



华中师范大学 信息管理学院

INFORMATION MANAGEMENT SCHOOL OF CENTRAL CHINA NORMAL UNIVERSITY



# 智慧图书馆的文献资源造像探讨

华中师范大学信息管理学院 李玉海

2021-05-14



# 汇报大纲

- 1 造像缘起
- 2 造像之材
- 3 造像之法
- 4 造像之果



# 1 造像缘起

## 1.1 造像之“急”

传统文  
献资源  
描述不  
足

资源**采购**不足、  
资源**描述**不足、  
资源**检索**不足  
资源**微**推送

智慧应  
用需求  
紧迫

更加细粒度  
的**个性需求**、  
**场景区分**等

文献资  
源生态  
环境不  
良

**数据资源**垄断、  
图书馆的价值  
认同



# 1 造像缘起

## 1.2 词解“造像”

①塑造物体形象。

②用泥塑成或用石头、木头、金属等雕成的形象。

③相片；照片。



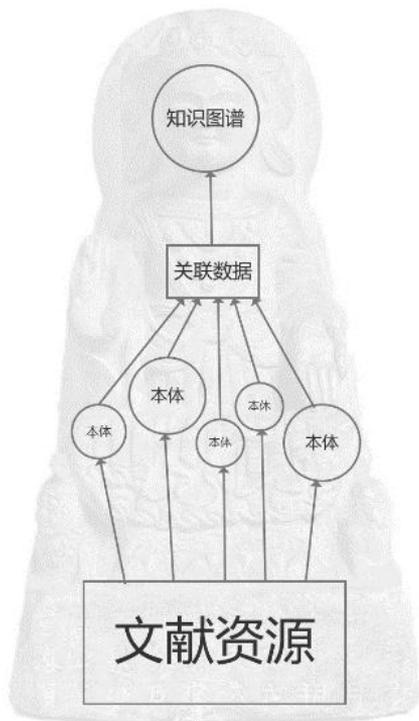
北周保定二年董道生造观音像



# 1 造像缘起

## 1.3 文献资源造像 n./v.

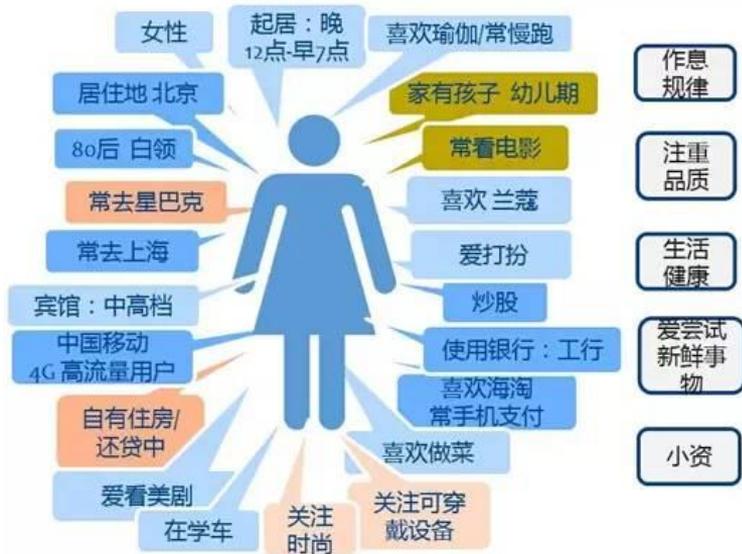
我们认为：文献资源造像是文献资源的**虚拟抽象**的代表，根据文献资源的类型、形式，从中抽取知识内容而构建一系列能够**被计算机所理解**和**关联**的**本体**，并最终呈现出具有**知识性、交互性、联接性**的知识图谱等多种新形式产品的一系列过程。



# 1 造像缘起

## 1.4 画像&造像

用户画像 (User Persona) 最早由Alan Copper提出, 认为用户画像是真实用户的**虚拟代表**, 根据用户**行为和动机**, 从中抽取每类用户的**共同特征**, 并设定名字、照片、场景等要素对其进行描述。



# 1 造像缘起

## 1.4 画像&造像

相同点	用户画像	资源造像
动静结合	用户的静态属性和动态属性	公开出版的静态信息资源和不断更新的网络信息资源
抽象	对现实用户的抽象	对文献资源内容的抽象
以技术为支撑	大数据、数据挖掘、可视化技术	关联数据、语义网技术



# 1 造像缘起

## 1.4 画像&造像

不同点	用户画像	资源造像
使用场景	业务层面	前端模型构建层面
目的	针对用户进行推荐、销售、评价产品	知识发现、节约查找时间
特性	基本性、移情性、真实性、独特性、目标性、数量和应用性	感知性、专业性、可变性、传递性、有用性、启发性



# 1 造像缘起

## 1.4 画像&造像

persona

primary  
基本性

empathy  
同理性

realistic  
真实性

singular  
独特性

objectives  
目标性

number  
数量性

applicable  
应用性

statue

sensitive  
感知性

technical  
专业性

alterable  
可变性

transferable  
传递性

useful  
有用性

enlightening  
启发性



## 2 造像之材

### 2.1 传统图书馆下的文献资源

以文献为载体

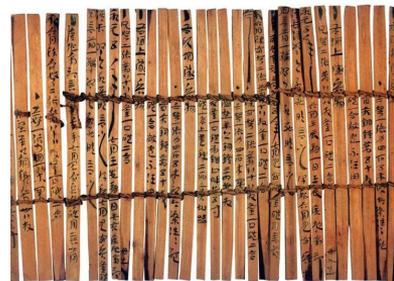


刻写型  
印刷型  
缩微型  
视听型



## 2 造像之材

### (1) 刻写型



■毛泽东手书《蝶恋花·答李淑一》

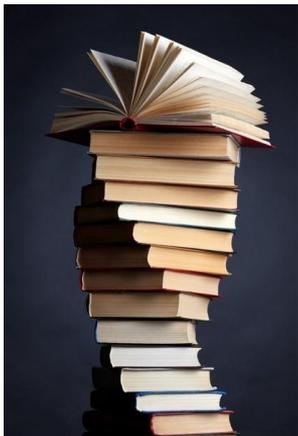


资料图片



## 2 造像之材

### (2) 印刷型



## 2 造像之材

### (3) 缩微型



## 2 造像之材

### (4) 视听型



## 2 造像之材

### 2.2 智慧图书馆的文献资源

#### (1) 口述历史资料



## 2 造像之材

### 2.2 智慧图书馆的文献资源

#### (2) 电子出版物





## 2 造像之材

### 2.2 智慧图书馆的文献资源

#### (4) 专利信息





## 2 造像之材

### 2.2 智慧图书馆的文献资源

以数字化手段为依托

刻写型  
印刷型  
缩微型  
视听型



口述历史资料  
电子出版物  
网络资源  
专利信息  
数字人文

.....



立体化资  
源体系



华中师范大学

信息管理学院

INFORMATION MANAGEMENT SCHOOL OF CENTRAL CHINA NORMAL UNIVERSITY

# 3 造像之法

## 3.1 文献资源造像&数字孪生

数字孪生( Digital Twin) 以数字化的方式建立物理实体的多维、多时空尺度、多学科、多物理量的动态虚拟模型来仿真和刻画物理实体在真实环境中的属性、行为、规则等。

文献资源造像将作为智慧图书馆资源建设的重要组成部分，其构建过程借鉴数字孪生的五维模型。



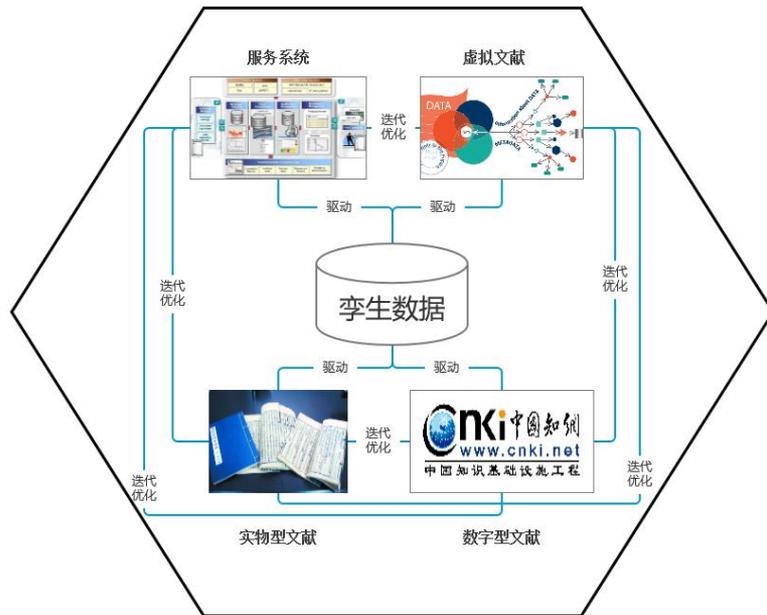
# 3 造像之法

## 3.1 文献资源造像&数字孪生

北航数字孪生技术研究团队提出了**数字孪生五维模型**的概念。主要包括：

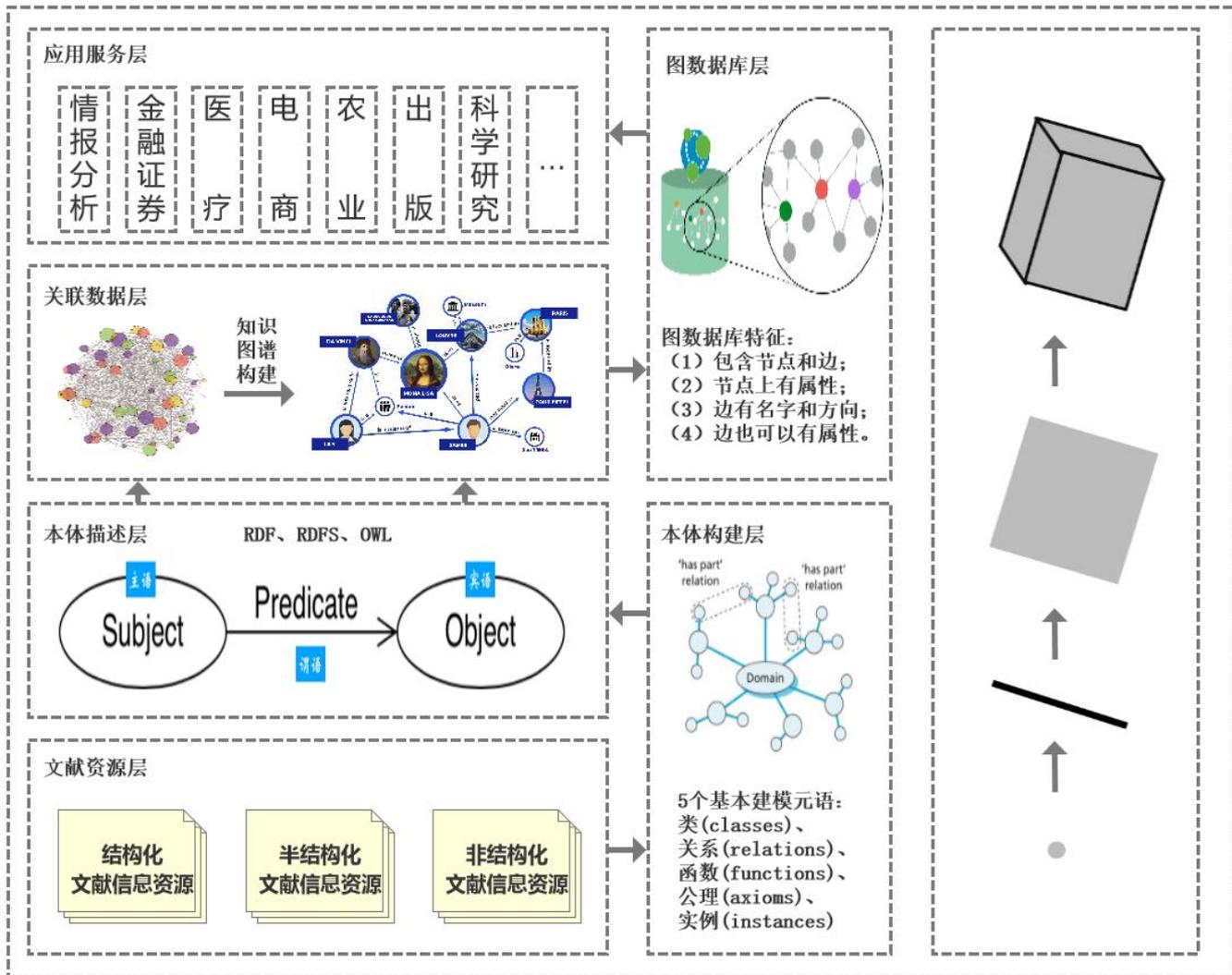
- (1) 物理实体, (2) 虚拟实体,
- (3) 服务, (4) 孪生数据,
- (5) 各组成部分间的连接。

但文献资源还要多一个维度——不具备实体的**数字型文献资源**



# 3 造像之法

## 3.2 构建框架



## 3 造像之法

### 3.3 本体的构建——一点

本体 (Ontology) 的概念源自于哲学领域，在哲学中的定义为“对世界上客观事物的系统描述，即存在论”。

在计算机科学与信息科学领域，理论上，本体是指一种“形式化的，对于共享概念体系的明确而又详细的说明”。本体实际上就是对特定领域之中某套概念及其相互之间关系的形式化表达 (formal representation)。

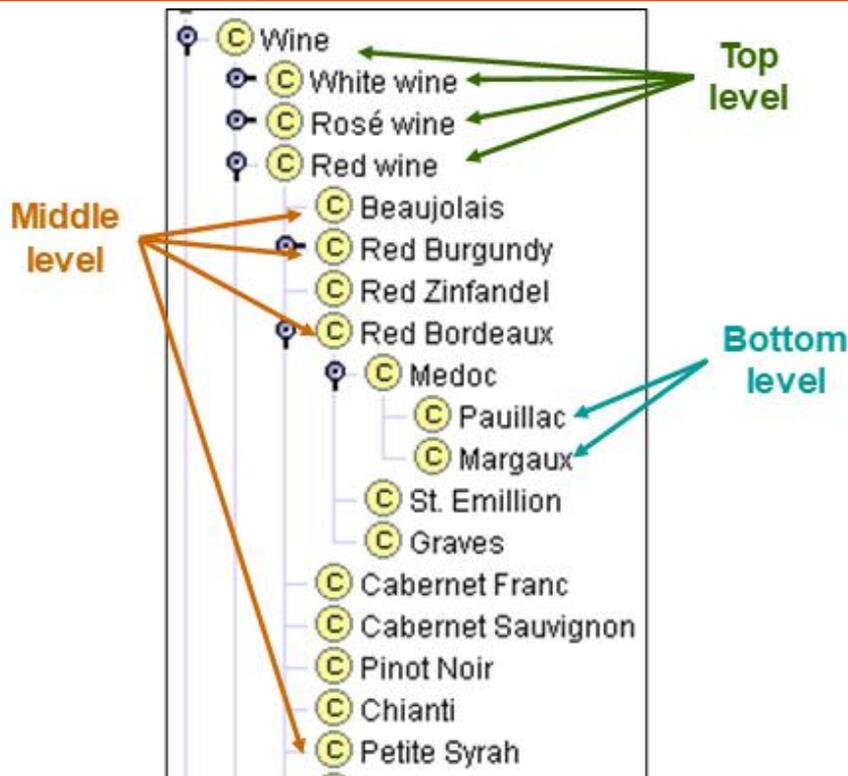


# 3 造像之法

## 3.3 本体的构建——一点

### 本体构建原则

- (1)明确性和客观性。
- (2)一致性。
- (3)可扩展性。
- (4)最小编码偏差。
- (5)最小本体承诺



## 3 造像之法

### 3.3 本体的构建——一点

本体的应用：

- (1) 信息检索：基于关键字的检索  语义检索
- (2) 信息集成：一定程度上解决语义异构的问题
- (3) 机器翻译：歧义消解
- (4) 知识获取：在具有语义意义的规则上挖掘知识



## 3 造像之法

### 3.4 本体的描述——线

**传统元数据描述的缺陷：**

**MARC元数据：**①MARC结构复杂， 字段重复；② MARC的格式设计存在不足（最初以磁带格式为标准）， 需借助专门软件读取。；

**DC元数据：**①是基于人的理解和认知建立的， 计算机难理解；  
②无法实现语义上的关联。

故需要新的描述语言来描述连接已有的元数据， 实现元数据的互操作并对所构建的本体进行有效的描述。



# 3 造像之法

## 3.4 本体的描述——线

### (1) RDF (Resource Description Framework, 资源描述框架)

RDF使用web标识符 (主体/主语) 来标记资源, 使用属性 (谓语) 和属性值 (客体/宾语) 来描述资源。这里的资源、属性和属性值就构成了一个陈述。

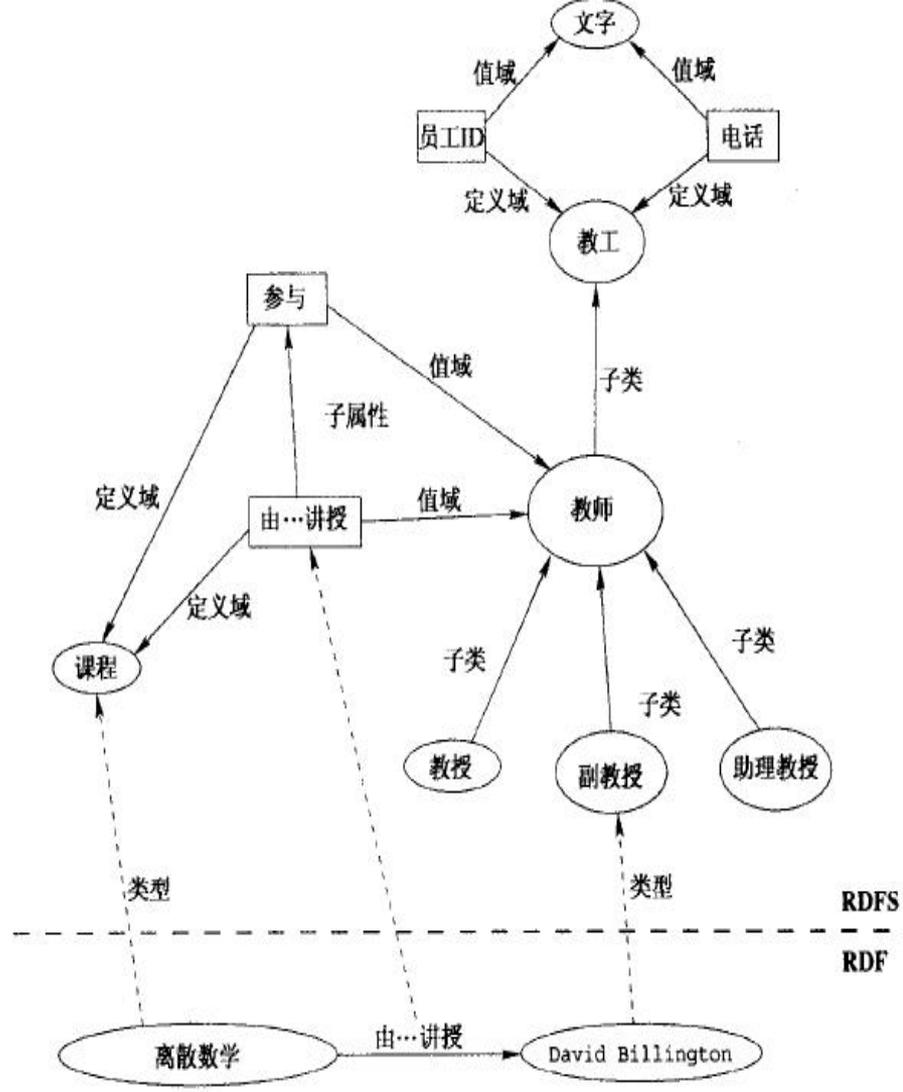


# 3 造像之法

## 3.4 本体的描述——线

### (2) RDFS (RDF Schema, RDF词汇描述语言)

在RDF基础上对其进行扩展，定义了类、属性、属性值来描述客观世界，并且通过定义域和值域来约束资源，更加形象化表达了知识。——RDF的“衣服”



## 3 造像之法

### 3.4 本体的描述——线

#### (3) OWL (Web Ontology Language, Web本体语言)

OWL是由W3C开发的网络本体语言，用来对本体进行语义描述。OWL保持了原有RDF、RDFS的兼容性，有保证率较好的语义表达能力，OWL分为三种子语言：OWL-Lite、OWL-DL和OWL-Full。——RDFS的扩展

作用：

1. 提供快速、灵活的数据建模能力。
2. 高效的自动推理。



## 3 造像之法

### 3.5 关联数据——一面

关联数据是国际互联网协会(W3C)推荐的一种规范, 用来发布和联接各类数据、信息和知识。

- 与RDF不同的是, 通过将URI (通用资源标识符) 与HTTP协议进行组合, 关联数据可以将用三元组来表示的资源发布到网上。
- 关联数据本身并不具备语义特征, 但可以通过在数据层面建立关联, 促进语义网络的形成。



## 3 造像之法

### 3.5 关联数据——一面

关联数据的发明人蒂姆·伯纳斯·李(Tim Berners-Lee)总结了四个原则:

- (1)使用URI 作为*任何事物的标识名称*, 不仅是标识文档;
- (2)使用HTTP URI, 使任何人都可以*参引 (dereference)*这一全局唯一的名称;
- (3)当有人访问名称时, 以*RDF 形式*提供有用的信息;
- (4)尽可能提供链接, *指向其他的URI*, 以使人们发现更多的相关信息。





## 3 造像之法

### 3.6 知识图谱——一体

知识图谱应用：

- (1) 智能语义搜索
- (2) 移动个人助理
- (3) 深度问答系统



## 3 造像之法

### 3.7 图数据库——库

由于无法高效表示复杂的关系，以及在大数据量和多服务器之上进行水平伸缩的限制，**关系型数据库**渐渐不适用。

NoSQL (Not Only SQL, 不限于SQL) 是一种解决方案，主要包括：

- (1) 键-值存储库 (Key-Value-stores) ；
- (2) BigTable实现 (BigTable-implementations) ；
- (3) 文档库 (Document-stores) ；
- (4) 图数据库 (Graph Database) ；



## 3 造像之法

### 3.7 图数据库——库

#### 图数据库 (Graph Database)

图数据库源起欧拉和图理论，是以“图”这种数据结构存储和查询数据，并非存储图片的数据库。它的数据模型主要是以节点和关系（边）来体现，也可处理键值对。具有以下特征：

- (1) 包含节点和边；
- (2) 节点上有属性（键值对）；
- (3) 边有名字和方向，并总是有一个开始节点和一个结束节点；
- (4) 边也可以有属性。



# 4 造像之果

## 4.1 使用场景全融合

### (1) 客户端与移动APP端融合

- 移动端的检索历史与电脑端同步
- 智能手环、智能音箱等智能设备新接口



# 4 造像之果

## 4.1 使用场景全融合

(2) 高校、企业、政府全融合

- **企业**：实现专利知识的共享，促进专利成果的转化
- **政府**：了解高校的研究新动态，促进政策的制定
- **高校**：以企业实际业务知识为抓手，促进对口人才的培养

期刊	编辑	作者	读者
文本	表单	AVG	



# 4 造像之果

## 4.2 知识利用全智能

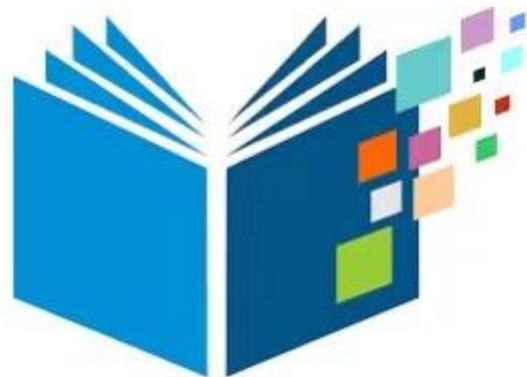
- (1) 新知识关联的推送服务
- (2) 个人知识图谱的构建，隐性知识的传递与服务
- (3) 通过知识关联的学科主题预测
- (4) 交叉学科的发展



## 4 造像之果

### 4.3 读者用户全覆盖

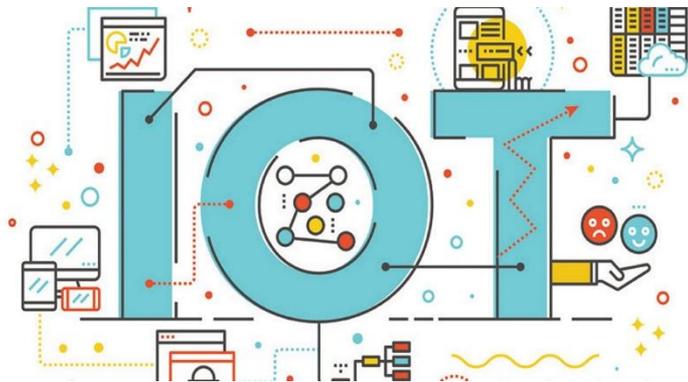
- (1) 青少年艺术、科学启蒙服务
- (2) 青年世界观塑造服务
- (3) 中青年职场业务知识服务
- (4) 老年医疗健康知识服务
- (5) 残障人士多媒体形式知识服务



# 4 造像之果

## 4.4 综合数据全利用

- (1) 用户交互行为数据，推荐好友。
- (2) 物联网设备采集数据，辅之以良好环境。
- (3) 浏览历史数据、点击数据、浏览时长数据等归纳与整理。



## 4 造像之果

### 4.5 管理业务全优化

- (1) 用户需求实时挖掘
- (2) 图书馆新业务
- (3) 馆员分工与价值认同
- (4) 用户满意度评估反馈



# 总 结

**有血有肉有灵魂的文献资源仍然是智慧图书馆核心组成部分。**

**智慧图书馆的文献资源一定有其本征特性，如何构造非常需要、也非常值得大家研究探索。**



# 参考文献

- [1] 张总. 石窟与造像中的塔形和碑式——佛教艺术中国化演变之例证[J]. 社会科学战线, 2021, (04): 113-119.
- [2] COOPER A. 交互设计之路[M]. 北京: 电子工业出版社, 2006.
- [3] 宋美琦, 陈烨, 张瑞. 用户画像研究述评[J]. 情报科学, 2019, 37(04): 171-177.
- [4] 肖珑, 陈凌, 冯项云, et al. 中文元数据标准框架及其应用[J]. 大学图书馆学报, 2001, (05): 29-35+91.
- [5] 夏天, 钱毅. 面向知识服务的档案数据语义化重组[J]. 档案学研究, 2021, (02): 36-44.
- [6] 陶飞, 刘蔚然, 张萌, et al. 数字孪生五维模型及十大领域应用[J]. 计算机集成制造系统, 2019, 25(01): 1-18.
- [7] 陶飞, 刘蔚然, 刘检华, et al. 数字孪生及其应用探索[J]. 计算机集成制造系统, 2018, 24(01): 1-18.
- [8] GRUBER T R. Toward principles for the design of ontologies used for knowledge sharing?[J]. International Journal of Human-Computer Studies, 1995, 43(5): 907-928.
- [9] BORST W N, BORST W N. Construction of Engineering Ontologies for Knowledge Sharing and Re use[D]. Centre for Telematics and Information Technology (CTIT), 1997.



# 参考文献

- [10] NOY N, MCGUINNESS D. Ontology Development 101: A Guide to Creating Your First Ontology[J]. Knowledge Systems Laboratory, 2001, 32.
- [11] 李善平, 尹奇韡, 胡玉杰, et al. 本体论研究综述[J]. 计算机研究与发展, 2004, (07): 1041-1052.
- [12] 严亚兰, 张金隆. 基于OWL的多媒体本体构建探索[J]. 计算机工程与应用, 2006, (28): 40-42+88.
- [13] 何海芸, 袁春风. 基于Ontology的领域知识构建技术综述[J]. 计算机应用研究, 2005, (03): 14-18.
- [14] 沈志宏, 张晓林. 关联数据及其应用现状综述[J]. 现代图书情报技术, 2010, (11): 1-9.
- [15] 刘炜. 关联数据:概念、技术及应用展望[J]. 大学图书馆学报, 2011, 29(02): 5-12.
- [16] 欧石燕. 面向关联数据的语义数字图书馆资源描述与组织框架设计与实现[J]. 中国图书馆学报, 2012, 38(06): 58-71.
- [17] 刘峤, 李杨, 段宏, et al. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(03): 582-600.
- [18] 赵子豪, 沈志宏. 一种适合多元异构图数据管理系统的交互分析框架[J]. 数据分析与知识发现, 2019, 3(10): 37-46.



# 精神家园

# 知识海洋



请各位批评指正！ 谢谢！



华中师范大学 信息管理学院  
INFORMATION MANAGEMENT SCHOOL OF CENTRAL CHINA NORMAL UNIVERSITY