Volume 4, General Issue (2017)

A Different Shade of Green: A Survey of Indonesian Higher Education Institutional Repositories

Toong Tjiek Liauw, Paul Genoni

INTRODUCTION Institutional repositories (IRs) are an accepted part of the open access landscape, and they have a particular role to play in supporting scholarly communication in developing countries, such as Indonesia. METHODS Content analysis was conducted of 52 Indonesian higher education institutional repository websites between November 2014 and February 2015. Assessment included the degrees of “openness” of repositories, the types of works collected, software used, exploration tools, existence of links to institutional website, the language used for access points, and the standard of metadata. The study also gathered qualitative indicators of local practices in the management and population of repositories. RESULTS Only 26.9% of the surveyed IRs provide all or most documents in full-text; the most widely included types of work are Theses and Dissertations (84.6%) and Published Works (80.8%), but there is also a high representation of Unpublished Works and University Records. Most IRs (90.3%) provide access points in the form of standardized subject headings, and English is widely used. DISCUSSION The characteristics of the content of the IRs surveyed suggests that many Indonesian IRs were conceived as a corporate information management system rather than as a genuine attempt to support open access. CONCLUSION The findings lead the authors to speculate that institutional repositories serving Indonesian higher education institutions are in their early adoption phase; and that initial drivers for them have been corporate information management, institutional prestige, and the need to combat plagiarism.
IMPLICATIONS FOR PRACTICE

1. Provides the first detailed survey of Indonesian higher education institutional repositories
2. Informs the field on the various ways institutional repositories have been used and local practices in Indonesian higher education sector
3. Informs the field on the efficacy of content analysis when applied to web-based dynamic contents

INTRODUCTION

Advances in information and communication technology have brought extraordinary changes in many areas of our personal and workplace endeavors, some of which have challenged and even transformed established ways of thinking and communicating. The scholarly community has been far from immune to these changes, and in many respects has been at the forefront of accepting new technologies and their sometimes unpredictable consequences. The emergence of the Internet as the ubiquitous point of access to scholarly publishing has, in particular, hastened the inevitable rise of digital publishing that has simultaneously both entrenched and challenged the hegemony of commercial publishers in the scholarly publishing sphere.

Academic communities, partly driven by concern about the increasing domination of scholarly communication by globalized commercial publishers, have devised and supported the concept of open access and put considerable resources into what has become known as the Open Access (OA) Movement. In a nutshell, the OA Movement has attempted to support both authors and users of scholarly communication by leveraging the power of digital storage and communication to cut out the middleman. An early milestone for the OA Movement was the Budapest Open Access Initiative Declaration (BOAID) in 2002. BOAID recommended both open-access journals (the so-called ‘Gold OA’) and author self-archiving (‘Green OA’) as strategies designed “to achieve open access to scholarly journal literature” (Chan et al., 2002, para. 5 & 6). Another significant development came with the 2003 conference convened by the Max-Planck Society that issued the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (Max-Planck Society, 2003). The Berlin Declaration is noteworthy in that whereas previous OA initiatives had foregrounded the Science disciplines, on this occasion the scope of OA was specifically refocused to include the Humanities. The Berlin Declaration also introduced the term ‘open access contributions’ instead of ‘open access publications,’ and thereby further broadened the scope of scholarly/scientific information that came within the purview of the OA Movement (Liauw, 2013). These developments in OA have supported an alternative form of scholarly communication...
that has enabled—to a certain degree—the availability of freely-accessible online scholarly content. As a result the strategy of self-archiving and the associated emergence of institutional repositories (IRs) have become key components of the scholarly communication landscape.

Both OA and IRs present important opportunities for research productivity and research communication in developing countries. As a result of the international adoption of IRs developing countries have access to scholarship they may not otherwise be able to afford. Just as importantly, IRs provide developing countries with the potential to make available the results of their research when they often find it difficult to use the established English language journals that are the standard international platform for the communication of peer-reviewed research. It is therefore important to understand the state of IRs in developing countries.

**Indonesian Higher Education**

Indonesia currently has 100 state higher education (HE) institutions and 2,972 private HE institutions (Direktorat Jenderal, 2014), serving a population of 255 million people in 2015 (Statistics Indonesia, 2015). Even though some academic publishers and/or aggregators provide discounted pricing for developing countries including Indonesia, affordability remains a major obstacle to accessing scholarly content. Indonesian access to commercial databases of scholarly content is typically limited to a small number of larger, state-funded institutions. Even the comparatively better-funded private institutions can afford only a small number of hardcopy journals, and very few can afford to subscribe to commercial journal databases. These conditions exist despite a Government mandate requiring state and local governments to allocate at least 20% of their budget for education (Dewan Perwakilan, 2002, Chapter XIII, Clause 31). Indonesia is therefore a potential beneficiary from OA as a form of free scholarly communication, with some prospect of compensating for the lack of access to scholarly content due to commercial paywalls.

**Institutional Repositories in the Indonesian Context**

Recognition of the importance of IRs in Indonesian HE was emphasized by the creation of Garuda (Garba Rujukan Digital or Digital Reference Portal) by the Directorate General of Higher Education (DIKTI) in 2009 (Farida, Tjakraatmadja, Firman & Basuki, 2015). Garuda provides an online union catalog that consolidates metadata from various HE IRs, and also provides server space and Internet bandwidth to HE institutions that cannot afford to establish and maintain their own IRs. The creation of Garuda was followed by the Ministry of National Education Act No. 17 / 2010, which mandates the use of Garuda or other forms of IRs to “upload electronically all scholarly works by students/lecturers/

The importance of IRs in Indonesia has been further reinforced by the increased attention given since 2006 to the Ranking Web of Universities. The webometrics used to rank international universities claim to provide “reliable, multidimensional, updated and useful information about the performance of universities from all over the world based on their web presence and impact” (Spanish National, n.d.a). In addition to institutional websites, repositories also play an important role in determining a university’s ranking on the Ranking Web of Universities. The importance of this ranking tool in Indonesia was established when the Directorate General of Higher Education commenced using it as an indicator to benchmark Indonesian HE institutions, and was also made apparent by the publication of the ranking on the Directorate’s and universities’ official websites (Universitas Gunadarma, 2012; Universitas Diponegoro, 2012; Koordinasi Perguruan, 2014).

Some local practices in Indonesian HE institutions deserve mention in order to help understand the context of this study:

- Most Indonesian HE institutions require a thesis for all levels of tertiary education, from bachelor to doctorate degree (some institutions even require it at the diploma/non-degree level).
- Theses stored in repositories are usually in separate chapters to compensate for low Internet bandwidth.
- A substantial number of articles available in Indonesian HE IRs are undergraduate theses that had been converted into journal-style articles to comply with DIKTI’s requirement that undergraduate students need to publish in scientific journal(s) as part of their degree (Santoso, 2012a). In clarifying this requirement, DIKTI has stated that the scientific journal(s) referred to in their directive was meant to indicate “online journals” (Santoso, 2012b). The Indonesian HE community interpreted this instruction as DIKTI’s approval to upload undergraduate theses to IRs in the form of journal-style articles summarizing a thesis.
- The majority of journals published in Indonesia are published by HE institutions, although there are also journals published by non-university research institutions and professional associations.

Outside HE institutions, two other national bodies have indicated their support for IRs in Indonesia. Firstly, the National Library of Indonesia has started to develop a national
portal for Indonesian IRs (Indonesia One Search), which was launched in March 2016 (National Library of Indonesia, 2016a). As of August 20, 2016 Indonesia One Search claims that it has harvested 5,262,957 records from 532 repositories across Indonesia (National Library of Indonesia, 2016b). Secondly, the Center for Scientific Documentation and Information—under the auspices of the Indonesian Institute of Science—had conducted a national workshop on “Data, Information, and Knowledge Management in Supporting the Development of Indonesian National Repository” in August 2016 (Lembaga Ilmu Pengetahuan Indonesia, 2016).

These various developments indicate that IRs have gained prominence in the Indonesian HE sector and have the potential to increase exposure to Indonesian research and publishing. This study therefore aims to establish benchmark evidence regarding the current state of IRs and their implementation within Indonesian HE and provide data that might also be reflective of practices in other developing countries as they look to take advantage of OA capability.

**Research Questions**

1. What is the current state of the adoption of institutional repositories in Indonesian HE institutions?

2. What imperatives are driving the current management of institutional repositories in Indonesian HE institutions?

3. What are the local practices of Indonesian HE institutions in populating and managing their institutional repositories?

**LITERATURE REVIEW**

**Overview on Institutional Repositories**

Repository implementation, in general, was one of the recommendations of the BOAID. It was the Green OA strategy that initially adopted the term ‘archive’ (self-archiving), which then evolved into the now widely used term ‘repository’ (Suber, 2016). There are broadly two types of repository: subject repositories and institutional repositories. The most widely known subject repository is perhaps the arXiv, a subject-based repository developed by Paul Ginsparg in 1991 and hosted at the Los Alamos National Laboratory with an initial focus on High Energy Physics, which has subsequently been hosted by Cornell University Library (Cornell University Library, 2013, para. 3 & 4). In addition
to being a subject repository, arXiv has also functioned as a publishing platform in physics and related fields.

Xia (2008) suggested that “[i]n regard to the mechanism of content recruitment, institutional repositories (IR) at their initial stage of development in the early 2000s followed what subject-based repositories (SR) had already practiced for many years ... although recent operations and content materials of IRs have been more diverse” (p. 489). The increased diversity of content means that in addition to published works IRs may include various other types of works, such as grey literature (unpublished works), corporate/institutional records, and special collections (Liauw, 2011, p. 166). The use of IRs to preserve corporate memory seems to be paramount for some organizations, whose corporate/institutional records are also considered to be grey literature (Onyancha, Al-Awah & Cole, 2012, p. 172). The inclusion of content in IRs in addition to formally published scholarly material suggests a departure from the initial intention of repositories as a Green OA response to the crisis in the scholarly communication landscape. However the inclusion of unpublished content does not mean that these repositories necessarily divert from the function of an IR. McDowell’s survey of United States’ IRs (2007, para 22) identified the breadth of materials collected therein:

ETDs [(Electronic Theses and Dissertations)]; e-prints (pre- or post-print articles); working papers and technical reports; conference proceedings and presentations; e-journals and e-books; learning objects; multimedia files (digital audio/video); datasets; pictures (images); digitized archival documents and university records (historical texts and primary sources); non-scholarly institutional publications; undergraduate student work; graduate student work (non-ETD); and course content (syllabi, assignments, lectures).

Taking a similarly broad view of IRs, Lynch (2003, p. 328) concluded that:

a mature and fully realized institutional repository will contain the intellectual works of faculty and students – both research and teaching materials – and also documentation of the activities of the institution itself in the form of records of events and performance and of the ongoing intellectual life of the institution.

Genoni (2004, p. 302) has similarly argued that as IRs are “designed to serve the needs and interests of the institutions that support them ... their content should be developed with local requirements foremost.” These local “needs and interests” may differ both between institutions, and between countries.
Previous national surveys of institutional repositories

In 2006 Oliver and Swain wrote that by assessing the relationship between IR content and investment in research and development “it may be possible to monitor the growth and distribution of innovation geographically around the world” (p. 4). Since that time researchers have produced numerous IR-based studies, including surveys or censuses of the state of IRs in various parts of the world. The Coalition of Networked Information (CNI) undertook a census of IRs by sending an email questionnaire to its member institutions in the United States and has gathered responses from 97 “doctoral universities” and 35 “liberal arts institutions” (Lynch & Lippincott, 2005, para. 5). In the same year CNI with the United Kingdom’s Joint Information Systems Committee (JISC) and the SURF Foundation conducted a census of IRs by sending a questionnaire to institutions in Australia, Canada, the United States, and ten European countries (van Westrienen & Lynch, 2005). In the following year the MIRACLE Project conducted another census of IRs via an online questionnaire targeting “academic library directors and senior library administrators” in the US, collecting responses from 446 institutions (Markey, et al., 2007, p. 13). Between November 2005 and March 2007 another census of United States’ IRs was conducted by analyzing relevant entries or records in several online platforms: DSpace Instances wiki; BePress Digital Commons repositories list; and the Registry of Open Access Repositories (ROAR ) (McDowell, 2007, para. 7). Between June-October 2006 another census of IRs was undertaken by means of a questionnaire distributed to various institutions in Europe, and gathered responses from 114 repositories in 17 countries (van Eijndhoven & van der Graaf, 2007, pp. 9-10). Primary Research Group, a company providing commercial research services, has conducted a series of surveys of international IR activity with a questionnaire to international HE and research institutions (2007; 2011; 2012), and finally yet another study reviewed the “worldwide growth of open access (OA) repositories, 2005 to 2012, using data collected by the OpenDOAR project” (Pinfield et al., 2014).

The early surveys of OA repositories therefore mostly investigated developed countries relying on data gathered from either questionnaires or registries of repositories, such as OpenDOAR, ROAR, and the Ranking Web of World Repositories (RWWR). This initial focus on developed countries is understandable since they were the early-adopters of repositories. More recently, however, similar surveys have started to emerge from developing countries. These include the pan-Asian surveys conducted by Abrizah, Noorhidawati & Kiran (2010); Nazim & Mukherjee (2011); and Tan, Abrizah & Noorhidawati (2013), which variously used data from OpenDOAR, RWWR, ROAR and OAIster. These studies have provided helpful early evidence regarding the developing characteristics of IRs in the Asian region.
Studies on IRs that focus on Indonesia are very limited. Toong Tjiek (2006) and Tjiek (2007) discussed the development of Desa Informasi (Information Village), the IR of Petra Christian University, as a case study. Farida, Tjakraatmadja, Firman & Basuki (2015) produced a study of IRs in Indonesia as part of a broader examination of HE knowledge management practices. In their study the authors conducted a survey in 2014, deriving data from Webometrics and OpenDOAR, that identified 42 Indonesian institutional repositories from various institutions such as HE, research, and hospitals. The survey has given a preliminary account of the status of IRs in Indonesia. The topic of IRs in Indonesia is also discussed by Liauw (2013) in his account of Green OA solutions in the context of scholarly communications in the Indonesian HE sector. There is, however, no previous extensive survey of Indonesian IRs that has been reported.

METHODS

Content analysis

This study utilized content analysis to analyze Indonesian HE IRs websites. As a “systematic analysis of text,” content analysis has a long history dating to the 17th century with church-related studies (theology), which evolved into “quantitative newspaper analysis” in the mass communication era in the beginning of the 20th century (Krippendorff, 2013, pp. 10-13). Content analysis has later been applied to new media, such as the Internet. Neuendorf (2002) considered the use of content analysis as a means of analyzing Internet websites and cited several instances of the emergence of this method of research.

Several limitations in this study need to be mentioned regarding the efficacy of content analysis when applied to IR websites. Firstly, the content analysis was applied to the metadata and documents contained in the IRs to gather information relevant to the characteristics and structure of the IRs. The content analysis was not applied to the individual works to gather information pertaining to the topic or subject of each work. As an example, when analyzing a repository no attempt was made to gather information on the subject areas covered by the works contained therein. Instead information was gathered on the various types of work represented (e.g. published, theses/dissertations, teaching materials, etc.).

Secondly, IR contents reside behind a database, which means that they are not always available in the form of static web pages that can be analyzed as a whole representation of the website. They need to be retrieved using the interface that enables users to explore the contents of the IR, either through the use of keywords/key-phrases in the search function or by browsing the IR’s hierarchical structure. It is also the nature of IRs to contain digital objects numbering from hundreds to hundreds of thousands, and these numbers keep changing (increasing or
decreasing) as the IRs are investigated. In this circumstance it is not possible to analyze the whole contents of a repository. A content analysis can only be completed by taking samples of the contents (records), which can then be used to formulate indicative conclusion(s).

A further potential limitation is that data collection was undertaken by a sole coder, whereas Neuendorf (2002, p. 51) has recommended the use of “at least two coders, to establish intercoder reliability.” Thus the study might not fully satisfy the objectivity requirement.

Data Collection

Some existing tools were used to gather Internet addresses of Indonesian IRs, with the three most useful tools found to be the Ranking Web of Repositories, OpenDOAR, and ROAR. The list obtained from the Ranking Web of Repositories was the Indonesian repositories ranking of July 2014. The lists from OpenDOAR and ROAR were produced by filtering the entries for only Indonesia (country) and removing duplicate entries.

A number of IRs were excluded from the survey for various reasons: for example the IR was undergoing a trial phase; some institutions have more than one IR, with a subsidiary repository that is far less comprehensive than the primary one; or the IR was integrated into the library OPAC, making it extremely difficult to conduct an assessment of the IR alone. After these exclusions, 58 IRs remained to be analyzed. Out of these 58, six were found to be inaccessible after several attempts on different dates during the period of the content analysis (November 19th, 2014 to February 1st, 2015). Therefore 52 IRs remained in the study.

Content analysis was undertaken by visiting each IR and gathering both quantitative and qualitative data. A coding schedule in a form of a table (using spreadsheet application) was prepared to capture the quantitative data and is available as an external dataset (Liauw, 2015), while the coding variables are listed in Table 1.

Specific characteristics of the IRs were collected from the IR websites through visual inspection. These included the following:

- The name of the IR software in use, with the intention to see which software is/are dominant
- The presence of search and/or browse functionalities—IRs with only search or browse functionality might suggest lack of development.
- The provision of link(s) from the IR website to the library and/or the main institutional website. The absence of such links may imply an isolation of the IR.
- The provision of access statistics, indicating some level of service for the content contributors, who would be able to see the level of usage of their contributed works.
The Year of establishment was included as additional information that might be useful in assessing whether older IRs also meant more mature IRs. This information was obtained in most cases from OpenDOAR and/or ROAR. Status of the HE institutions (state or private) and their locations geographically in different regions in Indonesia were also assessed in recognition that there have been significantly different stages of development in each region, which might have some influences in the maturity of IRs. These data were gathered from Pangkalan Data Pendidikan Tinggi (http://forlap.dikti.go.id/perguruanantinggi).

The number of digital objects in each IR indicated the size of the IRs and was manually acquired by either browsing the index for “Year” of publication, then summing the number of digital objects from each individual year (EPrints), or browsing the collections available, then summing up the number of digital objects from individual collections (DSpace and others). Similar information was also gathered from OpenDOAR and ROAR—where available—as comparisons. The size of the IRs could also be used as one of the indicators for the maturity of the IRs.

The characteristics used to evaluate and categorize content (types of work) in this study require further explanation. Content in IRs can be broadly categorized into published and unpublished works (grey literature). However in order to better understand how Indonesian HE institutions develop and utilize their IRs, it is necessary to describe the unpublished works using narrower categories, as shown in Table 3. In this study the published works category includes conference articles/papers (proceedings), including proceedings that were published by an author’s own institution. Although conference proceedings may not by some measures be considered to be publications, for the purpose of this research they are considered to have been through a standardized process of selection and editing, including in some cases peer-review. In this study the journal-style articles resulting from undergraduate theses are categorized as Theses and Dissertations since there is no peer-review in the production process (see Table 3).

Information on the types of work was gathered by taking at least three sample records from each smallest unit of (digital object) collection in the repository, and inspecting the metadata and the full-text documents. It was only necessary to detect the existence of the different types of work in the IRs, and no attempt was made to calculate item (digital objects) counts for each type of work.

Author naming conventions were assessed since one important aspect of an IR is the author’s formal affiliation with the institution. It is in the best interests of institutions (and authors) to ensure that there is no ambiguity concerning the identity of authors represented in an IR.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Options</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acronym</td>
<td>N/A</td>
<td>Text</td>
</tr>
<tr>
<td>Institution or IR Name</td>
<td>N/A</td>
<td>Text</td>
</tr>
<tr>
<td>IR Year (of establishment)</td>
<td>N/A</td>
<td>Numeric</td>
</tr>
<tr>
<td>Status</td>
<td>State</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Private</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Region</td>
<td>Java</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Bali-Nusa Tenggara</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Sumatra</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Kalimantan</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Sulawesi</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Maluku</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Papua</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td># Digital Objects</td>
<td>Manual</td>
<td>Numeric</td>
</tr>
<tr>
<td></td>
<td>OpenDOAR</td>
<td>Numeric</td>
</tr>
<tr>
<td></td>
<td>ROAR</td>
<td>Numeric</td>
</tr>
<tr>
<td>IR Software</td>
<td>Dspace</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Eprints</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>GDL (Ganesha Digital Library)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Other/In-house</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Exploration Tools</td>
<td>B (Browse)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>S (Search)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Links</td>
<td>LI (Link to Institutional Website)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>LL (Link to Library Website)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>NL (No Link to Either)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Access Statistics</td>
<td>Y (Yes/Available)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>N (No/Unavailable)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Types of Works</td>
<td>PUB (Published)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>UNPUB (Unpublished)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>THESES (Theses/Dissertations)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>TEACH (Teaching Materials)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>STDW (Student Works)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>UREC (University Records)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>SPEC (Special Collections)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>OTHER</td>
<td>Numeric (1 or empty)</td>
</tr>
</tbody>
</table>

*Table 1.* Coding variables for content analysis of Indonesian HE IRs. Table continued on next page.
In the course of this survey the practice with regard to the naming of authors was assessed using several criteria as mentioned in the results section. Naming practices were solely assessed on how consistently institutions implemented rules or policies regarding names of author in the repository metadata. In repositories using EPrints, this could easily be assessed by browsing the contents based on author, which is one of the default options provided by the software. Criteria used to assess the implementation of author naming convention were the consistency of:

- The formatting of names: for example, placement of first and last names, and/or the use of space and other punctuation marks
- The use of upper and lower case
- The use of academic and other titles in names, and the way they are used or written.
- The use of student or staff ID number in names, and the way they are used or written.

<table>
<thead>
<tr>
<th>Characteristics Recorded</th>
<th>Options</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author Naming Convention</td>
<td>Y (Yes)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>N (No)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Standardized Access Points</td>
<td>Standardized Subject Headings</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Free-text Keywords</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Mix</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Not Available</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Language of Access Points</td>
<td>English</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Indonesian</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Mix</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Public Availability of Full-Text</td>
<td>All/Most (n &gt; 90%)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Some (25% &lt;= n &lt;= 90%)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>Minimal (0% &lt; n &lt; 25%)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>No Full-Text (0%)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Openness</td>
<td>OA (Open Access) - Public Avail-ability of Full Text &gt; 90%</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>NOA (Not Open Access) - Public Availability of Full Text &lt;= 90%</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Source/List Used</td>
<td>WEBO (Webometrics)</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>OpenDOAR</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td></td>
<td>ROAR</td>
<td>Numeric (1 or empty)</td>
</tr>
<tr>
<td>Date of Inspection</td>
<td>N/A</td>
<td>Date</td>
</tr>
</tbody>
</table>

**Table 1.** Coding variables for content analysis of Indonesian HE IRs. Continuation from previous page.
Consistent implementation of the convention will help avoiding variations of name for the same person/author.

In order to facilitate discoverability of their contents, IRs need to provide adequate access to their records. Access points are typically incorporated into the metadata of the works collected in the IRs. Therefore an assessment on these access points provides useful information on how discoverable individual works might be within the IRs. Firstly, the availability of subject heading(s) and/or keyword(s) was assessed, along with the use of standardized entries for subject headings. Similar to the case for author naming convention, assessment on subject heading(s) was made solely on the basis on how consistently institutions implemented rules in their entries for subject headings. Secondly, the language used in the access points was also assessed. The premise was that the presence of English as subject heading(s) or keyword(s), of which contents were mostly in Indonesian, would suggest that the IRs intended their works to be utilized by a broader (international) audience.

Lastly, the availability of full-text documents (of the complete works) in the IRs was assessed based on the criteria laid out in the coding variables table (see Table 1). The “Openness” variable was added for the sole purpose of making it easier for the researchers to count the IRs with (presumed) OA policies in place (full-text found in >90% of the sampled records) and IRs without OA policies (full-text found in <=90% of the sampled records).

Additional qualitative assessment of each repository was undertaken and recorded in order to provide additional general information about the characteristics of individual repositories, and therefore the state of Indonesian IRs in general. Each repository was scrutinized in order to note any local practices in the management and population of IRs. As much as possible general descriptions were made on each cluster of contents (collection) in the repository along with examples to serve as evidence. The outcomes of this part of the survey are only provided as a summary in the results section below and incorporated into the discussion section, without reference to particular institutions.

**RESULTS**

The full list of the 52 IRs analyzed in this study is not presented in this article due to space limitations. The list is available as an external dataset (Liauw, 2015). Subset or summary tables and graphics are used instead to highlight aspects of the IRs surveyed.

The 52 IRs contained 547,451 digital objects and are located in all regions of Indonesia except for Maluku (Mollucans) and Papua, regions in which there are currently no IRs.
registered in Webometrics, OpenDOAR, or ROAR. Most IRs (76.9%) are concentrated in Java, where 56.82% of the nation’s population live (Statistics Indonesia, 2015). The prevalent language of works in the IRs is Indonesian with a number of works in English. Figure 1 illustrates the distribution of IRs in Indonesia and the population of each region.

![Figure 1. Distribution of IRs (in red) in Indonesia and the population of region (in black or white)](image)

With regard to the number of digital objects, the numbers gathered through manual counting generally support the data provided by OpenDOAR with 9 IRs having exact matches. Data accessed from ROAR during this study appeared to be out of date and was not useful for comparisons.

Among the 52 IRs surveyed, 31 (59.6%) are associated with state universities, and 21 (40.4%) with private HE institutions. The majority of the IRs surveyed are using EPrints (n=34, 65.3%), followed by DSpace (n=9, 17.3%), Ganesha Digital Library/GDL (n=4, 7.8%), and other or in-house-developed software (n=5, 9.6%). Figure 2 indicates the distribution of the number of IR based on IR software, while Figure 3 indicates the distribution of digital objects based on IR software (see Figures 2 & 3).

In terms of public availability of the full-text documents, based on the sample records retrieved during this study, 14 IRs (26.9%) provide all or most documents in full-text; 16 (30.7%) provide some or a representative number; 17 (32.7%) provide a small number only; and 5 (9.6%) do not provide any full-text documents (metadata only). The majority of the IRs (n=39; 75%) have not implemented any author naming conventions, and only 13 IRs (25%) have done so.
Figure 2. Distribution of Indonesian HE IRs based on IR software

Figure 3. Distribution of digital objects in Indonesian HE IRs based on IR Software

*The # of digital objects for University of Indonesia’s IR was taken from OpenDOAR data.
Most IRs (n=47, 90.3%) provide access points in the form of either standardized subject headings, free-text keywords (n=16, 30.7%), or both (n=15, 28.9%). Realizing that most (if not all) IRs will find it difficult to link their IR software to another database containing standardized subject headings, this study only assessed whether the subject headings used in an IR adhere to certain rules or guidelines, or simply allow free-text keywords to be typed in without reference to any guidelines. Some IRs might have separate field/metadata for subject headings and keywords, while some might place them both in the same field/metadata. Interestingly, there are 5 IRs (9.6%) that provide neither subject headings nor keywords. In terms of the language used for these access points, the majority (43 IRs, 82.6%) use English subject headings and/or keywords. Of these 43, 19 (36.5%) use English only, while 24 (46.1%) use English and Indonesian subject headings and/or keywords. Figure 4 indicates some characteristics of the IRs discussed above.

![Public Availability of Full-Text Document](image1.png)

![Use of Authors Naming Convention](image2.png)

**Figure 4.** Snapshots of some aspects of Indonesian HE IRs

In terms of exploration tools, almost all IRs (except one) have search and browse functionalities. One has a search function only. 28 IRs (53.84%) provide a link to neither the institutional website nor the library website, 8 (15.38%) provide a link to the institutional website only, 4 (7.7%) provide a link to the library website only, and 12 (23.08%) provide links to both. Most IRs (45 or 86.54%) do not provide access or usage statistics, with only 7 IRs (13.46%) providing these statistics.

The most widely included type of work are Theses and Dissertations (n=44, 84.6%), followed by Published Works (n=42, 80.8%). These are followed in turn by Unpublished Works (n=27, 51.9%) and University Records (n=26, 50%). Teaching Materials are only included in 20 (30.8%), and Student Works in 8 (15.4%). Special Collections comprise the least commonly encountered type of work of those specified, being found in one IR only (1.9%). Other types of work (those not fitting into the specified categories) were found in 10 (19.2%) of the IRs. Figure 5 reports the distribution of the types of work in Indonesian HE IRs (See Figure 5).
In terms of qualitative observations, it was found that a number of IRs provide only metadata and abstracts of theses, journal articles, etc. Other practices limiting access to the full copy of documents were observed. Some IRs:

- Upload only certain parts of the works
- Upload all parts of (whole) works (in separate files) but limit public access to some or significant parts of the works
- Impose an embargo period, often without a specified end-date (in some cases documents with an apparently expired embargo period still cannot be accessed)

Further, a number of apparently local and non-standard practices in the management and population of IRs were encountered, including:

- Compressed or password-protected files/documents
- Student documents that still included supervisor’s comments
- Textual documents stored as multiple image files
- Works that were separated into different records based on file format
- Broken links
- Confusing use of categories, such as categorizing promotional materials (institutional profiles) as books merely because the materials are in the form of book(let)s.

Figure 5. Distribution of the Types of Work in Indonesian HE IRs.
DISCUSSION

Based on the data and qualitative assessments described above, there are some noteworthy aspects of the results. Firstly, the distribution of IRs in Indonesia is consistent with the distribution of HE institutions, which generally reflects the distribution of the country’s population. The distribution of digital objects also corresponds to the population distribution in each regional area, as shown in Table 2.

<table>
<thead>
<tr>
<th>Region</th>
<th>% Population*</th>
<th>% IRs</th>
<th>% Digital Objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Java</td>
<td>56.82</td>
<td>76.9</td>
<td>78.06</td>
</tr>
<tr>
<td>Sumatera</td>
<td>21.64</td>
<td>15.4</td>
<td>17.56</td>
</tr>
<tr>
<td>Sulawesi</td>
<td>7.33</td>
<td>1.9</td>
<td>2.13</td>
</tr>
<tr>
<td>Kalimantan</td>
<td>6</td>
<td>3.9</td>
<td>1.92</td>
</tr>
<tr>
<td>Bali &amp; Nusa Tenggara</td>
<td>5.52</td>
<td>1.9</td>
<td>0.33</td>
</tr>
<tr>
<td>Maluku</td>
<td>1.12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Papua</td>
<td>1.57</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

*(Statistics Indonesia, 2015)

Table 2. Comparisons of population of people, IR, and digital objects in regions of Indonesia

Secondly, the exact matches in terms of the number of digital objects between the OpenDOAR data and manual counting for 9 IRs (7 use EPrints, 2 use DSpace) suggests that the method used in this study to manually count the number of digital objects in IRs is reliable. The differences that do occur between the OpenDOAR data and manual counting can be explained by recent updates by the respective IRs. It is the case, however, that the manual counting usually (56.25%) provided a smaller number of digital objects than OpenDOAR. This strongly suggests that some institutions have deleted records from their IRs since OpenDOAR last updated its records or at least removed digital objects from the records. One IR (Universitas Gunadarma) has recorded an 84.38% decrease in its number of digital objects, suggesting a change of software or a recommencement of the IR. It can be speculated that this implies that some Indonesian HE institutions are experimenting with the compositions of their IRs.

Thirdly, the finding that the majority of the IRs surveyed provide access points to facilitate discoverability (using at least keywords), and that the language used for these access points is mostly English, suggests that Indonesian HE institutions intend to make their IR’s content available to both domestic and international users. As many developing countries lack access to formal channels of scholarly communication, particularly in their native languages, IRs (and OA more generally) are seen as an important means of increasing the profile and
<table>
<thead>
<tr>
<th>Type</th>
<th>Criteria</th>
</tr>
</thead>
</table>
| **PUB**    | 1. Peer-reviewed journal articles (including articles published by the author's institution)  
2. Articles in conference proceedings (including papers published by the author's institution)  
3. Books or book sections (including books or sections published by the author's institution)                                                                                           |
| **UNPUB**  | 1. Unpublished/internal research reports (submitted internally for reports or for tenure requirements)  
2. Unpublished institutional research reports  
3. Audio/video recordings or text of professoriate inaugural speeches/orations                                                                                                                          |
| **THESIS** | 1. Non-degree (diploma) theses  
2. Undergraduate theses  
3. Master theses  
4. Ph.D. theses (Dissertations)  
5. Includes also journal-style articles derived from students theses and/or presentation slides of the thesis defense (if available)                                                                                       |
| **TEACH**  | 1. Classroom presentation slides  
2. Course/class-related (includes laboratory-based) materials                                                                                                                                                                                                                             |
| **STDW**   | 1. Works by students outside the scope covered by THESES and UREC  
2. Course/class-based works/assignments  
3. Includes visual resources (photographs, drawings, computer renderings, animations, etc.) produced by students as part of assignments  
4. Works produced by students as part of their participation in non-course-based activities such as competitions, internship and/or volunteer work, etc. |
| **UREC**   | The emphasis of this category is on the evidential value of the documents, instead of the informational value (topic/subject) of the works.  
1. Internal campus newspapers, magazines, bulletins, or newsletters, including student-published journals or media  
2. Promotional materials (flyers/posters/brochures) of campus notifications, events, facilities, services, programs/activities, or issue-based campaign  
3. Photography/audio/video recordings of campus notifications, events, facilities, services, programs/activities, or issue-based campaigns  
4. Newspaper clippings (includes screen captures in JPG/PDF of online news)  
5. Audio/video/text of general speeches (except professoriate inaugural speeches/orations, which are categorized as UNPUB)  
6. Institutional or departmental constitutions, by-laws, reports, guidelines, presentation slides, and other formal documents  
7. Documents (presentations, handouts, flyers, etc.) used in campus events/activities. |
| **SPEC**   | 1. Digitized or born-digital resources related to special/rare collections and/or local collections                                                                                                                                                                                                                                    |
| **OTHER**  | 1. Documents that are not related to the institution or its intellectual output or its unique collections (SPEC) but might be used or stored for reference purposes. Example: government documents, ministerial decrees, scanned books or eBooks; of which authors are not from the institution, etc.  
2. Documents that are not the main works but supplemental to the main works and cannot be categorized as UREC since the emphasis is on the informational value instead of evidential value. Example: conference presentations (of a research/paper/article). The presentation slides can be stored in the same record as the main works (articles) in the repositories. |

Table 3. Characteristics used to evaluate and categorize contents in the content analysis of Indonesian HE IRs
availability of their research outputs. It is a reality, however, of the global domination of English as the international language of scholarship that this will only be achieved if these outputs are widely discoverable, and that requires the use of English metadata.

Fourthly, only 14 (26.9%) of the IRs surveyed made all or most of the full-text documents freely-accessible to the public. In the original context of the OA Movement, where IRs were initially developed and implemented as a Green (self-archiving) strategy in response to the business model of commercial publishers, this result is far lower than might be expected. In this context, the very low degree of openness of Indonesian IRs is discouraging, unless we use a different perspective in understanding this trend. A consideration of the types of works available in these IRs might be helpful in this regard.

Regarding types of works, there are some notable characteristics of the IRs surveyed:
• The most common type of work is Theses & Dissertations (84.6% of IRs)
• Published Works (80.8%) are relatively common
• University Records are included in a relatively high proportion of IRs (50%), rivalling Unpublished Works (51.9%)
• Student Works (15.4%), Teaching Materials, Special Collections (1.9%), and other types of works are also present, but only minimally

The finding that Theses and Dissertations are the most common inclusion might suggest that they were the first type of work populating the IRs. This suggestion is in line with Lippincott’s (2006, p. 3) assertion that “an ETD [(Electronic Theses and Dissertations)] program has frequently served as the foundation or pilot for institutional repository content,” on the basis that they constitute the “low-hanging fruit.”

In the context of the origins of the OA Movement one would expect Published Works to be the dominant category included in IRs, but this is evidently not the case for Indonesian HE IRs. Even though the category of Published Works used in this study has been broadened to include conference articles/papers (proceedings), they nonetheless remain secondary to Theses and Dissertations. In observing the characteristics of the content of the IRs surveyed in terms of the visibility of different types of work and the very low degree of openness, it is highly likely that many Indonesian HE IRs were not conceived as a response to a scholarly communication problem, but rather as a corporate information management system, especially to manage students’ theses and dissertations. It can be concluded that the issue of OA is of secondary importance to many Indonesian HE IRs, and that the use of IRs as a management tool is a likely explanation for the lack of links connecting them to institutional websites and/or library websites.
The popularity of IRs in the Indonesian HE sector was significantly enhanced in 2006 when the Consejo Superior de Investigaciones Científicas released the Webometrics Ranking of World Universities (Spanish National, n.d.b.), which then be used by DIKTI as a tool to assess Indonesian HE institutions. Out of 52 IRs analyzed in this study, 3 were started in 2006, 41 were started after 2006, while the year of inception of 8 IRs cannot be determined. In other words, the study could not find any IR that was started before 2006. The majority of the IRs analyzed (36 IRs) were started in or after 2010. Also, one of the “quantitative web indicators” used by Webometrics is the availability of Rich Files, which is defined as “[f]iles in formats like Adobe Acrobat (.pdf), MS Word (.doc, .docx), MS Powerpoint (.ppt, .pptx) and PostScript (.ps & .eps) extracted from Google” (Spanish National, n.d. a). The Webometrics ranking therefore suggests that these 40 IRs (in the ranking) have—in one way or another—allowed their contents (full-text documents) to be crawled and indexed by Google (Aguillo, Ortega, Fernández & Utrilla, 2010), creating the appearance of open access. On the other hand, this study found a contradictory indicator in that only 26.9% (14 IRs) allow all or most of their contents (full-text documents) to be accessed publicly using their IR’s user interface. It is also relevant to note that the mandate to “upload electronically all scholarly works by students/lecturers/researchers/staff of any higher education institution” (Direktorat Jenderal, 2010, Chapter 7 Clause 2) was situated in a broader context of fighting plagiarism. This is apparent from the title of the relevant act; “The Prevention and Eradication of Plagiarism in Higher Education Institutions” (Direktorat Jenderal, 2010). Additionally, in his presentation titled “Open Access Repositories in Indonesia,” the initiator of Indonesia OneSearch and a consultant to the National Library of the Republic of Indonesia, laid out the roadmap of Indonesia One Search, noting that in addition to serving as a scholarly information portal, in its final phase (2018-2020) it will also serve as plagiarism checking services (Fahmi, 2016, p. 49). These indicators result in speculation by the authors that the push for OA in Indonesian HE IRs was initially driven more by prestige (ranking) and attempts to combat plagiarism, rather than the desire to make Indonesian research globally visible.

Finally, there are an apparent lack of institutional policies and quality control in most of the IRs surveyed. This has resulted in several conditions observed during the survey, such as:

- Different collection development policies and access policies used by academic departments in the same institution;
- Lack of standardized author naming conventions;
- Lack of awareness regarding privacy, confidentiality and copyright issues; and
- Sub-standard and occasionally chaotic contents.
CONCLUSION

The study found that the majority of IRs analyzed in this study were started in or after 2010, which indicates that IRs in Indonesian HE institutions are still in the early adoption phase. The practice of record or digital object deletion in a significant number of IRs analyzed (56.25%) suggests that the Indonesian HE institutions are in the stage of experimenting with their IRs. The study also found evidence of efforts by Indonesian HE that imply they intend to make their IR’s content available to a domestic as well as international audience. The dominant type of work is Theses and Dissertations, followed by Published Works. University Records and Unpublished Works are also relatively common, being found in about half of the IRs surveyed. A further notable finding is the quite small number of the IRs surveyed (26.9%) that make all or most of their contents accessible as full-text documents to the public.

The various characteristics identified as being common to Indonesian HE IRs surveyed leads the authors to speculate that some Indonesian HE IRs were implemented initially as corporate information management systems, rather than as contributors to a genuinely Green OA strategy. It is also very likely that institutional prestige in terms of webometrics’ ranking and the need to combat plagiarism have determined the growth and characteristics of OA in Indonesia more than the need to make Indonesian research visible and accessible. More detailed studies involving the investigation of individual IRs will be needed to establish whether those speculative conclusions are warranted.

Associated issues in Indonesian HE IRs identified in this study are the lack of adequate institutional policy frameworks and quality control, which have produced a substandard means of scholarly communication. These issues need to be addressed by the Indonesian HE community and DIKTI in order to establish IRs as a viable channel of scholarly communication that can help alleviate the various information divides that are associated with paywalls, language, and Indonesia’s economic circumstances as a developing country.

Despite these problems there are, however, reasons to be optimistic about the future of Indonesian HE IRs. The practice of using IRs to co-locate and distribute institutionally-affiliated publications is now well established and the first generation of development has seen them incorporated into the scholarly communication practice of many universities. The government incentives in this regard have been influential, but Indonesian IRs are now sufficiently widespread and integral to the routine management of HE research and teaching to suggest that they have a self-sustaining future. The problems that still need to be addressed should become the focus of second generation activity, and these include the various issues identified by this research that are necessary in order to regularize the structure, content and accessibility of IRs.
As Indonesia continues to focus considerable resources on improving its higher education and research sectors, an important next step will be to integrate its scholarly publishing within existing international standards and frameworks. This is likely to mean enhancing IRs in terms of their openness, focusing on the quality of material included, and lifting the metadata standards. These goals present challenges for developing countries, with issues including the affordability of infrastructure, the transgenerational timeframe required to build research culture and capacity, and the dominant role of English as the language of international scholarship. IRs nonetheless provide a singularly important means by which developing countries, including Indonesia, can support scholars and educators in making their research outputs as open, accessible and international as possible.

REFERENCES


Liauw, T. T. (2015). *Content analysis of Indonesian higher education institutional repositories (Nov 19, 2014 - Feb 01, 2015)*. [Dataset]. [https://doi.org/10.4225/06/565BACB4C3CBA](https://doi.org/10.4225/06/565BACB4C3CBA)


Suber, P. (2016). Someday, Altmetrics will no longer need ‘Alt’. Retrieved from https://plus.google.com/+PeterSuber/posts/hVYV3xVX3M1


