



PERPUSTAKAAN NASIONAL
REPUBLIK INDONESIA

Lokakarya Perpustakaan Nasional
31 Agustus 2016

Open Archive Initiatives (OAI)

Ismail Fahmi, PhD.

Inisiator

Indonesia OneSearch (IOS)

Ismail.fahmi@gmail.com





Ismail Fahmi, PhD.

Ismail.fahmi@gmail.com

- 1992 – 2007 S1, Teknik Elektro, ITB
- 2003 – 2004 S2, Computational Linguistics, Universitas Groningen, Belanda
- 2004 – 2009 S3, Computational Linguistics, Universitas Groningen, Belanda

- 2000 – 2003 Inisiator IndonesiaDLN (Digital Library Network pertama di Indonesia)
Mengembangkan Ganesha Digital Library (GDL)
Mendirikan Knowledge Management Research Group (KMRG) ITB
Membangun Digital Library ITB

- 2009 – Sekarang Engineer di **Weborama**, Perusahaan berbasis big data (Paris/Amsterdam)
- 2012 – Sekarang Co-Founder **Awesometrics**, Media Monitoring & Analytics Company
- 2014 – Sekarang Founder **PT. Media Kernels Indonesia**, a Natural Language Processing Company
- 2015 – Sekarang Konsultan **Perpustakaan Nasional**, Inisiator Indonesia OneSearch

Agenda

- Open Access
- Open Access Indonesia – Discovery System
- OAI PMH
- Standard Metadata
- Otomasi Pengolahan



A close-up portrait of a young man with short brown hair, smiling broadly. He is wearing a brown sweater over a white collared shirt. The background is a light blue wall with faint white geometric patterns.

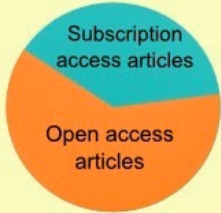
Jack Andraka

15 tahun, USA

HOW ONE
TEEN
INNOVATOR
IS CHANGING
THE WORLD

OPEN ACCESS: BENEFITS

Unique visitors

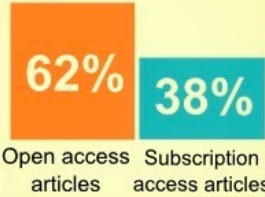


**Increased
Visibility**



**Increased
Citation Rates**

Citation Rates



"Jack Andraka, a 15 year old from USA, used free articles from Google Scholar to develop a highly innovative method of detecting a rare type of pancreatic cancer."



**Drives
Innovation**



**Global
Impact**

"Piya Sorcar, founder of Teach Aids made materials & research about HIV available for free globally. The material is now available in 15 languages and has reached over 70 countries, including numerous educators, governments and NGOs."

Open Access brings new audiences to scholarly content, such as educators, business owners, support workers and so on by removing barriers of affordability and access.



**Public
Access**



**Comply with
Funder Policies**

Funders, such as the Research Councils, European Commission and the Wellcome Trust, mandate for open access publication of research outputs.

Open Access Around The World

OpenDOAR

DOAJ DIRECTORY OF OPEN ACCESS JOURNALS

OpenRepository

Open Access
SOAJ Scientific Open Access Journals

SPARC*

ELSEVIER

OPEN ACCESS | MALAYSIA

Open access

OAK OPEN ACCESS KOREA

PLOS

OSJI OPEN SCIENCE JOURNAL

OPEN ACCESS

Karger Open access

open access india

doab directory of open access books



**“To have a robust national
open access repository
discovery system”**

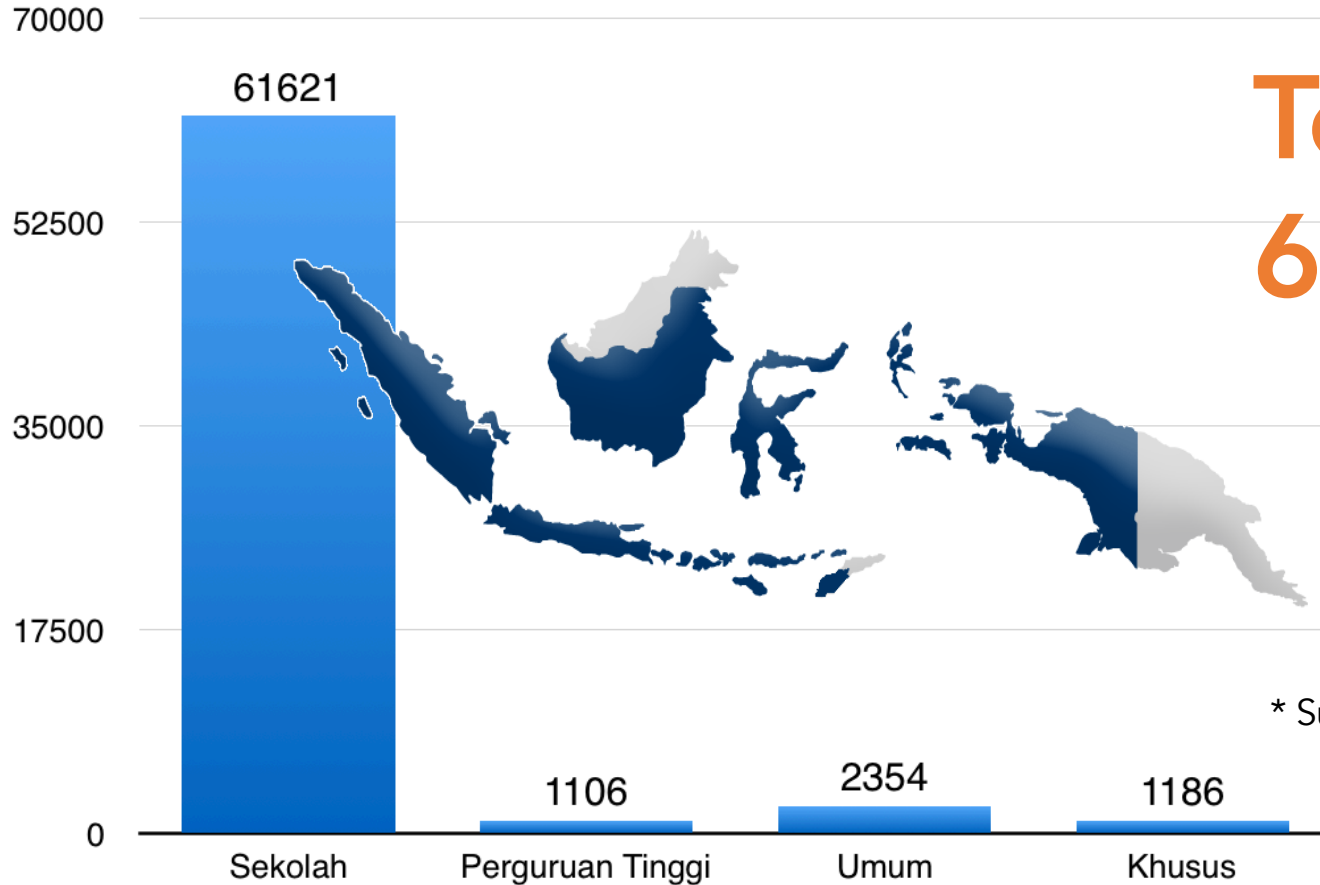
OAI Discovery System



Merupakan sebuah








Sumber Data: Perpustakaan



**Total:
66.000+**

* Sumber: Perpunas 2016

Jenis Data

	Katalog Buku	E-Journal	Digital Repository	Museum
	●			
		Discontinued		
		●		
		●		
	●	●	●	●

Standard Interoperability Protocol



~75%

Repository di
seluruh dunia
'OAI-Compliant'

OAI-PMH

a low-barrier interoperability framework

Skenario Harvesting



Any Platforms – Any Collections



Standard OAI-PMH



Harvester

Request Verbs:

- Identify
- ListMetadataformats
- ListSets
- ListIdentifiers
- ListRecords
- GetRecord



Provider
(Repository)

4.5 ListRecords

Summary and Usage Notes

This verb is used to harvest records from a repository. Optional arguments permit [selective harvest](#) repository's support for [deletions](#), a returned [header](#) may have a `status` attribute of "deleted" if a will be present for records with deleted status.

Arguments

- `from` an optional argument with a [UTCdatetime value](#), which specifies a lower bound for d
- `until` an optional argument with a [UTCdatetime value](#), which specifies an upper bound for
- `set` an optional argument with a [setSpec value](#), which specifies [set](#) criteria for [selective h](#)
- `resumptionToken` an [exclusive](#) argument with a value that is the [flow control](#) token returne
- `metadataPrefix` a [required](#) argument (unless the [exclusive](#) argument `resumptionToken` is [metadata part of the returned records](#). Records should be included only for items from which matching the [metadataPrefix](#) can be disseminated. The metadata formats supported by a request.

Error and Exception Conditions

- `badArgument` - The request includes illegal arguments or is missing required arguments.

Semua Harus Mengikuti Standard

<https://www.openarchives.org/OAI/openarchivesprotocol.html>

<http://an.oa.org/OAI-script?>

verb=ListRecords&

from=1998-01-15&set=physics:hep&

metadataPrefix=oai_rfc1807



<https://jurnal.uns.ac.id/index.php?>

journal=alchemy&

page=oai&

verb=ListRecords&

metadataPrefix=oai_dc



badArguments:
journal,page

Metadata

Paling banyak diimplementasikan
(berdasarkan data Indonesia OneSearch):

- MARC → marcxml
- DC → oai_dc

<http://an.oa.org/OAI-script?>

verb=ListRecords&metadataPrefix=oai_dc

<http://an.oa.org/OAI-script?>

verb=ListRecords&metadataPrefix=marcxml

Metadata Harus Standard

Semua platform software yang berbasis open source seperti **OJS, Eprints, Dspace, dan Koha**, sudah mendukung dan comply dengan standard OAI-PMH dan metadata.

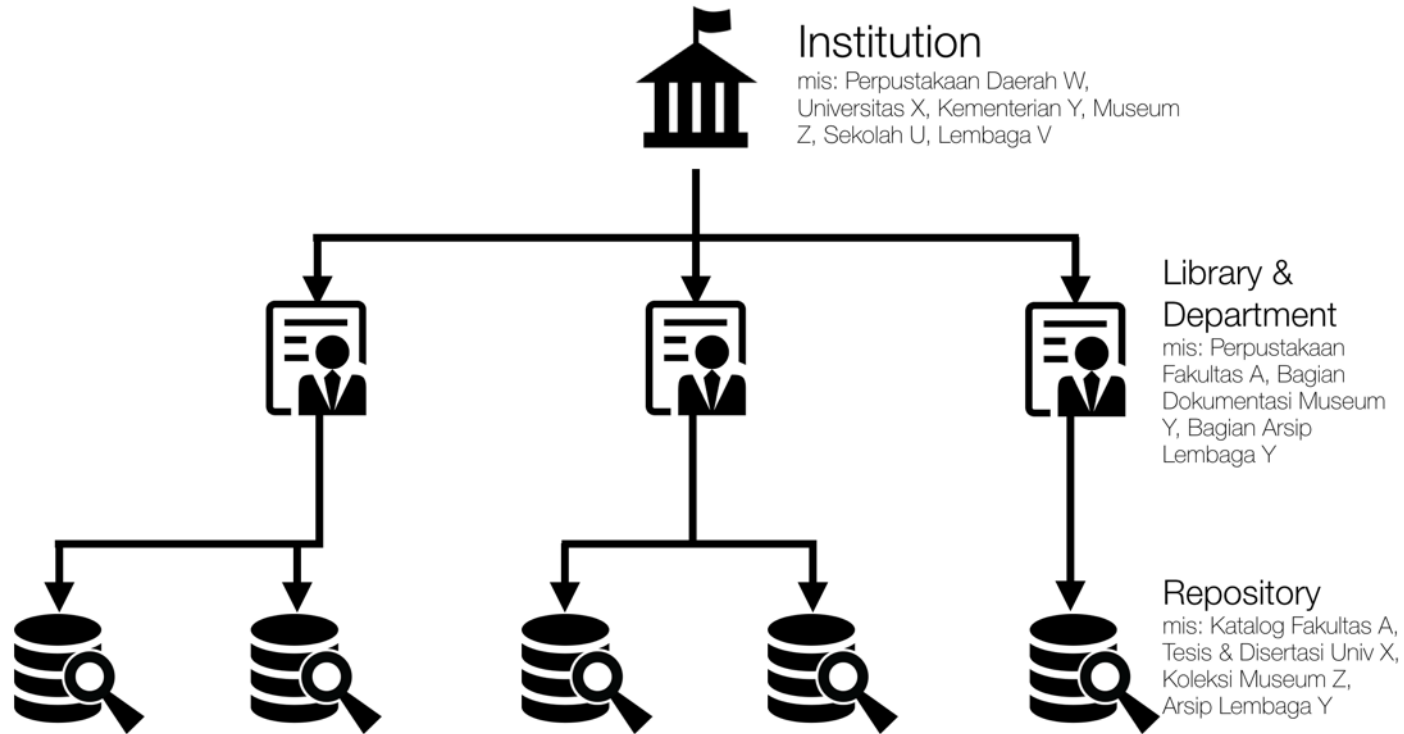
Software lokal cukup banyak yang tidak mendukung OAI-PMH, atau yang tidak comply.

SLIMS paling banyak digunakan di Indonesia, dan versi lama belum comply. Versi OAI-PMH untuk SLIMs dapat didownload di:

<http://wiki.onesearch.id/doku.php?id=oai-slims>



Struktur Informasi di IOS



Registrasi OAI di IOS

SYSTEM INFORMATION

Software Platform *	<input type="text" value="EPrints"/>
OAI Base URL *	<input type="text" value="e.g. http://journal.example.com/oai.php"/>
Metadata Prefix *	<input type="text" value="oai_dc"/>
Set (filter)	<input type="text" value="Set filter"/>
New Records Added *	<input type="text" value="Daily"/>

Base URL OAI

Benar:

<http://domainname.ac.id/oai.php>

Salah:

<http://domainname.ac.id/oai.php?verb=ListRecords&metadataPrefix=marcxml>

OAI dalam INLIS Lite

<http://203.176.180.116:12345/opac/oaipmh/oi.aspx>



ONLINE PUBLIC ACCESS CATALOG
Perpustakaan Umum Kota Depok

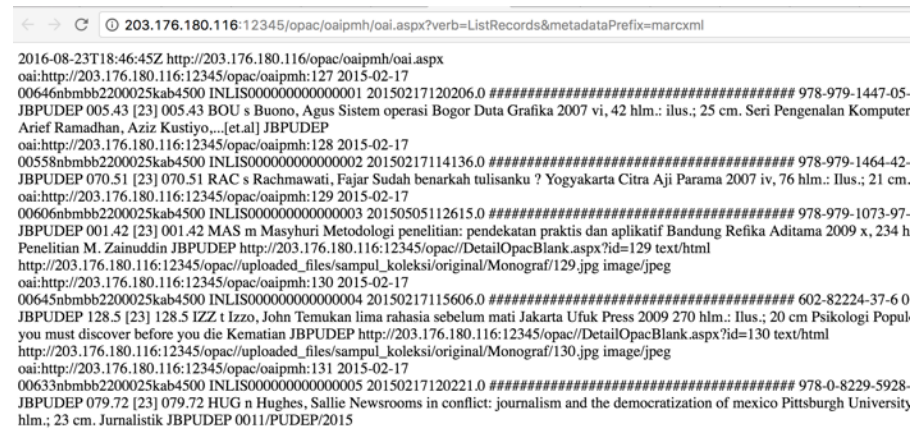
Pencarian Sederhana | Pencarian Lanjut | Pencarian Berbasis MARC | 10 Koleksi Sering Dipinjam | Tampung

Pencarian Sederhana

Kata Kunci : Judul Semua Jenis Bahan Cari

Panduan :

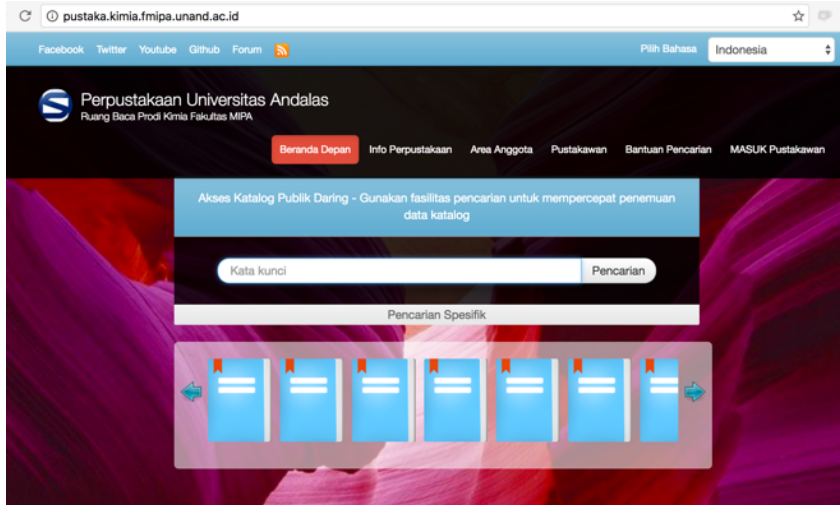
- Pencarian sederhana adalah pencarian koleksi dengan menggunakan hanya satu kriteria pencarian saja.
- Ketikkan kata kunci pencarian, misalnya : " Sosial kemasyarakatan "
- Pilih ruas yang dicari, misalnya : " Judul "
- Pilih jenis koleksi misalnya " Monograf(buku) ", atau biarkan pada pilihan " Semua Jenis Bahan "
- Klik tombol "Cari" atau tekan tombol Enter pada keyboard



2016-08-23T18:46:45Z <http://203.176.180.116/opac/oaipmh/oi.aspx>
oai:http://203.176.180.116:12345/opac/oaipmh:127 2015-02-17
00646nbmbb2200025kab4500 INLIS0000000000000001 20150217120206.0 ##### 978-979-1447-05-
JBPUDEP 005.43 [23] 005.43 BOU s Buono, Agus Sistem operasi Bogor Duta Grafika 2007 vi, 42 hlm.: ilus.; 25 cm. Seri Pengenalan Komputer
Arief Ramadhan, Aziz Kustiyo,...[et.al] JBPUDEP
oai:http://203.176.180.116:12345/opac/oaipmh:128 2015-02-17
00558nbmbb2200025kab4500 INLIS0000000000000002 20150217114136.0 ##### 978-979-1464-42-
JBPUDEP 070.51 [23] 070.51 RAC s Rachmawati, Fajar Sudah benarkah tulisanku ? Yogyakarta Citra Aji Parama 2007 iv, 76 hlm.: ilus.; 21 cm.
oai:http://203.176.180.116:12345/opac/oaipmh:129 2015-02-17
00606nbmbb2200025kab4500 INLIS0000000000000003 20150505112615.0 ##### 978-979-1073-97-
JBPUDEP 001.42 [23] 001.42 MAS m Masyhuri Metodologi penelitian: pendekatan praktis dan aplikatif Bandung Refika Aditama 2009 x, 234 h
Penelitian M. Zainuddin JBPUDEP <http://203.176.180.116:12345/opac/DetailOpacBlank.aspx?id=129 text/html>
http://203.176.180.116:12345/opac/uploaded_files/sampul_koleksi/original/Monograf/129.jpg image/jpeg
oai:http://203.176.180.116:12345/opac/oaipmh:130 2015-02-17
00645nbmbb2200025kab4500 INLIS0000000000000004 20150217115606.0 ##### 602-82224-37-6 0
JBPUDEP 128.5 [23] 128.5 IZZ t Izzo, John Temukan lima rahasia sebelum mati Jakarta Ufuk Press 2009 270 hlm.: ilus.; 20 cm Psikologi Popul
you must discover before you die Kematian JBPUDEP <http://203.176.180.116:12345/opac/DetailOpacBlank.aspx?id=130 text/html>
http://203.176.180.116:12345/opac/uploaded_files/sampul_koleksi/original/Monograf/130.jpg image/jpeg
oai:http://203.176.180.116:12345/opac/oaipmh:131 2015-02-17
00633nbmbb2200025kab4500 INLIS0000000000000005 20150217120221.0 ##### 978-0-8229-5928-
JBPUDEP 079.72 [23] 079.72 HUG n Hughes, Sallie Newsrooms in conflict: journalism and the democratization of mexico Pittsburgh University
hlm.; 23 cm. Jurnalistik JBPUDEP 0011/PUDEP/2015

OAI dalam SLIMs

<http://pustaka.kimia.fmipa.unand.ac.id/oai2.php>

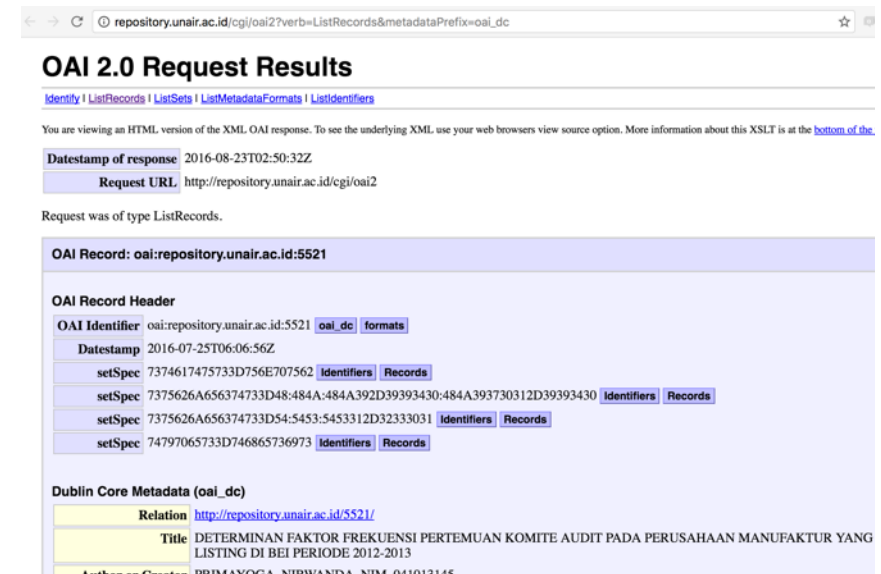


OAI dalam Eprints

<http://repository.unair.ac.id/cgi/oai2>



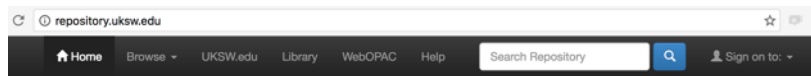
The screenshot shows the homepage of the Airlangga University Repository. At the top left is the university's logo. The main header reads "Airlangga University Repository". Below this is a navigation menu with "Home", "About", and "Browse". There are also links for "Login" and "Create Account". A green banner welcomes users to the repository, stating it is a digital collection of academic papers. Below the banner are several sections: "Latest Additions" with a link to view items added in the past week; "Search Repository" with a note to use the search field at the top; "Browse Repository" with a note to browse items by subject; and "About this Repository" with a note to get more information about the site. At the bottom right of the banner area, there are icons for Atom, RSS 1.0, and RSS 2.0.



The screenshot shows the "OAI 2.0 Request Results" page. The title is "OAI 2.0 Request Results". Below the title are links for "Identify", "ListRecords", "ListSets", "ListMetadataFormats", and "ListIdentifiers". A message states: "You are viewing an HTML version of the XML OAI response. To see the underlying XML use your web browsers view source option. More information about this XSLT is at the [bottom of the](#)". There are two fields: "Datestamp of response" with the value "2016-08-23T02:50:32Z" and "Request URL" with the value "http://repository.unair.ac.id/cgi/oai2". Below this, it says "Request was of type ListRecords." A section titled "OAI Record: oai:repository.unair.ac.id:5521" contains an "OAI Record Header" with the following details: "OAI Identifier" is "oai:repository.unair.ac.id:5521" with links for "oai_dc" and "formats"; "Datestamp" is "2016-07-25T06:06:56Z"; "setSpec" is "7374617475733D756E707562" with links for "Identifiers" and "Records"; "setSpec" is "7375626A656374733D48:484A:484A392D39393430:484A393730312D39393430" with links for "Identifiers" and "Records"; and "setSpec" is "7375626A656374733D54:5453:5453312D32333031" with links for "Identifiers" and "Records". Below the header is the "Dublin Core Metadata (oai_dc)" section with a table: "Relation" is "http://repository.unair.ac.id/5521/"; "Title" is "DETERMINAN FAKTOR FREKUENSI PERTEMUAN KOMITE AUDIT PADA PERUSAHAAN MANUFATUR YANG LISTING DI BEI PERIODE 2012-2013"; and "Author or Creator" is "PRIMAYOGA NIRWANDA NIM 041013125".

OAI dalam DSpace

<http://repository.uksw.edu/oai/request>



Communities in DSpace

Choose a community to browse its collections.

Final Year Projects	5657
Research Reports	213
UKSW-ana	1893

Discover

Author

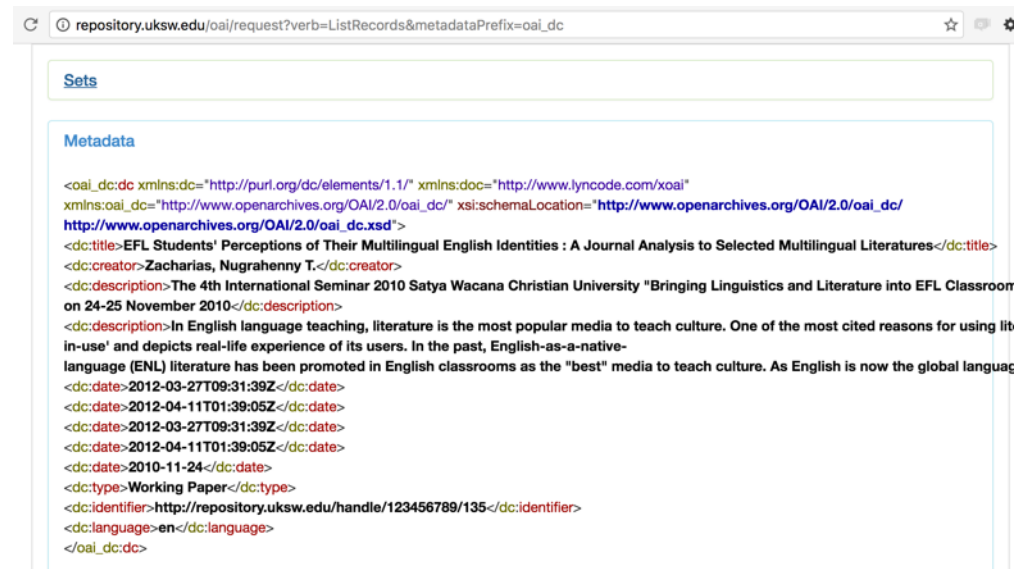
Parhusip, Hanna Arini	28
Setiawan, Andreas	28
Setiawan, Adi	26
Samiyono, David	24
Rondonuwu, Ferdy Samuel	23
Simanjuntak, Bistok Hasiholan	23
Martono, Yohanes	22
Setyawan, Iwan	22
Supatni	21

Subject

hasil belajar	777
hasil belajar IPA	134
hasil belajar matematika	134
matematika	60
IPA	74
motivasi belajar	56
prestasi belajar	42
Android	41
hasil belajar IPS	35
keaktifan	32

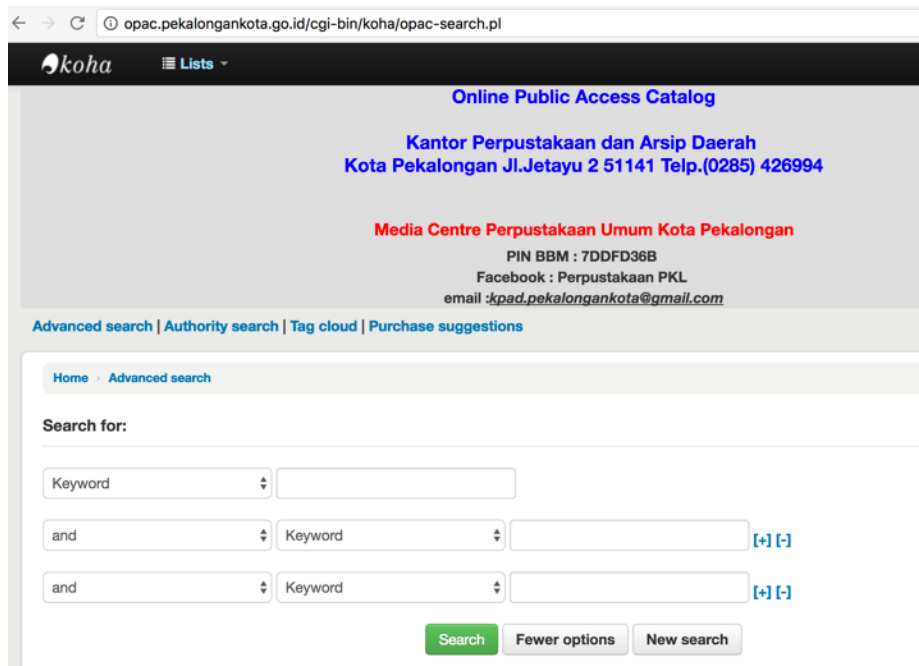
Date issued

2010 - 2016	7220
2000 - 2009	334
1990 - 1999	4
1970 - 1979	1



OAI dalam Koha

<http://opac.pekalongankota.go.id/cgi-bin/koha/oai.pl>



The screenshot shows the Koha Online Public Access Catalog homepage. At the top, there is a navigation bar with the Koha logo and a 'Lists' menu. Below this, the text 'Online Public Access Catalog' is displayed. The main content area features the name of the library: 'Kantor Perpustakaan dan Arsip Daerah Kota Pekalongan Jl. Jetayu 2 51141 Telp. (0285) 426994'. Below this, it identifies itself as the 'Media Centre Perpustakaan Umum Kota Pekalongan' and provides contact information: 'PIN BBM : 7DDFD36B', 'Facebook : Perpustakaan PKL', and 'email : kpad.pekalongankota@gmail.com'. At the bottom of the main content area, there are links for 'Advanced search', 'Authority search', 'Tag cloud', and 'Purchase suggestions'. Below the main content area, there is a search section with a 'Search for:' label and a search box. The search box contains the text 'Keyword' and a dropdown menu. Below the search box, there are two 'and' buttons and two more search boxes, each with a dropdown menu and a '+' '-' button. At the bottom of the search section, there are three buttons: 'Search', 'Fewer options', and 'New search'.



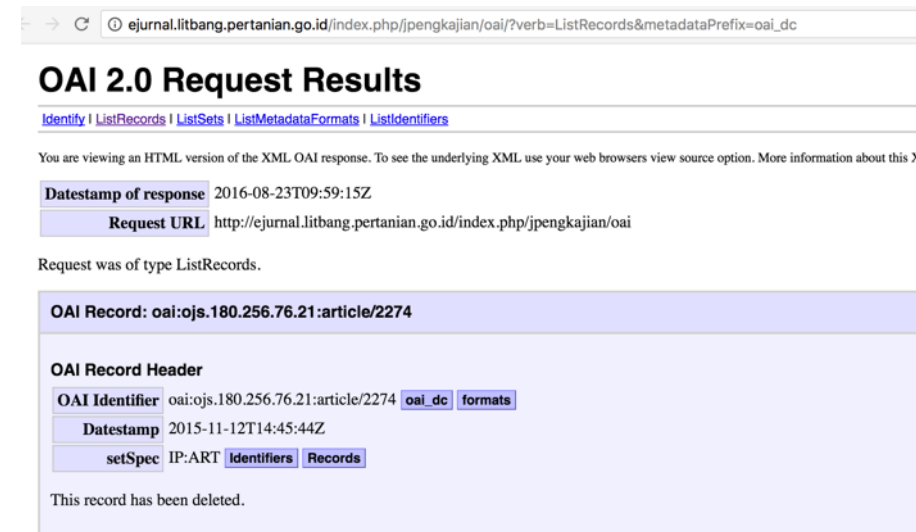
The screenshot shows an XML file view in a browser. The address bar displays the URL: <http://opac.pekalongankota.go.id/cgi-bin/koha/oai.pl?verb=ListRecords&metadataPrefix=marcxml>. The main content area shows the XML document tree. The root element is <OAI-PMH>, which contains a <request> element with the verb 'ListRecords' and the metadataPrefix 'marcxml'. The <request> element contains a <ListRecords> element, which contains a <record> element. The <record> element contains a <header> element with the identifier 'KOHA-PUKP:1' and a <datestamp> element with the value '2016-02-11T09:33:22Z'. The <record> element also contains a <metadata> element, which contains a <collection> element with the schemaLocation 'http://www.loc.gov/MARC21/slim' and the xsi:schemaLocation 'http://www.loc.gov/MARC21/slim http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd'. The <collection> element contains a <record> element, which contains a <leader> element with the value '00418nam a22001457a 4500' and a <controlfield> element with the tag '005' and the value '20160211093322.0'. The <record> element also contains a <datafield> element with the ind1 'a' and ind2 ' ' and tag '082', which contains a <subfield> element with the code 'a' and the value '492.7'. The <record> element also contains a <datafield> element with the ind1 'a' and ind2 ' ' and tag '100', which contains a <subfield> element with the code '9' and the value '1', and a <subfield> element with the code 'a' and the value 'Mukhliah'. The <record> element also contains a <datafield> element with the ind1 'a' and ind2 ' ' and tag '245', which contains a <subfield> element with the code 'a' and the value 'Dasar-dasar bahasa arab', and a <subfield> element with the code 'c' and the value 'Editor M. Alaika Slamulloh'. The <record> element also contains a <datafield> element with the ind1 'a' and ind2 ' ' and tag '260', which contains a <subfield> element with the code 'a' and the value 'Yogyakarta', a <subfield> element with the code 'b' and the value 'Insan Madani', and a <subfield> element with the code 'c' and the value '2007'. The <record> element also contains a <datafield> element with the ind1 'a' and ind2 ' ' and tag '300'.

OAI dalam OJS

<http://ejurnal.litbang.pertanian.go.id/index.php/jpengkajian/oai>



The screenshot shows the homepage of the journal 'Jurnal Pengkajian dan Pengembangan Teknologi Pertanian' (JPPTP). The header is green with the journal title and logo. Below the header is a navigation menu with links like HOME, ABOUT, LOGIN, REGISTER, SEARCH, CURRENT, ARCHIVES, EDITORIAL, TEAM, AUTHOR GUIDELINES, PUBLICATION ETHICS AND MALPRACTICE STATEMENT, FOCUS AND SCOPE. The main content area features the journal title and a brief description: 'Jurnal Pengkajian dan Pengembangan Teknologi Pertanian (JPPTP) adalah media ilmiah penyebaran hasil penelitian dan pengkajian teknologi pertanian untuk menunjang pembangunan pertanian wilayah. Jurnal ini memuat hasil penelitian dan pengkajian primer inovasi pertanian, khususnya yang bernuansa spesifik lokasi. Jurnal ini diterbitkan secara periodik tiga kali dalam satu tahun. JPPTP sudah diindeks oleh Google Scholar'. There is also a small image of the journal cover. On the right side, there are several utility boxes: 'OPEN JOURNAL SYSTEMS' with a 'Journal Help' link; a 'USER' login section with fields for Username and Password, a 'Remember me' checkbox, and a 'Login' button; a 'NOTIFICATIONS' section with links for 'View' and 'Subscribe'; a 'LANGUAGE' section with a 'Select Language' dropdown set to 'English' and a 'Submit' button; and a 'JOURNAL CONTENT' section with a 'Search' input field.



The screenshot shows the 'OAI 2.0 Request Results' page. The browser address bar shows the URL: http://ejurnal.litbang.pertanian.go.id/index.php/jpengkajian/oai?verb=ListRecords&metadataPrefix=oai_dc. The page title is 'OAI 2.0 Request Results'. Below the title are navigation links: [Identify](#) | [ListRecords](#) | [ListSets](#) | [ListMetadataFormats](#) | [ListIdentifiers](#). A message states: 'You are viewing an HTML version of the XML OAI response. To see the underlying XML use your web browsers view source option. More information about this X...'. The response data is displayed in a table-like format:

Datestamp of response	2016-08-23T09:59:15Z
Request URL	http://ejurnal.litbang.pertanian.go.id/index.php/jpengkajian/oai

Below this, it says 'Request was of type ListRecords.' The main record is shown in a light blue box:

OAI Record:	oai:ajs.180.256.76.21:article/2274
--------------------	--

Below the record is the 'OAI Record Header' section:

OAI Identifier	oai:ajs.180.256.76.21:article/2274 oai_dc formats
Datestamp	2015-11-12T14:45:44Z
setSpec	IP:ART Identifiers Records

A message at the bottom of the record box states: 'This record has been deleted.'

Otomasi Pengolahan

SYSTEM INFORMATION

Software Platform *

OAI Base URL *

Metadata Prefix

Set (filter)

New Records Added *

Browse Repositories

Search:

Types: Journal 2928 Its 198 Digital 184 Status: Pending 2611 Ok 219 Invalid 128 Deleted 82 Total: 3310

No	Description	Type	Software	Added	Status
1	OPAC PERPUSTAKAAN STAI SULTAN ABDURRAHMAN KEPRI Institution: STAI Sultan Abdulrahman Kepulauan Riau City: Kepulauan Riau OAI: http://opac.stai-sar.ac.id/ (oai_dc) Onesearch URL: IOS3539	its	SUMS	2016-08-04 21:33:40	Pending
2	Jurnal Teknologi dan Sistem Komputer Institution: Universitas Diponegoro City: Semarang OAI: http://jtsikom.undip.ac.id/index.php/jtsikom/oai (oai_dc) Onesearch URL: IOS3538	journal	Open Journal System (OJS)	2016-08-04 18:17:01	Pending



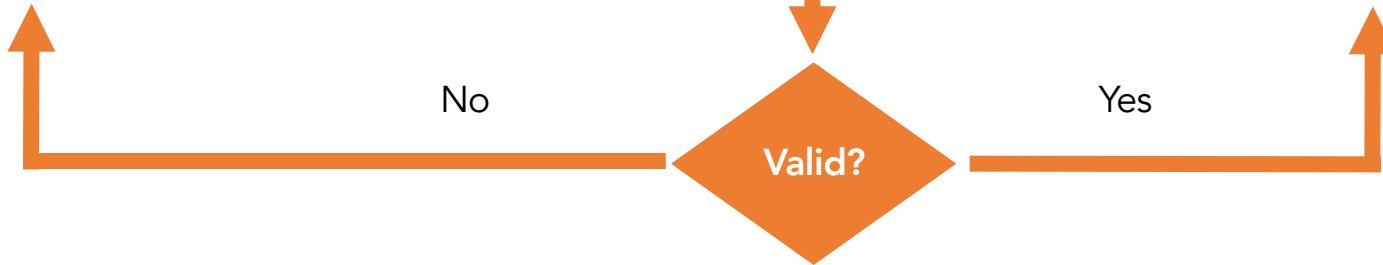
Repository Admin



Harvester Admin



OAI Harvester & Index



Deduplication

Proses deduplikasi dilakukan secara otomatis oleh sebuah program di dalam IOS. Setiap record yang mirip dengan record sebelumnya berdasarkan judul, penulis, penerbit, dan tahun, akan ditandai sebagai duplikat.

```
(
[source] => IOS3145
[context] => renormalize
[level] => INFO
[msg] => 10271 renormalized
[user] => fahmi
)
2016-08-22 19:22:18 [76150] [INFO] [deduplicate] Interrupt handler set
2016-08-22 19:22:18 [76150] [INFO] [deduplicate] Creating record list for 'IOS3145'
2016-08-22 19:22:18 [76150] [INFO] [deduplicate] Processing the records for 'IOS3145'
2016-08-22 19:24:45 [76150] [INFO] [deduplicate] 1000 records processed for 'IOS3145', 2 deduplicated, 7 records/sec
2016-08-22 19:27:00 [76150] [INFO] [deduplicate] 2000 records processed for 'IOS3145', 3 deduplicated, 7 records/sec
2016-08-22 19:29:17 [76150] [INFO] [deduplicate] 3000 records processed for 'IOS3145', 9 deduplicated, 7 records/sec
2016-08-22 19:31:39 [76150] [INFO] [deduplicate] 4000 records processed for 'IOS3145', 13 deduplicated, 7 records/sec
2016-08-22 19:34:04 [76150] [INFO] [deduplicate] 5000 records processed for 'IOS3145', 13 deduplicated, 7 records/sec
2016-08-22 19:36:37 [76150] [INFO] [deduplicate] 6000 records processed for 'IOS3145', 14 deduplicated, 7 records/sec
2016-08-22 19:38:57 [76150] [INFO] [deduplicate] 7000 records processed for 'IOS3145', 15 deduplicated, 7 records/sec
2016-08-22 19:41:14 [76150] [INFO] [deduplicate] 8000 records processed for 'IOS3145', 17 deduplicated, 7 records/sec
2016-08-22 19:43:25 [76150] [INFO] [deduplicate] 9000 records processed for 'IOS3145', 19 deduplicated, 7 records/sec
2016-08-22 19:45:34 [76150] [INFO] [deduplicate] 10000 records processed for 'IOS3145', 19 deduplicated, 7 records/sec
2016-08-22 19:46:09 [76150] [INFO] [deduplicate] Completed with 10271 records processed for 'IOS3145', 19 deduplicated
php ../onesearch/log.php --source=IOS3145 --user=fahmi --context=deduplicate --level=INFO --msg="10271 processed, 19 d
ated"
Array
(
[source] => IOS3145
[context] => deduplicate
[level] => INFO
[msg] => 10271 processed, 19 deduplicated
[user] => fahmi
)
```

Deduplication

Institution ▲

Poltekkes Kemenkes Banda Aceh ✓

Universitas Pesantren Tinggi Darul 'Ulum Jombang **10**

Collection ▲

OneSearch SLIMs UCS **2,919**

Katalog Perpustakaan Unipdu Jombang **10**

Format ▲

Book **2,858**

Book:Book **2,858**

WorkOfArt **59**

Select Page | with selected: Add to Book Bag

1



Human Physiology

Book

Published 2001

Location	Repository	IOS Number
Jakarta	OneSearch SLIMs UCS	IOS2846
Jombang	Katalog Perpustakaan Unipdu Jombang	IOS3145

Available online:

[Get online](#)

2



Ginekologi

Book

Published 1981

Location	Repository	IOS Number
Jakarta	OneSearch SLIMs UCS	IOS2846
Jombang	Katalog Perpustakaan Unipdu Jombang	IOS3145

Terimakasih

Ismail Fahmi, PhD

Email: ismail.fahmi@gmail.com

Hp: 0812 8908 3894