Big Data Management and Archival Practices: A Case Study of Omani Institutions Through Records Specialists' Perspective

Abderrazak Mkadmi[1], Faten Hamad[2], Mahmood Saif Almaqbali[3]

Abstract

This study delves into the current state and future potential of Big Data within Omani institutions, with a specific focus on insights provided by information and archives specialists. Employing a qualitative approach, the research conducts semi-structured interviews with 15 experts representing a diverse array of twelve Omani institutions, encompassing both public and private sectors. Participants were meticulously chosen based on their roles in records and archives management, supplemented by the involvement of IT specialists and administrative managers. The findings illuminate a nuanced understanding of Big Data within Omani institutions. While there exists a significant level of comprehension among staff regarding the complexities of Big Data, notable gaps persist in its practical application. Participants perceive Big Data as vast, diverse, and challenging data to manage, aligning with established scientific definitions. Formal education emerges as a critical avenue for knowledge acquisition, reflecting the evolving landscape where academic and professional domains intersect. Furthermore, interviews reveal the diverse nature of data within organizations, necessitating holistic management strategies and strategic investments in technology and security infrastructure. Key challenges such as data security, legal compliance, and the need for skilled human resources are underscored, emphasizing the necessity for specialized solutions and training programs. Moreover, the discussion extends to the future trajectory of records and information management, emphasizing the imperative of adapting to technological advancements and the continuous pursuit of education. Overall, this study offers valuable insights into the awareness, applications, transformations, governance procedures, and challenges associated with Big Data management within Omani institutions, as perceived by information and archives specialists.

Keywords: Big Data Management, Records Management, Omani Institutions, Records Specialists.

Introduction

The proliferation of big data has transformed archival and records management practices globally, presenting challenges that transcend geographical boundaries. The sheer volume, variety, and velocity of digital data require institutions to adopt new technologies, update governance frameworks, and develop specialized skills. The benefits of big data have been exploited in many fields such as business transactions, national security, education and health care. It improved predictions, saved money, improved efficiency, support strategies and operational intelligence, increased transparency, and eliminated waste, fraud, and abuse by utilizing the business value of enterprise data (TechAmerica Foundation, 2012). Furthermore, because more valuable information can be derived from the data, big data has the potential to significantly shorten the decision-making process and lead to better decisions (Hochtl et al., 2016; Hamad et al., 2022). Big Data has become a reality as a result of its ability to rapidly create

[1] Sultan Qaboos University, Oman / The University of Manouba, Tunisia, Email: a.mkadmi1@squ.edu.om
[2] Sultan Qaboos University, Oman / The University of Jordan, Jordan, Email: f.hamad@squ.edu.om
[3] Ministry of Justice and Legal Affairs (Director of the records Department), Email: maqbali.010@hotmail.com

and collect massive amounts of digital data in a variety of formats (Al-Hujran et al., 2015).

This shift is particularly pronounced in the field of archives and document management. There are many questions about how to effectively analyze and extract relevant information from these archives. The convergence of big data with new technological processing methods has resulted in the creation of "digital forests," which were often developed by individuals who were not concerned with sustainability or legal retention periods, or the conservation of these digital assets (Vieira, 2018). Archivists were initially not involved in this process, but have since realized the complexity of their tasks and the potential for big data to transform the field of archives. In recent years, there has been a growing awareness of the importance of data archiving, which serves at least two main functions: proof and memory. Archiving is necessary for protecting rights, justifying activity, complying with legal obligations, optimizing the cost of consulting and searching for documents, and preserving the integrity and authenticity of documents over time (Mkadmi, 2021). It is not simply a matter of storage, but rather saving data in an intelligent way to be able to use it over time.

In the context of archives and records management, "records" are traditionally defined as documents, data, or information that serve as evidence of an organization's activities, transactions, or decisions, and are systematically managed to remain accessible, authentic, and reliable over time, often governed by specific retention schedules and compliance requirements (ISO 15489, 2016). In contrast, "big data" refers to vast, high-volume, high-velocity, and diverse datasets generated from sources such as social media, sensors, and financial systems, which are often unstructured and not initially intended for archival retention (TechAmerica Foundation, 2012). However, big data can become part of an archival collection under certain conditions: when it holds contextual value (supporting decision-making or documenting organizational actions) (Diebold, 2003), when it is transformed into records (through processes like data summarization or reporting) (Borgman, 2015), or when legal and compliance requirements mandate its retention (e.g., for GDPR compliance or audit purposes) (Gartner, 2012). Additionally, big data may hold historical or research value, where it is preserved to document significant societal or technological trends (Hodgson & Marty, 2018). Distinguishing between traditional records and big data, and recognizing when big data should be archived, allows archivists to apply appropriate preservation, access, and retention strategies, ensuring that important data is integrated effectively into archival systems (Borgman, 2015).

The management and analysis of big data in archives presents both challenges and opportunities for researchers, archivists, and librarians (Borgman, 2015). The challenges include, the sheer volume of data, which can be difficult to store and manage efficiently (Hodgson and Marty, 2018), the diversity of data formats, which can make it difficult to analyze and interpret the data (Dell'Orto and Caviglia, 2018). There are also ethical and privacy concerns related to the handling of sensitive or personal information contained within the data (Dell'Orto and Caviglia, 2018; Hodgson and Marty, 2018, Hamad et al, 2021). Despite these challenges, big data in archives holds great potential for research and scholarship, as it allows for the examination of large, diverse datasets in ways that were previously not possible (Hightower and Dettwyler, 2016). While these issues are faced by institutions around the world, this paper uses Oman as a case study to examine how emerging economies are addressing these universal challenges. By focusing on the experiences of Omani institutions, the study provides insights into how global archives and records management professionals can navigate the complexities of big data management.

Importance of the Study

Despite these challenges, big data in archives holds great potential for research and scholarship, as it allows for the examination of large, diverse datasets in ways that were previously not possible. It can also facilitate more efficient and effective operations within cultural heritage institutions, and support the development of new digital tools and services, such as data visualization and analysis tools, digital humanities resources, and online exhibitions.

Oman, with its rapidly modernizing records management infrastructure, offers a unique lens through which to examine the global impact of big data on archival practices. This study derives its importance from the importance of the reality of big data and the problems it poses at the level of paths, applications, systems and practices of managing information and documents in Omani governmental and private institutions, due to the profound transformations it has caused in terms of data volume, diversity and speed of flow. This case study provides valuable lessons on technological adaptation, legal frameworks, and professional development, illustrating the broader relevance of the findings beyond Oman's borders.

Study Objectives

The study aims to identify the reality of big data in Omani institutions from the point of view of information and documentation specialists:

1.   Exploring the awareness level of big data management among information and documentation specialists.

2.   Identify the big data uses in Omani institutions.

3.   Identify the transformations brought about by big data on activities related to document and information management.

4.   Identify the procedures/systems used by Omani institutions to govern the management of big data.

5.   Identify the challenges facing information and documentation specialists in managing big data.

Research Questions

The study attempts to answer the following questions:

1.   What is the reality of big data in Omani institutions?

1.1 What is awareness level of information and documentation specialists of the importance and challenges of big data?

1.2 What are the uses of big data in Omani institutions?

1.3 How is big data managed by document and information specialists?

2.   What are the transformations in activities related to document and information management brought by big data?

3.   What is the rational of the management of big data from the perspective of the specialization of document and information management?

Literature Review

What is big data?

Big Data first appeared in academic literature in statistics and econometrics in the early 2000s, where it was used to describe "the outburst in the amount (and sometimes, quality) of available and potentially relevant data, largely as a result of recent and unprecedented advancements in data recording and storage technology" (Diebold, 2003). The literature about big data, reveals several definitions for it. For instance, Gartner has defined big data as: ''… *high-volume, high-velocity, and/or high-variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization''* (Gartner, 2012 in Philip Philip Chen and Zhang, 2014, p: 314-315). The TechAmerica Foundation (2012) referred Big Data as a is a term used to describe large amounts of high-velocity, complicated, and variable data that needs advanced techniques and technologies for data acquisition, storage, dissemination, management, and analysis. Later in 2014 De Mauro et al came with a definition of big data after examining existing literature on big data. They defined big data as information assets that gain Value using specific technology and analytical methods for transforming data that is characterized as being High Volume, Velocity and Variety that require into Value. Moreover, Big data is defined as data that is too large for commonly used software tools to capture, manage, and process in a reasonable amount of time. The concept of big data was necessitated by the increasing capacity of available information systems to facilitate the capture, processing, storage, and use of large volumes of variable but credible data in a timely enough manner to generate maximum value for users (Kwanya, 2014). The European Commission (2016) defines Big Data as large amounts of various types of data generated from various sources, such as people, machines, sensors, satellite imagery, digital pictures and videos, transition records, and GPS signals.

According to the available definitions of big data is defined by: a) size, the volume of datasets; b) complexity, structure, and different behaviors of the data; and c) technology to be used to process this volume of data with a complex nature (Ward and Baker, 2013). All definitions have pointed out three main characteristics of big data, namely, Volume, Velocity and Variety. researches have referred to this as 3Vs of big data. The volume refers to the amount of data generated by various resources that must be ingested, analyzed, and managed in order to make better decisions. Velocity refers to how quickly data is produced, transformed, and transferred, as well as how quickly these data must be received and processed meaning that data is accessible anytime, anywhere and in real time (Borodo et al, 2016). The Variety means a wide range of information/data sources; mobile devices, videos, chats, social media, etc… there are another tow Vs that are veracity and value that are recognized as big data attributes. The trustworthiness, authenticity, source and reputation, availability, and accountability of data are all examples of veracity. In other words, the analysis's accuracy is determined by the quality of the source data. The value aspect reflects the potential value of big data, which means that we may have access to big data, but we must convert it into useful data before it can be used. In general, information systems may be unable to handle increasing amounts of data (Li, Jiao, Zhang, & Xu, 2019). According to Chen et al. (2015), as the number of Internet users and web sites grows, so does the number of visitors and the scope of their use, and this trend appears to be accelerating resulting in the prevalence of big data.

Big data is data whose size is beyond the ability of commonly used software tools to capture, manage, and process within tolerable time. The concept of big data has been necessitated by the

growing capacity of the available information systems to facilitate the capture, processing, storage and use of large volumes of variable but credible data fast enough to generate optimum value for the users.

In an era of big data, archivists are finding it increasingly difficult to manage archives without the use of such techniques. These techniques provide numerous benefits that traditional archival approaches do not. Consider the potential of data sciences such as information visualization to transform archival preservation processes (Xu, 2011), or the opportunity to use visual analytics to transform archival representations and finding aids from two-dimensional hierarchies into multidimensional graphical representations with interactivity that supports the integration of human and machine analysis. In May 2016, the White House Federal Big Data Research and Development Strategic Plan (Big Data Senior Steering Group, 2016) discusses the creation of next-generation capabilities; the understanding of trustworthiness of data and resulting knowledge; the increase in the value of data through policies; the understanding of big data collections with regard to privacy, security, and ethics; and the creation of big data benchmarking centers.

Roles of archives and libraries, particularly in relation to big data are closely related while have different aspects. While both manage information, their core functions and challenges with big data differ significantly. Libraries primarily focus on providing access to information for current use, with dynamic and regularly updated collections, whereas archives are responsible for the long-term preservation of records that serve as evidence of organizational activities, emphasizing authenticity, integrity, and compliance with legal retention requirements (Borgman, 2015; ISO 15489, 2016). In managing big data, libraries are often concerned with handling large volumes of information for immediate retrieval and analysis, such as research data or digital publications. Archives, however, approach big data from a preservation perspective, ensuring that datasets with potential long-term value are maintained in ways that preserve their authenticity and usability over time (Hodgson & Marty, 2018). This distinction places different demands on archival systems, which must handle large, often unstructured datasets while adhering to strict governance standards. Moreover, the difference between big data in archives and big data archiving is crucial. Big data in archives refers to the challenge of integrating vast datasets, like social media or sensor data, into traditional archival systems (Borgman, 2015), whereas big data archiving involves developing new methods and infrastructures to store and manage big data itself, using tools such as distributed systems, cloud computing, and blockchain (Hodgson & Marty, 2018). This shift requires moving from traditional, localized archival models to scalable and decentralized storage solutions that can accommodate the size and complexity of big data (Vieira, 2018).

Big Data in Archives Environment

Big data in archives refers to the increasing amount of digital data being generated, collected, and preserved by archives and libraries (Hightower and Dettwyler, 2016). This includes a wide range of data types, such as text documents, emails, social media posts, audio and video recordings, and more (Dell'Orto and Caviglia, 2018). Big data in archives allows for the examination of large, diverse datasets in ways that were previously not possible, enabling new insights and discoveries (Dell'Orto and Caviglia, 2018). For example, big data analytics techniques can be used to identify patterns and trends within the data, and to visualize and communicate the results in meaningful ways (Hightower and Dettwyler, 2016). In addition to its potential for research and scholarship, big data in archives has the potential to facilitate more

efficient and effective operations within cultural heritage institutions (Hodgson and Marty, 2018). For example, big data analytics techniques can be used to identify and prioritize collections for digitization, to optimize storage and access, and to improve the preservation of digital assets (Borgman, 2015). Furthermore, big data in archives has the potential to support the development of new digital tools and services, such as data visualization and analysis tools, digital humanities resources, and online exhibitions (Hightower and Dettwyler, 2016). These tools and services can help to increase the accessibility and visibility of cultural heritage collections, and to support a wider range of research and learning activities (Dell'Orto and Caviglia, 2018).

The transformation of archival practices began with the digitization and virtualization of heritage collections and archival records, allowing them to be accessed through dedicated interfaces and circulated in networks (Felton, 2017). This shift challenged and weakened the traditional roles and responsibilities of archivists and also introduced legal and technological considerations that expanded the number of individuals involved in the practice of archiving. Some have even questioned whether information professionals, such as computer scientists, statisticians, and business creators, should also play a role in this evolving landscape (INRIA, 2014).

In the context of the digital age, Françoise Banat-Berger (2010) has argued that the increasing use of digital technology in archives offers opportunities for archivists to gain more visibility and for archives to become not just sources of historical and scientific value, but also tools of good governance within various organizations and vehicles for conveying civic values (Nougaret, 2017). However, archivists today and in the future face the significant challenge of dealing with large volumes of digital data. According to Servais and Mirguet (2015), this data should not be treated as fixed or "heritage" data, but rather as active data that must be subject to good governance requirements to be effective. As a result, archivists must be involved in the design phase of these processes and ensure that document governance requirements are considered.

Similarly, in the context of big data, archivists must work to classify and identify data to make it usable and reusable while also protecting personal information. To do this, they must consolidate all relevant metadata, including context, meaning, and conservation information, from the outset. This added value to the data and to the data archiving process, while also reducing the cost of conservation and making it easier for producers to use the data (Servais & Mirguet, 2015). One potential solution for managing the large volume of data that requires significant processing power is the use of crowdsourcing, which involves outsourcing tasks to a large group of people (Ben Love, 2018).

However, big data presents a number of challenges for archives, including the volume and velocity of data, the diversity of data sources and formats, and the complexity of data management. One challenge of big data in archives is the sheer volume of data that must be managed. The exponential growth of digital data has led to an explosion in the amount of data that must be stored, accessed, and analyzed. This poses significant challenges for archives, as traditional data management approaches may not be sufficient to handle such large amounts of data (Hodgson & Marty, 2018).

Another challenge is the diverse and rapidly changing nature of data sources and formats. With the proliferation of devices and platforms generating data, archives must deal with a wide range of data types and formats, including structured and unstructured data, as well as data from social media, sensors, and other sources (Hightower & Dettwyler, 2016). This diversity can make it

difficult for archives to manage data effectively and ensure its integrity over time.

Finally, the complexity of data management is a significant challenge for archives dealing with big data. The sheer volume and diversity of data can make it difficult to apply traditional data management techniques, such as classification and organization, and can require the use of new technologies and approaches, such as cloud computing and blockchain (Dell'Orto & Caviglia, 2018). In addition, the legal and ethical issues surrounding the use and dissemination of big data must be carefully considered (Borgman, 2015).

Data governance practices enhance records management by improving data quality, security, and accessibility. Strong governance frameworks ensure that records are managed efficiently, minimizing the risks of data loss or corruption, while also ensuring compliance with legal and ethical requirements (ISO 15489, 2016). Furthermore, emerging technologies such as AI and machine learning are reshaping the management of big data in archives by automating processes like data classification, enhancing the accuracy of retrieval systems, and providing predictive analytics to improve decision-making (Borgman, 2015). These technologies are especially useful in handling the unstructured nature of big data, which is a significant challenge in records management (Dell'Orto & Caviglia, 2018).

Big Data Archiving

The characteristics of big data, including its volume and speed, have significant impacts on how it is archived. Big data is often created by various actors in large, unstructured volumes and is constantly in motion. It is frequently analyzed and exploited for marketing and financial purposes by companies such as American web giants, who use it to better understand their customers and their needs (Vieira, 2018). This makes the task of archiving big data more complex, as traditional archiving tools and solutions that rely on standards are unable to effectively manage such large amounts of data and archivists are often not present at the time of creation. Despite these challenges, several projects have emerged to archive big data using technologies such as cloud computing and blockchain.

Viana and Sato (2014), cited by Vieira (2018), developed a functional model for archiving big data based on the Open Archival Information System (OAIS) functional model and cloud computing. This generic and abstract model emphasizes sustainability in the traditional archival sense and allows external applications to submit or retrieve information in the form of a submitted information package (SIP) or an archival information package (AIP). The architecture of this model consists of three abstract layers: the first layer is that of big data, which includes both structured (SQL) and unstructured (NoSQL) data with associated metadata; the second layer is that of the archives, where archiving, research, preservation, confidentiality, and communication activities occur; and the third layer is a storage layer based on the cloud. However, the developers of this model express reservations about the use of the cloud and suggest that without a coherent archiving model, it may not be a reliable solution. They therefore propose the use of a "long-term digital preservation as a service" (LDPaaS) approach.

Another project, called the European Archival Records and Knowledge (E-ARK) project, was launched in 2013 and aims to improve methods and technologies for digital archiving on a Europe-wide scale. E-ARK is a comprehensive big data research project that manages data and documents throughout their lifecycle, inspired by the OAIS model. It includes a Pre-Ingest component that ensures interoperability between various information systems and allows for the creation of SIPs (submitted information packages). The Ingest component includes the E-ARK

packages that have been submitted. The Preservation component includes tools for managing the data, transforming the submitted packages into archived packages (AIPs) by adding metadata according to OAIS, PREMIS, and METS standards. The Access component includes tools for management, organization, and research, as well as tools for creating broadcast packages (DIPs) and user interfaces. E-ARK tries to take advantage of both traditional electronic archiving systems and big data by maintaining the integrity and authenticity of the data while also addressing scalability issues through the use of the Hadoop platform, including the HDFS and MapReduce models and the HBase non-relational database.

Other big data archiving projects in the public domain include the ARCHANGEL project developed by the British National Archives. This decentralized platform uses blockchain technology to ensure the integrity of large amounts of digital data stored in public archives through cryptography, giving users confidence in the archives. This trust is established by transitioning to a distributed ledger, which guarantees the provenance, immutability, and integrity of the archived documents (Collomosse et al., 2010).

In France, several public bodies have also been interested in archiving projects using blockchain technology since 2014, including the Post Office, which has experimented with this technology while recognizing its limitations, particularly in terms of legal issues. The legal aspect of blockchain technology, including its use for timestamping, proof of existence, or proof of integrity, is indeed problematic as it is not currently recognized by courts (Ducellier, 2017). Other projects in various fields, including healthcare (e.g. Blockchain Health, which aims to preserve and make medical data available in emergencies) and land ownership (e.g. BITLAND, a project in Ghana that allows citizens to register and manage land titles and related transactions to resolve social conflicts), have also been developed using blockchain technology (Mkadmi, 2021).

Why Oman?

While the literature discusses various aspects of big data management, including the integration of emerging technologies such as cloud computing and artificial intelligence into information and records management systems, there is a significant gap in understanding how these technologies can be applied within the context of Omani institutions and similar emerging economies. Moreover, while studies have addressed the technical aspects of big data management, few have explored the legal, regulatory, and institutional challenges that impede the integration of big data governance frameworks into traditional records management systems. There is also limited research on the impact of AI-driven tools and the specific role of data governance in enhancing records management practices, particularly in regions where regulatory frameworks are still evolving to accommodate digital transformations. Oman offers a unique perspective on big data management due to its recent digital transformation efforts and its focus on modernizing records management in both public and private sectors. Like many countries, Oman faces significant challenges in managing rapidly growing digital data, ensuring data security, and integrating new technologies into existing archival frameworks. By studying Oman, we can gain valuable insights into how emerging economies are handling the global shift towards big data management in archives.

Research Methodology

This study employs a qualitative research methodology, using Oman as a case study to explore broader challenges of big data management in archives and records management. Semi-

structured interviews were conducted with 15 experts from 12 Omani institutions, both public and private, to gain insights into the practical realities of big data management within these organizations. The institutions were selected for their involvement in document and archives management, representing a range of sectors that face similar challenges to those encountered globally. While the focus is on Oman, the selection of institutions reflects commonalities in how organizations across different sectors are responding to the challenges of big data. The individuals were strategically chosen based on their distinct functions and pivotal roles within their respective institutions, employing a purposive sampling method to ensure a comprehensive and representative insight into the diverse perspectives prevalent in the Omani context. They were chosen based on their significance in records management and their involvement in big data initiatives. The case of Oman provides a localized perspective on issues that are universal in nature, such as data governance, security, and the integration of new technologies into archival systems.

P	Institution	Function	Code	Status
1	Oman Flour Mills	Assistant Archivist	I1	private
2		Administrative Officer	I2	private
3	Oman Broadband	Documents and archives specialist	I3	private
4	Directorate General Of Civil Status (Royal Oman Police)	Documents and archives specialist	I4	public
5	General Office of the Ministry of Justice and Legal Affairs	Director of the Councils and Committees Department	I5	public
6	National Records and Archives Authority	Documents and archives specialist	I6	public
7	Omani State Council	Documents and archives specialist	I7	public
8	Ministry of Justice and Legal Affairs	Director of Records Department	I8	public
9	Namaa Holding	Documents and archives specialist	I9	private
10	Oman Airports	Director of Records Department	I10	private
11	DUCRAY	Chief executive officer	I11	private
12	General Secretariat of the Tender Board	Documents and archives specialist	I12	public
13	Asyad Group	Documents and archives specialist	I13	private
14	Environment Agency	Documents and archives specialist	I14	public
15	Ministry of Heritage and Tourism	Director of Records Department	I15	public

Table 1: Sample Details

Interviews were conducted in a semi-formal manner by an archivist, with each session spanning approximately 60 minutes. Employing a structured set of guiding questions, the interviews

maintained flexibility to delve into emerging topics that surfaced organically during discussions. To facilitate comprehensive analysis, the conversations were recorded and subsequently transcribed. Initially conducted in Arabic, the dialogues underwent translation into English to ensure broader accessibility.

The study employs Braun and Clarke's (2006) version of thematic analysis, allowing for the systematic identification and coding of prevalent themes and patterns embedded within the interview transcripts. This method facilitated a nuanced exploration of participants' perspectives and experiences, offering a profound understanding of the ongoing landscape of digital preservation and accessibility in Oman.

Results Analysis

The analysis of the results is done according to four themes which are illustrated in the following table:

Big Data in Omani Institutions				
Awareness of Big Data	**Big Data Management Practices**	**Challenges in Managing Big Data**	**Technology and Tools**	**Transformations in Document and Information Management**
- Staff understanding of Big Data. - Sources of knowledge (e.g., training, conferences)	- Data storage methods. - Handling unstructured data.	- Legal compliance issues. - Lack of skilled personnel	- Use of cloud computing, AI tools like Hadoop, and big data analytics platforms.	- Real-time performance monitoring. - Proactive decision-making enabled by data

Table2: Structure for the Thematic Map

Awareness of Big Data

Most respondents confirmed their awareness of the term "big data" through various channels and sources including university studies, seminar attendance, scientific conferences, readings, and training courses related to document and information management. The majority provided similar definitions for big data, cloud computing, and artificial intelligence, indicating widespread understanding and awareness of these technologies. One respondent (I1) described big data as;

"*data collected with greater diversity, reaching increasing storage units at a faster rate, considered a massive amount of data and information from various sources.*"

Another respondent (I2) defined it as "*a massive amount of data and information from different sources,*" he also emphasizing that big data is difficult to handle and process using traditional techniques.

Several respondents (I3, I4, I6, I9, I12, I13, I14, and I15) highlighted the exceptional analytical aspects of big data, considering it as the aggregation, organization, and linking of existing data within an organization to facilitate the analysis process. One respondent (I5) emphasized big

data as;

"*a vast collection of data that requires extraordinary capabilities to manage and analyze. It is challenging to handle using traditional processing and storage techniques due to the continuous generation of a massive amount of information.*"

Some respondents have linked big data with artificial intelligence, for instance one (I11) defined big data as "*data that can be analyzed to train artificial intelligence.*" Consistent definitions were provided for cloud computing and artificial intelligence, with cloud computing seen as an online data storage platform, and artificial intelligence viewed as a computer science field dedicated to solving cognitive problems associated with human intelligence (I1 and I2).

In response to a question about the types of data circulating within the organization, the responses from those interviewed vary in terms of the nature of this data. Emphasis was placed on customer data, products, markets, and financial data (I1 and I2), databases, and file systems (I3). Additionally, there was a focus on data related to the electricity sector, encompassing electricity transmission, distribution, purchase, and supply, as well as the water sector involving procurement, production, and distribution. Other data included sewage sector information, connections within these sectors, and invoices, among other aspects (I9). Administrative, financial, legal data, passenger information, and technical data related to operations and maintenance were also highlighted (I10 and I11). Furthermore, digital office data and email, constituting over 90% of the remaining data, were discussed (I13).

Technologies used in document management

The respondents affirmed the need for specialized software, such as Hadoop, to collect and process data efficiently, particularly data related to social media sites (I13). In the same context, responses mainly focused on electronic document and email management (A1 and I2), such as the File Net system (I9). The use of artificial intelligence in interactive information search operations (I3), and the Internet of Things (IoT) as a group of "*devices and systems that the Internet enables to understand and connect to each other, facilitating collaboration*" (I14), were also highlighted. One respondent (I15) confirmed using the Otter.ai application for editing text versions of conversations and managing them within the application, as well as reading recordings at different speeds and importing audio and video files. Others mentioned using ChatGPT for language translation, while some explored applications in social media (Facebook, Twitter, Instagram, WhatsApp), electronic browsers (Google, Google Maps), cloud computing, government agency applications (Deem, Service, Development, Zakat, etc.), and banking applications such as Bank Muscat and Bank Nizwa (I5).

Challenges of Big Data

Regarding the challenges posed by Big Data, responses were primarily centered on the management of data and information in terms of their form, circulation, and secure preservation, as well as the search for programs necessary to process them, along with the need for a specialized human framework in this field (I1, I5, I7, and I12). The establishment of a specialized service to handle these data, employing skills and providing the necessary devices and tools (I2 and I7) were also reported as challenges. Respondents (I8, I10, and I11) focused on legal and legislative challenges to limit data leaks, where (I8) states;

"*There are legal and legislative challenges hindering the use of certain data on electronic systems, and a lack of flexibility before.*"

Some managers appeared to be reluctant to include or deal with subjects in electronic systems due to their confidentiality and security classification, and some institutions do not receive data in electronic form for fear of data and information leakage. There are also challenges related to the lack of qualified human frameworks.

Concerning data processing, there is no significant financial support for handling data and preparing the infrastructure of devices, networks, and technical programs within the organization. Other challenges mentioned by respondents include the backup and processing of data, maintaining their quality, and determining how to apply retention periods (I13, I14, and I15).

Big Data Applications in Omani Institutions

The responses gathered highlight that data sources in Omani institutions primarily stem from administrative work, customer interactions, or citizens engaging with businesses and government entities, social media platforms, email, etc. Most government sector workers emphasized that data originates from the administrative units of the state apparatus performing their functions. In this context, (I5) specifies that;

"*most data and information available to various administrative units result from the exercise of their functional competencies for which they were created in accordance with the requirements of public interest and stipulated laws and regulations in the Sultanate of Oman, in addition to practical experience and accumulated knowledge in cooperation with other countries in various fields.*"

(I6) further noted that "*big data has multiple sources, including employee databases, civil status and phone number records, email addresses, and multiple accounts*." For private sector workers, big data often originates from the nature of their activities and targeted social media sites. (I13) emphasizes that;

"*big data currently comes from social media sites directly targeted by commercial enterprises, as the information is processed by custom software that can lead to the creation of a new product or deduce specific geographical areas' needs for accessories or cosmetics, reflecting real consumer needs.*"

These data can also be utilized in medical and educational sectors, impacting people globally, providing valuable information that, if processed correctly, is easy to control and direct according to needs. (I10) indicated that megadata originates from "*airline activities and the Civil Aviation Authority (air navigation): flight operation and management, equipment and airport device maintenance, and passenger control.*" Additionally, it comes from "*organization-owned programs such as billing programs, control systems (Scada), and tender and procurement programs,*" (I9).

Big Data Handling

 In the realm of the handling and processing of big data, the responses indicated that big data is being handled in the same way as handling documents using both traditional and modern practices (I5, I8, I9, I10). Documents pass through a comprehensive lifecycle encompassing creation, classification, storage (both physical and digital), and eventual long-term preservation (I1, I2, I5) on servers (I1, I2, I7, I8, I10, I13, I15) and/or cloud storage (I12) while some are still paper based storage system (I5, I8). In line with this, I13 pointed out that big data in handled using the same standard applied for documents. I14 confirmed that they use a specialized tool

called "big data processing tool" that clean analyses and visualize and store big data.

The emphasis was placed on the use of duplicate storage and cloud-based solutions (I1, I2, and I12), with a focus on ensuring data confidentiality and security during archiving and data disposal in accordance with the tables. Conservation and compliance requirements (I1) emphasized that data availability complies with the law on documents and archives, as indicated by (I2);

 "*data availability is in accordance with the law on documents and archives and procedures followed and specified based on the type of data, and it is further reinforced through seminars and conferences to confirm and correct historical information and to benefit by directing the future.*"

Regarding the equipment and devices used to manage big data, responses primarily centered around general programs developed by Google, Microsoft, Oracle, and other open-source programs. (I7) mentioned SPSS, EXCEL POWERPOINT, and SQL programs, while (I8) referred to Microsoft, Windows, firewalls, antivirus programs, video communication programs. (I14) added the Python programming language, Power BI applications, SAS statistical software suite and Excel.

Respondent (I5) highlighted that the organization relies on several programs to manage and process big data, such as:

- "*Correspondence Management System: an internal electronic system through which subjects received by the department are referred to the relevant department for study until fully submitted for approval, and then sent to the competent authority.*

- *Lawyers and Experts System: an electronic system available to workers in the lawyers' and experts' sectors to receive requests for registration or re-registration in the lawyers' and experts' board, as well as other requests, and then track all their procedures.*

- *Ministry Website: Data resulting from the Ministry's exercise of the competencies for which it was created are published so that everyone is aware, with the Ministry's website considered the official reference for those working in the field of law in the Sultanate of Oman*".

The only respondent (I13) admitted that their organization lacks software to manage and process big data.

Document and Information Management Operational Transformation

Apparently, the integration of big data into information and document management has revolutionized the operational transformations in the Omani institutions in several ways. 40% of respondents indicated that big data management system is inherently utilized within their document and information management system (I3, I6, I8, I9, I10, I15). For instance, I15 said:

"The institution has an electronic document and document management system (Wosool). This system provides a secure environment for storing and managing the institution's data, which provides it with a central repository to store and organize its documents, which facilitates the management and control of access to important information and ensures the accuracy and security of data."

While, I4 indicated that big data is partly included in the document and information management by providing the data of specific departments where the data is presented in an integrated manner

with the other departments. I13 indicated that their institution is planning for the future integration and compatibility of the document management system and electronic documents with other business management systems such as human, financial and operational resources management in the organization. The responses generally reflect that although the institutional document and information management system holds and process all institutional data the responses do not perceive it as big data. This was indicated by respondents (I1, I2, I5, I7, I12, I13, I14). One institution (I11) indicated that they still do not have document and information management system.

Big Data Operational Transformations

Big data integration has brought out several transformations. It has increased operational efficiency of institutions. This transformation has enabled real-time monitoring of performance indicators, improving the ability to track effectiveness. Big data has significantly impacted document management by enhancing continuous performance monitoring, as I3 indicated "*it helps to monitor performance indicators continuously and directly*". It has improved data preservation and enabled quick responses during data analysis, for instance I6 pointed out this by saying, "*facilitating tangible information through quick responses to analysis results and indicators*". This integration has led to more proactive decision-making, as accessing information has become quicker and easier, facilitating speedy decision-making through easier access to information (I5, I6, I8, I13, I14, I15). In line with this, I8 emphasized that electronic document management systems are crucial for organizing and retrieving data, reducing paper usage, and supporting digital transformation. While I13 and 114 affirmed that big data management enhanced information and data acquisition that positively facilitating valuable information. The respondent I15 affirmed that "*Big data is considered a wealth of information due to the results and positives that this data provides, which has made organizations race to benefit from it*".

The ability to harness big data for decision-making and policy development is transforming how organizations and governments operate. By analyzing large datasets, institutions can identify trends, forecast future scenarios, and make informed decisions. Big data analysis is particularly beneficial in crafting laws and regulations, as it allows policymakers to base their decisions on empirical data, thereby enhancing the relevance and effectiveness of new policies (I15). It also expedites problem solving, research, developments and economics policies.

However, integrating big data with traditional systems remains a challenge, calling for new methodologies in data and document management necessitating the need to transform infrastructure and human skills and competencies. In this regards, I7 affirmed the need for new systems, software, and technologies for efficient big data and document management. These changes indicate a significant shift towards a data-centric approach in managing information and documents that requires financial allocation and training planning to enable specialists to handle big data. I2 indicated the evolving role of records specialists to integrate and handle big data, emphasizing the importance of data specialists and collaboration between different roles. Furthermore, I10 has pointed out that big data integration with document management systems has brought alongside the use of artificial intelligence for data analysis, that in turn has altered the roles of documentation specialists, emphasizing data specialists' importance. Big data has necessitated a careful balance between complying with legal requirements in document management and ensuring data availability for authorized use. For instance, respondents I12 and I15 indicated that legal frameworks and systems for managing documents have been

established, reflecting big data's influence in shaping document management practices (I12, I15).

In Oman, several laws and regulations, notably the Documents and Archives Law (Royal Decree No. 2007/60) and the Law on Classification of State Documents (Royal Decree No. 2001/118), along with ministerial decisions and circulars from the National Documents and Archives Authority as explicitly mentioned by I5, I8 and I15, govern document management. I6, I9 and I10 stated that they manage big data in compliance with laws and procedures effective in Oman. I3 emphasized that documents are not part of the organization's big data and made accessible internally and to relevant departments. Documents availability is governed by its classification that is four levels (top secret, secret, discreet, limited) to serve security purposes (I14). I13 described the process of managing institutional data and availability procedure, the responded stated that;

"*Raw data in business management systems is managed per approved workflows and legal texts, and when converted into documents like PDF for example, they are managed through an electronic document management system and document retention period as per legal requirements*.

Overall, the availability and access to this data are also governed by the Omani Documents and Archives Law. Intestinally, I12 reported that the do not follow the Omani Documents and Archives Law to manage document availability; instead they base their document availability procedures on laws and legislations related to electronic data and transactions.

Omani institutions operate under the Omani Documents and Archives Law, yet they face challenges in adapting this legal framework to accommodate the rapid pace of digital transformation. Globally, archival institutions are similarly constrained by legal frameworks that need to evolve to address the complexities of big data management. The need for clear data governance policies—covering data security, privacy, access control, and compliance—applies universally, with many countries striving to update their legal frameworks in line with international standards such as the GDPR (regulations governing personal data protection in the European Union) or ISO 15489 (An international standard for records management that provides guidelines on how to create, store, and manage records efficiently and securely.).

Revolutionize Operational Performance

Integration and Optimization

Data integration of across institutions, supported by modern technology and coherent legal and procedural frameworks, is revolutionizing the way organizations handle and utilize data. The responses focused on need for developing advanced data management systems, coupled with the strategic use of big data, enabling organizations to make more informed decisions, optimize operations, and maintain a competitive edge in an increasingly data-driven world. The responses generally focused on creating a more unified and efficient data management system. Responses suggest a need for an interconnected framework where data from various institutions can be consolidated and accessed seamlessly (I7). This integration would allow for better data utilization, leading to more informed decision-making. In this regards, I3 marks the integration of data across institutions through direct connections to a central government station and a unified platform for data display as significant advancement in data management and utilization.

Respondents (26%) indicated the use of modern technology, especially in the processing and analyzing of data, as critical for integration. The advent of large storage solutions and the

development of worker skills in line with contemporary scientific advancements are pivotal. In line with this, I5 stated that big data can be utilized by "*using advanced and modern softwares for data processing and analysis, facilitating data storage and capacity building of employees related to latest skills and technological advancements*" These technologies facilitate the handling of vast amounts of data. It also enhances the accuracy and efficiency of data analysis. Focusing on key indicators and regular assessments of company management underscores the increasing importance of data-driven strategies in organizational operations and policy-making. This is particularly relevant in customer classification, where nuanced data can be processed quickly to identify patterns and trends that inform strategic decisions (I1).

Establishing coherent work procedures in line with the Documents and Archives Law is crucial for ensuring legal compliance and operational efficiency (I6). This requires institutions to have clear data management plans, including the acquisition of systems and software dedicated to managing, processing, and preserving data (I5, I7, I11). Furthermore, I8 made an argument related to the significance of the continuous development and updating of software systems to keep pace with the rapidly evolving digital landscape, which increasingly relies on cloud computing. This approach not only ensures that data is managed efficiently but also that it remains accessible and secure, aligning with legal and ethical standards.

Moreover, I9 has pointed out the importance of data preservation methods, such as classification, coding, and the elimination of redundant storage, for maintaining data integrity and relevance. I13 elaborated further on this point stating,

"*important data and big data is transferred to authoritative bodies namely National Documents and Archives Authority and stored in secure document and electronic document management systems guarantees that it can be effectively processed and analyzed to meet organizational objectives*".

This highlights the need for a comprehensive methodology to manage data throughout its lifecycle, from creation to eventual archiving or destruction.

The formation of specialized teams or departments focused on developing big data management policies and guides is a strategic move (I12). I12 elaborated that "*These teams, by collaborating with private sector entities and specialists, can devise effective strategies and suggest initiatives to maximize the utility of big data*". This collaborative approach ensures that a wide range of expertise and perspectives inform the development of big data policies, leading to more effective and innovative solutions.

Infrastructure Investments

Responses indicate that substantial investments in both hardware and software are crucial. This includes not only state-of-the-art data processing tools but also infrastructure that can support large-scale data storage and analysis. Electronic document management systems (EDMS) have emerged as pivotal tools in the realm of data management and utilization. These systems offer a structured approach to managing the flow and storage of documents in digital form, which is increasingly important in today's data-driven world. EDMS represent a strategic approach to making data a cornerstone of organizational efficiency, policy development and decision-making. The integration of these systems with technologies like AI and cloud computing, coupled with the emphasis on training and awareness, enables organizations to harness the full potential of big data in a globally connected world.

The respondent I14 pointed out the impact of EDMS on organizational spending is noteworthy. By shifting from paper-based to electronic document management, organizations can significantly reduce the costs associated with printing, copying, and physical storage. This practice aligns with environmentally sustainable practices by reducing paper consumption and also translates into direct financial savings. Additionally, the reduced need for physical transportation of documents streamlines operations and cuts down on associated logistical expenses.

Big data is strongly associated with other digital technologies like cloud computing and artificial intelligence. Big data, characterized by its volume, variety, velocity, and veracity, requires robust infrastructure for effective storage and processing. I14 focused on the need for investing in cloud computing to offer scalable and flexible solutions, enabling organizations to handle large datasets without the need for substantial physical infrastructure. Furthermore, both I10 and I14 elaborated on the importance of integrating artificial intelligence to facilitate the extraction of actionable insights from vast pools of data, making big data an invaluable asset for decision-making, strategic planning, and policy formulation. I10 also mentioned the value of implementing AI, cloud computing, and blockchain technology to further support the integrity, security, and accessibility of data. A key aspect of maximizing the benefits of EDMS and big data is the awareness and education regarding document and data management (I9). Organizations must ensure that their personnel are well-versed in managing digital documents, including the efficient handling, storage, and disposal of data.

To support the investment in big data further, I5 affirmed the adoption of open software for data processing and analysis and training specialists in these areas is crucial to fully harness the potential of big data. I6 focused on the importance of aligning data management practices with the organization's long-term vision and objectives to ensure consistency and effectiveness in operations. Furthermore, I7 stated; "*acquisition of EDMS and integrate it with other organizational systems*" is required. This enhances cohesion and efficiency allowing seamless data flow and accessibility, crucial for timely decision-making and operational agility. Also, collaborations between government units, such as National Center for Statistics, are essential for effective data collection and management, enabling a more coordinated approach to national data handling (I7). Lastly, I8 suggested the development of independent national systems like the Arab Cooperation Council underscores the importance of data dominance and regional cooperation. Such systems enable countries to manage and exchange data in alignment with their unique regulatory and cultural contexts.

Data Protection and Security

The importance of safeguarding big data is a recurring theme in the responses. This involves implementing advanced security measures and protocols to protect data integrity and privacy. The focus is on developing systems that are not only robust in terms of security but also adaptable to emerging threats. For instance, I1, I3, I6, I7 and I8 confirmed the importance of implementing an advanced security system, including key verification schemes, content leak detection, and anonymization systems to guard against unauthorized access and cyber threats. The respondent I6 emphasized the need for establishing strict laws for data processing, coupled with robust security procedures, is another crucial step. I6 articulated further on the need for "*…periodic reviews of laws and continuous risk assessments…*". This approach not only enhances data security but also fosters a sense of responsibility and ownership among national stakeholders. I8 articulated on the need for organizations to adopt a clear information and data

security policy, approved by the head of the unit, and ensure that all employees adhere to it. This policy should encompass the acquisition of systems and software designed to protect data and counteract security breaches. Furthermore, I12 elaborated on the need for implementing effective strategies, policies, and control tools in place to prevent unauthorized modification or access data.

Encrypting data is a fundamental security measure, ensuring that it can only be decrypted with a specific password (I1, I7, I14). Moreover, I14 affirmed the importance of enhanced network security measures, such as firewalls, intrusion detection systems, and Security Information and Event Management (SIEM) tools, in preventing unauthorized access to an organization's network. Regular audits and risk assessments are crucial in identifying vulnerabilities within big data infrastructure. Furthermore, some respondents indicated the aspects related to human factor. In this regards, I14 stated; "*Also, employees can be an important source of security vulnerabilities, often unintentionally. Therefore, it is necessary to educate and train them on best practices.*" In line with this, I8 emphasized the need to raise their employee's awareness about issue related to protection their personal data, as he stated; "*Continuously educate the organization's employees to maintain their personal data, and update usernames and passwords regularly*". Mostly, respondents indicate that storing their institutional big data on servers and data centers that facilitate large storage capacities and adhere to high security and safety is essential. In addition, advanced software for protection and provisions of backup copies are necessary to recover lost data in cases of natural disasters or human errors (I13, I15).

Big data protection and security encompass a multifaceted approach involving technological, legal, and educational strategies. The implementation of advanced security systems, continuous employee training, and the establishment of robust policies and administrative structures are crucial in protecting data against the evolving landscape of cyber threats.

Document and Information Management Future Requirements

Technological Transformation

The future of the document management sector appears bright, with promising opportunities to leverage artificial intelligence techniques, which are intimately connected with big data management systems. In this regard, I13 explained that;

"*The focus of the modern document specialization at the present time lies on artificial intelligence techniques that are directly related to big data management systems... studying and analyzing data will enable us to reach a bright future that enables documents to be managed in the correct way and according to the future need that is witnessing an inflation and explosion of information and data in various sectors.*

Apparently respondent visualize that AI could enable effective management of documents tailored to future needs in an era marked by an explosion of information and data across various sectors. Furthermore, document management can benefit from vast data in public sectors (I14). In particular, document management can benefit from the extensive data related to citizens' transactions and records and therefore, through interlinking this data and facilitating services through an integrated network, updated and comprehensive information can be made readily available.

Database management is a critical field in both business and technology, significantly contributing to data storage and retrieval. Respondent I15 to visualize the future of document

management in light of database management systems. It facilitates a reliable basis for decision-making and simplifies work processes and flow. Furthermore, a notable shift towards comprehensive cloud databases and the transformative impact of artificial intelligence on data analytics are key trends to keep track of. Moreover, and to solidify the future of database management, it is important to closely relate it to data security and compliance with evolving data security standards, enhancing database performance in the long term.

As the field of information technology continues to develop at a staggering pace, document management procedures must evolve correspondingly to remain effective and flexible, particularly in the aspects of data preservation, storage, analysis, and processing. In this regards I5 believes;

"*that the importance of document and archive management will be very high provided that document management procedures in general develop so that they are in line with the tremendous development taking place in the field of information technology*"

I6 and I7 believe that the future of big data is promising. The respondent based his/her argument on the growing volume of information, driven by the advent of big data, that is facilitated by the availability of sophisticated software and applications that is transforming the way data is processed and analyzed, saving previously expended time and effort. Documents, serving as repositories of information, have the potential to be mined, analyzed, and utilized effectively, thus forming a fundamental component of organizational infrastructure (I7). The output from these processes can offer invaluable insights for decision-making and strategic planning.

Document and archive management is undergoing a profound transformation, propelled by advancements in information technology. Responses suggest a shift towards more digital-centric practices, where traditional document management evolves to incorporate advanced data processing and analytics capabilities. The future of document management is promising, provided that these systems can adapt to the changing landscape of data. I1 pointed out the secure and organized documentation of information, coupled with the ability to smoothly access, preserve, and process vast amounts of data, as pivotal in today's digital age. The increasing volume of data, its rapid circulation globally, and the constant evolution in data management techniques highlight the critical importance of maintaining privacy and managing unstructured data effectively.

Educational and Training Requirements

Enhancing educational opportunities in the field of documents management, especially in the context of the electronic information revolution, is multifaceted and complex. It encompasses several key areas, including expanding access to educational materials, automating document handling, and integrating information technology into the curriculum. The responses call for specialized training programs focusing on data analysis, management, and the latest technological trends, ensuring that professionals are equipped to handle the complexities of big data. Training will empower them to efficiently handle big data, which is becoming increasingly vital across various sectors. For instance, I13 expanded on leveraging international expertise and best practices as a key strategy for document management in the ear of big data empowerment. Specialized courses and workshops targeting diverse societal segments can facilitate the transfer of global knowledge and skills into the local training and education sector, thereby enriching and developing the information and documents sector.

While the role of documentation specialists in the realm of big data is not yet fully defined, their

skills and expertise can be developed and adapted to meet the demands of big data management. Continuous training and the creation of new roles in this field will be essential. I8 argued that to capitalize on these opportunities, there is a pressing need to train and qualify personnel in the field of document management and modern archives; "*It is necessary to train and qualify human cadres in the field of modern document and archive management…*". This training should incorporate specializations related to advanced technologies, reflecting the rapid changes in modern technology; "*…to involve specializations related to the field and advanced technologies in light of the rapid changes in the world of modern technology*". Additionally, the enactment of legal legislation governing big data is essential to ensure ethical and efficient handling of information. The respondents I9 and I10 capitalized on the value of the documentation sector playing a crucial role in data management, facilitating the organization and handling of data of all types and sizes. Furthermore, I12 elaborated on the future role on documentation specialist in the fields of big data. I12 articulated;

"*Documentation specialist can be invested to be part of the big data management system because he possesses skills that can be developed and adapted in the fields of big data with continuous training and the creation of new jobs in this regard*".

The necessity for education and training to evolve in response to the rapid advancements in technology and data management is a crucial consideration in today's information-driven society. I9 articulated on the need for transforming traditional methods of education a capitalize on electronic systems, artificial intelligence, and big data management integration into document management. This shift is essential not only for keeping pace with current developments but also for preparing individuals to effectively manage and utilize the burgeoning volumes of data. Furthermore, I12 emphasize on the need for adding courses specifically tailored to big data management within document specializations and implementing specialized programs for document management professionals.

However, there are significant challenges in realizing the full potential of big data. One major issue is the lack of specialized skills and competencies in data science. I14 has pointed out that;

"*…major challenges facing the benefit of big data, including the lack of specialized skills and competencies trained in data science and capable of dealing with it efficiently…*"

The respondent added another challenge related to inadequate infrastructure and systems that are capable of effectively receiving, managing, and analyzing large data sets. To address these challenges, I15 articulated the imperativeness of increasing the number of specializations focused on information technology, artificial intelligence, and data science. Furthermore, creating more applications and conducting precise studies focusing on technology in big data management can also contribute to overcoming these hurdles. Also, organizing specialized conferences and workshops on big data management can foster a deeper understanding of these issues and promote the sharing of knowledge and best practices (I8) to further facilitate awareness of the latest developments in the information and documents sector. As we are in the era of the electronic information revolution, it is vital for educators in the field of documents and archives to focus on aspects related to information technology, such as information circulation, analysis, processing, preservation, availability, and security (I1, I5). To further support, the training process, I3 mentioned the need for developing platforms characterized by user-friendliness, making training and development more accessible records specialists. The emphasis should be on promoting educational opportunities for all, expanding access to data and documents, and enhancing the automation and understanding of documents. This approach is

critical for building new and diverse paths in the field of documents and manuscripts and for ensuring that the workforce is equipped to handle the challenges and opportunities of the information age.

The shortage of skilled professionals in Oman capable of managing big data effectively mirrors a global challenge. Institutions worldwide require archivists trained not only in traditional records management but also in digital preservation techniques, data science, and cybersecurity. The importance of continuous professional development in these areas is a critical takeaway for global archival practices.

Summary of results

We try through the following table (table 2), to show the main specific views shared by participants for each sub-theme.

Awareness of Big Data	Big Data Management Practices	Challenges in Managing Big Data	Technology and Tools	Transformations in Document and Information Management
: 10 out of 15 (I1, I2, I3, I4, I6, I9, I12, I13, I14, I15) mentioned awareness of Big Data, understanding it as a massive and diverse set of data that presents challenges for traditional data management techniques	8 out of 15 (I1, I2, I5, I7, I9, I10, I12, I13) discussed the integration of Big Data into document and information management systems, with some relying on electronic systems, others still using paper-based storage	6 out of 15 (I8, I10, I11, I13, I14, I15) highlighted legal and regulatory challenges related to data confidentiality, security, and processing capabilities, along with a shortage of skilled professionals in data science	5 out of 15 (I3, I6, I8, I13, I15) emphasized the use of specialized tools like Hadoop, cloud computing, and AI to manage and analyze Big Data.	7 out of 15 (I3, I6, I8, I9, I10, I13, I15) noted that Big Data integration has transformed their institutions' operational efficiencies, enabling real-time monitoring and more proactive decision-making

Table 3: Specific views shared by participants for each sub-theme

As shown in the table 2, a good level of knowledge about Big Data is available to the majority of the sample participants 10 out of 15, while only 8 have integrated data management practices into their documentary systems and a significant share of them is still dependent on paper storage. The challenges encountered include, among others, legal and security issues by 6 participants, and the use of specialized technological tools as indicated by 5 participants. Finally, 7 out of 15 observed a change in practices owing to the implementation of Big Data; it enhanced efficiency and real-time decision-making. As for the representation of key themes and their inter-relationships about Big Data management within the Omani institution, it can be illustrated by the figure below. It will help in showing how big data impacts different elements of records and information management in Omani institutions through. Awareness, management practice, challenges, and technology and transformation all have a relationship and show the influence of
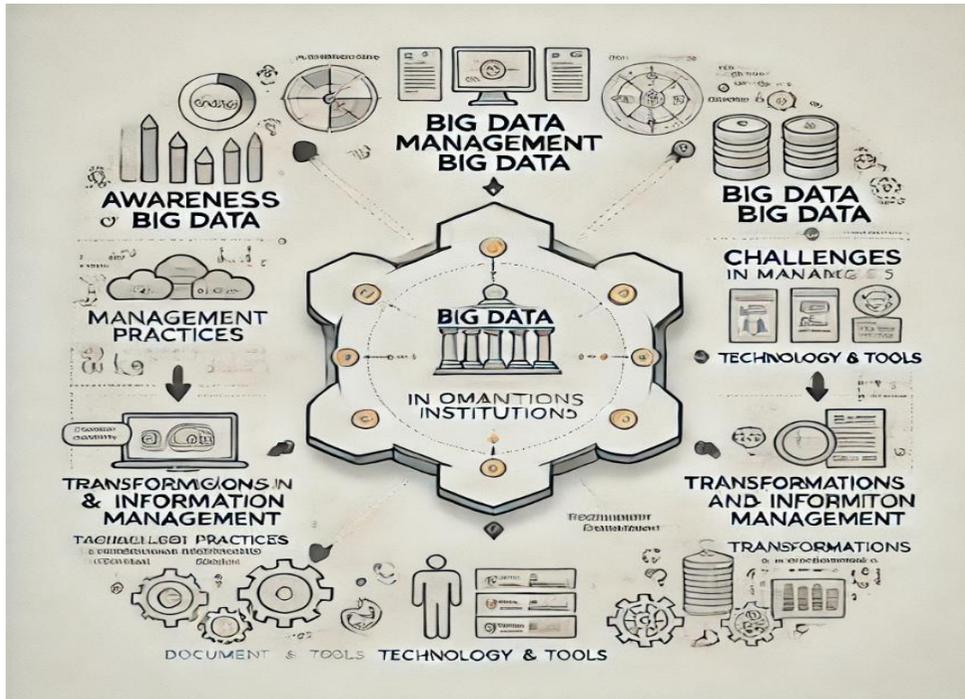
one aspect to another.

Fig1: Interconnected Relationships Between the Key Themes of Big Data Management in Omani Institutions

This schema summarizes the interleaved relationships that exist between the main factors of Big Data management in Omani institutions. The central influence is Big Data in Omani Institutions, which in turn shapes and is shaped by many factors.

Big Data awareness will, in turn, drive challenges in managing Big Data and transformations in document and information management. The more aware the institution becomes of Big Data, the more it will be driven to transform its management due to challenges in skills gaps and legal issues.

Big data management changes are directly related to document and information management transformations where the challenges, be it data security or lack of sufficient competencies, serve as a driver for operational improvements.

Big Data management practices are linked inextricably with Technology and Tools, each influencing the other. Good management practices depend upon the use of advanced technologies like cloud computing, AI, and data analytics, which in turn evolve according to needs and practices for Big Data management.

Finally, Technology and Tools have emerged as strong enablers in conducting Transformations in the Document and Information Management area-new tools enable institutions to track performance, manage data effectively, and decide on real-time insights. This seamless relationship between technology and transformation reflects the important role of innovation in adapting Big Data challenges.

Discussion

The awareness and implications of Big Data in Omani institutions, as derived from various respondents, provide a comprehensive perspective encompassing definitions, applications, challenges, and the evolving landscape of document and information management. The essence of Big Data is recognized as a massive aggregation of data from diverse sources, with notable characteristics such as substantial volume, high velocity, and significant variety, posing challenges to traditional data processing techniques. Respondents from various backgrounds highlight its scope, which includes data types like customer, market, financial, and operational data, among others. The linkage of Big Data with technologies like artificial intelligence and cloud computing underscores the technological sophistication required for its management.

The insights gleaned from our interviews align closely with prevailing definitions in the scholarly literature. For instance, Gartner (2012) characterizes Big Data as information assets characterized by substantial volume, high velocity, and/or significant variety. Similarly, TechAmerica (2012) depicts it as an overarching term encapsulating immense, high-velocity, complex, and diverse datasets, demanding advanced methodologies and technologies for their acquisition, storage, dissemination, management, and analysis. This conceptualization resonates with the definition put forth by the European Commission (2016), which portrays Big Data as substantial quantities of diverse data types sourced from various origins. Additionally, Kwanya (2014) posits that the essence of big data embodies the evolution of information systems facilitating the capture, processing, storage, and utilization of extensive and diversified datasets, thereby maximizing value generation for users.

The interviews reveal an intricate relationship between Big Data, artificial intelligence highlighting the technological advancements in this domain. However, challenges like data security, legal and regulatory constraints, and the need for skilled human resources in data science and programming are also prominent. A robust technology infrastructure supporting Big Data initiatives is vital, encompassing high-performance storage systems, analytics platforms, intuitive data visualization tools, and secure network connectivity. Policies and procedures ensuring data integrity, security, and confidentiality throughout its lifecycle are also crucial. This comes as an outcome of the diverse array of datasets within organizations, encompassing customer data, financial records, technical specifications, and sector-specific data. This diversity underscores the vast scale and intricate complexity of data traversing organizational frameworks, necessitating a holistic approach to its management and analysis. The magnitude of this data deluge presents a significant challenge for institutions, both public and private, in their administrative endeavors, specifically in effectively managing these voluminous datasets. To address this challenge, institutions must make strategic investments in appropriate technologies and infrastructure, including distributed database management systems, cloud storage solutions, advanced analytics tools, and robust information security protocols. In line with this, Teboul (2021), has noted that data serves as the bedrock of artificial intelligence, constituting the fundamental raw material that nourishes these systems, equipping them with the requisite data for autonomous learning, comprehension, and decision-making.

Omani institutions have implemented various technological solutions to manage the vast volumes of data generated by modern organizations. These institutions rely on cloud computing, AI, and specialized big data tools such as Hadoop for managing and analyzing data. However, the integration of these technologies remains challenging, especially when balancing innovation with the need for data security and compliance with national laws. This reflects a broader global

trend where archives worldwide must embrace emerging technologies to manage the growing complexities of digital data. The challenges in managing, securing, and processing Big Data necessitate specialized expertise and technological infrastructure. The integration of Big Data with existing systems demands a renovation of infrastructure to accommodate its scale and complexity. This necessitates an overall development of existing infrastructure, where big data tools come into play. Technologies like Hadoop and cloud computing offer necessary scalability and processing power, prompting a rethink in data storage, management, and security approaches. This is a challenge faced by archival institutions worldwide as they grapple with incorporating emerging technologies into traditional archival frameworks. The findings from Oman emphasize the need for scalable solutions that are adaptable to diverse institutional settings. Cloud storage, for example, offers a universally applicable solution for institutions handling vast datasets, while AI enhances data analysis capabilities, applicable in various archival settings globally.

Utilizing Big Data effectively transforms organizational decision-making by enabling the analysis of large, complex datasets, uncovering previously inaccessible insights. This capability is instrumental in identifying opportunities, weaknesses, and crafting suitable solutions. Additionally, it allows for the provision of better services and enhances competitive advantages. This strategic use of data is particularly relevant in sectors like risk management and security, where informed decisions can have significant implications. In same context, Philip-chen and Zhang (2014) and Hamad et al. (2022) have pointed out that the effective use of big data should transform economies presenting productive growth. In addition, with the growing adoption of cloud storage, document retention becomes more flexible and diverse. Looking forward, the introduction of an electronic document management system is set to further streamline these processes. This lined up with other researchers (Philip-chen & Zhang, 2016, Federer, 2016, Fister, 2015) as they all affirmed the use of data to provide institutional decision-makers with strategic information related to investments and operational enhancement.

Moreover, the handling of big data introduces additional layers of complexity, requiring specialized tools for storage, analysis, filtering, and visualization, ensuring that large volumes of data are managed efficiently and effectively. This comprehensive approach marks a significant shift towards more technologically advanced, secure, and efficient document management strategies. Specialized staff who use modern technology, including sophisticated software and secure servers, facilitate this process. The collaboration between documentation experts and IT professionals is crucial for effective document lifecycle management. As such, it becomes evident that a nuanced understanding of Big Data is indispensable for unlocking its transformative potential and deriving actionable insights. Amazon Web Services (2017), has pointed out that in today's fiercely competitive marketplace, businesses must leverage data-driven insights to make informed decisions that drive tangible outcomes, whether it be augmenting revenue streams, enhancing customer retention rates, or optimizing product quality. In this context, Big Data emerges as a pivotal enabler, playing a central role in achieving these strategic objectives by furnishing organizations with actionable intelligence gleaned from vast reservoirs of data (Hamad et al., 2021).

A common challenge highlighted by Omani institutions is the shortage of skilled professionals capable of managing big data effectively. The findings further emphasize the need for the increased awareness, specialized training, and comprehensive strategies for leveraging Big Data in operational and decision-making processes. Traditionally data security has primarily revolved around three fundamental pillars, delineated in descending order of importance as

confidentiality, integrity, and availability. Researchers asserted that privacy is one of the most challenging aspect of big data (Sudarsan, 2015; Borodo et al., 2016, Wang, et. al, 2016, Talha, 2022; Hamad et al, 2021).

The respondents envision a future where document and information management is increasingly important, requiring adaptation to rapid information technology developments. Continuous education and training in data analysis and management, protection of primary data sources, and adoption of advanced protection systems are emphasized as crucial for securing Big Data. Proficiency in data analysis, data science, programming, and project management is indispensable for maximizing the benefits of Big Data initiatives. Consequently, institutions must invest in training and upskilling their workforce, in addition to recruiting specialized talent capable of adeptly designing, implementing, and managing Big Data projects. Respondents emphasized the need for continuous training in data science, AI, and digital preservation techniques. This issue is not unique to Oman, as the global archives and records management profession also faces a growing demand for archivists with expertise in big data, cybersecurity, and information technology. Oman's experience highlights the need for continuous professional development in this field.

In Oman, archival practices are governed by national legislation, including the Omani Documents and Archives Law, which sets the guidelines for records and data management. Like many countries, Oman faces the difficulty of updating legal frameworks to accommodate the complexities of big data. Globally, archival institutions encounter similar challenges, where outdated legal frameworks struggle to keep up with the rapid pace of technological change.

This comprehensive understanding of Big Data management in Omani institutions highlights systemic approaches for integration and optimization, emphasizing interconnected data systems for efficiency and accuracy. The vision for document and information management in the era of Big Data is forward-thinking, recognizing the need for the sector to evolve alongside technological advancements and stressing the importance of upskilling in data analysis and management. Apparently, Big Data in Omani institutions presents a multi-dimensional impact, necessitating technological proficiency, security measures, and specialized training for efficient utilization and to harness its potential in shaping organizational strategies and public policies.

Establishing a Universal Framework of Data Governance and Digital Preservation

Global Relevance of Big Data Challenges

The challenges Omani institutions face in managing big data—such as technological adaptation, data governance, and skills gaps—reflect issues encountered by archival institutions worldwide. This underscores the need for a universal framework that can be tailored to diverse institutional contexts. Based on the insights from Oman, we propose a framework that can guide global archival and records management practices.

Key Components of the Universal Framework

A. Technological Infrastructure Oman's experience highlights the necessity of investing in scalable technological infrastructure. Globally, institutions should prioritize:

- **Cloud Solutions** for long-term data preservation, ensuring secure and accessible storage.

- **Blockchain** for data integrity, providing decentralized, secure records management.

B. Legal and Ethical Considerations Aligning legal frameworks with technological advancements is critical for effective big data management. Globally:

- **International Standards** like GDPR and ISO 15489 should be adopted to ensure legal compliance across institutions.

- **Ethical Guidelines** must balance data accessibility with privacy, particularly when handling personal information.

C. Digital Preservation Strategies Long-term digital preservation is essential for sustaining archival records:

- **Adoption of Models** like OAIS ensures the accessibility and authenticity of digital records.

- **Sustainable Strategies** such as regular data migration and distributed storage systems are vital for preventing data obsolescence.

D. Professional Development Oman's skills gap reflects a broader global issue. Institutions worldwide should focus on:

- **Continuous Professional Development** in digital preservation, cybersecurity, and data analytics.

- **Interdisciplinary Collaboration** between archivists, IT professionals, and legal experts to strengthen big data governance.

Global Adaptation of the Framework

This framework is adaptable to various institutional and national contexts. While Oman provides a specific example, its principles—technological investment, legal and ethical alignment, sustainable preservation, and capacity building—can be universally applied. Emerging economies can follow Oman's lead by focusing on scalable solutions, while more developed economies may invest in advanced digital preservation technologies.

Implications for Global Archives and Records Management

The Omani case illustrates both unique and shared challenges in big data management. By applying these lessons, global institutions can build robust systems that prioritize scalability, compliance, and professional development, enabling them to manage digital records more effectively.

While our study yields valuable insights into the realm of big data in Oman, acknowledging the awareness among archivists and managers regarding the significance and challenges posed by this evolving scenario, there exist certain limitations. The constrained sample size may not adequately capture the diverse array of perspectives and experiences within the realm of Big Data expertise. Additionally, the interviews were conducted within specific sectors and organizations, implying potential limitations in generalizing the findings to other domains. Nevertheless, our study furnishes crucial information on the practicalities of big data in Oman, shedding light on the awareness levels of archivists and managers concerning the implications and challenges arising in this transformative landscape. These insights form a compelling foundation for prospective research endeavors, offering avenues for further exploration and contributing to the evolving discourse on digital preservation and information management.

Research recommendations

Oman as a Microcosm of Global Big Data Challenges

The challenges faced by Omani institutions mirror those faced by archives and records management professionals globally. The rapid growth of digital data, the need for scalable technological solutions, and the integration of big data into archival frameworks are universal issues that extend far beyond Oman. By focusing on Oman as a case study, this research sheds light on how institutions in various countries can approach similar challenges, providing valuable insights for the global archival community.

Lessons for Global Archives and Records Management

The findings from Oman offer several key lessons for global archives and records management professionals:

1.   **Strategic Technological Investments**: Omani institutions' use of cloud computing and AI highlights the importance of making strategic investments in technology that can support big data management. This is a lesson that can be applied globally, where institutions are increasingly reliant on scalable digital solutions to handle growing data volumes.

2.   **Legal Frameworks for Big Data**: The experience of Omani institutions in navigating outdated legal frameworks reflects a common issue faced by archival institutions globally. As big data continues to evolve, there is a need for legal systems worldwide to adapt, ensuring that records and data management practices remain compliant with modern technological realities.

3.   **Addressing the Skills Gap**: Oman's shortage of trained professionals in big data management emphasizes the need for a global response to the skills gap in the archives profession. Continuous professional development, training in digital tools, and a focus on data governance are essential for building a workforce capable of meeting the demands of big data in records management.

Conclusion

The case study of Omani institutions provides valuable insights into the broader challenges of big data management in archives and records management. Although the findings are rooted in the Omani context, the lessons learned have universal applicability. Institutions worldwide face similar challenges, from technological adaptation to legal compliance and the skills gap in managing digital data. By examining the Omani experience, this study offers practical solutions that can be applied to archives and records management systems across different cultural and institutional contexts, contributing to a more comprehensive understanding of big data management globally. The article provides a detailed overview of the Big Data landscape within Omani institutions, elucidating the diverse sources, management strategies, and challenges associated with its adoption and utilization. Various data sources, spanning administrative records to social media platforms, not only serve as invaluable repositories but also fuel innovation and decision-making processes across sectors. Interviewees' understanding of Big Data, fostered through formal training and professional development, underscores its growing significance in shaping organizational strategies. Despite its transformative potential, challenges such as legal constraints and resource scarcity necessitate strategic interventions for effective management. The study underscores a notable shift towards leveraging Big Data to enhance operational efficiencies, decision-making, and policy development within Omani organizations. Integration into document and information management systems has revolutionized operations,

leading to increased efficiency and better service delivery. However, challenges persist, including the need for new methodologies, infrastructure investments, and training to manage Big Data effectively. Looking ahead, optimism abounds regarding Big Data's potential to further transform organizational decision-making processes, necessitating investments in infrastructure and skill development. In conclusion, the article emphasizes the importance of a proactive approach to Big Data, advocating for strategic and holistic strategies to harness its potential fully. By doing so, organizations can bolster competitiveness, foster innovation, and thrive in an increasingly data-centric landscape, ensuring sustained success in the future.

This study outlines a universal framework for data governance and digital preservation, drawing on the Omani case to inform global archival practices. By integrating technological solutions, legal frameworks, ethical guidelines, and professional development, archives worldwide can enhance their capacity to manage big data securely and sustainably, ensuring long-term access and relevance in the digital age.

References

Amazon Web Services (2017). LE BIG DATA DANS LE CLOUD Comment relever les cinq principaux défis du Big Data – Tout sur le cloud. Online : https://www.infomania-services.fr/controle/file/180831085754000000_4466289594_1535705874.pdf

Banat-Berger, F. (2010). Les archives et la révolution numérique. Le Débat, 158, 70-82. https://doi.org/10.3917/deba.158.0070

Ben Amor, F.; Mkadmi, A. (2018). Archives in the Age of Big Data: the challenges of archiving Big digital data. Proceedings of Digital Tools and Uses Congress (DTUC'18), Paris, France, 7 pages. https://doi.org/10.1145/3240117.3240139

Big Data Senior Steering Group. (2016). The federal big data research and development strategic plan. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1019&context=scholcom

Billenness, Clive (2016). An introduction to the E-ARK General Model. In: E-ARK Conference 6 december 2016, Budapest. [online]: https://www.eark-project.com/resources/conference-presentations/finconfpres.html

Borgman, C. L. (Ed.). (2015). Big Data, Little Data, No Data: Scholarship in the Networked World. Cambridge, MA: MIT Press.

Borodo, S. M., Shamsuddin, S. M., & Hasan, S. (2016). Big Data Platforms and Techniques. Indonesian Journal of Electrical Engineering and Computer Science, 1(1), 191-200.

Commission E. The EU Data Protection Reform and Big Data: Factsheet 2016 [https://publications.europa.eu/en/publication-detail/-/publication/51fc3ba6-e601-11e7-9749-01aa75ed71a1.

Dell'Orto, C., & Caviglia, G. (Eds.). (2018). Big Data in the Humanities and Social Sciences: Opportunities and Challenges. Cham, Switzerland: Springer.

Diebold, F. X. (2003, February). Big data dynamic factor models for macroeconomic measurement and forecasting. In Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress of the Econometric Society,"(edited by M. Dewatripont, LP Hansen and S. Turnovsky) (pp. 115-122).

Ducellier, Philippe (2017). DOCAPOST complète la Blockchain pour l'adapter à l'archivage de documents. Online: https://www.lemagit.fr/actualites/450426156/DOCAPOST-complete-la-Blockchain-pour-ladapter-a-larchivage-de-documents

Federer, L. (2016). Research data management in the age of big data: Roles and opportunities for librarians. Information Services & Use, 36(1-2), 35-43.

Felton, A. (2017). Le potentiel des technologies du web sémantique pour l'archivage dans un contexte de gouvernance de l'information : le cas du Système d'Archivage Electronique Mutualisé (SAEM) girondin. Mémoire pour l'obtention du Master esDOC, Sous la direction de David Guillemin, Université de Poitiers. Online : http://nuxeo.edel.univ-poitiers.fr/nuxeo/site/esupversions/458d0df5-3904-476e-9425-8ac5e7af4a59

Fister, B. (2015). Big Data or Big Brother? Data, Ethics, and Academic Libraries. Reftrieved 15-12-2017 from: http://barbarafister.com/LIbigdata.pdf.

Forbes (2018). Big Data : Le Milieu Académique Est-Il Encore Pertinent ? online : https://www.forbes.fr/technologie/big-data-academie-pertinence/

Gartner, E.S. (2012). 10 Critical Tech Trends for the Next Five Years, October 2012. http://www.forbes.com/sites/ericsavitz/2012/10/22/gartner-10-critical-tech-trends-for-the-next-five-years/. In: Philip Chen, C. L., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information Sciences, 275, 314-347.

Hamad, F., Al-Aamr, R., Jabbar, S. A., & Fakhuri, H. (2021). Business intelligence in academic libraries in Jordan: Opportunities and challenges. IFLA journal, 47(1), 37-50.

Hamad, F., Fakhuri, H., & Abdel Jabbar, S. (2022). Big data opportunities and challenges for analytics strategies in Jordanian Academic Libraries. New Review of Academic Librarianship, 28(1), 37-60.

Hightower, M. A., & Dettwyler, K. L. (Eds.). (2016). Big Data in Libraries, Archives, and Museums. Chicago: Association of College and Research Libraries.

Höchtl, J., Parycek, P., & Schöllhammer, R. (2016). Big data in the policy cycle: Policy decision making in the digital era. Journal of Organizational Computing and Electronic Commerce, 26(1-2), 147-169.

Hodgson, J. P. E., & Marty, P. F. (Eds.). (2018). Big Data, Big Challenges: Data Management in Libraries, Archives, and Museums. Chicago: Association of College and Research Libraries.

INRIA (2014). Big data: nouvelles partitions de l'information actes du séminaire IST Inria, octobre 2014. Louvain-la-Neuve [Paris] : De Boeck ADBS. ISBN 978-2804189150.

ISO 15489. (2016). International Organization for Standardization. Information and Documentation — Records Management.

Collomosse, T. Bui, A. Brown, J. Sheridan, A. Green, M. Bell, and J. Fawcett, J. Higgins O. Thereaux (2010). ARCHANGEL: Trusted Archives of Digital Public Documents. In Proceedings of ACM Conference (Submitted to ACM DocEng'18). ACM, New York, NY, USA, Article 39, 4 pages. https://doi.org/ 10.1145/

Kwanya, T. (2014, September). Big data in land records management in Kenya: a fit and viability analysis. In International Conference on Knowledge Management in Organizations (pp. 15-24). Springer, Cham.

Marone, Reine Marie Ndéla & Mbengue, Moustapha (2021). L'archivage à l'ère des Big Data : quels nouveaux paradigmes pour l'Archive et quels rôles pour l'archiviste ? Revue maghrébine de documentation et d'information, n°30, pp.45-67

Mkadmi, A. (2021). Archives in the digital age: preservation and the right to be forgotten. wiley. https://iste.co.uk/book.php?id=1763

Nougaret, Christine (2017). Une stratégie nationale pour la collecte et l'accès aux archives à l'ère numérique. FranceArchives [en ligne]. [Consulté le 31 juillet 2017]. Disponible à l'adresse : https://francearchives.fr/fr/article/28204701

Philip Chen, C. L., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. Information Sciences, 275, 314-347.

SENN, A. (2016). Enabling 21st century archives. Paths and traces. In: E-ARK Conference 6 december 2016, Budapest. [online]: https://www.eark-project.com/resources/conference-

presentations/finconfpres.html

Servais, Paul et Mirguet, Françoise (2015). L'archive dans quinze ans : vers de nouveaux fondements. Louvain-la-Neuve : Academia-l'Harmattan. Publications des Archives de l'Université catholique de Louvain, 33. ISBN 978-2-8061-0225-6. Z, page 250

Sudarsan, Sithu & Jetley, Raoul & Ramaswamy, Srini. (2015). Security and Privacy of Big Data. In book: Big Data (ISBN: 978-81-322-2493-8. pp.121-136). DOI: 10.1007/978-81-322-2494-5_5.Talha, Mohamed (2022). Big Data entre Qualité & Sécurité de données. Informatique [cs]. Université Cadi Ayyad (Marrakech, Maroc), 2022. Français. ffNNT : ff. fftel-04011007

Teboul, B. (2021). Les enjeux de la convergence de la Data, de l'IA, du Cloud, de la Blockchain, de l'IoT et de la Cybersécurité. Online: https://www.europeanscientist.com/fr/opinion/les-enjeux-de-la-convergence-de-la-data-de-lia-du-cloud-de-la-blockchain-de-liot-et-de-la-cybersecurite/

TechAmerica Foundation: Federal Big Data Commission. (2012). Demystifying Big Data: A Practical Guide To Transforming The Business of Government. Retrieved 15/11/127, from: https://bigdatawg.nist.gov/_uploadfiles/M0068_v1_3903747095.pdf

Vieira, A.. (2018). Archiving Big Data: A Challenge for Today's Archivists. Journal of Digital Preservation.

Wang, C., Xu, S., Chen, L., & Chen, X. (2016, June). Exposing library data with big data technology: A review. In Computer and Information Science (ICIS), 2016 IEEE/ACIS 15th International Conference on (pp. 1-6). IEEE.

Wang, C., Xu, S., Chen, L., & Chen, X. (2016, June). Exposing library data with big data technology: A review. In Computer and Information Science (ICIS), 2016 IEEE/ACIS 15th International Conference on (pp. 1-6). IEEE.

Xu, W., Esteva, M., Jain, S. D., & Jain, V. (2011, October). Analysis of large digital collections with interactive visualization. In 2011 IEEE Conference on Visual Analytics Science and Technology (VAST) (pp. 241-250). IEEE.

Appendix:

Interview guide

Within the scope of research activities in the Department of Information Studies at the College of Arts and Social Sciences at Sultan Qaboos University, we are conducting a study on the **Reality of big data management in Omani institutions from the perspective of information and records specialists.** Therefore, we kindly ask you to cooperate with us by answering some questions, noting that the data will only be used for scientific research purposes. We thank you for your kind cooperation with us.

1. General Data

Date: .....

Interview Place: ..... Interview Duration: .....

Name of the institution: .....

Date of establishment: ....

Areas of expertise .....

Respondent's name: ..... Position: .....

2. Interview topics:

Axis 1: Identifying the extent of awareness of the organization's employees of the term Big Data

- What are the types of data circulating in the organization?

- Have you heard of the term Big Data before?

- Do you have knowledge of cloud computing? Hadoop? Blockchain? Artificial intelligence? Data Lake, etc.?

- What challenges do you think it poses at the organization level?

- What, in your opinion, are the methods and strategies that must be available in the organization to manage and benefit from Big Data?

Axis Two: The Reality of Big Data and Its Uses in Omani Institutions

- What are the sources of big data?

- What are the equipment and software used in processing and analyzing big data?

- What are the fields and activities that provide this big data among the activities carried out by the institution?

- How is big data managed and processed?

- What are the methods of making big data available and how are they invested?

- What are the methods you follow to store data?

Axis Three: The Impact of Big Data on Activities Related to Records and Information Management

- Does the Document and Information Management System in the institution include big data management?

- What are the most important transformations brought about by big data at the level of information and document management?

- How is this data processed for archival purposes and how is it saved and stored?

- How do you manage the availability of this data according to the legal texts related to document management?

Axis Four: The visions proposed by Omani institutions to rationalize big data management from the perspective of the specialization of Records and information management?

- How, in your opinion, can big data be managed to exploit it optimally and in a way that serves the objectives of the institution?

- In your opinion, what investments should be made in the organization to rationalize the management and exploitation of this data?

- What is the future of the document and information management sector in the era of big data?

- What is your vision of the education and training requirements in the information and documents sector in light of the changing data management requirements?

- What are the ways to protect and secure big data?