

## 2010 IIPC Overview

The International Internet Preservation Consortium (IIPC) was founded in 2003 by 12 national libraries from Europe, North America, Australia and the Internet Archive. The mission of the IIPC is to acquire, preserve and make accessible knowledge and information from the Internet for future generations everywhere, promoting global exchange and international relations. In 2007, 15 new institutions joined the consortium and in 2008 the total reached 38 members. In 2010 there are 39 members, including universities, national archives and foundations.

### Structure of the IIPC

The IIPC accomplishes its mission through Working Groups that produce work packages based on the needs and concerns of the Working Group members. Overall strategic direction, approval of new members and budgets for projects are provided by a Steering Committee made up of 15 member institutions, including a Chair/coordination institution. Coordination between members and Working Groups are accomplished by a Program Officer, a Treasurer and a Communications Officer. More information on IIPC leadership roles, as described in the Membership Agreement found on our public Web site: <http://netpreserve.org/membership/IIPCagreement.pdf>.

### Harvesting Working Group

The Harvesting Working Group's primary focus is the development of web harvesting technologies, particularly around the Internet Archive's Heritrix web crawler. The major areas of work include a smart crawler. Other areas of focus include:

- Development of a smart crawler and improving harvesting performance
- Development and support of the WARC file format
- Best practices and databases for sharing crawl information in bulk or selective harvesting
- Feature requests for crawler
- Harvesting the deep web
- Harvesting video and streaming media

### Access Working Group

The Access Working Group (AWG) focuses on issues relevant to providing access to web archives. In addition to technical research and development, the group recognizes the legal, ethical and economic aspects of access. Furthermore, end-user, administrative and curatorial access to web archives all form parts of the group's considerations. Other focus areas include:

- Understanding and defining user requirements for access
- Resources discovery including full-text and innovative ways of searching web archives
- Access to multimedia content within archived websites

- Tools for analysis of structure and content of web archives
- Identification and documentation of web archive use cases
- Technology watch

## Preservation Working Group

The Preservation Working Group (PWG) focus is on policy, practices and resources in support of preserving the content and accessibility of web archives. The PWG aims to understand and report on how approaches used for other kind of digital resources might be used with web archives, as well as the special characteristics of web archives that might require new approaches. It will provide recommendations for additions or enhancements to tools, standards, practice guidelines, and possible further studies/research. Other focus areas include:

- Characterize large scale web archives in order to
  - Identify relevant approaches, standards and practices already used for preservation of other digital assets
  - Report on how they might be used with archived web resources and/or
  - Identify the gaps and promote new approaches.
- Enhancements or additions to tools, standards, practices, guidelines, testing, and possible further studies/research.
- The unique requirements to preserve archived web resources not achieved by other preservation programs for digital assets.

## 2010 Leadership

Coordinating Institution/Chair	Raju Buddhharaju - National Library Board, Singapore
Program Officer	Aaron Binns - Internet Archive
Treasurer	Gildas Illien - Bibliothèque nationale de France
Membership & Communications	Abigail Potter - Library of Congress
Harvest WG leaders	Kristinn Sigurdsson - National and University Library of Iceland
	Lewis Crawford - The British Library
Access WG leaders	Kris Carpenter Negulescu - Internet Archive
	Helen Hockx-Yu - The British Library
Preservation WG leaders	Clément Oury - Bibliothèque nationale de France
	Gina Jones - Library of Congress

## For more information

Public Website: <http://netpreserve.org>

Members forum: <http://netpreserve.org/forum/> (login required)

Contact: Abigail Potter, Communications Chair at [abpo@lov.gov](mailto:abpo@lov.gov)