

The CLOCKSS Archive:

*Un archivio affidabile governato da una “community”
e impegnato a favore dell’open access*

Challenges in Digital Preservation:

Sfide dell’archiviazione digitale

Local Choice / Global Concerns

Scelta locale, attenzione mondiale



- presented by Randy Kiefer

Definitions / Definizioni

Digital Preservation / Conservazione digitale

- refers to the processes and procedures required to ensure content remains accessible well into the future.
- riguarda i processi e le procedure necessarie per assicurare nel futuro l'accesso ai contenuti digitali

Permanent Access (or perpetual access)

Accesso Permanente (o accesso perpetuo)

- is most commonly associated with e-journal licence clauses designed to provide assurance of continued access to subscribed material in certain circumstances, including post-cancellation...
- Solitamente associato alle clausole di licenza degli e-journals per assicurare che l'accesso al materiale sottoscritto continui in particolari situazioni, tra cui la *post-cancellation*



CLOCKSS

Nel 2006, molti importanti editori accademici e biblioteche universitarie si sono uniti per creare un archivio a livello mondiale e far fronte alla sfida e alla responsabilità di dover conservare nel lungo termine il patrimonio della ricerca in formato elettronico, a beneficio dell'intera comunità. Il modello di collaborazione è basato sia su una governance da parte di una community, sia sull'impegno a favore dell'*open access*. Oggi CLOCKSS (Controlled Lots of Copies Keep Stuff Safe) è un'associazione senza fini di lucro governata in modo trasparente e democratico da partecipanti distribuiti in tutto il mondo.

In 2006, several of the world's largest scholarly publishers and top research libraries, challenged by the responsibility to preserve digital scholarly assets for the good of the entire community, joined forces to build a global archive for the very long term. Their unique collaboration emphasized community governance and a commitment to open access. Today, CLOCKSS (Controlled Lots of Copies Keep Stuff Safe) is a not-for-profit organization governed transparently and democratically by participants around the world.



What are the problems? / Quali sono i problemi?

- **Libraries and publishers move to e-only**
- **Biblioteche ed editori scelgono l'e-only**
 - Potential space savings / Potenziale risparmio degli spazi
 - Improves access / migliorare l'accesso
 - 24/7
 - greater number of titles / maggior numero di titoli
 - Increases usage / aumentare l'utilizzo
- **Different business model**
- **Diversi modelli di business**
 - Licensing access /accesso alla licenza
 - Concerns over long-term access / preoccupazione per l'accesso a lungo termine
 - Concerns over preservation / attenzione verso la conservazione



Beginning in 2011 – **Inizio nel 2011**

- **I had worked for a society publisher from 1999 to 2010.**
- **Ho lavorato per un editore aziendale dal 1999 al 2010.**
 - I understood academic publishing.
 - **Ho imparato a conoscere l'editoria accademica.**
- **The CLOCKSS Archive was officially two years old.**
- **L'Archivio CLOCKSS aveva ufficialmente due anni.**
 - Number of Library Supporters = **86**
 - Number of Participating Publishers = **33**
 - Number of Archive Nodes installed = **7**
- **My first office assistant**
- **(Il mio primo collaboratore)**

Charlie (our Labrador)



Funders, librarians and publishers have a unique vantage point on research

Each year

- 3 million articles submitted
- 300,000 peer reviewers
- 1.5 million articles published
- 30 million readers
- 2 billion digital article downloads
- 30 million article citations



The CLOCKSS Archive defines its designated community in three parts:

La community dell'Archivio CLOCKSS è composta da tre soggetti:

1. The scholars, students and readers of electronic academic content.
I ricercatori, gli studenti e i lettori dei materiali accademici elettronici.

2. The libraries who purchase and manage this content on behalf of the scholars, students and readers.

Le biblioteche che acquistano e gestiscono questi contenuti per conto degli studiosi, degli studenti e dei lettori.

3. The publishers of this content.
Gli editori di questi contenuti.



Community Governed - Governato da una *community*

CLOCKSS è progettato per condividere la responsabilità dell'archiviazione a livello di comunità accademica globale. Editori e biblioteche hanno uguale peso nel decidere procedure, priorità e quando dichiarare un contenuto *triggered*. CLOCKSS si impegna ad essere una struttura trasparente e governata da una *community* per una maggiore sostenibilità nel futuro, adattandosi ai cambiamenti tecnologici e di priorità.

CLOCKSS is designed to share the responsibility of archiving across the global academic community. Publishers and librarians have equal say in deciding procedures, priorities, and when to trigger content. CLOCKSS is committed to a transparent community-based governance structure, to be most sustainable in the future as technologies and priorities change.



The CLOCKSS Archive

Services provided to the community / servizi offerti alla community dall'archivio CLOCKSS

The CLOCKSS Archive delivers services to each component of its designated community by providing a sustainable, geographically distributed dark archive that ensures the long-term survival of Web-based, scholarly publications.

L'Archivio CLOCKSS offre un servizio a tutti componenti della propria comunità, fornendo un *dark archive* sostenibile, distribuito in diverse aree geografiche, che assicura la sopravvivenza a lungo termine delle pubblicazioni scientifiche pubblicate sul web.

- Scholars, students and readers are provided with free, open access to content that would otherwise have become inaccessible. **I ricercatori, gli studenti e i lettori possono disporre gratuitamente di contenuti aperti che altrimenti diverrebbero inaccessibili.**
- Librarians are reassured that the content which they purchase will remain accessible to their readers. **I bibliotecari sono sicuri che i contenuti acquistati saranno sempre accessibili dai propri utenti.**
- Publishers are relieved of the responsibility of providing for access to their content in the event that they no longer do so. **Gli editori sono sollevati dalla responsabilità di fornire l'accesso ai propri contenuti anche qualora non li sero più.**



CLOCKSS Archive Nodes

Global Stewardship and Preservation

Gestione e conservazione mondiale

- Australian National University * Australia
- Humboldt University – Berlin * Germany
- Indiana University * USA
- National Institute of Informatics * Japan
- OCLC * USA
- Rice University * USA
- Stanford University * USA
- Università Cattolica del Sacro Cuore * Italy
- University of Alberta * Canada
- University of Edinburgh * United Kingdom
- University of Hong Kong * Hong Kong
- University of Virginia * USA



Global Approach: Decentralized Preservation

Strategia globale: Conservazione decentrata

Libraries preserving content around the globe

Biblioteche nel mondo che conservano i contenuti

- Reinforcing social value as memory organizations
- Rafforza il valore sociale di custodi della memoria
- Insuring against geo-social and geophysical risks
- Rassicura rispetto a rischi geo-sociali e geofisici



What are the Goals of Preservation?

Quali sono gli obiettivi della conservazione?

- A lofty goal: preserve record of scholarship / **Un obiettivo alto: conservare il patrimonio della ricerca**
 - Part of the mission of all research libraries to the world / **Parte della mission di tutte le biblioteche accademiche del mondo**
 - A responsibility shared by the whole community / **Responsabilità condivisa dall'intera community**
 - Thinking hundreds of years ahead / **prevedere con centinaia di anni di anticipo**
- A practical goal: keep your access / **Un obiettivo pratico: mantenere il proprio accesso**
 - Part of the mission of your library to your university / **parte della mission della biblioteca nei confronti della propria università**
 - A responsibility within your library. Thinking tens of years ahead / **Una responsabilità all'interno della propria biblioteca: prevedere con decine di anni di anticipo**



Both goals must be served

Entrambi gli obiettivi devono essere serviti

Digital Archiving / Archiviazione digitale

- Digital archiving was designed for e-journals
- L'archiviazione digitale è stata progettata per gli e-journals
 - Because journals were online first
 - perché le riviste sono state le prime ad essere pubblicate online
 - HighWire Press, Stanford University, 1995



Now, Ebooks are growing rapidly

- The large majority of academic libraries provide ebooks, and the average number of ebooks available in academic libraries that do provide them was 33,830.
- La grande maggioranza delle biblioteche accademiche forniscono e-book e la media di e-book disponibile nelle biblioteche accademiche che li forniscono è di 33,830

From: "The Growing Importance of Ebooks in U.S. Library Collections", Sept 2010

http://www.libraryjournal.com/lj/home/887020-264ebook_summit_kicks_off_with.html.csp



Preservation Concerns?

Preoccupazioni sulla conservazione?

- Is it just like ejournal archiving?
- E uguale all'archiviazione degli e-Journal?
- What's different? What's the same?
- Quali sono le differenze e quali le similitudini?



Archiving Costs / Costi di archiviazione

- *The same for ebooks and ejournals*
- *Sono gli stessi!*
 - Good news!
 - Should make it possible to archive many e-books for a very low cost
 - Questo dovrebbe permettere l'archiviazione di molti e-book ad un costo molto basso



1st Challenge: Archiving Terms And Rights

1. Sfida: Termini e diritti di archiviazione

- *Different for ebooks and ejournals:*
- *Diversi tra e-book e e-journal*
 - Author reversion clause / *Clausola sulla successione dei diritti di autore*
 - publishing rights may revert from the publisher to the author / *I diritti possono passare dall'editore all'autore*
- *Approach:*
- *Strategia*
 - Maintaining more details – authors, rights – tie in with authoritative databases to chase ownership
 - *Registrare più dettagli possibili – autori, diritti – collegarsi a database autorevoli per risalire alla proprietà*



2nd Challenge: Where are the Ebooks?

2. Sfida: Dove sono gli eBook?

- *Most ebooks are on various platforms*
- *La maggior parte degli e-book sono su diverse piattaforme*
 - The platforms are usually with third parties / *Le piattaforme sono di parti terze*
 - Possibly, not all of the collection is in one place / *E' normale che le collezioni non si trovino tutte in un unico posto*
 - Various formats – which one to preserve? / *Formati diversi, quali conservare?*
- *Approach / Strategia:*
 - Collecting the files, including all formats, directly from the publisher or choosing one of the platforms to harvest. Publishers should maintain a copy of all of their files.
 - *Raccogliere i file con tutti i formati, direttamente dall'editore o scegliere una delle piattaforme di raccolta. Gli editori manterranno una copia di tutti i loro file*



Conclusion

- Ebook archiving is somewhat different than ejournals
- L'archiviazione degli e-book è un po' diversa da quella degli e-journal
 - Technical challenge is the same
 - La sfida tecnologica è la stessa
 - The level of detail is greater / il livello di dettaglio è maggiore
 - The due diligence before trigger is more fragmented
 - La «*due diligence*» prima della dichiarazione di contenuto *triggered* è più frammentata
 - Author rights / Diritti di autore
 - Aggregation contracts / contratti di aggregazione
 - The market place and business model has not yet settled down in the same way the journal market has. Preservation will continue to adapt.
 - Il mercato e il modello di business non è ancora stabile come quello degli e-journal. La conservazione subirà continui adattamenti



Now in 2013

Adesso, 2013

- **I have worked for CLOCKSS since 2011.**
- Lavoro per CLOCKSS dal 2011.

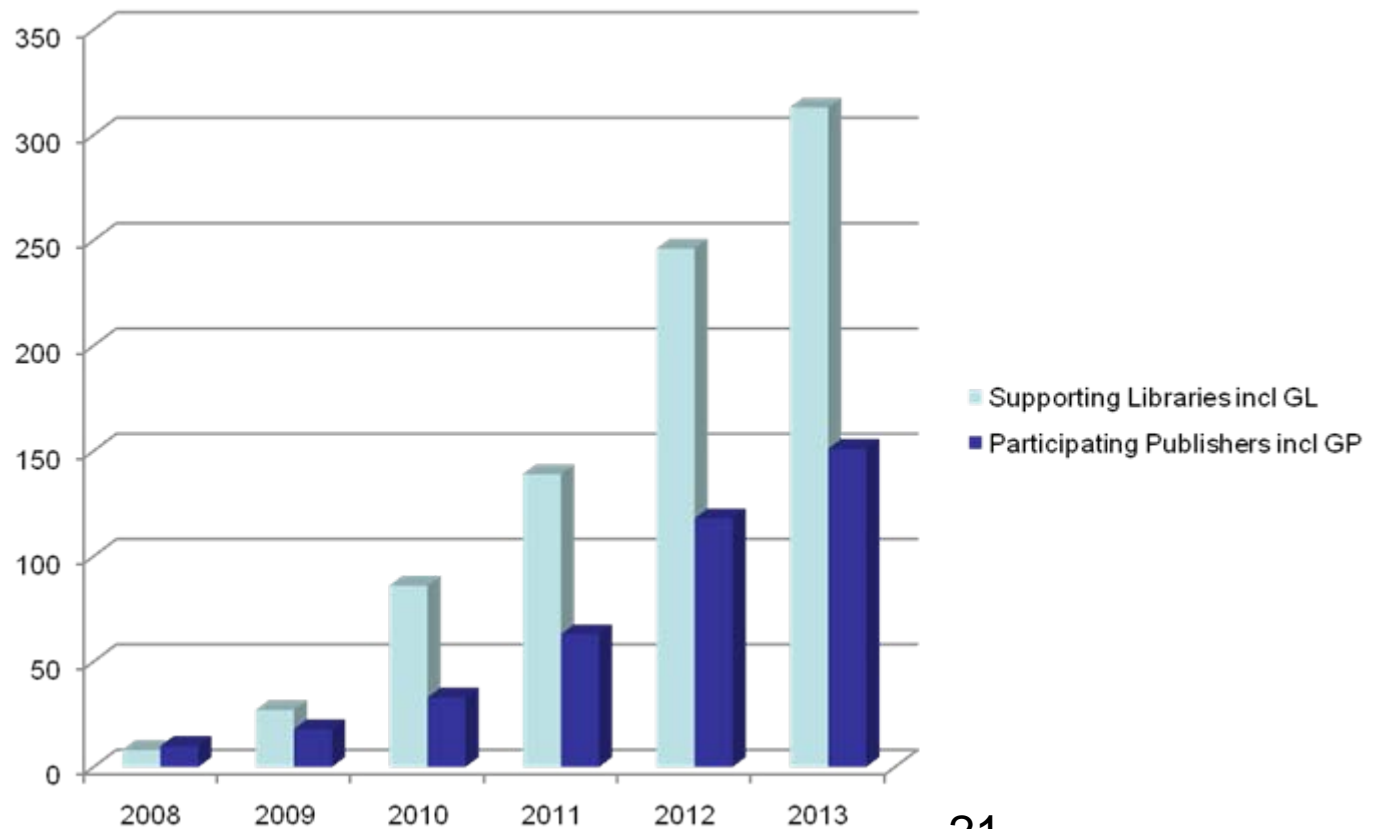
- **I understand digital preservation of academic publications.**
- Ho imparato a conoscere la conservazione digitale delle pubblicazioni accademiche



Now in 2013

Adesso, 2013

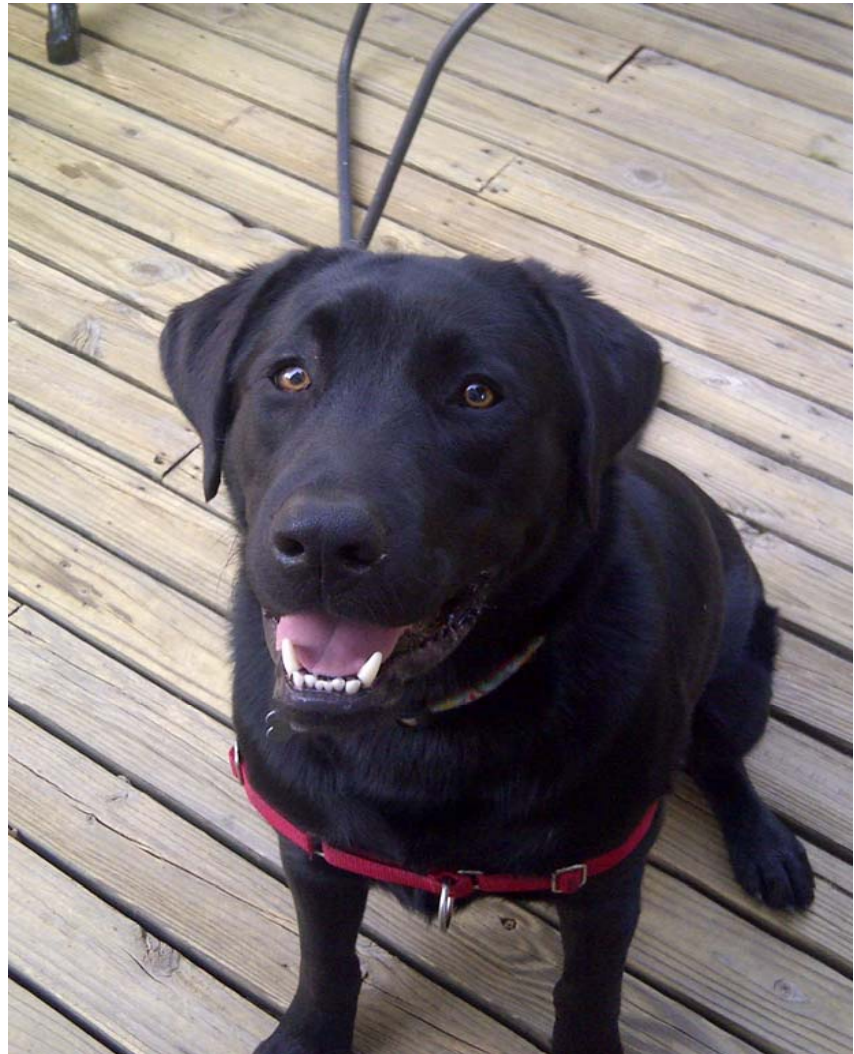
- ❖ Supporting Libraries = **313**
- ❖ Participating Publishers = **151**
- ❖ Number of Archive Nodes installed = **12**



Now in 2013

Ora nel 2013

- ❖ And Charlie is still my office assistant
- ❖ E Charlie è ancora il mio collaboratore
- ❖ He is now two years old
- ❖ Adesso ha due anni



Thank You!
Grazie!

Questions or Comments?

Domande o commenti?

Randy Kiefer
Executive Director
Direttore Esecutivo

The CLOCKSS Archive
L'Archivio CLOCKSS

[mailto: randy.kiefer@clockss.org](mailto:randy.kiefer@clockss.org)

