

UAlbany Arclight Update and Outlook

Gregory Wiedeman (gwiedeman@albany.edu)

November 19, 2018

The M.E. Grenander Department of Special Collections & Archives at UAlbany has had a local Arclight instance up since March 2018. It has proven to be very stable deploying via Passenger and Nginx on a CentOS server. We have a separate Solr 7.2.1 server running on Windows Server 2012.

We successfully indexed all of our 677 EADs, which are unaltered ArchivesSpace v1.5.1 exports. We experienced two types of EAD index errors, and rectified these by improving description in ArchivesSpace. One was a long and inappropriate Chronological list in a Historical Note, more problematic than a standard <chronlist>. The other issue was with invalid date ranges, like 1956-9999, which was a problematic local practice.

Our 677 indexed EADs make up about 230,000 Solr docs. We only have 8GB of RAM allocated to Solr, and while it's minimally-usable right now, we have to do some tuning and potentially add more memory. We have been experiencing timeout errors during indexing which are rectified by optimizing the index. I expect that tuning Solr will fix this issue.

Overall, we find Arclight to be a major step forward in the display and usability of archival description. Arclight seems to make complex archival hierarchies more usable by allowing users to dive quickly into the lower levels. It leverages the strengths of archival description by presenting materials within the context of scope notes and inherited description. Still, there are persistent usability concerns as we find that users are still challenged to understand their place within a collection's hierarchy and navigate through it. With the results of our user testing, we plan to make some more minor design changes before launching, and hope to re-examine the layout and navigation to a component's contents in future iterations, hopefully based on more testing and the experiences of others.

While it is imperfect, as-is Arclight is a substantial improvement to our current XTF system in both design, flexibility, and maintainability. We find it substantially easier and more efficient to edit ERB templates than working in XTF, and the Nginx/Passenger/Rails stack with a separate Solr to be more stable and dependable than XTF and Tomcat. Despite limited technical staff, we are confident in maintaining Arclight long term. We are now currently implementing a production instance and hope to go live with Arclight as our collections access system in February 2019.

[Technical Issues](#)

[Design Issues](#)

[Unresolved Issues](#)

[Local Modifications](#)

[Planned Local Modifications](#)

[Future Priorities](#)

[Public API Access](#)

Technical Issues

Supporting Multiple Dates

Arclight does not support multiple dates in components, and will only display the first date. This is an indexing issue, as only the first date is added to `normalized_date`, which is then appended to `normalized_title`. This means it will still include other dates when faceting by dates, just not display them. Allowing multiple dates is a longstanding local practice here and is supported by DACS. Our solution to this is here: <https://github.com/sul-dlss/arclight/pull/491>



New York State Coalition for Criminal Justice Records, 1971-1986

Correctional Restitution, 1976, 1980, 1983

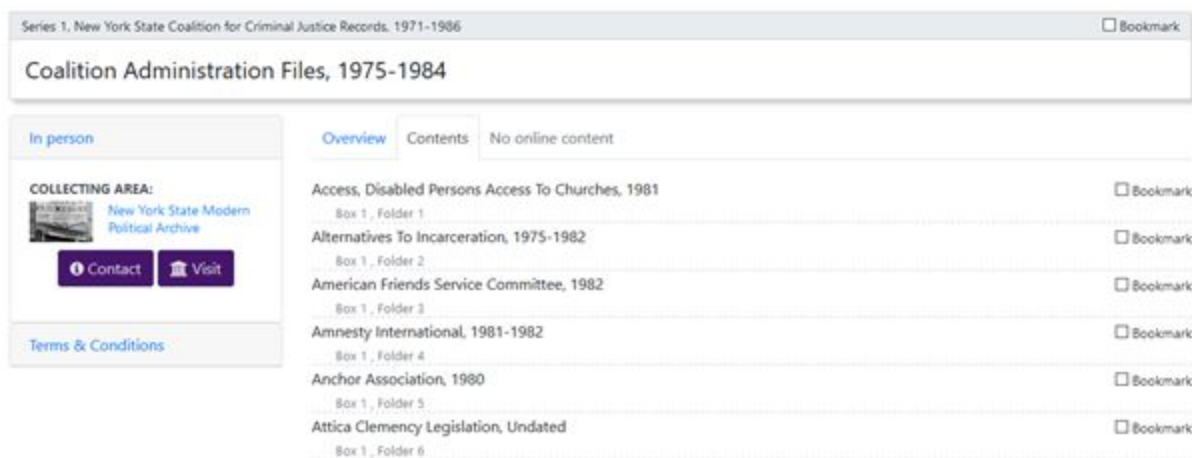
Displaying Component-Level Contents

Arclight only shows child components for collections, and lower level components (such as file series) have no access to their children without returning to the collection level or searching. It was relatively easy to implement this and performance is the same as navigating down from the collection level. Our method for this is here: <https://github.com/sul-dlss/arclight/pull/489>

We did however find that we could not easily add Google Form request links here in the long list view, as rendering this partial potentially 100s+ times was really inefficient and created performance issues:

<https://github.com/UAlbanyArchives/arclight-UAlbany/commit/5f8797a1d51f5cf867cbb80f458eb2971d29c262>

Obviously, adding component contents is likely to display very long list, which might be problematic for UX reasons. We use a fixed header to alleviate this, but this is an imperfect fix. We also had to do some custom styling to make the titles and containers smaller, and this could still use more attention.



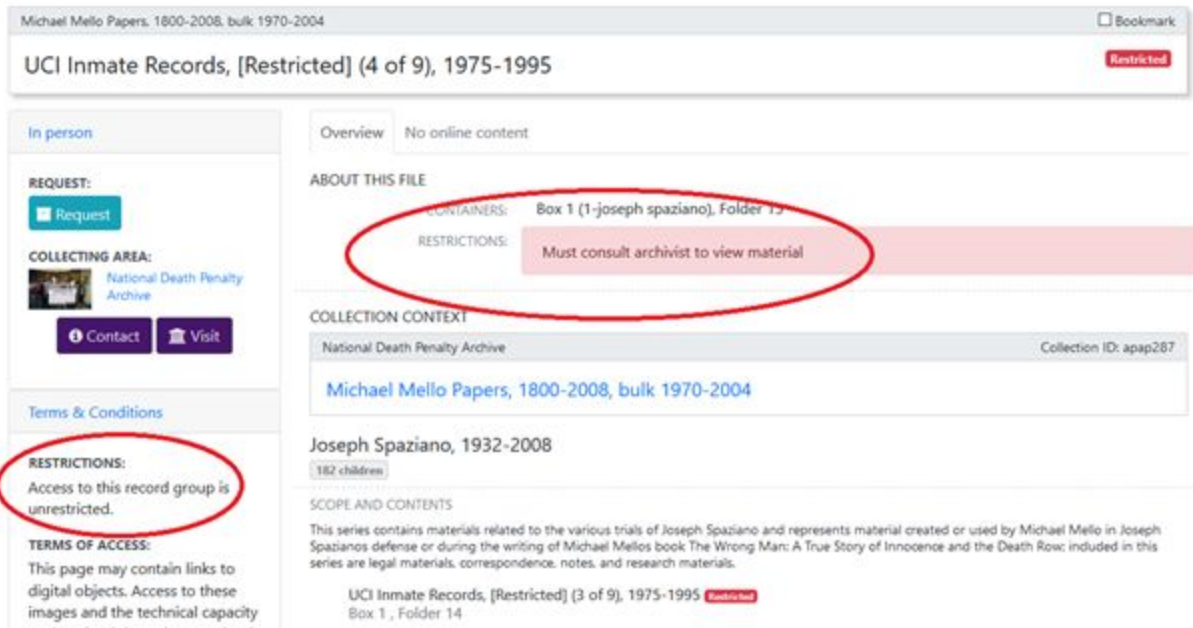
Lower Level Access Restrictions Repeats

Lower level `<accessrestrict>` tags will each show up in a parent's "Terms & Conditions" card. Thus, if a series does not have an access note, but some of its children do, it will display the restrictions of all its children.

This is an indexing issue, and fixing it potentially creates a design issue. Child access restrictions are added to a components `accessrestrict_ssm` field. The fix to `solr_ead_indexer_ext.rb` here fixes that:

<https://github.com/UAlbanyArchives/arclight-UAlbany/commit/f66382b7f6b7f1ef1793d7377317b330385ae2c8#diff-ab385a02a8028462403d51dff1f3d568>

However, this potentially creates a design issue, as a child note should override its parent, but if a parent has a note that states that its unrestricted, it still shows up in the "Terms & Conditions" card for lower level components that have an overriding access restriction note. The change to `_context_card.html.erb` in that commit overrides the parent restriction in the "Terms & Conditions" card when an access restriction note is present, but this might be problematic from a UX perspective, as the restriction in the "Terms & Conditions" card refers to parent restrictions when this is not the case.



Design Issues

Lauren Puzier, UAlbany's User Experience Librarian did some user testing our test Arclight implementation and wrote up a report here:

https://docs.google.com/document/d/1etPr-b1B13nWuDa0eXgKrYXFazZ5vYJ_NS7VmbrBxl/e/dit

Contents Tab

We find the biggest design issue to be the layout of the component pages. We find that users' eyes are drawn towards the middle of the page and often miss vital information. For collection-level pages, the title may not stand out as much as it should. We added a box-shadow to that panel, and while it may help, it is likely a "band-aid" for a broader problem.

We find that users often miss the Contents tab, so we tried replacing the existing Bootstrap Pills with Tabs. This did make the layout more transparent and users were more likely to understand that there is more content behind the "Overview" tab, but lack of color/contrast made them less likely to spot it. We think that the "Contents" tab is very important, as it allows users to navigate down the collection hierarchy. While we hope to experiment with this layout in future iterations, we do not have actionable ideas for improvement. The multi-tab view seems like the least bad option.

No Single Page Display

During their very help evaluation of Arclight, the Bentley library stated their desire for a complete collection view on a single page. We do not think this is necessary, as we think that this is a view that archivists are very familiar with, but is problematic for many users. While providing the option of a full collection view in itself would not be detrimental, conversations focusing on single-page views could potentially distract from making real usability advances that benefit everyone.

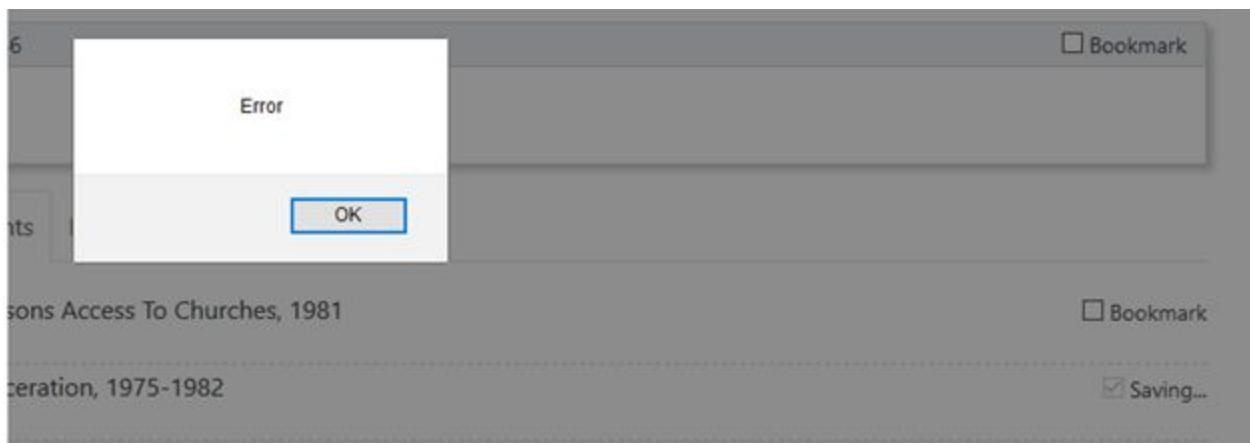
Unresolved Issues

Online Content Badges

Only collections display badges when a child has online content. This is an indexing issue that we haven't resolved yet. If a lower level component has online content, the `has_online_content_ssm` field is true, yet its parent series will be false, but its top level collection will be true.

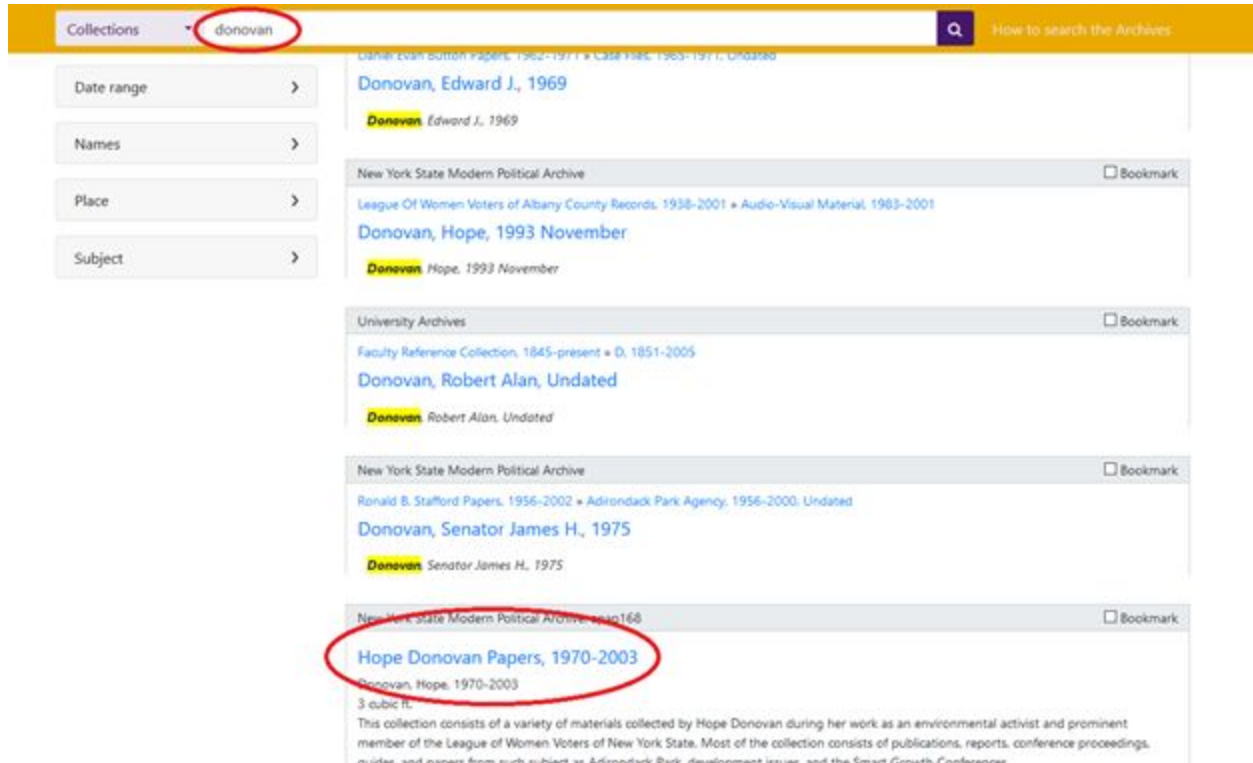
Bookmarks Checkbox

Sometimes checking or unchecking a bookmarks checkbox returns a JS alert error. This seems to be always fixed by a refresh.



Search Behavior

Collections are not prioritized as we would have expected. Searching for a known collection often returns lower level materials from other collections over the collection we expect. Its might be more of an issue for archivists than end-users.



Local Modifications

The largest changes we made were adding local headers and footers, styling changes, and some minor changes to views. Most of this section is specific to UAlbany, but I wrote them up to see if anything would be generalizable enough to incorporate in to Arclight itself.

Using Repositories for Collecting Areas

We do not have a use for the multitenancy of Arclight, yet since we have defined collecting areas, we decided to try using this feature to categorize these as repositories. It was technically very simple to alter repositories.yml for this, though we made additional changes to the view and added a tagline in the yml file. Functionally it works well and faceting by collecting area is very useful. Our testing did find that users had some confusion with collecting areas as they did comprehending other hierarchical levels.

	<p>New York State Modern Political Archive</p> <p><i>Elected officials, interest groups, and activists from New York State.</i></p> <p>The New York State Modern Political Archive (NYSMPA) was established in 1982 to document the work of individuals and private interest groups concerned with New York State public policy issues in the 20th century. Originally named the Archives of Public Affairs and Policy, the NYSMPA collects, preserves, and facilitates access to primary sources pertaining to New York State</p> <p>view more</p>	<p>281 collections</p> <p>View more</p>
	<p>National Death Penalty Archive</p> <p><i>Researchers, writers, activists, and records on capital punishment in the United States.</i></p> <p>The National Death Penalty Archive (NDPA) is a partnership between the University at Albany Libraries and the Capital Punishment Research Initiative (CPRI) at the University's School of Criminal Justice. In 1999, researchers at the School of Criminal Justice formally established the CPRI. Its overarching goals were research and education -- initiate capital punishment</p> <p>view more</p>	<p>34 collections</p> <p>View more</p>
	<p>German and Jewish Intellectual Émigré Collections</p> <p><i>Personal and professional papers of German-speaking Émigré in the social sciences, humanities, and the arts and the organizations which assisted those who fled the Nazi regime.</i></p> <p>In recognition of the serious scholarly interest in the mass migration of German speaking exiles from the Nazi regime, a German and Jewish Intellectual Émigré Collection was established in 1976 at the University at Albany, State University of New York. This growing collection has been developed since the 1970s through the efforts of the University Libraries and Professor John M.</p> <p>view more</p>	<p>87 collections</p> <p>View more</p>
	<p>Business, Literary, and Local History Manuscripts</p> <p><i>Manuscripts, records, and papers primarily related to businesses and people of New York and New England.</i></p>	<p>30 collections</p>

Indented Breadcrumbs

We thought indenting the breadcrumbs would solve the problem of really long component names, and would help users better understand the hierarchy of materials. From our testing you can see that getting users to comprehend the hierarchy is still problematic. The commit for this is here:

<https://github.com/UAlbanyArchives/arclight-UAlbany/commit/bc63a1973dfb3affdc71862ece04c8b9e4c5a096>

New York State Modern Political Archive

- » Michelle Crone Papers, 1927-2000, bulk 1982-1995
- » Activism, 1973-2000, Undated
- » General Activism, 1975-2000, Undated
- » New York State Activism, 1978-1995, Undated

[Bookmarks](#) 0 [History](#)

Subseries 1.1.2, Michelle Crone Papers, 1927-2000, bulk 1982-1995 Bookmark

New York State Activism, 1978-1995, Undated

Support for Creative Commons URIs in Usage Restrictions

We added support for Creative commons license URIs in Usage restriction notes. The language is not finalized, but this does demonstrate how easily customizable Arlight is. Arlight might consider supporting Creative Commons or RightStatements URIs like this in the future.

ACCESS AND USE

CONDITIONS GOVERNING
ACCESS:
TERMS OF USE:

Access to this collection is unrestricted unless otherwise denoted.



The materials in this collection that were **created by "System Dynamics Society"** are licenced under a [Creative Commons BY Attribution 4.0 International License](#).

This does not include materials collected or gathered by "System Dynamics Society", whose copyright is most likely retained by the original creator.

Please [Contact Us](#) for questions or additional rights.

[Terms & Conditions](#)

RESTRICTIONS:
Access to this collection is unrestricted unless otherwise denoted.

TERMS OF ACCESS:

The materials in this collection that were **created by "System Dynamics Society"** are licenced under a [Creative Commons BY Attribution 4.0 International License](#).

This does not include materials collected or gathered by "System Dynamics Society", whose copyright is most likely retained by the original creator.

Please [Contact Us](#) for questions or additional rights.

BACKGRI

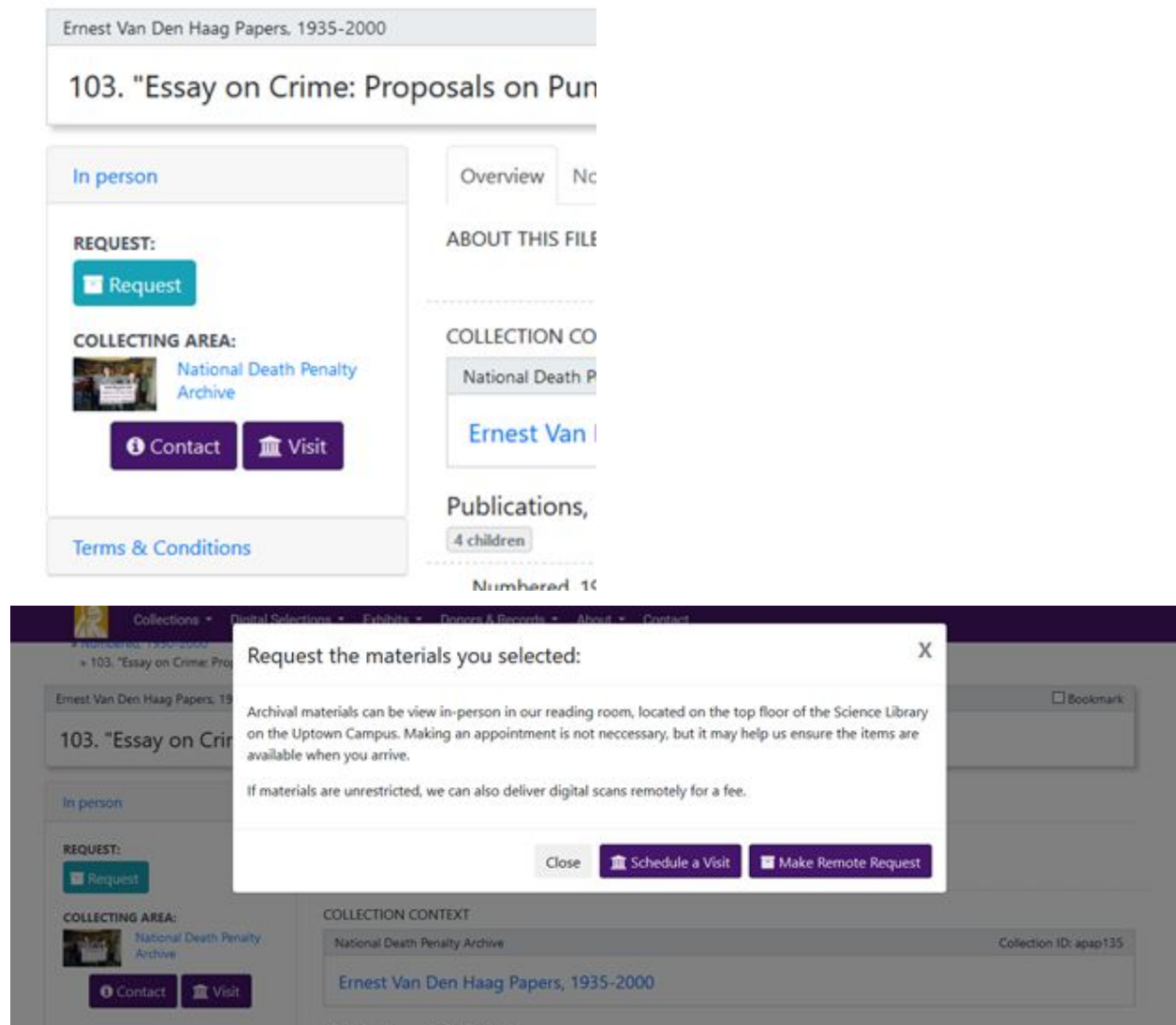
BIOGRJ

ACQUI

INDEXEC

Left side “In Person” Card

We altered the information in this area to display the collecting area and buttons for those seeking more information. For items without any children (folders, series without more detailed description, etc.) it also displays a request button that calls a modal with more information on how to view materials. Users had difficulty in discovering and understanding this information, so we will probably revisit this.



Restriction tags

We try to prioritize making access restrictions obvious and transparent to users. In the past we've relied on a number of imperfect practices, like adding "[RESTRICTED]" at the end of a

folder title, or using a collection-level access restriction note to describe specific lower level restrictions that only apply to certain parts of the collection. With some minor modifications, Arclight allows us to effectively display restrictions using lower-level notes and inheritance as archival description requires. For non-collection level items with accessrestrict_ssm information, we added Bootstrap badge-danger tags next to titles, and alter-danger in the “About This File” section. Additionally, if a non-collection level parent has an accessrestrict_ssm note, but its children do not, it inherits the parent’s restriction as an alert-warning. Some of this still relies on string matching to distinguish restrictions from notes that say “is unrestricted” or similar, so long-term this should require local standards for restrictions or potentially URIs. The code for our changes is here:

<https://github.com/UAlbanyArchives/arclight-UAlbany/commit/529d8dbbe62ef21c13041a0c86b828d4777cdc2b> and here:

<https://github.com/UAlbanyArchives/arclight-UAlbany/commit/96d414b13a86d8618234ed1879de2cc95dcf1bb9>

« Previous | 1 - 10 of 436 | Next »

Sort by relevance ▾ 10 per page ▾

New York State Modern Political Archive Bookmark

Conservative Party Of New York State Records, 1962-2004 » Clippings File, 1962-1987

Willowbrook, 1972-1973

Willowbrook, 1972-1973

New York State Modern Political Archive Bookmark

Willowbrook Review Panel Records, 1968-1981 » Audits and Surveys [RESTRICTED], 1975-1980 » Willowbrook Audits [RESTRICTED], 1975-1980

Willowbrook Audit [RESTRICTED], 1980 April Restricted

Willowbrook Audit [RESTRICTED], 1980 April

New York State Modern Political Archive Bookmark


Willowbrook Review Panel Records, 1968-1981 » Audits and Surveys [RESTRICTED], 1975-1980 » Willowbrook Audits [RESTRICTED], 1975-1980

Willowbrook Audits [RESTRICTED], 1977-1978 Restricted

Subseries 3.1, Willowbrook Review Panel Records, 1968-1981 Bookmark

Willowbrook Audits [RESTRICTED], 1975-1980 Restricted

In person

COLLECTING AREA:
 New York State Modern Political Archive
[Contact](#) [Visit](#)

[Terms & Conditions](#)

Overview Contents No online content


- Surveys, Dental Clinic [RESTRICTED], 1975 Restricted Bookmark
Box 1, Folder 1
- Surveys, Dining Room Conditions [RESTRICTED], 1975 Restricted Bookmark
Box 1, Folder 2
- Surveys, Food Service and Nutrition [RESTRICTED], 1975 Restricted Bookmark
Box 1, Folder 3
- Surveys, Policy and Procedure Manual [RESTRICTED], 1975 Restricted Bookmark
Box 1, Folder 4
- Surveys, New Personal Service Items [RESTRICTED], 1975 Restricted Bookmark
Box 1, Folder 5
- Surveys, Deficiencies and Plans of Correction [RESTRICTED], 1975-1976 Restricted Bookmark
Box 1, Folder 6
- Education Reinstatement Audit [RESTRICTED], 1975-1976 Restricted Bookmark
Box 1, Folder 7

Willowbrook Review Panel Records, 1968-1981 Bookmark

Surveys, Food Service and Nutrition [RESTRICTED], 1975 Restricted

In person

REQUEST:
[Request](#)

COLLECTING AREA:
 New York State Modern Political Archive
[Contact](#) [Visit](#)

[Terms & Conditions](#)

Overview No online content

ABOUT THIS FILE

CONTAINERS: Box 1 (3-audits and surveys [restricted]), Folder 3

RESTRICTIONS: Must consult archivist before viewing this material.

COLLECTION CONTEXT

New York State Modern Political Archive Collection ID: apap127

[Willowbrook Review Panel Records, 1968-1981, bulk 1975-1981](#)

Office of the President Records, 1827 - 2017 July 17 Bookmark

Committee Appointments and Related Materials, 1947-1962

In person

REQUEST:
[Request](#)

COLLECTING AREA:
 University Archives
[Contact](#) [Visit](#)

[Terms & Conditions](#)

Overview No online content

ABOUT THIS FILE

CONTAINERS: Box 11, Folder 5

PARENT RESTRICTIONS: This file series is semi-processed and may contain some records that must be withheld due to public records laws. Please consult an archivist for more information.

COLLECTION CONTEXT

University Archives Collection ID: ua000

[Office of the President Records, 1827 - 2017 July 17](#) [view content](#)

Planned Local Modifications

DAO Integration

For digital objects, for at least the medium term future we are planning on presenting digital objects in another system, with only Links from the Online Content tab. Since the Hyrax repository we are implementing exposes JSON with an item's data and thumbnail, we are planning on replacing the link in the Online Content tab that's in Arclight now with a stock thumbnail, and have a CORS request replace it with data and the thumbnail from Hyrax.

Future Priorities

Analytics

We utilize analytics data to inform processing decisions. Currently we are planning to use an existing proprietary system to parse the logs from Nginx. It would be ideal if Arclight had a view that listed and graphed the most heavily utilized pages.

Aeon/Requests

We do not plan to implement Aeon, but modifying Arclight so it sends an Aeon-friendly post request would be useful. Right now, we customized the Bookmarks page with a "Schedule a Visit" button that sends a GET request that partially completes a Drupal form with all of the IDs on that page as params. This only minimally works due to URL limits and items on other pages, but it is as far as we could get without modifying the controller. If Arclight it setup to deliver a POST request with this data in a format that Aeon uses, we could set up our Drupal form to manage that as well.

Bookmarks

1 - 3 of 3

Sort by relevance ▾

10 per page ▾



Clear Bookmarks

Schedule a Visit to view all of the material on this page:

Schedule a Visit

National Death Penalty Archive	<input checked="" type="checkbox"/> In Bookmarks
Virginians for Alternatives to the Death Penalty, 1937-2006 » Subject Files, 1937-2007	
1000th Execution, 2005	
National Death Penalty Archive	<input checked="" type="checkbox"/> In Bookmarks
The Abraham Bonowitz Papers, 1977-2015 » Communications, 1983-2013	
10.10 is A World Day, 10 years of World Coalition against the Death Penalty, 2012	
National Death Penalty Archive	<input checked="" type="checkbox"/> In Bookmarks
Ernest Van Den Haag Papers, 1935-2000 » Publications, 1949-2000 » Numbered, 1950-2000	
104. "History as Factualized Fiction," In "Philosophy and History", edited by Sidney Hook, New York Press, 1963	

Public API Access

We think it is worthwhile to mention that one of the features we found most useful in Arclight is that it functions as a completely open API. Arclight serves data in JSON as well as a human-readable view by merely appending “?format=json” to the end of each URL. This allows for computational use for archival description and is extremely useful for easy system integration.

Computational Access Potential

Arclight’s public API has the potential to allow end-users to parse archival description computationally at scale. Not only can users request collection-level description in JSON, but this call (`/catalog?f[parent_ssi][]=$ID&view=hierarchy&format=json`) will allow them to iterate over child components. With the addition of a call to return a list of collection-level URIs, users could potentially iterate over all of the archival description in Arclight. Providing this access separately from the ArchivesSpace API is also good practice, as Arclight should only contain publically available description, and does not require authentication. We expect this feature of Arclight to become progressively more useful as archivists potentially incorporate licenses for description (<https://github.com/saa-ts-dacs/dacs/pull/14>), and need to provide computational access to their collections as data within the context of archival description.

Systems Integration

Our planned digital repository might serve as a small example of how useful Arclight can be for system integration. We plan to use Cross-Origin Resource Sharing (CORS) to access archival description from Arclight with client side JavaScript. This is useful for archivists uploading digitized content with existing archival description, as entering a component's ID can retrieve titles, dates, parent description, and other useful information from Arclight without any authentication. We plan to include IDs of parents, and when objects are rendered, another JavaScript call can return that parent's title, dates, and scope and content notes directly from Arclight. There is a tremendous potential for Arclight to allow archivists to utilize description in new and powerful ways.