M21HS0303-AppliedEthics/Essay/2019115007/CHD

Abstract:
The term big data is not a fancy word anymore. Many companies, both large scale and small scale, are trying to use big data analysis as a service for their clients as well to improve themselves as a start-up or company with the tremendous amount of data generated and available these days. In the modern world, where some quintillion bytes of data is being produced every day, playing around with big data and its relational nature is dangerous for individuals and communities. There can be complex implications like loss of privacy, increase in cyber threats, scams etc.
Thus, implementing and following some rules and regulations by the individuals and organisations involved with big data analysis is essential. So researchers are trying to create an objective ethical framework for making proper use of data. In this paper, i will try to throw some light on some ethical arguments and methods concerning the debates on issues regarding Big Data analysis. Also, I will conclude this essay with the term paper question as instructed.

As mentioned in the abstract, many individuals and organisations use big data methods to perform their tasks. Big data analysis is nothing but yet another scientific invention. It is hard to marginalise one definition for scientific innovation; a scientific design can be a thesis, a machine, a method or an idea. Roughly we can marginalise the definition as a by-product of the processes in which researchers and scientists follow standard scientific procedures.

With this definition of scientific innovation, it is not hard to judge where big data stands in this academic hierarchy. Let's look at it from a different perspective, big data algorithms (just like any other idea) and the ability of modern computers to provide the platform for computational purposes (just like any other machine); all are scientific innovations themselves. Therefore all the questions concerning the morality of practices being followed in big data can be answered in the same approach as any other scientific invention. I am trying to say that there is nothing unique about data ethics; we can take a similar approach to the ethics of invention.

I argue that questions concerning the morality of any entity become challenging to answer whenever there is concurrent existence of instrumental harm and intrinsic good. (this is not the only possibility tho) This example helps to understand what I meant by the simultaneous presence of instrumental harm and intrinsic /instrumental good. Consider an action A1 that is intrinsically good that is good in itself. Action A1 being intrinsically good does not stop it from causing instrumental harm to another agent in the system (that is, it can cause harm to that agent eventually). In these situations, the agent who wishes to perform A1 finds him/herself under two or more inconsistent moral requirements, where one requirement sacrifices the other.

We can also extend this same example to data ethics. Analysing big data can be intrinsically good because it provides us with the actual truth. Nevertheless, it can also cause instrumental harm to the subjects from where the data is collected; (privacy issues exclusively). Respecting others privacy stops the analyst from getting closer to the truth. Getting closer to the truth prevents the analyst from respecting others privacy!.

Let us look at this dilemma from a different point of view. It is evident that we have to analyse as much information as possible for getting closer to the truth, and for collecting that, we must use as much data as possible. Data collected only by ethical means is limited compared to the quantity of data that can be collected by ethical and unethical standards. Whenever we use data collected only through righteous means (that is, with the owner's consent), we are not giving our best to reach closer to the truth (because we are choosing to leave the data collected from unethical means). To make my point, by not getting closer to the truth, who knows, we might be sacrificing some groundbreaking advancements in fields like medicine, biology, healthcare, finance, etc. These lost advancements might help doctors save many lives, might help financial minsters of a country to reduce their country's poverty significantly. However, suppose we include (use) the data collected from unethical means. In that case, we are encouraging the immoral means of data collection or causing harm to all the people for where the data is being collected unethically.

Usually, morally dilemmas like those described above can be solved by traditional techniques such as refuting the puzzle, using a value theory approach, or finding alternative solutions.[**1**] The names of the methods are not quite revealing. Let me explain each of them in a few words. The term 'Refuting the puzzle' can be synonymised with breaking the tie. In this approach, we will be searching for fallacies and misconceptions which might be used in the base reasoning for defining the dilemma. If we identify these fallacies, that particular moral problem is not a problem anymore.

The value theory approach inspired by utilitarian ways gives us the policy to choose the action which maximises the overall utility of the subjects involved. Depending on the problem, the sign of the utility might change, i.e., choosing greater good or lesser evil as a utility. The main drawback of this utilitarian approach in solving moral problems is that it is practically impossible to assign utilities to moral actions in ways similar to how MBA graduates assign utility values to activities available for a company [6]. In almost all cases, the involved agent count spans the population of earth! One cannot force someone to approximate the utility values. Due to this reason, sometimes it is better to be self-satisfied with an alternate possibility to escape from

this dilemma. I am going to use this sense of reasoning in my term paper to build up arguments for my final question; I felt it is worth mentioning them in this essay.

The credibility of big data analysis can also be criticised with concepts related to harm; intrinsic harm, instrumental harm, and pro tanto harm. To judge whether an event is overall harmful or not, it is essential that the pro tanto harm that single person experiences dominate the benefits experienced by the person. Yes, this sounds like a utilitarian approach, and one might argue that "this guy criticised utilitarian approaches in the previous paragraph and now using the approach to make his premises!". No bro, these might appear to be the same, but they have a slight difference in nature. In the above paragraph, we are trying to maximise the overall expected utility, and that is too hard to estimate due to the population size. But in this case, we are going to look at how one particular person is going to be affected by unethical data collection practices.

By definition, we are not supposed to care about the intent behind the action for analysing harm. If we try to reason in such as way that a consequence has more pro tanto harm than benefits, your job is done. It doesn't matter what the intentions behind the action (reaching closer to the truth) are. But we simply cannot say this action is harmful because it makes the person unhappy; or harmful because it is immoral. If we do that, we are accidentally begging the question. Instead, we are supposed to identify why something is harmful clearly. That is, we have to find one particular attribute that is common in all harms, and we should provide arguments to prove the fact that unfair data collection methods have the exact attributes. Okay, no more spoilers; the rest will be discussed in the term paper.

To conclude, my term paper question is,

- "*In the base argument of ethical dilemma on data ethics described in the essay, give your own arguments on which side you support. It is essential to use the concepts and methodology described in the essay. How are you going to extend your reasoning to support the need for an objective framework which you mentioned earlier?*"

References:

[1]
https://corporatefinanceinstitute.com/resources/knowledge/other/ethical-dilemma/

[2]
L. Floridi and M. Taddeo, "What is data ethics?," *Phil. Trans. R. Soc. A.*, vol. 374, no. 2083, p. 20160360, Dec. 2016, DOI: 10.1098/rsta.2016.0360.

[3]
A. Bilal, S. Wingreen, and R. Sharma, "Virtue Ethics as a Solution to the Privacy Paradox and Trust in Emerging Technologies," Apr. 2020. DOI: 10.1145/3388176.3388196.

[4]
S. Leonelli, "Locating ethics in data science: responsibility and accountability in global and distributed knowledge production systems," *Phil. Trans. R. Soc. A.*, vol. 374, no. 2083, p. 20160122, Dec. 2016, DOI: 10.1098/rsta.2016.0122.

[5]
"Big Data- Can Virtue Ethics Play a Role unimelb.pdf." Accessed: Oct. 19th, 2021. [Online].Available:https://minerva-access.unimelb.edu.au/bitstream/handle/11343/227494/BigData-%20Can%20Virtue%20Ethics%20Play%20a%20Role%20unimelb.pdf?sequence=1&isAllowed=y

[6]
"Statistical Methods in Artificial Intelligence lecture transcript"

[7]
http://blog.practicalethics.ox.ac.uk/2020/03/oxford-uehiro-prize-in-practical-ethics-can-science-ethically-make-use-of-data-which-was-gathered-by-unethical-means/