

Implementing Hydra at the Digital Library of Georgia

Michael Kanning / mak@uga.edu



What is the Digital Library of Georgia?

- The Digital Library of Georgia (DLG) is a gateway to Georgia's history and culture found in digitized books, manuscripts, photographs, government documents, newspapers, maps, audio, video, and other resources.
- An initiative of GALILEO, Georgia's statewide library cooperative, housed at the University of Georgia libraries.
- A **service hub** for the Digital Public Library of America (DPLA).
- Made up of over 415,000 items across over 500 collections from more than 100 institutions.
- A website enabling searching of record metadata and browsing of collections, with links to the original items at their home institution's website.

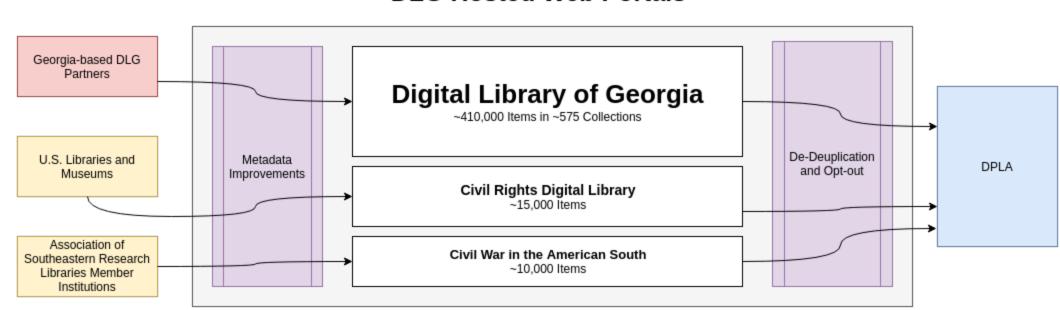
Current Infrastructure

- Metadata is stored in an internally-developed schemaless database called "Ultimate".
- Static HTML pages for records are generated nightly by a set of scripts and processes called "Sitegen".
- Search Indexes are regenerated nightly.

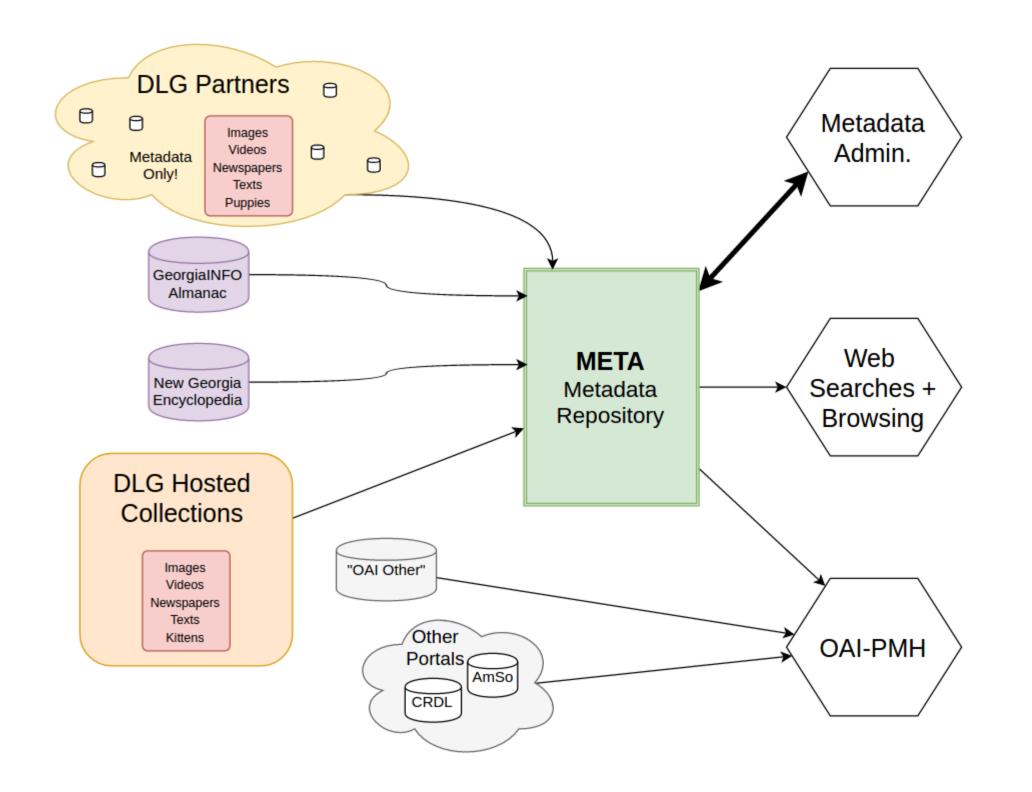
Growing Pains

- The current stack has been in place for over ten years.
- Learning curve for new developers working on the DLG is steep.
- Processing time for generating static HTML files is increasing with the size of the DLG.
- Loading facets and performing searches takes a long time and can impact UX.
- Records cannot be added or updated in real time, only after pages are regenerated.

DLG Hosted Web Portals



The Metadata Repository



Rescue Phase

- Currently developing a Blacklight-based site to replace the metadata store that enables DLG administration, user searches and browsing, and eventually the DPLA feed.
- Metadata is stored in PostgresDB and indexed to Solr using Sunspot.
- This will help the DLG be sustainable as it continues to grow, while moving in the direction of longer-term goals, while also introducing new features for both public users and DLG staff.
- Improved UX, speed, and on-demand additions and updates.
- Improved metadata (spatial indexing, QDC fields, etc.)
- Fully exploit Blacklight plug-in ecosystem (Gallery, Maps, etc.).
- Improve staff processes related to adding new collections and updating existing metadata.
- Utilize well-supported, mature open-source tools to help encourage sustainability and future development.

Special Snowflake?

- The DLG organization manages 3 distinctive web portals: The DLG, Civil Rights Digital Library and the Civil War in the American South. Records can be discoverable through one or many of these portals, may or may not be pushed to the DPLA, and may or may not be from a source in the state of Georgia.
- As part of its role as a DPLA Service Hub, the DLG aggregates, manages and feeds to the DPLA content that comes from other southeastern states, but is not contained within any of the DLG web portal ["Other OAI"].
- Primarily, the DLG is responsible for the aggregation and management of metadata. Only a small portion of records in the DLG have their corresponding DOs managed by the DLG. Access to original DOs can sometimes be an issue, making things like deriviative generation a hurdle.

Future Directions

- Looking to implement Hydra as a more modern, sustainable and community-supported platform.
- Implement new technologies and standards to increase visibility, accessibility and interoperability of our metadata and digital objects.
- Centralize repositories for video, audio, text, newspapers, etc. into one.
- Play nice with statewide partner systems, including other IRs and DCs.

