

# The ISIS Files

*a Hyrax implementation*

August 14, 2020

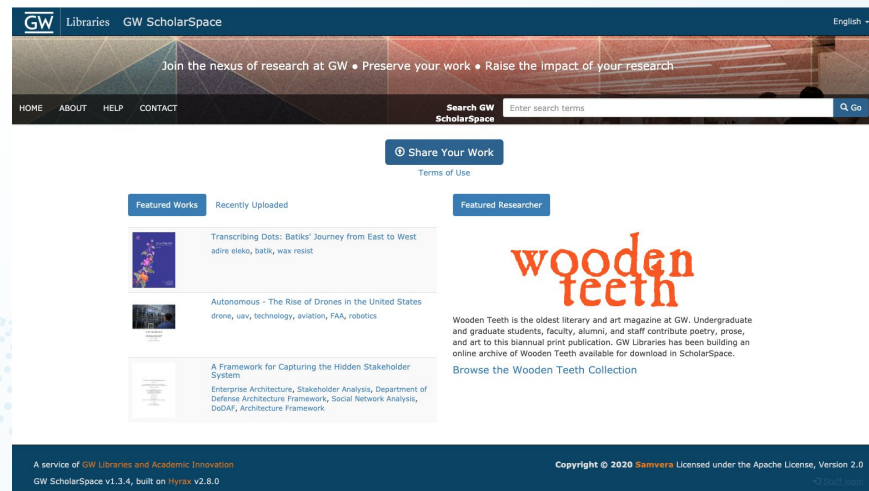
Dan Kerchner, Senior Software Developer & Librarian  
George Washington University

# 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/>







Libraries &  
Academic Innovation

# Program on Extremism

THE GEORGE WASHINGTON UNIVERSITY

**The New York Times**

THE  
ANDREW W.

**MELLON**

FOUNDATION

# 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

The background of the slide is a solid blue color. Overlaid on this background are several wavy, horizontal lines composed of small, dark blue dots. These lines create a sense of motion and depth, with some lines appearing more prominent than others. The lines start from the left edge and extend towards the right, with some curving upwards and others downwards.

# 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

The background is a solid blue color. Overlaid on this background are several wavy, dotted lines in a darker shade of blue. These lines form a series of peaks and valleys, resembling a stylized wave or a series of connected dots that create a fluid, organic shape. The dots are small and closely spaced, giving the waves a textured appearance.

## 'ISIS Files' launch: Thousands of documents reveal the terror group's inner workings

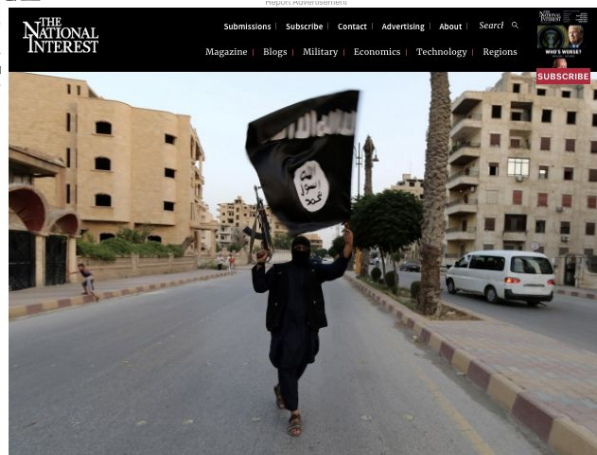
PAUL SZOLDRA JUL 3, 2020 11:36 AM EDT



(YouTube/ISIS via Business Insider)

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, those documents are beginning to appear on the [ISIS files website](#), which is based around themes such as ideology, war spoils, agriculture programs, and religious police files.



June 29, 2020 | Topic: Security | Region: Middle East | Blog Brand: Middle East Watch | Tags: ISIS, Islamic State, History, Middle East, Syria, Iraq

## What Was ISIS's Real Estate Market Like? GW University Releases "The ISIS Files"

Lot number 128. Previous owner: apostate.

by Matthew Pettit

Report Advertisement

**G**eorge Washington University has released the first batch of its "ISIS" trove of bureaucratic documents captured from the Islamic State during its offensive in Mosul.

Masters in Global Policy LEARN MORE >

## SETTING THE RECORDS STRAIGHT IN IRAQ

MICHAEL P. BRILL

COMMENTARY

JULY 17, 2020



Homeland Security News Wire

## Homeland Security News Wire

BIOMETRICS BORDER/IMMIG. BUSINESS CYBERSECURITY  
PUBLIC SAFETY PUBLIC HEALTH REGIONAL SCI-TECH

## Terrorism GW Launches ISIS Files Digital Repository

Published 3 July 2020



**The George Washington University on Monday launched its ISIS Files repository. The virtual public repository features a selection of the 15,000 digitized pages from the documents collected in Iraq by New York Times journalist Rukmini Callimachi and a team of Iraqi translators.**

The George Washington University on Monday launched its **ISIS Files repository**. The virtual public repository features a selection of the 15,000 digitized pages from the documents collected in Iraq by New York Times journalist Rukmini Callimachi and a team of Iraqi translators. GW's **Program on Extremism** formed a partnership with the New York Times in 2018 to make these documents available to the public. The first collections released by GW focus on ISIS ideology and real estate and taxation.

The Program on Extremism worked with the Times and GW's Libraries and Academic Innovation division to digitize, translate, and analyze the documents. They then were posted on a new website



1 4 11

Show this thread



1 2 7

Show this thread





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

The background of the slide is a solid blue color. Overlaid on this background are several wavy, horizontal lines composed of small, dark blue dots. These lines create a sense of motion and depth, with some lines appearing closer and more densely packed than others, creating a 3D effect. The lines flow from the left side of the slide towards the right, with some curving upwards and others downwards.

# 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](mailto:kerchner@gwu.edu)*