

Title

The cost of data: Environmental impact of health care data storage.

Problem statement

Already in 2018, data centers were estimated to take up 200 terawatt hour of energy, accounting for 1% of the worlds' energy use which translates to 0.3% of the worlds CO2 emissions. The energy consumption is projected to increase dramatically in the coming decade, up to 2% of the worlds' energy use by 2030.

Health care is becoming increasingly data-supported. The reasons are obvious when looking at all the advantages of digital information. With digital patient records, information is more easily shared amongst health care professionals and with the patients themselves. The use of remote sensors allows for precise patient monitoring without hospital visits. And data based algorithms hold the promise of improved personalized therapeutic monitoring, disease prevention and early diagnosis of disease.

More is better, seems to be the motto when digital information is concerned. However, there have been serious concerns about the negative side effects. Most discussed are ethical reasons such as privacy and biased algorithms, however the environmental impact is largely ignored. With the expected increase in data-storage and climate crisis in mind, this issue needs to be addressed.

In this research project, we would like to tackle this knowledge gap in order to make well-informed decisions about how much data storage is actually enough and what could be done to mitigate possible negative environmental impacts. We will do this by taking a hospital data storage as a case study.

Research question(s)

How can the environmental impact of a single hospital data storage be mitigated?

Possible sub questions include:

1. How much data is currently being stored and what is projected in the near future?
2. How is data stored now, and what is the projection for the near future?
3. What is the environmental footprint of data, depending on storage type
4. What alternatives are there to data storage?
5. How much data storage is necessary? (think about health care quality, regulations but also how much data do algorithms actually need to function and how much is redundant)



Expected type of work

This research subject includes: This is a project allowing for much student creativity in how to address the research question. Possible work includes database analysis, surveys, interviews with health practitioners, regulators, data information officers, computer scientists and industrial ecologists, as well as data storage companies and other possible third parties. Visual mapping of data flows. Calculations of environmental impacts (including but not limited to CO₂ emissions).

Remarks

We encourage motivated, proactive and independently working students with any study background to apply and co-create the research design as part of the project.

References

- Lucivero F. Big Data, Big Waste? A Reflection on the Environmental Sustainability of Big Data Initiatives Science and Engineering Ethics volume 26, pages1009–1030 (2020) <https://link.springer.com/article/10.1007/s11948-019-00171-7>
- Koomey JG, 2008 Worldwide electricity used in data centers Environ. Res. Lett. 3
- Koot M, Wijnhoven F, Usage impact on data center electricity needs: A system dynamic forecasting model Applied Energy Volume 291, 1 June 2021
- Jones N, How to stop data centres from gobbling up the world's electricity The energy-efficiency drive at the information factories that serve us Facebook, Google and Bitcoin. Nature. 12 September 2018
- Tahir A et al, A Systematic Review on Cloud Storage Mechanisms Concerning e-Healthcare Systems Sensors. 2020 Sep; 20(18): 5392. <https://doi.org/10.3390/s20185392>
- International Energy Agency website: <https://www.iea.org/reports/data-centres-and-data-transmission-networks>

Commissioner details

Organization / Department: Medical Delta, LUMC

Name: Rosa van den Berg, Frank Willem Jansen

Email: rosa.vandenberg@medicaldelta.nl; f.w.jansen@lumc.nl