

数字人文视野下的方志数据库建设

2019年7月

目 录

- 一、数字人文的影响
- 二、数字人文发展趋势
- 三、地方志数据库建设情况
- 四、国家数字方志馆建设构想

一、数字人文的影响

- (一) 数字人文改变了人们的阅读方式
- (二) 数字人文扩展了人们对资料的需求
- (三) 数字人文深刻影响了人文社科研究

二、数字人文发展趋势

数字化 —— 数据库化

（一）数字化

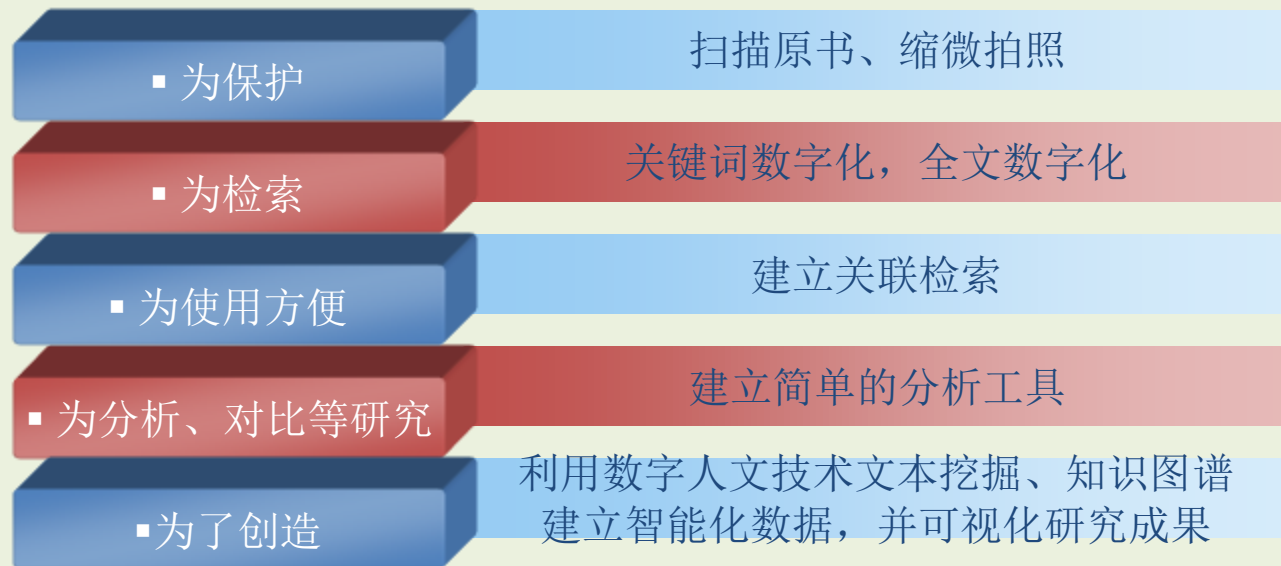
- 从纸本到电子本；
- 传统文本的数字化、数字作品的保存和简单分析与展示层面；

（二）数据库化

- 从电子文本到数据库；
- 满足计量统计、定位查询、可视化呈现、数据关联特征的结构化数据。

这种结构化数据是一种智慧数据，**具有可解释性**：是一种人和机器都能读懂的显性编码方案；**应带有自描述机制**：符合特定的语法和语义结构，拥有某种形式规范，支持逻辑推理，具有一定的粒度，可独立性发布；**有可溯源性**：能够满足意义上和证据上追溯的要求，以及演化过程表示的需要。**支持灵活性再利用**：包括多层次、多结构、领域、多学科、多场景、重组、链接、表示等再利用。

文献数字化发展历程



三、全国地方志数据库的建设

□资料

旧方志

中国地方志联合目录+中国地方志集成

新方志

志书+年鉴+地情书等

□商业性数据库

□地方志机构数据库

地方志的价值越来越受到重视

- 地方志的功能：存史、资政、育人

- 地方志具有独特价值

 - 全面系统性；资料证据性；百科知识性；

 - 国情调研性

- 具有可挖掘性

 - 服务发展、宣传地方

在中国出现的一系列地方志，无论从它们的广度来看，还是从它们的有系统的全面性方面来看，都是任何国家的同类文献所不能比拟的。”——[英]伟烈亚力

全国地方志资料情况资源统计

全国地方志系统资源统计

(截至2018年底,个别截至2019年4月)

首轮志书	5807册
二轮志书	3200.5册
地方综合年鉴	22946种
行业年鉴、部门年鉴、专业年鉴	8683册
行业志、部门志、专业志	24356册
乡镇村志、街道社区志	5779册
山水名胜古迹志	576册
地情书	12387册
教材、理论著述、工具书	1269册
历代方志整理(主要是旧志)	3408册
合计	88414.5册

商业性地方志数据库

数据库	新方志	旧方志	综合年鉴	专业年鉴
万方（地方志数据库）	42100	6826		
中国数字方志库（籍古轩）		10000种 （10万册）		
超星地方志数据库	3万余册	50067卷		
爱如生中国方志库		10000种左右		
中国知网年鉴数据库			4374（种，多种算一册）	1203种 （8975册）
国家图书馆（数字方志）		6529册		
中国地方志数据库（华中师范大学）	31483		中国地方志数据库（华中师范大学）	
华东师范大学数字方志集成平台	40576	27177		

市场运作方式，机制灵活；有源源不断的收入和投入；技术敏感有优势

全国各地地方志机构数据库的建设

1. 数据资源：方志、年鉴、地情书

2. 数据库

全国有21个省级地方志机构开展了地方志数字化工作（其中有两个省因为意识形态检查，暂时关闭）

综合志书：4310册

其他志书（部门行业、乡镇村志、特色志）：1020种（册）

地方综合年鉴：3064（册）

其他年鉴（专业行业年鉴、其他）：204册

旧志古籍：781册

地情资料：1038册

期刊：631期

其他（地方史等）：72册

合计：11120种（册、期）

一些大型城市开展地方志数据库建设

如有杭州、宁波、厦门、济南、青岛、郑州、武汉、长沙、南宁、桂林、成都、西安、兰州、银川等14个辖区市，数字化地方志达6475册。

这些有部分与省级库相冲突，但仍有约三分之一为省级库所无。

当前方志数据库存在的问题

- 问题一：文献存储格式不统一
 - 问题二：文献数字化程度不统一
 - 问题三：存储数据库不统一
 - 问题四：全文检索平台不统一
- 利用不方便

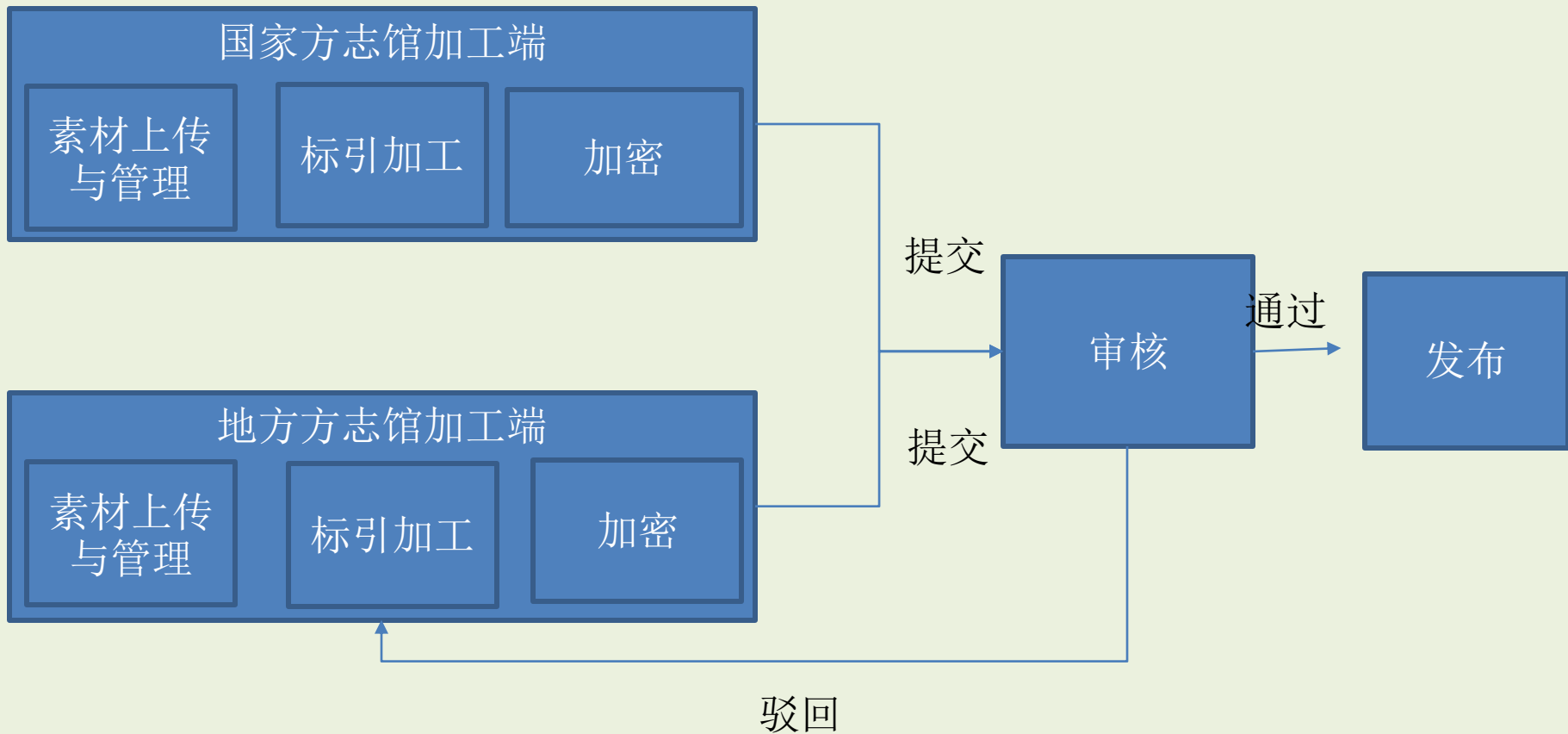
四、国家数字方志馆的建设

建设原则

- 共建共享公益
- “国家—地方”两级建设机制
- 确保存储的安全性和使用的便利性

平台基本架构





(一) 分布式加工平台

- 实现国家方志馆的资源上传、标引、加密与管理
- 实现地方方志馆的资源上传、标引、加密、管理与提交
- 实现国家方志馆对地方方志馆提交的资源的审核与管理

管理系统：机构&用户管理

机构管理

- 新建/编辑/删除机构
- 向机构授权

用户管理

- 新建/编辑/删除用户
- 向用户授权

权限管理

- 权限分配
- 权限开关

管理系统：门户管理

国家方志馆门户管理

地方方志馆门户管理

国家馆栏目建设

地方馆栏目管理

地方馆内容建设

国家馆内容建设

管理系统：专题管理

专题名称、题图
管理

专题栏目管理

专题内容管理

管理系统： 统计管理

- 数据分析
- 数据收集
- 数据可视化与报表

管理系统：安全管理

加密

密钥发放

加密处理

授权

按资源

按用户

Folio

数据
接口

- 成为文献应用单位的数据支撑平台

应用
接口

- 向文献应用单位提供基于OKAPI的微服务

应用系统·门户

- 目标
 - 实现国家方志馆门户和地方方志馆门户的访问
 - 实现全文检索文献
 - 实现带有DRM的web浏览器端志书、年鉴和期刊的在线阅读
 - 实现音视频、图片资源的在线播放
 - 实现虚拟馆的访问（外联）

应用系统：门户

频道

- 资源频道
- 新闻动态
- 虚拟馆
- 地方馆（仅国家馆有此频道）

检索检索

- 全文检索
- 逻辑检索
- 分面
- 知识图谱

专题

- 国家馆专题
- 地方馆专题

应用系统：阅读与播放

兼容的

- 操作系统:Windows 7及以上, macOS Mojave
- 浏览器: IE11/EDGE/Chrome/Firefox/Safari及使用上述内核的第三方浏览器

安全的

- 只针对被授权的用户展示全文
- 防下载

免安装的

- 无需安装程序, 直接在web浏览器中打开
- 无需安装插件

国家数据库建设的优势

- 各地已经有了数据化的基础
- 有顶层设计
- 便于协作共建：各地投入
- 有实体馆做保障。

国家方志馆建设规划：

- 将国家方志馆打造成为国家最大的地情信息文献中心。
- 馆藏图书资料达到10万册。
- 到2025年，国家数字方志馆平台建设成熟，集中全国各级地方志工作机构数字资源，打造成全国最大的地方志资源数据库，实现全国数字方志资源的共享共用。

国家数字方志馆建设的问题

□1.起步晚，很多具体问题需要在建设中解决。

如标准统一问题、兼容性问题、如何调动地方积极性问题

□2.基础设施、人才等方面还比较薄弱。

人才需要既懂人文，又懂技术

□3.投入需要逐年申请

□4.安全、保密等不确定性因素

□5.机构改革，带来了一定实施困难

最终目标

- 实现全国方志资源库的互联互通
- 实现全国方志资源的共享

谢谢大家