Collaborative Collections

Alex Thurman

Web Archiving Collaboration: New Tools and Models
June 4, 2015
A GLOBAL NETWORK OF EXPERTS ARCHIVING THE WEB FOR FUTURE GENERATIONS.

Learn More About IIPC
<table>
<thead>
<tr>
<th>IIPC member region</th>
<th>#</th>
<th>IIPC member type</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe</td>
<td>26</td>
<td>national/regional libraries</td>
<td>33</td>
</tr>
<tr>
<td>North America</td>
<td>14</td>
<td>universities</td>
<td>8</td>
</tr>
<tr>
<td>Asia</td>
<td>4</td>
<td>other non-profits</td>
<td>5</td>
</tr>
<tr>
<td>Oceania</td>
<td>2</td>
<td>national archives</td>
<td>1</td>
</tr>
<tr>
<td>North Africa/Middle East</td>
<td>2</td>
<td>commercial</td>
<td>2</td>
</tr>
<tr>
<td>South America</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
IIPC members, United States
MISSION & GOALS

Mission
The mission of the International Internet Preservation Consortium (IIPC) is to acquire, preserve and make accessible knowledge and information from the Internet for future generations everywhere, promoting global exchange and international relations.

Goals
To achieve its mission, the IIPC is working to accomplish the following goals:

➢ To enable the collection of a rich body of Internet content from around the world to be preserved in a way that it can be archived, secured and accessed over time.
➢ To foster the development and use of common tools, techniques and standards that enable the creation of international archives.
➢ To encourage and support national libraries, archives and research organisations everywhere to address Internet archiving and preservation.
IIPC Olympics Collections
(4550+ seeds in total, 23TB of data)
Collaborative collections
Proposal

Announced at 2014 General Assembly in Paris

Make IIPC collecting more robust and sustainable

Purchase Archive-It subscription?

Proposal was popular with members

Need to consider risks, evaluate all vendor options
Collaborative collections

Benefits

- Access
- Agility
- Collaboration (more institutions, between GAs)
- Promotion/Advocacy
- Researcher engagement
Varying legal environments

Able to archive national domains and provide online access
- National and University Library of Iceland
- National Library Board Singapore
- Portuguese Web Archive

Have legal deposit for domain harvesting, but no copyright exception to provide online access without site owner permission
- Bibliotheque nationale de France
- British Library
- National Library of the Czech Republic

No legal deposit requirement or explicit copyright exception for web archiving, so permission-based harvesting and online access
- Columbia University
- Koninklijke Bibliotheek
- National Library of Australia
Collaborative collections

Risks

- Legal
- Sustainability of IIPC project funding
- Long-term preservation
- Competition with or duplication of existing member collections
Collaborative collections
Structure and vendor choice

IIPC Content Development Group

Working group for shaping, harvesting, describing collections

Roles include: Leader(s); Collection administrator; Lead curator(s); Contributing curator(s)

CDG identified service requirements (including data capture, descriptive metadata, access, preservation, cost) and nine services were considered

Archive-It selected (metadata, common platforms, all-inclusive costs including storage and access)

Archive-It subscription effective April 2015. Crawl budget of 50,000,000 documents and 5TB
2014 Winter Olympics
Collected by: International Internet Preservation Consortium
Archived since: Dec, 2013
Description: A collection of web coverage of the 2014 Sochi Winter Olympics, compiled by members of the IIPC
Creator: International Internet Preservation Consortium

Title: RT - Sochi Olympic Games 2014 Torch Relay
URL: http://2014.rt.com/torch-relay

Description: Photos of the torch relay in cities across Russia. Khabarovsk - Torch Relay Captured 36 times between Jan 6, 2014 and Feb 26, 2014
Videos: 609 Videos Captured
Subject: Media - Independent
Language: English
Country: Russia (Federation)
Sport: Alpine Skiing, Biathlon, Bobsleigh, Cross-Country Skiing, Curling, Figure Skating, Freestyle Skiing, Ice Hockey, Luge, Nordic Combined, Short Track, Skeleton, Ski Jumping, Speed Skating, Snowboard

Title: Adam Barwood
URL: http://adambarwood.com
5-21 AUGUST 2016

10,500 ATHLETES FROM 205 COUNTRIES

WILL COMPETE IN THE FIRST OLYMPIC GAMES IN SOUTH AMERICA

306 MEDAL EVENTS

136 WOMEN'S
9 MIXED
161 MEN'S
Collection topic criteria

- Of high interest to IIPC members
- Does not map to any one member’s responsibility or mandate
- Adds research value via broader perspective
- Transnational in scope, but not necessarily “global”
World War I Centenary Commemoration collection

Several members planning WWI collections (BnF, BL, LC, NLA, LAC)

No member has responsibility for comprehensive multinational collection

IIPC joint collection could be duplicative but improves access and adds research value via broader perspective

CDG now consolidating seed lists from these member collections and other interested members
EXPLORE OUR COLLECTION

Library of Congress WWI collection seed
HVERDAGSLIV I KRIGENS SKYGGE
KILDER OM DANMARK UNDER 1. VERDENSKRIG

"Du Purling-Land, som hygger dig i Smug, mens hele Verden brænder om din Vugge"

Fra Jeppe Aakjærds digt "Historiens Sang", 1916


"Hverdagsliv i krigens skygge" er beregnet til undervisning på ungdomsuddannelser. Ved at undersøge de historiske kilder på disse sider, kan du blive klogere på de skygger, krigen kastede ind over.

Netarchive.dk WWI collection seed
Nous recherchons en permanence documents, archives et photographies sur la Guerre des gaz ; je recherche également les objets se rapportant à ce thème (masques à gaz).

Vous pouvez nous contacter à l'adresse suivante : Guerredesgaz@aol.com

Dernières mises à jour
Women’s Land Army

The Women’s Land Army (WLA) began during World War 1 when many men left farming to fight in the trenches. Women took over in the fields to help grow the food the country needed. Food couldn’t be imported and we needed to be self-sufficient. When Britain went to war again in 1939 the Land Army was called up a second time. By 1944 there were 80,000 women working on farms all across the country.

They did every kind of farm work from driving tractors to milking cows and harvesting wheat. They also worked in the forests cutting down trees because wood was also in very short supply. Despite their efforts and the vital part they played in the war, the WLA were not officially recognised for their work until 2008 when they were issued with a badge from DEFRA, the Department for Environment, Food and Rural Affairs.

British Library WWI collection seed
“Everything has to start again:” Some Thoughts on the Centennial of the Assassination of Franz Ferdinand

Today, as many of you know, marks 100 years since the assassination of Austro-Hungarian Archduke Franz Ferdinand and his malignantly wife Sophie on the streets of Sarajevo. While some marvel at the
Library of Congress WWI collection seed
‘The Army

During the First World War, the British Army faced great hardships and challenges. German Zeppelin raids on London were frequent, and the war in the western front was particularly brutal. However, the British Army maintained its strength and resilience, with many soldiers volunteering and serving with distinction.

British Library WWI collection seed
International Cooperation Organizations collection

Value of an ongoing thematic collection that resonates with IIPC’s mission of promoting global exchange and international relations

Makes more sense for IIPC to build collection of international cooperation organization websites than it would for individual members

Can be started with comprehensive collection of the small but important top level domain “.int”

Small collection of international organizations allows for experimenting with rich metadata, linked data
The domain name int is a sponsored top-level domain (sTLD) in the Domain Name System of the Internet. Its name is derived from the word international, characterizing its use for international organizations and treaty-related purposes.[1] The first use of this domain was by NATO, which had previously been assigned the top-level domain nato.

According to Internet Assigned Numbers Authority (IANA) policy, based on RFC 1591, the sTLD int is reserved for international treaty-based organisations, United Nations agencies, and organisations or entities having Observer status at the UN.[2] int is considered to have the strictest application policies of all TLDs, as it implies that the holder is a subject of international law. For this reason, the application procedure requires the applicant to provide evidence that it is indeed treaty-based by providing a United Nations treaty registration number and that it has independent legal status.
A GLOBAL NETWORK OF EXPERTS ARCHIVING THE WEB FOR FUTURE GENERATIONS.

Learn More About IIPC