Powerful Discovery information Service tools for Big Datain institutions for higher education: A case study of KSA Universities

Dr. Mohammed A. Abdulla[1]; Furkan Yousaf[2]

Prince Mohammad Bin Fahd University[1]; Prince Mohammad Bin Fahd University[2]
E-mail: mabdulla@pmu.edu.sa[1]; fyousaf@pmu.edu.sa[2]

ABSTRACT

The concept of discovery tools is growing tremendously within the environment of higher education institutions. In the presence of discovery tools,Big Data (BD) also represents an important stage in the development of information and communication technology systems. It reflects an enormouslylarge volume of BD that is larger than the capacity of traditional computer software and mechanisms to store, process and distribute.Thus the BD limiting information specialists to develop sophisticated alternative solutions that enable them to controlinformation, which can be termed as information discovery tools.

Based on the above KSA universities got the opportunity to establish discovery tools that help universities to control, organize and retrieve information and BD to benefit from them in achieving their educational and research objectives. This study comes in the search and disclosure of the powerful discovery tools to explore the information and data in currently implemented in KSA universities and to reveal their strength and effectiveness and the extent of discovery tools influence in the information literacy of the educational institution's societies. While mentioning the barriers and obstacles facing KSA universities in dealing with BD. In the current scenario, many academic institutions in KSA still on process to accept and implementation of discovery tools to achieve their objectives.

Keywords:Big Data, Discovery tools, Information Retrieval, Academic Institutions

1. INTRODUCTION

The problematic of computer interactions with the BD in the educational sectors, and the BD poses challenges to the traditional computer systems used in education especially the processes of extracting and analyzing important information, therefore the need to help in trying to reach new solutions and innovative for data processing in order to overcome the obstacles posed by BD such as the huge of data on the growth of the generations and their stages of study is extremely important so that there are some solutions that can be adopted and operate.The introduction of hybrid cloud storage systems helps data storage and processing, however, this type of solution is not accepted by most institutions, especially those with sensitive data.They are concerned about the possibility of problems related to data security since the data will be transferred from the internal storage area of the institution to a cloud storage area with the participation of many individuals and entities is a public storage area(Al Akalabi, Ali, 2018).On the other hand, there are several discovery tools and techniques that are used to analyze BD such as Hadoop, Map Reduce, HPCC. Hadoop is one of the most famous of these tool. It is a software or open-source software platform written in Java language for storing and processing BD in a distributed format such as storing BD on multiple devices and then distributing process on these devices to speed up the processing result. One of the most popular users of Hadoop are Amazon, Apple, AVG, eBay, Electronic Arts, Facebook, Google, IBM, LinkedIn,

Microsoft, New York Times, Twitter, and Yahoo(C. Pettey, 2012).

The challenges and techniques used in the BD management have confirmed that the issue of BD has become an affecting issue and has become the concern of many researchers and decision-makers in academic institutions.The high acceleration in the growth of information and BD has caused the troubles great for humans, from here it can be concluded that the BD contains many untapped treasures that have not yet been discovered.The huge data store many valuable benefits and knowledge potential and useful (C.L. Philip Chen, & C.Y. Zhang, 2014).

The Kingdom of Saudi Arabia is the leader in the gulf region took initiative and accept the updated trend of data science, BD, and open data. KSA'sintuitive bright, and enlightened minded government, semi-government, private cooperate and academic sector work through and understand the power of BD processing, analyzing and conclude valuable results from BD.

This study examines the status of implementation of BD discovery tools in higher education institutions in KSA, and the effectiveness of the discovery tools to explore, retrieve, modify and analyze data to achieve the objectives of these educational and research institutions. The discovery tools are used in KSA higher education institutions to extract data as a technique aimed at finding knowledge from BD, the availability of discovery tools based on smart algorithms that support e-learning which is the most driving solution in "BD analysis " that would be the future of learning and education in KSA. On the other hand, the paper outlines the information literacy level of the KSA educational institution's community on accessing and usage of discovery services working for BD and its barriers. Through this, it will guide information professionals in either KSA educational institution's community to require a sufficient level of training to use discovery tools effectively.

2. DESIGN / METHODOLOGY / APPROACH

In KSA, a total of 10 academic institutions is operating. Due to a number of universities, the researcher has considered to include all university and keep it as a sample of the study. A qualitative method of research wasused therefore; the descriptive method was also used to analyze

discovery tools content to identify the properties of open BD pages on the portals. The document analysis method of data and information collection techniqueswas used. Analyzing documents incorporates coding content into themes similar to all KSA educational institutions community as the focus group, specifically,such universities implemented discovery tools through the exploration of websites and analysis discovery tools related documents.The current discovery tools used search within those tools and conduct of tests that kept it more effective and powerful.In addition, conducting the interview of information professionals, direct observation of university discovery tools websites used to achieve targeted results (Meshael, Sultan and Ayesha,Mukthar,2017).

3. Objective

3.1 To check the implementation status of discovery tools for BD in academic institutions of KSA.

3.2 To know the effectiveness and the extent of discovery tools influence in the information literacy of the educational institution's societies

3.3 To explain the power of discovery tools and discovery services for accessing and usage of information and BD in Academic institutions of KSA

3.4 To identify the barriers facing by KSA educational institutions community in dealing with discovery services for BD.

4. RESEARCH QUESTION

4.1 What is the implementation status of discovery tools for BD in academic institutions of KSA?

4.2 What is the effectiveness of discovery tools' influence in the information literacy of the educational institution's societies?

4.3 What is the power of discovery tools and discovery services for accessing and usage of information and BD in Academic institutions of KSA?

4.4 What are the barriers facing by the faculty member to get information literacy instruction on discovery services for BD?

4.5 What are the benefits of the use of the discovery tools for BD in the academic libraries for the development of information services provided to the KSA educational institutions community?

5. LITERATURE REVIEW

Universities students of the 21st century are technology-oriented, always looking to activate one search team and get the best result of their demand. Where the world leading web browsers and search interface making their dreams true. The dream of information centers in higher academic institutions to provide a one-stop academic search interface gone true with the intent of discovery tools. Potentially, it has fulfilled the demand for academic needs. Discovery tools are playing a critical role in the academic institutions by dominating and effectively presenting the information from BD. The catalogs of old models and indexing system have been replaced and the power of navigating right information within multi-databases has been aligned with the single search model of discovering tool (A. Karadia and S. Pati.2015).

There are many definitions for discovery tools services,among them the definition provided by JISC "A discovery tool is a developmental tool that students and staff can use to self-assess their digital competences, identify their strengths and opportunities to develop their skills further to enhance their practice of learning" (Building digital capability, (JISC. 2019). Discovery tools as web software that searches journal articles and library OPAC metadata in a unified index and presents search results in one single interface. This differs from federated search software, which searches multiple databases and combinations the results with joined search software (ProQuest. 2019). Nowadays the discovery tools offer a real information service in higher academic institutions form integrated search experience in academic libraries and information technology departments.

The researchers sought for "BD" definitions, some researchers define BD is high volume, high velocity, and high variety of information assets. Some definitions also include veracity, which refers to the need to ensure the data is correct (which is an increasingly difficult task in BD), high volume, high velocity, and high variety information assets. (X. Shacklock.2016). BD is a collection of data from traditional and digital sources inside and outside your company that represents a source for ongoing discovery and analysis (L. Arthur. 2015). Gartner, described as "a leading information technology research and consulting firm", defining BD as "high-volume, high velocity and high-variety information assets

that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" (C. Pettey, 2012). "BD" refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze (A. Karadia and S. Pati, 2015). From all these definitions we conclude that the BD is a term used to describe a very large data set or a degree of complexity or that requires a great deal of rapid processing (sometimes called volume/diversity/speed problems) which becomes difficult or impossible to deal with using the traditional data rules and traditional analytical tools. The data processing of this size requires software that works in a balanced way on tens or hundreds or thousands of servers.

Recently BD of higher education institutions is increasing. There are many computers, laboratories in university and research centers that are connected to the internet on campuswhere many researchers, faculty members and students of these academic institutions are using their personal devices for storing the information. usually,their devices are connected with the universities' internal network. They access thedata that belong to educational institutions, resulting in a flood of data that is difficult to deal with traditional methods.
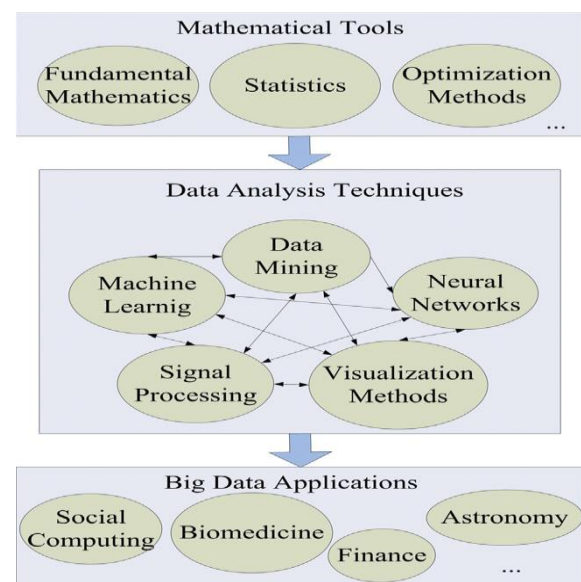

Figure 1: BD techniques

Discovery tools can be considered as an appropriate solution to invest time, money and effort.In the BD analytics initiative, building a researching strategy,

retrieving and analyzing BD with the design of the data architecture to achieve the goals of educational institutions.

Finally, BD has a deep relationship with E-Science (T. Hey and A. E. Trefethen, 2002) which is computationally intensive science which usually implemented in distributed computing systems. Many issues on BD applications can be resolved by E-Science which requires network computing (B. Jacob, M. Brown, K. Fukui, and N. Trivedi 2005). E-Sciences include particle physics, bioinformatics, earth sciences, and social simulations. How to probe knowledge from the data produced by large-scale scientific simulation? It is a certain BDproblem in which the answer is still unsatisfiable or unknown(C.L. Philip Chen, & C.Y. Zhang, 2014).

Examples of discovery tools in higher academic institutions:

5.1 Knowledge discovery and Data mining tools:

Knowledge Discovery in Database (KDD) is a complex process that some might think will stop when collecting and managing data, but rather to analyze, anticipate and predict; what happens in the future. Data mining is part of the knowledge discovery, and its process is the most comprehensive. The knowledge discovery process includes the following steps:

- **Data Discovery**: The data collection phase includes the detection, identification, and characterization of available data.
- **Data Cleaning**: at this point, data that is irrelevant is removed, and conflicting data and inconsistent data are deleted.
- **Data Integration**: similar and relevant data from multiple data sources are collected and merged at this stage.
- **Data Selection**: at this point, appropriate data is identified and retrieved from the dataset.
- **Data Transformation**: at this point, the data is converted into custom forms suitable for search and retrieval procedures by means of the complete summary or aggregation processes.

Thus, data mining is an essential step to the application of intelligent methods in order to detect interesting data patterns hidden in BD sets. However, some organizations find that the term "data mining" has become more popular to refer to the process in which the entire knowledge

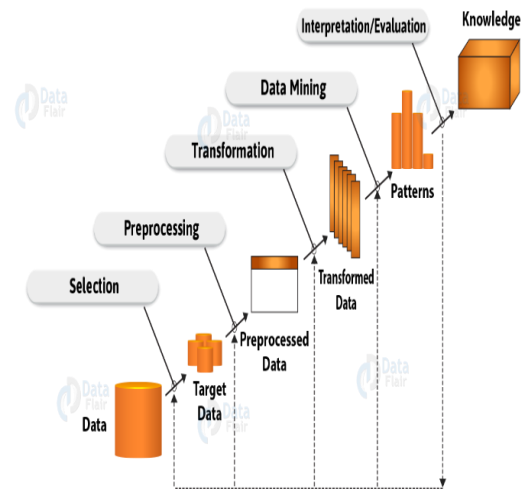discovery, below arethe elementsof data Mining and Knowledge Discovery (Naseej.com 2019).

Figure 2: Elements of Data Mining and Knowledge Discovery, 2019

Through websites survey of many higher education institutions,BD with reference to anunlimited number of research and studies; the applications of BD analytics for higher education are onthe initial stage. For example:

- Significant data analyses are being introduced into education to increase the number of local students enrolled in additional education and to measure the performance of schools and universities, according to (Sutton, Mark, 2017).
- A cloud and BD based workflow have been introduced, analyzing for overall quality and improving higher education. This framework serves as a system for resuming work, proposing changes to the course outline and the curriculum (Ahmed, Shakeel, and Hemant Kumar Mehta, 2015).
- BD analyses help the KSA government to link the skills and they need for the future economy and this will translate into programs and produce opportunities. They work with Deloitte Consulting, a company that specializes in helping people to find a job instead of taking social welfare, as part of the KSA, TAQAT program. According to Julie Mercer, a leading partner in the global industry, educational consultancy, Deloitte University plans to "create an environment and curriculum that supports

broader political, economic and social development in a challenging world."

5.2 My library services:

As library users become more and more comfortable with each new innovation especially, in the web-based world. Academic libraries have found themselves struggling to offer a knowledge-based database that seems as easy as Google search engine.Yet directs beneficiaries to the information resources maintained or licensed by the academic libraries. A discovery tool is a lens through which the beneficiaries sees the library's physical and digital collections (Mary Pagliero Popp, 2012).

The best model of my library service is provided by the Saudi Digital Library (SDL), which works to build a large and sophisticated digital library that includes various disciplines that support the educational process and meet the needs of beneficiaries in a higher education institution, KSA.

The E -portal of the SDL is a real search manager that provides all search services through a single-window

including Arabic databases, English databases, eBooks, KSA universities theses, dissertations Archive, trial databases and open access.Through this portal, users can obtain results that coverall areas of interest from various and reliable sources.
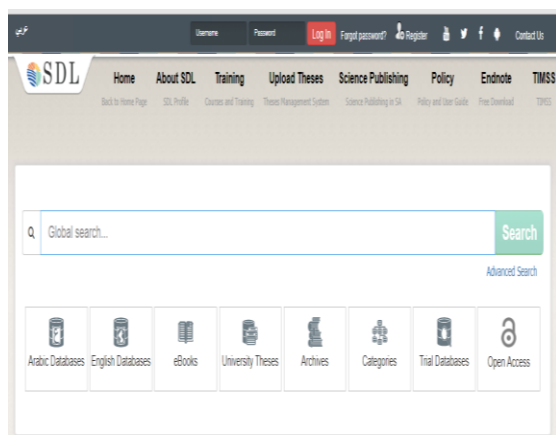
5.3 Semantic indexing (OPAC)

OPAC is an online library catalog, created from local bibliographical data sources include discipline-specific digital repositories, digital collections, applications that extract and adapt catalog data for a particular purpose forthe audience (Mary,Pagliero Popp, 2012).OPAC used to find data quickly. Finding the required information within that vast BD may be impossible without using indexing. On the other hand, with an exponential increase in the volume and variety of data, the indexing of raw data has begun to lose its efficiency. Different types and diversity increase the size of the index.Later introduce metadata and linked data. Therefore, instead of indexing raw data, in-depth learning provides a simplified and abstract representation of this data to be indexed. The in-depth learning algorithms are trained to extract semantic features from data. The semantic indexing algorithms begin using the output of the in-depth learning algorithms as input rather than raw data. Hence, semantic indexing is used to classify, organize, and represent different types of documents. However, with in-depth learning ability to represent different types of files, semantic indexing can be used to handle such types of files(In-depth learning and big data, 2016).

The best sample for indexing applications as a discovery tool to explore the BD found in the Arabic Union Catalogue (ARUC).ARUC is a cooperative non-profit project aimed at creating a cooperative environment for Arab libraries in particular in order to reduce the cost of cataloging the Arabic information resources through the process of shared cataloging that requires standardization cataloging practices within Arab libraries and adoption of international standards in the Bibliographical description. This is achieving the development of the level of conduct within the Arab libraries through high-quality records made available to libraries inside and outside the Arab world.

6. Discovery tools and IR models

Information retrieval (IR) models enable the researcher to formulate the research. There are different kinds of IR models that support researchers to extract information from BD such as the Boolean model, vector space model, region model, and probabilistic model, two Poisson, Language Model, Bayesian Network and Google's Page Rank Model (Irfan, S. and B. Babu, 2016). Boolean is the earliest IR model almost all libraries and information centers use to teach Boolean operators in digital resource orientation and information literacy sessions. Library discovery tools have adopted this model and gave AND, OR and NOT facility under advance searching option.

Georgetown University Law Center subscribed Serials Solutions Summon system and assign local names as "OneSearch". They have

incorporate library discovery tools (LDT) with the Information literacy program. Krause who teaching library instruction, during information literacy session. He teaches library users how OneSearch works. He searches "Genetic Screening" and found 214000 documents. That was the devastating number of resources. So he starts guiding LDS has many features. It has covered many latest IR models that filter results and how its advance search feature works to retrieve accurate results(Krause, M, 2016).

From the beginning days, the concept of LDT gone popular because it offered easy access, link library local data with external data, and that can refine data from BD. The different commercial firms and open source got involved and developed applications by adopted standard IR models(Krause, M.2016;Breeding. M. 2010). The most popular LDT are Summon by Serials Solutions, Encore from Innovative, WorldCat Local by OCLC, AquaBrowser from Medialab Solutions, Primo from Ex Libris, BiblioCommons, EBSCO Discovery Service, Visualizer from VTLS, and Enterprise from SirsiDynix (Primo Content Index, 2019).

7. DISCOVERY TOOLS FOR BD IN KSA UNIVERSITIES

The development of Higher education in KSA is at its peak forthe last 10 years. Many new public and private universities have started their business and their enrollment of students also increases tremendously. In 2009, a total of 20 universities was working in KSA. Now the strength of universities reached up to 41, among them, 29 are of Government universities and 12 of them are private universities. Where 13 government and 7 military colleges started their work within the same period of time. Statistics of KSA General Authority for Statistics show that students enrolment increase up to 40%. At the time of 2009, the total enrolment was 272854 and after 10 years of gap it reached up to 586,179 among them 379179 students were studying within the tertiary of KSA and rest of 207,000 were studying abroad on scholarship. This 586,179 is 2.8% of the total population of BD (A. Mukthar and M. Sultan, 2017).

Today's researchers are overwhelmed with information and data, according to a UNESCO publication report the researchers working in scientific fields with BD; wonder how best to manage and control their massive data exchange with research labs. Where draw the dividing line

between transparency of data for the science and social development, and between the risks of exploding data that cannot be controlled.It takesa long time for higher education institutions to have resources to effectively control the BD storage and analysis process. Withthe advent of next-generation technologies like Hadoop, this is no longer an issueUNESCO, (2015).

One of the most distinctive applications for the use of LDT and BDtechnology; experiencedby King Saud University, which recognized the importance of the knowledge industry and analysis of the BD.Therefore,established many units that contribute to the investment of knowledge and work on analyzing the BD. Thedecision-maker support canenable such applications in institutions. LDT is animportant application that produces valuable information, storage, processing and analysis fromBD that is a real challenge for KSA universities. King Saud University has established a system for planning and information management called (ITQAN).This system ITQAN aims to provide accurate and reliable data that support the process of preparation, planning, and development which helps to make the right decision in a timely manner. In addition to facilitating the process of accessing, analyzed, and reducing the data accurately ITQAN system framework following the best IR model. The structure model of ITQAN is below:
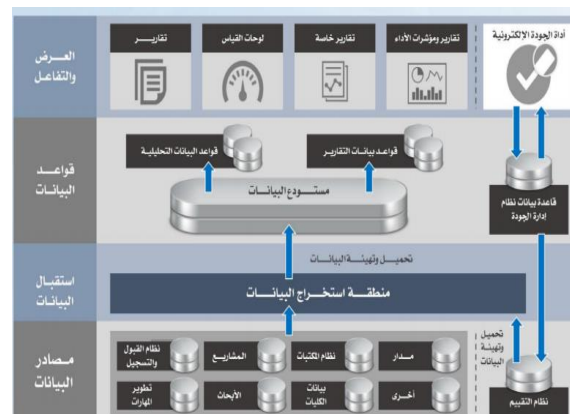

Figure 3: ITQAN System Guide

Another experience led by the Knowledge Technology Foundation. The renowned Arabic open source discovery and unified search portal consist of a comprehensive and customizable user interface that supports smart search methods, deep browsing tools, and operational user processes across all information systemsand repositories of descriptive data of all kinds.

7.1 Data analysis tools

Data analysis tools process and analyze the BD through which users can able to discovervalued information, by comparison, classification, approach, linking, analytical, and organizational. These things could achieve the desired results and predetermined targets.

7.2 WorldCat Discovery tool

WorldCat discovery (WCD) tool is one of the best library resources navigation tools that can support or integrate with any kind of library system. This tool balances the user's needs versus technology advances and provides efficient resources from a complex matrix (WorldCat Discovery, 2019).

WCD has a single-user search model interface. It offers two levels of application; the first level of application deals with a basic level of searching functionality and others can have the functionality to deal with a complex search. There single users interface is covering more than 1.8 billion physicals, digital, electronic content. It also has the flexibility to add the local content of the subscribed library. As worldwide they have the ability to manage 285 different world-leading publishers' data (WorldCat Discovery. 2019).

7.3 Ebsco Discovery Service (EDS)

EBSCO is a pioneer in providing information centers and libraries with resources including the e-Resources management system; Ebsconet®, the world's first free online research service; EBSCOhost® including full-text databases, the topics guide, the medical reference point of care, historical digital archives, and e-books. EBSCO provides more than 375 research databases.Approximately 450,000 e-books, and subscription management services for more than 360,000 unique magazines, and 57,000 electronic journals. EBSCO also serves the content for all kinds of researchers through an information center of tens of thousands of full-text newspapers and magazines from well-known publishers (academic content, medical, pre-university, public and academic libraries, government agencies, etc.). EBSCO also offers an EBSCO Discovery Service (EDS)that provides each organization with a fast search engine to reach its full range.The EDS service is superior to other research services,

offering superior evidence and full-text search of newspapers and magazines (EBSCO. 2019).

SDL information services department explores that EDS is an engineering research source and training. The database includes content that supports and gives the information needed by engineers at all levels. It includes research, planning, product development, etc. The collections provide complete full-text coverage, which is relevant to many engineering disciplines which include 3000 publications with full text and relevant journals in engineering source, research and studies catalog, conference papers books and others (EBSCO. 2019a).

7.4 Vufind Discovery Service

There are many challenges and difficulties for institutions as they work to deal with BD, and the most prominent of these challenges can be summarized below:

Arabic portal for discovery and integrated search combining the features of metasearch (to simultaneously search all the sources of the library at the same time, different formats, places, or ports, paper or electronic).It providing the services of the general social catalog, where rich social services such as branding, evaluation, reviews, citations, and comments are available to users(VUfind Arabic2019).The Arabic portal for discovery and integrated search consists of a range of distinctive services that help with the adaptability and immediate operation, such as browsing the topic headers, supporting content enrichment, supporting all citation management tools, as well as supporting

RSS, QR, code. In addition to the Arabic Unified Search Portal,it supports all types of searchessuch as wildcard search, fuzzy search, spatial search. It can be said that the portal complies with exploration systems and software (EDS/Primo/Summon/WorldCat), Vufind works on the Apache Lucent Solr, Solr main features include a powerful full-text search engine, multi-faceted search, dynamic data aggregation, database integrity, rich handling of documents (such as Word, PDF), and spatial and temporal search. Solr can also be distributed if you need to distribute the index load on more than one server, or if it is distributed on a farm with multiple servers.

Solr is highly capable of modifying its size, providing distributed search, index copying,and supporting navigation features across many of the world's largest and most popular internet sites.

Vufine is one of the most valuable library application in KSA. Majmaah University unified catalog of libraries, the catalog of KSA public libraries on the Arabic Koha system, the digital repository of Taibah University, Library Research Guides Plus, Arabic CORAL, Arabic ORCID (Open Researcher and Contributor ID) are the users of Vufind application.

7.5 Summon Discovery Service

The summon discovery tool by ProQuest aims to improve the value of libraries by giving extraordinary research experience. ProQuest-summon is not only offering a single search engine of all library resources. It also makes a strong professional relationship with the library and users. It helps users to keep their resources more discoverable and provide ground to connect with their library professionals. Summon results are more

comprehensive that strengthen the role of the learning resource center (ProQuest. 2019).

King Abdulaziz City of Science and Technology uses the discovery search engine tool (summon) as a central search engine in the global information resources common to it under the name of "One Click", which includes the titles of scientific documents for more than 7800 publishers and 94000 journals and periodicals containing a total of 5 million scientific electronic documents. In the summon discovery tool user can include several features during the search process where the user interface contains one field of search that is similar to the design of the Google search engine. This one search box improves the quality of performance and escalates access to global information sources. (KACST Library, 2019)

7.6 Encore Discovery (Innovative interface)

Innovative interfaces developed the Encore Discovery tool, "a new union search and access tool" launched on 26 May 2006, that "Control Web 2.0 technologies and builds on the foundation of the Millennium integrated library technology platform" (A. Fifarek 2007). The Millennium integrated library system a number of subsystems that are dealt with through the interfaces of information professionals.The Millennium cataloging subsystem; the Millennium circulation subsystem; the Millennium acquisitions subsystem; The Millennium serials sub-system; and the online catalog subsystem (OPAC), The Millennium system uses the Inspire discovery tool which established and operated by the 27 Member States of the European Union, this discovery services search for specific data and specific services on the basis of the asset of consistent metadata and show the metadata asset. (Initial Operating Capability Task Force for Network Services, 2011)

This system has evolved and taken another name which called the Sierra system is designed to go beyond the traditional functions of academic and research libraries and to enable the library specialist to achieve their strategic tasks represented in providing a superior knowledge service to the beneficiaries. The Sierra library management system provides the latest operating options, databases, and open software design to provide an integrated system that facilitates everyday tasks and makes them more efficient and integrated with other

systems. It is a discovery tool used in many KSA universities, The beneficiary found it easier and more efficient.The use of the additional system in Sierra called Encore that access all linked resources, such as electronic or printed material, and users cansearch and explore in order to achieve research needs. King Fahd University (KFUPM), Bisha University, Burayhah Colleges and Umm Al Qura University using the Encore discovery toolprovide their information services.


Encore discovery tool enables libraries to manage the latest operating options, databases and open software design, that provide an integrated system to facilitate the performance of traditional daily tasks and make them more effective and integrated with other systems.The Sierra system is highly efficient in the management of printed and electronic materials, integrated systems for digital asset management, SQL relational database, connectivity with other applications using API. It hasthe ability to work on smart devices.Integrated management screen provides a complete image of library operations that application is easy-to-use and allows staff to more flexible and effective in issuing periodic reports (A. Fifarek 2007).

7.6 Primo Central (Ex Libris Group)

Ex Libris group developed its next-generation library resources search engine called "primo" in 2005 and officially they released in 2007. The 3rd version was released in 2010 from that on thousands of libraries started to subscribe primo. The primo has launched its platform for libraries to integrate scattered resources and fit into one search box. It can harvest local library bibliographic collection, local library digital repository resources, and blended with remote index results. Late 2010, they extended and improve the primo discovery by given large pre-harvested index of articles till level. From 2010 to on, Ex Libris releasing primo interface and

application enhancements almost every 3 months and major comes after 14 to 16 months (J. Vaughan. 2011).

Early fifty subscribed universities got the change to search 300 million items on primo central index platform that was obtained from different world largest publishers like Association of Computing Machinery, Gale, Ebrary, BioOne, IGI Global, Oxford University Press, Lexis Nexis, Springer and many more. Ex Libris obtained a central index of open-source publishers like Directory of open access Journals (DOAJ), arXiv.org, HathiTrust, and others and incorporated with Primo Central. By the time the primo central index is growing (J. Vaughan. 2011). The primo is one of the mega aggregators of open access and
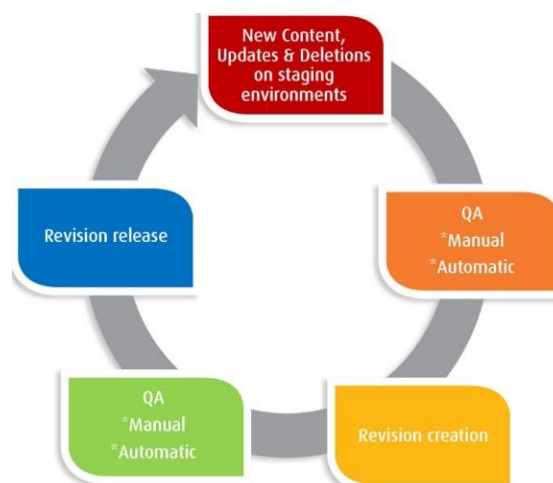

commercial content. Primo discovery is broad and deep that covers eBooks, articles, newspapers, audio files, bibliographic records of printed books, video files, legal documents, government files and many more. Ex Libris expert librarians keeping primo central index updated by carefully adding and deleting resources. The process enriching primo central index illustrated (J. Vaughan. 2011; Primo Content Index, 2019).

7.7 Internet search engine

Search engines are the electronic resources that help internet users to find online information available on the web. Users can navigate the web by using internet search engines by using different algorithms and techniques (Amanda Spink, Dietmar Wolfram, B. J. Jansen, and TefkoSaracevic. 2001), World Wide Web's (WWW) search engines are the most important tools for searching and retrieve

online information, Google is most famous search engine on the internet, The word Google was launched in the 1990s on a search engine on the computer and is intended to search for information on a subject, People like to search their needed information on Google as it offers smart interface, structures and ease of use functionality. It performs better than the other search engines (M. Kaur, N. Bhatia, and S. Singh 2011).
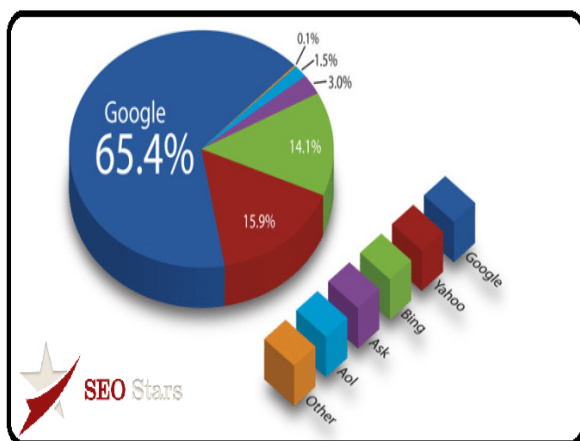
In the higher education institutions, Google scholar and KSA Digital Library (SDL), are the most used search engines by faculty members in KSA universities.

8.   INFORMATION   LITERACY USAGE OF DISCOVERY SERVICES FOR BD & BARRIERS

These days, ignoring data may be costlier than the maintenance process costs. The delivery of real-time or quasi-real information is one of the distinguishing features of massive data analysis. Therefore, idle is avoided whenever and wherever possible, it has become necessary to create a culture that respects the data and makes data analysis an integral part of every decision taken by the academic institution.

With the drastic growth in an academic publication, it is very hard to acquire and manage BD for every library. Libraries in the world are producing and managing their own content and possibly linking their data with other libraries. There is not a suitable plan to integrate all libraries BD. Millions of corporate companies like Proquest, IEEE, Emerald, Elsevier and etc. are managing digital research publications and sailing it to academic, government and law bodies. ProQuest claims to have 90,000 authoritative sources, more than 6 billion digital pages that cover the spans of

six centuries, 20 million web pages and 450,000 electronic books(ProQuest, 2019). There are many more digital publishing companies clams of having billions of digital literally pages. Most libraries subscribing different content from a different publisher and linked all publisher data and local content with one platform called discovery tool. The usage of discovery tools and finding resources from BD required training.

The teaching of information literacy in academic libraries has been changed from the past generation. Instruction librarians have chosen different approaches to guide users to search, locate, and identify reliable and relevant information. Most librarians are offered a training session and explain the discovery services (Seeber, K.P. 2015). During information literacy sessions, instruction librarians for the sake of improving users' search skills, they explain information retrieval (IR) techniques like Boolean operator, focusing/limiting a search, phrase searching, truncation/wildcard searching, fuzzy searching, proximity searching, weighted searching, query expansion, etc.

The time constraint is given for LRC orientation, information literacy, and information research skills, aninstruction librarian rather got a short opportunity to demonstrate discovery services features, IR techniques and complex informative steps within one session (Cmor, D. and X. Li 2012). Dartmouth college one of the early adopter of LDS, they selected summon and done thorough evaluation and implementation in 2009. Even after given library discovery information service usage instruction, they realized that half of our users still not preferred to use discovery tools to filter down BD(Schuetze, S. 2018).

The boolean operating system is one of the best and very commonly found in any LDS. It still has some barriers. It is lacking to rank retrieved documents (Irfan, S. and B. Babu, 2016).

9. FINDINGS

Many information professionals and education leaders believe that the spirit of education, especially higher education, is threatened by the advances in technology or bythe globalization of the education markets and the entry of new elements into the arena of educational programs. Digital culture threatens the characteristics of the heart of higher education, which includes the social life of students in society.In professional circles and in the vitality of the mind; it includes the promotion of

social mobility and the provision of safety for dialogue and educational excellence

BD can help provide insights to support student's learning needs. For instance, learning analytics as a fundamental component of BD in higher education provide researchers with opportunities to carry out real-time analysis of learning activities. By performing retrospective analysis of student data, predictive models can be created to examine students at risk and provide the appropriate intervention, hence enabling instructors to adapt their teaching or initiate tutoring, tailored assignments and continuous assessment (Ben, Daniel. 2015).

There are many opportunities and challenges facing higher education institutions in the use of BD exploration tools.The current techniques and technologies applied to higher education institutions in KSA are constantly evolving to raise information awareness.The importance of these tools in supporting higher education by providing information and data that help institutions in planningeconomic and social development. This importance also creates the knowledge society which is the vision of the KSA 2030. There are many possible techniques to solve the huge data problem at higher educational institutions in KSA.Cloud computing is supporting data centers and information centers at universities, information security, infrastructure solutions and, unified collaboration.Although these technologies do not still under development.In the near future, there is no doubt that today's and future BD problems will benefit from such technological advances (C.L. Philip Chen, and C.Y. Zhang, 2014).

According to Deloitte 2016, more than 3,000 government officials from around the world attended a World Government Summit in Dubai.Universities in the Middle East, including KSA higher education institutions,to provide education in a unique style for students that do not fit the different requirements.To deal with the challenges associated with increasing student numbers and the need for digital transformation. The summit stressed indicators and forecasts resulting from student-specific analyses can transform the services offered to students to become more personalized, proactive and inexpensive compared to individual mentoring sessions. This is allowing special mentors, and teachers to focus on students who need to follow-up. This will greatly

contribute to the data and information exploration tools by empowering students, researchers and all members of higher education institutions to have quick access to information and knowledge that serve their research and educational purposes.They can achieve better results in their scientific and educational careers (WorldGovernmentSummit, 2019).

10. ACKNOWLEDGMENTS

11. REFERENCES

[1] A. Fifarek (2007). "The Birth of Catalog 2.0: Innovative Interfaces' Encore Discovery Platform," *Library Hi Tech News,* vol. 24, no. 5, pp. 13-15.

[2] Ahmed, Shakeel, and Hemant,K. Mehta (2015). "On Applying Big Data and Cloud Computing for Quality Improvement in Higher Education." Proc. of Int'l Cont. on Advances in Big Data Analytics.

[3] A. Karadia and S. Pati. (2015). "Discovery Tools and Services for Academic Libraries.

[4] Ali, Alklabi (2018).The importance of big data analysis in decision-making at King Saud University: a case study on "Mastering" system at King Saud University,The 24th SLA/AGC Annual Conference & Exhibition:Big Data and Its Investment Prospects:The Way towards Knowledge Integration, 6-8.

[5] Amanda Spink, Dietmar Wolfram, B. J. Jansen, and Tefko Saracevic. (2001). "Searching the Web: The Public and Their Queries". Journal of the American Society for Information Science and Technology, Vol. 52 Issue 3.

[6] A. Mukthar and M. Sultan. (2017). "Big Data Analytics for Higher Education in Saudi Arabia," International Journal of Computer Science and Information Security, vol. 15, no. 6, pp. 256-277.

[7] Ben Daniel. (2015). "Big Data and analytics in higher education: Opportunities and

challenges".British Journal of Educational Technology, pp 904–920, Vol 46 No 5.

[8] B. Jacob, M. Brown, K. Fukui, and N. Trivedi (2005). "Introduction to grid computing," IBM Redbooks, pp. 3-6.

[9] Breeding. M. (2010). The state of the art in library discovery in Computers in libraries. 2010. p. 31-34.

[10] Building digital capability, JISC. (2019). [Online]. Available: https://digitalcapability.jisc.ac.uk/. [Accessed: 21- Apr- 2019].

[11] C.L. Philip Chen, &C.Y. Zhang.(2014). "Data-intensive applications, challenges, techniques, and technologies: Asurvey on Big Data," Information Sciences journal. vol. 275,pp.314 –319.

[12] Cmor, D. and X. Li (2012). Beyond boolean, towards thinking: Discovery systems and information literacy. Library management, 33(8/9): p. 450-457.

[13] C. Pettey, (2012). "Gartner says big data creates big jobs: 4.4 million IT jobs globally to support big data by 2015," in Analysts Discuss Key Issues Facing the IT Industry During Gartner Symposium/ITxpo, 2012.

[14] EBSCO. (2019). Improving research around the world, Information Services, Inc. www.ebsco.com, 2019. [Online]. Available: https://www.ebsco.com/. [Accessed: April 2019].

[15] EBSCO. (2019a). EBSCO Electronic Information Services Launches Engineering Source™ for Engineering research and training, ", Saudi Digital Library, 2019. [Online]. Available: https://sdl.edu.sa/SDLPortal /en/post.aspx?p=2684. [Accessed: 22- Apr- 2019].

[16] In-depth learning and big data. (2016). Retrieved January 22, 2020, from https://www.wikiwand.com/ar/

[17] Irfan, S. and B. Babu (2016). Information retrieval in big data using evolutionary computation: A survey. in 2016 International Conference on Computing, Communication, and Automation (ICCCA). IEEE.

[18] J. Vaughan. (2011). "Ex Libris Primo Central," Library Technology Reports, vol. 47, no. 1, pp. 39-47.

[19] Krause, M. (2016). How Library Discovery Platforms Have Changed Teaching Interdisciplinary Research, in Law Library Lights. p. 1.

[20] L. Arthur. (2015). "What Is Big Data?", Forbes.com, 2019. [Online]. Available: https://www.forbes.com/sites/lisaarthur/2013/0 8/15/what-is-big-data/#45fa1cf35c85. [Accessed: April 2019].

[21] Mary, Pagliero (2012). PoppPlanning and Implementing Resource Discovery Tools in Academic Libraries, Information Science Reference.

[22] Meshael, Sultan (2017). Ayesha Mukthar. Big Data Analytics for Higher Education in Saudi Arabia.The International Journal of Computer Science and Information Security (IJCSIS), Vol. 15, No. 6.

[23] M. Kaur, N. Bhatia, and S. Singh (2011). "Web search engines evaluation based on features and end-user experience," International Journal of Enterprise Computing and Business Systems, vol. 1, no. 2, pp. 1-19.

[24] Naseej.com (2019). Data mining and knowledge extraction,Blog.naseej.com, 2019. [Online]. Available: http://blog.naseej.com/2014/03/27. [Accessed: April 2019].

[25] Parry, M. (2010). After losing users in catalogs, libraries find better search software. The Education Digest, 75(6): p. 54.

[26] Primo Content Index, (2019). Ex Libris. [Online]. Available: https://www.exlibrisgroup.com/products/primo library-discovery/content-index/. Accessed April 2019.

[27] ProQuest. (2019). *Who We Are*. Available from:https://www.proquest.com/about/who- we-are.html.

[28] ProQuest. (2019a). Products & Services - The Summon® Service, Proquest.com, [Online]. Available: https://www.proquest.com/products-

services/The-Summon-Service.html#returntop. [Accessed: April 2019].

[29] Schuetze, S. (2018).*Impact of Library Discovery Services: The Surveys Speak*. [cited 2018 12 December]; Available from: https://www.exlibrisgroup.com/impact-of-library-discovery-services-the-surveys-speak/.

[30] Seeber, K.P. (2015). *Teaching "format as a process" in an era of Web-scale discovery.* Reference Services Review, 43(1): p. 19-30.

[31] Sutton, Mark. "Analytics sets directions in education."ITP.net . ITP.net, 2 Apr.2016. Web. 24 Apr. 2017.

[32] T. Hey and A. E. Trefethen. (2002). "The UK e-science core program and the grid," in International Conference on Computational Science, pp. 3-21: Springer.

[33] UNESCO. (2015). Director-General, UNESCO science report: towards 2030, 2nd ed. United Nations Educational, Scientific and Cultural Organization, 2015.

[34] VUfind Arabic (2019). Kwareict, Kwareict.com, 2019. [Online]. Available: http://www.kwareict.com/ar/taxonomy/term /18?page=1. [Accessed: April 2019].

[35] WorldCat Discovery. (2019). OCLC, 2019. [Online]. Available: https://www.oclc.org/en/worldcat-discovery.html. [Accessed: April 2019].

[36] WorldGovernmentSummit, (2019). World Summit of Governments: Executive Education Sessions [Online]. Available: https://www.worldgovernmentsummit.org/ar/ [Accessed: April 2019].

[37] W. Discovery. (2019). Imagination thrives on access to knowledge. Available: https://www.oclc.org/content/dam/oclc/service s/brochures/215405-WWBE_WorldCat-Discovery-Product-Brochure.pdf, accessed April 2019.

[38] X. Shacklock. (2016). "From Pricks to click: The potential Data and analysis in Higher Education," p.76 Available: https://www.policyconnect.org.uk/hec/sites/sit e_hec/files/report/419/fieldreportdownload/fro mbrickstoclicks-hecreportforweb.pdf