know and to anticipate audience tastes becomes crucial for predicting successful returns on investment" (Havens, 2014).

  Whether people are doing a search, playing a game or viewing a show, a record is made of how many people begin or stop an activity online – like each time people are pausing, rewinding and fast-forwarding when they watch audio-visual material, for example. Hence, thousands of viewers pausing while they watch a particular TV show online would generate ample data for logarithms to crunch and for inferences to be made about viewers' levels of interest at particular points of that show. The latter inferences can then subsequently be taken into account by the TV show's producer and by the TV network when preparing new episodes or making future programming choices in order to more effectively fulfill viewers' interests and expectations. With increasingly more people watching content streamed online than on physical Blu-ray Discs and DVDs, "big bets are now being informed by Big Data, and no one knows more about audiences than Netflix. A third of the downloads on the Internet during peak periods on any given day are devoted to movies streamed by Netflix" (Carr, 2013). Hence, Netflix's proprietary algorithms apparently scrutinized trillions of Netflix data points which permitted "a better understanding and targeting of viewers' tastes and of how effectively the company filtered its glut of subscriber data to create and distribute" (Havens, 2014) exclusively online its popular award-winning

series *House of Cards* (2013-), as well as validated Netflix's initial investment of US$100 million for the show to go into production. Carr (2013) further clarifies how TV shows and movies on Netflix are:

> annotated with hundreds of tags – metadata descriptors – inserted by viewers commissioned to describe the talent, the action, the tone and the genre, among many other things. In the past, those tags were used to recommend other shows from the long tail of content on the service, essentially building profiles based on the preferences of individual subscribers. But now Netflix is commissioning original content because it knows what people want before they do.

And, like with the making of *House of Cards*, Netflix apparently implemented a similar approach when creating another one of its many successful TV shows: *Orange Is The New Black* (2013-).

Big Data and Journalism

As "the traditional dual product marketplace of selling content to audiences and audience attention to advertisers continues to break down, the notion of an alternative dual-product marketplace in which user data replaces audience attention represents another direction" (Napoli, 2016, p.4) that the news industry needs to follow. News writing, as the most basic form of journalism work, has already started evolving in a major way with the involvement of non-human content producers in generating news stories. News content is now being created according to what has been labeled 'automated journalism'. The term refers to "algorithmic processes that convert data into narrative news texts with limited to no human intervention beyond the initial programming" (Carlson, 2015). Some consider data-oriented practices like 'automated journalism' to be highly disruptive with regards to 'traditional' compositional forms of news, journalistic labor and authority – as a result of which have emerged "fundamental tensions not only about the work practices of human journalists but also about what a future of automated journalism may portend for larger understandings of what journalism is and how it ought to operate" (Lewis, 2015, p.325). But for others however, like *The New York Times* creative director Alexis Lloyd, "the future of computational journalism and automation will – and should – be a collaborative one, where you have machines and people working together in a conversational way" (quoted in Lecompte, 2015).

Journalists working for such news providers as *ProPublica*, *Forbes*, *The New York Times*, *Oregon Public Broadcasting* and *YahooNews* are now recurrently using algorithms in the process of telling stories about business, sports, education, public safety, and so on (Lecompte, 2015). Automation tools

> support news organizations in their push to develop new storytelling formats that highlight the relationships between news events and help provide readers with richer context. Most of these efforts require large amounts of detailed metadata that can help link together stories that have in common people, places, or ideas. Adding metadata is a

> frustrating task for most reporters, who are typically more concerned with crafting their story than dissecting it. Automation is a way to expand the use of metadata – without putting an extra burden on reporters and editors (Lecompte, 2015).

Automation tools effectively support news gathering by tracking social media datasets and enabling journalists to become aware of specific issues, to pay more attention to particular conversations online, and to subsequently react and produce news content faster.

Big Data, Advertising and Public Relations

It is now almost impossible for contemporary advertising and public relations practitioners to operate successfully without engaging with, making sense of and leveraging Big Data. Every project undertaken essentially involves collecting, analyzing and interpreting data culled from various primary and secondary sources. As the practice of advertising and public relations strives to not only provide but also demonstrate 'value-addedness', all proposed strategies and content have to be based on data and their accurate (and ethical) interpretation. Social media channels for instance cannot be effectively managed simply by means of posting creative content alone: social listening and monitoring in real-time are also required to identify consumer concerns.

The practice of advertising and public relations has been shifting from generalist to specialist realms with a place for different skills set, but each in their individual category must outperform to stay connected and relevant. Practitioners hence increasingly have to address audience segments with specific interest points in order to be captivating, to connect immediately, and to hold audience attention. As such, being able to read the audience to know what each particular segment cares most about becomes a vital skill. The vast amounts of data generated by social media platforms give the opportunity to advertising and public relations agencies to listen in to the many thousands of things being voiced online about their brands and clients. Big Data analytics and insights help to find out a number of things about online consumers: who they are, how they feel, what they want, where they get their information from, and how they choose to buy anything.

In advertising and public relations, Big Data get generated as a result of reputation measurement, brand image measurement, the monitoring of traditional and social media, and the monitoring of issues in real-time. Global communications firm Edelman, for example, uses a combination of monitoring and analysis technologies that include Radian6, Brandtology and iSentia as well as its own proprietary technologies like TweetLevel and BlogLevel to operate in real-time. Global public relations firm Weber Shandwick for its part launched in Asia-Pacific its own Mediaco monitoring tool which relies heavily on performance analytics with regards to how different audiences are consuming content across multiple channels. According to the head of digital at Weber Shandwick Asia-Pacific Jonathan

Wade, "the real benefits of real-time data are in maximizing engagement and results with media and therefore ensuring budgets and resources are optimized for the best return on investment. This also means our audiences get the content they want, when they want it" (quoted in Benjamin, 2014).

Real-time monitoring tools hence enable stories to be tracked not only as they hit mainstream media but also as they start to develop even earlier through social media channels. To meet the needs of the client in today's digital world, advertising and public relations agencies are hence making serious investments in real-time planning, monitoring and their related technologies. By enabling a deeper understanding of customers, consumers and users, Big Data analytics and insights assist advertising and public relations agencies in choosing the appropriate social media platforms to spread information as well as to fine-tune social media messages.

The Need to Sync Communication and Media Education with Professional Practice

In different informal discussions and conversations between the author and senior practitioners in the contemporary communication and media industries, a commonly professed opinion is that many fresh communication and media graduates display in entry-level jobs limited preparedness with regards to engaging with Big Data as well as to utilizing various planning and monitoring technologies. This is perceived to be the result of a gap in the overall skill sets being imparted to students who, for their part, would often claim that the study programs they were enrolled in did not equip them with such skills. Neill & Schauster (2015, p.12) quote a U.S-based digital consultant as saying that students are ill-prepared in the area of tools and technologies used in contemporary public relations and advertising campaign strategies that now include programmatic buying, content amplification, social listening, monitoring and insights: "when they land at an agency, […] there's at least a one-year learning curve to just look inside mostly vast software, service-based tools and platforms that are now driving the business today." In order for Big Data to be used in the most relevant and effective way, the right interpretation logic needs to be deployed through a combination of algorithmic and human analysis. Whether denoted as 'Big Data' or 'algorithmic culture', "the integration of digital recording, distribution, and data analysis technologies" (Havens, 2014) in the operations of the communication and media industries is an actuality that contemporary students simply have to be familiarized with.

There are so many instances of algorithms, automation and artificial intelligence being applied on a daily basis to productively support work and to facilitate content adaptation and customization for different types of audiences and consumers in the media industries. As a result, "the ability to effectively analyze large quantities of data is spilling over into a wide range of media industry career paths that previously did not require this particular skill set" (Napoli, 2016, p.3). But learning about Big Data usage, analytics, insights and related applications is however still a limited aspect of many

contemporary communication and media programs curricula which may appear to be out-of-sync with how Big Data analytics and insights are at the center stage of contemporary processes of content production in the communication and media industries. Many communication and media programs hardly offer opportunities for students to develop the kind of analytical skills and technical competencies relevant to engage with Big Data productively and to acquire a better understanding of highly complex channels of communication and of the behaviors of globally dispersed highly interactive online audiences.

The relative 'new-ness' of Big Data inevitably means a short history about how the right skills and relevant competencies should be taught. Duncan, Caywood & Newsom (1993) have however convincingly argued that integrated and flexible curricula can help higher education programs keep up with trends in the advertising and public relations industry, for instance, while enabling the adequate professional and personal development of students. And there is a legitimate case to be made about implementing a shift in the teaching of communication and media programs: from a lecture-based format to a modular learning-oriented approach that is predominantly inquiry-based and that considers students as scientists who "develop hypotheses, design and conduct experiments, collect and interpret data, and write about their results" (Handelsman *et al.*, 2004, p.521).

Educators and educational institutions need to reflect on and review the kind of attributes and competencies they are fostering in future professionals who will operate at various levels of the communication and media industries. Curricula for communication and media programs require some re-orientation to enable the nurturing of skills that allow students during their studies to access, analyze, interact with and apply Big Data to generate content in similar ways as in the contemporary communication and media industries. In being trained to deal with a fast-evolving computer-mediated world, communication and media students need to be taught to manage and make sense of large volumes of data. Such activity requires some knowledge of statistics and the ability to extract insights, in combination with a solid knowledge of the 'traditional' fundamentals of the communication and media industries (such as how public opinion building works, how to use techniques of persuasion, media effects, and so on).

Communication and media students need to develop adequate numerical competency in order to understand, analyze, interpret, and ultimately leverage the kind of information they have access to, in addition to the core competency of the written word traditionally expected of those engaging in processes of communication and media content production. This can be achieved by including math-intensive courses that involve statistical calculation and analysis as core components in the curricula of communication and media programs. And if those programs enable basic skills to be developed in prescriptive analytics, diagnostic analytics, predictive analytics and outcome analytics, communication and media

students would wholly comprehend not just how and why "large-scale datasets and their collection, analysis, and interpretation are becoming increasingly salient for making sense of and deriving value from digital information" (Lewis, 2015, p.321), but also in what manner such information can be utilized in contemporary practices for producing communication and media content. Communication and media program curricula that continue not to cultivate skills in mathematics, statistics and data analysis are not likely to produce the kind of graduate attributes that are in sync with Big Data-oriented work carried out in the contemporary communication and media industries.

The proliferation of Big Data has influenced the emergence of many new roles in the communication and media industries that depend on math-centric skills for the purpose of data management and analysis processes relating to different types of audiences, such as social media analysis, social media listening and programmatic buying. "The simultaneous evolution of big data applications and new audience formations via the Internet signifies the extent to which we may find ourselves increasingly studying data about audiences, instead of the audiences themselves" (Athique, 2018, p.71). As new media tactics and production practices emerge in response to new audiences, technologies, tools, platforms and channels, and as social media analytics and various forms of Big Data become even more influential on creative decisions, skills in mathematics and statistics are now additionally needed by communication and media students. They will be expected to know how to gather, make sense of, assess, and apply different types of numerical data as part of doing work in the communication and media industries.

Writing and presentation skills are still considered essential for students to develop and for educators to nurture with regards to the practice of advertising and public relations. But advertising and public relations professionals are increasingly expected to additionally have adequate skills in mathematics and statistics – especially because of the prominence of data and analytics across a multiplicity of media platforms and channels – which would enable them to evaluate trust scores for active social media subscribers (by using social network data analytics for example). Communication and media programs should familiarize students with such content management and social listening tools like Sysomos, Hootsuite and Radian6, with content amplification tools like OutBrain, and with media planning and buying tools like comScore, DART, Trade Desk, Rubicon, Pubmatic, ThinkVine and AgilOne. This should allow future advertising and public relations practitioners to comprehend how software is used to automate the purchase, placement, and optimization of media inventory via a bidding system: student projects should hence involve engagement with programmatic media buying, marketing and advertising tools like DoubleClick Campaign Manager, Admedo and Paid Search. They should also learn the principles of Boolean search, by being made to use Sysomos software for instance, in the process of learning how to look for information that is relevant to their projects. And they also need to be made to practice text mining from Twitter data for the

purpose of making insights about brand associations. Future reporters and journalists should be taught how automation technologies are being incorporated by editors into the work flow of the newsroom so that they can concentrate on the fundamental aspects of being reporters and have extra time to work out the best way to tell a story. This can be done by incorporating particular automation technologies into the practical assignments of journalism courses, such as BBC's Juicer to streamline media work flows, Reuter's News Tracer to track down breaking news, The Washington Post's Knowledge Map to dig out media insights and Narrative Science's Quill platform to turn raw data into intelligent stories (Underwood, 2018). And with regards to television production, future television scriptwriters should be made aware of how algorithms now analyze trillions of data points to enable a better perception and more precise targeting of the taste of viewers, and they should be given projects that would require them to create scripts and audio-visual content based on the analysis of data about specific types of audiences.

Conclusion

A key transformation currently happening and impacting on workers in the communication and media industries is the transition from an era of data paucity to an era of data overabundance about the audience whereby

> different individuals and organizations are using data to try and gain acquiescence and advantage over other players. Many believe that the massive amount of data on viewership that Netflix collects, for instance, including more than fifty data points about the consumption of every film and television show, gives Netflix substantial capacity to predict what viewers want, far more than traditional Nielsen ratings ever gave broadcasters (Carr, 2013).

The speedy propagation of digital information technologies and the extensive reliance on Big Data to generate content have overwhelmed workers in the communication and media industries with immense quantities of information about audiences. But Big Data also allow communication and media practitioners to be able to display almost immediately the impact and worth of their work to senior management: following the release and distribution of content, various kinds of metrics become rapidly available to enable the evaluation in real-time of how it is being viewed/consumed.

While the impact of Big Data on professional practice is clearly recognized, as illustrated by how the ways of working are being transformed in the contemporary communication and media industries, Henke *et al.* (2016) pertinently comment however that adapting to "an era of data-driven decision-making is not always a simple proposition. Some companies have invested heavily in technology but have not yet changed their organizations so they can make the most of these investments. Many are struggling to develop the talent, business processes, and organizational muscle to capture real value from analytics." It should also be noted that as a research discipline, Big Data

and Big Data analysis "are still evolving and not yet established, thus, a comprehensible understanding of the phenomenon, its definition and classification is yet to be fully established" (Sivarajah *et al.*, 2017, p.264). While Big Data might still be considered an emerging phenomenon, its impact on professional practice in the communication and media industries has made it nevertheless very relevant to current communication and media-related academic research.

Because of how Big Data are already seriously impacting on ways of doing and of knowing, as well as on the negotiation of value within the communication and media industries, analytical and strategic thinking skill sets have become a premium that really need more serious nurturing. Big Data hence need to be more present in the curricula of contemporary communication and media programs so as to better prepare students in effectively using Big Data analytics and insights when developing and producing content.

This paper has argued for the need to review communication and media education curricula in the light of the: perceived contemporary state of communication and media education in relation to Big Data usage; questionable preparedness of fresh university graduates to deal with Big Data in entry-level jobs; predilection of the contemporary communication and media industries for recruiting workers who are able to analyze, interpret and use big data in an intelligent and effective way. Teaching, the role of communication and media educators and the discipline's overall purpose should hence be re-oriented towards better equipping communication and media students with updated competencies and skill sets that would make them a lot more industry-ready in professional worlds already extensively driven by Big Data usage, analytics and insights.

To follow up on the ideas and points brought up here, the intended subsequent phase will be to work towards substantiating and validating the argument proposed in this theoretical paper by means of qualitative data collection and analysis – which will involve in-depth personal interviews in particular. As explicated by Davis (2012), such interviews are adequate and relevant when motivation, attitudes and behaviors need to scrutinized, while Hatzios & Lariscy (2008) indicate that they can also produce a much deeper level of detail. The scope of data to be collected and analyzed will ultimately depend on budgetary considerations as well as on the availability of the communication and media professionals concerned.

References

Asano, E. (2017, January 4). How Much Time Do People Spend On Social Media?
Social Media Today. Retrieved from
https://www.socialmediatoday.com/marketing/how-much-time-do-people-spend-social-media-infographic.

Athique, A. (2018). The Dynamics and Potentials of Big Data for Audience
Research. *Media, Culture & Society*, 40 (1), 59-74.

Benjamin, K. (2014, January 15). Big Data and Planning: Real-Time PR Needs
Talent as well as Tools. *PR Week*. Retrieved from
http://www.prweek.com/article/1290000/big-data-planning-real-time-pr-needs-talent-tools.

Bennett, S. (2014, October 24). The Average Internet User Has 5 Social Media
Accounts. *Adweek*. Retrieved from
http://www.adweek.com/digital/social-media-accounts/.

Carr, D. (2013, February 24). Giving Viewers What They Want. *The New York
Times*. Retrieved from http://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html.

Carlson, M. (2015). The Robotic Reporter – Automated Journalism and the
Redefinition of Labor, Compositional Forms, and Journalistic Authority. *Digital Journalism*, 3 (3), 416-31.

Davis, J. (2012). *Advertising Research: Theory and Practice (2nd ed.)*. Upper Saddle
River: Prentice Hall.

Duncan, C.R., Caywood, C.L. & Newsome, D. (1993). *Preparing Advertising and
Public Relations Students for the Communications Industry in the 21st Century: a Report of the Task Force on Integrated Communications*. Evanston: Department of Integrated Marketing Communication, Northwestern University.

Edwards, R. & Fenwick, T. (2016). Digital Analytics in Professional Work and
Learning. *Studies in Continuing Education*, 38 (2), 213-27.

Fan, W., & Gordon, M. (2014). The Power of Social Media Analytics.
Communications of the ACM, 57 (6), 74-81.

Gartner Inc – Big Data. (2012-2018). In *Gartner Inc*. Retrieved from
https://www.gartner.com/it-glossary/big-data.

Handelsman, J., Ebert-May, D., Beichner, R., & Bruns, P. (2004). Scientific Teaching. *Science*,
304, 500-25.

Hatzios, A. & Lariscy, R.W. (2008). Perceptions of Utility and Importance of
International Public Relations Education Among Educators and Practitioners. *Journalism & Mass Communication Educator*, 63 (3), 241-58.

Havens, T. (2014). Media Programming in an Era of Big Data. *Media Industries*,
1 (2). Retrieved from
http://www.mediaindustriesjournal.org/index.php/mij/article/view/43/82.

Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B., &
Sethupathy, G. (2016, December). *The Age of Analytics: Competing in a Data-Driven World*. McKinsey Global Institute. Retrieved from http://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world.

HubSpot Research – Messaging Apps Have Over 5B Monthly Active Users. (2017,
July). *HubSpot Research*. Retrieved from
https://research.hubspot.com/charts/messaging-apps-have-over-4b-monthly-activeusers.

King, R. (2012, August 10). Panel: Netflix, StubHub, IBM Execs Discuss Value of
Big Data. *ZDNet*. Retrieved from
http://www.zdnet.com/article/panel-netflix-stubhub-ibm-execs-discuss-value-of-big-data/.

Lecompte, C. (2015, September 1). Automation in the Newsroom. *NiemanReports*.
Retrieved from
http://niemanreports.org/articles/automation-in-the-newsroom/.

Lewis, S.C. (2015). Introduction – Journalism in an Era of Big Data: Cases,
Concepts, and Critiques. *Digital Journalism*, 3 (3), 321-30.

Napoli, P. (2016). Special Issue Introduction: Big Data and Media Management.

International Journal on Media Management, 18 (1), 1-7.

Neill, M., & Schauster, E. (2015). Gaps in Advertising and Public Relations
Education: Perspectives of Agency Leaders. *Journal of Advertising Education, 19 (2), 5-17.*

Richterich, A. (2018). *The Big Data Agenda: Data Ethics and Critical Data Studies*.
London: University of Westminster Press.

Redvall, E.N. (2016). Film and Media Production as a Screen Idea System. In P.
McIntyre, J. Fulton, & E. Paton (Eds.), *The Creative System in Action: Understanding Cultural Production and Practice* (pp.139-54). Basingstoke: Palgrave Macmillan.

Rodríguez-Mazahua, L., Rodríguez-Enríquez, C.A., Sánchez-Cervantes, J.L.,
Cervantes, J., García-Alcaraz, J.L., Alor-Hernández, G. (2016). A General Perspective of Big Data: Applications, Tools, Challenges and Trends. *Journal of Supercomputing*, 72, 3073-113.

Sivarajah, U., Kamal, M.M., Irani, Z., & Weerakkody, V. (2017). Critical Analysis
of Big Data Challenges and Analytical Methods. *Journal of Business Research*, 70, 263-86.

Tableau Software – Top 10 Big Data Trends 2017. (2017, January). *Tableau
Software*. Retrieved from
https://www.tableau.com/resource/top-10-big-data-trends-2017.

The Guardian – Editorial. (2018, April 17). The Guardian View on Facebook's
Business: a Danger to Democracy? *The Guardian*. Retrieved from https://www.theguardian.com/commentisfree/2018/apr/17/the-guardian-view-on-facebooks-business-a-danger-to-democracy.

Underwood, C. (2018, January 17). Automated Journalism – AI Applications at
New York Times, Reuters, and Other Media Giants. *Techemergence*. Retrieved from https://www.techemergence.com/automated-journalism-applications/.

Wagner, K. (2017, September 7). Two-Thirds of Americans Are Now Getting News
From Social Media. *Recode*. Retrieved from https://www.recode.net/2017/9/7/16270900/social-media-news-americans-facebook-twitter.

Xu, Z., Frankwick, G., & Ramirez, E. (2016). Effects of Big Data Analytics and
Traditional Marketing Analytics on New Product Success: a Knowledge Fusion Perspective. *Journal of Business Research*, 69, 1562-66.

RAMA VENKATASAWMY, PhD is currently teaching in the Bachelor of Communication program at the Singapore campus of the University of Newcastle, Australia. More details are available at: http://www.newcastle.edu.au/profile/rama-venkatasawmy. Email: Rama.Venkatasawmy@newcastle.edu.au