ISSN: 0975-833X

Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 5, Issue, 11, pp. 3529-3530, November, 2013

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

REVIEW ARTICLE

TACKLING DATA PROLIFERATION

***Dr. A. Vijayakumar and K. Muraleedharan**

Karpagam University, Coimbatore

ARTICLE INFO

Article History:
Received 19th September, 2013
Received in revised form
29th September, 2013
Accepted 04th October, 2013
Published online 19th November, 2013

Key words:
Information Explosion,
Raw data,
Storage of data,
Data filtration,
Information Lifecycle.

ABSTRACT

It is often termed that the present day society as an 'information society', where information is being generated every now and then. This new addition of information is ultimately leads to the modification of the existing knowledge or the creation of new knowledge which definitely has to pass on to the needy. The 'knowledge boom' or the' information explosion' has paused a situation, where a genuine information seeker feels some sort of "poverty" in the midst of plenty, where the information is abundant, and how it could be make best use of it? And here the real problem arises-how far the accumulated information is accurate and apt. Hence a mere collection of information is not enough for utility, but the 'raw' information is to be processed, to yield systematic and integrated information needed for a specific purpose. The fact that fruit ful education is in process in every sect of schools where students are undergoing apt and up-to-date education. The dynamic nature of modern education envisages right and accurate information to the student at the right direction.

Copyright © Dr. A. Vijayakumar and K. Muraleedharan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

'Data proliferation' is a term coined with excess outbreak of information. The pouring in of data structured as well as unstructured from all sources, right from the very generation onwards has resulted in the excess quantum of raw data. In most cases these 'raw data' has resulted in the abundance of data which has of little usage for research and development activities. The amount of usability of these data determines the strength of the research and development being taken place in a country. The Research and Development activities are truly remarkable in the advent of ICT boom. It happens, so quickly, taken such commanding control of the benefits of the communication technology and assumes audacious responsibility for the research endeavors. Yet for all its ambitious and aspiring ways, it has not, in essence, moved so very far from our ancestry. While capable of incredible sophistication of thoughts and discovery, still the barriers or certain obstacles that prevent the Researchers from accumulating information from all corners and still crave certainty and simplicity and shy away from complexity and ambiguity. And still tend to default, whenever it can, to black - and-white, "either or" logics that allow to discern neat patterns and constant comfortable worldviews. Data handling has been a matter of concern, right from the beginning of the ICT boom. For example The US military has documented it as a problem since August 1971. Considering the excessive documentation submitted during the acquisition of strategic armory systems. Still the problem persists and efforts are being made to minimize the problems associated with it (Data Proliferation). It is evident that now the storage, retrieval and dissemination of information is more prone to automated modulation and being taken place in automated surroundings.

**Corresponding author: Dr. A. Vijayakumar*
Karpagam University, Coimbatore

Data handling

With the sprung up of digital storage devices, storage of data has been easier when compared to paper documentation. Thus there has been a general feeling among the user community that digital storage and retrieval of information has of much more accurate and speedy, where it provide some sort of accuracy in information management. But in reality it is myth that abundance of data or information will not cope up with the actual information need.

Problems

The problem of data proliferation is pausing serious complexities on all aspects of business endeavors. This is due to the emergence of comparatively of less expensive data storage techniques. Instead of being made a judicious analysis and application of the data, more often they are being 'dumped' in to secondary storage devices without having a proper 'filtration'. This may pauses problems that could have an adverse effect on the effective functioning of Health Services, Police and Security forces, local and National Governments and other organizations.

Reasons

- Information retrieval is becoming a tedious affair. At Xerox, on an average it takes employees more than one hour per week to find hard-copy documents, costing $2,152

a year to manage and store them. For businesses with more than 10 employees, this increases to almost two hours per week at $5,760 per year. In large networks of primary and secondary data storage, problems finding in electronic data are analogous to problems finding hard copy data.

- Data loss and legal liability when data is disorganized, not properly replicated, or cannot be found in a timely manner. In April 2005, Ameritrade Holding Corporation told 200,000 current and past customers that a tape containing confidential information had been lost or destroyed in transit. In May of the same year, Time Warner Incorporated reported that 40 tapes containing personal data on 600,000 current and former employees had been lost en route to a storage facility. In March 2005, a Florida judge hearing a $2.7 billion lawsuit against Morgan Stanley issued an "adverse inference order" against the company for "willful and gross abuse of its discovery obligations." The judge cited Morgan Stanley for repeatedly finding misplaced tapes of e-mail messages long after the company had claimed that it had turned overall such tapes to the court.
- Although, the data storage devises are becoming more prone to user friendly, and simple to operate, still it requires skilled man power requirements to make optimum utilization and for effective management of the system. Though the mechanism ensures simplicity and accuracy in operation, it requires timely and updated hardware operational mechanism.
- Often the networks links are finding inadequate to meet the requirements of the excess traffic due to excessive use by the users to retrieve information. This has pauses a tacit situation in information retrieval world over.
- The hardware maintenance cost in terms of electricity and alternative power solutions may be marginally increasing progressively. For example a 100 terabyte system may have cost around $35,040 a year to run, besides cooling cost. This may have resulted in budgetary provisions.

Solutions

- Better utilization and application of modern technology.
- Data 'filtration' and reduction in duplicate data.
- Improvement of metadata structures.
- Improvement files and storage transfer structures.
- 'Proper' user 'education' programmes.
- Proper running mechanism of information life cycle management solutions to eliminate low – potential information to save time and resources to enhance the efficiency of knowledge management.

Procedures

- The proper maintenance of backup
- Proper Digital Asset Management
- Ensure a proper Disk storage mechanism
- Ensure a scientific Document management system
- Ensure Hierarchical storage management modules
- Attention to Information Lifecycle Management
- Maintenance of Information repository
- Maintenance of Institutional repositories
- Maintenance of Magnetic tape data storage

Conclusion

It is obvious that the quantum of knowledge is spontaneously increasing in every branches of knowledge, which ultimately leads to the growth of new knowledge or the modification of the existing knowledge that eventually led to a situation where information handling is becoming more and more complex. This complexity is worsening with 'data proliferation' which often mislead the users from their intellectual pursuits. Accurate corrective measures will certainly minimize the adverse implications of data proliferation in its least optimum level. Modern ICT techniques and computing peripherals enable the information professionals to tackle the situation.

REFERENCES

1. /\"Downsizing the digital attic" *(http://web.archive.orgl web/20* 11 *0722194032Ihttp:// www.deloitte.co.uk dddd* ITMTPredictions/technology Downsizing-digital-attic-data-storage.cfin). *Deloitte Technology Predictions.* Archived from the original *(*http://www.deloitte.co.uk lTMTPredictions/technologyl *Downsizing-d*igital-attic-data-storage.cfin) on July 22, 2011.
2. /*abed* "The Toxic *(http://web.archive.orglweb/ 2009020 6010415/http://www-3.ibm.co*mlsystems Iresources/ systems_storage_solutions~dCtoxic_tb.pdf) Terabyte", IBM Global Technology Services, July 2006
3. /*ab* Evolution of the Data Proliferation Problem within Major Air Force Acquisition Programs. (http:// stinet.dtic. milloai/oai?&verb=getRecord&metadataPrefix=html&identifier=AD0892652)
4. /\ Data Proliferation: Stop That (http://www. thic.orglpdf lJun02/dod.rroderique.020612.pdf)
5. /\ "Dealing with data proliferation"; Vawn Himmelsbach. it business.ca: Canadian Technology News, September 19, 2006 (http://www.itbusiness.caliticlientlenihomelNews. asp?id=40615&cid= 13)
6. /\ "Data: Lost, Stolen or Strayed", Computer World, Security (http://www.computerwcrld.com/printthis *1200510,4814,103541,00.html)*

****