



Sobia Fatima¹, Abdul Karim Jamali² & Razia Abbas³

¹MPhil Scholar, Department of Information Management, University of Sargodha Pakistan.

Email: Sobiafatima1911@gmail.com

²Librarian, Topian Spirit Foundation Academic Campus Rashida bad, Tando Allahyar.

Email: akjamalibaloch203@gmail.com

³Student, University of Karachi, Karachi. Email: raziaabbaxx@gmail.com

KEYWORDS	ABSTRACT
Big Data, information services, digital libraries, user experience, data ethics, knowledge management	The rapid development of digital technologies has led to a proliferation of large, varied datasets, transforming how information services are created and delivered. This article explores the significance of Big Data in knowledge ecosystems, particularly the 5Vs: Volume, Velocity, Variety, Veracity, and Value, in libraries, archives, and knowledge management. As data volume and variety grow, concerns about data quality, management, and ethical governance also increase. Predictive analytics, data-driven learning, and real-time processing are becoming more common in information institutions to enhance service delivery, such as through personalized recommendations and more flexible resource allocation. The study is based on a literature review, recent case studies, and quantitative findings from survey results, and looks at how information is being transformed from Information Management and Knowledge Management to data-driven information and data-driven decision-making. It also considers concerns related to algorithmic bias and privacy. The results indicated that the effectiveness of using Big Data in information services depends not only on the infrastructure to support it, but also on data quality and institutions' ability to convert available data into useful information for decision-making and service enhancement.
ARTICLE HISTORY Date of Submission: 23-02-2026 Date of Acceptance: 27-03-2026 Date of Publication: 31-03-2026	
	 2026 Journal of Social Sciences Development
Corresponding Author	Sobia Fatima
Email	Sobiafatima1911@gmail.com
DOI	https://doi.org/10.59075/jssd.v5i1.398
Volume/Issue	5(1), 2026, 372-391

Introduction

The nature of generated data is continuously changing, as are the ways organizations collect, manage, and use it. The growing scale and changing patterns of data have made Big Data an important part of modern information systems. This development is commonly explained through the three main characteristics of Big Data: volume, velocity, and veracity, which represent the increasing amount of data, the speed at which it is generated and processed, and the reliability or accuracy of the information being handled (Kitchin, 2014).

In today's digital world, as organizations increasingly use data to transform decisions, operations, and strategies, Big Data analytics has become a key component of corporate decision-making, operational improvement, and strategic innovation (Chen et al., 2012).

The impact is especially noticeable in terms of information services. They are no longer just repositories of digital libraries or archives, knowledge management systems, or online information platforms. They operate in a networked, digital environment where users create, share, and use information (Borgman, 2015). Information services can also offer real-time, personalized, and predictive solutions as machine learning, artificial intelligence, and cloud-based infrastructure expand their capabilities.

However, the explosion of data-driven information services raises concerns beyond technology. Privacy, algorithmic bias, surveillance, and unequal access to digital resources are also important factors in the evolution of sustainable digital ecosystems (Boyd & Crawford, 2012). This increased reliance on data requires responsible information handling, transparency, and accountability. While much has been written on Big Data, fewer empirical studies have examined the combined impacts of the five dimensions within a single framework for measuring an information service's performance. This study seeks to fill this gap by proposing and testing the relationship between Big Data capabilities and Information Services Performance.

Literature Review

Volume in Information Services

Digital information has changed how data is produced, stored, managed, and accessed. Big Data is a key element in information services today, as organizations produce and gather information on a much larger scale. This expansion opens new opportunities for libraries and information centers but also poses challenges for storage, organization, preservation, and accessibility. To manage digital resources and to ensure reliable, accessible information for users, it is therefore important to clearly understand the characteristics of Big Data (volume and variety).

Today's digital collections, institutional repositories, and online databases have grown rapidly because to the abundance of data produced. Academic libraries and information centers can process terabytes, or even petabytes, of information, such as research publications, multimedia resources, data sets, and user-generated content (Gandomi & Haider, 2015). Processing these large

collections requires adequate storage, expansive databases, and systems that can distribute data processing across resources. Organizing, searching, and long-term maintenance of large digital collections can be increasingly challenging without adequate infrastructure and can impact the services provided to information.

Variety in Information Services.

Variety refers to the diversity of data formats and types. In the context of big data, this includes structured, unstructured, and semi-structured data. Examples include text, images, videos, sensor outputs, and more. This diversity creates a major challenge for information services, as heterogeneous data requires more complex storage, processing, integration, and retrieval strategies than homogeneous data (Salman et al., 2020). Library and information institutions need to adopt flexible data structures and metadata schemes that can effectively handle large collections of diverse data, allowing users to find, retrieve, and use information efficiently while correctly managing data variety. This often requires sophisticated computing tools and analysis methods to understand and identify unstructured or semi-structured content and to design information retrieval systems, user interfaces, and service strategies accordingly.

Velocity in Information Services

Velocity in Big Data is the speed at which data is generated, gathered, and analyzed. To facilitate early insight into the situation and early action. In a world of information technology and social media, high speeds require parallel and distributed processing. Today, data streams in at high speeds and continuously through parallel and distributed processing interactions, transactional systems, and sensor networks. At unprecedented speeds (Sagiroglu & Sinanc, 2013), high-velocity data streams need to be propagated, analyzed, and responded to in real time or near real time using information systems to keep them relevant and useful. Real-time processing enables companies to see patterns as they are happening. They can predict user needs and adjust their service dynamically. In information services, this capability has far-reaching implications: library systems can update catalog data in real time, recommendation systems can adapt to user activity, and digital platforms can provide instant alerts and interactive experiences (Hashem et al., 2015). For instance, real-time data received from users' queries or engagement indicators enables improvisations and adaptations to enhance usability and responsiveness. Effectively managing data velocity often requires stream-computing frameworks and technologies related to in-memory processing. These minimize data-handling latency (Grolinger et al., 2014).

Veracity in Information Services

Volume, variety, and velocity refer to the quantitative characteristics of data, while veracity refers to their qualitative characteristics and addresses uncertainty, reliability, and trustworthiness. In information services, veracity is of paramount importance when it comes to digital information. Libraries and knowledge repositories are sources of authoritative knowledge and research, so the

abundance of user-generated content, social feeds, and automatic reloads is a factor to consider. In digital data collection, significant noise, bias, and potential inaccuracies have been introduced due to data collection, where the effects of individuals can lead to changes in ecosystems (Janssen et al., 2017). To ensure your data is accurate, you will need to validate it thoroughly through processes, anomaly detection, and strong governance frameworks to weed out misinformation and maintain archival records. Recent evidence points to the importance of veracity not just as a technical data problem; it's a basic question of trust in an information retrieval system. As specialized AI search services become increasingly tool- and Large Language Model (LLM) driven, the risk of "hallucinated" or incorrect citations grows, and new layers of citation needs are added; verification is underway (Lund et al., 2023). Information professionals are now challenged to develop sophisticated, scalable quality assurance processes that can differentiate high-quality, scholarly publications from questionable publishers or gray literature that has not been subject to review.

Value in Information Services

The fifth V, Value, represents the ultimate objective of Big Data initiatives: the ability to extract meaningful insights that drive decision-making and improve service delivery. For information services, value comes not from accumulating data, but from analyzing it to enhance user experience, optimize collection development, and demonstrate return on investment (ROI) to stakeholders (Mikalef et al., 2018). By analyzing usage patterns, search behaviors, and citation networks, libraries can transition from intuition-based management to evidence-based strategies. Value creation manifests in cost-effective resource allocation, where predictive analytics help institutions acquire materials most likely to be used, reducing waste in acquisition budgets. Furthermore, the value extends to users through personalized discovery pathways that save research time and surface relevant, cross-disciplinary connections that traditional cataloging might miss (Affleck, 2021). Consequently, the "Value" dimension bridges the gap between raw technical infrastructure and tangible organizational outcomes.

Conceptual Framework

The conceptual model looks at how the five dimensions of Big Data (Volume, Variety, Velocity, Veracity, Value) relate to Information Services Performance. It considers how these dimensions can support effective, reliable, timely, accessible, and high-quality information services.

This framework is adopted from the Technology-Organization-Environment (TOE) model, which focuses on how technological capabilities, organizational factors, and the environment influence organizations' adoption and use of technology. Here, the five dimensions of Big Data are treated as technological capabilities that can contribute to Information Services Performance.

Conceptual Model

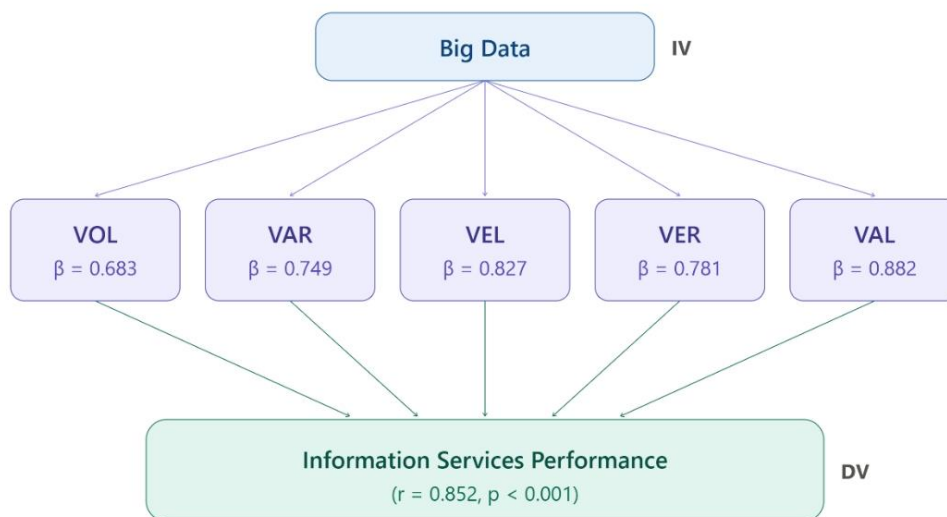


Figure 1. Conceptual Model illustrating the impact of 5Vs Big Data on Information Services Performance.

Research Objectives

The study examines the multidimensional impact of Big Data on information services performance. The specific objectives are:

- RO1. To examine the impact of Big Data on Information Services Performance.
- RO2. To examine information professionals' use of Big Data in information management.
- RO3. To assess applications of Big Data in information services for their future adoption.

Research Hypotheses

- H1. Big Data Volume has a significant relationship with information service quality.
- H2. Big Data Variety significantly affects information accessibility.
- H3. Real-time Data Velocity affects information service performance.
- H4. Big Data Veracity contributes to the reliability of information services.
- H5. Big Data Value extraction positively influences strategic decision-making.

Research Methodology

Research Design

This study used a quantitative, cross-sectional survey design to examine the relationships between Volume, Variety, Velocity, Veracity, Value, and information service performance. The study emphasizes information service providers, such as researchers, information professionals, and librarians.

This research adopted the TOE (Technology-Organization-Environment) framework and the Information Systems Success Model (DeLone & McLean, 2003), modified for the context of Big Data in information services.

Population and Sampling

The information professionals, librarians, and researchers were selected using convenience sampling; they are directly involved in providing information, managing information, or conducting information research and have relevant experience with the Big Data phenomenon. The method was designed to ensure respondents were accessible, willing to participate, and able to provide meaningful responses.

Research Instrument

The data were gathered using a structured questionnaire subdivided into different sections. Part A presented the participants' demographic information (gender, qualification, designation, and experience). The rest of the sections measured five aspects of Big Data – Volume, Variety, Velocity, Veracity, Value – and Information Services Performance.

All items were measured on a 5-point Likert scale, from 5 (Strongly Agree) to 1 (Strongly Disagree). The questionnaire items were designed based to the study's purpose and the literature on Big Data and information services.

Data Collection Method

A Google Forms online questionnaire was sent by e-mail to library and information professionals. The data collection period was fixed, and the responses were collected and analyzed in SPSS.

Variables of the Study

The study includes the following variables:

- **Independent Variables:** Big Data dimensions (Volume, Variety, Velocity, Veracity, Value)
- **Dependent Variable:** Information Services Performance

This study aims to examine how the variables, in terms of Big Data characteristics, relate to service effectiveness.

Data Analysis Techniques

The data were analyzed using SPSS (Statistical Package for the Social Sciences. Descriptive statistics such as means, frequencies, and standard deviations were used to report respondents' characteristics and their responses to the items. Inferential techniques included Pearson correlation and simple linear regression to analyze relationships among the five Big Data dimensions and Information Services Performance, as well as to test hypotheses H1 to H5.

Validity and Reliability

The study instrument showed very good internal consistency. The Cronbach's alpha value for the 25 items of the Big Data scale was 0.972, and 0.910 for the 5 items of Information Services Performance. Both values fall in the 'Excellent' category.

Table 1. Internal Consistency of Big Data and Information Services Measures

Construct	No. of Items	Cronbach's Alpha (α)	Reliability
Big Data	25	0.972	Excellent
Information Services Performance	5	0.910	Excellent

Results

Socio-Demographic Information

The demographic results indicate that the highest number of respondents (48.1%) were from the University/Academic Libraries, followed by Research Centers/Special Libraries (23.4%) and Public Libraries (20.8%). By designation, Librarians ranked highest (45.5%), followed by Information Professionals (29.9%) and Researchers (24.7%). Regarding experience, most respondents (45.5%) had 0-5 years, and 31.2% had 6-10 years. The largest institutional size category was small institutions (41.6%). Geographically, the largest share of respondents were from Punjab (28.6%), followed by Sindh (26.0%) and Islamabad Capital Territory (19.5%). The sample is broadly representative of the various institution types, roles and experience, and sizes and regions, enabling a wide range of views on Big Data and Information Services Performance.

Table 2. Demographic Characteristics of Respondents

Professional Profile	Frequency	Percent
Type of Institution		
University / Academic Library	37	48.1
National Library / Archive	6	7.8
Research Center / Special Library	18	23.4
Public Library	16	20.8
Designation		

	Librarian	35	45.5
	Information Professional	23	29.9
	Researcher	19	24.7
Years of Experience			
	0–5 years	35	45.5
	6–10 years	24	31.2
	11–15 years	13	16.9
	16–20 years	5	6.5
Size of Institution			
	Small (under 1,000 users)	32	41.6
	Medium (1,000–5,000 users)	26	33.8
	Large (5,000+ users)	19	24.7
Province/Region			
	Sindh	20	26.0
	Punjab	22	28.6
	Balochistan	8	10.4
	Khyber Pakhtunkhwa	12	15.6
	Islamabad Capital Territory	15	19.5

Descriptive Statistics

Volume

The highest mean within the Volume dimension was for the item “Organizing large datasets requires specialized skills” ($M = 3.78, SD = 1.008$). The second highest was for “Large datasets at my institute aid in analyzing research” ($M = 3.48, SD = 1.059$).

The mean of the item “I can easily retrieve relevant information from large datasets for service delivery” was 3.40 ($SD = 1.003$), while the mean of the item “My institution has enough storage to keep large amounts of data” was 3.39 ($SD = 1.053$). The lowest average was for “Maintaining backups for large datasets is challenging” ($M = 3.34, SD = 1.021$).

Variety

The highest mean score ($M = 3.66, SD = 0.954$) was for Multiple data sources are effectively combined to support research activities for Variety. The second was “Systems support the integration of diverse data types (e.g., text, images, and video)” ($M = 3.64, SD = 1.038$), and the third was “Data variety improves the coverage of information services” ($M = 3.62, SD = 0.946$).

The mean score for "Diverse data sources improve completeness of information services" was 3.51 (SD = 0.941), and for "Institutional tools enable efficient handling of diverse data formats" was 3.48 (SD = 1.008), the lowest mean score.

Velocity

"The highest-rated item for Velocity was Data processing helps identify emerging trends (M = 3.60, SD = 0.936). The second most frequent responses were for "Frequent data updates don't cause a problem with system performance" (M = 3.53, SD = 0.954) and "Newly added data files are correctly recorded in the system" (M = 3.51, SD = 1.034).

"The system supports near real-time data processing" had a mean of 3.45 (SD = 0.911). The lowest mean was for the statement "The system recommends things based on what the user does" (M = 3.40, SD = 1.067).

Veracity

Data accuracy is ensured through regular verification (M = 3.52, SD = 1.046), and historical data is documented with its sources to ensure accuracy (M = 3.52, SD = 0.982), which had the highest mean scores in the Veracity dimension.

The mean score of 'Staff are trained to uphold high standards in data management' was 3.40 (SD = 1.079). The lowest mean scores were obtained for 'Automated systems detect misleading information' (M = 3.39, SD = 0.989) and 'Outdated data is systematically excluded from use in organizational workflows' (M = 3.39, SD = 0.989).

Value

The item with the highest mean (M = 3.65, SD = 0.914) in the Value dimension was "Data analysis supports effective resource allocation and budgeting." "Big data is used to show value and impact to stakeholders" scored a mean of 3.60 (SD = 1.016), and "Analytics are used to gain actionable insight from available data" scored a mean of 3.52 (SD = 0.968).

The mean score for 'Predictive analytics informs decisions on collection development' was 3.51 (SD = 0.927). The lowest mean was found for "Data analytics enables the discovery of the underlying patterns in a set of data" (M = 3.40, SD = 0.892).

Overall, the results suggest that data analytics is perceived as valuable for resource allocation, budgeting, and demonstrating organizational impact.

Big Data Practices

The responses to the five items related to usage were generally moderate to positive. The mean score of “Big Data tools efficiently access research-relevant information” was 3.34 (SD = 1.021). The mean of “Services meet or exceed expectations” was 3.22 (SD = 0.883), and “Systems adapt to users’ workflows” was 3.27 (SD = 0.955).

The mean for “Users trust the accuracy and reliability of data” was 3.27 (SD = 0.927), and for “Personalized features enhance overall experience”, the highest mean was found in this group at 3.42 (SD = 1.018).

Information Service Performance

The highest mean (M = 3.51, SD = 0.995) was found for Information Services Performance, “User engagement has increased (e.g., visits, downloads, citations). The mean for “Services have become more innovative” was 3.42 with an SD of 1.056, and the mean for “Big Data capabilities have improved resource utilization and collection development” was 3.42 with an SD of 1.128.

The mean score for “Strategic decisions are informed by knowledge of data” was 3.34 (SD = 1.046). The lowest mean was found to be “Operational efficiency has improved through Big Data adoption” (M = 3.27, SD = 1.008).

Table 3. Descriptive Statistical Analysis of Big Data and Information Services

Construct	Item	Mean	SD	Interpretation
Volume (VOL)	My institution has enough storage to keep large amounts of data	3.39	1.053	Moderate
	I can easily retrieve relevant information from large datasets for service delivery	3.40	1.003	Moderate
	Maintaining backups for large datasets is challenging	3.34	1.021	Moderate
	Organizing large datasets requires specialized skills	3.78	1.008	High
	Large datasets at my institution contribute to research analysis	3.48	1.059	High
Variety (VAR)	Multiple data sources are effectively combined to support research activities	3.66	0.954	High
	Systems support the integration of diverse data types (e.g., text, images, and video)	3.64	1.038	High
	Institutional tools enable efficient handling of diverse data formats	3.62	0.946	High

Velocity (VEL)	Diverse data sources improve the completeness of information services	3.51	0.941	High
	Data variety improves the coverage of information services	3.48	1.008	High
	The system supports near-real-time data processing	3.60	0.936	High
	Newly added data is accurately recorded in the system	3.53	0.954	High
	The system provides recommendations based on user behavior	3.51	1.034	High
	Data processing helps identify emerging trends	3.45	0.911	High
Veracity (VER)	System performance is maintained during frequent data updates	3.40	1.067	Moderate
	Data accuracy is ensured through regular verification	3.52	1.046	High
	Automated systems detect misleading information	3.52	0.982	High
	Staff are trained to uphold high standards in data management	3.40	1.079	Moderate
	Outdated data is systematically excluded from use in organizational workflows	3.39	0.989	Moderate
	Historical data is documented with its sources to ensure accuracy	3.39	0.989	Moderate
Value (VAL)	Analytics are used to derive actionable insights from available data	3.65	0.914	High
	Predictive analytics inform decisions on collection development	3.60	1.016	High
	Data analysis supports effective resource allocation and budgeting	3.52	0.968	High
	Big data is utilized to demonstrate value and impact to stakeholders	3.51	0.927	High
	Data analytics supports the identification of underlying patterns within datasets	3.40	0.892	Moderate
	Big Data tools help me efficiently access research-relevant information	3.42	1.018	High
Big Data Practices				

Information Performance (ISP)	Services	The services meet or exceed my expectations	3.34	1.021	Moderate
		The system adapts well to my workflow and tasks	3.27	0.955	Moderate
		I trust the accuracy and reliability of the data provided	3.27	0.927	Moderate
		Personalized features enhance my overall experience	3.22	0.883	Moderate
		Operational efficiency has improved through Big Data adoption	3.51	0.995	High
		User engagement has increased (e.g., visits, downloads, citations)	3.42	1.056	High
		Strategic decisions are better informed by data-driven insights	3.42	1.128	High
		Services have become more innovative	3.34	1.046	Moderate
		Big Data capabilities have improved resource utilization and collection development	3.27	1.008	Moderate

Correlation

To examine the relationship between Big Data and Information Services Performance, we used Pearson correlation analysis. A Pearson correlation coefficient of $r = 0.852$, $p < 0.001$, and $N = 77$ was obtained from the analysis.

The result shows a strong, positive relationship between the two variables, Big Data and Information Services Performance. That is, an increase in Big Data adoption is correlated with an increase in Information Services Performance. The value of 0.852 represents a very strong linear relationship between the two variables.

Table 4. Correlation between Big Data and Information Services

Variables	BIGDATA	ISP
BIG DATA	Pearson Correlation = 1	
ISP	Pearson Correlation = .852**	1

Correlation is significant at the 0.01 level (2-tailed).

The 5 dimensions also showed strong positive correlations among themselves and with Information Services Performance. All mentioned correlations are statistically significant ($p < 0.001$). The strongest correlation with Information Services Performance is with Value ($r = 0.882$), indicating a

strong association between the capacity to gain value from Big Data and information service performance.

Table 5. Correlation between 5Vs of Big Data and Information Services

Constructs	VOL	VAR	VEL	VER	VAL	ISP
VOL	1					
VAR	.861**	1				
VEL	.748**	.866**	1			
VER	.676**	.791**	.820**	1		
VAL	.746**	.828**	.892**	.831**	1	
ISP	.683**	.749**	.827**	.781**	.882**	1

Correlation is significant at the 0.01 level (2-tailed).

Hypotheses

The regression analysis found all five dimensions of Big Data (Volume, Variety, Velocity, Veracity, and Value) to be positive and statistically significant. The results of the associations between the variables showed that Variety was a significant positive predictor of Information Services Performance ($\beta = 0.749$, $t = 9.795$, $p < 0.001$), and Volume was also a significant positive predictor of Information Services Performance ($\beta = 0.683$, $t = 8.091$, $p < 0.001$). Velocity exhibited a strong positive effect ($\beta = 0.827$, $t = 12.715$, $p < 0.001$), followed by Veracity ($\beta = 0.781$, $t = 10.841$, $p < 0.001$). Value ($\beta = 0.882$, $t = 16.200$, $p < 0.001$) was the most positive contributor of all five dimensions. The results show that all of the five dimensions of Big Data have statistically significant positive influences on Information Services Performance, and Value is the most significant dimension. Thus, the hypothesis is confirmed.

Table 6. Empirical Findings for the Proposed Hypotheses

Big Data	B	Std. Error	β (Beta)	t	Sig.	Decision
Volume (VOL)	0.733	0.091	0.683	8.091	<0.001	Supported
Variety (VAR)	0.770	0.079	0.749	9.795	<0.001	Supported
Velocity (VEL)	0.921	0.072	0.827	12.715	<0.001	Supported
Veracity (VER)	0.831	0.077	0.781	10.841	<0.001	Supported
Value (VAL)	0.974	0.060	0.882	16.200	<0.001	Supported

Research Hypotheses

Based on the conceptual framework and literature review, the following hypotheses are proposed:

H1: Big Data Volume has a significant relationship with information service quality.

Theoretical Support: (Elgendy & Elragal, 2014) provide a theoretical overview of data volume, the vast amount of data created in scientific and educational pursuits, business operations, and human interactions. They note that as data grows, data storage and processing become increasingly important. With the advent of new storage technologies and decreasing storage costs, it is becoming easier to store data but still difficult to process huge amounts of data. In addition, (Geczy, 2014) notes that as data expands by orders of magnitude, data storage systems must become more efficient and cloud-based solutions must be used for larger data sets. Yet, as mentioned, (C. L. P. Chen & Zhang, 2014) data volumes are still rising and are outpacing the growth of computational processing power needed for data processing, management, and decision-making.

H2: Big Data variety significantly affects information accessibility.

Theoretical Support: According to (Geczy, 2014) Theoretical support for data variety is the different forms and levels of organization in which data exists. Structured data is organized in a specific format and can be stored and processed effectively in traditional databases, while semi-structured and unstructured data are less organized and harder to process and analyze. Structured Data: Traditional Relational Database Management Systems (RDBMS) are generally used for structured data and require expensive infrastructure. With the rise in data diversity, Big Data has brought new challenges to traditional data management systems, as it has become increasingly difficult to process information from various sources—including emails, mobiles, and distributed systems—and requires more sophisticated tools to handle it (C. L. P. Chen & Zhang, 2014).

H3: Real-time data velocity affects information service performance.

Theoretical Support: (Salehian & Yan, 2016) point out that Big Data has three main characteristics, namely: volume – how much data there is, velocity – the speed at which the data is generated, transmitted, and processed, and variety – the diversity of data types. They highlight the need for batch, near-real-time, and real-time processing to meet urgency needs at different levels. Data velocity, crucial for effective and responsive information services, enables faster processing and helps meet time-sensitive requirements.

H4: Big Data Veracity contributes to the reliability of information services.

Theoretical Support: (Mills et al., 2012) refer to big data veracity as the degree of accuracy, reliability, and trustworthiness (ART) of big data for decision-making. They make sure to point out that data can vary in quality because of inconsistencies, incompleteness, ambiguity, delays, deception, and estimation. Hence, it is essential to have adequate data quality and security plans in place to guarantee the reliability of information, safeguard sensitive information, and adhere to relevant regulations.

H5: Big Data Value extraction positively influences strategic decision-making.

Theoretical Support: (Kayyali et al., 2013) outline the reasons why Big Data is valuable based on the tangible benefits that can be derived from using the data to gain insight and then applying that insight. (La Valle et al., 2011) also highlights the competitive and operational benefits organizations can gain by transforming Big Data into actionable insights. This lens highlights the importance of Big Data, based on an organization's ability to transform raw data into information that informs decision-making.

Big Data Applications in Information Services

Information services use Big Data analytics for various functions. These apps range from preserving and administering digital collections to providing user services and supporting planning efforts.

Digital Libraries and Repositories

Digital libraries use technologies like distributed file systems and NoSQL databases to manage massive quantities of digital resources. These systems enable long-term storage of digitized manuscripts, audio archives, and born-digital datasets at scale. Semantic web technologies and linked data principles can also be used to better connect resources in large metadata catalogs by linking them via contextual relationships (Khan et al., 2022).

These methods enable repositories to be more interconnected and may facilitate the creation of dynamic knowledge graphs, not only based on traditional catalog structures.

Personalized Recommendation Systems

Information services can use recommendation systems based on data such as circulation history, download activity, or search queries to find resources relevant to particular users. For this purpose, libraries use collaborative filtering and content-based filtering techniques.

In academic settings, such systems can help researchers navigate vast amounts of literature to identify pertinent studies and those that might not receive much attention (Padmavathi & Suresh, 2023). A recommendation system can thus support users and increase the visibility of less-used collection items.

Predictive Analytics for Collection Development

Predictive analytics data can show collection managers where demand is likely to occur in the future. Circulation records, interlibrary loan requests, and course syllabi can be used to reveal usage patterns and predict the need for specific courses or titles.

This data can support acquisition decisions so libraries can focus investments on materials they anticipate will be used more and avoid purchasing resources that may be used less. Predictive modeling can further support the "just-in-time" acquisition model, in which resources are made available when needed rather than kept on file for perceived secondary topics (Xu et al., 2022).

Information Service Provider Analysis

In addition, Big Data enables information service providers to analyze behavior patterns in offline and online settings. This can provide better data to guide staffing, hours, space, and interface decisions.

A technology-based Patron Counting system based on network-connection data records the usage of space. The system offers an affordable way to gather data that can inform decisions on staffing, hours of operation, and space design. In contrast, (Jiang et al., 2017) analyzed millions of clickstream records from a library's online catalog to assess information-search behaviors and to offer a framework for advancing interface design.

These examples show how information providers can use data to track existing services and to inform service planning and redesign.

Case Studies and Examples

The British Library's Digital Scholarship

Big Data and computational methods have become part of the British Library's work through its Digital Scholarship and BL Labs programs. These schemes support digital collections in which both content and interfaces are treated as datasets that can be subjected to computational research, in addition to standard reading and browsing.

The library offers datasets produced from digital copies of books, newspapers, maps, and other cultural heritage materials. This collection of resources enables researchers to use computational tools to work with large historical collections. For instance, a project between the British Library and University College London digitized about 60,000 books and analyzed them using high-performance computing for humanities research.

This example shows how the library can directly support computational research and DH without stopping or reducing traditional access and discovery services.

OCLC's World Share Collection Evaluation

OCLC uses big data from cooperating libraries (WorldCAT) to support collection analysis and benchmarking. It enables libraries to compare their collections with other peer and benchmark libraries and spot topics or books that are unique or shared across collections.

This type of comparison can support collection development by helping libraries identify gaps, overlaps, and unique collection holdings. Shared-print programs and initiatives also use WorldCAT-based data and analytics, providing libraries with the information they need to decide which print resources to keep, share, and/or deselect. This helps to maintain library resources and maximize space use.

Limitations

- Limited engagement with Big Data practices in Pakistan may have affected respondent participation.
- Geographic coverage may limit generalizability across Pakistan.
- Convenience sampling may limit population representation.
- The sample of 77 respondents may limit the detection of significant relationships.

Discussion

Personalized recommended systems represent a broader shift in what information providers do. Rather than acting as gatekeepers of content, they now help users to locate and interpret it. Collaborative filtering and content-based filtering can be used to find resources that interest to the user and also to introduce lesser-used resources to their interest (Padmavathi & Suresh, 2023).

This is especially significant for the long-tail issue in libraries, where a relatively small percentage of content can account for a large percentage of use. Recommendation systems offer a means to be able to make a wider variety of resources visible.

The tide is turning in predictive analytics. According to (Xu et al., 2022), Data-driven methods enable information providers to shift from reactive acquisition decisions to more proactive resource allocation. Rather than relying only on past usage, institutions can also account for expected demand over time when allocating acquisition resources.

Academic libraries report similar benefits (Khan et al., 2022). The results indicate that Big Data analytics can enhance service delivery, especially in applications where staff have relied on intuition or limited circulation reports. So, personalization and predictive analytics go hand in hand. So, personalization and predictive analytics go hand in hand. Recommendation systems work at the point of user interaction, while predictive analytics can influence resource decisions and collection development prior to users' interaction.

Behavioral data can also be used to understand limited use of resources. Clickstream and behavioral data (Li & Liu, 2021) can show poor discoverability, weak metadata, or interface design issues can be identified. This helps because if a resource is not being used, it does not mean it is not useful; it may simply mean people are unable to find and/or use it.

The results of both studies suggest a broader range of information service use enabled by Big Data. The data extends throughout various uses – to enhance individual interactions, to inform patterns of use, to assist with resource allocation and to inform institutional decisions. Thus, the usefulness of

Big Data depends on how effectively organizations use the data available to them and transform it into information for action (Mikalef et al., 2018).

Conclusion

The organization and delivery of information services is being fundamentally transformed by Big Data. As libraries and archives, along with other information organizations, shift their purpose from repository to more interconnected, data-driven contexts, the five dimensions of Volume, Variety, Velocity, Veracity, and Value can help explain this change.

This study's results showed a positive relationship between Big Data and Information Services Performance. Examples of data use in applications include personalized recommendation systems, predictive analyses, digital repositories, and behavior analysis, which demonstrate the value of data to both user-facing services and institutional decision-making.

With the growing uptake of these technologies also come responsibilities. However, technical development should be matched by attention to privacy, data quality, algorithmic bias, and governance. As a result, the future of information services will hinge on the responsible use of new analytics, as well as on whether those analytics are advanced.

The future of AI development in information retrieval and privacy-preserving analytics models needs to be explored, with a focus on explainable AI. The lesson learned from Big Data in information services is that these technologies will deliver value only in the long haul if they contribute to equitable, reliable, and meaningful access to knowledge for institutions.

References

- Affleck, M. A. (2021). The quantified library: Big data and the future of service evaluation. *Journal of Library Administration*, 61(3), 324–341. <https://doi.org/10.1080/01930826.2021.1883372>
- Borgman, C. L. (2015). *Big Data, Little Data, No Data: Scholarship in the Networked World*. MIT Press.
- Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Chen, C. L. P., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, 275, 314–347.
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165–1188. <https://doi.org/10.2307/41703503>
- DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, 19(4), 9–30. <https://doi.org/10.1080/07421222.2003.11045748>

- Elgendy, N., & Elragal, A. (2014). Big data analytics: A literature review paper. *Industrial Conference on Data Mining*, 214–227.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Geczy, P. (2014). Big data characteristics. *The Macrotheme Review*, 3(6), 94–104.
- Grolinger, K., Hayes, M., & Capretz, M. A. M. (2014). Data management in cloud environments: NoSQL and NewSQL data stores. *Journal of Cloud Computing: Advances, Systems and Applications*, 3(22). <https://doi.org/10.1186/s13677-014-0022-y>
- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, 47, 98–115. <https://doi.org/10.1016/j.is.2014.07.006>
- Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, 70, 338–345. <https://doi.org/10.1016/j.jbusres.2016.08.007>
- Jiang, T., Chi, Y., & Gao, H. (2017). A clickstream data analysis of Chinese academic library OPAC users' information behavior. *Library & Information Science Research*, 39(3), 213–223. <https://doi.org/10.1016/j.lisr.2017.07.004>
- Kayyali, B., Knott, D., & Kuiken, S. Van. (2013). The big-data revolution in US health care: Accelerating value and innovation.
- Khan, S. A., Bhatti, R., & Khan, A. A. (2022). The impact of Big Data analytics on the service delivery of academic libraries. *Journal of Documentation*, 78(4), 890–908. <https://doi.org/10.1108/JD-06-2021-0118>
- Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. Sage Publications.
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review*, 52(2), 21.
- Li, L., & Liu, J. (2021). User behavior analysis in digital libraries based on big data technology. *Information Processing & Management*, 58(6), 102713. <https://doi.org/10.1016/j.ipm.2021.102713>
- Lund, B. D., Wang, T., & Mannuru, N. R. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, 74(5), 570–581. <https://doi.org/10.1002/asi.24750>
- Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: A systematic literature review and research agenda. *Information Systems and E-Business Management*, 16(3), 547–578. <https://doi.org/10.1007/s10257-017-0362-y>
- Mills, S., Lucas, S., Irakliotis, L., Rappa, M., Carlson, T., & Perlowitz, B. (2012). *Demystifying big data: A practical guide to transforming the business of government*.

- Padmavathi, T., & Suresh, J. (2023). AI-driven personalized recommendation systems in digital libraries: A review. *Library Hi Tech*, 41(2), 345–367. <https://doi.org/10.1108/LHT-09-2022-0412>
- Sagiroglu, S., & Sinanc, D. (2013). Big data: A review. 2013 International Conference on Collaboration Technologies and Systems (CTS), 42–47. <https://doi.org/10.1109/CTS.2013.6567202>
- Salehian, S., & Yan, Y. (2016). Comparison of spark resource managers and distributed file systems. 2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), and Sustainable Computing and Communications (SustainCom), 567–572.
- Salman, M. S., Abdullah, M. K. J., & Sahid, S. N. Z. (2020). Big Data analytic concepts in libraries: A systematic literature review. *International Journal of Academic Research in Progressive Education and Development*, 9(2), 237–249.
- Xu, F., Du, J. T., & Chu, S. K. W. (2022). Data-driven collection management in academic libraries: A systematic review. *Journal of Academic Librarianship*, 48(4), 102543. <https://doi.org/10.1016/j.jacalib.2022.102543>