



# ICIS2022Penang

<https://fim.uitm.edu.my/index.php/research/conference/342-icis-2022>

## 5th International Conference on Information Science

Royale Chulan, Penang, Malaysia, 19-21 Sep 2022

Organised by Faculty of Information Management, UiTM, Malaysia



## Records Management Practices: A solution in dealing with big data

Nor Sakila Asnawi<sup>1</sup>, Irwan Kamaruddin Abd Kadir<sup>2</sup>, Azmi Ab Rahman<sup>3</sup>

Alwi Mohd Yunus<sup>4</sup>

<sup>1</sup> ERS Energy Sdn Bhd, Kuala Lumpur, Malaysia.

<sup>2</sup> Faculty of Information Management, Universiti Teknologi MARA, Puncak Perdana, Selangor Malaysia.

[norsakila@ers.my](mailto:norsakila@ers.my), [irwan@uitm.edu.my](mailto:irwan@uitm.edu.my), [ar\\_azmi@uitm.edu.my](mailto:ar_azmi@uitm.edu.my), [alwiyunus@uitm.edu.my](mailto:alwiyunus@uitm.edu.my)  
Tel: +6019-2869249

---

### Abstract

Big data in the Internet of Things (IoT) led to various issues and problems. Therefore, this study aims to provide a guideline through Records Management practices. This study is qualitative with the records professional who deals with big data and records management from various fields. Organizations involved in emerging big data will be chosen as respondents. The study intended to develop guidelines from the current Records Management standard, best practices and guidelines in managing big data. This will offer new research and view on the ability of Records Management as a solution in managing big data.

Keywords: Big Data, Records Management, Internet of Things

eISSN: 2398-4287 © 2022. The Authors. Published for AMER ABRA cE-Bs by E-International Publishing House, Ltd., UK. This is an open-access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). Peer-review under responsibility of AMER (Association of Malaysian Environment-Behavior Researchers), ABRA (Association of Behavioral Researchers on Asians), and cE-Bs (Centre for Environment-Behavior Studies), Faculty of Architecture, Planning & Surveying, Universiti Teknologi MARA, Malaysia.  
DOI: <https://doi.org/10.21834/ebpj.v7i1S110.4124>

---

### 1.0 Introduction

Big data keep increasing from time to time. Many data are created today on various platforms, especially the Internet of Things (IoT). The increase of big data was given a significant impact and challenges to organizations. Big data analytics seems like the best way to manage big data. However, the question is, why the big data keep increasing? This is because the creation of big data analytics that more support big data led to a massive increase in big data. Through that perspective, this paper tries to find a new solution to managing big data rather than supporting the increase. The increase of big data today was led to many issues. Each organization faces problems in managing big data. Due to that reason, this paper would like to see either Records Management practices can manage big data or not. The writers believe that records management standards can help control emerging issues in big data.

With increasing information that leads to the increase of big data, many organizations face difficulties in capturing opportunities in big data. Moreover, the increase in big data also affected the organization's performance, where organizations had difficulty capturing valuable information among a massive amount of data. Information created in the Internet of Things (IoT) today is also one of the major causes of the increasing big data. Much information created today from various formats, either structured, unstructured, and many more, led to challenges for the organizations. Many big data analytics was created as a solution for big data. However, it is more on supporting big data, such as storage to keep and handle big data. This is not enough to manage and control big data. A new solution must be created to control the increasing big data and manage issues related to big data. A good solution in big data not only can support the increase of big

eISSN: 2398-4287 © 2022. The Authors. Published for AMER ABRA cE-Bs by E-International Publishing House, Ltd., UK. This is an open-access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). Peer-review under responsibility of AMER (Association of Malaysian Environment-Behavior Researchers), ABRA (Association of Behavioral Researchers on Asians), and cE-Bs (Centre for Environment-Behavior Studies), Faculty of Architecture, Planning & Surveying, Universiti Teknologi MARA, Malaysia.

data but also must be able to manage and control it. Hence, Records Management is crucial today in managing information. Therefore, the records management standard, ISO 15489-1:2016 (E), was chosen to see whether elements on it can manage the big data. This study aims to provide a guideline that can help manage significant data issues through Records Management practices.

## 2.0 The Increase of Big Data In Malaysia through the Internet of Things

Internet of Things (IoT) is one of the big sources that lead to the increase of big data. Much of the information was created daily, every second, and every hour. This is also happening in Malaysia. According to DataReportal (2019), Malaysia has the highest number of internet users based on the population that surfs the internet for daily usage. Around 76% of Malaysian are internet users from 32.26 million population. They use the internet frequently, have mobile phone connections, and are active in social media. Out of that population, 83% or 26 million are several users that are active in social media; 83% or 26.69 million people are internet users, and around 127% or 40.69 million people compare to the current population that has and uses mobile phone connection to perform their daily task.

The increase of big data through IoT needs to be taken into consideration. This is because it can lead to an increase in big data. This has been proven by Yasmeen Shahdad (2018), who mentioned that the advancement of technologies through the internet and network resources due to the IoT could increase the amount of data we call big data. The results of increasing big data through IoT can cause another issue. Ang & Loh (2016) discussed this problem seriously and identified that the internet would not be a censor. They are considering laws to sieve "other nonsense and pornography" on the internet. He also added that they need to observe other countries implement new internet.

## 3.0 Issues In Big Data

The increase of big data through IoT led to the failure to capture opportunities and insight. Joashi (2018) also believes that big data can give challenges in capturing opportunities and insight to the organization if the big data is not adequately managed and controlled. Moreover, Joashi (2018) stated that the problem occurs with the increase of big data the organizations when it is challenging to capture valuable information. This is because valuable information is vital to capture opportunities and extract actionable data. However, the increase in big data was challenging for an organization to identify valuable information. Joashi (2018) added that the increase in big data also leads to difficulties for an organization to allocate resources. This is supported by Somani & Deka (2017), who believe the increase in big data led to difficulties in capturing, processing, storing, and analyzing all the data. This can happen in massive amounts of data, such as medical, financial, environmental, and many more. Kadadi et. al. (2014) also believe that this issue can become more severe if the organization finds it challenging to manage data due to the many data that must be handled simultaneously.

The increase of big data can also give difficulties to specific users in understanding the data when too much data is created daily. This can be seen starting from the sourced background of data until it can be analyzed. (Unhelkar, 2017). He also believes that this problem can give difficulties in producing business value when users lack an understanding of the results in fragmented and lack localized use of analytics in big data. This has been supported by Almeida (2017) that these issues can be the worst challenges in big data when too many companies are still struggling with this new world of information.

The big data issue can be seen when there is an issue related to storing the data. This issue has been proven when Somani & Deka (2017), I.B.M. stated that every day in our daily life, there are about 2.5 quintillion bytes of data, which will increase to 50,000 gigabytes. This statistic was recorded in 2018, and now in 2020, we can imagine how much data was created. The increasing data led to the issue of storing data that needs huge storage. Moreover, some organizations prefer traditional storage rather than cloud storage. (Somani & Deka, 2017).

The data increase also happens due to the velocity of the data in IoT (Internet of Things). Garyaev & Garyaeva (2019) believe that the increase of big data today is due to the information technology where data created, either structured or unstructured, can lead to the failure of capturing insight and opportunities by organizations. Moreover, many companies today also deal with petabytes of data. This can be seen when Google process data every day up to 30 petabytes, and Facebook receives more than 10 million photos every day. (Garyaev & Garyaeva, 2019). Somani & Deka (2017) also agree that the speed of data today can lead to a significant loss when some organizations have difficulties analyzing and processing the data. Moreover, Garyaev & Garyaeva (2019) also mentioned that by 2020 the total of data would keep increasing up to 44 zettabytes. This is a serious issue that needs to be tackled.

Besides, Mohammed et. al. (2016) also believes that IoT, more often unstructured and semi-structured, can bring significant challenges to the world. This is because He believes that the size of data can increase up to one trillion by 2030 due to different types of data generated in IoT.

These serious issues need to be a worry when Malaysia was known as the higher user on the internet in Asia (DataReportal,2019). New Straits Times (2019) also posted that Malaysia was ranked as the highest user in Southeast Asia and the top five on the global stage. Singapore and the Philippines followed this. There are various amounts of data across all platforms. Moreover, the latest DataReportal (2019) also stated that Malaysian people spend their daily life with an average of eight hours online (5 minutes) on the internet. This shows that IoT is important but, at the same time, can have a significant impact on human lives. (Kundhavai & Sridevi, 2016). Kadadi et. al. (2014) also believe that data in IoT can lead to productivity, workloads, and a workflow that is difficult to enhance. This is because the Kadadi et. al. (2014) data has functioned in drawing and analyzing the conclusion for the estimation of the future and growth of projects. This was agreed by Joashi (2018) that believes when an organization fails to extract value in big data, it will lead to failure in identifying the right objectives for the organization. In the end, it will have an impact on the growth of the organization.

Organizations also often fail to manage their data when there is a lot of data derived from various sources (variety of data). (Kadadi, 2015). The Insights on governance risk and compliance (2014) added that big data includes information that is collected in an internet-enabled device, video, voice recordings, and social media, that continues structured or unstructured data. According to Su (2018), unstructured data is more challenging to control than structured data. This is because structured data is information that fits nicely in relational databases compared to unstructured data. Su (2018) also added that unstructured data is challenging to handle due to the various formats of data and the bites and bytes, meaning that not easy to define. In addition Nik Azliza, Alwi and Irwan (2021) stated that as the data come from numerous and variety of sources, it makes it more complex and different.

Other than that, there is also an issue with data quality. Somani & Deka (2017) believes that the increase in data can affect the data quality related to data veracity. This is because, Demchenko et al. (2013) believe that data quality or veracity is needed in defining the integrity, trustworthiness, and authenticity of data. Therefore, when the data is kept increasing, this will lead to difficulty in achieving data accuracy. Moreover, Somani & Deka (2017), mentioned that the issues become serious when some organizations do not have "trust" in information to make decisions. Some organization is more focused on technical aspect like storage of data rather than maintaining the quality of data. (Somani & Deka, 2017). Somani believes that organizations should focus on maintaining the data quality first rather than supporting the emergence of big data by having large storage to keep the data.

Besides, the increase in big data also can lead to data security issues. Oussous et. al. (2018) stated that the world today is facing 1.8 zettabytes of information, and by 2020 it will generate 50 times that amount. A zettabyte here is equal to 1000 exabytes. Therefore, due to those reasons, many issues related to security and privacy emerged. Moreover, Kundhavai & Sridevi (2016) also mentioned that there are data security issues on the internet that traditional security methods cannot control. This is also supported by Heba Aly et al. (2015) believe user information is no longer private due to the big data that kept increasing on the internet

#### 4.0 Records Management In Managing Big Data

To overcome issues in big data due to the emergence of data on the internet, it is crucial to have proper management in managing big data. Almeida (2017) believes that to develop a solution for managing big data, it is vital to consider a few key points, including the governance of data, quality of information, manipulation, and integration. Moreover, Wulff & Wunck (2016) also stated that challenges today in big data are not only at a technical level but also at the data level itself. Therefore, he believes that having appropriate information or data processing can be an excellent solution to the efficiency and effectiveness of data integration.

Therefore, in this study, Records Management was chosen to see whether Records Management Standards can manage big data. Based on the problem of failure in capturing opportunities and insight due to the increase of big data in IoT, this paper would like to see how Records Management can manage big data. This is because Kundhavai & Sridevi (2016) mentioned that the solution of big data analytics today, such as SQL in Hadoop, Streaming Big Data Analytics, Hive, and many more in more on support the increase of big data and improvement of IoT rather than control and manage the emerging of big data. This showed that big data will keep increasing and give more challenges in the future if the big data is not appropriately managed and controlled.

To manage the big data, the Records Management standard, ISO 15489-1:2016(E), was chosen to see whether it is suitable for managing big data. The reason for choosing this standard is because this standard is a necessary standard that is well-known and recognized worldwide as an excellent baseline established in records management. (Pember., M. 2006). The standard also helps provide a blueprint for establishing, monitoring, structuring, and auditing best practices in records management programs. Besides, this standard also allows any organization in this world to be more efficient and effective in the retrieval of information. This includes decision-making, productivity, accountability, and reducing information risk.

Records Management, ISO 15489, applied in various records formats created from various media, either involving public data or individual data. Xiaomi & Jiao (2004) also believes this standard is significant since it is the first international standard related to Records management. They stated this standard is also needed and important in setting the parameter within the records management program.

Due to the strength of this ISO 15489 standard. The latest edition, ISO 15489-1:2016(E), was selected to manage and control big data. The previous ISO 15489 is AS/ISO 15489-1:2001 Information and documentation - Records Management Part 1 and 2, which is a standard that sets out the principles, processes, and practices associated with records management and sets out accepted good practices within the industry.

**ISO 15489-1:2016(E)** is the latest standard that is also part of International Standards, and Technical Reports that capture and manage records. (ISO 15489, 2016). However, based on Convery, N (2016), the latest version focuses more on records systems and controls. The latest standard is also more theoretical than the previous version in 2001 and more practical. Besides, he also mentioned that some of the terminologies in the latest standard were changed, such as "Appraisal," which focused on the value of records for continued support in the business. The methodologies used in the new standard are seen as the most significant change to help establish the records management program (Convery, 2016). This new standard seems can contribute to managing and controlling big data. As stated by Convery (2016), the new standard can help manage structured and unstructured data that can be invaluable to all organizations by applying a theoretical framework that helps in each business activity, including the contexts and process.

#### 5.0 The Elements In Records Management For Managing Big Data

Referring to the Records Management standard, ISO 15489-1:2016(E), some elements in this standard have been taken out to see whether it is suitable or not for managing big data issues. There are **eight (8)** elements in ISO 15489-1:2016(E) that was chosen as per below:

Table 1 Elements in ISO 15489-1:2016 (E)

No	ISO 15489-1:2016(E)
1	Creation
2	Capture
3	Classification and indexing,
4	Access control
5	Storing records
6	Use and reuse
7	Migration/conversion,
8	Disposition.

Based on the table above, ISO 15489-1:2016 (E), **Creations** are related to the records created to perform the business activity. The creation of records should involve content and metadata. Appraising records from business, legal, and other requirements are needed in deciding on records creations. This all must document the circumstances of their creation. According to the function of creation in ISO15489 (2016), the writers believe that identifying what records or information is needed based on business activity can also help identify what records can be created. This can help control the overwhelming volume of information created every day without knowing the actual purposes of the information creation.

Next is **capture**. According to ISO 15489-1:2016 (E), capturing happens when there is an appraisal to identify necessary information to be kept and managed over time by capturing them in a records system. Capturing is functioned to identify unique elements in data that are valuable. According to the function of capture in ISO15489 (2016), the writers believe capturing metadata can help identify valuable information where needed in decision-making. Capturing can also find the relationship between the records with other records. Hence, it will help to keep the valuable information only.

Classification based on ISO 15489-1:2016 (E) links the records with the business context by associating them with categories in a business classification scheme. Records classification can be applied by an individual or any level at any time. Due to that, the writers believe classification can help classify the information in the big data area that has quality and value to the organization. Hence, indexing can also help retrieve valuable information for correct purposes at correct times and people.

According to the ISO 15489-1:2016 (E), **access control** is about managing the authorized proceed for those who can access the records, which can help to allow certain people to access the information that has value; the writers believe this element can help to protect the valuable information missing, misuse, and damage in the big data area. However, the access control rules depend on the requirement, needs of an organization, business activity, and risk associated with the business activity.

Other than that, it is an element storing records. Based on ISO 15489-1:2016 (E), **storing records** is recorded in any format or media is kept properly and correctly. This is functioned to avoid any loss or destruction, including theft and disaster, and to avoid unauthorized access. Therefore, The writers believe that by storing records, valuable information can be used longer by keeping it at a secure place or platform.

According to the function of **use and reuse** in ISO 15489-1:2016 (E), the writers believe use and reuse elements are also needed to make sure valuable information is more long-lasting. Therefore, information usability can also be maintained by creating backup information, converting it into alternative formats, or developing a disaster plan.

Next is **migration or conversion**. Based on ISO 15489-1:2016 (E), records migration is about transforming records from analogs to digital format (digitization). The writers believe migration or conversion elements are needed for a variety of big data. This is because big data comes from various sources, especially on internet platforms or media. Therefore, it is essential to have this element to make sure the valuable information can be used for a long-term period, moreover, due to the increase in technology nowadays. It is essential to have these elements.

Finally, is the disposition **element**. According to ISO 15489-1:2016 (E), a disposition is about destroying metadata and records. The disposition also transfers control of records to the organization that is assumed responsible for business activity. The disposition also functions to control the permanent retention of external archives. Due to that, the writers believe disposition is an essential element in the volatility of big data. Since the increase of big data leads to challenges in terms of storage and many more, it is crucial to know the period of information that can be kept. From this element, only valuable data will be kept, while other information that is no longer needed and exceeds the retention period will be destroyed or disposed of.

Moreover, as we can see today, many tools in big data analytics were created to support the emergence of big data. Therefore, through the disposition element, the writers would like to see whether the big data can be controlled and minimized by disposing of unnecessary data or data that is no longer needed for the organization. Only valuable data will be kept for future reference.

## 6.0 Research Design

To know whether the Records Management elements are suitable or not for managing big data, the writers will use qualitative data collection. This will explain in research design that is divided into three (3) phases. In phase one (1), the writers do some surveys and analyses through qualitative data collection. In this phase, big data issues that emerged in IoT were discussed. This follows by analysis and discusses the current records management elements suitable for managing big data. In this phase, the writers gain an idea to develop the research questions and matrix analysis to see which elements in the records management standard can tackle and handle big data issues.

The research design's second (2) phase is focused on data collection. In this phase, the writers choose purposive sampling by narrow criteria of selecting informants from an organization dealing with big data emergence. In this phase, the writers come out with guidelines on the research questions that must be asked during the interview.

Lastly, in phase three (3), the data that has been collected in phase two (2) will be analyzed and interpreted. Through the data collection, the writers will discuss the current elements in records management and big data and whether it meets the objective of the study or not. This is a crucial phase where the writers can find the best guideline in records management that is suitable for managing big data.

## 7.0 Findings

Through proper research design, the study can develop and provide a guideline in Records Management in dealing with the big data environment. There is a possibility to have new elements in Records Management that can also manage big data issues. This needs to be discovered by in-depth conversations and interviews with the records professionals as well analyzing most current and available documents. The findings of the study will greatly give impact to the Information and Records Management community on its importance in managing big data. Moreover, it opens up the door for future research on the ability and potential of Records Management in the world of information.

## 8.0 Conclusion

In a nutshell, the increase in the amount of big data is unstoppable that a proper approach must be in place. The failure to manage big data can give significant impacts on organizations. Information professionals and other researchers need to collaborate to identify the best way to manage and control the surge of big data. This paper foresees some elements in Records Management Standard, ISO 15489-1:2016 (E) and some identified models as well as other frameworks suitable for managing the said issues. The needs to come out with more in-depth study will give more views and probably solutions to the challenges currently and ahead of us.

## References

- Almeida, F. (2017). Benefits, Challenges and Tools of Big Data Management. *Journal of Systems Integration*, 4, 12-20. DOI:10.20470/JSI.V8I4.311
- Aly, H., Elmogy, M., & Barakat, S. (2015). Big data on internet of things: applications, architecture, technologies, techniques, and future directions. *Int. J. Comput. Sci. Eng.*, 4(6), 300-3013.
- Ang, P.H., & Loh, C.M. (2016). *Internet development in Asia*. Wayback Machine. [https://web.archive.org/web/20160103054146/www.isoc.org/inet96/proceedings/h1/h1\\_1.htm](https://web.archive.org/web/20160103054146/www.isoc.org/inet96/proceedings/h1/h1_1.htm)
- Convery, N. (2016). *There is a new version of the international standard for records management (ISO 15489-1:2016)*. <https://infrastructure.jiscinvolve.org/wp/2016/06/27/there-is-a-new-version-ofthe-international-standard-for-records-management-iso-15489-12016/>
- Digital 2019: Global digital overview — DataReportal – Global digital insights*. (2019, April 13). DataReportal – Global Digital Insights. <https://datareportal.com/reports/digital-2019-global-digital-overview>
- Garyaev, N., & Garyaeva, V. (2019). Big data technology in construction. *E3S Web of Conferences*. <https://doi.org/10.1051/e3sconf/20199701032>
- Joashi, N. (2018). *Problems with big data that we failed to notice*. <https://www.allerin.com/blog/problems-with-big-data-that-we-failed-to-notice>
- Kadadi, A., Agrawal, R., Nyamful, C., & Atiq, R. (2014). Challenges of data integration and interoperability in big data. *Proceedings - 2014 IEEE International Conference on Big Data, IEEE Big Data 2014*. <https://doi.org/10.1109/BigData.2014.7004486>
- Kundhavai, K. R., & Sridevi, S. (2016). IoT and Big Data-The Current and Future Technologies: A Review. *International Journal of Computer Science and Mobile Computing*.
- Mohammed, A. F., Humbe, V. T., & Chowhan, S. S. (2016). A review of big data environment and its related technologies. *2016 International Conference on Information Communication and Embedded Systems, ICICES 2016*. <https://doi.org/10.1109/ICICES.2016.7518904>
- New Strait Times. (2019, January). *Malaysia ranks top 5 globally in mobile social media penetration, highest in region*. <https://www.nst.com.my/lifestyle/bots/2019/01/456119/malaysia-rankstop5-globallymobile-social-media-penetration-higest>
- Arifin, N. A. N., Yunus, A. M., & Kadir, I. K. A. (2021). The role of big data in the healthcare industry. *Journal of Islamic*, 6(36), 235-245.
- Oussou, A., Benjelloun, F. Z., Ait Lahcen, A., & Belfkih, S. (2018). Big Data technologies: A survey. In *Journal of King Saud University - Computer and Information Sciences*. <https://doi.org/10.1016/j.jksuci.2017.06.001>
- Pember, M. (2006). Sorting out the standards: What every records and information professional should know. In *Records Management Journal*. <https://doi.org/10.1108/09565690610654765>
- Somani, A. K., & Deka, G. C. (2017). Big data analytics: Tools and technology for effective planning. In *Big Data Analytics: Tools and Technology for Effective Planning*. <https://doi.org/10.1201/b21822>

Su, X. (2018). *Introduction to Big Data*. Ing. Luis Armando García Eliseo – Docente | Consultor en Tecnologías de la Información.  
<https://lagesoft.files.wordpress.com/2018/11/bd-introduction-to-big-data.pdf>

Unhelkar, B. (2017). Big data strategies for agile business: Framework, practices, and transformation roadmap. In *Big Data Strategies for Agile Business: Framework, Practices, and Transformation Roadmap*. <https://doi.org/10.4324/9781315120232>

Wulff, A., & Wunck, C. (2016). Integration of business process management and Big Data technologies. *Proceedings of the International Conference on Industrial Engineering and Operations Management*.

Xiaomi, X., & Jiao, H. (2004). Assessing records management in China against ISO 15489 and the implications. *Records Management Journal*.  
<https://doi.org/10.1108/09565690410528929>

Yasmeen Shahdad, S., Miraj Khan, F., Meraj Bilfaqih, S., Sultana, H., & Ashfaq Hussain, M. (2018). IMPACT OF BIG DATA IN INTERNET OF THINGS (IoT). *Ijct*.