Digital Preservation Activities
in National Library of China(NLC)

Li Chunming

National library of China

1

International Conference on Preservation of Digital Objects

GÖTTINGEN

Outline

➢ strategic Aim of NLC

➢ Research in Progress

➢ Current Activities

➢ Others

2

National Library of China(NLC)
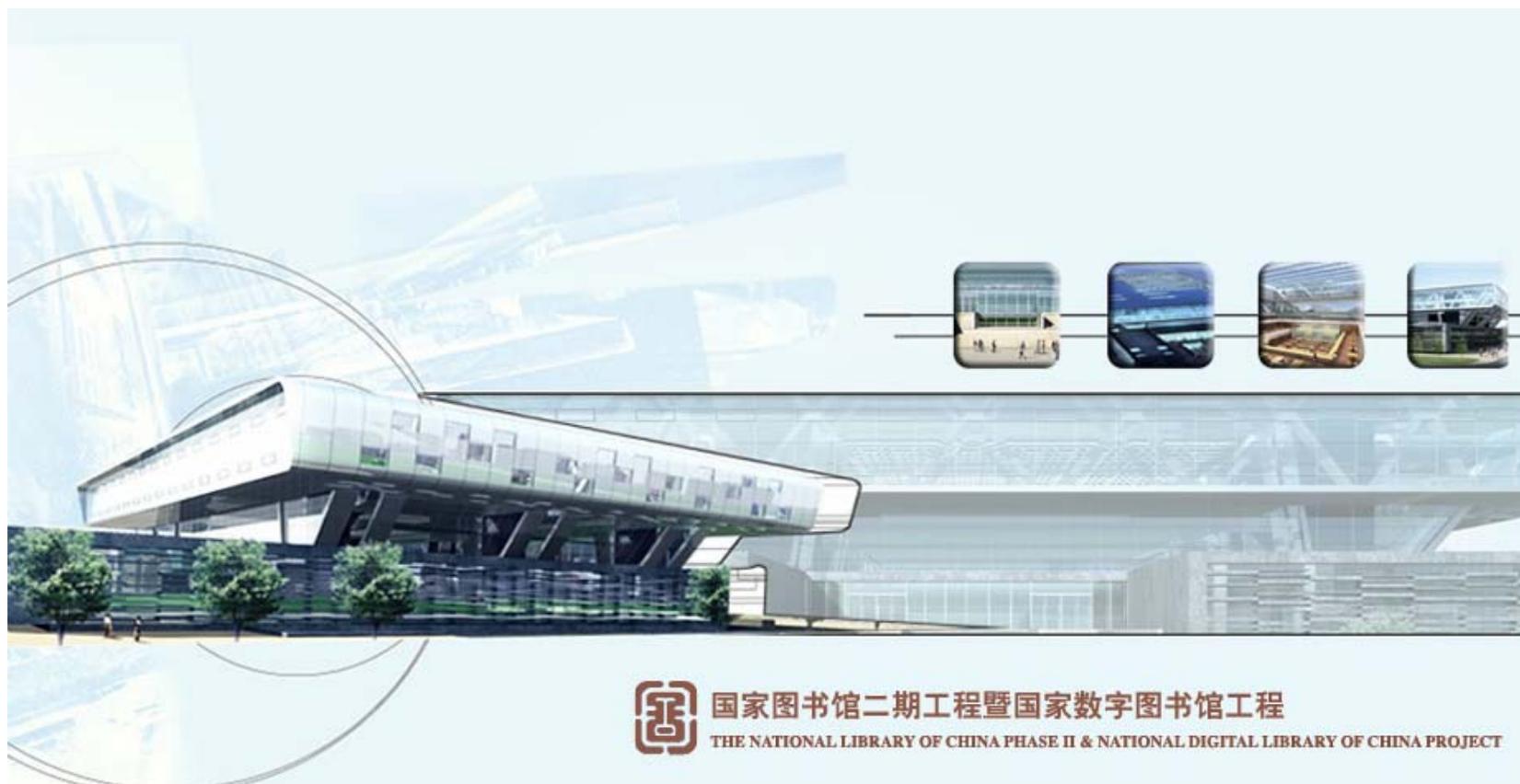

3

strategic Aim of NLC

✦ **Maximize collection of Chinese digital heritage**

✦ **Selected collection of foreign academic resources**

✦ **Resolution under international standards**

✦ **Utilization of advanced IT technologies**

✦ **Enhancement of multilateral cooperation**

4

National Digital Library of China Project

国家图书馆二期工程暨国家数字图书馆工程
THE NATIONAL LIBRARY OF CHINA PHASE II & NATIONAL DIGITAL LIBRARY OF CHINA PROJECT

(National Library of China Phase II will be completed in 2007)

5

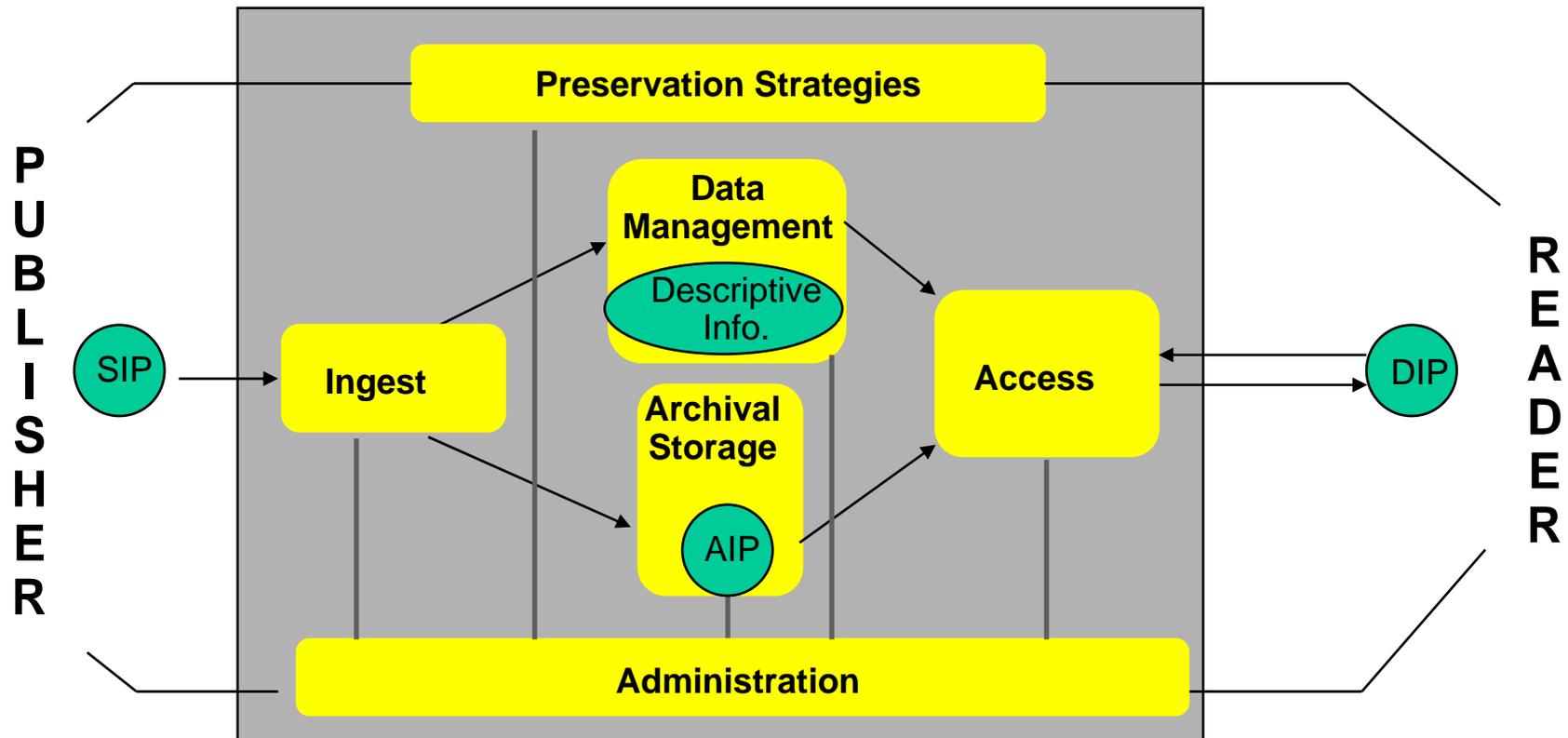Sep. 16, 2005   International Conference on Preservation of Digital Objects

GÖTTINGEN

Research in Progress

6

OAIS and Software

PUBLISHER

READER

Preservation Strategies

Data Management

Descriptive Info.

SIP

Ingest

Archival Storage

AIP

Access

DIP

Administration

SIP = Submission Information Package

AIP = Archival Information Package

DIP = Dissemination Information Package

7

Metadata schemes

MARC, DC, etc. Comprised in METS

Search systems

Compliant with OAI , Z39.50, X-server

8

Preservation strategies

- Secure storage

- Backing up

- Specification of naked data structure

- Normalization

9

Current Activities

- **Obtain policy support and Raise public awareness**

- **DiNeR: Digital Newspaper Repository**

- **Progress on WICP project**

GÖTTINGEN

Obtain policy support and public recognition

- Ask the government for policy support on permanent digital resource preservation

- Let the public realize the significance of digital resource preservation with multiple methods

11

Digital Newspaper Repository

- Newspapers are a major publishing source.

- Accessibility to these collections is very limited.

- The traditional search methods are too slow and difficult.

- The digital collections of large daily newspapers made the access to newspapers more effective and efficient.

12

13

GÖTTINGEN

What kind of titles were selected?

2,119 newspapers in total

213 national newspapers

766 provincial newspapers

898 prefecture and city newspapers

242 county newspapers

Websites of Newspapers 1,800

Focused on 80 websites with PDF files

14

Intellectual property processing

- Chinese law stipulates that the National Library is entitled to get donation of e-publication.

- Contacting for authorization to obtain the right.

15

Capturing, collecting and service

- **The capturing tool: Flashget**

- **Collecting**

- **Metadata Generation**

- **Object organization**

- **Access service**

16

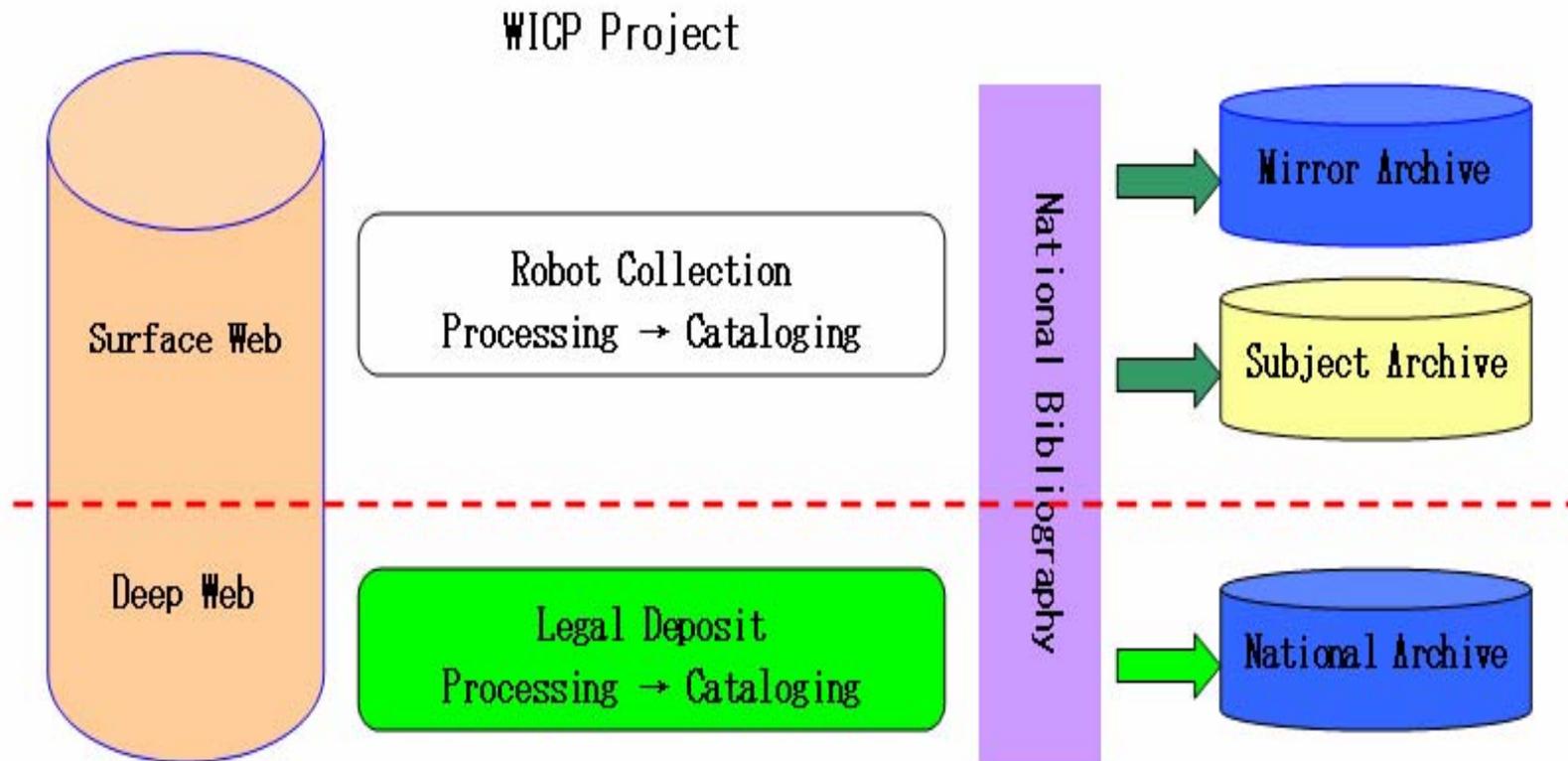**Sep. 16, 2005** **International Conference on Preservation of Digital Objects**

GÖTTINGEN

Future Development

- Content Enhancement

- Content Management

- Opening to the public

18

Progress on WICP project

WICP Project

GÖTTINGEN

Partnerships

- TianWang, Peking University

- Tsinghua Tongfang

- BeiDa Founder

- WanFang Data

20

中国 Web 信息博物馆

Web InfoMall

English

Web InfoMall 简介

"中国Web信息博物馆"是在国家 973和985项目支持下，北京大学网络实验室开发建设的中国网页历史信息存储与展示系统。目前已经维护有10亿以中文为主的网页，并以平均每月一千万网页的速度扩大规模。

历史网页回放

http:// **GO!**

Web InfoMall 信息服务

▶ 输入URL，浏览永久保存的历史网页，欣赏旧时网页的风采

▶ 畅游昔日网站，随意纵横比照，品味网络世界的兴衰变迁

▶ 关注重大历史事件，将发展进程历历览尽，感受时代的进步

▶ 申请网页数据，研究深层联系，挖掘信息世界的潜在秘密

○ Web InfoMall是一项服务社会的公益事业

○ 汇集历史网页，再现大千变化，包容万千信息探索无穷智慧

Web InfoMall 数据量

▶ 共有自2001年以来 1,017,104,240 网页在线（17,106 GB） 详细 >>>

历史事件专题回放

党的十六大专题
聚焦伊拉克战争
2003 年 4月底广东网上信息的一次快照与分析
抗击非典的日子
中国首次载人航天
NEW 其他..
建立自己的主题..

信息产品与服务

中国Web网页全文
中国Web网页链接URL
天网搜索引擎用户日志
中国IP地址信息大全
中文网页分类训练集
理念与规则
公共许可证
申请使用数据

相关链接

天网:Web InfoMall工作的基础
燕穹: 在Web InfoMall基础上的科学研究和增值信息服务体系
CNDS:Web InfoMall的摇篮

典型历史网页展示

1996 年中国网页拾零
1997 年中国网页拾零

关于我们

使用帮助 问题解答
Web InfoMall论坛
联系我们

Thank you very much for your attention.

My Email: licm@nlc.gov.cn

National Library of China

http://www.nlc.gov.cn

22

International Conference on Preservation of Digital Objects

GÖTTINGEN