

Live Demonstration of Using Big Data Sources in Artistic Practice

Steven Aishman
Royal College of Art
London, UK
saishman1@gmail.com

1. INTRODUCTION

From algorithmic financial software that is driving the perceived value of human owned assets to personal medical data which is used to make national policy decisions, Big Data is beginning to effect more of human lives than ever before. However, Big Data and algorithms are not neutral. They shape behaviour primarily through their ability to conceal or reveal and subsequently, they shape how the world is perceived. Why are certain pieces of data collected, parsed, isolated, visualised and then distributed? Who gets to make decisions about how Big Data is used?

This paper/demonstration explores the relationship between Big Data and the visual arts. While many Big Data sets are free and available for public use, the technical barrier to sourcing and then distilling Big Data sets into a useful form can be daunting to visual artists who may want to use the data for their own art. This demonstration will cover how to locate Big Data sources, parse the data, and use the data in an artistic visualisation. Specifically, the demonstration will use live astronomical data provided by NASA which will be accessed through NASA's API. The demonstration will also show how to use a flexible programming language designed for visual arts, Processing, to parse and visualise the data. The demonstration does not require any experience with either Big Data or programming.

2. BIG DATA ALGORITHMS ARE NOT NEUTRAL

The algorithms that are used to collect, analyse, parse and visualise Big Data are not neutral. They shape behaviour primarily through their ability to conceal or reveal and subsequently, they shape how the world is perceived. Facebook's new feed is an example of this. The algorithm individually decides what new information to post to users' feeds. The algorithm is changed on a weekly basis; it is never publicly revealed and it deals with such

complex volumes of Big Data that even the engineers who design it cannot accurately predict how it will function. The use of an algorithm itself may not be an issue, but the fact that Facebook's news feed has become the most important element for Facebook to sell its \$12.5 billion worth of advertising makes the use of the algorithm increasingly important. (Luckerson 2015, pp. 1–2) The news feed algorithm presents news that the algorithm believes the user will be interested in. This causes some significant ethical problems as even users who are aware that they are only receiving one type of news may not even notice that what appears to be a chorus of voices reverberating throughout a news cycle is an echo chamber as the algorithm continues to feed the user only the stories that are "liked" or comment on. The danger of this algorithmic effect was seen in 2014 when Facebook's news feed did not rank the story Michael Brown being killed by police in Ferguson, MO very high because the algorithm decided the story was not getting enough "likes" and comments. (Tufekci 2016, p. 4)

3. ARTISTS AS PARTICIPANTS IN VISUALISATION-BASED DATA DISCOVERY METHODS

Visualisation-based data discovery methods provide a host of opportunities including allowing for better analysis of data, improving decision making and improving the understanding of the relationships between pieces of data, among others. However, it is important to note that the primary users of Big Data are "businesses, public administrations numerous industrial and not-to-profit sectors, and scientific research" (Sivarajah et al. 2017). This means that Big Data is primarily deployed in a top down manner by institutions that have the resources dedicated to Big Data visualisation. These resources are dedicated to visualising data in a way that meets institutional needs.

Artists provide a unique and necessary opportunity to disrupt the hierarchy of data visualisation. The exhibition “Big Bang Data” at Somerset House (3 December 2015 – 20 March 2016) was comprised of artists and designers who used data as a basis for their visualisations. Their visualisations could construct relationships and display data in an array of forms that challenge mainstream, institutionalised methods of display. The exhibition catalogue states that Big Data can be used to create a society that is “fairer, more stable and efficient” or Big Data can be “wielded as a means of unprecedented mass surveillance and commodification” (Somerset House Trust 2016). Jake Porway from DataKind has said, “As more data become less costly and technology breaks barriers to acquisition and analysis, the opportunity to deliver actionable information for civic purposes grows. This might be termed the ‘common good’ challenge for Big Data.” (Akerkar 2014, p. 124). Artists using Big Data in their practice provide an opportunity use Big Data in ways that may not be possible if the production of Big Data visualisations is left solely to large institutions.

4. DEMYSTIFYING BIG DATA ACCESS TO ARTISTS

While large data sets may be made available for public access, there is often a perceived technical barrier. When asked what the main barriers were that hindered communities like artists from using Big Data in social capital projects, Roger Barga from Microsoft Research replied, “The greatest barrier may be social. Such projects require community awareness to bring people to take action and often champion to frame technical challenges in a way that is approachable by the community.” (Akerkar 2014, p. 125) With that spirit in mind, this demonstration seeks to de-mystify the process of accessing and visually displaying Big Data so artists will feel more comfortable in using Big Data as a potential source for their artistic practice.

5. METHODS TO LOCATE, PARSE, DISPLAY

This demonstration covers many of the opportunities for access to public Big Data sources through APIs. These include: Data.gov (the US government’s information on everything from climate to crime), US Census Bureau (information on the lives of US citizens covering population data, geographic data and education), European Union Open Data Portal (EU data), Data.gov.uk (UK data), NHS Health and Social Care Information

Centre, World Health Organisation, Facebook, Google Public data explorer, among others.

The demonstration will then use NASA’s API for meteor impacts as an example of Big Data that can be accessed for public use. The demonstration will use Processing, to parse and visualise the data.

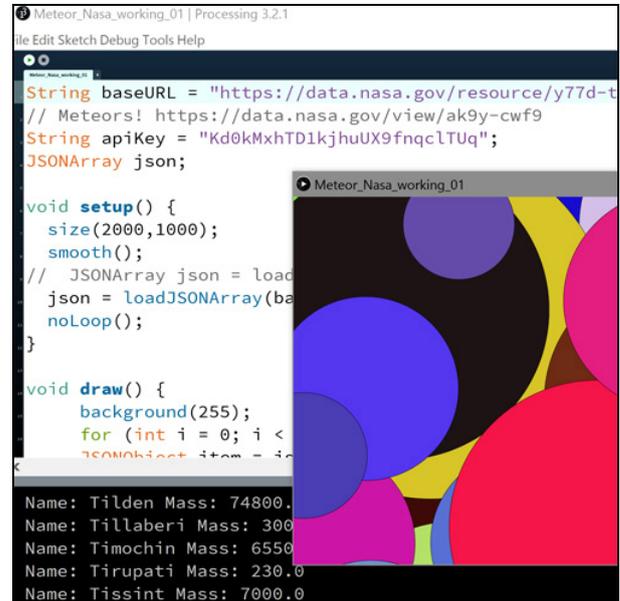


Figure 1: Image of using Processing to visualise mass of meteor impacts from NASA’s live API feed.

6. REFERENCES

- Akerkar, R. (2014) Big data computing. CRC Press, Boca Raton (FL).
- Luckerson, V. (2015) Here’s how your Facebook news feed actually works. *Time*, 9 July. <http://time.com/3950525/facebook-news-feed-algorithm/> (accessed 14 March 2017).
- Sivarajah, U., Kamal, M. M., Irani, Z., and Weerakkody, V. (2017) Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, pp. 263–286.
- Somerset House Trust (2016) *Big Bang Data* [Exhibition]. Somerset House, London, UK.
- Tufekci, Z. (2016) The real bias built in at Facebook. *The New York Times*, 19 May. <http://www.nytimes.com/2016/05/19/opinion/the-real-bias-built-in-at-facebook.html> (accessed 14 March 2017)