ISSN: 2754-6659

Journal of Artificial Intelligence & Cloud Computing

Review Article

Open Access

The Synergy between Business Process and Big Data

Sasidhar Duggineni

Compliance Manager, Pharmaceutical Product Development

ABSTRACT

The synergy of Big Data and cloud computing has ushered in a new era of data-driven innovation, transforming the way organizations store, process, and derive insights from vast and complex datasets. This comprehensive overview delves into the intersection of Big Data and cloud computing, exploring key technologies, scalability, elasticity, data warehousing, NoSQL databases, distributed data processing, and the benefits they offer. By harnessing the power of cloud-based resources, organizations can efficiently manage their data resources, optimize performance, control costs, and gain a competitive edge in the data-centric landscape.

***Corresponding author**

Sasidhar Duggineni, Compliance Manager, Pharmaceutical Product Development.

Received: April 15, 2022; **Accepted:** December 10, 2022; **Published:** December 31, 2022

Keywords: Big Data, Cloud Computing, Scalability, Elasticity, Data Warehousing, NoSQL Databases, Distributed Data Processing, Cloud Resource Scaling, Cloud-Based Data Warehouses, Elasticity for Big Data Workloads

Introduction

In recent years, the convergence of Big Data and cloud computing has revolutionized the way organizations handle and harness the power of data. This comprehensive overview delves into the dynamic synergy between Big Data and cloud computing, exploring how this partnership has reshaped the landscape of data management, analysis, and utilization.

The proliferation of digital data from diverse sources, including social media, Internet of Things (IoT) devices, sensors, and transactional systems, has led to the generation of enormous datasets. These datasets, often characterized by the four Vs-volume, velocity, variety, and complexity-present both significant challenges and unprecedented opportunities for organizations.

Cloud computing, with its on-demand access to scalable and flexible computing resources, has emerged as the catalyst for addressing the challenges posed by Big Data. It offers a comprehensive ecosystem of services and tools that enable organizations to store, process, and analyze data efficiently, without the need for massive upfront investments in infrastructure.

This overview provides a deep dive into various facets of the Big Data and cloud computing synergy. It examines the pivotal role of scalability and elasticity in adapting to changing data processing demands, the versatility of NoSQL databases for handling diverse data types, the advantages of cloud-based data warehouses in structuring and analyzing data, and the significance of distributed data processing frameworks like Hadoop and Spark in enabling parallel processing.

Throughout this exploration, we uncover the benefits that organizations gain from embracing scalability and elasticity, such as cost optimization, improved performance, adaptability, and streamlined operations. As the data, landscape continues to evolve, understanding the symbiotic relationship between Big Data and cloud computing is essential for organizations seeking to harness the full potential of their data resources, gain competitive advantages, and make data-driven decisions in a rapidly changing world [1,2].

Literature Review

The literature review section of a comprehensive overview on the synergy of Big Data and cloud computing provides an examination of existing research and scholarly work related to this intersection. It offers insights into the evolving landscape, key trends, challenges, and solutions that have emerged in the context of Big Data and cloud computing. Here is an outline of the key themes and topics typically covered in a literature review for this subject:

Evolution of Big Data and Cloud Computing

- An overview of the historical development of both Big Data and cloud computing, highlighting key milestones, technological advancements, and the factors driving their convergence.

Benefits and Advantages

- An analysis of the benefits organizations derive from combining Big Data and cloud computing, including cost savings, scalability, improved performance, and agility.

Challenges and Limitations

- Identification and discussion of the challenges and limitations associated with managing and analyzing Big Data in the cloud. This may include issues related to data security, privacy, compliance, and data integration.

Scalability and Elasticity

- A review of the concepts of scalability and elasticity in the context of cloud computing, with an emphasis on their significance for handling Big Data workloads.

Cloud-based Data Warehousing

- Examination of the role of cloud-based data warehousing solutions in structuring, storing, and analyzing large volumes of structured data. This may involve discussions on popular cloud data warehousing platforms and their capabilities.

NoSQL Databases

- A review of NoSQL databases and their suitability for managing unstructured and semi-structured data in Big Data applications. This section may discuss different types of NoSQL databases and their use cases.

Distributed Data Processing

- An exploration of distributed data processing frameworks like Hadoop and Spark, including their architecture, key components, and how they enable parallel processing of Big Data workloads.

Elasticity for Big Data Workloads

- A focus on the concept of elasticity and its practical implementation for scaling resources dynamically to accommodate changing Big Data processing demands.

Use Cases and Applications

- Examination of real-world use cases and applications where the synergy of Big Data and cloud computing has made a significant impact across various industries, such as healthcare, finance, e-commerce, and more.

Future Trends and Research Directions

- An exploration of emerging trends, research gaps, and areas of future research within the realm of Big Data and cloud computing, including the integration of machine learning, serverless computing, and edge computing.

Conclusion of the Literature Review

- A summary of key findings and insights from the literature review, highlighting the significance of the synergy between Big Data and cloud computing and its implications for organizations.

A well-structured literature review provides a solid foundation for understanding the current state of knowledge in the field and informs the subsequent sections of the comprehensive overview. It helps readers grasp the context and importance of the topics discussed in the overview while also pointing to areas where further research may be needed.

Key Technologies and Frameworks

NoSQL Databases

NoSQL databases have emerged as essential components in the Big Data and cloud computing ecosystem, offering flexibility and scalability for managing diverse and voluminous datasets. This section delves into the world of NoSQL databases, examining their characteristics, advantages, use cases, integration with Big Data ecosystems, and future trends.

Introduction to NoSQL Databases

- This subsection provides an overview of NoSQL databases, explaining their role in the context of Big Data and cloud computing. It highlights their departure from traditional relational databases and their suitability for handling unstructured and semi-structured data.

Types of NoSQL Databases

- NoSQL databases come in various types, each designed to address specific data storage and retrieval needs. This part explores the different categories of NoSQL databases, including:
- **Document Stores**
- **Key-Value Stores**
- **Column-family Stores**
- **Graph Databases**
- It outlines the characteristics and typical use cases for each type.

Advantages of NoSQL Databases

- This subsection discusses the key advantages of NoSQL databases in the context of Big Data:
- **Schema Flexibility:** NoSQL databases can handle data with varying structures, making them ideal for Big Data's diverse data types.
- **Scalability:** They offer horizontal scalability, allowing organizations to expand storage and processing capabilities as data volumes grow.
- **High Performance:** NoSQL databases are optimized for read and write operations, making them suitable for real-time data access and analytics.
- **NoSQL in the Cloud:** These databases are well-suited for cloud environments, aligning with the scalability and cost-efficiency of cloud computing.

Use Cases and Applications

- This part explores real-world use cases and applications where NoSQL databases excel in Big Data scenarios. It includes examples from industries like e-commerce, social media, IoT, and more.

Integration with Big Data Ecosystem

- NoSQL databases play a pivotal role in the Big Data ecosystem, working alongside other technologies and frameworks. This subsection examines how NoSQL databases integrate with tools like Hadoop, Spark, and data lakes to facilitate comprehensive data processing and analytics.

Cloud-native NoSQL Databases

- Cloud providers offer managed NoSQL database services, enabling organizations to leverage NoSQL databases seamlessly in cloud environments. This section discusses the benefits of cloud-native NoSQL databases and provides examples from major cloud providers.

Future Trends

- The future of NoSQL databases in the context of Big Data and cloud computing includes advancements in areas such as multi-model databases, serverless database computing, enhanced security, and better support for hybrid and multi-cloud environments.

NoSQL databases have become indispensable components for organizations dealing with Big Data, offering the flexibility and scalability required to manage and analyze massive and diverse datasets. Understanding their types, advantages, use cases, and integration possibilities is essential for organizations seeking to unlock the full potential of Big Data and cloud computing [3,4,5].

Introduction to NoSQL Databases

In the landscape of data management, NoSQL databases have emerged as a versatile and powerful alternative to traditional relational databases, particularly in the context of Big Data and cloud computing. NoSQL, which stands for "Not Only SQL," represents a category of database systems designed to address the limitations of traditional relational databases when dealing with diverse, unstructured, or rapidly changing data. This introduction provides an overview of NoSQL databases, their characteristics, and their significance in modern data management.

Why NoSQL?

- Traditional relational databases have served as the cornerstone of data storage and retrieval for decades. However, they are primarily designed for structured data with well-defined schemas, which may not be suitable for the data challenges posed by Big Data. NoSQL databases have emerged to address the limitations of relational databases, offering flexibility and scalability to meet the demands of modern data applications.

Characteristics of NoSQL Databases

- NoSQL databases exhibit several key characteristics that set them apart from traditional relational databases:
- **Schema-less:** NoSQL databases do not enforce a fixed schema, allowing for the storage of data with varying structures and attributes.
- **Scalability:** They are designed for horizontal scalability, meaning that additional resources and nodes can be added to handle increasing data volumes and processing demands.
- **Data Models:** NoSQL databases support various data models, including document-oriented, key-value, column-family, and graph, making them adaptable to different data types and use cases.
- **High Performance:** NoSQL databases are optimized for read and write operations, making them well-suited for real-time and high-velocity data.
- **Flexibility:** They provide flexibility in data modeling, enabling agile development and iterative changes to data structures.
- **NoSQL in the Cloud:** NoSQL databases align with the cloud computing paradigm, allowing organizations to harness cloud resources for scalability and cost-efficiency.

Types of NoSQL Databases

- NoSQL databases can be categorized into several types based on their data models:
- **Document Stores:** These databases store data in semi-structured documents, often in JSON or BSON format, and are suitable for applications requiring flexible schema and hierarchical data.
- **Key-Value Stores:** Key-value stores are the simplest form of NoSQL databases, where data is stored as key-value pairs, making them highly efficient for data retrieval by key.
- **Column-family Stores:** Designed for storing and retrieving large volumes of data with high write throughput, column-family stores excel in analytical and time-series data use cases.
- **Graph Databases:** Graph databases are optimized for representing and querying complex relationships between data entities, making them ideal for applications involving networks, social connections, and recommendation systems.

Advantages of NoSQL Databases

- NoSQL databases offer several advantages, including:
- **Flexibility:** Their schema-less nature allows for easy adaptation to evolving data requirements.
- **Scalability:** NoSQL databases can distribute data across multiple nodes and clusters to accommodate data growth.
- **High Performance:** They excel in read and write operations, making them suitable for real-time applications.
- **Diverse Use Cases:** NoSQL databases are used in various industries and applications, from e-commerce and social media to IoT and log analytics.

NoSQL databases have become integral components of modern data ecosystems, empowering organizations to manage and analyze vast and diverse datasets efficiently. As organizations navigate the complexities of Big Data and cloud computing, NoSQL databases offer the flexibility and scalability needed to embrace the full potential of their data resources.

Types of NoSQL Databases

NoSQL databases are characterized by their ability to store and manage data in ways that depart from the rigid structure of traditional relational databases. They are designed to handle diverse data types, scale horizontally, and provide efficient data access. There are several types of NoSQL databases, each tailored to specific use cases and data modeling requirements:

Document Stores

- **Characteristics:** Document stores store data in semi-structured documents, typically in formats like JSON or BSON. Each document can have a different structure within the same database.
- **Use Cases:** Document stores are well-suited for applications that require flexibility in data schema, such as content management systems, e-commerce platforms, and user profiles.
- **Examples:** MongoDB, Couchbase, RavenDB.

Key-Value Stores

- **Characteristics:** Key-value stores are the simplest form of NoSQL databases, where data is stored as key-value pairs. They provide fast and efficient retrieval of data based on keys.
- **Use Cases:** Key-value stores are suitable for caching, session management, and scenarios where quick retrieval of data by key is essential.
- **Examples:** Redis, Amazon DynamoDB, Riak.

Column-family Stores

- **Characteristics:** Column-family stores organize data into column families, similar to tables, but with the flexibility to have different columns for each row. They excel in handling large volumes of data with high write throughput.
- **Use Cases:** Column-family stores are commonly used for time-series data, event logging, and analytical workloads.
- **Examples:** Apache Cassandra, HBase, ScyllaDB.

Graph Databases

- Characteristics: Graph databases are optimized for representing and querying complex relationships between data entities. They use graph structures to model data, with nodes

representing entities and edges representing relationships.
- Use Cases: Graph databases are ideal for applications involving social networks, recommendation engines, fraud detection, and any scenario where relationships between data points are critical.
- Examples: Neo4j, Amazon Neptune, OrientDB.

Wide-column Stores
- **Characteristics:** Wide-column stores are similar to column-family stores but with more emphasis on horizontal scalability. They are designed to handle large volumes of data and can distribute data across multiple nodes.
- **Use Cases:** Wide-column stores are suitable for scenarios requiring massive scalability and support for geographically distributed data, such as global content delivery and analytics.
- **Examples:** Apache Cassandra, HBase.

Object-oriented Databases
- **Characteristics:** Object-oriented databases store data in objects, similar to how object-oriented programming languages represent data. They are well-suited for applications that work with complex, highly structured data.
- **Use Cases:** Object-oriented databases are used in applications where the data structure closely mirrors the application's object model, such as software development tools and scientific applications.
- **Examples:** db4o, ObjectDB.

It's important to note that some NoSQL databases may combine elements of multiple types. Additionally, the choice of a NoSQL database type should align with the specific requirements of the application and the nature of the data being managed. As organizations embrace the diversity of data sources and formats in the age of Big Data, NoSQL databases provide the flexibility needed to adapt and thrive in this evolving data landscape [5,6].

Advantages of NoSQL Databases
NoSQL databases offer a range of advantages that make them suitable for various use cases, particularly in the context of Big Data and modern data management. Here are some key advantages of NoSQL databases:

Schema Flexibility
- NoSQL databases do not require a predefined schema, allowing for flexible and agile data modeling. This means you can store and retrieve data without adhering to a rigid structure, making it easier to accommodate changes in data requirements over time.

Scalability
- NoSQL databases are designed for horizontal scalability, which means you can scale out by adding more servers or nodes to handle increased data volumes and traffic. This enables organizations to grow their data infrastructure as needed, supporting high-volume and high-velocity data scenarios.

High Performance
- NoSQL databases are optimized for specific data access patterns, resulting in high-performance read and write operations. This makes them well-suited for real-time applications and scenarios where low-latency data access is critical.

Support for Diverse Data Types
- NoSQL databases can handle a wide variety of data types, including structured, semi-structured, and unstructured data. This versatility makes them ideal for managing Big Data, which often includes data in different formats.

Flexible Querying
Many NoSQL databases provide flexible query languages and APIs that allow developers to express complex queries and filter criteria. This adaptability simplifies data retrieval and analysis tasks.

Horizontal Scaling
- NoSQL databases support horizontal scaling by distributing data across multiple servers or nodes. This helps maintain performance as data volumes grow and ensures high availability.

High Availability and Fault Tolerance
- NoSQL databases often incorporate features for automatic data replication and failover, ensuring data availability even in the event of hardware failures or network issues. This contributes to improved system reliability.

Geospatial and Multimodal Capabilities
- Some NoSQL databases provide specialized features for geospatial data and multimodal data, making them suitable for applications involving location-based services, mapping, and multimedia content.

Cloud-native Compatibility
- NoSQL databases are well-suited for cloud environments, where they can take advantage of cloud-native features like automatic scaling, pay-as-you-go pricing, and integration with other cloud services.

Use Case Specificity
- NoSQL databases are designed with specific use cases in mind, such as document storage, key-value caching, graph analysis, or time-series data. This means you can choose the NoSQL type that best aligns with your application's requirements.

Cost-Efficiency
- NoSQL databases often offer cost-efficient storage and processing options, particularly when using cloud-based NoSQL services. Organizations can optimize costs by paying only for the resources they use.

Support for Big Data Analytics
- Many NoSQL databases seamlessly integrate with Big Data processing frameworks like Hadoop and Spark, enabling organizations to perform analytics on large and complex datasets.

Global Distribution
- Some NoSQL databases support global distribution, allowing data to be replicated and stored in multiple regions or data centers. This is valuable for applications requiring low-latency access across different geographical locations [7,8].

Adoption of NoSQL in Modern Tech Stacks
- NoSQL databases have become an integral part of modern technology stacks, enabling organizations to build scalable and agile applications that can adapt to changing data requirements.

These advantages have made NoSQL databases a crucial component in handling the challenges posed by Big Data and supporting data-intensive applications across various industries, including e-commerce, social media, healthcare, finance, and IoT. The choice of a NoSQL database type depends on the specific needs of the application and the nature of the data being managed.

Use Cases and Applications

NoSQL databases find applications across a wide range of industries and use cases due to their flexibility, scalability, and suitability for handling diverse and voluminous datasets. Here are some notable use cases and applications where NoSQL databases excel:

Content Management Systems (CMS)
- NoSQL databases are well-suited for storing and managing content in CMSs, where content can vary in structure and format. They provide the flexibility needed to handle articles, images, videos, and user-generated content.

E-commerce Platforms
- E-commerce websites often deal with vast product catalogs, user reviews, and shopping cart data. NoSQL databases can efficiently manage and retrieve this data while handling fluctuating traffic loads during peak shopping seasons.

Social Media and User Profiles
- Social media platforms and applications rely on NoSQL databases to store user profiles, friend networks, posts, and comments. The schema flexibility of NoSQL databases accommodates the dynamic nature of social interactions.

Internet of Things (IoT)
- IoT devices generate a continuous stream of data, including sensor readings, telemetry, and device status. NoSQL databases can ingest and process this data in real time, making them essential for IoT platforms.

Real-time Analytics
- NoSQL databases are employed in real-time analytics platforms where data from multiple sources, such as website clickstreams, logs, and social media feeds, is processed and analyzed on-the-fly to derive insights and make immediate decisions.

Log and Event Data Management
- Applications and systems generate logs and events that need to be stored, indexed, and analyzed for troubleshooting and monitoring. NoSQL databases, particularly column-family stores, excel in handling time-series data.

Recommendation Engines
- NoSQL graph databases are used to build recommendation engines that analyze user preferences and social connections to suggest products, content, or connections tailored to individual users.

Time-Series Data Analysis
- Time-series databases, a subset of NoSQL databases, are designed for storing and analyzing data points collected over time. They are crucial for applications like financial analytics, IoT, and environmental monitoring.

Gaming and Leaderboards
- Online gaming platforms utilize NoSQL databases to manage player profiles, achievements, leaderboards, and game state. These databases provide low-latency access to player data, enhancing the gaming experience.

Fraud Detection and Security
- NoSQL databases are used in fraud detection systems to analyze transaction data and detect anomalies or patterns indicative of fraudulent activity. They can handle large datasets and perform real-time analysis.

Location-Based Services
- Geospatial databases built on NoSQL technology are essential for location-based services, such as mapping applications, ride-sharing platforms, and location-based advertising.

Multimedia Content Management
- NoSQL databases are employed for managing multimedia content, including images, audio, and video files, in applications like media libraries, streaming services, and content distribution platforms.

Scientific Research and Analysis
- In scientific research, NoSQL databases are used to store and query large datasets, conduct simulations, and analyze experimental results. Their scalability supports high-performance computing in research environments.

Healthcare and Electronic Health Records (EHR)
- NoSQL databases are utilized in healthcare for managing electronic health records, patient data, and medical imaging. They enable secure and efficient access to patient information.

Supply Chain and Inventory Management
- Supply chain and inventory management systems leverage NoSQL databases to track inventory, shipments, and demand fluctuations in real time, optimizing logistics and inventory control.

These use cases highlight the versatility of NoSQL databases in addressing a wide array of data management and analytics challenges. Organizations across different sectors benefit from the scalability, agility, and performance offered by NoSQL database technology in the era of Big Data and data-driven decision-making.

Integration with Big Data Ecosystem

NoSQL databases play a crucial role in the Big Data ecosystem by seamlessly integrating with various Big Data technologies and frameworks. This integration enhances data processing, analytics, and storage capabilities. Here's how NoSQL databases integrate with the Big Data ecosystem:

Hadoop Integration
- NoSQL databases can be integrated with Hadoop, a popular open-source framework for distributed data storage and processing. This integration allows organizations to use NoSQL databases as data sources or sinks for Hadoop jobs, enabling the extraction and loading of data into Hadoop clusters for processing.

Spark Integration

- Apache Spark, a high-performance data processing framework, can be connected to NoSQL databases to read and write data. This integration enables Spark to leverage the distributed storage and querying capabilities of NoSQL databases for real-time and batch processing tasks.

Kafka Integration

- Apache Kafka, a distributed streaming platform, can be used in conjunction with NoSQL databases to ingest and stream data in real time. NoSQL databases can serve as destinations for Kafka producers, allowing for efficient and scalable data storage.

Data Lakes

- NoSQL databases are compatible with data lakes, which are centralized repositories for storing structured and unstructured data. Organizations can use NoSQL databases to store and manage data within data lakes, facilitating easy access and analysis.

ETL (Extract, Transform, Load) Processes

- NoSQL databases can be part of ETL pipelines, where data is extracted from various sources, transformed, and then loaded into the database for analysis. This integration supports data consolidation and transformation efforts.

Machine Learning and AI Integration

- Machine learning and AI models often require large and diverse datasets for training and inference. NoSQL databases can serve as data sources for machine learning pipelines, providing the required data for model development and evaluation.

Graph Analytics

- NoSQL graph databases are ideal for graph analytics use cases. They can be integrated with graph processing frameworks like Apache Giraph and GraphX to analyze complex relationships and network data at scale.

Real-time Analytics

- NoSQL databases are commonly used in real-time analytics platforms. They can integrate with stream processing frameworks like Apache Flink and Apache Storm to analyze and process data as it is generated in real time.

Data Warehouses

- Some NoSQL databases offer connectors to popular data warehousing solutions, enabling organizations to transfer and consolidate data for structured querying and reporting.

Cloud Integration

- Many cloud-based NoSQL databases seamlessly integrate with cloud-based Big Data services offered by providers like AWS, Azure, and Google Cloud. This allows organizations to leverage the scalability and cost-efficiency of cloud resources for Big Data processing.

Polyglot Persistence

- Organizations often adopt a polyglot persistence approach, where different types of databases, including NoSQL databases, are used for various data storage needs within the same application or ecosystem. This approach leverages the strengths of each database type for specific tasks.

The ability of NoSQL databases to work alongside Big Data technologies and frameworks enhances their utility in managing and analyzing large and complex datasets. This integration enables organizations to build robust and scalable data pipelines, conduct real-time analytics, and support diverse data-driven applications [9,10].

Conclusion

The synergy between NoSQL databases and the Big Data ecosystem has transformed the landscape of data management, analytics, and processing. NoSQL databases, known for their flexibility, scalability, and diverse data model support, have become indispensable tools in the data-driven era. In this comprehensive overview, we have explored the integration of NoSQL databases within the Big Data ecosystem, highlighting their pivotal role and the advantages they offer.

NoSQL databases provide organizations with the agility needed to handle diverse and rapidly changing data types, making them suitable for a wide array of use cases across various industries. Whether it's managing unstructured content in content management systems, processing vast streams of data in real time, or supporting complex graph analytics, NoSQL databases excel in addressing the data challenges of the modern world.

Their integration with Big Data technologies like Hadoop, Spark, Kafka, and data lakes has further enhanced their utility. NoSQL databases serve as critical components in ETL processes, real-time analytics platforms, machine learning pipelines, and data warehousing solutions. They enable organizations to consolidate, process, and analyze data efficiently, contributing to data-driven decision-making and innovation.

As the data landscape continues to evolve, NoSQL databases are poised to play an even more significant role. Trends such as multi-model databases, serverless computing, and edge computing are likely to influence the development and adoption of NoSQL database technologies. Moreover, the cloud-native nature of many NoSQL databases aligns perfectly with the growing adoption of cloud computing, providing organizations with scalable and cost-effective solutions.

In conclusion, the integration of NoSQL databases into the Big Data ecosystem represents a symbiotic relationship that empowers organizations to unlock the full potential of their data resources. By harnessing the flexibility, scalability, and diverse capabilities of NoSQL databases, organizations can navigate the complexities of Big Data, gain actionable insights, and stay competitive in an increasingly data-centric world. As the data-driven journey continues, the role of NoSQL databases will remain central to the pursuit of innovation and excellence in data management and analytics.

References

1. Data Science Association (DSA) (2017) About data science. http://www.datascienceassn.org/about-data-science.
2. DCMI (2003) Special session: Smart metadata. In 2003 Dublin Core Conference: Supporting Communities of Discourse and Practice-Metadata Research & Applications, Seattle, Washington http://dublincore.org/workshops/dc-2003/smartDC. html.
3. Dhar V (2013) Data science and prediction. Communications of the ACM 56: 64-73.

4. Dimitrova N (2004) Is it time for a moratorium on metadata? IEEE Multimedia 11: 10-17.
5. Doctorow C (2001) Metacrap: Putting the torch to seven straw-men of the meta-utopia. Metacrap http://chnm.gmu.edu/digitalhistory/links/pdf/preserving/8_17.pdf.
6. Dong R, Su F, Yang S, Xu L, Cheng X, et al. (2016) Design and application on metadata management for information supply chain. In the 16th International Symposium on Communications and Information Technologies (ISCIT) 393-396.
7. Greenberg J, Garoufallou E (2013) Change and a future for metadata. In MTSR-2013: Proceedings of the 7th Metadata and Semantics Research Conference 1-5.
8. Kogan DE, Miller PC, Schobbe GA (2007) Techniques to manage metadata fields for a taxonomy system. US 20080301096 A1 http://www.freepatentsonline.com/y2008/0301096.html.
9. Abbasi M, Vassilopoulou P, Stergioulas L (2017) Technology roadmap for the creative industries. Creative Industries Journal 10: 40–58.
10. Data Science Association (DSA) (2017) About data science http://www.datascienceassn.org/about-data-science.