Collaborative Discussion 1

Initial Post

Initial Post

by Sahr Solar Sumana - Sunday, 12 November 2023, 2:17 PM

We are constantly reminded of data becoming one of the most valuable commodities in the modern age, "data has become an essential commodity with the power to influence economic growth and our social prosperity" (James, 2023). Technology has now developed leading to the internet of things (IOT) where data can now be gathered from items, we use everyday "including your PC, mobile phone, smart watch, and anything else that connects to the Internet and sends or receives data" (Tejada, 2023).

There are limitations to collecting data via IOT. The sheer size of the data collected can lead to the tasking challenge of cleaning all of the data attained in order to understand the contents of the data gathered. Not only would this cost a lot of money but it would cost a lot of time and energy too. On the other hand, once data sourced through IOT has been cleaned an opportunity to understand your data will then present itself, Wickham (2014) stated "a structured format of data will enable analysis; it's easier to manipulate, model and visualise data when tidy", thus appreciating the value of your data.

Extreme values and inaccurate data appearing from faulty equipment is another limitation to data from IOT, which can lead to incorrect hypothesis being generated because of these values. Although "preventing errors within data is cheaper than fixing them" (Huxley, 2020). Preventing these limitations is also tough since devices linked to the internet are not impossible to stop from failing which can produce bad data.

In summary data collection via IOT is highly advantageous because it provides information that cannot otherwise be gathered as quickly and accurately through other methods such as surveys however, it is important to covert data to clean data and prevent errors to maximise the value of this data.

Reference List:

Huxley, K. (2020) 'Data Cleaning', *SAGE Research Methods Foundations* [Preprint]. doi:10.4135/9781526421036842861.

James, F. (2023) *Future-proofing our most valuable commodity, Office for National Statistics.* Available at: https://blog.ons.gov.uk/2023/08/18/future-proofing-our-most-valuable-commodity/ (Accessed: 12 November 2023).

Tejada, Z. (2023) *Big Data Architectures - Azure Architecture Center, Azure Architecture Center | Microsoft Learn.* Available at: https://learn.microsoft.com/en-us/azure/architecture/databases/guide/big-data-architectures (Accessed: 12 November 2023).

Wickham, H. (2014) 'Tidy Data', *Journal of Statistical Software,* 59(10). doi:10.18637/jss.v059.i10.

Response to Peer #1

Initial Post

by Bahar Yatman - Wednesday, 22 November 2023, 9:40 PM

The Internet of Things (IoT) is a network of interconnected devices that communicates and exchanges data with other IoT devices as well as the cloud. Consumer goods, digital and mechanical machinery, or any combination of these, are examples of IoT devices (Gillis, 2023). The Internet of Things has altered many aspects of our lives and businesses; automation, efficiency, cost-saving measures, and improved decision-making are the main advantages of this innovation. However, the way that data is collected has changed significantly as a result of the massive amounts of data generated by IoT devices. As a result, IoT also has disadvantages like any other technological breakthrough, such as complexity, interoperability, data privacy, and security (Prajapati, 2023). Thus, in order to address these problems and manage the varied and dynamic nature of IoT data, it is essential to have an appropriate processing plan and cutting-edge analytical tools. Additionally, IoT data requires an extensive cleaning process to deal with issues like extreme outliers, missing values, and data collection errors (Huxley, 2020).

All in all, IoT integration with big data has the potential to transform a number of industries, such as agriculture, finance, transportation and logistics, retail, manufacturing, and healthcare (Peranzo, 2023), but it also presents a number of risks and difficulties. As previously stated, data cleaning is a complicated process, so in order to make reliable data-driven decisions, researchers must employ transparent and robust methods.

References:

Gillis, A.S. (2023) *What is IOT (internet of things) and how does it work?: Definition from TechTarget, IoT Agenda.* Available at: https://www.techtarget.com/iotagenda/definition/Internet-of-Things-IoT [Accessed: 14 November 2023].

Huxley, K. (2020) Data Cleaning, Sage Research Methods. Available at: https://methods.sagepub.com/foundations/data-cleaning [Accessed: 22 November 2023].

Peranzo, P. (2023) *8 sectors that can benefit the most from IOT development, Imaginovation.* Available at: https://imaginovation.net/blog/8-sectors-benefit-from-iot-development-in-2021/ [Accessed: 22 November 2023].

Prajapati, G. (2023) *Pros and cons of internet of things (IOT), DRC Systems.* Available at: https://www.drcsystems.com/blogs/pros-and-cons-of-internet-of-things-iot/ [Accessed: 14 November 2023].

Peer Response

by Sahr Solar Sumana - Tuesday, 13 February 2024, 5:05 PM

In this initial post it was stated that the IOT has improved and altered many aspects of our everyday lives by allowing processes to become more efficient, and by also implementing automation within both personal and business based processes. Due to the significant volume of data collected by sensors and devices from the IOT data is not standardised as it normally would be, Huxley (2020) argued this point by highlighting the limitation of "data that has not been collected for research purposes, which can be messy and unstructured". This can make the cleaning, normalisation and processing of the collected data extremely difficult and time consuming. Although the IOT provides the benefit of fast and accurate data it could be argued that there is a factor of too much data, how much data is too much data? "Too much data results in too much noise and compromises the performance, profitability and security of any enterprise" (Eizenberg, 2023). To build upon this and the initial post's point of data cleaning being a complicated process, there needs to be a rule of measure which can determine how much data is too much when running data cleaning activities, and that can be difficult when the IOT can provide such a large amount of data.

Reference List:

Eizenberg, E. (2023) Council post: Why, yes, there is such a thing as too much data (and why you should care), Forbes. Available at: https://www.forbes.com/sites/forbestechcouncil/2022/01/03/why-yes-there-is-such-a-thing-as-too-much-data-and-why-you-should-care/ (Accessed: 13 February 2024).

Huxley, K. (2020) Data Cleaning, Sage Research Methods. Available at: https://methods.sagepub.com/foundations/data-cleaning (Accessed: 13 February 2024).

Maximum rating: -   Permalink   Show parent   Reply

Response to Peer #2

Initial Post

by Simbiso Makunde - Thursday, 23 November 2023, 6:44 PM

Internet of Things (IoT) is a term which refers to devices and gadgets having embedded sensors which are connected to the internet in a Wireless Sensor Network (WSN). This allows for real-time data collection from systems such as smart cars, smart watches and home systems to mention a few.

According to Rahumani et al (2021), IoT gives us the ability to monitor our environment at will .However the Big Data requires cloud services on account of their large-capacity storage platforms, analysis software as well as web services which serve as user interfaces for user requests.

The concept of IoT is essential for optimising energy efficiency, production quality and costs in order to improve customer satisfaction and business decision making. IoT is at the core of modern information systems for Education, Industries, Healthcare, and Retail fields.

Despite all of the opportunities associated with the large-scale IoT process of data collection, there are also some challenges in the data collection and management processes. One of which is that there are data privacy and security risks from having all the devices connected together in a WSN.

In addition, conformity and interoperability present a significant challenge with different manufacturers producing the devices and systems .in order to ensure a productive WSN, they should be able to communicate smoothly and comply with industry regulations (Scalerandi, 2023)

References:

Rahmani, A.M., Bayramov, S. & Kiani Kalejahi, B. (2022) Internet of Things Applications: Opportunities and Threats. *Wireless Person Communication* 122: 451–476. DOI: **https://doi.org/10.1007/s11277-021-08907-0**

Scalerandi,D. (2023) Leveraging IoT For Business Growth: Opportunities And Challenges Available from: **https://www.forbes.com/sites/forbesbusinesscouncil/2023/10/13/leveraging-iot-for-business-growth-opportunities-and-challenges/amp/** [Accessed 16 November 2023].

Peer Response

by Sahr Solar Sumana - Monday, 12 February 2024, 5:10 PM

It is understood that the concept and integration of the IOT has a multitude of benefits as listed by Simbiso including being able to automate the collection of big data, and have that data collected and processed in real time.

Although the limitation of cloud services due to large capacity storage requirements has been listed, there is an outstanding limitation of data cleaning and data integrity requirements. Hellerstein (2008) has stated that "sensor deployment often affects data quality, and many sensors are subject to errors including miscalibration and interference from unintended signals furthermore, any procedure that integrates data from multiple sources can lead to errors". To extend upon this statement a major downside is that data from the IOT can be susceptible to erroneous entries and with the scale at which data is collected a huge number of erroneous entries can be recorded and not be picked up upon until later stages of the data lifecycle such as during the cleaning or analytical phases.

I agree with Simbiso's point of devices and systems being able to communicate with one another to ensure a productive WSN. Not making this a priority can lead to missing data or erroneous entries as mentioned before, a WSN where devices are not in sync with one another can lead to extreme values where the distribution of data is not normal. With the IOT it may be harder to screen data due to the size of the data and how time consuming it can be Huxley (2020) has stated that "data screening process is often a neglected stage of data preparation".

Reference List:

Hellerstein, J. (2008) *Quantitative Data Cleaning for Large Databases* [Preprint].

Huxley (2020) 'Data Cleaning', *SAGE Research Methods Foundations* [Preprint]. doi:10.4135/9781526421036842861.