



## Making Sense of Data Journalism from Bangladesh Perspective

Md Aminul Islam\*

Department of Journalism, Communication and Media Studies, Varendra University, Bangladesh

\*Corresponding author: Md Aminul Islam, Professor, Department of Journalism, Communication and Media Studies, Varendra University, Bangladesh, Tel: 0721-751274; E-mail: aminul.vu@gmail.com; aminul@vu.edu.bd

Received date: Nov 12, 2018; Accepted date: Nov 15, 2018; Published date: Nov 22, 2018

Copyright: © 2018 Islam MA. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Citation: Islam MA. Making Sense of Data Journalism from Bangladesh Perspective. Global Media Journal 2018, 16:31.

### Abstract

With emergence of information and communication technologies, our ways of living and interactions, making decisions and comprehending reality are being changed. The fundamental aspect of life is being changed by adding a quantitative dimension to it. The traditional logic of journalism has dramatically been changed by infusing computational thinking into the newsrooms. It has also changed the linear storytelling into a more interactive and engaging way. Troves of available data are offering us a new perspective on reality to be understood, explained and communicated. Most of the scholarships on data journalism have been done from the context of western and developed world. Little has been done from the context of a developing country like Bangladesh. When the role of data grows as a basis for reporting and storytelling, answers to the questions relating to the use or utilization of data, skills needed to do this, platforms and tools of data journalism, potential change in newsrooms needed to be sought. This article tries to answer the questions from the perspective of Bangladesh. It also explore the philosophical turns in understanding the recording, retrieving, explaining, interpreting and representing the reality through the lenses of Big Data and data journalism.

undoubtedly one of the hottest trends in the journalism, as well as in journalism studies, and getting much attention in professional journalistic practices and academia. Several books have been published and articles about data journalism are getting spaces in the scientific journals in the field. Most of the works have been done from the context of western and developed world. Little has been done from the context of a developing country like Bangladesh. On the other hand, with increase of the academic publications, debates, theorizing and practices of data journalism, some questions are being raised: If the role of data grows as a basis for reporting and storytelling, then what are the skills needed to do this? What can be done with data? How will newsrooms change? What are the platforms and tools needed to do this? The case of Bangladesh's news media industry is no exception of the questions. When the role of data grows as a basis for reporting and storytelling, answers to the questions relating to the use or utilization of data, skills needed to do this, platforms and tools of data journalism, potential change in newsrooms needed to be sought. This article tries to answer the questions from the perspective of Bangladesh. It also explores the philosophical turns in understanding the recording, retrieving, explaining, interpreting and representing the reality through the lenses of Big Data and data journalism.

### Data and Big Data

The word 'data' means 'given' in Latin, in the sense of a fact [2]. It refers to a description of something that can be recorded, analyzed and recognized [3]. Anything countable can be counted as data. Anything that a computer processes is data. In a general sense, 'data' is any collection of numbers gathered on a spreadsheet [4]. A data set would be considered as 'big data' if it is huge in volume, high in velocity, diverse in variety, and represents a veracity [5,6]. At the heart of the four Vs, there are new frontiers of possibilities to understand, analyze, interpret, and represent the reality. At its core, the big data is all about finding out a pattern in course of actions, issues, event and other social processes [7]. It is about finding connections, connections with the people around us, and connections between people's behavior and outcomes [8]. It is fact about people's behavior [9,10] instead of their beliefs. It's about the behavior of customers, employees, and prospects for a new business. Big Data may be originated from things like

**Keywords:** Data; Big data; Data journalism

### Introduction

Information and communication technologies have fundamentally transformed the way we make sense, understand and explore the world around us. Every day ordinary people are confronted with troves of statistically-based information. Data-processing tools are becoming an integral part of the newsroom's environment around the world. Understanding how data and computational journalism are affecting news norms, practices and organizations is essential to adapt to the new and ever-changing information, communication and media ecology. In a simple form, data journalism is quantitative form of journalism [1]. It is

location data of the cell phone or punch of credit card and the things people post on Facebook, searches on Google or share on twitter. So, increasingly it is becoming all about the real behavior of the people. By analyzing the data, scientists can tell an enormous amount about a person. They can tell whether the person would pay back bank loans, get any disease or involved any criminal activities. They can do this because thoughts, emotions and actions of person are largely determined by his/her social context [11]. If one aspect of a person's behavior is understood, the rest can be predicted by analyzing and comparing the person to the people in his or her social contacts. Because people are so enmeshed in the surrounding social fabric that it determines the sorts of things that they think are normal, and what behaviors they will learn from each other. Big data shows us the connections that cause these events. Big data gives us the possibility of understanding how these systems of people and machines work, and whether they're stable.

With emergence of information and communication technologies, our ways of living and interactions, making decisions and comprehending reality are being changed. The fundamental aspect of life is being changed by adding a quantitative dimension to it, and big data is all about seeing and understanding relations within and among pieces of information [12]. It means harnessing information to build useful insights of issues, events and course of actions. It has transformed the way we understand and organize the society. The world of big data is shaking up everything from business and science to healthcare, government, education, economics, the humanities, and every aspects of society (). Nature of business, market and society are being changed [13]. Big data often messy, varies in the quality, and is distributed among countless servers around the world. Exactness requires carefully curate data. With enough data, numbers can speak for themselves. Properly analyzed data can help to make sense and prediction of spread of pandemic, natural disaster or threat of terrorist attack, and thus it can help to prevent the spread of diseases, loss of wealth and lives.

We are living in the age of datafication [14,15]. The demarcation line between primitive societies and advanced societies is the ability to record information [16,17]. The amount of stored information grows four times faster than the world economy [18]. Today, Facebook knows our likes, Amazon can recommend the ideal books, Google can rank the most relevant websites, LinkedIn defines whom we know. Now about 2.5 quintillion bytes of data are created each day in the world; more than 3.7 billion people use the internet worldwide; on average, Google now processes 3.5 billion searches per day; Snapchat users share 527,760 photos; users watch 4,146,600 YouTube videos; 456,000 tweets are sent on Twitter; Instagram users post 46,740 photos per minute; 1.5 billion people are active on Facebook daily; and 156 million emails are sent daily [19].

### Data journalism and its scope

What 'word's means in traditional journalism, 'data' means data journalism. In traditional form of journalism, a story is a

narrative that is true, interesting and relevant to its audience. While, data journalism is the new way of understanding, interpreting and making sense of the world through careful analysis of vast amounts of data'. It is use of data as a tool to get closer to the truth of what is going on in the world. It is all about an interaction between journalists and computer scientists, statisticians and graphic designers, and bridging the gap between statisticians and wordsmiths.

In the traditional form, journalism is all about putting together information, facts, opinion, quotes, and data to explain the significance of world events and their context to create a narrative. Meanwhile, data journalism is about blending of digital data abundance and analytic techniques for data-driven insights [20]. The insight is gained through careful use of algorithms, computation, and quantification [21]. The process of data journalism include: access to or production of structured data-sets, using the data-sets for research of a story, discuss data/data source in published piece, provide raw-data for the users along with the story [22]. It means figuring out how to get the data, how to understand it, and how to find the story.

However, there is a clear distinction between 'data' and 'story'. Data itself is a form of information. To extract information, a journalist has to go through a process, usually first cleaning up messy formats, structuring and sorting the data into a readable format before it can be visualized. Story is an interactive form of communication, where information is brought into a context that people can understand, remember, discuss and tell others about. Data journalists use large databases, infographics, data visualization and interactive visualization to tell stories. It is about uncovering truth through mining trove of information the public do not have enough time and skills to do themselves, interrogating it, interpreting it, making sense of it and sharing it with the audience. Data journalism is about looking for something that is categorizable, quantifiable and comparable in any news topic. It also can be understood as the media's attempt to adapt and respond to the changes in our information environment — including more interactive, multi-dimensional story-telling, enabling readers to explore the sources underlying the news and encouraging them to participate in the process of creating and evaluating stories.

We live in a data-driven world. Everything of human action is data—even toilets [23,24]. Today, the banks, the telephone companies, the medical companies, advertisers, insurance companies, national security agencies have the most valuable data. With rapid advancement of technology, tectonic changes are taking place in services such as personal recommendations, identity certification without passwords, personal and public services for transportation, health, and so forth. Every time we send a message, make a call, or complete a transaction, we leave digital traces. Surveillance, biometrics, automation, data creeping, or profiling consumer behavior, all offer opportunities and challenges to news reporting. Data can be used to provide deeper insights into what is happening around us and how it might affect us. From telling people what has happened, data can help journalists provide the analysis

and information that they need to be able to make sense of the important issues of the day. If the data is used carefully and accurately, it can give us unprecedented scope to understand our society, and improve the way we live and work. Data can improve a complex story; can help journalists speak truth to power. Combined with traditional reporting techniques, data can help tell stories in more compelling and innovative ways and give citizens actionable information. A data journalist can help his/her audiences to make sense of the government's initiatives whether those are working or failing. The journalist can tell stories by investigating the pattern of the rates of poverty, infant mortality, crime, changes in GDP and other social indicators, budgetary allocations, currency exchange rates, rise and fall of crop productions etc.

### History and forms of data journalism

In journalism, reporting is an invention of 19th century [25]. A critical investigation into the history and models of journalistic practices reveal that epistemologically journalism is guided by two ideals: the ideal of the “story” and the ideal of the “information”. The ideal of the “story” telling is rooted in the narrative structure and expert human knowledge, which is qualitative in nature. While, the ideal of the “information” providing is grounded in mathematical data, facts, statistical analysis, finding correlations and patterns, which is quantitative in nature.

Since mid of the 20th century, a shift is being happening in the philosophy of journalism—from qualitative to quantitative. From qualitative view, journalism is explanatory and interpretative. It seeks to explain the world rather measures it. As human being is conditioned to look for causes, it searches for causality of issues and events of the world. In search of truth or reaching the truth, it first builds hypothesis then collects relevant information to justify it. Primarily it deals with words. Journalists go to the field to gather information by observation and interaction with people. They record what they find, and then analyze, interpret it to show how the world function. Historically, traditional journalism has been built around two elements—textual and visual.

Meanwhile, availability of computer has made possible to harness vast quantities of data opening the door to new ways of understanding and exploring the world. The certainties that we believed in are changing. A transition is taking place from hypothesis-driven world to data-driven world. The world is shifting from causation to correlation. However, the use of data in telling stories is not new one rather it is the continuation of human kind's ancient quest to measure, record and analyze the world.

Today what we now call data journalism is computer-assisted reporting (CAR) [26]. The use of computers in journalism dates back to the 1950s. US media organization CBS first used CAR in 1952 to predict the result of the presidential election. It was the first organized and systematic approach to using computers to collect and analyze data to tell news stories. Since the 1960s, many journalists analyzed databases of public records with scientific methods. Philip Meyer pioneered the CAR and the application of social and behavioral

science research methods to the practice of journalism. The other pioneers in CAR are: Donald Barlett, James Steele, Adrian Holovaty, and Oren Etzioni. In the early 1970s the term precision journalism was coined to describe this type of news-gathering [27]. Systematic study of data journalism had begun in the 90s of the past century mainly in USA. Scholars have coined multiple terms to describe the phenomenon. Among the terms are: data Journalism, data-driven journalism, computer-assisted reporting, database journalism [28], structured journalism, computational journalism, database journalism [29]. All the terms have very similar professional and epistemological roots—quantitative and computational form story telling. The quantitative forms of journalism try to discover patterns and correlations in the data that offer novel and invaluable insights. The insights bring the certainties about what, not why. The correlations may not tell us precisely why something is happening, but they alert that it is happening. We do not always need to know the cause of a phenomenon; rather we can let data speak for itself.

### Necessary skills for data journalism

The skills necessary to succeed in the workplace are changing. Making sense of big data, particularly unstructured data, is the key skill needed for a journalist today. A data journalist must be able to curate, verify, analyze and synthesise the trove of data, provide context, clarity and, find truth in the expanding amount of digital content in the world [30]. It does not a replace of traditional skills of journalism; rather it demands new set of skills for searching, understanding, and analyzing, visualizing and programming data. Today's journalists need an overlapping set of competencies drawn from different fields. They should have skills and knowledge in the statistical methods of social scientists, the mapping tools of GIS, the visualization arts of statistics and graphic design, and a host of skills that have their own job descriptions and promotion tracks among computer scientists: web development, general-purpose programming, database administration, systems engineering, data mining (even, cryptography). So, today journalist is a complete package of a statistician, software programmer, info-graphic designer as well as a storyteller.

Data is so widely available and so strategically important that the scare thing is knowledge to extract wisdom from it. After all, a goldmine is not worth anything if gold can be extracted. Today, expertise in database management, data science, analytic and machine learning algorithm are the key to professional success in journalism. In making news, a data journalist should be able to explore news stories from datasets, create a bigger picture by linking the datasets, seek patterns in issues-events and actions, analyze data behind the stories, connect data to stories, put stories into context, publish datasets with stories, and present visual representation of information. Moreover, a data journalist should have the following skills: ability to differentiate ‘data’ and ‘story’; have literacy in numbers and computation; able to understand exactly how the quantification process works in the society; able to blend technological and journalistic competences; able to find data to support stories; able to

finding story ideas with data analysis; able to deal with messy data; able to tell stories with visualization and able to tell individual stories that reflect overall change.

## Locating Data Journalism in Bangladesh

Emerging news norms, practices and organizations are the key indicators to understand the phenomenon of data journalism and computational journalistic field in a country. The norms and practices vary depending on journalists' educational backgrounds and skills, use of tools and techniques, and organizational goals. As part of the work, several informal in-depth interviews have been conducted among mid-career and senior journalists based in Dhaka, the capital of Bangladesh. Unstructured observation of news stories published in Dhaka-based newspaper also helped to understand the phenomenon.

Data journalism is still uncommon in Bangladesh, so is computational journalistic field in the country. Data journalism is at an early stage. Bangladeshi media occasionally, now more frequently, does stories based on data. But they are not exploring the field in a focused manner. But there are few in-depth investigative reporting based on extensive data analysis. In most cases, Dhaka-based news outlets do stories by using the data amassed by some international organizations such as World Bank and WHO. Most data based stories are done on business and economic issues and sports storytelling. Bangladesh Bank, Bangladesh Bureau of Statistics and several other government bodies are the main sources of data used in the stories. But efforts to correlate/juxtapose data from different datasets to create an interesting observation are not evident in the news stories. Meanwhile, most of the data-based stories are not actually what data journalism should be. Because data journalism is not merely about crunching numbers and interesting findings, rather story is more important than just the analysis and presentation of data.

Access to open and public data is crucial to meaningful data journalism. Open data is hardly available in Bangladesh. A lack of open information laws, very complex, messy and poorly organized public data and low digital literacy are also the barrier to the practice of data journalism in the country. Meanwhile, culture of mistrust is one of the big problems for data journalism in Bangladesh. Government officials and journalists mistrust each other. Journalists often accuse authorities of cover-ups and acting secretly, while officials tend to perceive journalists as highly unreliable when relaying facts and information. On the other hand traditional newspapers do not data stories section as they have not really taken to data journalism seriously. Most of the news media outlets find it financially unviable to have its own research and statistics wing.

behind puzzling columns of numbers. Putting the numbers in context and analyzing those with insights can lead to uncovering stories. Journalists need to recognize the stories behind the data, know how to analyze the data and understand how to use. Interviews with journalists show that

despite having confronted with a lot of data, many of them do not know how to analyze it and how to turn it into a story. The main challenges they face are a shortage of time and the need for training and developing data journalism skills.

## Conclusion

The ability to record information is one of the lines of demarcation between primitive and advanced societies. Written language enabled early civilizations to measure reality, record it and retrieve later, which facilitated the creation of data. The core philosophy of big data is a continuation of human kind's ancient quest to measure, record and analyze the world. Tectonic shift have been taking place in the philosophical bases of recording and retrieving data, and measuring the reality in the world. The shifts are: from causation to correlation, from a hypothesis-driven world to a data-driven world, from the search of causality to the search of computation and quantification, and less reliance on normative assumptions. In seeking truth and representing the reality, traditionally we search for causality. Because, our minds, reasoning and cognitive process have been programmed for looking the cause of issues, events and actions around us. But big data offer us to discover patterns and correlations in the phenomena of the world. The correlations do not tell us precisely why something is happening, but they alert us that it is happening. Big data is about what, not why. We do not always need to know the cause of a phenomenon; rather we can let data speak for itself.

Human decisions are based on information that is small, exact and casual in nature. Harnessing vast quantities of data opens the door to new ways of understanding and exploring the world. Quantification, the process that generate data, is the first step to revealing trends and patterns in society that can then become a story. Using data means there is less guesswork about what the facts are. A combination of tools and receptive mindset has been making possible mining the reality in a new way. The main question at the heart of quantitative thinking is "compared to what?"

Journalism is a craft. It is craft of searching, filtering and transforming a maze of aspects into information. Information then is defined as data that we can act upon. With development of new information and communication technologies, the world is going toward an ecosystem of open data. In the ecosystem, there are trove of loosely connected bytes and pieces of small information which are easy to reuse and easy to recombine, where there are numerous creator, producer, contributors and maintainers and which is distributed and decentralized.

Today we live in a numerate society where words, location, and interactions are becoming data. Data journalism has disrupted the traditional logic of journalism by infusing computational thinking into the newsrooms [31]. It has also changed the linear storytelling into a more interactive and engaging way to offer content. Not it is presumed that the world is understandable with number and math. Troves of available data are offering us a new perspective on reality to



be understood, explained and communicated. Traditional journalism defines fact as information and quotes from official sources [32]. But the fundamental aspects of traditional ways of gathering, filtering, and disseminating news, journalistic institutions, business models, distribution channels, and audiences are being changed radically with the emergence of internet, cloud computing, mobile devices, and open source software. In the developed world, today journalists are creating news apps with interactive features and telling compelling stories by harnessing the vast quantity of data that help people to make sense of and decision about their real-world, explore more data and act upon the insights derived from it.

Doing data journalism is all about three major shifts in mindset: first, ability analyze vast amount of data about a topic. Second, willingness to embrace data's real-world messiness rather than privilege precision and third, a growing respect for correlation rather than a continuing quest for elusive causality. The challenge lies in carefully filtering and analyzing large amounts of information. It is not enough simply to gather data; in order to yield meaningful predictions; the data must be placed in an analytical framework so that the audience can make sense of the world around them.

Data-driven decisions can augment or overrule human judgment. Statistical analyses force people to reconsider their instincts. The data can reveal what people want to read about better than the instincts of experienced journalists. In dealing with data, one of the great questions is: who owns and produces the data? Another important issue with big data is that since this data is mostly about people, there are enormous issues about privacy [33], data ownership, and data control, un-careful handling of big data easily produce false correlations. As result of this, there is a danger of being precisely [34] in telling the truth, representing the reality.

## References

1. Gray J, Bounegru L, Chambers L (2012) *The data journalism handbook*. Cambridge: O'Reilly.
2. Rogers S (2013) *Facts are sacred*. London: Faber and Faber.
3. Sato A (2014) *Applied Data-Centric Social Sciences Concepts, Data, Computation, and Theory*. Japan: Springer.
4. Herzog D (2015) *Data literacy (1stedn)*, New York: SAGE Publications, Inc.
5. Hilbert M (2015) Big Data for Development: A Review of Promises and Challenges. *Development Policy Review* 34: 135-174.
6. Bhadani A, Jothimani D (2016) Big data: Challenges, opportunities and realities, In: Singh MK, Kumar DG (eds.) *Effective Big Data Management and Opportunities for Implementation*. Pennsylvania, USA, IGI Global, pp: 1-24.
7. Dumbill E (2013) Making Sense of Big Data. *Big Data* 1: 1-2.
8. Pentland AS (2014) *Living Big Data*.
9. Shmueli G (2016) Analyzing Behavioral Big Data: Methodological, Practical, Ethical, and Moral Issues. *SSRN Electronic Journal*.
10. Zwilling M (2018) What Can Big Data Ever Tell Us About Human Behavior?
11. Baez S, García A, Ibáñez A (2018) How Does Social Context Influence Our Brain and Behavior?. *Frontiers For Young Minds*.
12. Mayer-Schonberger V, Cukier K (2013) *Big data: A revolution that will transform how we live, work, and think*, London, John Murray.
13. Sinclair TM (2015) *Economic Forecasts in the Age of Big Data*.
14. Arsenault A (2017) The datafication of media: Big data and the media industries. *International Journal Of Media & Cultural Politics* 13: 7-24.
15. Lycett M (2013) 'Datafication': making sense of (big) data in a complex world. *European Journal Of Information Systems* 22: 381-386.
16. Mayer-Schonberger V, Cukier K (2013) *Big Data: A Revolution That will Transform How We Live, Work, and Think*, UK: John Murray (Publishers)
17. Parasie S (2014) Data-Driven Revelation? Epistemological tensions in investigative journalism in the age of "big data". *Digital Journalism* 3: 364-380.
18. Hilbert M, Lopez P (2011) The World's Technological Capacity to Store, Communicate, and Compute Information. *Science* 332: 60-65.
19. Marr B (2018) How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read.
20. Lewis S (2014) Journalism in an Era of Big Data: Cases, Concepts, and Critiques. *Digital Journalism* 3: 321-330.
21. Lewis S, Westlund O (2014) Big Data and Journalism. *Digital Journalism* 3: 447-466.
22. Meyer P (2002) *Precision Journalism: A Reporter's Introduction to Social Science Methods*. Lanham (Maryland): Rowman & Littlefield.
23. Lun K (2018) The Datafication of Everything - Even Toilets. *Yearbook Of Medical Informatics* 27: 234-236.
24. Fritsch K (2018) Towards an Emancipatory Understanding of Widespread Datafication. *SSRN Electronic Journal*.
25. Schudson M (1999) *The New Journalism*. In: David C, Paul H (eds.) *Communication in History: Technology, Culture and Society*, New York: Longman
26. Houston B (2003) *Computer-Assisted Reporting: A Practical Guide*. New York: Bedford/St. Martin's.
27. Meyer P (1991) *The new precision journalism*. Bloomington: Indiana University Press.
28. Appelgren E, Nygren G (2014) Data Journalism in Sweden. *Digital Journalism* 2: 394-405.
29. Anderson C (2012) Towards a sociology of computational and algorithmic journalism. *New Media & Society* 15: 1005-1021.
30. Vallance-Jones F, McKie D, Wolfe-Wylie W, McGregor G (2017) *The Data Journalist: Getting the Story (1st ed.)*. Oxford: Oxford University Pres.
31. Porlezza C, Porlezza C (2018) How Data Journalism Is Changing Newsrooms... Slowly.
32. Hermida A (2010) Twittering the News - The Emergence of Ambient Journalism. *Journalism Practice* 4: 297-308.

33. Mai J (2016) Big data privacy: The datafication of personal information. *The Information Society* 32: 192-199.
34. McFarland D, McFarland H (2015) Big Data and the danger of being precisely inaccurate. *Big Data & Society* 2: 205395171560249.