Background
Samvera/Hyrax at GWU

- 2015 - Built "GW ScholarSpace" IR on Sufia 6/7
- 2017 - Rebuilt on Hyrax, with **GwWork** and **GwEtd** work types
- Built batch loader
- As of today, 10K+ works in GW ScholarSpace (Hyrax 2.8.0)

https://scholarspace.library.gwu.edu/
ISIS Files - Documents

- 15,000 pages
- ~ 1500 files (original TIFFs)
- Initial batch: 62 documents

Workflow: Digitize

Sort/Contextualize (Review by Fellows) →
Translate →
Create Metadata →
Redact →
→ PDF + accessibility fix → Ingest
Hyrax as a platform
Choosing a platform

● Evaluated:
  ○ Hyrax
  ○ Drupal
  ○ WordPress
  ○ Murkutu
Important Discriminators

- Out-of-the-box feature set
  - "Works" with files + customizable metadata
  - Search/Indexing of text documents
  - Image viewing*

- Accessibility
- Design
- Preservation
- Security
Other Considerations

- In-house skills/experience
- Effort level
- If Hyrax: v2.x or v3.0.0-rc1
  - timeline (started Sept. 2019, release June 2020)
  - appearance/accessibility/security
- Community (Slack, etc.)
Demo
https://isisfiles.gwu.edu/
Public Reception
'ISIS Files' launch: Thousands of documents reveal the terror group's inner workings

PAUL SZOLLOSY | JUL 3, 2020 11:20 AM EDT

The long-awaited 'ISIS Files' have launched online at the George Washington University, offering researchers and others a window into the inner workings of the terror group that called itself "The Islamic State."

The New York Times and GWU announced a partnership in 2018 to digitize, translate, and analyze more than 15,000 pages of internal ISIS documents. Now, these documents are beginning to appear on the ISIS Files website, which is based around themes such as intelligence, war strategy, and cultural norms, and releases online files.

GWU has released the first batch of its "ISIS" series of bureaucratic documents captured from the Islamic State's Iraqi stronghold in Mosul.

George Washington University has released the first batch of its "ISIS" series of bureaucratic documents captured from the Islamic State's Iraqi stronghold in Mosul.

The New York Times has released these documents to the world in an effort to help explain how the Islamic State operated and how its members communicated with each other.

GWU's Program on Extremism, which is working with the Times and the National Security Archive, has been working to digitize and translate the documents and make them available to the public. The first collections released by GWU focus on ISIS ideology and real estate and taxation.

The Program on Extremism worked with the Times and GWU's Libraries and Academic Innovation division to digitize, translate, and analyze the documents. They then released a new website with the collection already made available to the public. The first collections released by GWU focus on ISIS ideology and real estate and taxation.
Technical workarounds, enhancements, future challenges
Deployment and Security

- AWS
- GW full security review + remediations
- Other risk reductions :-(

isisfiles-hyrax  Private

CSS 0 0 30 0 Updated on Jun 26
Customizations

- Open PDF in-browser
- Menu enhancements
- Cookie popup (GDPR compliance)
- Accessibility & Security fixes
- reCAPTCHA - authentication and contact form
- UI simplification
Future directions
Next...

- More content!
- Hyrax 3.0
- Use more features: IIIF, etc.
- Rethink i18n of metadata values
- Add UI languages (Arabic, maybe?)
- Batch loading
- Technical debt
- Community debt
Takeaways
Uniquenesses

- Provenance of documents
- Sensitive nature of docs and metadata - ethical questions
- Workflow around docs
- Appearance & features
Critical Success Factors

- Samvera community :-) especially on Slack
- Hyrax feature set and maturity
- Prior experience with Hyrax

*Hyrax was the easy part!*
Thanks!

Dan Kerchner
kerchner@gwu.edu