Archive-It collection on “Occupy Movement 2011/2012”
Archiving Web Content
Archive-It

- Web archiving service first deployed at the Internet Archive in 2006
- In 2007, started to collect “at risk” web content on spontaneous events that occur in the US and the world.
- Web content needs to be documented and archived for historical and cultural purposes
- Curators use Archive-It to add websites, metadata and set up automated crawls
- Digital Collections are all publicly available
“Occupy Movement 2011/2012” collection

Collection is publicly available at:
http://archive-it.org/collections/2950

Organized into Website Groups:
  – Blogs, International, News Sites and Articles, Other Sites, Social Media
“Occupy Movement 2011/2012” collection

Collection was created Nov 30, 2011:

• Web sites selections were just filtering in.

• Working with “Activist Archivists”, groups from NYU, OWS, and other individuals.

• Named the collection “Occupy Movement”, to include content from around the world.

• Staff at the the Internet Archive created a blog post to generate visibility and seed submissions for the collection.
Current Crawling Activity on “Occupy Movement 2011/2012”

- Have included 770 websites to be crawled
- Captured 17 million documents
- Archived 637 gigabytes of data
- Crawling daily, weekly, and monthly
Managing the “Occupy Movement 2011/2012” Collection

• Seed submissions:
  – bulk and single website submission from curators of content and other individuals
  – scraped and included websites from community generated lists (e.g. “We All Occupy”, “Occupy Feeds”)

• Monitor and check crawls:
  – looking for crawler traps
  – adding crawling rules to capture content where needed
Global Content

“Occupy Clermont-Ferrand”
France

“Mi smo 99%”
Serbia
• http://wayback.archive-it.org/2950/20120210032944/http://occupyserbia.org/
Unique Content

We, the undersigned writers and all who will join us, support Occupy Wall Street and the Occupy Movement around the world.

“Occupy Writers”
Static websites with unique content that may not be maintained

News & Blogs

News Articles:
Article about arrest of Occupy Protestors


Special Interest Groups:
Article about destruction of OWS “People’s Library”

Images & Video

Photo Albums of Events:
“Occupy Long Beach October 18 2011”

“25 best Occupy photos of 2011”

Citizen Video: Pepper spray victim in Birmingham
Thank you!
The Archive-It Team
www.archiveit.org
graham@archive.org