



中国科学院国家科学图书馆  
National Science Library, Chinese Academy of Sciences



# Constructing a national S&T literature preservation system

---

**Zhenxin Wu**  
**National Science Library,**  
**Chinese Academy of Sciences**



# Outline

---

- Archiving Activity at the NSLC
  - Part of the Chinese Academy of Sciences
- NSLC Digital Preservation System (brief view)
- Constructing a National Co-operation Preservation Network [Pilot Project]

## DPC of CAS

- NSLC has been working to build digital infrastructure for digital resources management
  - long-term preservation is regarded as an important part
  - Digital preservation centre of CAS works on this target.





# Resources policy of preservation

---

- Commercial resources
  - e-journal, e-Book, Protocols, etc.
- Free resources
  - OA e-journal
  - Web information (Web Archive)
- Research output of CAS
  - More than 100 institutions



## Progress report: Subscribed content

- Springer (journal、eBook、Protocols), BMC, IOP, RSC (journal、eBook), NPG, PNAS, Wiley (journal、eBook),
- VIP (Chinese): >14000, all with ISSN, not e-only
- dnc.las.ac.cn

用户名:  密码:  [登录](#)

### 数字科技文献长期保存中心 Digital Preservation Center

[首页](#)

[机制与政策](#)

[合作保存战略](#)

[保存服务](#)

[关于我们](#)

[出版物](#)

[存档统计](#)

[数据存档](#)



#### 长期保存，为了未来的持久访问！

伴随着信息技术的迅速发展，越来越多的人类知识成果可以在网络上以数字形式被获取，然而，这些数字资源却转瞬即逝。作为加工、管理和保存这些数字资源的长期方法，数字存档计划被认为具有永恒的价值。自从上个世纪90年代开始，不少国家，像澳大利亚、美国以及欧洲国家就在保存这些资源方面展开了持久的努力。而作为这些国家的长期规划，这些项目一般都是由相应的国家图书馆和其他机构和组织共同合作开展的。...

[了解详情](#)



动态追踪

Dynamic tracking

[RSS](#)



合作保存

Preservation

[RSS](#)



新闻

News

[RSS](#)

2015年4月8日下午，中国科学院文献情报中心（以下简称文献中心）与Springer科学与商业媒体集团（以下简称Springer）正式签署Springer电子图书长期保存合作扩展意向书。该意向书是继2011年Springer电子期刊长期保存合作扩展系列协议成功签署之后，文献中心与

国科图将与全国图书馆共建共享数字文献资源长期保存服务。（1）高效益的长期保存依赖合作。数字文献资源长期保存涉及大量资源、海量内容、复杂技术与管理问题、长期经济与管理负担，任何单一机构难以完全由自己对所需的所有资源进行长期保存。通过共建共享长期

- 长期保存领域动态监测2014年第三期 [研究报告]
- 长期保存领域动态监测2013年第十二期 [研究报告]
- 长期保存领域动态监测2013年第十一期 [研究报告]
- 长期保存领域动态监测2013年第十期 [研究报告]

# Progress report: Open Access content

- Selected 4055 high quality OA journals,
- Harvested full text from 1765 OA journals
- [gooa.las.ac.cn](http://gooa.las.ac.cn)

The screenshot displays the Go OA website interface. At the top, the logo "Go OA" is followed by the text "开放论文一站式发现" and "开放期刊投稿分析". Navigation links include "首页", "OA期刊", "期刊投稿推荐", "开放接口", "帮助", and "关于GoOA". A search bar is present with a "检索" button. Below the search bar, there are statistics: "1765种 GoOA遴选期刊", "145家 OA出版社", and "268384篇 OA论文". The main content area shows search results under the "GoOA期刊遴选标准" tab. The first result is a case report: "Case report of renal cell carcinoma in automobile manufacturing factory worker due to trichloroethylene exposure in Korea". The second result is "FGF21 protects against ox-LDL induced apoptosis through suppressing CHOP expression in THP1 macrophage derived foam cells". On the left side, there are two sections: "热点概念" (Hot Concepts) with terms like "gene expression(987)", "Human(911)", "aging(834)", "angiogenesis(429)", "biomarker(410)", "DNA methylation(367)", "teaching In Astronomy(357)", "antioxidant(339)", "evolution(338)", and "nuclear Physics(326)"; and "来源期刊" (Source Journals) with terms like "BMC Bioinformatics(5035)", "BMC Cancer(4453)", "BMC Public Health(4185)", "Indian Journal of Ophthalmology(4102)", "BMC Research Notes(3703)", "BMC Infectious Diseases(3635)", "Journal of Medical Case Reports(3278)", "Progress In Electromagnetics Research(268)", and "BMC Musculoskeletal Disorders(2498)".



# Progress report: Websites

- Web information (Web Archive)
  - Finished the archive system development
  - 2000+ institutions websites Will be Harvested periodically

国际重要科研机构信息Web存档

首页 站点浏览 系统简介

搜索

### Web存档信息查找

INTERNET ARCHIVE  
**WayBackMachine**

http://

### 存档情况统计

<b>228</b> 存档站点个数	<b>17T</b> 存档数据总量	<b>118,510,377</b> 存档总页面数
能源 9,407,707 / 29 <input type="button" value="+MORE"/>	生物与环境 22,994,599 / 39 <input type="button" value="+MORE"/>	食品与健康 8,518,731 / 22 <input type="button" value="+MORE"/>
存档时间: 2014-03-27—2015-08-19	存档时间: 2014-07-02—2015-08-24	存档时间: 2014-07-07—2015-07-30
欧盟生物燃料技术平台 The European Bio...	气候和能源解决方... Center for Climat...	国际食品科技联盟 International Uni...
		英国研究理事会 Research Council...

## Current progress

---

- IR Grid : another important part
  - Has made good progress  
<http://www.irgrid.ac.cn/>
  - integrated 98 IRs 、 items=653,089、 Full text=500,566
  - Journal articles(>75%), dissertation/thesis, conference papers, patent documents, presentations, report, etc.





[ 浏览 ] BROWSE

研究所知识库列表

- ▶ 学科主题
- ▶ 内容类型
- ▶ 所有条目



[ 分类导航 ]  
FACETED BROWSING

▶ 研究所知识库列表 + MORE

高能物理研... [53797]  
工程热物理... [37890]  
金属研究所 [28225]  
物理研究所 [26498]  
长春光学精... [23679]

▶ 作者 + MORE

zhang [2728]  
li [2499]  
wang [2494]  
liu [2142]  
others [1749]

▶ 内容类型 + MORE

期刊论文 [432801]  
会议论文 [62437]  
学位论文 [55622]

## 最新采集 Recent Additions

- ▣ [20世纪80年代末以来中国土地利用变化的基本特征..](#)
- ▣ [20世纪中国知名科学家学术成就概览 力学卷](#)
- ▣ [New ent-kaurane diterpenoids from Isodon rube..](#)
- ▣ [Activities of prenylphenol derivatives from f..](#)
- ▣ [Two new guaianolides and a new daucene deriva..](#)
- ▣ [Contribution of a combination of ponacidin an..](#)
- ▣ [Effects of typhoon Matsa on ionospheric TEC](#)
- ▣ [Asymptotic and numerical solutions of the ini..](#)
- ▣ [New determination of the position of the puls..](#)
- ▣ [A new Legendre-type polynomial and its applic..](#)

CAS IR GRID 是定期采集中国科学院研究所IR的数据建立的联合机构知识库服务网络。

## 成员单位 Communities & Collections

选择成员单位并浏览其专题

- 力学研究所
- 软件研究所
- 地质与地球物理研究所
- 植物研究所
- 计算技术研究所
- 上海天文台
- 上海微系统与信息技术研究所
- 华南植物园
- 半导体研究所
- 地理科学与资源研究所
- 遥感应用研究所
- 古脊椎动物与古人类研究所
- 化学研究所
- 上海有机化学研究所
- 上海光学精密机械研究所
- 昆明动物研究所

## 中国科学院开放获取政策问答

### 2014年度热门下载论文

#### [ 利用排行 ]

浏览		下载
总排行	年排行	月排行
1.	<a href="#">Influences of aerodynamic load..</a>	[17]
2.	<a href="#">Effects of heavy metals on pla..</a>	[16]
3.	<a href="#">监利县淡水鱼资源综合开发研究</a>	[11]
4.	<a href="#">氯化挥发法综合处理含金高铁硫酸..</a>	[9]
5.	<a href="#">River Water Quality Assessment..</a>	[8]
6.	<a href="#">水生生物毒性试验方法</a>	[8]
7.	<a href="#">巧用工具，助力科研ENDNOTE X7介..</a>	[8]
8.	<a href="#">Physical modeling of local sco..</a>	[8]
9.	<a href="#">Comparative mechanisms of phot..</a>	[8]
10.	<a href="#">Underlying principle of effici..</a>	[8]

#### [ 汇总数据 ]

集成IR: 98  
 数据总量: 653089      全文总量: 500566  
 英文数量: 288953      全文开放量: 244729  
 累积浏览量: 99874341      累积下载量: 12470219  
 院外浏览量: 97406678      院外下载量: 11923095  
 国外浏览量: 18503399      国外下载量: 5033746

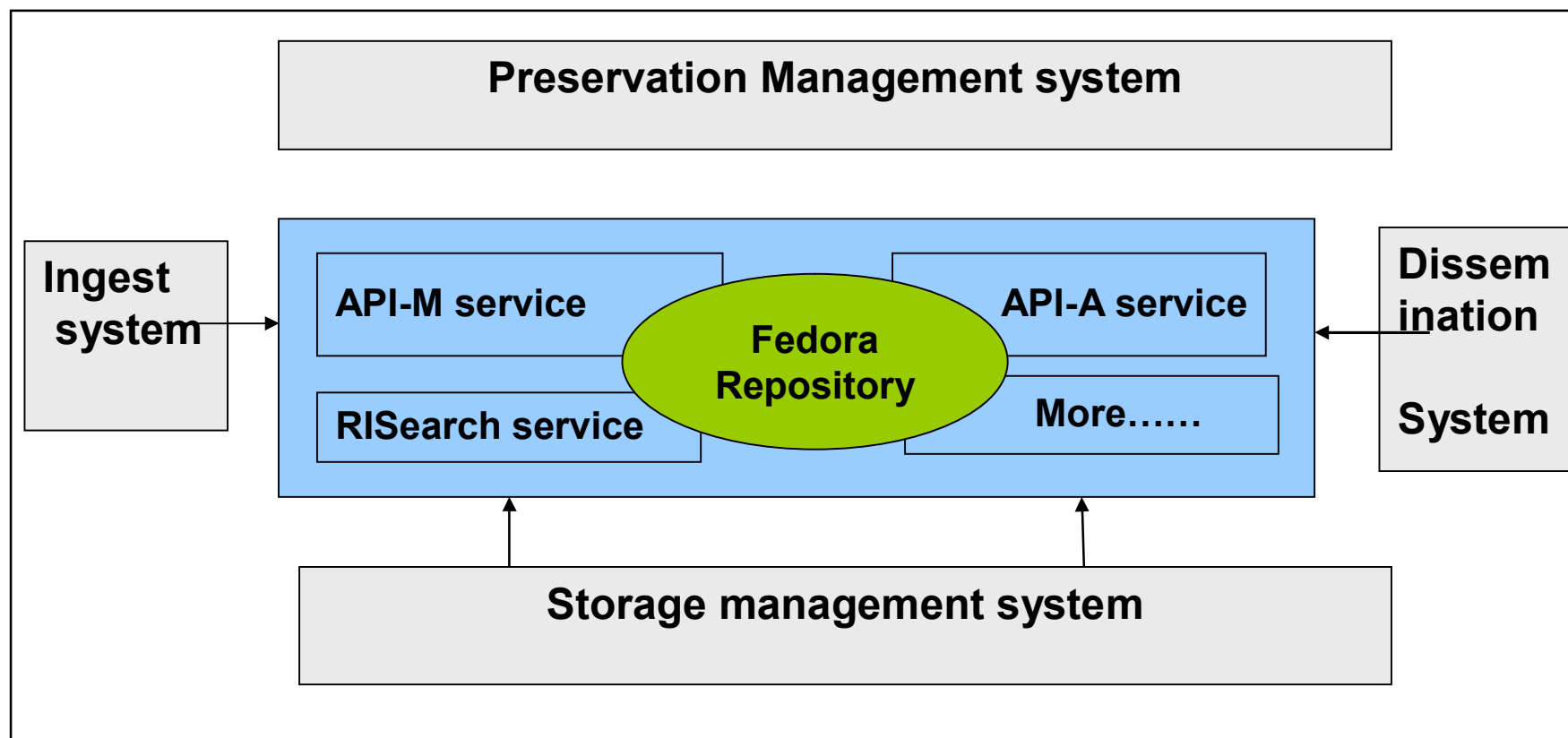


# Outline

---

- Archiving Activity at the NSLC
  - Part of the Chinese Academy of Sciences
- NSLC Digital Preservation System (brief view)
- Constructing a National Co-operation Preservation Network [Pilot Project]

## System framework based on OAIS





# Digital Preservation System of NSLC

---

- Three parts of preservation platform
  - Ingest & preservation management platform
  - A public access platform after a trigger event
    - Browse\search\download
  - A management platform for cooperator
    - Audit archived data
    - Check ingest report
    - Check statistic report
    - Monitor the running status of platform

Preservation Strategic Plan      Cooperation Strategic Plan

SIP Submission

Receipt

Check & Process

Ingest

Backup

Preservation Agreement Management

Audit & Certification

Data Object Management

Change Monitor

Format Management

Statistics & Report

User Management

Agreement & Subscription Management

Audit & Certification

Archive Report

Browse

User-Log Management

Public Access Report



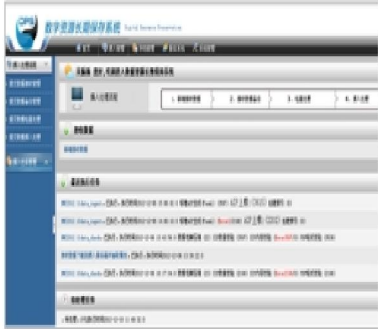
Browse

Retrieval

Full Text Download

Access Monitor

Public Access Statistics



Access Rights Management Based On Subscription Agreement

Trigger Event

Access Rights Management Based On Subscription Agreement

Publishers

Cooperative Preservation Institution





# Constructing a National Co-operation Preservation Network

---

- A pilot project funded by Ministry of Science and Technology
- 2014-2015
- Focus on STM literatures preservation



## Constructing a National Co-operative Preservation Network

---

- Geared to national strategic needs, meeting a national responsibility
  - Agree a common resource policy
  - Share preservation activity
- Workgroups meet challenges and resolve problems
  - Strategy and planning; negotiation and copyright management; administration; technology&system; audit and certification



## Constructing a National Co-operative Preservation Network

---

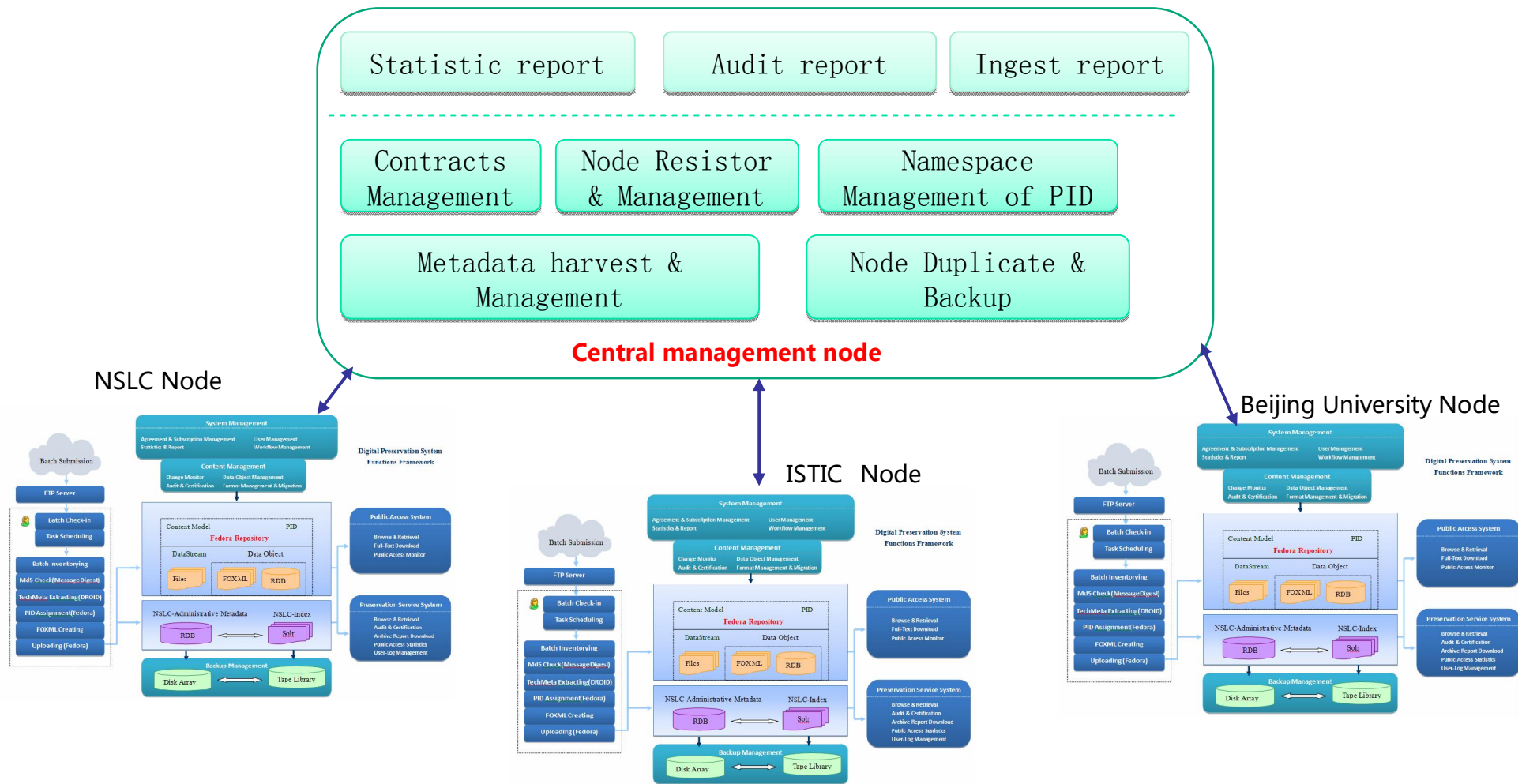
- Ministry of science and technology provides some financial support
  - Mainly for hardware, software, system operation and maintenance , and personnel training
- Other costs will be covered by the archive institutions themselves



## Constructing a National Co-operative Preservation Network

- The technology system and network for preservation
  - One central management node and multiple archive nodes
  - Each node in a different institution
  - Each archive node acts independently: collection and functionally
  - All deploy NSLC digital preservation system
  - They are each other's backup

# A National Co-operation Preservation Network

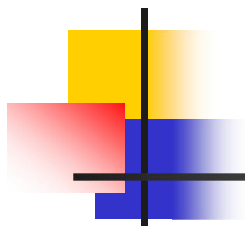


# Constructing a National Co-operation Preservation Network

- Central management node is in building
  - Contracts registry
  - Nodes registry
  - PID namespace management
  - Metadata of archived data harvest
  - Duplicate & Backup between nodes
  - Ingest report\Statistic report\Audit report

# Constructing a National Co-operation Preservation Network

- Currently three nodes in this network
  - NSLC: ours
  - ISTIC, Institute of Scientific and Technical Information of China
    - Began in January 2015
    - Retrospective data
  - Beijing University Library
    - Will start in September 2015 (this month)
    - Resources in basic sciences, interdisciplinary fields, and high-tech fields, social science



# Thank You !