

**ARCHIVE-IT**

**Partner Meeting 2019 | Austin, TX**

Archive-It  
13th Partner Meeting  
2006-2019  
[teen years achieved]



PARAMOUNT PICTURES PRESENTS THIRTY THE 13TH A SEAN S. CUNNINGHAM FILM STARRING DREW PALMER ANDRIENNE YONG HARRY GREGORY LAURIE CHATHAM MARK HELESON JEANINE TAYLOR BOBBI WITRODA  
KEVIN BACON DIRECTOR OF PHOTOGRAPHY GARY ABRAMS MUSIC BY HARRY MARAFREDDI ASSOCIATE PRODUCER STEPHEN MOWER EXECUTIVE PRODUCER ALVIN BELLER WRITTEN BY VICTOR MILLER  
PRODUCED AND DIRECTED BY SEAN S. CUNNINGHAM A GEORGETOWN PRODUCTIONS INC. PRODUCTION © 2006 BY PARAMOUNT PICTURES (SMP) ALL RIGHTS RESERVED A PARAMOUNT RELEASE

**R** RESTRICTED  
PARENTS STRONGLY CAUTIONED  
SOME MATERIAL MAY BE INAPPROPRIATE FOR CHILDREN UNDER 17



880073

Friday the 13th

# Landscape Thoughts

## Wide

- Sustainability Questions
- Economic Forecasts
- Cloud, Hosted, E2E

## Focused:

- “Secondary” Outputs
- Degrees of Distribution
- Capture/Replay/Affiliated



# Staff Updates & Shout Outs

## NEW AIT TEAM MEMBERS

- Pete Mitchell ([pete@archive.org](mailto:pete@archive.org)), Technical Support Specialist
- Madeline Carruthers ([madeline@archive.org](mailto:madeline@archive.org)), Partner Coordinator

## OTHER NOTABLES

- Here ye: James, Jillian, Karl, Kyrie, Lori, Sylvie
- Wish you were here: Adam, Barbara, Mouse, Neil, Noah, Connor, Helge
- Gone but not forgotten: Maria, Mary



# Community Updates New Partners (47)

Cape Ann Museum  
WGBH  
Museum of Contemporary Art  
Scottish Poetry Library  
Corning Museum of Glass  
University of Pennsylvania  
Greenpeace International  
Joshua Meyer-Gutbrod (Digital Campaigns Project)  
Institut français du Proche-Orient  
North American Coordinating Council on Japanese Library Resources  
Senator John Heinz History Center  
99 Antennas  
Fashion Institution of Technology  
Norwich University  
News Measures Research Project (Branchhead Consulting)

University of Colorado - Boulder  
Congresswoman Niki Tsongas  
Pepperdine University  
Indiana University Purdue University Indianapolis (IUPUI)  
Iowa University SILS  
California State Library  
MacEwan University  
Arizona State University  
New York City Department of Records  
Marine Biological Laboratory  
Worcester Polytechnic Institute  
Expatriate Archive Centre  
Los Angeles County Metropolitan Transportation Authority  
University of Georgia  
Publications Office of the European Union  
Humboldt State University

Phillips Academy - Andover  
East Carolina University  
University of Adelaide  
Heinrich-Heine-Universität Düsseldorf  
Hyetert  
Yale University Libraries IT  
Center for Cooperative Media  
California Digital Library  
Abilene Christian University  
University of Northern Colorado  
North Park University  
Art Gallery of Ontario  
Wichita State University  
Algoma University  
Beckman Research Institute of City of Hope  
LIBER Europe

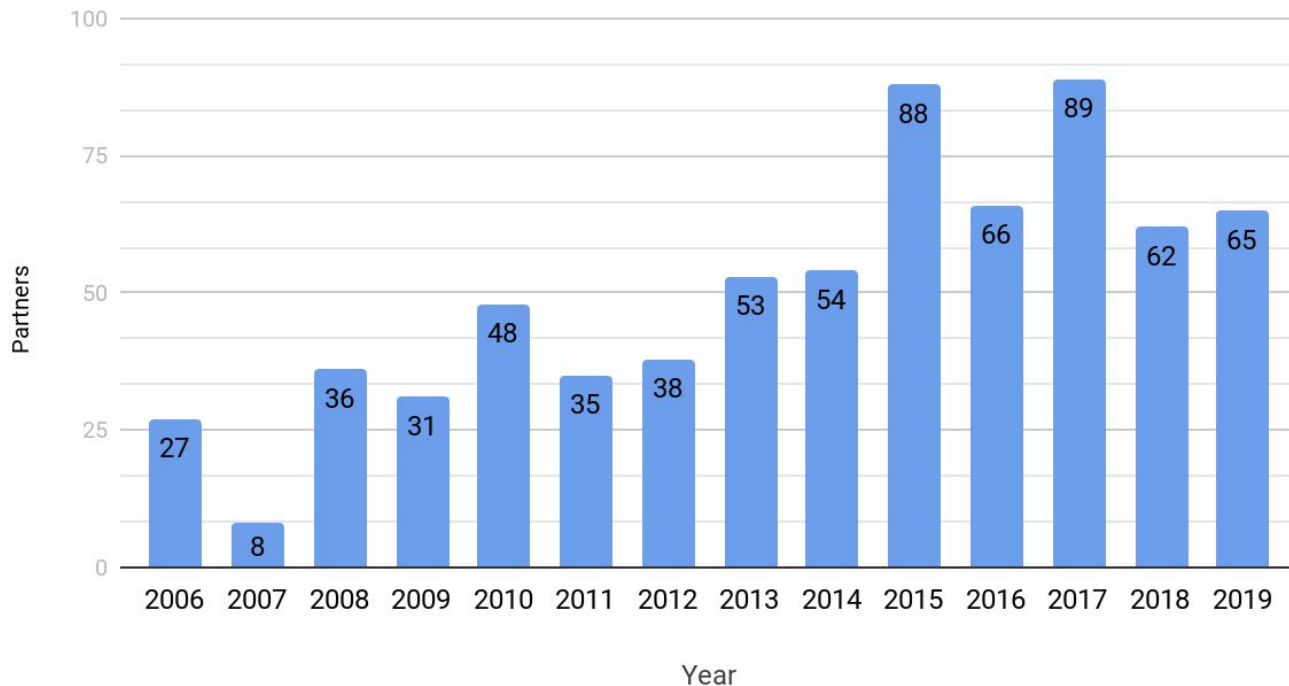
**2006-today: 664 Archive-It partners | 29 in 2019 so far**

#aitTX | @archiveitorg



# AIT Community Updates

Partners Per Year

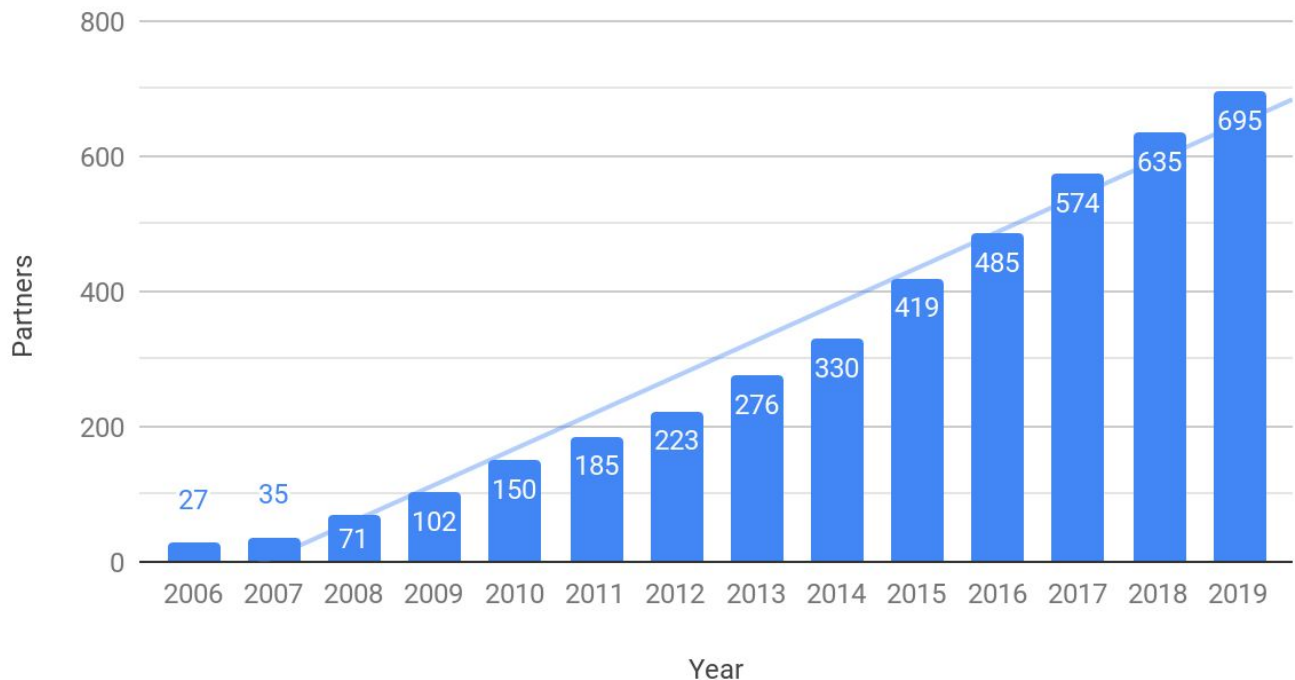


#aitTX | @archiveitorg



# AIT Community Updates

Total Partners by Year

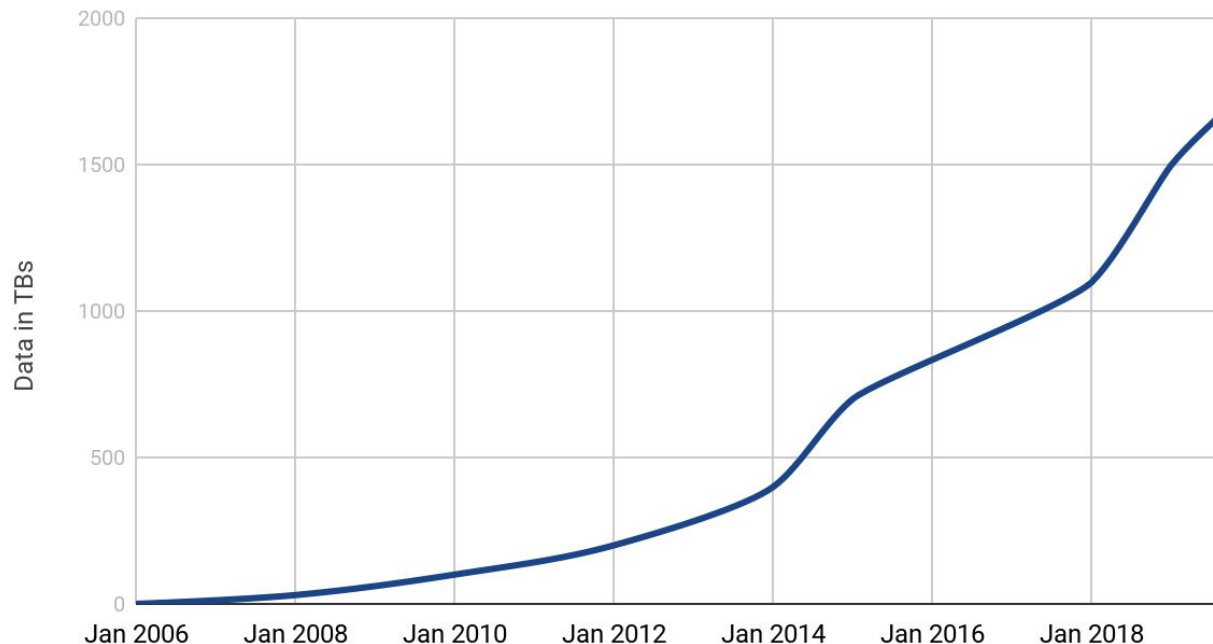


#aitTX | @archiveitorg



# AIT Community Updates

Archive-It Total Size of All WARC's



#aitTX | @archiveitorg



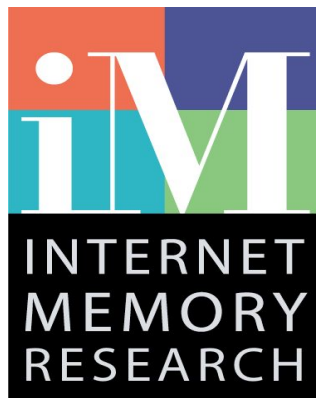


# AIT stats-a-thon

- Total URLs: 40 Billion URLs preserved
  - +33% from 30B 1Y ago
- Total Data: 1660 TB
  - +34% from 1235 TB 1Y ago
- Total Collections: 11,800 public+private
  - +16% from 10,100 1Y ago
- Total Seeds: 1.76 Million
  - +15% from 1.53M 1Y ago
- Total Crawls: 488,000
  - +38% from 388,000 1Y ago



# AIT Community Updates



## Consortial & Other Partnerships

- Ongoing consortia: Consortia Canada, LIPA, CDL, KAIC, Yale, Stanford, BSZ, Harvard, Five Colleges
- Transitioned IMF partners - additional international and managed partnerships
- Discussions with ArchivesSpace, DuraSpace, and more about service integrations/partnerships

#aitTX | @archiveitorg



# Grants & Strategic Activities



“Continuing Education for Web Archiving”  
- Workshops and training for technical tools



“Global Events & Trends Archive” (GETAR)  
- Events archiving, data mining, K-12 education



“Building Infrastructure and Integrations for Open Data Preservation and Access”  
- Data sharing, access and preservation



“Broader Web-Scale Provision of Parallel Corpora for European Languages”  
- Machine language data mining



Society of California Archivists #aitTX | @archiveitorg



# Grants & Strategic Activities



HOME ABOUT PARTICIPANTS CURRICULUM NEWS

## Community Webs

Empowering Public Libraries to Create Community History  
Web Archives



Since 2017 the Cohort of 27 Public Libraries:

- Created over 200 community focused collections
- Collected over 27TB of data
- Spoke at over 15 different conferences

At the end of the grant we have:

- A strong group of leaders in the field
- A free online course on web archiving geared toward public libraries, developed with the input of the cohort

<https://communitywebs.archive-it.org/>

#aitTX | @archiveitorg



# Grants & Strategic Activities

About Archive-It   Blog   Projects   Conferences

## Advancing Art Libraries and Curated Web Archives

A National Forum

The Institute of Museum and Library Services (IMLS) awarded a collaborative grant to the [New York Art Resources Consortium](#) and the Internet Archive's Archive-It staff to host a national forum event, along with associated workshops and stakeholder meetings, to catalyze collaboration among art libraries in the stewardship of historically valuable art-related materials published on the web. This National Leadership Grant in the Curating Collections program category to conduct a National Forum and affiliated meetings builds on NYARC's and Archive-It's work together expanding web archiving amongst art and museum libraries and archives.



The image shows three logos: the Archive-It logo (a blue triangle with a white 'i' and the text 'ARCHIVE-IT'), the NYARC logo (the text 'nyarc' in a light blue font), and the logo for the Institute of Museum and Library Services (a green starburst pattern followed by the text 'INSTITUTE of Museum and Library SERVICES').

- Symposium and half-day workshops among 50 librarians, archivists, and technologists.
- A follow-up meeting for roadmapping with stakeholders..
- Roadmap of a collaborative collecting model for art materials on the web

<https://archive-it.org/blog/learn-more/art-libraries/>

#aitTX | @archiveitorg



# AIT Special & Affiliate Projects

- Edu: 10 MLIS program partners
- Training & Support:
  - Special Topics Webinar Trainings
- Partner meet(ings)(ups):
  - New models for partner engagement
- Professional Participation:
  - IIPC Training WG & SC Member
  - DPC Web Archiving & Preservation Working Group
  - 30+ conference presentations/workshops/etc



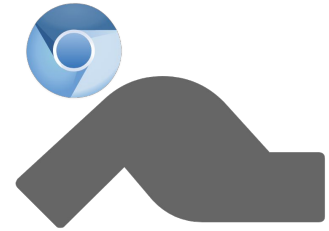
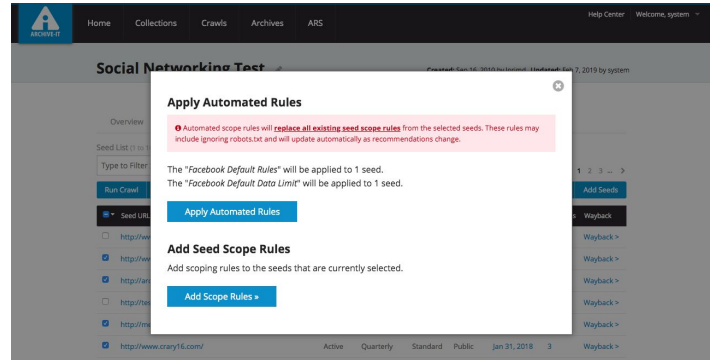
# AIT Special & Affiliate Projects

- Pitch a Collection Contest Winners:
  - Five Collections from 3 partner orgs
  - Topics include: Beer and Brewing, Mission History, and San Diego Refugees
  - Collected 92 seeds and 120GB data since Nov 2018
- Examples of IA-Sponsored event collections
  - Puerto Rican Protests 2019, UPR and UConn
  - Malibu Community Woolsey Fire, Pepperdine University
  - 2018 Mexican Government Web Archive - El Colegio de Mexico



# Archive-It New Features

- Brozzler for all!
- Bulk seed scoping rules
- Per-seed WARC6s for all crawls
- Custom user agents
- Redirection Service
- Token-based authentication for data download





# Archive-It backend changes

## Completed:

- Operational Resiliency
- Postgres
- Phase 1 - Renewals Streamlining

## In Progress/upcoming:

- OutbackCDX - Faster wayback!
- Crawlboss - additional crawl customization/scalability
- Rules Engine - more flexibility in access restrictions/rewrites



# AIT Current & Planned Development



- Partner API - Advanced training Aug 21!
- Crawl management/video replay redesign
- Python Wayback
- Partner Wayback Controls
- Public Site redesign
- Renewals in web application
- Data budget enforcement in UI



# Final Thoughts

- AIT Application Extensibility
- Beyond Libraries/Archives
- Beyond Crawling
- Interoperability/Integration
- Promotion/Use



#aitTX | @archiveitorg