

# BIG DATA . . . small story

Bernard C K Choi <sup>1,2,3</sup>

## BIG DATA: WHAT IT IS

In modern society, a huge volume of data (big data) are being collected and accumulated, and growing exponentially.<sup>1 2</sup> Data come from the internet, social media, medical records, customer databases, massive open data and other data sources. Big data are huge in quantity (volume), fast in production sometimes in real time (velocity), structured and unstructured in data types (variety), inconsistent and changing (variability), uncertain in data quality (veracity) and potentially useful (value).<sup>3 4</sup>

Big data, when analysed properly, can improve corporate performance, increase productivity, provide insights and predict future scenarios for better decisions in medicine, public health, science and technology, business and other sectors.<sup>5 6</sup> Big data, when misused, can lead to data breach and privacy concerns<sup>6</sup> and, when misinterpreted<sup>7</sup> or manipulated,<sup>8</sup> can cause false predictions leading to potentially disastrous consequences.

## SMALL STORY: PIZZA DELIVERY

Here is a small story that might help understand the impacts of big data. It is in part inspired by some ideas already posted on the internet.<sup>9 10</sup> The story is about a customer who orders pizza by phone, and the interesting exchange that ensues, delightful and not so delightful, that potentially can happen during the big data era.

**Pizza:** [*The phone rings*] Thanks for calling. How may I help you?

**Customer:** I would like to order a pizza.

**Pizza:** What is your customer number?

**Customer:** XY1357

**Pizza:** Hi, Mr. Lee. Your address is 789 Main Street. Your phone number is 987–1234. Which credit card do you want to use—the one ending in 1122, or 3210, or 2468?

**Customer:** How do you know my credit cards?

**Pizza:** We are connected to the CRM system—Customer Relationship

Management. You want the usual?

**Customer:** You know my usual?

**Pizza:** According to our records, you ordered pizza with pepperoni and mushrooms nine times in the last 2 months.

**Customer:** I want a seafood pizza this time.

**Pizza:** Seafood is not good for you.

**Customer:** Why?

**Pizza:** According to your medical records, you have high cholesterol and high blood pressure.

**Customer:** What would you recommend?

**Pizza:** You should try our new low-fat healthy pizza.

**Customer:** Are you sure I will like it?

**Pizza:** Last week you borrowed a cookbook from the Central Library on 'Low-Fat Healthy Recipes'.

**Customer:** OK, I will order a large 'low-fat healthy pizza'. Is it big enough for five people?

**Pizza:** It is big enough for your family. Facebook tells us that you have a spouse and two boys aged 6 and 9. Your mother also lives with you. Please make sure that she keeps her portion small. We see she had heart bypass surgery done just last month.

**Customer:** OK, OK. Can I pay by credit card now?

**Pizza:** Let me check your credit cards ... Sorry, you have to pay cash. All three of your credit cards have exceeded the credit limits.

**Customer:** OK, I will go get some cash at a bank machine.

**Pizza:** According to the records, you have already exceeded your daily withdrawal limit for today.

**Customer:** Then deliver the pizza to my home. I have cash at home.

**Pizza:** Don't worry about the money. We can directly access your bank account, where you have direct paycheque deposit. This account is the best.

**Customer:** How long will it take for the delivery?

**Pizza:** About 30 min. Or you can pick it up. It is faster.

**Customer:** Why?

**Pizza:** According to our GPS—Global Positioning System—you are in a car right now on Main Street making a right turn onto First Avenue. It will take you 4 min and 32 s to get here.

**Customer:** This is excessive and awful! It seems like you are spying on me every moment.

**Pizza:** I'm sorry. But we are doing this just to help you!

**Customer:** What?! [*hangs up*]

## BE PREPARED FOR THE BIG DATA ERA

The short story on pizza delivery is rooted in reality, perhaps a bit exaggerated at this time of our technology, but is totally feasible because the big data era advances so rapidly. This thought-provoking story exposes some potential benefits and issues of big data. There are questions that are tough to answer. When you phone a stranger, would you like that person to know every detail of your life? Would you like your whereabouts being tracked in real time over your life? What if your personal information stored in big data has been modified? Meal delivery based on customers' food preferences may improve customer service. Linking eating habit information with medical data may improve a person's lifestyle behaviour for a longer and better life. But when life choices are dictated by automated computer algorithms, life becomes mechanical and loses the human touch. Financial, such as personal banking, information as part of big data can facilitate money transfer and fraud prevention but, on the other hand, can be hacked and misused.

Should big data stop at some line to give people some personal space? Who would decide where that line should be, and enforce and ensure there are consequences when the line is crossed? It is a dilemma. On the one hand, big data are felt to be doing the tracing and tracking of the society to help people, livelihood and the economy. On the other hand, big data are felt to be spying on the society. A big issue for big data is, while individuals' information is collected for big data, it is difficult if not impossible for individuals to ask big data for their own individual information, or even know the extent to which their individual information is being collected and stored in big data.

The big data era is here whether we like it or not. So keep thinking, and keep asking questions, before big data get too big and out of control.

**Acknowledgements** The author would like to thank Stephanie Prince Ware, Sylvain Desmarais, Alejandra Dubois, Melissa Baker, Laura Hilary Thompson and Anthony Coles for their suggestions on an earlier draft of the manuscript.

**Contributors** BCKC conceptualised the project, reviewed the literature, developed and perfected the small story based on a few initial ideas found on the internet, and wrote and revised various drafts of the manuscript.

<sup>1</sup>Public Health Agency of Canada, Ottawa, Canada

<sup>2</sup>University of Toronto, Toronto, Canada

<sup>3</sup>University of Ottawa, Ottawa, Canada

**Correspondence to** Bernard C K Choi, Public Health Agency of Canada, 785 Carling Avenue, Ottawa, Ontario, Canada; bernard.choi@canada.ca

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Disclaimer** The views expressed in this paper are those of the author and do not necessarily reflect those of the author's affiliations including the Public Health Agency of Canada (PHAC) and the Government of Canada.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; internally peer reviewed.



**OPEN ACCESS**

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made

indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.



**To cite** Choi BCK. *J Epidemiol Community Health* 2021;**75**:309–310.

Received 5 October 2020  
Accepted 13 October 2020  
Published Online First 28 October 2020

*J Epidemiol Community Health* 2021;**75**:309–310.  
doi:10.1136/jech-2020-215752

**ORCID iD**  
Bernard C K Choi <http://orcid.org/0000-0002-0508-355X>

**REFERENCES**

1 Daas PJH, Puts MJ, Buelens B, et al. Big data as a source for official statistics. *J Official Stat* 2015;31:249–62.

2 Dash S, Shakyawar SK, Sharma M, et al. Big data in healthcare: management, analysis and future prospects. *J Big Data* 2019;6:54.

3 Jin X, Wah BW, Cheng X, et al. Significance and challenges of big data research. *Big Data Res* 2015;2:59–64.

4 Vijayarani S, Sharmila S. Research in big data: an overview. *Inf Eng Int J* 2016;4:1–20.

5 Manyika J, Manyika J, Chui M, et al. *Big data: the next frontier for innovation, competition, and productivity*. Seoul: McKinsey Global Institute, 2011.

6 Pastorino R, De Vito C, Migliara G, et al. Benefits and challenges of big data in healthcare: an overview of the European initiatives. *Eur J Public Health* 2019;29:23–7.

7 Bates DW, Saria S, Ohno-Machado L, et al. Big data in health care: using analytics to identify and manage high-risk and high-cost patients. *Health Aff* 2014;33:1123–31.

8 Lepore J. Scientists use big data to sway elections and predict riots: welcome to the 1960s. *Nature* 2020;585:348–50.

9 Anonymous. A joke about the era of big data (posted 5 Jun 2014) (Chinese). 2014. Available [http://blog.sina.com.cn/s/blog\\_cf74b6be0101dn7x.html](http://blog.sina.com.cn/s/blog_cf74b6be0101dn7x.html) (accessed 3 Oct 2020)

10 Anonymous. Google's pizza. (Posted 18 May 2017). Available <http://www.jokesoftheday.net/joke-Google-s-pizza/2017051897> (accessed 3 Oct 2020)