

On Implementing Hydra for Special Collections at Yale

Eric James programmer/analyst eric.james@yale.edu

June 9, 2015



Legacy



- AMEEL (A Middle Eastern Electronic Library) (2007)
- JSS (Joel Sumner Smith Slavic Collection) (2010)
- YFAD (Yale Finding Aid Database) (2009)

Stack:

/tomcat/fedora/solr/gsearch/

Main issues:

Digitization workflow

Content models (fedora 3)

Arabic OCR (VERUS)

Slavic language indexing (solr)

EAD style sheets



AMEEL

2005-2010 Yale University Library

Send comments to AMEEL

This file last modified Sat Mar 19 14:31:28 EDT 2011



YALE UNIVERSITY LIBRARY Arabic and Middle Eastern Electronic Library

Query: Search **Available Materials** Instructions **±** Gazettes Hover over Available Materials or Databases content in the left panel **±** Journals for specific instructions. ■ Manuscripts ± University of London **±** Yale University Library ■ References ± Dictionaries ± Encyclopaedia of Islam, THREE ■ Manuscript Catalogs 🛨 A list of Arabic manuscripts in Princeton University Arabic manuscripts in the Yale University Library ➡ Bibliothecæ Bodleianæ codicum manuscriptorum orientalium ... catalogus ... 🛨 Catalogus codicum arabicorum bibliothecae academiae Lugduno-Batavae / auctoribus M. J. De Goeje et M. Th. Houtsma. ☐ Catalogus codicum manuscriptorum orientalium qui in Museo Britannico asservantur : partem secundam codices Arabicos amplectentem / curaverunt W. Cureton et Ch. Rieu. ☐ Catalogus codicum manuscriptorum orientalium qui in Museo Britannico asservantur : partem secundam codices Arabicos amplectentem / curaverunt W. Cureton et Ch. Rieu. ± Supplement of the Catalogue of the Arabic manuscripts in the British museum; by Charles Rieu, PH.D. Printed by order of the Trustees. 🗄 Verzeichnis der arabischen Handschriften der Königlichen Universitätsbibliothek zu Tübingen. I. Von Dr. Christian Seybold ... Databases Middle East Research Services (MERS) MENAlib AMEEL Home Conditions of Use Yale University Library



Joel Sumner Smith



Yale UNIVERSITY LIBRARY Digital Collections	Research Tools	Libraries & Collections	About the Library	Library Services	
Joel Sumner Smith Collection					
Query: Sear	ch				
There are 508 book/volumes starting with the letter S. Page 1 of 26					
Next Page Last Page					
ı. <u>Sbornik "Nivy"</u>					
2. Sbornik "Nivy"					
3. Sbornik kriticheskikh stateĭ o N. A. Nekrasoviĉ` Sostavil V. Zelinsktĭ.					
4. Sbornik piŝŝen bukovinskago naroda. : [A collection of Bukovina folk songs] / Sostavil A. Lonachevskii. (Iz mater'iāla, dostavlennago G.L. Kupchankon v IUgo-zapadnyī otdiēl Imperatorskago russkago geograficheskago obshchestva).					
5. Sbornik russkikh narodnykh liricheskikh pesen					
6. Sbornik russkikh narodnykh piêsen / sostavlennyi M. Balakirevym.					
7. Sbornik stateř, chitannykh v Otdřělenři russkago řážyka ř slovesnosti. Volume 1					
8. Sbornik stateř, chitannykh v Otdiělenii russkago iážyka i slovesnosti. Volume 2					
9. Sbornik stateĭ, chitannykh v Otdíĕlenři russkago iážyka ř slovesnosti. Volume 3					
10. <u>Sbornik stateř, chitannykh v Otdřelenii russkago řážyka i slovesnosti. Volume 4</u>					
11. Sbornik stateř, chitannykh v Otdiélenii russkago jážyka ř slovesnosti. Volume 5					
12. <u>Sbornik stateř, chitannykh v Otdiělenři russkago řážyka ř slovesnosti. Volume 6</u>					
13. <u>Sbornik stateř, chitannykh v Otdiělenři russkago ižžyka ř slovesnosti. Volume 7</u>					
14. <u>Sbornik sviëdiëhii po knizhno-literaturnomu diëlu za 1866 god / Izdanie knigoprodavtŝà A. Cherenina.</u>					
15. <u>Sbornik. Volume 10</u>					
16. Sbornik. Volume 11					
17. Sbornik. Volume 12					
18. Sbornik. Volume 13					
19. Sbornik. Volume 14					
20. Shornik. Volume 15					

Search this Site / Orbis Library Catalog / Contact Us / Yale Info / Yale University
Contact: Carolyn Caizzi. Certifying Authority: Elizabeth Beaudin.
© 2011 Yale University Library. This file last modified October 26.2011.





Yale UNIVERSITY LIBRARY

Finding Aid Database Help Feedback



Browse by Repository Manuscripts and Archives

Arts Library Beinecke Library Divinity Library Medical Historical Library Music Library Visual Resources Collection Yale Center for British Art

Lewis Walpole Library

Yale Peabody Museum

Welcome to the Yale Finding Aid Database

This database consists of finding aids for archival and manuscript materials at Yale University. Their purpose is to provide information about the creation, historical context, arrangement, and content of archival and manuscript collections as well as the information necessary for users to identify and request the portions relevant to their research. These finding aids vary in style, depth of detail, and amount of materials covered, but their common purpose is to describe and enable access to archival and manuscript collections.

Finding aids represent holdings from:

- · Manuscripts and Archives (Sterling Memorial Library)
- Robert B. Haas Family Arts Library Special Collections
- Beinecke Rare Book and Manuscript Library
- · Yale Divinity School Library
- Historical Library, Harvey Cushing/John Hay Whitney Medical Library
- · Irving S. Gilmore Music Library
- · Visual Resources Collection
- · Yale Center for British Art, Department of Rare Books and Manuscripts
- The Lewis Walpole Library
- · Yale Peabody Museum of Natural History

For more information on conducting research with archival and manuscript materials at Yale, see Primary Sources at Yale.

Medieval and Renaissance Manuscripts at the Beinecke Rare Book and Manuscript Library are now searchable in a separate database.

Requesting Materials

Requests to view materials in the reading rooms of the Beinecke Library and Manuscripts and Archives must be submitted through this online system. Links to the request system are found at the box level in each Beinecke Library and Manuscripts and Archives finding aid. Requests to view materials at the Divinity Library and the Arts Library must be submitted through separate online forms: Divinity Library request form, Arts Library request form. To request materials at other special collections repositories please use the contact information listed in each finding aid.

Search this Site | Orbis Library Catalog | Feedback | Yale Info | Yale University | Ask! a Librarian

© 2013 Yale University Library.



Current hydra(blacklight) 1.2^6 objects



Yale university library



Limit your search Creator Date **Content Type** Topic Language **Digital Collection** Henry A. Kissinger Papers 15918 Lewis Walpole Library 7703 Day Missions Periodicals 3088 Israel Sack Furniture Archive 2682 Day Missions Annual Reports 1738 1464 Yale Indian Papers 850 Drama School Collection 416 Sanborn Fire Insurance Maps 187 Maurice Durand Han Nom Arabic and Persian Medical **Books and Manuscripts** 72 Persian Philological Texts 22 Access Restrictions **Call Number** Repository **Date Created**

Yale University Library Digital Collections

Welcome to the new Yale University Library Digital Collections repository.

This interface will ultimately replace the many individual digital collections now available from the Yale University Library.

During this period of development, we will add both new collections and new features and are very interested in your feedback.

These collections are currently in the repository:

Arabic and Persian Medical Books and Manuscripts

Arabic and Persian manuscripts and books, as well as early translations of Arabic and Persian works dating from 1300 to 1921, from the Medical Historical Library, Cushing/Whitney Medical Library.

Connecticut Sanborn Fire Insurance Maps

Over 6,500 maps in 470 atlases for Connecticut towns dating from 1880 to 1970 at 1:600 scale. These highly detailed maps provide building, outbuilding, and property footprints, labeled streets, addresses, and information about building materials and construction features for urban and residential areas of Connecticut.

Day Missions Collection: Annual Reports

Annual reports of mission agencies and institutions document educational, medical, and religious work worldwide. Dating primarily from the 1830s through the 1930s, these reports form part of the Yale Divinity Library's Day Missions Collection, the preeminent North American collection for documentation of the history of missions and world Christianity.

Day Missions Collection: Periodicals

Periodicals published by mission agencies and worldwide religious organizations provide documentation of educational, medical, and religious work, and eyewitness accounts of events and conditions. These periodicals form part of the Yale Divinity Library's Day Missions Collection, the preeminent North American collection for documentation of the history of missions and world Christianity.

Israel Sack Furniture Archive

The Israel Sack Furniture Archive is a comprehensive digital database of American decorative arts, consisting of over 7,000 records cataloging the material objects bought and sold by the firm of Israel Sack, Inc. As the premier antiques firm representing early American furniture for much of the twentieth century, Israel Sack, Inc. was an influential corporation, operating from 1905 to 2002 in Boston and then New York City. The Israel Sack Furniture Archive will provide students, scholars and decorative arts enthusiasts alike with an unparalleled resource of comparative material for the study of early American furniture.

Lewis Walpole Digital Images Collection

Highlights from the Lewis Walpole Library's eighteenth-century British collections: visual caricatures and satires; Horace Walpole and Strawberry Hill—views, contents, and extra-illustrated copies of Walpole's "Description of the Villa"; selected topographical, portrait, and historical prints and drawings; ephemera such as trade cards, advertisements, bookplates, and playbills, and more.

Maurice Durand Han Nom Handwritten and Woodblock Manuscripts

Original woodblock or brush ink texts and translations of Han Nom texts into modern Romanized Vietnamese collected by Maurice Durand, a prominent Vietnamese/French scholar of Han Nom from the mid-20th century. Han Nom script uses classical Chinese characters to represent Sino-Vietnamese vocabulary and some native Vietnamese words, while other words are represented using locally created characters based on the Chinese model.

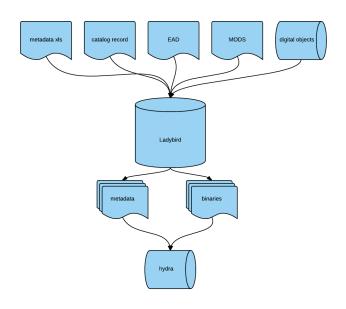
Persian Philological Texts

Selected older Persian philology texts which originate from South Asia, are rare European translations, or are reprints. All of the selected items are in the "endangered" preservation category. Many have extensive marginalia from scholars, or are from the personal libraries of Yale scholars such as Edward Salisbury.



Ingest





Ingest



- Rake task constantly polling for content in ladybird queue table (hydra_publish and child table hydra_publish_path)
- These tables have properties, timestamps, and file locations
- The metadata files (desc, access, rights) and content (binaries) ingested from mounted disk
- Uses content model (Simple, Complex Child, Complex Parent, Complex childUnstruct)



- Able to run concurrently, 1->5 instances improved throughput from 1.93 to 0.77 sec/obj
- Concurrency required use of stored procedures with SQL insert transfer rather than use of SQL updates due to locking issues



- Hydra_publish table property proliferation (viewOpt, ingestServer,handle,priority, attempts, hierarchyLevel, # of digital children)
- Frequent metadata updates a pain point mistakes, metadata schema changes (1 ex: ISO dates for date slider facet)



- Errors happen
- Use database error table for quick lookup
- Use well labeled and concise logging (grep is your friend)



- Pluggable conditional workflow sequences,
- Quick turnaround to add features such as handles, and OCR solr fieldtype conditionals

Contextual Navigation



- Scale of Henry Kissinger Papers (13000 containers, 7 layers)
- Breadcrumbs
- Context tree
- Search within

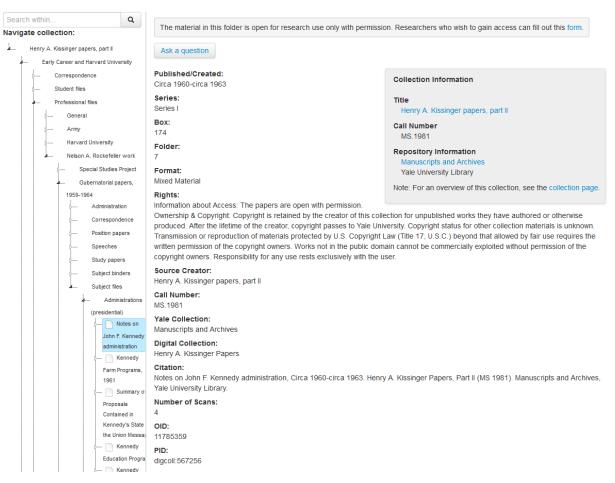
Contextual Navigation





Henry A. Kissinger papers, part II > Early Career and Harvard University > Professional files > Nelson A. Rockefeller work > Gubernatorial papers, 1959-1964 > Subject files > Administrations (presidential) > Notes on John F. Kennedy administration

Notes on John F. Kennedy administration



Context tree



- Javascript jstree implementation
- Backed by web service within hydra that leverages solr to create json nexting
- AUTH/Z baked in for filtering selective material
- Lazy loading (chunking via toplevel, direct selection, sibling and hierarchy context supplementation, and blocks

Breadcrumbs and search within



- 2 fields directly indexed leveraging hierarchical relationships
- Breadcrumbs (component titles and links)
- Hierarchy (space separated list of PIDs going down the hierarchy ending in a wildcard
- So "digcoll:parent digcoll:child*" is used as a filter to search within grandchildren like "digcoll:parent digcoll:child digcoll:grandchildX"

Full text search



- Default access
- Full access
- Selected access
- Default search in solr fulltext_open field
- Full search in solr fulltext_open AND fulltext_restricted fields
- Selected search in solr fulltext_open OR (fulltext_restricted AND (folder PID whitelist))

Image Viewer



- FAIL: riiif openseadragon (slow and required caching maintenance)
- jpegs were satisfactory in terms of resolution and zoom
- Home grown image server exposing images exposed by fedora 3 REST API
- Thumbnail in search results page
- Thumbnail strip on show page
- Single image page w/ ocr (on/off)
- Fulltext (all folder content displayed vertically)
- Thumbnails (all folder content as thumbnails)
- PDF download
- Component level AUTH/Z (thumbnail,jpg,ocr,metadata, PDF)

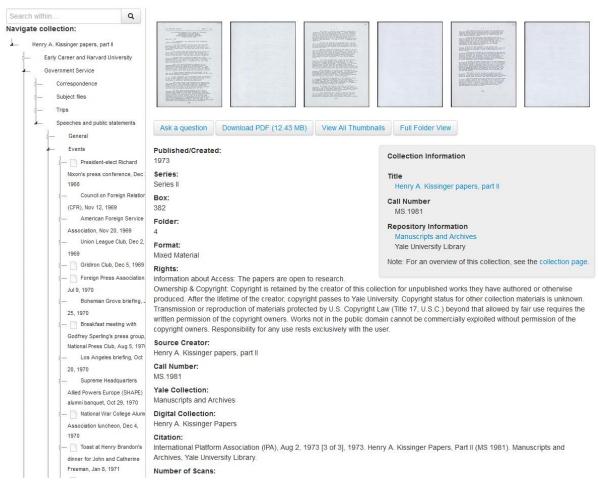
Show Page





Henry A. Kissinger papers, part II > Government Service > Speeches and public statements > Events > International Platform Association (IPA), Aug 2, 1973 3 of 3

International Platform Association (IPA), Aug 2, 1973 [3 of 3]



Single Item OCR





Henry A. Kissinger papers, part II > Government Service > Speeches and public statements > Events > International Platform Association (IPA), Aug 2, 1973 [3 of 3] > Image Viewer

International Platform Association (IPA), Aug 2, 1973 [3 of 3]- Image 1

Title

International Platform Association (IPA), Aug 2, 1973 [3 of 3]

Date 1973

Container

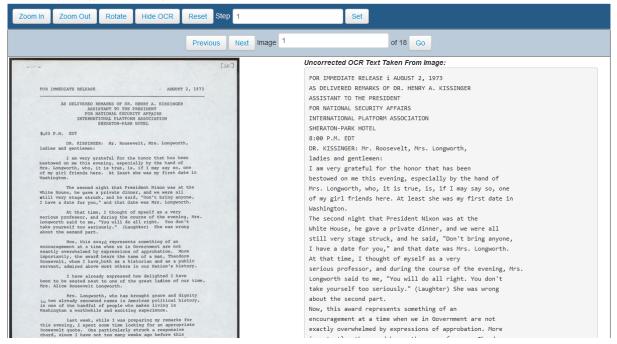
Box 382

Folder 4

Preferred Citation

International Platform Association (IPA), Aug 2, 1973 [3 of 3], 1973. Henry A. Kissinger Papers, Part II (MS 1981). Manuscripts and Archives, Yale University Library.

Download PDF (12.43 MB) | View All Thumbnails | Full Folder View





Single Item full image



Henry A. Kissinger papers, part II > Government Service > Speeches and public statements > Events > International Platform Association (IPA), Aug 2, 1973 [3 of 3] > Image Viewer

International Platform Association (IPA), Aug 2, 1973 [3 of 3]- Image 1

International Platform Association (IPA), Aug 2, 1973 [3 of 3]

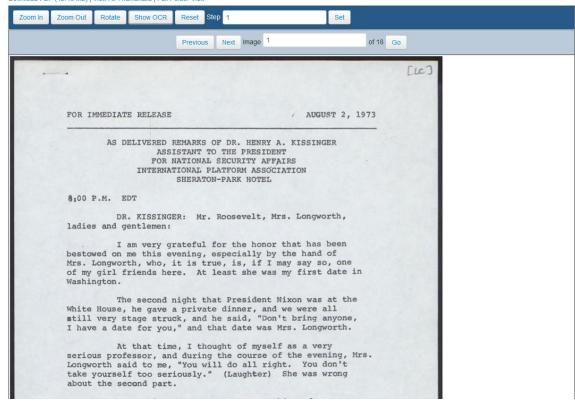
Date
1973

Container
Box 382
Folder 4

Preferred Citation

International Platform Association (IPA), Aug 2, 1973 [3 of 3], 1973. Henry A. Kissinger Papers, Part II (MS 1981). Manuscripts and Archives, Yale University Library.

Download PDF (12.43 MB) | View All Thumbnails | Full Folder View



AUTH/N SSO (openidconnect OAUTH2)



Login

Local Login	External Login		
Username	Log In with CAS Google		
Username			
Password			
Remember My Login			
Login			
Register			

Component level AUTH/Z datastream



```
-<schema>
 -<object sourceID="11775674">
      <rule type="Open Access" code="thumbnail"/>
      <rul><rule type="Open Access" code="thumbnail_dl"/>
      <rul>< rule type="Open Access" code="ipg"/>
      <rule type="Open Access" code="jpg dl"/>
    - <rule type="AD Group Restriction" code="tif">
        <item>LibraryLadybirdAdmins</item>
    - <rule type="AD Group Restriction" code="tif dl">
        <item>LibraryLadybirdAdmins</item>
      <rul><rule type="Open Access" code="jp2"/>
      <rul>< rule type="Open Access" code="jp2 dl"/></ri>
      <rul><rule type="Open Access" code="pdf"/>
      <rule type="Open Access" code="pdf dl"/>
      <rul><rule type="Open Access" code="ocr"/>
      <rule type="Open Access" code="ocr dl"/>
      <rule type="Open Access" code="undefined"/>
      <rul>< rule type="Open Access" code="undefined d1"/>
      <rul><rule type="Open Access" code="metadata"/>
      <rul>< rule type="Open Access" code="metadata_d1"/>
      <rule type="Open Access" code="premis"/>
      <rule type="Open Access" code="premis d1"/>
      <rul><rule type="Open Access" code="ead"/>
      <rule type="Open Access" code="ead dl"/>
    -<rule type="AD Group Restriction" code="tif master">
        <item>LibraryLadybirdAdmins</item>
    -<rul>-<rule type="AD Group Restriction" code="tif_master_dl">
        <item>LibraryLadybirdAdmins</item>
      </rule>
   </object>
 </schema>
```

AUTH/Z flow and restriction types



Check_user_session (verifies email,session,IP)

Check object AUTH/Z datastream(w/ PID, and component)

- OpenAccess
- Yale Only (netid or IP range)
- IP Restriction (IP on a list for object)
- NetID Restriction (netid on a list for object)
- AD Group Restriction (AD group on a list for object)
- AeonRegistration*

AeonRegistration



- 1. User granted permission to certain folders of digital content
- Upon user login, an aeon AUTH/Z endpoint is called that returns JSON with PID of whitelisted folders
- 3. This JSON content is persisted to an aeon_assets table
- 4. When AUTH/Z occurs for a component of an object with type "Aeon Registration", the aeon_assets table is checked for permissions related to user





ITS INFORMATION TECHNOLOGY SERVICES