Web Archiving in National Diet Library

Yukiko Saito

Director General

International Library of Children's Literature

National Diet Library

CDNLAO Apr. 20, 2009

1. Today's Topics

- Forthcoming issue: Web archiving (coordinated by NDL)

- Web archiving project in NDL

- How to promote Web archiving in CDNLAO

2. Web Archiving in NDL

☆ **WARP (Web Archiving Project) (2002-)**

　　Websites: 2,448 titles (of central and local government, corporations

　　　　　　　and organizations, universities, events, etc.)

　　Online journals: 1,827 titles

• Selective harvesting based on contracts with right holders

☆ **Institutionalization of Web archiving**

• Comprehensive collection of websites in Japan

　　　　↓

• Government, research institutions, organizations

　　　　↓

• Only government information and academic information of
national and public universities, etc.

☆ **Scope of Harvesting**
　→Who collects ? What & How?

☆ **Technical issues**

　→**What is long-term preservation?**

☆ *Needs of International Cooperation*

4. International Efforts for Web Archiving

- Scope of acquisition and methods for Web archiving vary among countries

- IIPC（International Internet Preservation Consortium）

- Trilateral Cooperation between the National Library of China, the National Library of Korea and NDL

Questionnaire survey

1) Invite participation in the survey and comments on its

   questionnaire items

2) Fix questionnaire items

3) Conduct the survey

4) Share the results

5) Sort out collaboration themes from the results

1. How widely is the Internet spread in your country?

    What is the Internet diffusion rate?

    Do you have an approach for spreading Internet use?

2. Do you have a legal deposit or other legal systems for Web archiving?

   If not, please let us know the reason and future prospect.

3. Do you or will you have a Web archiving project?

   If you have, what is the scope of acquisition?

   Do you have a Web archiving system?

4. What are the challenges or obstacles to implementing the Web archiving project?

   Do you have legal or technical problems or others?

5. Does your country have a framework to share roles in acquisition, preservation and provision of Web information?

   In that case, what is your role?

   What institutions are you working together with?

6. What is the scope of acquisition of Web information in your country?

    What are the criteria to identify the scope?

    For example, do you employ bulk archiving or selective archiving of specified fields?

    What is the frequency of archiving?

7.   How do you think we can cooperate in the Web archiving?

What do you expect for other CDNLAO members in cooperating in Web archiving?

8. How do you provide or utilize accumulated Web contents?

   Please let us know how you use, how you provide them on the Internet, etc. Please give us examples.