

国家数字科技文献资源长期保存体系

National Digital Preservation Program

郑建程

国家科技图书文献中心

NDPP项目组

2018年中国高校图书馆发展论坛

昆明 · 2018年6月7日

引言

数字文献资源长期保存是一种风险防范，需要确保被保存资源在任何技术、经济、市场和管理条件下长期可获得和可使用。数字文献长期保存又是一项涉及复杂权利、复杂技术、复杂过程和大量经济与管理投入的长期工作。



https://www.unesco.nl/sites/default/files/uploads/Comm_Info/digital_roadmap_report.pdf

国家数字科技文献资源长期保存体系

National Digital Preservation Program

- 2013年科技部批准建设
- 国家科技图书文献中心（NSTL）资助、领导和管理
- 首批保存机构
 - 中国科学院文献情报中心
 - 中国科学技术信息研究所
 - 北京大学图书馆
- 国内参与机构
- 实现保存资源

确认签署《数字文献资源长期保存共同声明》的**216**个学术机构和图书馆。

中文期刊 >18,000 种，中文期刊论文 >5,000 万篇
外文期刊 >4,000 种，外文期刊论文 >440 万篇
外文电子书 42324 本，外文事实性数据库 >50,000 篇
预印本库 (arXiv) 1,389,474 篇

国家数字科技文献资源长期保存体系

National Digital Preservation Program

一、建设背景

二、主要战略

三、实施机制

四、后续工作

一、建设背景

- 数字文献已经成为学术信息的主流和首要形态
 - e-only 文献资源的增长
 - 学术信息资源开放获取（Open Access）成为许多领域的主流，OA文献资源不再有对应的印刷形态
- 数字文献已经成为学术机构的主流和首要“馆藏”
 - 从用户使用角度的“馆藏”
 - 图书馆的各类电子文献（期刊、图书、会议录、学位论文等）
 - 开放获取的各类文献资源、网络信息资源
 - 从图书馆采购角度的馆藏
 - 资源形态从p+e 向e-first+p，再向e-only的发展
 - 图书馆特别是研究型图书馆采购电子资源的经费比例愈来愈大

一、建设背景

- 数字文献已经成为学术机构的主流和首要“馆藏”
 - 从图书馆采购角度的馆藏
 - 中科院系统（赵艳，201710），建成以数字资源为主的保障体系（数据库171个，外文电子期刊1.82万种，外文电子图书17.7万种，外文电子学位论文54.31万篇，中文电子期刊1.7万种，中文电子图书35万余种，中文电子学位论文279万篇）
 - 高校CALIS系统（DRAA，肖珑，201805），累计集团引进数据库165个，2017年参团8754馆次，全文数据库占67%。

一、建设背景

- 数字文献仅仅是数字知识资源的一小部分
 - 数字化是信息与知识的新常态
 - 数字资源是互联网、物联网和新业态基础
 - 科学数据、数字科研资源
 - 数字图像和音视频资源
 - 新闻资源、社交媒体资源
 - 开放教育资源
 - 数字人文资源
 - 电子政务资源
 - 个人和社群数字资源.....



INTERNET TRENDS 2016 – CODE CONFERENCE

Mary Meeker
June 1, 2016

kpcb.com/InternetTrends

KPCB | KLEINER
PERKINS
CAUFIELD
BYERS

Data = A New Growth Platform...
Powering New Services / Systems / Apps

一、建设背景

■ 数字内容长期可用面临的挑战

- 介质的长期可靠性
- 格式的长期可靠性（不断变化的数据格式）
 - 谁能读出过时的数据格式？
 - 能否读出“富媒体数字对象”包含的全部内容？
- 内容的完整性
 - 什么是“内容”实体？如何保障其“动态化完整性”
- 内容的可使用性
 - 可获得，可理解，可互操作，可复用
- 内容的长期可获得性
 - 技术系统、网络链接、提供者管理、“政治”的可靠性
 - 自然灾害、人为破坏

一、建设背景

■ 数字资源长期保存面临的挑战

- 订购模式只获得网络使用权（信任出版商？）
 - 订阅结束怎么办？受到许多威胁性的摆布。
- 永久获取权（perpetual access rights）
 - 年度付费、系统升级、权属变更、地缘政治等
- 本地镜像站点（local mirror sites）
 - 少数（法理上技术上由出版商控制，没有保存和服务权利）
- 本地裸数据（local raw data）
 - 难以获得权利，成本，持续可服务性
 - 采购合同的限制（禁止系统化下载和拷贝），成本？

一、建设背景

- 数字资源长期保存面临的挑战
 - 数字内容实质上处于濒于损毁的状态。
 - 数字知识内容长期保存已成为关键战略需求。

一、建设背景

■ 数字信息资源长期保存的多种机制

The Keepers Registry

Supporting long-term access to journal content

[Home](#) | [Journals](#) | [Publishers](#) | [Archiving Agencies](#) | [Member Services](#) | [FAQ](#) | [About](#) | [Help](#)

Discover who is looking after your e-journals

Search

Input a search term (e.g. free text, or a title or ISSN(s))

Search

Search examples

ecology

finds records mentioning the word *ecology*

science medicine

finds only those records mentioning both science and medicine

science or medicine

finds records mentioning either science or medicine, or both

"science and medicine"

(including the quotation marks) finds records containing the exact phrase *science and medicine*

- Hide search examples

+ Show search examples

Glossary and definitions

The following organisations are the Keepers of digital content, working on your behalf to ensure long-term access to the scholarly and cultural record. They provide the registry with information on their archival holdings, ordered by most recent update (date of which is shown):

- [Archaeology Data Service](#) (24 May 2018)
- [Global LOCKSS Network](#) (23 May 2018)
- [CLOCKSS Archive](#) (22 May 2018)
- [Cariniana Network](#) (21 May 2018)

Working Together

Find out how publishers, research libraries and national libraries can support archiving initiatives to ensure the future of the digital scholarly record. [Read more.](#)

Current Statistics

- Serials with volumes reported as 'ingested and archived' by at least one Keeper: 38399
- Serials with volumes reported as 'ingested and archived' by three or more: 14888

Work on other serials (and volumes) is 'in progress'.

News and Events

- 06 Nov 2017: [Webinar - Introduction to the Keepers Registry](#)
- 01 Nov 2017: [Working Together to Ensure the Future of the Digital Scholarly Record](#)
- 26 May 2017: [Swiss National Library joins the Keepers Registry](#)
- 21 Dec 2016: [IARLA endorses Working Together statement](#)

The Keepers Registry

Supporting long-term access to journal content

[Home](#) | [Journals](#) | [Publishers](#) | [Archiving Agencies](#) | [Member Services](#) | [FAQ](#) | [About](#) | [Help](#)

Information about archiving agencies

The term 'Archiving Agency' represents an initiative that is running a programme for the metadata on the journals in their programme available to the Keepers Registry. The data information obtained from the ISSN Register.

The Keepers Registry is collecting data from the following archiving agencies.

- ↓ [Archaeology Data Service](#)
- ↓ [British Library](#)
- ↓ [Cariniana Network](#)
- ↓ [CLOCKSS Archive](#)
- ↓ [e-Depot](#)
- ↓ [Global LOCKSS Network](#)
- ↓ [HathiTrust](#)
- ↓ [Library of Congress](#)
- ↓ [National Science Library, Chinese Academy of Sciences](#)
- ↓ [PKP Preservation Network](#)
- ↓ [Portico](#)
- ↓ [Scholars Portal](#)
- ↓ [Swiss National Library](#)

Each participating agency has provided a description of their organisation's approach to access conditions. In order to understand each archiving agency in detail, it is recommended to visit their websites.

一、建设背景

■ 数字信息资源长期保存的多种机制

■ 荷兰国家图书馆, KB

- E-depot with Legal deposit into KB by “domestic” publishers

■ British Library, BL

- eSerials and Newspapers; Legal Deposit Libraries (Non-Print Publications) Regulations 2013

■ German National Library (DNB)

- AREDO; eSerials

■ Library of Congress

- eSerials (including e-only journals) as legal deposit

国家级机构, 国家公共投入, 多利用法定存缴, 多侧重本国出版物, 在公共服务上有明确限制。

一、建设背景

■ 数字信息资源长期保存的多种机制

■ 荷兰国家图书馆，KB

■ Access Policy:

- Provide on-site access to all archived publications of the publishers
- In specific circumstances (trigger events) access can be broadened to a wider audience:
 - when a publisher stops operations, or
 - when a publisher ceases to publish a title, or
 - when a publisher no longer offers back issues, or
 - upon catastrophic and sustained failure of a publisher's platform

一、建设背景

■ 数字信息资源长期保存的多种机制

■ 荷兰国家图书馆，KB

■ Financial Support:

- Internal funding under Acquisition & Processing, and IT
- Earmarked funding from Ministry of E+C+S

■ Legal Support:

- Agreements with Dutch Publishers Association and Publishers

一、建设背景

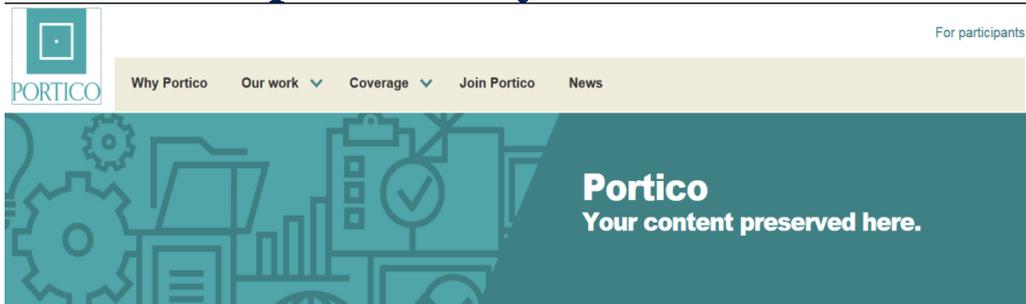
■ 数字信息资源长期保存的多种机制

■ Portico第三方机制

■ A digital DARK ARCHIVE service provided by ITHAKA

■ Governance

- Board of Trustees
- Advisory committee



Libraries and publishers can be confident that their e-journals, e-books, and digitized collections will be accessible to researchers, scholars, and students in the future.

89 million
preserved items



Titles and collections

Download e-journal holdings in [Excel](#), [KBART](#), or [ONIX-PH](#) and e-book holdings in [Excel](#) or search all content below. You may also [request a holdings comparison](#).

Search titles and ISSNs by keyword:

Show: E-journals E-books Digital Collections

28,644
e-journals committed

1,241,499
e-books committed

187
collections committed

NEWS

All News

May 22, 2018
Universitas Islam Indonesia is now preserving 25+ OA e-journals with Portico

May 15, 2018

Search for titles in the Portico archive

Search titles by keyword:

E-journals E-books Collections

The Online Journal of
CURRENT CLINICAL TRIALS

Help us complete a very early online publication

Do you have print or digital copies of articles from The Online Journal of

一、建设背景

■ 数字信息资源长期保存的多种机制

■ Portico第三方机制

■ 价格机制

- 图书馆按年度电子期刊、电子图书采购支出分档缴费
- 出版社按年度电子期刊、电子图书销售收入分档缴费

Total LME	Annual Archive €	Total LME	Annual Archive Support
\$25-30 million	\$25,462	\$25-30 million	\$7,638
\$20-25 million	\$20,157	\$20-25 million	\$6,047
\$15-20 million	\$18,566	\$15-20 million	\$5,570
\$13-15 million	\$17,080	\$13-15 million	\$5,124
\$11-13 million	\$16,126	\$11-13 million	\$4,838
\$9-11 million	\$15,171	\$9-11 million	\$4,552
\$7-9 million	\$13,792	\$7-9 million	\$4,138
\$5-7 million	\$12,200	\$5-7 million	\$3,661
\$4-5 million	\$10,609	\$4-5 million	\$3,183
\$3-4 million	\$9,442	\$3-4 million	\$2,833
\$2-3 million	\$8,275	\$2-3 million	\$2,482
\$1-2 million	\$7,108	\$1-2 million	\$2,132
\$750,000-1 million	\$5,729	\$750,000-1 million	\$1,719
\$500-750,000	\$4,456	\$500-750,000	\$1,337
\$250-500,000	\$3,289	\$250-500,000	\$987
\$150,000-250,000	\$1,591	\$150,000-250,000	\$478
<\$150,000	1.03% of LME	<\$150,000	30% of E-journal fee

Annual journals revenue	Annual archive contribution	E-books	
>\$200 million	\$81,960	Annual fees for e-books are based on total e-book revenues (electronic sales, licensing, and advertising) as follows:	
\$100-200 million	\$54,635		
\$50-100 million	\$27,320	Annual e-book revenue	Annual archive contribution
\$10-50 million	\$16,440	\$10-50 million	\$15,000
\$5-10 million	\$5,460	\$5-10 million	\$10,000
\$1-5 million	\$2,735	\$1-5 million	\$7,500
\$500k-1 million	\$1,085	\$500k-1 million	\$5,000
\$250-500k	\$500	\$250-500k	\$2,500
\$0-250k	\$250	<\$250k	\$1,000

一、建设背景

■ 数字信息资源长期保存的多种机制

■ Portico第三方机制

■ 使用（服务）机制

■ Triggered Content

- Participating libraries, including their clients, gain access to preserved content when specific titles no longer available from the publisher or any other source.

■ Trigger events include:

- Cessation of a publisher's operations
- Discontinuation of a title by a publisher
- Back issues no longer offered by a publisher
- Catastrophic and sustained failure of a publisher's delivery platform

一、建设背景

■ 数字信息资源长期保存的多种机制

■ Portico第三方机制

■ **Triggered titles**

The following titles, in whole or in part, have been triggered and are available to all Portico participants through our website:

- [Advances in Tumor Virology](#) (*Open Access*)
- [Annals of Clinical Psychiatry](#)
- [Arbitration Law Reports and Review](#)
- [Astronomy Education Review](#) (*Open Access*)
- [Australian and New Zealand Journal of Audiology](#)
- [Auto/Biography](#)
- [BoneKEy Reports](#)
- [Brief Treatment and Crisis Intervention](#)
- [Cell & Tissue Transplantation & Therapy](#) (*Open Access*)
- [Cell Communication Insights](#) (*Open Access*)
- [Clinical Medicine Insights: Geriatrics](#) (*Open Access*)
- [Clinical Medicine Reviews in Cardiology](#) (*Open Access*)
- [Clinical Medicine Reviews in Oncology](#) (*Open Access*)
- [Clinical Medicine Reviews in Patient Care](#) (*Open Access*)

一、建设背景

■ 数字信息资源长期保存的多种机制

■ CLOCKSS联盟型机制

■ Community-governed

- Supporting libraries
- Participating publishers

■ Distributed preservation

- 12 nodes around the world

■ Non-profit organization

- Board of Directors
- Advisory Council

■ Supporting Library Agreement

■ Participating Publisher Agreement



HOME

The CLOCKSS Archive

A Trusted Community-Governed Archive

[CLOCKSS Announces the Participation of Additional Publishers](#)

[CLOCKSS Archive and Code Ocean Announce Preservation Partnership](#)

[Joint Statement from Collaborating Digital Preservation Organization](#)

CLOCKSS Makes Guidelines Available:

Web Harvest Guidelines

■ [CLOCKSS Web Harvest Guidelines](#)

File Transfer Guidelines

■ [CLOCKSS File Transfer Guidelines](#)

NAVIGATION

- [Home](#)
- [Benefits of Participating in CLOCKSS](#)
- [How CLOCKSS Works](#)
- [Frequently Asked Questions](#)
- [News](#)
- [How to Participate](#)
- [Supporting Libraries](#)
- [Participating Publishers](#)
- [Board of Directors](#)
- [Triggered Content](#)
- [LOCKSS \(Lots of Copies Keep Stuff Safe\)](#)

How CLOCKSS WORKS

CLOCKSS is a not-for-profit joint venture started by libraries and publishers committed to ensuring long-term access to scholarly publications in digital format. As libraries migrate from print to online-only publications, they expect assurances from publishers that their shared investments are protected and preserved for generations to come. The CLOCKSS archive provides this assurance via its secure network of content that can be accessed only when a trigger event is deemed to have occurred. CLOCKSS is unique because it makes all content triggered from the archive freely available to the world.

Technical Overview

The following is a step-by-step overview of how CLOCKSS works.

一、建设背景

- 数字信息资源长期保存的多种机制
 - CLOCKSS联盟型机制
 - CLOCKSS价格机制（按年度分档缴费）

Supporting Libraries

- Libraries Materials Budgets -	- Annual Contribution -
Under \$1 million	\$465
\$1 - 2 million	\$615
\$2 - 3 million	\$1,235
\$3 - 4 million	\$1,855
\$4 - 5 million	\$2,470
\$5 - 7 million	\$3,090
\$7 - 9 million	\$4,320
\$9 - 11 million	\$5,560
\$11 - 13 million	\$6,795
\$13 - 15 million	\$8,030
\$15 - 20 million	\$9,265
\$20 - 25 million	\$12,350
Over \$25 million	\$15,440

Participating Publishers

The 2018 annual contribution below covers e-journals, e-books, and any other materials you'd like to archive.

- Total Publishing Revenue -	- Annual Contribution -
Under \$250,000	\$232
\$250,000 - 500,000	\$465
\$500,000 - 1 million	\$1,090
\$1 - 5 million	\$2,180
\$5 - 10 million	\$4,360
\$10 - 15 million	\$6,540
\$15 - 20 million	\$8,720
\$20 - 30 million	\$10,900
\$30 - 40 million	\$13,080
\$40 - 50 million	\$15,260
\$50 - 200 million	\$16,365
Over \$200 million	\$27,275

一、建设背景

■ 数字信息资源长期保存的多种机制

■ CLOCKSS联盟型机制

■ CLOCKSS使用（服务）机制

How does the CLOCKSS board define a trigger event?

Trigger events include situations of non-availability of archived content in which:

Publisher No Longer in Business

The publisher is no longer in business or is no longer in the business of publishing content or providing access to previously published content and there are no successor interests or reversions or transfers of rights;

Title No Longer Offered

The publisher has stopped publishing and is no longer providing access to the content and there are no successor interests or reversion or transfer of rights;

Back Issues No Longer Available

The publisher has stopped offering or providing access to some or all of the back issues of the content and there are no successor interests or reversion or transfer of rights; or

Catastrophic Failure

While still publishing content, the publisher is not able to provide access to the content electronically due to technical or similar catastrophic and permanent failure.

Advances in Tumor Virology

Libertas Academica and SAGE Publications
eISSN: 1179-5654. Triggered: 2017.

Annals of Clinical Psychiatry

Taylor & Francis

Print ISSN: 1040-1237, eISSN: 1547-3325. Triggered: N

Archives of Family Medicine

The American Medical Association (AMA)

eISSN: 1063-3987. Triggered: January 2012.

Autism Insights

Libertas Academica

eISSN: 1179-5964. Triggered: February 2015.

Auto/Biography

SAGE Publications

Print ISSN: 0967-5507, eISSN: 1479-9715. Triggered: A

BoneKEy Reports

The International Bone & Mineral Society

Print ISSN: 2047-6396. Triggered: 2017.

Brief Treatment and Crisis Intervention

Oxford University Press

Print ISSN: 1474-3310, eISSN 1474-3329. Triggered: D

Cell & Tissue Transplantation & Therapy

Libertas Academica and SAGE Publications

eISSN: 1179-0687. Triggered: 2017.

Cell Biology Insights

Libertas Academica

eISSN: 1178-6272. Triggered: 2015.

Cell Communication Insights

Libertas Academica and SAGE Publications

国家数字科技文献资源长期保存体系

National Digital Preservation Program

一、建设背景

二、主要战略

三、实施机制

四、后续工作

二、主要战略

(1) 保存什么资源——保存资源规划与选择

- 数字信息体系的复杂内容
 - 学术信息、社会信息、商业信息、娱乐信息.....?
 - 数字文献、网站内容、科学数据、数字音视频.....?
 - 原生数字文献、数字化文献.....?
 - 正式出版物、非正式出版物.....?
 - 商业出版物、开放获取出版物.....?
 - 期刊、图书、会议录、专利.....?
- 选择因素
 - 服务对象、机构责任、采购范围、合作可能.....
 - 资源重要性、资源使用度、资源可用性风险.....
 - 权益可获得性、技术成熟性、管理成熟性.....
 - 政治考虑：本土、可控、权益平衡.....

二、主要战略

(2) 如何保存——长期保存机制选择

- 国家法定存缴（BL、LC、KB、DNB等）
 - 无力于国外文献；缺乏刚性法律保障？多样性挑战
- 国家长期保存体系
 - 作为国家资源保障体系的一部分（NSTL机制？）
- 联盟型或商业化长期保存体系
 - CLOCKSS, Europeana, Portico
- 专业或机构知识库、专门数字化文献库
 - PMC/PMC International/arXiv.org/HathiTrust
-
- 如何选择、否定因素是什么？
 - 本土化、可控性、成本？

二、主要战略

(3) 保存什么内容——保存内容规划与选择

- 数字信息保存的深度
 - 保存bits & bytes — 介质保存，保证可读出
 - 保存格式 — 内容结构保存，保证可读懂
 - 内容描述和组织信息 — 元数据（多层次）
 - 内容关联关系与结构保存 — 内容细粒结构和关联机制，保证可读懂内容关系和内容关联网络网络
 - 内容处理机制保存 — 保存相应工具和平台，保存内容呈现机制与结果
 - 内容管理机制保存 — 保存内容组织、处理和服务的管理要求与机制，支持可持续复用
- 选择因素：重要性、可授权性、技术复杂度、成本……
- 面对不同数字内容的不同选择

二、主要战略

(4) 有权进行保存吗——长期保存权利保障

- 数字文献资源长期保存权
 - 保存权；保存处理权；保存服务权；合作保存权。
- 长期保存权的来源
 - 采购资源本身就意味着长期保存权？（是的！采购的是长期可利用的知识，不是采购纸张或临时可用的知识）
 - 专门法律保障
 - 著作权保护限制与例外→图书馆保存权
 - 网络信息传播保护条例→图书馆保存权
 - 合同安排：与出版社的长期保存协议
 - 市场博弈
- 通过具有法律效力的机制落实权益和获取权益的组织过程
- 诚信尽责保护出版社的合法权益

二、主要战略

(5) 谁来保存——合规保存机构

- 通过分工合作的合规保存机构来具体实施期保存
- 合规机构条件：
 - 公共事业法人；
 - 本身采购数字资源并提供服务；
 - 支持NDPP长期保存战略目标和运行管理要求；
 - 承担或参加被保存资源的保存谈判；
 - 承担长期保存的法律责任、接受公共认证审计；
 - 触发事件下提供被保存资源的公共服务，支持特殊情况下继承保存；
 - 遵守长期保存体系的法律、技术和管理规范；
 - 提供长期保存系统的基础设施、人员和管理支持；
 - 互为备份支持；
 - 当兼顾行业和地区分布。

二、主要战略

(6) 采用什么技术体系——OAIS参考模型

ISO 14721:2012 (CCSDS 650.0-P-1.1) Preview

Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model

This standard was last reviewed and confirmed in 2018. Therefore this version

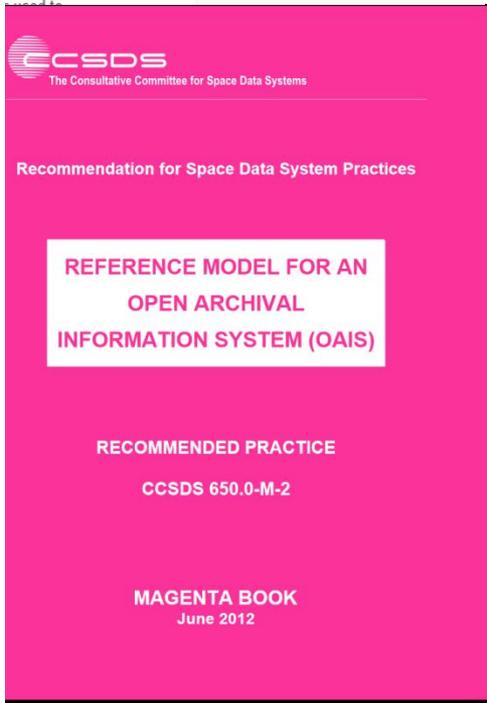
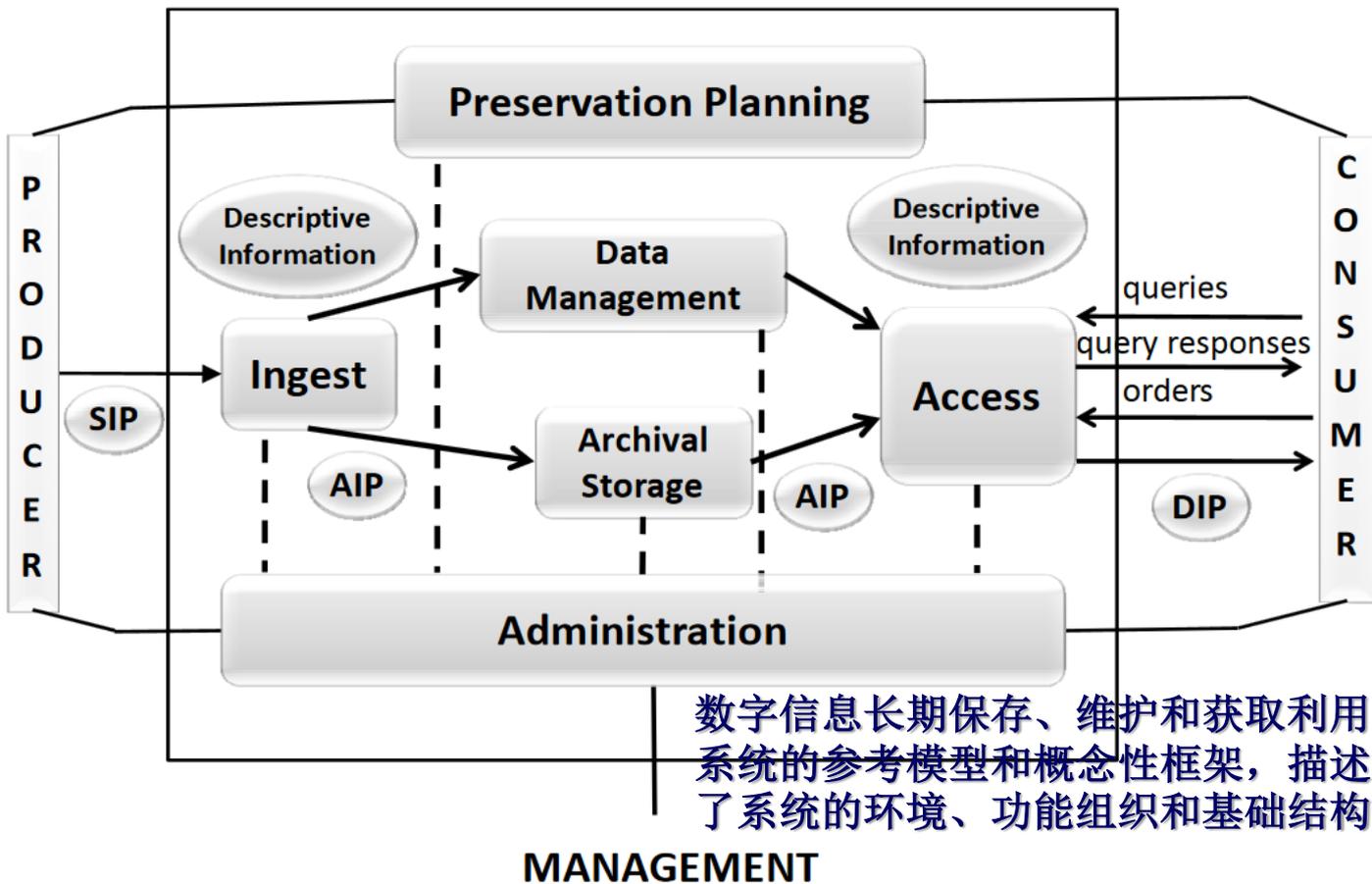
Buy this standard

Format

PDF English

Paper English

CHF 198



二、主要战略

(7) 什么情况下能向谁提供服务——触发条件

- 长期保存本质上是Dark Archive，只在资源提供者无法提供正常服务时才由保存者向订购者提供检索和获取服务
- 服务触发条件（Trigger Events）
 - **Publisher No Longer in Business**
 - **Title No Longer Offered**
 - **Back Issues No Longer Available**
 - **Catastrophic or sustained failures**
 - **(no successor interests or reversion or transfer of rights)**
 - 是否应包括停止订购？
- 触发事件的后的服务：被保存内容的检索与获取
- 服务管理：用户管理与使用审计，非法使用监控
- 触发条件消失后处置：触发条件消失的确认，提供者恢复服务
- 什么时候可对全社会公开服务：市场完全失灵？

二、主要战略

(8) 保障长期保存本身的可靠性——可信赖性审计与认证

- 作为保障性系统本身必须可靠
 - 要保存的内容确实得到了保存、并能在需要时提供可靠服务
 - 整个保存过程和系统能长期可靠地做到上述要求
- 长期保存系统可靠性
 - 权益安排可靠性
 - 技术可靠性
 - 系统可靠性
 - 管理可靠性
 - 经济可靠性
 -

二、主要战略

(8) 保障长期保存本身的可靠性——可信赖性审计与认证

- 长期保存服务的可靠性审计
 - 审计什么：保存审计和保存机制审计
 - 审计标准：**ISO 16363 (CCSDS 652.0-M-1) : 2012** (Audit and Certification of Trustworthy Digital Repositories, 可信赖性数字存储系统审计与认证)，包括对技术基础设施风险管理、安全风险管方面的认证标准。（依据哪些技术、系统建设、信息安全标准，运行维护、管理，必要的升级、更新等）
 - 谁进行审计：自查审计、第三方审计、用户审计... ..
 - 审计细粒度：Journal Title、Volume、Issue、Article、Detailed Data?
 - 审计问题的问责机制？

二、主要战略

(9) 如何选择技术方法与平台——可靠系统

- 长期保存关键技术、流程与平台
- 确保可靠、安全、高效、可管理的长期保存系统
 - 保证数字媒介本身的可靠性
 - 保证数字内容本身的完整性
 - 保证摄入、保存和服务过程的准确、高效和可管理
 - 保证整体过程的规范性和可审计性
 - 保证整体系统的可迁移性和可互操作性

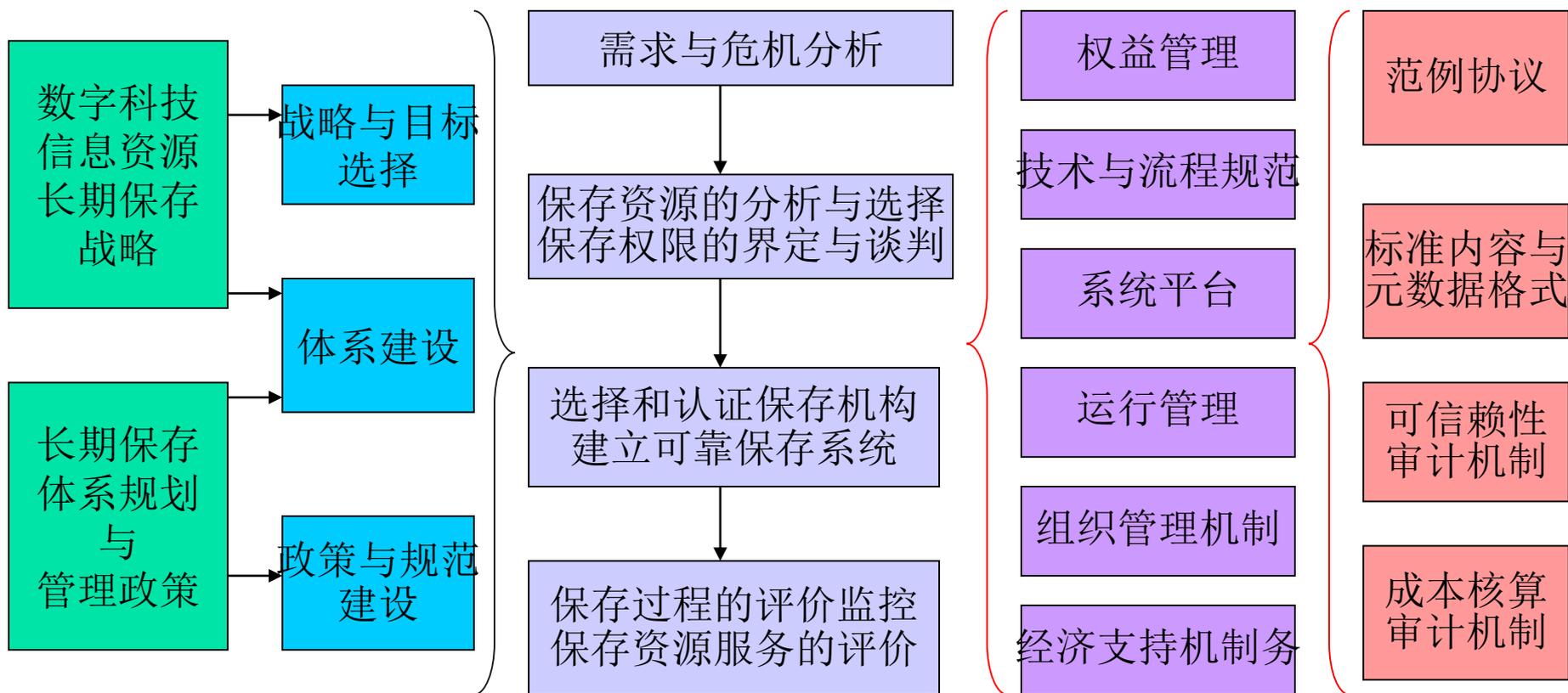
二、主要战略

(9) 如何选择技术方法与平台——可靠系统

- 确保技术与管理可靠、技术与运行可互操作
 - 遵循OAIS模型
 - 支持标准数据格式或格式识别与转换
 - 支持标准元数据描述或元数据转换与增强
 - 支持信息系统安全管理规范
 - 支持可信赖系统审计规范
 - 支持标准的检索获取协议及其使用监控与审计

二、主要战略

(10) 规划体系



国家数字科技文献资源长期保存体系

National Digital Preservation Program

一、建设背景

二、主要战略

三、实施机制

四、后续工作

三、实施机制

(1) 发展过程

- 2004 中科院和NSTL举办国际iPres会议
- 2004 NSTL启动数字化科技信息资源长期保存研究项目
- 2007 中科院启动数字科技文献长期保存服务系统建设，研究和试验数字期刊与数字图书长期保存
- 2009 中科院与Springer 签署长期保存协议并开始存储其电子期刊，随后又与NPG、Wiley等签署长期保存协议
- 2013 科技部正式同意由NSTL牵头组织实施NDPP
- 2014 NSTL开始国家数字科技文献长期保存体系建设
 - 系统建设、资源保存、规范规则制订
- 示范系统，2014，中科院节点、中信所节点
- 基本系统，2015-2016，扩展建设节点（北大）
- 体系完善，2017-2020，覆盖所有重要资源

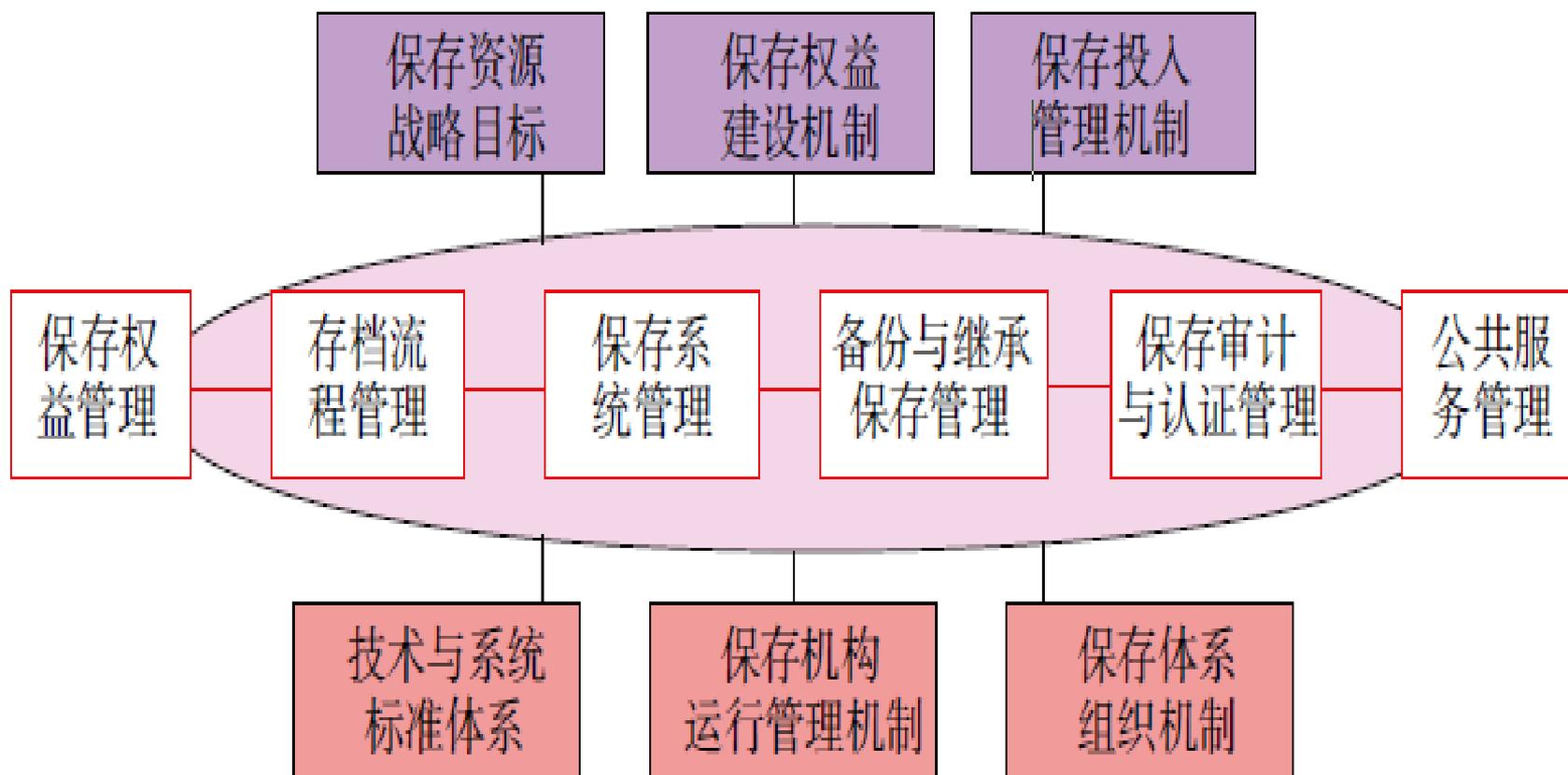
三、实施机制

(2) NDPP运行原则

- **国家主导**：作为国家科技文献保障平台的有机部分，由国家投资，由NSTL 管理；
- **联合参与**：动员国内图书馆合作参与，并选择若干图书馆作为“合规保存机构”联合承担保存任务；
- **可靠管理**：保证重要数字科技文献资源在大陆本地可靠长期保存，保证各个长期保存系统的可靠运行和可靠监管；
- **公共服务**：长期保存体系本身作为公益性公共服务由国家支持。

三、实施机制

(3) 总体实施体系



三、实施机制

(4) 保存目标

- 作为国家科技信息资源战略保障，长期保存我国科技创新用户所需的主要数字科技文献资源，同时积极参与国家教育文化社会各领域所需的其他数字文献资源的长期保存。
- 主要选择使用价值高、学术价值高、风险高、保存可操作性强的数字科技文献资源优先部署长期保存，并创造条件保存其他具有保存价值和消除风险作用的数字文献资源。
- **使用价值**：当前及长期需求量（采购机构和实际使用量）
- **学术价值**：资源在科学研究和科技发展中的长期学术价值、且保证系统性覆盖重要资源；
- **风险程度**：依赖程度、内容或使用渠道不可替代性及其风险程度；
- **可操作性**：权益安排可接受度、内容格式规范化程度、所需保存技术的掌握程度、保存机制的可负担性、合作协调能力等。

三、实施机制

(5) 保存资源

- 首要目标范围：自然科学、工程技术领域、交叉领域
- 其他目标范围：与自然科学、工程技术密切相关的管理、经济、行业、市场、金融等领域
- 目标资源类型：
 - 国际主要综合型学术出版商的科技期刊和科技会议录数据库及这类数据库的回溯数据内容；
 - 国际主要专业学会/协会和大学等专门出版商的科技期刊和科技会议录数据库及这类数据库的回溯数据内容；
 - 国内主要科技期刊、科技会议和学位论文数据库；
 - 主要的开放获取期刊、会议录、专业知识库等数据库；
 - 重要数字科技专著数据库和重要数字工具书资源；
 - 其他重要的数字知识资源（例如学位论文、科技报告等）。

三、实施机制

(6) 权利声明、获取与管理

- 2015年《数字文献资源长期保存共同声明》
 - <https://www.nstl.gov.cn/facade/news/newsInfo.do?act=toNewsContent&id=127244>
- 图书馆对所采购数字文献资源拥有长期保存的权利。
- 图书馆采购数字文献资源实质上是采购知识内容为用户提供当前和未来服务，因此长期可靠利用是这种采购行为的内在含义和基础条件。
- 没有可靠的长期保存，图书馆对所采购知识内容的长期利用权就不完整和不可靠，因此长期保存是图书馆持续进行资源采购的必要基础。
- 长期保存权包括：数据存档权、数据处理权、数据服务权、合作保存权和继承保存权。后两者是保存权的自然延伸，也是长期保存权利平等实现的基础。

三、实施机制

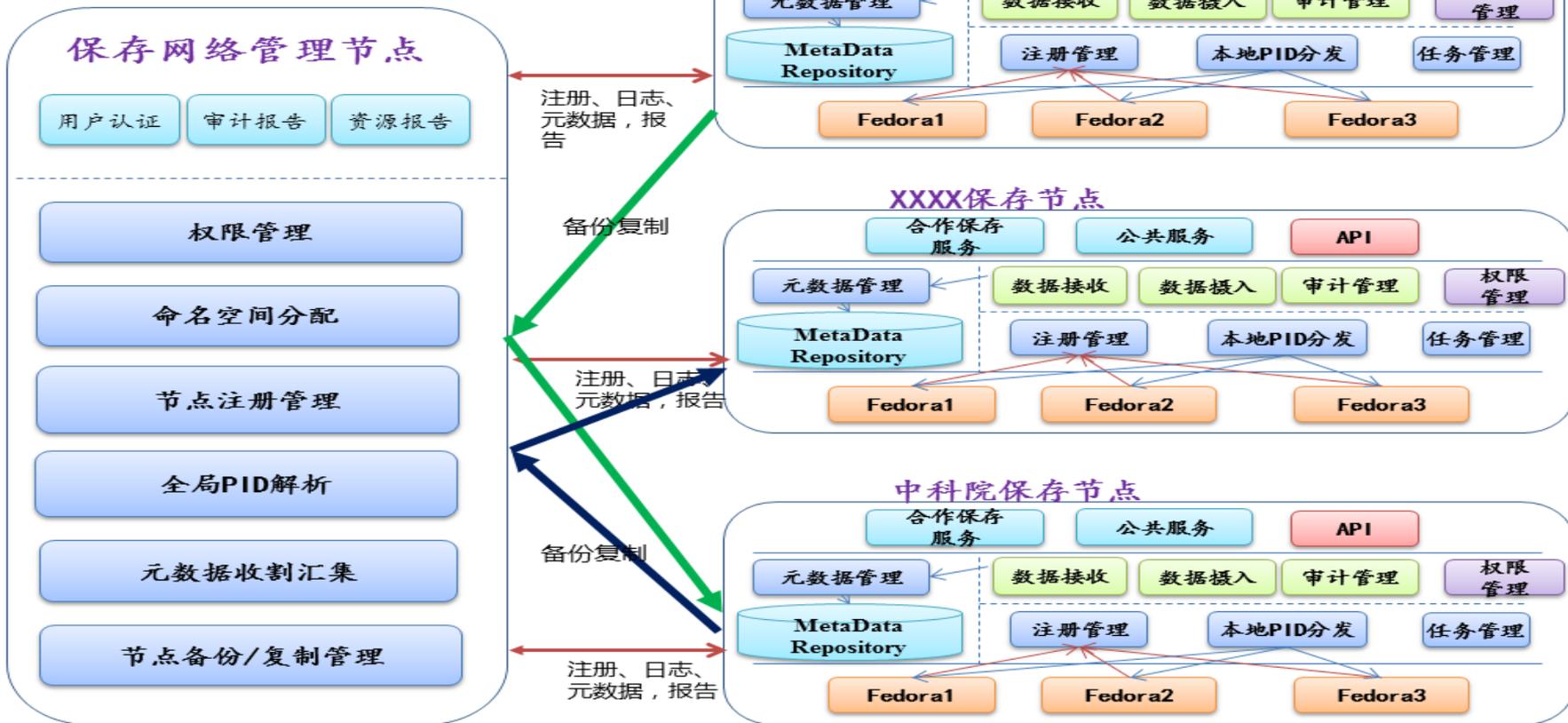
(6) 权利声明、获取与管理

- NDPP受NSTL委托、代表国家和签署共同声明的图书馆行使上述长期保存权。
- NDPP充分维护出版社的合法权益，承担诚信尽责保护出版商合法权益的义务，承诺建立可信赖的机制保证利益相关方遵守在长期保存中的权利和义务。
- NDPP建立行使长期保存权的制度化机制
 - 推动国家确立“使用公共资金采购数字文献资源的机构”长期保存责任，机构应要求所采购资源在中国本土得到可靠长期保存，应将长期保存权利纳入与出版社签署的采购合同中。
 - 权益谈判机制（委托或联合合规保存机构）
 - 根据分工向出版商提出长期保存要约；
 - 结合资源订购谈判进行长期保存协议谈判；
 - 签署长期保存协议、保存数据摄入技术协议，NDPP登记；
 - NDPP对长期保存协议的执行进行检查和保存审计。

三、实施机制

(7) 技术体系

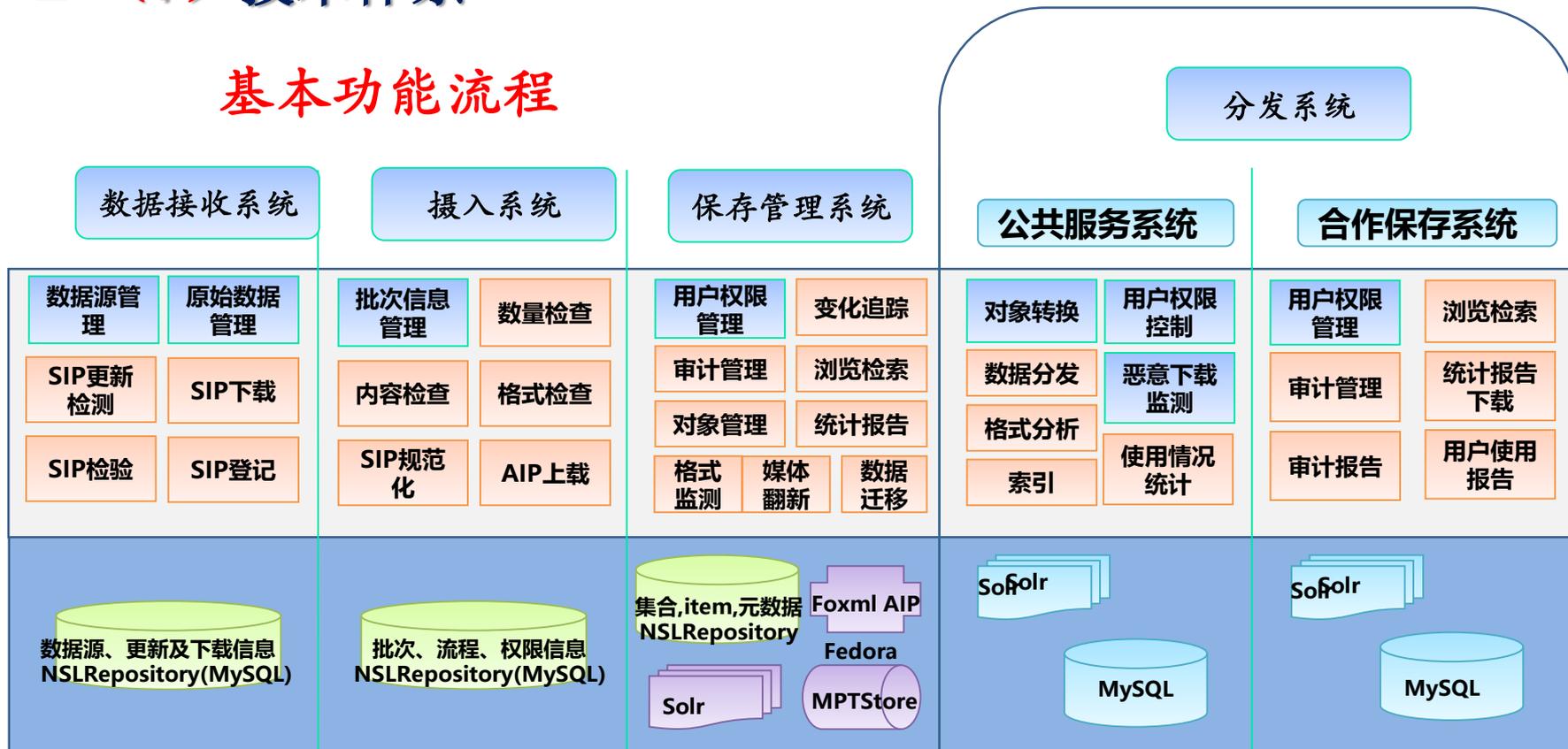
国家数字科技文献资源 长期保存网络



三、实施机制

(7) 技术体系

基本功能流程



分发系统

公共服务系统

合作保存系统

三、实施机制

(8) 已保存资源

- 维普中文期刊
- IOP期刊
- RSC期刊
- PNAS期刊
- OUP期刊
- Springer科技图书
- Springer回溯图书
- IOP图书
- Springer Protocols库
- NSTL以国家许可方式订购的学术期刊（45家出版社、600多种期刊）
- 万方科技资源数据库
- ProQuest学位论文、Emerald期刊
- Taylor & Francis期刊（协议谈判中）、CUP期刊库（协议签署中）
- Springer期刊
- BioMed Central期刊
- Nature期刊
- Wiley期刊
- AGU期刊
- RSC图书
- Wiley图书
- Springer Protocols回溯
- arXiv预印本

三、实施机制

(9) 可信赖审计与检查

- 保存机构保存进展报告
 - 月度报告：应保存多少、接收到多少、实际保存多少
- 保存机构自我保存审计报告
 - 年度自我保存审计报告：针对每种资源，系统自动统计
 - 应保存数量、接收到的数量、实际保存数量、累积保存数量
- NDPP组织第三方机构进行保存审计
 - 遴选非保存机构的10家合作保存机构
 - 远程随机抽取审计对象（期刊、卷期、篇）
 - 实时对比期刊网站
 - 统计分析多少应保存对象得到可靠保存

三、实施机制

(9) 可信赖审计与检查

- NDPP组织对合规保存机构进行可信赖认证
 - 按照长期保存系统可信赖认证标准（参考ISO 16262）
 - 组织非合规保存机构和其他保存机构专家
 - 现场进行保存审计、检查管理制度、检查保存日志等

国家数字科技文献资源长期保存体系

National Digital Preservation Program

一、建设背景

二、主要战略

三、实施机制

四、后续工作

四、后续工作

(1) 营造Preservation-enforcing环境

- 加强利用国家力量推动长期保存
 - 落实《互联网信息服务管理办法》
 - 落实公共订购机构长期保存责任
 - 国家资金和激励措施优先支持P-friendly出版社
- 加强所有订购机构对长期保存的责任和推动力
 - 了解和支持NDPP
 - 订购时理直气壮、持之以恒、理利兼施地要求保存权
- 加强保存权利的联合谈判
 - NSTL+NDPP+合规保存机构+订购联盟共同谈判

四、后续工作

(2) 扩大NDPP保存范围

- 扩大学协会出版社资源
- 扩大重点综合出版社资源
- 扩大电子图书出版资源
- 扩大非书刊类数字文献资源
- 试验工具书数据库
- 与人文社科数字文献资源责任机构的联合保存
- 适当扩大合规保存机构
 - 必须获得尚未得到保存的资源的保存权利
 - 或者获得特殊资源的保存权利

四、后续工作

(3) 扩展研发工作

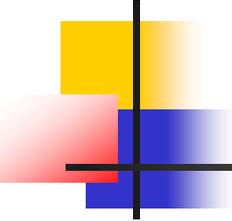
- 利用被保存资源进行数据挖掘的权利？
- 新型数字内容资源的长期保存研究与实验
 - 科学数据
 - 数字图像和音视频资源
 - 社交媒体资源、个人和社群数字资源
 - 计算机辅助设计和数字艺术资源
 - 复合数字内容产品
 - 智能数字内容产品
 - 数字人文资源
 - 数字教育资源
 - 虚拟现实内容资源
 - 下半年学术研讨会

(征文通知: <http://www.sal.edu.cn/view/new.aspx?id=12323>)

四、后续工作

共同的挑战、共同的使命、共同努力

- 每个机构都面临数字信息资源长期保存的挑战
 - 你已经面临数字内容为主的信息资源体系
 - 你如何规划基于数字环境的资源建设政策？
 - 你如何规划基于资源生命周期的数字内容可持续保存利用机制？
 - 你如何参加NDPP：
 - 参与规划、参与谈判、参与培训、参与宣传
 - 参与对合规保存机构的保存审计和可信赖机构认证审计
- 每个机构及其服务社区和用户都面临数字内容长期保存需求
 - 如何支持科研人员保存科研数据？
 - 如何支持教师保存课程内容资源？
 - 如何支持所服务社区、团体、活动保存数字知识资产？
 - 如何支持个人保存家庭或个人数字生活记录？

A decorative graphic on the left side of the slide, featuring overlapping yellow, red, and blue squares with a black crosshair.

国家数字科技文献资源长期保存体系

National Digital Preservation Program

谢谢!

<http://www.ndpp.ac.cn/>