International Programs, Office of; Records ua550

This finding aid was produced using ArchivesSpace on March 06, 2018.

English
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary Information</td>
<td>3</td>
</tr>
<tr>
<td>Biographical History</td>
<td>3</td>
</tr>
<tr>
<td>Scope and Contents</td>
<td>4</td>
</tr>
<tr>
<td>Arrangement of the Collection</td>
<td>4</td>
</tr>
<tr>
<td>Administrative Information</td>
<td>4</td>
</tr>
<tr>
<td>Controlled Access Headings</td>
<td>5</td>
</tr>
<tr>
<td>Physical Characteristics and Technical Requirements</td>
<td>5</td>
</tr>
<tr>
<td>Collection Inventory</td>
<td>5</td>
</tr>
<tr>
<td>Website for the Center for International Education and Global Strategy</td>
<td>5</td>
</tr>
</tbody>
</table>
Summary Information

| Repository: M.E. Grenander Department of Special Collections & Archives |
| Creator: Office of International Programs |
| Title: Office of International Programs Records |
| ID: ua550 |
| Date [inclusive]: 2008 July 23 - 2017 May 17 |
| Physical Description: 20 cubic ft. |
| Physical Description: 28 captures |
| Physical Location: The materials are located onsite in the department. |
| Language of the Material: English |
| Language of the Material: English |
| Mixed Materials [Collection]: X |
| Abstract: The collection consists of inactive records from the University at Albany's Office of International Programs. |

Preferred Citation

Preferred citation for this material is as follows:

Identification of specific item, series, box, folder, Office of International Programs Records. M.E. Grenander Department of Special Collections and Archives, University Libraries, University at Albany, State University of New York.

Biographical History

The University at Albany's Office of International Programs was established in 1964. Until 1971, this office was called the Office of International Studies and World Affairs.
Scope and Contents

This collection consists of inactive records of University at Albany's Office of International Programs including: correspondence, budgets and financial statements, proposals and brochures for overseas programs and summer language programs, correspondence, and newspaper clippings.

Arrangement of the Collection

The collection is unprocessed and is likely disorganized. Individual items may be difficult to find.

Administrative Information

Publication Statement
M.E. Grenander Department of Special Collections & Archives

Revision Description
EAD file created Encoded in EAD by Conversion and encoding by ArchProteus 2014 2014

Processing Information
Processed in 2015 by unprocessed.

Acquisition Information
All items in this manuscript group were transferred to the University Libraries, M.E. Grenander Department of Special Collections and Archives.
Access

Access to this collection is restricted because it is unprocessed. Portions of the collection may contain recent administrative records and/or personally identifiable information. While it is likely that portions of the collection may be viewed, access must be managed by an archivist.

Copyright

The researcher assumes full responsibility for conforming with the laws of copyright. Whenever possible, the M.E. Grenander Department of Special Collections and Archives will provide information about copyright owners and other restrictions, but the legal determination ultimately rests with the researcher. Requests for permission to publish material from this collection should be discussed with the Head of Special Collections and Archives.

Controlled Access Headings

- Web Archives
- UAlbany Academic Administration
- Wages -- New York (State)
- State University of New York at Albany-Office of International Programs

Physical Characteristics and Technical Requirements

Web Archives

Collection Inventory

Website for the Center for International Education and Global Strategy, 2008
July 23 - 2017 May 17

Date [inclusive]: 2008 July 23 - 2017 May 17
Physical Description: 28 captures
URL

http://www.albany.edu/international/

Access Requirements

The item contains web archives preserved as WARC files. They must be access though web archival replay tools such as the "Wayback Machine." The links here direct you to files hosted by the Internet Archive, but you may also request WARC files.

Immediate Source of Acquisition:

Immediate Source of Acquisition

Web crawling is managed through the Internet Archive's Archive-It service.

This item is not a seed within an Archive-it collection, but captures of this page are found within:

collection: University at Albany, SUNY Website

created_date: 2012-09-17T13:15:00Z

created_by: bkeough

state: ACTIVE

last_updated_date: 2016-09-26T23:38:05Z

last_updated_by: system

Controlled Access Headings:

• Web Collection

<table>
<thead>
<tr>
<th>Title/Description</th>
<th>Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>University at Albany - Office of International Education</td>
<td></td>
</tr>
<tr>
<td>Date [inclusive]: 2008 July 23</td>
<td></td>
</tr>
<tr>
<td>Digital Object: University at Albany - Office of International Education</td>
<td></td>
</tr>
</tbody>
</table>

Scope and Contents

Meta tag for description: University at Albany press releases.

Immediate Source of Acquisition:

Immediate Source of Acquisition

This crawl was performed by the Internet Archive, not the UAlbany web archiving program, so the provenance is unknown.

University at Albany - Office of International Education

Date [inclusive]: 2008 September 7
Digital Object: University at Albany - Office of International Education

Scope and Contents

Meta tag for description: University at Albany press releases.

Immediate Source of Acquisition:

Immediate Source of Acquisition

This crawl was performed by the Internet Archive, not the UAlbany web archiving program, so the provenance is unknown.

University at Albany - Office of International Education

Date [inclusive]: 2013 May 8

Digital Object: University at Albany - Office of International Education

Scope and Contents

Meta tag for description: University at Albany press releases.

Immediate Source of Acquisition:

Immediate Source of Acquisition

This item had no crawl rules.
The CDX index for this capture did not contain crawl information.

WARC file for University at Albany - Office of International Education

Date [inclusive]: 2013 May 8

Immediate Source of Acquisition:

**Immediate Source of Acquisition**

This item had no crawl rules.

The CDX index for this capture did not contain crawl information.

**Physical Characteristics and Technical Requirements**

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

University at Albany - Office of International Education

Date [inclusive]: 2013 May 8

Digital Object: [University at Albany - Office of International Education](#)

**Scope and Contents**

Meta tag for description: University at Albany press releases.

Immediate Source of Acquisition:

**Immediate Source of Acquisition**

This item had no crawl rules.

The CDX index for this capture did not contain crawl information.

WARC file for University at Albany - Office of International Education

Date [inclusive]: 2013 May 8

Immediate Source of Acquisition:

**Immediate Source of Acquisition**

This item had no crawl rules.
The CDX index for this capture did not contain crawl information.

**Physical Characteristics and Technical Requirements**

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

---

&lt;title&gt; not present

**Date [inclusive]:** 2015 October 18

**Digital Object:** &lt;title&gt; not present

**Immediate Source of Acquisition:**

**Immediate Source of Acquisition**

crawl: 179047

This item had no crawl rules.

**Crawl Times**

start_date: 2015-10-17T16:56:32Z

original_start_date: 2015-10-17T16:56:32Z

last_resumption: None

processing_end_date: 2015-10-22T19:45:11Z

date: 2015-10-22T19:02:36Z

elapsed_ms: 439557452

**Crawl Types**

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

**Crawl Limits**

time_limit: 432000

document_limit: None
Crawl Results

status: FINISHED_TIME_LIMIT
discovered_count: 390211
novel_count: 36614
duplicate_count: 202550
resumption_count: 0
queued_count: 151047
downloaded_count: 239164
download_failures: 276
warc_revisit_count: 202548
warc_url_count: 239075
total_data_in_kbs: 61075441
duplicate_bytes: 53769145120
warc_compressed_bytes: 6127282356

Crawl Technical Details

doc_rate: 0.54
kb_rate: 138.0

WARC file for &lt;title&gt; not present
Date [inclusive]: 2015 October 18

Immediate Source of Acquisition:

This item had no crawl rules.

Crawl Times

start_date: 2015-10-17T16:56:32Z
original_start_date: 2015-10-17T16:56:32Z
last_resumption: None
processing_end_date: 2015-10-22T19:45:11Z
end_date: 2015-10-22T19:02:36Z
elapsed_ms: 439557452

Crawl Types
type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 390211
novel_count: 36614
duplicate_count: 202550
resumtion_count: 0
queued_count: 151047
downloaded_count: 239164
download_failures: 276
warc_revisit_count: 202548
warc_url_count: 239075
total_data_in_kbs: 61075441
duplicate_bytes: 53769145120
warc_compressed_bytes: 6127282356

Crawl Technical Details
Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

Date [inclusive]: 2015 October 18
Digital Object: &lt;title&gt; not present
Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 179047

This item had no crawl rules.

Crawl Times

start_date: 2015-10-17T16:56:32Z
original_start_date: 2015-10-17T16:56:32Z
last_resumption: None
processing_end_date: 2015-10-22T19:45:11Z
end_date: 2015-10-22T19:02:36Z
elapsed_ms: 439557452

Crawl Types

type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits

time_limit: 432000
document_limit: None
Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 390211

novel_count: 36614

duplicate_count: 202550

resumption_count: 0

queued_count: 151047

downloaded_count: 239164

download_failures: 276

warc_revisit_count: 202548

warc_url_count: 239075

total_data_in_kbs: 61075441

duplicate_bytes: 53769145120

warc_compressed_bytes: 6127282356

Crawl Technical Details

doc_rate: 0.54

kb_rate: 138.0

WARC file for &lt;title&gt; not present

Date [inclusive]: 2015 October 18

Immediate Source of Acquisition:

Immediate Source of Acquisition


crawl: 179047

This item had no crawl rules.

Crawl Times

start_date: 2015-10-17T16:56:32Z

original_start_date: 2015-10-17T16:56:32Z

last_resumption: None
processing_end_date: 2015-10-22T19:45:11Z
end_date: 2015-10-22T19:02:36Z
elapsed_ms: 439557452

Crawl Types
type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 390211
novel_count: 36614
duplicate_count: 202550
resumption_count: 0
queued_count: 151047
downloaded_count: 239164
download_failures: 276
warc_revisit_count: 202548
warc_url_count: 239075
total_data_in_kbs: 61075441
duplicate_bytes: 53769145120
warc_compressed_bytes: 6127282356

Crawl Technical Details
Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 415196
novel_count: 17867
duplicate_count: 229881
resumption_count: 0
queued_count: 167448
downloaded_count: 247748
download_failures: 242
warc_revisit_count: 229876
warc_url_count: 247648
total_data_in_kbs: 61913750
duplicate_bytes: 62091532709
warc_compressed_bytes: 256259119

Crawl Technical Details
doc_rate: 0.57
kb_rate: 143.0

WARC file for The Center for International Education and Global Strategy - University at Albany - SUNY
Date [inclusive]: 2015 November 17

Immediate Source of Acquisition:

Immediate Source of Acquisition
crawl: 183657

This item had no crawl rules.

Crawl Times
start_date: 2015-11-17T17:41:09Z
original_start_date: 2015-11-17T17:41:09Z
last_resumption: None
processing_end_date: 2015-12-03T05:50:07Z
end_date: 2015-11-30T19:18:26Z
elapsed_ms: 432098049
Crawl Types
type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False
Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None
Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 415196
novel_count: 17867
duplicate_count: 229881
resumption_count: 0
queued_count: 167448
downloaded_count: 247748
download_failures: 242
warc_revisit_count: 229876
warc_url_count: 247648
total_data_in_kbs: 61913750
duplicate_bytes: 62091532709
Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

Physical Characteristics and Technical Requirements

The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2015 December 12

Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Immediate Source of Acquisition: This crawl was performed by the Internet Archive, not the UAlbany web archiving program, so the provenance is unknown.

The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2015 December 17

Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Immediate Source of Acquisition: This item had no crawl rules.

Crawl Times

start_date: 2015-12-17T16:55:04Z
original_start_date: 2015-12-17T16:55:04Z
last_resumption: None
processing_end_date: 2015-12-22T21:57:26Z
end_date: 2015-12-22T21:40:44Z
elapsed_ms: 449129872

Crawl Types

type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits

time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT
discovered_count: 304232
novel_count: 13877
duplicate_count: 191764
resumption_count: 0
queued_count: 98591

downloaded_count: 205641
download_failures: 234
warc_revisit_count: 191759
warc_url_count: 205547
total_data_in_kbs: 60257781
duplicate_bytes: 60588343215
warc_compressed_bytes: 34726906

Crawl Technical Details

doc_rate: 0.46
WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2015 December 17

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 187970

This item had no crawl rules.

Crawl Times

start_date: 2015-12-17T16:55:04Z

original_start_date: 2015-12-17T16:55:04Z

last_resumption: None

processing_end_date: 2015-12-22T21:57:26Z

date: 2015-12-22T21:40:44Z

elapsed_ms: 449129872

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 304232

novel_count: 13877
Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 January 17

Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 191947

This item had no crawl rules.

Crawl Times

start_date: 2016-01-17T16:55:00Z

original_start_date: 2016-01-17T16:55:00Z
last_resumption: None
processing_end_date: 2016-01-22T17:27:17Z
date: 2016-01-22T16:56:58Z
elapsed_ms: 432114027
Crawl Types
type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False
Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None
Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 573952
novel_count: 20681
duplicate_count: 273347
resumption_count: 0
queued_count: 279924
downloaded_count: 294028
download_failures: 181
warc_revisit_count: 273342
warc_url_count: 293928
total_data_in_kbs: 63947333
duplicate_bytes: 64312716410
warc_compressed_bytes: 176478309
Crawl Technical Details

doc_rate: 0.68

kb_rate: 147.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY
Date [inclusive]: 2016 January 17

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 191947

This item had no crawl rules.

Crawl Times

start_date: 2016-01-17T16:55:00Z

original_start_date: 2016-01-17T16:55:00Z

last_resumption: None

processing_end_date: 2016-01-22T17:27:17Z

date: 2016-01-22T16:56:58Z

elapsed_ms: 432114027

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT
 Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

The Center for International Education and Global Strategy
CIEGS - University at Albany - SUNY
Date [inclusive]: 2016 February 18
Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY
Immediate Source of Acquisition:
Immediate Source of Acquisition

crawl: 196723

Crawl Rules
Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times

start_date: 2016-02-17T16:57:12Z

original_start_date: 2016-02-17T16:57:12Z

last_resumption: None

processing_end_date: 2016-02-23T04:10:10Z

end_date: 2016-02-22T23:20:27Z

elapsed_ms: 454987781

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 7638127

novel_count: 795824

duplicate_count: 806868

resumption_count: 0

queued_count: 6035435

downloaded_count: 1602692

download_failures: 248
warc_revisit_count: 806856
warc_url_count: 1602561
total_data_in_kbs: 226292461
duplicate_bytes: 193554304269
warc_compressed_bytes: 633147239

Crawl Technical Details

doc_rate: 3.52
kb_rate: 497.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY
Date [inclusive]: 2016 February 18

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 196723

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times
start_date: 2016-02-17T16:57:12Z
original_start_date: 2016-02-17T16:57:12Z
last_resumption: None
processing_end_date: 2016-02-23T04:10:10Z
end_date: 2016-02-22T23:20:27Z
elapsed_ms: 454987781

Crawl Types
type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits
Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 7638127

novel_count: 795824

duplicate_count: 806868

resumption_count: 0

queued_count: 6035435

downloaded_count: 1602692

download_failures: 248

warc_revisit_count: 806856

warc_url_count: 1602561

total_data_in_kbs: 226292461

duplicate_bytes: 193554304269

warc_compressed_bytes: 633147239

Crawl Technical Details

doc_rate: 3.52

kb_rate: 497.0

Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

The Center for International Education and Global Strategy
CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 March 19
Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 201840

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for iibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times

start_date: 2016-03-17T16:54:59Z

original_start_date: 2016-03-17T16:54:59Z

last_resumption: None

processing_end_date: 2016-03-22T18:02:10Z

date: 2016-03-22T16:56:57Z

elapsed_ms: 432107388

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 639552

novel_count: 138564

duplicate_count: 332568
resumption_count: 0
queued_count: 168420
downloaded_count: 471132
download_failures: 1411433
warc_revisit_count: 332563
warc_url_count: 471037
total_data_in_kbs: 116820623
duplicate_bytes: 113225737713
warc_compressed_bytes: 765920585

Crawl Technical Details

doc_rate: 1.09
kb_rate: 270.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 March 19

Immediate Source of Acquisition:

**Immediate Source of Acquisition**

crawl: 201840

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times

start_date: 2016-03-17T16:54:59Z

original_start_date: 2016-03-17T16:54:59Z

last_resumption: None

processing_end_date: 2016-03-22T18:02:10Z

derelative_time: 432107388

type: MONTHLY

recurrence_type: MONTHLY
Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request.
remotely. Please consult an archivist if you are interested in advanced research with web archives.

The Center for International Education and Global Strategy
CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 April 18

Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 207692

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times

start_date: 2016-04-17T16:55:13Z

original_start_date: 2016-04-17T16:55:13Z

last_resumption: None

processing_end_date: 2016-04-22T22:29:09Z

date: 2016-04-22T18:00:08Z

elapsed_ms: 435852385

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT
discovered_count: 7165982

novel_count: 582992

duplicate_count: 879835

resumption_count: 0

queued_count: 5703155

downloaded_count: 1462827

download_failures: 284

warc_revisit_count: 879824

warc_url_count: 1462700

total_data_in_kbs: 227675763

duplicate_bytes: 214936501775

warc_compressed_bytes: 813502772

Crawl Technical Details

doc_rate: 3.36

kb_rate: 522.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 April 18

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 207692

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times

start_date: 2016-04-17T16:55:13Z

original_start_date: 2016-04-17T16:55:13Z

last_resumption: None

processing_end_date: 2016-04-22T22:29:09Z

end_date: 2016-04-22T18:00:08Z

elapsed_ms: 435852385
Crawl Types

type: MONTHLY
recurrence_type: MONTHLY

pdfs_only: False
test: False

Crawl Limits

time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT
discovered_count: 7165982
novel_count: 582992
duplicate_count: 879835
resumption_count: 0
queued_count: 5703155
downloaded_count: 1462827
download_failures: 284
warc_revisit_count: 879824
warc_url_count: 1462700
total_data_in_kbs: 227675763
duplicate_bytes: 214936501775
warc_compressed_bytes: 813502772

Crawl Technical Details
doc_rate: 3.36
kb_rate: 522.0
Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.
Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times

start_date: 2016-05-17T16:55:02Z

original_start_date: 2016-05-17T16:55:02Z

last_resumption: None

processing_end_date: 2016-05-24T10:38:52Z

end_date: 2016-05-22T22:20:27Z

elapsed_ms: 451503421

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None
crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 8685839

novel_count: 568136

duplicate_count: 1043240

resumption_count: 0

queued_count: 7074463

downloaded_count: 1611376

download_failures: 229

warc_revisit_count: 1043205

warc_url_count: 1611247

total_data_in_kbs: 246906729

duplicate_bytes: 230750229955

warc_compressed_bytes: 2352496320

Crawl Technical Details

doc_rate: 3.57

kb_rate: 546.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 May 18

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 212937

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)
Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times

start_date: 2016-05-17T16:55:02Z

original_start_date: 2016-05-17T16:55:02Z

last_resumption: None
processing_end_date: 2016-05-24T10:38:52Z
end_date: 2016-05-22T22:20:27Z
elapsed_ms: 451503421

Crawl Types
type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 8685839
novel_count: 568136
duplicate_count: 1043240
resumption_count: 0
queued_count: 7074463
downloaded_count: 1611376
download_failures: 229
warc_revisit_count: 1043205
warc_url_count: 1611247
total_data_in_kbs: 246906729
duplicate_bytes: 23075022955
warc_compressed_bytes: 2352496320

Crawl Technical Details
Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.
Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times

start_date: 2016-06-17T16:55:03Z
original_start_date: 2016-06-17T16:55:03Z
last_resumption: None
processing_end_date: 2016-06-23T01:46:32Z
end_date: 2016-06-22T19:28:00Z
elapsed_ms: 441165235

Crawl Types

type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits
time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 8887536

novel_count: 557229

duplicate_count: 1067492

resumption_count: 0

queued_count: 7262815

downloaded_count: 1624721

download_failures: 247

warc_revisit_count: 1067450

warc_url_count: 1624592

total_data_in_kbs: 251943396

duplicate_bytes: 239433481575

warc_compressed_bytes: 1787914405

Crawl Technical Details

doc_rate: 3.68

kb_rate: 571.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 June 17

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 220038

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)
start_date: 2016-06-17T16:55:03Z
original_start_date: 2016-06-17T16:55:03Z
last_resumption: None
processing_end_date: 2016-06-23T01:46:32Z
end_date: 2016-06-22T19:28:00Z
elapsed_ms: 441165235
Crawl Types
type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False
Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None
Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 8887536
novel_count: 557229
duplicate_count: 1067492
resumption_count: 0
queued_count: 7262815
downloaded_count: 1624721
download_failures: 247
warc_revisit_count: 1067450
warc_url_count: 1624592
total_data_in_kbs: 251943396
Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.
Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times

start_date: 2016-07-17T16:55:03Z

original_start_date: 2016-07-17T16:55:03Z

last_resumption: None

processing_end_date: 2016-07-22T22:52:35Z

end_date: 2016-07-22T17:55:28Z

elapsed_ms: 435383503

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 6792375
novel_count: 212820
duplicate_count: 1098346
resumption_count: 0
queued_count: 5481209
downloaded_count: 1311166
download_failures: 256
warc_revisit_count: 1098317
warc_url_count: 1311042
total_data_in_kbs: 212249210
duplicate_bytes: 207609316437
warc_compressed_bytes: 675349961

Crawl Technical Details
doc_rate: 3.01
kb_rate: 487.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY
Date [inclusive]: 2016 July 17
Immediate Source of Acquisition:
Immediate Source of Acquisition

crawl: 225859

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Crawl Times

- start_date: 2016-07-17T16:55:03Z
- original_start_date: 2016-07-17T16:55:03Z
- last_resumption: None
- processing_end_date: 2016-07-22T22:52:35Z
- end_date: 2016-07-22T17:55:28Z
- elapsed_ms: 435383503

Crawl Types

- type: MONTHLY
- recurrence_type: MONTHLY
- pdfs_only: False
- test: False

Crawl Limits

- time_limit: 432000
- document_limit: None
- byte_limit: None
- crawl_stop_requested: None

Crawl Results

- status: FINISHED_TIME_LIMIT
- discovered_count: 6792375
- novel_count: 212820
- duplicate_count: 1098346
- resumption_count: 0
- queued_count: 5481209
- downloaded_count: 1311166
download_failures: 256
warc_revisit_count: 1098317
warc_url_count: 1311042
total_data_in_kbs: 212249210
duplicate_bytes: 20760931637
warc_compressed_bytes: 675349961

Crawl Technical Details

doc_rate: 3.01
kb_rate: 487.0

Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

---

The Center for International Education and Global Strategy  
CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 July 21

Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Immediate Source of Acquisition:

Immediate Source of Acquisition

This crawl was performed by the Internet Archive, not the UAlbany web archiving program, so the provenance is unknown.

---

The Center for International Education and Global Strategy  
CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 August 17

Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 231044

Crawl Rules
Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2016-08-17T16:58:06Z
original_start_date: 2016-08-17T16:58:06Z
last_resumption: None
processing_end_date: 2016-08-23T01:20:50Z
end_date: 2016-08-22T23:33:45Z
elapsed_ms: 455703830

Crawl Types

type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits

time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT
discovered_count: 1189467
novel_count: 184079
duplicate_count: 736427
resumption_count: 0
queued_count: 268961
downloaded_count: 920506
download_failures: 267
warc_revisit_count: 736404
warc_url_count: 920384
total_data_in_kbs: 202254345
duplicate_bytes: 198029314926
warc_compressed_bytes: 397770039

Crawl Technical Details

doc_rate: 2.02
kb_rate: 443.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY
Date [inclusive]: 2016 August 17

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 231044

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)
Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)
Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2016-08-17T16:58:06Z

original_start_date: 2016-08-17T16:58:06Z

last_resumption: None

processing_end_date: 2016-08-23T01:20:50Z

date: 2016-08-22T23:33:45Z

eelapsed_ms: 455703830

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

ddfs_only: False

test: False
Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 1189467

novel_count: 184079

duplicate_count: 736427

resumption_count: 0

queued_count: 268961

downloaded_count: 920506

download_failures: 267

warc_revisit_count: 736404

warc_url_count: 920384

total_data_in_kbs: 202254345

duplicate_bytes: 198029314926

warc_compressed_bytes: 397770039

Crawl Technical Details

doc_rate: 2.02

kb_rate: 443.0

Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.
The Center for International Education and Global Strategy
CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 August 23

Immediate Source of Acquisition:

This crawl was performed by the Internet Archive, not the UAlbany web archiving program, so the provenance is unknown.

The Center for International Education and Global Strategy
CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 September 17

Immediate Source of Acquisition:

crawl: 236726

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2016-09-17T16:55:32Z

original_start_date: 2016-09-17T16:55:32Z

last_resumption: None

processing_end_date: 2016-09-22T19:45:41Z

end_date: 2016-09-22T17:03:18Z

elapsed_ms: 432208662

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False
Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 1414002

novel_count: 188377

duplicate_count: 898376

resumption_count: 0

queued_count: 327249

downloaded_count: 1086753

download_failures: 287

warc_revisit_count: 898345

warc_url_count: 1086634

total_data_in_kbs: 205611728

duplicate_bytes: 200652210118

warc_compressed_bytes: 309191153

Crawl Technical Details

doc_rate: 2.51

kb_rate: 475.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 September 17

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 236726

Crawl Rules
Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2016-09-17T16:55:32Z
original_start_date: 2016-09-17T16:55:32Z
last_resumption: None
processing_end_date: 2016-09-22T19:45:41Z
end_date: 2016-09-22T17:03:18Z
elapsed_ms: 432208662

Crawl Types

type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits

time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT
discovered_count: 1414002
novel_count: 188377
duplicate_count: 898376
resumption_count: 0
queued_count: 327249
downloaded_count: 1086753
Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

Crawl Technical Details

doc_rate: 2.51

kb_rate: 475.0

Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

The Center for International Education and Global Strategy
CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 October 17

Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 243096

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)
Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2016-10-17T16:56:12Z

original_start_date: 2016-10-17T16:56:12Z

last_resumption: None
processing_end_date: 2016-10-22T20:32:46Z
end_date: 2016-10-22T17:07:30Z
elapsed_ms: 432270469

Crawl Types
type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 1819420
novel_count: 557303
duplicate_count: 900445
resumption_count: 0
queued_count: 361672
downloaded_count: 1457748
download_failures: 296
warc_revisit_count: 900407
warc_url_count: 1457649
total_data_in_kbs: 215568357
duplicate_bytes: 201221539907
warc_compressed_bytes: 1485925187

Crawl Technical Details
WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 October 17

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 243096

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2016-10-17T16:56:12Z

original_start_date: 2016-10-17T16:56:12Z

last_resumption: None

processing_end_date: 2016-10-22T20:32:46Z

end_date: 2016-10-22T17:07:30Z

elapsed_ms: 432270469

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results
status: FINISHED_TIME_LIMIT

discovered_count: 1819420

novel_count: 557303

duplicate_count: 900445

resumption_count: 0

queued_count: 361672

downloaded_count: 1457748

download_failures: 296

warc_revisit_count: 900407

warc_url_count: 1457649

total_data_in_kbs: 215568357

duplicate_bytes: 201221539907

warc_compressed_bytes: 1485925187

Crawl Technical Details

doc_rate: 3.37

kb_rate: 498.0

**Physical Characteristics and Technical Requirements**

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

---

The Center for International Education and Global Strategy
CIEGS - University at Albany - SUNY

*Date [inclusive]:* 2016 November 17

*Digital Object:* The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

*Immediate Source of Acquisition:*

*Immediate Source of Acquisition*

crawl: 249672

Crawl Rules

- Page 89-
Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2016-11-17T16:57:23Z
original_start_date: 2016-11-17T16:57:23Z
last_resumption: None
processing_end_date: 2016-11-22T23:26:30Z
end_date: 2016-11-22T18:15:59Z
elapsed_ms: 436686458

Crawl Types

type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits

time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT
discovered_count: 6455397
novel_count: 904107
duplicate_count: 647491
resumption_count: 0
queued_count: 4903799
downloaded_count: 1551598
download_failures: 287
warc_revisit_count: 647445
warc_url_count: 1551482
total_data_in_kbs: 234676600
duplicate_bytes: 166348006398
warc_compressed_bytes: 22645478722

Crawl Technical Details
doc_rate: 3.55
kb_rate: 537.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 November 17

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 249672

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)
Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)
Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2016-11-17T16:57:23Z

original_start_date: 2016-11-17T16:57:23Z

last_resumption: None

processing_end_date: 2016-11-22T23:26:30Z

end_date: 2016-11-22T18:15:59Z

elapsed_ms: 436686458

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False
Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 6455397

novel_count: 904107

duplicate_count: 647491

resumption_count: 0

queued_count: 4903799

downloaded_count: 1551598

download_failures: 287

warc_revisit_count: 647445

warc_url_count: 1551482

total_data_in_kbs: 234676600

duplicate_bytes: 166348006398

warc_compressed_bytes: 22645478722

Crawl Technical Details

doc_rate: 3.55

kb_rate: 537.0

Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.
Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2016-12-17T16:54:59Z

original_start_date: 2016-12-17T16:54:59Z

last_resumption: None

processing_end_date: 2016-12-22T23:02:36Z

end_date: 2016-12-22T17:39:38Z

elapsed_ms: 434651176

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stopRequested: None

Crawl Results

status: FINISHED_TIME_LIMIT
discovered_count: 5605147
novel_count: 672626
duplicate_count: 721586
resumption_count: 0
queued_count: 4210935
downloaded_count: 1394212
download_failures: 247
warc_revisit_count: 721559
warc_url_count: 1394095
total_data_in_kbs: 173161585
duplicate_bytes: 152072050167
warc_compressed_bytes: 4226018506

Crawl Technical Details

doc_rate: 3.21

kb_rate: 398.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2016 December 17

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 255339

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2016-12-17T16:54:59Z

original_start_date: 2016-12-17T16:54:59Z

last_resumption: None

processing_end_date: 2016-12-22T23:02:36Z

end_date: 2016-12-22T17:39:38Z
elapsed_ms: 434651176

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 5605147

novel_count: 672626

duplicate_count: 721586

resumption_count: 0

queued_count: 4210935

downloaded_count: 1394212

download_failures: 247

warc_revisit_count: 721559

warc_url_count: 1394095

total_data_in_kbs: 173161585

duplicate_bytes: 15207205016

warc_compressed_bytes: 4226018506

Crawl Technical Details

doc_rate: 3.21

kb_rate: 398.0
Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Date [inclusive]: 2017 January 17

Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 261028

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2017-01-17T17:03:42Z
original_start_date: 2017-01-17T17:03:42Z
last_resumption: None
processing_end_date: 2017-01-23T00:48:05Z
end_date: 2017-01-22T19:31:33Z
elapsed_ms: 440845477

Crawl Types

type: MONTHLY
recurrence_type: MONTHLY

pdfs_only: False
test: False

Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 6856243
novel_count: 810250
duplicate_count: 844972
resumption_count: 0
queued_count: 5201021
downloaded_count: 1655222
download_failures: 351
warc_revisit_count: 844920
warc_url_count: 1655090
total_data_in_kbs: 260833049
duplicate_bytes: 210669517964
warc_compressed_bytes: 15598206209

Crawl Technical Details
doc_rate: 3.75
kb_rate: 591.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY
Date [inclusive]: 2017 January 17

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 261028

Crawl Rules
Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)
Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times
start_date: 2017-01-17T17:03:42Z
original_start_date: 2017-01-17T17:03:42Z
last_resumption: None
processing_end_date: 2017-01-23T00:48:05Z
end_date: 2017-01-22T19:31:33Z
elapsed_ms: 440845477

Crawl Types

type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits

time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT
discovered_count: 6856243
novel_count: 810250
duplicate_count: 844972
resumption_count: 0
queued_count: 5201021
downloaded_count: 1655222
download_failures: 351
warc_revisit_count: 844920
warc_url_count: 1655090
total_data_in_kbs: 260833049
Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.
Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2017-04-17T16:55:09Z
original_start_date: 2017-04-17T16:55:09Z
last_resumption: None
processing_end_date: 2017-04-22T22:22:12Z
end_date: 2017-04-22T17:55:16Z
elapsed_ms: 435220728

Crawl Types
type: MONTHLY
recurrence_type: MONTHLY
pdfs_only: False
test: False

Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None

Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 8221578
novel_count: 736180
duplicate_count: 1311488
resumption_count: 0
queued_count: 6173910
downloaded_count: 2047668
download_failures: 335
warc_revisit_count: 1311435
warc_url_count: 2047538
total_data_in_kbs: 284745035
duplicate_bytes: 261915565851
warc_compressed_bytes: 2411311996

Crawl Technical Details
doc_rate: 4.7
kb_rate: 654.0

WARC file for The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY
Date [inclusive]: 2017 April 17
Immediate Source of Acquisition:
Immediate Source of Acquisition

crawl: 291352

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Crawl Times

start_date: 2017-04-17T16:55:09Z

original_start_date: 2017-04-17T16:55:09Z

last_resumption: None

processing_end_date: 2017-04-22T22:22:12Z

date: 2017-04-22T17:55:16Z

elapsed_ms: 435220728

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000

document_limit: None

byte_limit: None

crawl_stop_requested: None

Crawl Results

status: FINISHED_TIME_LIMIT

discovered_count: 8221578

novel_count: 736180

duplicate_count: 1311488

resumption_count: 0

queued_count: 6173910
Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

---

The Center for International Education and Global Strategy
CIEGS - University at Albany - SUNY

Date [inclusive]: 2017 May 17

Digital Object: The Center for International Education and Global Strategy CIEGS - University at Albany - SUNY

Immediate Source of Acquisition:

Immediate Source of Acquisition

crawl: 301465

Crawl Rules

Ignore Robots.txt for www.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ualbanysports.com (last updated 2016-02-11)

Ignore Robots.txt for library.albany.edu (last updated 2017-05-19)
Ignore Robots.txt for alumni.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for asrc.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for atmos.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for bioinformatics.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cela.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for choose.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cs.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for csda.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for ibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Host Rule Type C CONTAINS meq.library.albany.edu:8080/archive/search? (last updated 2017-05-11)

Crawl Times

start_date: 2017-05-17T16:55:17Z
original_start_date: 2017-05-17T16:55:17Z
last_resumption: None
processing_end_date: 2017-05-23T00:29:43Z
end_date: 2017-05-22T21:46:25Z
elapsed_ms: 449440076
Crawl Types
type: MONTHLY
recurrence_type: MONTHLY
pdffs_only: False
test: False
Crawl Limits
time_limit: 432000
document_limit: None
byte_limit: None
crawl_stop_requested: None
Crawl Results
status: FINISHED_TIME_LIMIT
discovered_count: 2923335
novel_count: 388249
duplicate_count: 1238165
resumption_count: 0
queued_count: 1296921
downloaded_count: 1626414
download_failures: 336
warc_revisit_count: 1238113
warc_url_count: 1626286
total_data_in_kbs: 265327774
duplicate_bytes: 251274786133
Ignore Robots.txt for imls.ctg.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for www.ctg.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for cwig.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for events.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for hr.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for iibl.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for illiad.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for liblogs.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for libguides.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for scholarsarchive.library.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for listserv.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for m.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for omega.math.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for mumford.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for nyjm.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for pdp.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for resnet.albany.edu (last updated 2016-02-11)
Ignore Robots.txt for rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for cyberphysics.rit.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for rna.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for slsc.albany.edu (last updated 2017-05-19)

Ignore Robots.txt for uaems.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for uapps.albany.edu (last updated 2016-02-11)

Ignore Robots.txt for wiki.albany.edu (last updated 2016-02-11)

Block host dev.library.albany.edu

Host Rule Type C CONTAINS
meg.library.albany.edu:8080/archive/search? (last updated 2017-05-11)

Crawl Times

start_date: 2017-05-17T16:55:17Z

original_start_date: 2017-05-17T16:55:17Z

last_resumption: None

processing_end_date: 2017-05-23T00:29:43Z

date: 2017-05-22T21:46:25Z

elapsed_ms: 449440076

Crawl Types

type: MONTHLY

recurrence_type: MONTHLY

pdfs_only: False

test: False

Crawl Limits

time_limit: 432000
Physical Characteristics and Technical Requirements

Researchers interested in data analysis with web archives may request a WARC file. WARC files are very large and difficult to work with. Your request may take time to process, and we may be unable to deliver your request remotely. Please consult an archivist if you are interested in advanced research with web archives.

^ Return to Table of Contents