



The Rise of Digital Preservation Consortia

Thib Guicherd-Callin
LOCKSS Program
Stanford University Libraries
thib@cs.stanford.edu

Seminário Internacional de Preservação Digital
May 7-9, 2014
Brasília, Brazil



LOCKSS Program

- “Lots Of Copies Keep Stuff Safe”
- Started in 1998
- Flexible, distributed, fault-tolerant digital preservation software
- Implement paper business model in Web world
 - Ownership of copies
 - Massive replication



LOCKSS Program

- 150 research libraries worldwide
- Self-sustained by libraries (LOCKSS Alliance)
- Library staff (Stanford University, Internet Archive)
- Format-agnostic
- Sourced from Web harvest, file transfer, OAI-PMH, proprietary publisher APIs
- Research papers
- 2014 LITA/Library Hi Tech Award



LOCKSS Program

- LOCKSS preservation primer
 - LOCKSS plugins
 - Autonomous harvesting
 - LCAP peer-to-peer audit-and-repair protocol
 - Filtering and normalization
 - Metadata extraction
 - Proxy integration
 - Content server
 - Link resolver integration
 - etc.



Then and Now

- Focus on e-journals
- Mostly large research libraries
- Static or dynamic Web
- Punctual access



Then and Now

- Rapid evolution of peer-reviewed publishing
 - E-books, government documents, digitized collections, electronic theses, institutional repositories...
 - Data sets, workflows, post-publication review, social media...
- Regional, national, special-interest digital preservation communities
- Interactive Web
- Cross-sectional access



Global Initiatives

- Global LOCKSS Network (GLN)
 - Individual libraries take responsibility for custody of content
 - Loosely-coupled relationships among libraries
 - Run by Stanford University
 - Mostly for preservation and post-cancellation access of e-journals and e-books
 - Over 500 participating publishers





Global Initiatives

- MetaArchive Cooperative
 - Non-profit run by Educopia Institute
 - Global membership of cultural memory organizations
 - Preserving high-value, locally-created digital materials
 - ~40 terabytes





Global Initiatives

- CLOCKSS Archive
 - Dark archive
 - 24-member board of directors (half publishers, half libraries)
 - 12-node preservation network (worldwide)
 - When preserved content is unavailable from any source, triggered out of dark archive and made available to everyone for free





Regional and National Initiatives

- Alabama Digital Preservation Network (ADPN)
 - Consortium of Alabama institutions
 - Historic archival materials, image collections
 - Model of governance for small PLNs



THE ALABAMA DIGITAL PRESERVATION NETWORK
PRESERVING ALABAMA'S DIGITAL RESOURCES



Regional and National Initiatives

- Council of Prairie and Pacific University Libraries (COPPUL) PLN
 - University libraries in Western Canada
 - E-journals from member universities and small presses in Western Canada
 - Institutional repositories
 - CONTENTdm
 - DSpace



Regional and National Initiatives

- US Government Documents (USDOCS) PLN
 - ~30 libraries, US Government Printing Office
 - Ensuring the transition of the Federal Depository Library Program (FDLP) to the electronic era
 - Preserving US government documents from the Federal Digital System (FDSys)





Regional and National Initiatives

- Data Preservation Alliance for the Social Sciences (Data-PASS)
 - Collaborative effort to preserve social science data sets
 - Opinion polls, voting records, income surveys, social network data, government statistics, economic indices, GIS data...
 - Led by the Institute for Quantitative Social Science at Harvard University





Regional and National Initiatives

- Public Knowledge Project (PKP) PLN
 - *see presentation by Mark Jordan*



Regional and National Initiatives

- Other regional, national and special-interest initiatives





Regional and National Initiatives

- Rede Brasileira de Serviços de Preservação Digital CARINIANA
 - Led by Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)
 - Preserving over 1,000 Brazilian open access journals from Sistema Eletrônico de Editoração de Revistas (SEER), an Open Journal Systems (OJS) variant
 - Institutional repositories
 - DSpace





Benefits of National Consortia

- Distributed geographically
- Part of existing collaborative structure
- Aligned preservation goals
- Motivated by legal mandate
- Content resides on national soil



Benefits of National Consortia

- Able to negotiate agreements with publishers
 - Publishers that do not currently participate in global initiatives
- Able to negotiate agreements with hardware suppliers
- More equitable allocation of resources
- Complex access rules managed internally