

The Development of Indonesia One Search: Retrieving Information Resources in Indonesia

Muhammad Azwar

*Faculty of Adab and Humanity
Syarif Hidayatullah State Islamic University Jakarta
Jl. Ir. H. Juanda No.95 Ciputat, Indonesia
muh.azwar@uinjkt.ac.id*

Abstract— The article discusses about the development of Indonesia One Search portal. It is applied to retrieve a large number of references in all libraries, journals, and repositories from various institutions in Indonesia. The information resources are obtained by harvesting technique from membership online repositories. With Indonesia One Search, every people can easily access digital collection by watching and downloading full text, video learning materials, and etc. The goal of the article is to acknowledge the development of the One Search in the future. This is a qualitative study with a descriptive approach. The key person of this research is Ismail Fahmi, Ph.D, the initiator of Indonesia One Search, and it is also supported with various literatures relating to the digital library technology. This portal develops through three stages. The first stage is the bibliographical metadata gathering process from various repositories (started in 2015). The second stage involves the crawling PDF full text in every related collection, the text analysis using NLP technology (Natural Language Processing), information extraction & knowledge mapping based on NLP, and the research mapping among institutions (2017-2018). The third stage involves the system and anti-plagiarism services by applying the plagiarism clarification (2019-2020). Indonesia One Search website answers the complications experienced by the learners and studious society in obtaining the access of scientific information by establishing an open and integrated access from various institution and preventing the possibility of plagiarism which is currently occurring in Indonesia.

Keywords— Portal of Indonesia One Search; Retrieving; Information Resources; Libraries; Journals; and Repositories.

I. INTRODUCTION

The current information age encourages every institution to apply technology for various conveniences. Institutions seek to provide information services in the form of scientific information searches online. With their libraries, they display the Online Public Access Catalogue (OPAC), publish electronic journals based on Open Access Journal (OJS), and in particular place their publications in repositories and some other institutions display museum collections and even archives.[1]

Indonesia has an enormous number of libraries. There are approximately more than twenty five thousands of libraries spread across Indonesia, consisting of twenty-two thousand school libraries, one thousand five hundred public libraries, one thousand specialized libraries, and eight hundred and forty-five college libraries. Unfortunately, not all libraries are able to present the technology and display all the information available in the library.

In a library institution, there are many information search

portals, such as OPAC, digital libraries, e-resources, as well as e-journals. However, all these portals are separate from each other, so users must visit each portal one by one when wanting to access it. There is no connection between libraries that enables the users to take advantage of collections across libraries. Similarly, this also happens to the information resources owned by universities, such as repositories, electronic journal managers, archival managers, and museums. The scientific information presented in the repository, e-journal, archive, and museum by each institution can only be enjoyed by the community separately.

INDONESIA ONE SEARCH

The development of technology and information continue to impact on the growth of information resources and library service facilities. By utilizing information technology, libraries need to establish an information and knowledge power centre.[2]

The Indonesian regulation no. 43 year 2007 on the library service improvement mandate is done by utilizing library network system based on information and communication technology. With the cooperation between libraries, it is expected that the limitations can be overcome so that the library can access a wider range of information resources according to the needs of the users.[3]

The need for a library network system based on information and communication technology is due to the fact of significant growth in printed and electronic publications, the growth of science and technology, the demands of users, the effectiveness and efficiency of time, energy, funds, and various other resources. The establishment of this network is important that users can optimally access it quickly and easily even remotely and it provides with up-to-date information that can be used flexibly to meet their needs.

Since March 2015, the National Library of Indonesia has initiated to build a digital library network portal that is "Indonesia One Search" which is expected to answer various obstacles that are still found in the existing digital library network. The socialization of Indonesia One Search is still intensively conducted in the hope of getting responses and inputs to improve the capacity of information resources and expand access.[4]

Indonesia One Search is a web portal of bibliographic and a full-text collection of which data is collected through harvesting methods from the member's online repository. This portal can also be referred to as "Summon" for the entire library of electronic libraries in Indonesia. Similarly, to Summon and WorldCat which index bibliographic collections and present them through a single search interface, Indonesia One Search also allows users to search and access the entire collection of electronic libraries in Indonesia through a single door. With Indonesia One Search, everyone can

easily access the various digital collections of all libraries in Indonesia incorporated in One Search in various formats, download full text, view videos, listen to audio, access learning materials, and etc.[5]

Indonesia One Search is able to integrate various existing information resources in the form of OPAC, e-Journal, repository, archives, and museum. This portal has the ability to integrate a variety of collections from various applications (multi-platform). Library catalogues or OPAC use applications such as, Senayan Library Management Systems (SLiMS), KOHA, Inlis Lite, and My SIPIS Pro. Digital libraries include GDL, Green Stone Digital Library. Applications for repositories cover, such as Eprints, and DSpace. Applications for e-journal management are the Open Access Journal (OJS). Applications for archives comprises of Omeka, ICA-ATOM, and ATOM. Omeka appears as apps for museums. These overall applications can be integrated in the Indonesian One Search portal. The official website of Indonesia One Search can be accessed through <http://onesearch.id>.

VUFIND

Indonesia One Search portal uses an Open Source-based app called VuFind. This VuFind app is a portal designed and developed for library development. The purpose of VuFind is to make it easier for library users to search and browse all the information resources in libraries, including libraries, journals, digital libraries, and Institutional Repository, Institutional Bibliography, and other collections and information resources.[6] VuFind is made and designed modularly, so that the users can implement the system in a basic way or use all the components in it. Due to its open source nature, its users can modify or add modules as needed.[6]

This VuFind has a fairly high interoperability because it uses OAI-PMH or the Open Archive Initiative's Protocol for Metadata Harvesting protocol. OAI-PMH is a protocol that can take (harvest) various repository metadata used to share information resources scattered across institutions. This protocol is very useful in integrating information resources from various applications.

In addition to having good interoperability, VuFind is tested in certain domains, such as its ease to use (user friendly), its powerful search engine, its potential to be enriched with faceted search, its complete information, support to mobile devices, provision of information about authors, citation indexes, and statistics that are important and interesting for contributors. This application is also supported by a system that allows portals to grow for the long term (sustainable). VuFind official website can be accessed <http://vufind.org> [7]



Fig. 1: Portal Of Indonesia One Search

Indonesia One Search Portal has been modified in such a way by the development team initiated by Ismail Fahmi as a consultant of National Library. This portal is not merely aimed

at integrating all information resources in Indonesia, but it also has several other purposes. Since its launch in 2015, the portal has a road map. This paper discusses what kind of development of Indonesia One Search portal will be in the future.



Fig. 2: Official Website Of VuFind

II. METHODS

This research is a qualitative method with a descriptive approach.[8] The subject in this research is Ismail Fahmi as consultant of National Library of Indonesia, and the initiator of Indonesia One Search Portal. The information on the development Indonesian One Search portal is explored directly by researchers as a research instrument through interview techniques, participatory observation, and documentation. They observe the Portal of Indonesia One Search directly through <http://onesearch.id>.

The data were analysed since they were collected during the study through describing and presenting the data reduction level, and category; enriching explanation of findings, drawing temporary conclusions in each of the findings, arranged and categorized in the form of relevant themes. The data validity and credibility were done by extending observations, increasing perseverance, triangulation, peer discussions, and reference adequacy.[9]

III. RESULT AND DISCUSSION

Indonesia One Search Portal is still in development progress. There are 3 phases of development, namely: phase 1 (2015-2016) One Search Basic, phase 2 (2017-2018) One Search Advanced, and phase 3 (2019-2020) No Plagiarism as described on Table 1.

TABLE 1

DEVELOPMENT MAP OF INDONESIA ONE SEARCH		
Stage 1 : One Search Basic	Stage 2 : One Search Advanced	Stage 3: No Plagiarisme
Input : <ul style="list-style-type: none">• Bibliography• Perpusnas thematic databases	Input : <ul style="list-style-type: none">• PDF Fulltext	Input : <ul style="list-style-type: none">• PDF Fulltext• Wikipedia (Indonesian)• Online News (Indonesian)
Process : <ul style="list-style-type: none">• Harvesting• Search & facet Bibliography indexing	Process : <ul style="list-style-type: none">• PDF Crawling• PDF to text and image converting• Fulltext indexing• Linguistics processing• Semantic indexing• Content analysis	Process : <ul style="list-style-type: none">• Wikipedia crawling• Online news crawling• Document finger print indexing• Similarity analyzer• Similarity report builder
Output : <ul style="list-style-type: none">• Search• Bibliography• Search Summon OAI Manager	Output : <ul style="list-style-type: none">• View fulltext online (e-reader)• Content: analysis and research• Citation index View Collection and usage statistics.	Output : <ul style="list-style-type: none">• Document upload• Document similarity detection report admin

Stage 1: One Search Basic

Phase 1 is implemented for 2 years from 2015 to 2016. This stage is the basic stage. This stage includes harvesting metadata activities, preparing interoperability between applications using OAI-PMH, metadata prefix using MARC XML and OAI_DC, specifying server prototypes for web portal needs, building database and registration repository form, and designing complete and analytical reports.[10]

This basic stage is the process of collecting (harvesting) bibliographic metadata from various institutions spread throughout Indonesia which includes Information Library System (ILS), such as INLIS Lite, SLiMS, KOHA; Journal, such as OJS; Digital Library and repository, such as DSpace, Eprints, Omeka; and summon, the integration of searches to several online databases subscribed by the National Library. At this stage, there are also some non-technical activities, such as management strategy and reliability (sustainability) of Indonesia One Search portal, building and receiving input from the community, form teams, and building interagency cooperation and socialization.

Stage 2: One Search Advanced

Phase 2 is implemented for 2 years from 2017 to 2018. This stage is an advanced stage. This stage includes PDF full text crawling activities, application of Natural Language Processing (NLP) technology, content analysis to full text, authority index, and citation parsing and indexing. Indonesia One Search crawls PDF Full text from all of thesis, dissertation, research report, journal articles, and other resources from all institutions in Indonesia. It makes easy for end-users to find information resources with PDF full text format.[11]

NLP or Natural Language Processing technology is a field of computer science, artificial intelligence and computational linguistics concerned with the interactions between computers and human natural languages. The main goal of NLP is making a machine that can understand human language and then give suitable response. Indonesia One Search will be developed in order to analyse content using the Natural Language Processing. Information will be extracted and knowledge will be mapped based on NLP. It helps end-users to find information needed and also analyse it easily.[12] At this stage there also some non-technical activities, such as socialization and campaigns on the importance of full content sharing, and demonstration of the benefits derived from content analysis for users. An important goal of this second phase is that users get a lot of convenience in searching for information and research.[13]

Stage 3: No Plagiarism

Phase 3 is implemented for 2 years from 2019 to 2020. This stage is the stage of No Plagiarism. This stage involves building a plagiarism detector system, indexing news, Wikipedia, blogs, and full-text in the One One's portal for plagiarism detection, providing server infrastructure and data centres for plagiarism detectors, building interfaces, as well as for college users in higher education, research institutions, and individuals.

At this current stage, some non-technical activities are implemented, such as the trial and socialization of No Plagiarism towards lecturers and students in universities, and socializing and campaigning against plagiarism. An important goal of this phase is to build a reliable system of plagiarism detectors and be able to index all articles and information not only English but also Indonesian. This will greatly help detect and prevent plagiarism in Indonesia.[14]

IV. CONCLUSION

Indonesia One Search is a search portal for all public collections from libraries, museums, archives, and electronic sources in Indonesia. This portal is expected to be a solution that allows users to find the source of information because it is only through one door. This portal is through 3 stages of development, namely basic, advanced, and no plagiarism. Through this development, it is expected that users receive various facilities in searching for information, searching for relevant content, and avoiding plagiarism.

REFERENCES

- [1] P. L. Pendit, *Perpustakaan digital : perspektif perpustakaan perguruan tinggi indonesia*. Jakarta: Sagung Seto, 2007.
- [2] H. Darmawan, "Perpusnas Luncurkan Portal Jejaring Perpustakaan Indonesia One Search," 2016. [Daring]. Tersedia pada: <http://www.pnri.go.id/2016/03/perpusnas-luncurkan-portal-jejaring-perpustakaan-indonesia-one-search/> [Diakses: 16-Okt-2017].
- [3] G. Chowdhury dan S. Chowdhury, *Introduction to digital libraries*. London: Facet Publishing, 2003.
- [4] "Undang-Undang Nomor 43 Tahun 2007 tentang Perpustakaan." 2007.
- [5] "Perpustakaan Nasional RI, "onesearch [Indonesia OneSearch Wiki]," 2015. [Daring]. Tersedia pada: <http://wiki.onesearch.id/doku.php?id=onesearch> [Diakses: 16-Okt-2017].
- [6] J. Emmanuel, "Usability of the VuFind Next- Generation Online Catalog," *Inf Technol Libr*, vol. 30, no. 1, hlm. 44–52, 2012.
- [7] Sugiyono, *Metode Penelitian Pendidikan*. Bandung: Alfabeta, 2010.
- [8] L. J. Moleong, *Metode Kualitatif*. Bandung: Remaja Rosdakarya, 2004.
- [9] Sugiyono, *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. Bandung: Alfabeta, 2010.
- [10] I. Fahmi, "Indonesia Onesearch: Registration, Harvesting, Indexing, Searching, a...," [Daring]. Tersedia pada: https://www.slideshare.net/IsmailFahmi3/indonesia-onesearch-registration-harvesting-indexing-searching-and-community-virtual-hosting?qid=971816ea-d193-4ce4-bb76-0415f6d025b9&v=&b=&from_search=14 [Diakses: 16-Okt-2017].
- [11] I. Fahmi, "Repositori Akses Terbuka di Indonesia," 2016. [Daring]. Tersedia pada: https://www.slideshare.net/IsmailFahmi3/repositori-akses-terbuka-di-indonesia-pdf?qid=9d831886-dee1-45ee-8c23-271fa8954e09&v=&b=&from_search=20 [Diakses: 16-Okt-2017].
- [12] A. Chopra, A. Prashar, dan S. Chandresh, "Natural Language Processing," *Int J Technol Enhanc Emerg Eng Res*, vol. 1, no. 4, hlm. 131–134, 2013.
- [13] I. Fahmi, "Pengembangan Repository Perpustakaan Melalui Indonesia OneSearch," 2017. [Daring]. Tersedia pada: https://www.slideshare.net/IsmailFahmi3/pengembangan-an-repository-perpustakaan-melalui-indonesia-onesearch?qid=971816ea-d193-4ce4-bb76-0415f6d025b9&v=&b=&from_search=16 [Diakses: 16-Okt-2017].
- [14] I. Fahmi, "Indonesia OneSearch: Satu Pintu Pencarian Untuk Seluruh Perpustakaan ...," 2015. [Daring]. Tersedia pada: <https://www.slideshare.net/IsmailFahmi3/indonesia-onesearch-satu-pintu-pencarian-untuk-seluruh-perpustakaan-berbasis-slism-di-indonesia> [Diakses: 16-Okt-2017].