Data Journalist, A Hope for Indonesia's Quality Journalism in the Digital Era?

Pratiwi Utami[1]

[1]Monash University, Victoria, Australia

Abstract: The declining of quality journalism as well as the the information overload in the era of big data have provoked the occurrence of data journalism. In short, data journalism is journalism done based on vast and diverse data. It produces data-driven – rather than statement-driven – news reports. Data journalism aims to bring back the quality journalism and help people translate big data by creating in-depth, meaningful and contextual narratives. This paper will uncover the data journalism practice in Indonesia, by elaborating the professional practice of data journalism in two online news sites, Katadata and Visual Interaktif Kompas (VIK). Interview to both sites' key persons as well as Indonesian data journalism academician was executed to figure out the development of data journalism in a unique context of Indonesia. This paper also discusses the data journalists' challenges and chances in improving the quality of online journalism in the future. Based on the interview and observation to the current practice of online journalism in Indonesia as well as a reflection on the news sites' content, it is visible that quality content can have a place in the middle of – as well as promote betterment in – "speedy but shallow" culture in Indonesian online media industry.

Keywords: Big Data, Data Journalism, Online Journalism Industry

Introduction

"The challenges and opportunities presented by the digital revolution continue to disrupt journalism," said Alex Howard from O'Reilly Media (Data Journalism Handbook 2012a, para. 17). Obviously, the emergence of Internet requires a journalist to adapt with new tools and the nature it brings in the news-gathering process. The Internet allows a reporter to post news, videos, or images within seconds. However, it drives the growth of online journalism genre, of which perceived as troublesome (see, for example, Campbell 2012; Margianto & Syaefullah 2013; Sukarno 2015; Himawan 2011; Holiday 2015). Don Campbell, in his column in USA Today, concerned a new culture brought by online climate to journalism world. He criticised the trend of fast but inaccurate news production in online media (2012, para. 9-10). Such media are racing to publish news as soon as they can, using "get it first!" principle rather than "get it right!". They strive to be the first among the others to tell the world the latest issue and tend to neglect the validity of the information as well as the ethics (Sukarno 2015, para. 8). *Speed-driven journalism*, thus this practice is known. It is identical with fast, short, and brief news, providing room for a journalist to update the news anytime. Online media adopt this practice solely for the sake of the speed.

The Internet also contributes to the information overload. According to Domo's infographic about the amount of data created on the web (2014), there are 277,000 Tweets, 2.46 million Facebook content shared, and 240 million e-mail sent every minute. In the same time span, Internet users upload 72 hours of video on YouTube and send four million search queries to Google. This big data are available on the web and openly accessible to the public. However, given how much of information floating, users may find it overwhelming and not all of them are able to interpret the flooding data. The information becomes overabundant and somewhat useless. Considering the declining of the journalism quality as well as the meaningless information overload, data

Corresponding Author E-mail: puta3@student/monash.edu/pratiwi.utami@yahoo.com

journalism occurs to bring back the quality journalism and help people translate the big data by creating in-depth, meaningful and contextual narratives.

Data journalism becomes popular during 2009-2010 started in Canada and also Europe, but the practice of computer-assisted reporting had already existed back in 1952. CBS network in the US worked with computer-mainframe experts to predict the outcome of the presidential election (Houston 2015, para. 2). In Indonesia, data journalism is relatively very young. It was Katadata.co.id that began this practice in 2012. This online news site produces data-driven news supported with infographic as the data visualisation. It also releases the results of their data processing related to recent issues of economic, political, and business in Indonesia and the world. The data releases are open and accessible to the public on their website. Another online media that also adopt the data journalism practice in this country is Visual Interactive Kompas (VIK) through vik.kompas.com. To date, both sites are consistently presenting data-driven news and begin to gain popularity.

This essay will uncover the data journalism practice in Indonesia. The first parts of this paper will discuss data journalism and its development in the unique context of Indonesia. Hereafter, the discussion will be more specific in elaborating the professional practice of data journalism in Katadata and VIK. The final part of the essay will discuss the data journalists' challenges and chances in improving the quality of online journalism in the future. Based on the interviews with Kadatada and VIK's key persons and observation to the current practice of online journalism in Indonesia, it is visible that quality content can have a place in the middle of – as well as promote betterment in – "speedy but shallow" culture in Indonesian online media industry.

What is data journalism?

According to Oxford Dictionary, journalism is '[t]he activity or profession of writing for newspapers, magazines, or news websites or preparing news to be broadcast', while data means '[f]acts and statistics collected together for reference or analysis'. In short, data journalism is 'journalism done with data' (Bradshaw 2012, para. 1). A question arises from this definition: is this term overrated, because journalism is supposed to be *always* based on data?

Data journalism is not *just* about data; it relates to big data. It advances along with the availability of *open data* accessible to public and processable using *open source* software. Mirko Lorenz in his presentation for Innovation-Journalism Conference (2010) illustrated that data-driven journalism is a workflow from digging into a set of big data, filtering or mining for "nuggets" information, visualising the information using graphic of multimedia, connecting a storytelling with the statistics, and creating media that have values for the reader. Similar to Lorenz, Bradshaw illuminates that data journalism is a combination of the traditional "nose for news" and the 'ability to tell a compelling story, with the sheer scale and range of digital information now available' (Bradshaw 2012, para. 4).

This term became popular in 2010 when The Guardian and New York Times analysed and published the data released by WikiLeaks. These Excel spreadsheets contained 92 thousand columns drew the US military and allies involvement in the Afghanistan War from 2004 to 2009 (Wendratama 2015, p. 21). At that time, the journalists from both media were confused about what to do with it, although they were aware of the value and the importance of the extensive data. They started to classify data by topic, place, and event, after which they pulled one key story to explain the typical of the war: attack against the West military was executed using improvised explosive devices (IED). This key then became the entrance to the investigation to the sea of data (Wendratama 2015, p. 21).

Another essential thing in data journalism practice is visualisation. In this practice, a clear and concise narration is equipped with attracting design to provide more engaging experience to the audiences. The Las Vegas Sun's 2010 ″Do No Harm″ series on hospital care packed the data in different elements of visualisation: an interactive

46

graphic for the reader to see in what hospital the surgical injuries happened more often; other interactive graphic which allows users to categorise data to see where people are getting hurt; and a map completed with a timeline 'that shows infections spreading hospital by hospital' (Data Journalism Handbook 2012b, para. 2). Based on this variable, data journalism can achieve its best potential if it appears on online interactive media (Wendratama 2015, p. 21). Arguably, there are three keywords in understanding data journalism: big data, visualisation and online media.

So, why does data journalism matter? What is the difference with other types of journalism? Perhaps it is the potentiality in the fuse of "news intuition" in journalism tradition and the ability to write compelling stories based on vast and diverse data. Sarah Cohen, a US journalist and a Pulitzer winner for investigative reporting in 2002, convinces that data analysis can be beneficial to concede "a story's shape" in the news (Lorenz 2012, para. 5). In another sense, data are used to distinguish data journalism from the statement-driven conventional journalism (Yudiantika 2016, para. 7). The term "conventional journalism" here means the product of journalism from media, especially online media, which is created without depth, citing sources without providing a clear context as well as accurate and complete data, aims merely to catch clicks (Yudiantika 2016, para. 7).

Data journalism shoots to shape new services in the public sphere, to help consumers, managers, and politicians to understand the patterns and make a decision on the findings (Yudiantika 2015, para. 6). It is expected to create a better storytelling in the news and bring dry data closer to the reader's lives. Data journalism can help contesting the online journalists' "instant and shallow" image, thus positioning journalists in a more relevant role for the community with a new approach (Yudiantika 2015, para. 6).

New kids on the block: Data journalism practices in Indonesia

Data journalism practice in Indonesia was pioneered by Katadata.co.id. "Pioneer" here means a news channel who has started practicing data journalism consistently in reporting a large data sets. Formerly, Tempo magazine and Kompas daily have already been using data and infographic to present the news. Both are major print media in Indonesia and have an established research division and archive bank. However, according to Metta Dharmasaputra, a former Tempo journalist and current CEO of Katadata, what both media do in their research division is only at the stage of data collection, not data mining or processing, let alone big data operating.[i] Furthermore, data journalism is about online media.[ii] Tempo, the only Indonesian media involved in The Panama Paper report, indeed released the results of the colossal investigation initiated by International Consortium of Investigative Journalism (ICIJ) in *Tempo Investigasi* portal, investigasi.tempo.co.id. Interactive infographic produced by ICIJ also included in the articles. However, *Tempo Investigasi* website has not been consistently put data visualisation in other news on the site. Therefore, Katadata was conceivably the first online news site that implement data journalism.

Katadata is an online media and research company in economic and business. Established on 1 April 2012 in Jakarta, Katadata provides news, information, data, and research results in effective yet attractive packaging to make it easier to be understood (Katadata n.d.). Accordingly, the presentation of news and information is not only with text but also with interesting visualisation. Katadata's team consists of experienced journalists and researchers in energy, financial, macroeconomics, trade, and infrastructure. This site focuses on providing a selection of *quality content* news, composed in depth and comprehensively (Katadata n.d.). It also releases "Databoks" portal on the website as an effective mean to publish a variety of data and information to the public, as well as a source of data users can retrieve easily.

Dharmasaputra conveys that Katadata was founded to challenge the notion that investigative work in journalism must be executed with field reporting. He believes narrated data will do, as evidenced by ICIJ through The Panama Papers. In addition, Dharmasaputra and his colleagues are concerned about the quality of media that has

not shown a significant improvement although Indonesia has been experiencing the freedom of the press since 1998.

> "It can be dangerous, especially when it comes to crucial matters like the economic sector. If the news is only statement-driven, based on what regulators and politicians have said and are not supported by an analysis of valid data, there will be distortions of information. However, if we use data, the news will have a stronger basis for argument. In respect to this matter, I think we need to fulfill public's right to data-based information."[iii]

Following Katadata, Kompas Media group recently launched Visual Interactive Kompas (VIK) through vik.kompas.com. VIK contains in-depth reporting about current issues in Indonesia with an attractive display. It is an extension of regular online news site Kompas.com and served as Kompas media group's efforts in seeking for depth, meaning, and perspective to an event. VIK aspires to give a new colour in Indonesian journalism by presenting narratives enriched with graphics, photos, video, and audio (VIK n.d.).

The actors and the industry

The data journalist's routines in both Katadata and VIK are similar with typical journalists' daily activities, be it print, electronic, or online. Each day begins with Editorial Division meeting; the reporters and editors discuss what to cover by considering issues or agendas that have news values. Based on an interview with Bambang Priyo Jatmiko, one of the producers in VIK, in every meeting, the Data and Research Division staffs are involved to specify what data they should provide. They can also propose a topic based on any interesting data they have. In regards to topic selection, VIK selects what to display in the website based on the most-visited news in Kompas.com.[iv] Visit rates in Kompas.com indicate that readers are interested in particular topics; thus follow-up reports through VIK are necessary.

After determining an issue, the Editorial Division will explore all aspects of the issue and determines what data and illustrations are operated to make the news easier to understand. Next point is task assigning; the editorial team decides who does what. Reporters conduct field reporting and interview while editors work with data. In this stage, Data and Research Division escorts the editorial work by pulling published data government such as ministry websites or the Central Bureau of Statistics. In Katadata, data cleaning and filtering will be done by a computer program.[v] Once the reporter and editor finish composing the news, the editorial team discusses with Artistic/Design Division to assess which photos, videos, and infographics fit best with the article. After the layout process, then news are uploaded.

To date, Katadata hires 50 staffs. There are 11 personals in Editorial Division and 15 persons in Data and Research Division, with the balanced proportion between male and female journalists. With that squad, Katadata produces 20-25 news every day. Regarding journalist's skill and qualification, both Katadata and VIK do not require a journalist to have an academic journalism background. A candidate only needs an undergraduate degree from any major with minimum six-month experience in writing or being a journalist. Dharmasaputra expresses his concern regarding this.

> "This is actually my criticism to journalism qualifications in general di Indonesia. Until now there are no schools or faculties of journalism in Indonesian higher education. Journalism is only a part of communications and media studies, and consequently, most of the Indonesian journalists do not have a basic knowledge in journalism. It is very different from journalism practices in the West, where every journalist must have a degree in journalism to be able to work in this field."[vi]

Furthermore, a degree in Statistics or Mathematics is also not a requirement for data journalist in Indonesia. In most cases, they are learning by doing. Katadata and VIK prefer to appoint senior journalists who have 10-15 year experience to do the data processing and translating. The reason is that their experience makes them knowledgeable and competent to understand the issue and data more comprehensively. However, to upgrade the

editorial team's skill, Katadata and VIK collaborate with other parties. Katadata teams up with the Department of Statistics, Bogor Agricultural University in an upcoming work about the analysis of investment potential in Indonesia. VIK sends the journalists to workshops held by the Alliance of Independent Journalists. In the future, Katadata plans to employ data analysts, either to train the team or to work in the company.

The organisational structure in Katadata and VIK is also similar to other media in Indonesia: the separate division between editorial, data and research, and artistic or design. The communication within the organisation is conducted at three levels, namely internal Editorial Division, Editorial with the management (Chief Editor, Managing Editor), and Editorial with the Artistic. Regular discussions are usually carried out in internal divisions (between editors and reporters) as well as with the artistic staff. Meanwhile, communication with the higher management is not intensive.

Challenges and chances

Talking about the development of data journalism in Indonesia also means discussing the trail of online journalism. The first online journalism practice in this country was www.detik.com in 1998. This site introduced a news style in journalism: fast and to the point. Aiming to be the fastest, Detik.com produces short news with incomplete 5W + 1H elements (Margianto & Syaefullah 2013, pp. 17-18). This practice has destroyed the ecosystem of online media, as well as damaging public's perception about online journalism.[vii] Data journalist can take a role to mend this situation.

There are three visible roles of data journalist in Indonesia. First is to bring a change in online journalism culture in Indonesia, from statement-driven to data-driven practice. Dharmasaputra believes that data-driven news is more valid and accurate than the statement-driven ones. It is because quotes from a source are not neutral. In contrast, data has no interest thus can validate facts in the news, which surely must be accompanied with fine news writing. The second role is to fulfill digital society's need for quality information. Public have outraged with the poor quality of online journalism. In 2015, there were 95 complaints submitted to Indonesian Press Council; 75% of it was rooted in inaccuracy, imbalance, and opinion adding (Putera 2016, para. 3). To avoid such complaints, Katadata and VIK implement Journalistic Code of Ethics applied to all journalists in Indonesia as well as Guidelines on Online Media Reporting by the Press Council. However, in Katadata, journalists are required to be careful when using a personal account on the social media as it can influence the credibility of the journalist and the media they work for.[viii]

The third role of data journalist in Indonesia is to create meaningful, coherent, and contextual stories about the data. Lewis (2015) denotes that 'in the developed world of digital information technologies, we are situated in a moment of data deluge' which contributes to the 'overwhelming volume and variety of digital information' (p. 322). Data journalism can help readers making sense of the large sets of data. According to Dharmasaputra, Indonesian readers are quite interested in data-driven news.[ix] Katadata's infographic story about the comparison of the military budget among ASEAN countries has attracted almost one million visitors. Its coverage about dual-citizenship of Arcandra Tahar (Vice Minister of Energy and Mineral Resources) obtained 125,000 hits and was referred by other media.

To accomplish these roles, data journalists are advantaged by a great number of Internet users in Indonesia. According to Indonesian Internet Service Providers Association (APJII)'s survey (2016), 51.8% or a half of Indonesian total populations are Internet users and 97.4% of them are active social media consumer (Katadata 2016, para. 2-9). For Katadata, social media acts a tool to disseminate news and infographic. This company creates a fan page account on Facebook which has gained 220,000 likes. It means, everytime Katadata posts latest news or infographic, the information will be delivered to at least 220,000 readers. Dharmasaputra predicts that the high rate of Internet penetration and government programs in developing Internet infrastructure will overcome the digital gap in Indonesia.

Other opportunities for data journalists are the nature of online media that removes space and time limitation. News can last for a long time on the website, especially if it gets a high rate on visit and share. The ready availability of communication and information technology can facilitate the data processing. Multimedia features such as video and audio will complement text and graphics in the article. For the readers, an attractive and interactive news will help them comprehending the issues as well as interpreting the texts. However, on the other side of the equation, data journalists stumble upon internal and external factors in their development. From internal side, the biggest problem is changing journalists' mindset from statement-driven to data-driven practices. Not all journalists are willing to learn new things, let alone passionate to work with data.[x]

From external side, media companies in journalism industry still assess online reporter's performance based on the quantity of news he or she produces and how many clicks the story gains, rather than the quality. Also, the industry still hesitates to devote fund in investigative journalism, and such work is not inexpensive not only in Indonesia but other parts of the world. Hirohisa Hanawa, the producer of News Department di NHK, Tokyo, admits that the company spent seven million yen to start data journalism (Utomo 2015, para. 25). Dharmasaputra also mentions aggregator content as one of the challenging factors. Aggregator contents are websites or applications that are pulling information link from other websites and allow users to view plentiful information from one place (Rouse 2011, para.1). This kind of site challenges data journalist because it effortlessly gathers trending news and get visits from the users, therefore are appealing to advertisers. It considerably disregards serious journalistic works behind that news.

Conclusions: The future of data journalism in Indonesia

Data journalism is still young; it demonstrates a simple form of an infographic or audio-visual features. Katadata has not provided extensive in-depth reporting or taken a sophisticated investigative yet. Likewise, although VIK presents *long-form* news with multimedia, it is not always equipped with a supportive and interactive infographics. In terms of visualisation, Indonesian data journalist has not taken the most advantage of Internet's interactive nature and innovatively brought a more personal experience for the readers. Nevertheless, Katadata and VIK have recognised data journalism as not a solely quantified journalism or computerised newsroom. Tabary, Provost, and Trottier (2015) posit that 'data journalism is based on the use of quantitative data, but beyond the simple use of quantitative objects, there is the transformation of the journalistic product and its production process' (p. 68). Data journalism is supposed to be able to drive developed practice and process in online journalism. In fact, today Katadata ranks 933rd in Alexa.com. For a news site with only 20-25 news produced per day, it is fairly impressive.

The data journalism practice has started and should be appreciated. Katadata and VIK do have potency for the improvement of online journalism in Indonesia; at least, both media serve as the examples of how a long-form quality content can still be compelling and selling. The emergence of other online news sites with long-form reporting such as x.detik.com and tirto.id indicates that online journalism in Indonesia starts moving towards the better practices. In the future, Katadata and VIK are aiming for the high-quality journalism in Indonesia. However, to get there, it takes some preconditions, among them is the changing mindset of online journalists in writing the news and dealing with data. The journalists should not avoid data because it may generate quality reports. They also need to start utilising online applications such as import.io, easel.ly, infogr.am, and storymap.knightlab.com to help them processing the data, or piktochart.com, canva.com, and venngage.com to visualise it. In addition, there must be changes in the industry's culture and habits in measuring the performance of journalists and online media. Their performances should be evaluated based on the quality, not always the quantity.

"The challenges and opportunities presented by the digital revolution continue to disrupt journalism," said Alex Howard. However, "technology is easy. Journalism is hard," said David Ho from the Wall Street Journal (Zhang 2012, para. 1). People are frequently blaming information and communication technology for the

decreasing of journalism's quality. However, media types are only platforms, while the journalists have always been obliged to implement ethical and responsible work. Hence, as long as the information is still one of the staple foods for modern people, data journalist will always exist to serve the society.

Acknowledgement

References

Bradshaw, P 2012, 'What is data journalism', in J Gray, L Bounegru, & L Chamber, *Data journalism handbook*, viewed 24 October 2016, http://datajournalismhandbook.org/1.0/en/introduction_2.html.

Campbell, D 2012, 'Column: Dangers of speed vs. accuracy', USA Today, 24 April, viewed 26 October 2016, http://usatoday30.usatoday.com/news/opinion/forum/story/2012-04-24/journalism-social-media-giffords-paterno-nikki-haley/54513448/1.

Data Journalism Handbook 2012a, Why is data journalism important?, viewed 24 October 2016, http://datajournalismhandbook.org/1.0/en/introduction_2.html.

Data Journalism Handbook 2012b, Some favourite examples, viewed 24 October 2016, http://datajournalismhandbook.org/1.0/en/introduction_3.html.

Domo 2014, Data never sleeps 2.0, viewed 28 October 2016, https://www.domo.com/learn/data-never-sleeps-2.

Himawan, IQ 2011. 'Ethical problem in online journalism', Jurnal Dewan Pers, vol. 4, pp. 35-44, viewed 30 October 2016, http://dewanpers.or.id/assets/media/file/publikasi/jurnal/878550.Jurnal%20Dewan%20Pers%20edisi%20Ke-4.pdf.

Holiday, R 2014, 'Journalism's biggest problem is not what you'd expect—and it's entirely fixable', Observer, viewed 29 October 2016, http://observer.com/2015/04/journalisms-sourcing-problem/.

Houston, B 2015, 'Fifty years of journalism and data: a brief history', Global Investigative Journalism Network, viewed 26 October 2016, http://gijn.org/2015/11/12/fifty-years-of-journalism-and-data-a-brief-history/.

Margianto, JH & Syaefullah, A 2013, Online Media: Readers, Profit, and Ethics, Alliance of Independence Journalists – Indonesia, Jakarta, Indonesia.

Katadata n.d., About Katadata, viewed 24 October 2016, http://katadata.co.id/pagebottom/7/tentang-katadata.

Katadata 2016b, Half of Indonesian population are Internet users, 65 percent in Java, viewed 24 October 2016, https://translate.google.com/#id/en/Separuh%20Penduduk%20Indonesia%20Pengguna%20Internet%2C%2065%20Persen%20di%20Jawa.

Lewis, SC 2015, 'Journalism in an era of big data', Digital Journalism, vol. 3, no. 3, pp. 321-330, doi: 10.1080/21670811.2014.976399

Lorenz, M 2010, 'Data-driven journalism: what is there to learn?', SlideShare, viewed 29 October 2016, http://www.slideshare.net/mirkolorenz/datadriven-journalism-what-is-there-to-learn.

Lorenz, M 2012, 'Why journalists should use data?', Data Journalism Handbook, viewed 24 October 2016, http://datajournalismhandbook.org/1.0/en/introduction_1.html.

Putera, S 2016, 'On online journalism: where are we going?', *Digi-Journalism*, viewed 28 October 2016, http://digi-journalism.or.id/mau-dibawa-ke-mana-jurnalisme-online-kita/.

Rouse, M 2011, 'Content aggregator', *Tech Target*, viewed 30 October 2016, http://searchsoa.techtarget.com/definition/content-aggregator.

Sukarno, A 2015, 'Online journalism: fast but inaccurate', *Digi-Journalism*, viewed 28 October 2016, http://digi-journalism.or.id/jurnalisme-online-cepat-tapi-tidak-akurat/.

Tabary, C, Provost, A & Trottier, A 2016, 'Data journalism's actors, practices and skills: A case study from Quebec', *Journalism,* vol. 17, no. 1, pp. 66-84, doi: 10.1177/1464884915593245

Utomo, YW 2015, 'Statistics, computer and future journalism', *Kompas.com*, viewed 27 October 2016, http://sains.kompas.com/read/2015/06/15/13021221/Statistika.Komputer.dan.Jurnalisme.Masa.Depan.

Wendratama, E 2015, *National Health Insurance data journalism: Data journalism handbook of National Health Insurance*, MediaLink, Jakarta.

Yudiantika, AR 2016, 'Data journalism and big data', *Pindai*, viewed 27 October 2016, http://pindai.org/2016/06/27/jurnalisme-data-dan-big-data/.

Zhang, M 2012, 'David Ho on the daily and creating WSJ's iPad app', *FishbowlNY*, viewed 31 October 2016, http://www.adweek.com/fishbowlny/david-ho-wall-street-journal/70653.

---

[i] In a personal interview via Skype, October 25, 2016

[ii] E. Wendratama, the author of "National Health Insurance data journalism: Data journalism handbook of National Health Insurance" (2015), in a personal interview via e-mail, October 10, 2016.

[iii] In a personal interview via Skype, October 25, 2016.

[iv] In a personal interview via e-mail, October 27, 2016.

[v] In a personal interview via Skype, October 25, 2016.

[vi] Ibid

[vii] Wendratama, in a personal interview via e-mail, October 10, 2016

[viii] Dharmasaputra, in a personal interview via Skype, October 25, 2016

[ix] Ibid.

[x] Jatmiko, in a personal interview via e-mail, October 27, 2016