

Letter to the Editor

Big data in third world countries: Do the means justify the end?

Vijayant Govinda Gupta, Girdhar Singh Bora and Ravimohan S. Mavuduru*

Department of Urology, Nehru Hospital, Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh, India

Received 4 March 2016

Accepted 20 March 2016

Abstract. The amount of data that is being generated in health care is enormous, especially with the widespread availability of multimedia devices. However, the data keeping in developing countries is still unstructured. Lots of this Big Data is therefore vulnerable to loss. In the current report we highlight the problem and suggest the need for cost effective solution.

Keywords: Big data, data storage, data retrieval

“In god we trust, all others (must) bring data”.

With these words, W. Deming, an American statistician, exhorted the global industry to move into the new age of “Big Data”. Big data refers to data so large and voluminous that traditional data processing applications are inadequate to analyze them. This data when analyzed or mined can reveal insights and new knowledge which is especially relevant in the field of healthcare. This allows for better utilization of resources, increasing competitiveness and profitability along with betterment of care for the patients. Data in healthcare comes in many forms. The majority of healthcare data generated today is unstructured, e.g. We generate about approximately 10 Gigabyte (GB) of structured data per day. However there is a higher volume of data that is stored unstructured, collected at an individual level (data like laparoscopic/open surgical videos, radiological images, clinical photographs etc.). Partly this is due to the wide availability of multimedia devices and lack of facility for a central storage of the data-base. Even for this small structured data we are facing problems to store and ways to retrieve it as and when required. If we wish to store all the data (both structured and unstructured) the volume will be huge (approximately 20 terabytes a year) and problems of storage and retrieval will be compounded. This is an example of a medium sized (100–150 beds) healthcare facility in a resource constrained setting in a developing country like ours (1). As a solution to this problem, we have started to save data in optical disk drives and flash drives to store it individually. For larger data we rely on HDD and SSD on multiple individual computer terminals.

*Corresponding author: Ravimohan S. Mavuduru, Department of Urology, Nehru Hospital, PGIMER Chandigarh, India. Tel.: +91 9417532955, +91 7087009360; E-mail: ravismi2003@yahoo.com.

There is very limited information available on data storage options for small to medium healthcare facilities across the world. An ideal solution to our needs then would be a hybrid structure with three levels. Level one would be individual drives on terminals which would be backed up on a locally installed Network associate RAID system of 100 TB capacity (Level two). Level three would be a cloud storage solution offering a capacity of 10 TB per month, where long term data can be backed up.

The estimated cost for such a system would be 20000 USD (approx. 1.4 million INR) of fixed costs and 1000 USD per month (approx. 70000 INR) of maintenance costs. The question then is, does it make economic sense for developing data storage infrastructure? Does it provide tangible, measurable benefit that can be directly observed on hospital bottom lines? The science behind big data in health care is robust, the evidence to prove that data analytics is improving health care is substantial. But, for healthcare facilities in resource constrained settings, do the needs justify the ends? This is the perspective that needs further research.

References

- [1] Hess I, Srikantan KS. Recommended Variables for the Multiple Stratification of General Hospitals. *Health Services Research*. 1970; 5(1): 12-24.
- [2] Miller JD. Sharing clinical research data in the United States under the health insurance portability and accountability act and the privacy rule. *Trials*. 2010; 11: 112. doi: 10.1186/1745-6215-11-112.