



IIPC Harvesting Practices Report

Michaela Mayr

Austrian National Library





Project History

- Survey developed by National Library Board Singapore
- Opened to members in April 2010
- Austrian National Library volunteered to take over the project
- Started to compile the report in November 2010
- IIPC Harvesting Practices Report Version 1 published May 2011



Goals

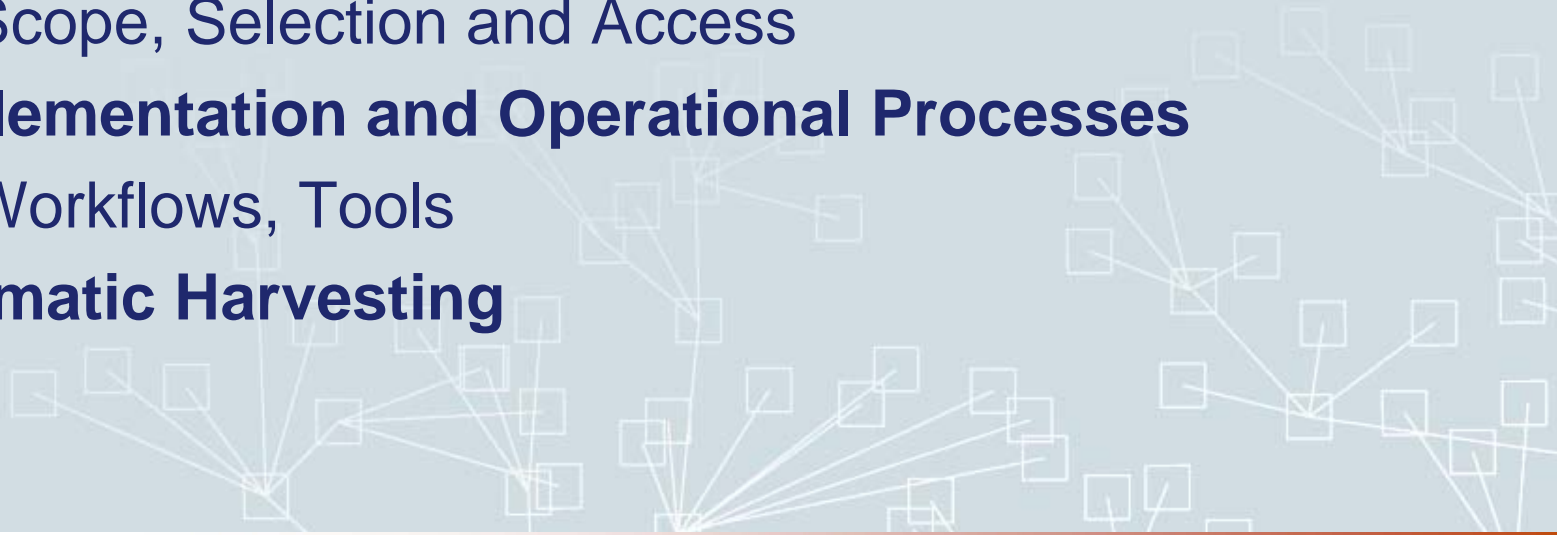
- From members for members
- Overview of current web archiving practices
- Encourage and support you to address various web archiving issues
- No manual for the implementation of National Domain Crawls





Structure of the Report

- 97 questions, 4 sections
- **Harvesting Project Information**
 - Background Information, Resources, Web Archives
- **Policy and Appraisal**
 - Scope, Selection and Access
- **Implementation and Operational Processes**
 - Workflows, Tools
- **Thematic Harvesting**



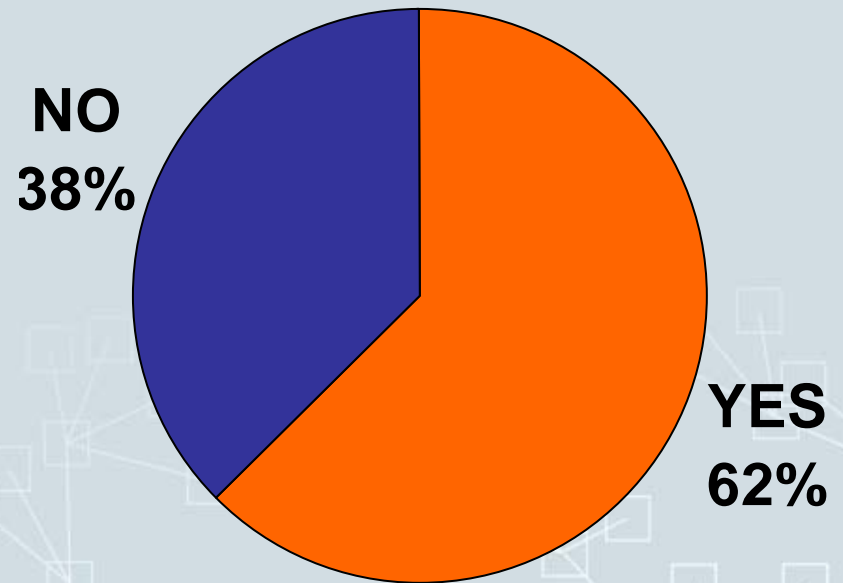
16+ Responders – Thank you!

Institution	Abbr.	Region
1. The National Library of Israel	IL	Asia
2. National Diet Library, Japan	JP	
3. National Library of Korea	KR	
4. Österreichische Nationalbibliothek (Austrian National Library)	AT	Europe
5. Netarchive.dk (Royal Library/State and University Library, Aarhus)	DK	
6. Biblioteca Nacional de España (National Library of Spain)	ES	
7. Bibliothèque nationale de France (National Library of France)	FR	
8. British Library	GB	
9. Koninklijke Bibliotheek (National Library of The Netherlands)	NL	
10. Nasjonalbiblioteket (The National Library of Norway)	NO	
11. Kungliga biblioteket (National Library of Sweden)	SE	
12. National and University Library of Slovenia	SI	
13. Bibliothèque et Archives Nationales du Québec (BANQ)	CA	North America
14. California Digital Library	CDL	
15. Internet Archive	IA	
16. National Library of New Zealand	NZ	Oceania



Some Facts...

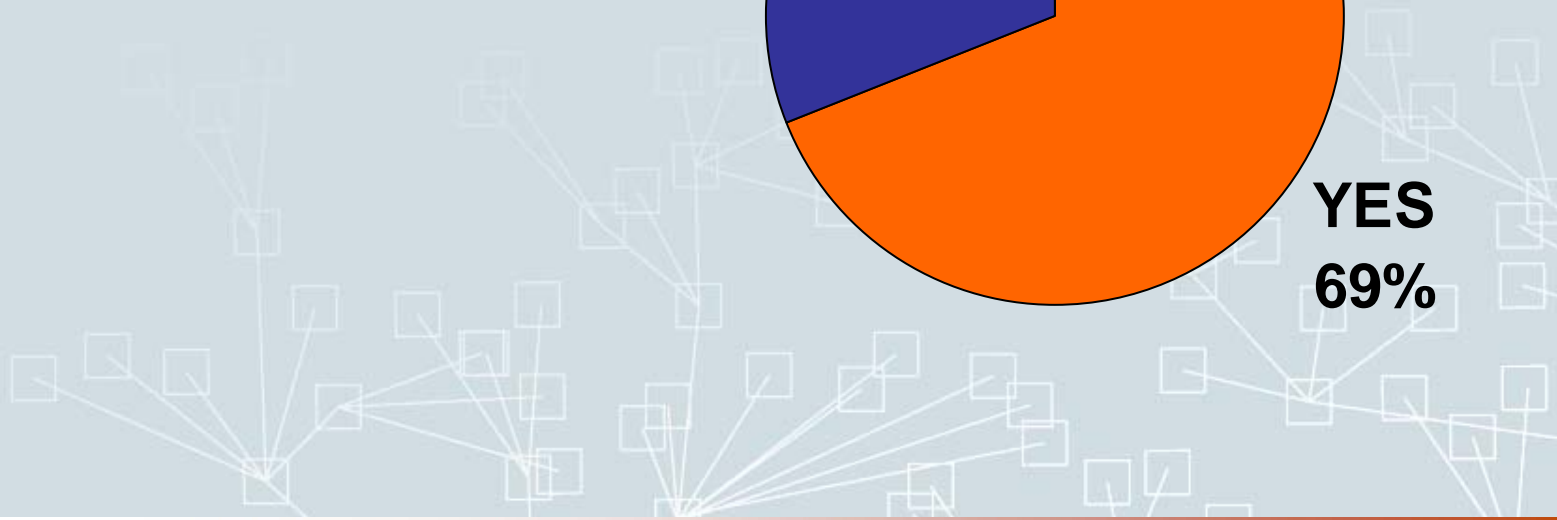
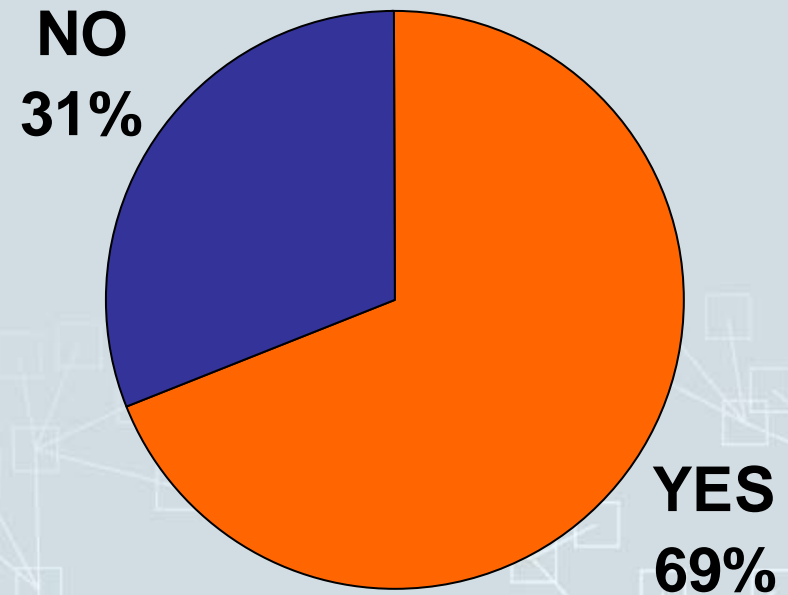
→ Department or Team focusing on Web Archiving?





Some Facts...

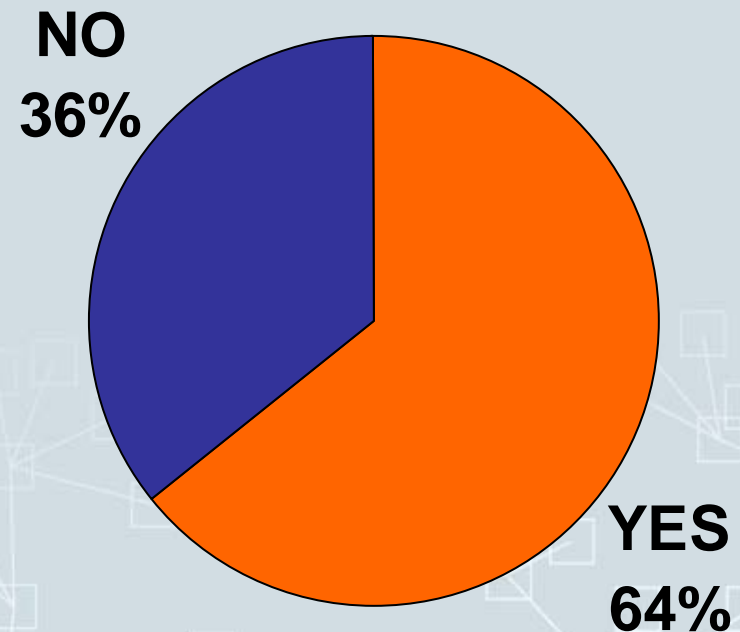
→ Statutory Act in place for Web Archiving?





Some Facts...

→ Do you perform
Domain
Harvesting?



Some Facts...

→ What requires the most manpower, time and cost?

	Mentions
Selection of Sites (Scope)	6
Development of tools	5
Monitoring	4
QA of results	3
Problematic sites, technical issues	2
Preparation and planning	2
Data Management and Maintenance	2
Policies	2
Preservation	1



Some Facts...

→ Favourite Features of Web Archives?

- Ability to browse the past
- Ability to recover lost sites when they are accidentally lost from the live web
- Personal support by librarians and web-archiving team
- Online access
- Extra depth provided by thematic collections
- Usability of the services
- Site comparison features





More in the report...





Contact me

- Missing surveys until end of May
- Feedback
- Interest in Follow-Ups
- Michaela.Mayr@onb.ac.at
- Updated version in June

