



international internet preservation consortium

# **Preservation Working Group report**

Gina Jones, Library of Congress

Clément Oury, National library of France

# Goals of the PWG

- Several tools have been designed in order to ensure preservation of different kind of digital documents :
  - Software
  - Metadata schemes
  - Best practices and guidelines
- We archives have specific needs that you need to understand in order to
  - Identify what is relevant in policies, practices, and tools already in use
  - Identify the gaps
  - Promote web archives specific needs.
- The PWG is a group of experts who
  - work collaboratively to produce studies and technical analyses
  - act as a forum for best practices and guidance
  - identify and assess development needs
- Source: [http://www.netpreserve.org/about/PWG\\_TOR\\_final.doc](http://www.netpreserve.org/about/PWG_TOR_final.doc)

# Working fields

- **objectives** and concepts of preserving archived web resources;
- preservation **metadata** for archived web resources (capture, packaging, usability);
- preservation **workflows** and **digital repository** functions and requirements;
- preservation **strategies** for long-term access to archived web resources (migration, emulation...);
- **technical documentation** (notably web technical environment documentation);
- evaluation of digital **preservation tools** and gaps for use by web archiving institutions;
- **organizational issues** (costs, sustainability, promotion, skills,...).

# Main 2010-2011 deliverables

- Risk assessment database
  - Build a list a risks for web archives from NLA table of threats
  - Identify risks bibliography
  - List of risk mitigation strategies
  - Build a database and an online interface for institutional risk assessments
- Web environment database
  - Survey institutional web / web archives access environments
  - Create an online database
  - Start describing web software and formats
  - See: <http://www.ignaciogc.com/netpreserve/>



international internet preservation consortium

# PWG database

iipc · netpreserve beta

[Home](#)

[Institutions](#)

[Environments](#)

[Software](#)

[Formats](#)

[Risks](#)

[Risk Controls](#)

[Bibliography](#)

[F.A.Q.](#)

Home

## Preservation Working Group



[Institutions](#)



[Environments](#)



[Software](#)



[Risks](#)




[Bibliography](#)



[Statistics](#)

me » [Environments](#)

## Environments

 [Add new environment](#)

1 2 3 4 » » » »

Showing 1 to 3 of 20

### National Library of Australia (2010)

NLA Reading Room - Environment for onsite use.

#### Operating system

Windows XP Professional Service Pack 3

#### Browsers

Internet Explorer 7.0

#### Applications

WinZip

Windows Media Player 11

Microsoft Office Powerpoint 2007

Microsoft Office Excel 2007

Microsoft Office Word 2007

#### Plugins

Adobe SVG Viewer

AlternaTIFF 2.0

DJVu Control 4.5

MrSID Browser Plugin 1.3

Java 1.4.2

Adobe Flash Player 10

Adobe Shockwave Player 10.1

QuickTime 7.0

RealPlayer 11

# Main 2010-2011 deliverables

- Information package survey
  - An information package: a consistent set of data with associated metadata that is ingested, stored and disseminated by a digital repository
- Share and exchange information packages from different institutions and different harvesting models
- Compare information packages in order to identify similar patterns and potentially best practices

# Our goals for this week

- Present outcomes of past work packages
- Decide WP statuses
  - Close some of them, continue the others, revise tasks, identify new ones
- Work on the public part of the database
  - What is interesting for the public?
  - How do we repurpose the data?
  - What else should we be looking at?

# Agenda

- **Wednesday** – Review and discussion of accomplishments/WP's status
- Review PWG database
  - Discuss relationship with UDFR/Pronom
  - Discuss repurposing PWG data for public IIPC site
- **Thursday** – Risk Assessment
  - Review of database
  - Review of testing assessments
  - Future work (analysis/output tools, policy development)
- Information Packages.
  - Conclusions from survey
  - Review of other projects outside IIPC related to web archives preservation