

# 关联数据 (Linked Data) 在家谱数字人文服务中的应用

上海图书馆 夏翠娟

2018年6月 @ \*\*\*\*大学

# 什么是关联数据 (Linked Data)

关联数据 (Linked Data)  
是一种在Web上组织、发布、获  
取、利用数据的方式

# 关联数据的四原则



# HTTP URI

IRI: 内部资源标识符

URI: 统一资源标识符 (isbn)

URL: 统一资源定位符

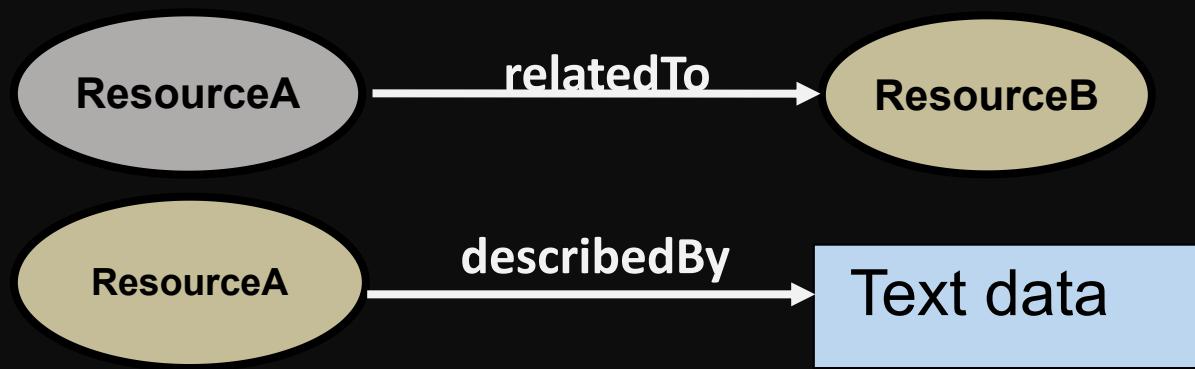
[https://www.baidu.com/s?ie=utf-8&f=8&rsv\\_bp=0&rsv\\_idx=1&tn=baidu&wd=胡适](https://www.baidu.com/s?ie=utf-8&f=8&rsv_bp=0&rsv_idx=1&tn=baidu&wd=胡适)

HTTP URI: 唯一标识、唯一定位、永久稳定

[http://dbpedia.org/resource/Lu\\_Xun](http://dbpedia.org/resource/Lu_Xun)

主体(subject)    谓词(predict)    客体(object)

抽象模型  
Abstract Model

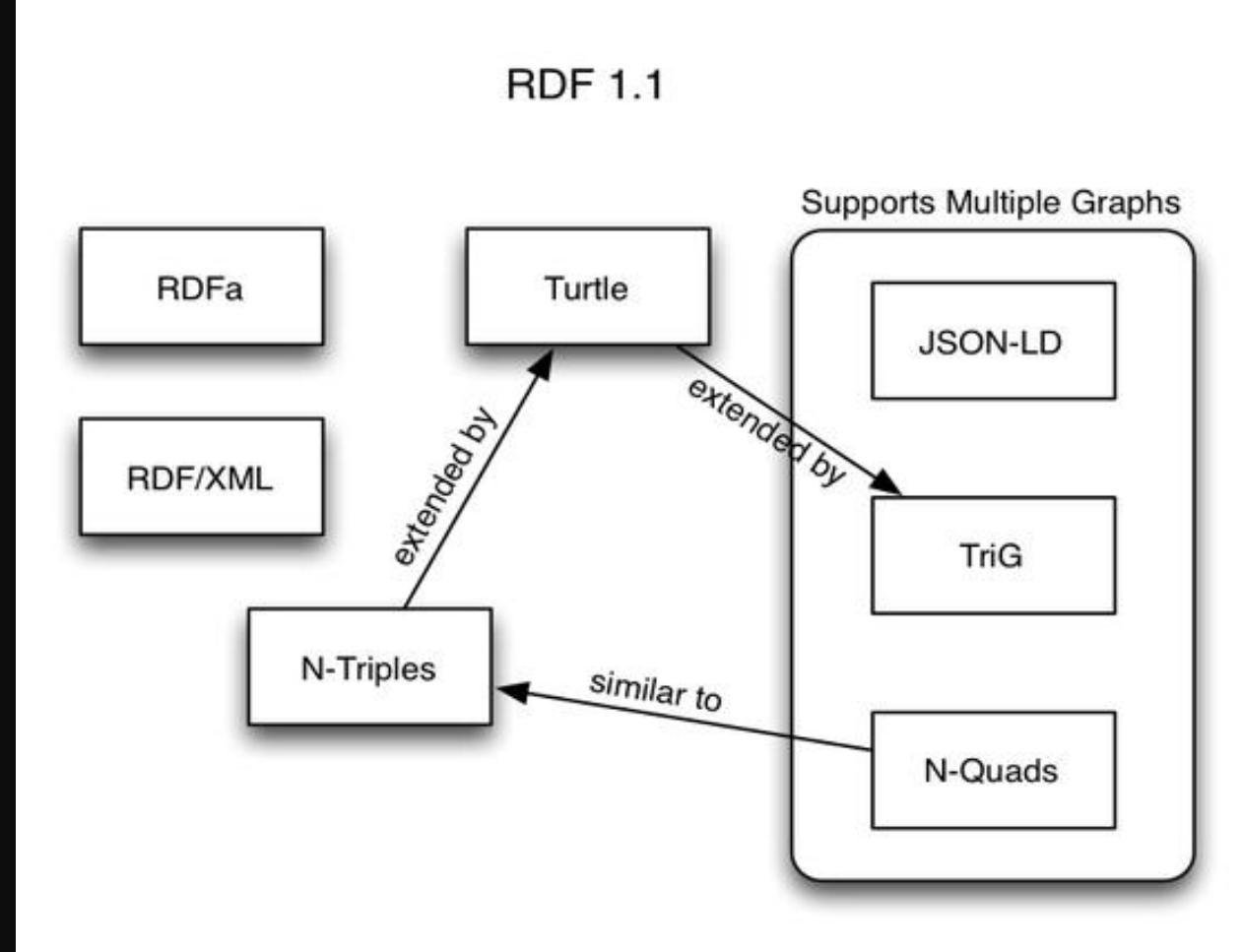


数据实例  
Data Instance

```
@prefix foaf: <http://xmlns.com/foaf/0.1/> .  
<http://data.library.sh.cn/jp/entity/person/etrd44w3m3g1vncn> a shl:Person;  
    foaf:familyName <http://data.library.sh.cn/authority/familyname/68n959cf8zdfkz3v>;  
    foaf:name “夏翠娟”@zh_cn.
```

数据编码  
序列化格式  
Serialization

# RDF 序列化 格式



<http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>

# RDF 序列化格 式— RDF/XML

## EXAMPLE 3

### Complete description of all graph paths

```
<rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
  <ex:editor>
    <rdf:Description>
      <ex:homePage>
        <rdf:Description rdf:about="http://purl.org/net/dajobe/">
          </rdf:Description>
        </ex:homePage>
      </rdf:Description>
    </ex:editor>
  </rdf:Description>

<rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
  <ex:editor>
    <rdf:Description>
      <ex:fullName>Dave Beckett</ex:fullName>
    </rdf:Description>
  </ex:editor>
  </rdf:Description>

<rdf:Description rdf:about="http://www.w3.org/TR/rdf-syntax-grammar">
  <dc:title>RDF 1.1 XML Syntax</dc:title>
</rdf:Description>
```

# RDF 序列化 格式— Turtle

## EXAMPLE 1

```
@base <http://example.org/> .  
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .  
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .  
@prefix foaf: <http://xmlns.com/foaf/0.1/> .  
@prefix rel: <http://www.perceive.net/schemas/relationship/> .  
  
<#green-goblin>  
    rel:enemyOf <#spiderman> ;  
    a foaf:Person ;    # in the context of the Marvel universe  
    foaf:name "Green Goblin".  
  
<#spiderman>  
    rel:enemyOf <#green-goblin> ;  
    a foaf:Person ;  
    foaf:name "Spiderman", "Ч е л о в е к - п а у к"@ru .
```

# RDF 序列化 格式— RDFa

本网页资源是一个由  
schema.org定义的博客  
[ ] rdfs:type schema:Blog

## EXAMPLE 1: Example of an HTML+RDFa 1.1 document

```
<!DOCTYPE html>
<html lang="en">
  <head>
    <title>Example Document</title>
  </head>
  <body vocab="http://schema.org/">
    <p typeof="Blog">
      Welcome to my <a property="url" href="http://example.org/">blog</a>.
    </p>
  </body>
</html>
```

## EXAMPLE 2: Turtle output of Example Document

```
[] a <http://schema.org/Blog>;
<http://schema.org/url> <http://example.org/> .
```

# RDF 序列化格式—JSON-LD

```
{  
  "@context":  
  {  
    "xsd": "http://www.w3.org/2001/XMLSchema#",  
    "foaf": "http://xmlns.com/foaf/0.1/",  
    "foaf:homepage": { "@type": "@id" },  
    "picture": { "@id": "foaf:depiction", "@type": "@id" }  
  },  
  "@id": "http://me.markus-lanthaler.com/",  
  "@type": "foaf:Person",  
  "foaf:name": "Markus Lanthaler",  
  "foaf:homepage": "http://www.markus-lanthaler.com/",  
  "picture": "http://twitter.com/account/profile_image/markuslanthaler"  
}
```

# RDF 查询语言—SPARQL

```
PREFIX bf: <http://bibframe.org/vocab/>
SELECT ?title
FROM <http://loc.gov/catalog/dataset.rdf>
WHERE {
    ?x bf:title ?title .
}
ORDER BY ?title
Limit 10
```

跨网域  
Web成为一  
个global的  
数据库

基于RDF三元组模型  
直接面向知识的查询

SPARQL 协议与RDF查询语言  
( SPARQL Protocol and RDF Query Language )

# 关联数据的技术堆栈

资源标识和定位：HTTP URI

领域建模：Web Ontology

数据抽象模型：RDF

数据编码格式：XML/Turtle/N3/Json-LD

数据存储：GraphDB

数据查询：SPARQL

# 为什么要用关联数据

# 从文档的网络web of documents 到数据的网络 (web of data)

胡适

编辑

+ 收藏 14994 | 416

胡适（1891年12月17日-1962年2月24日），原名嗣糜，学名洪骍，字希疆，笔名胡适，字适之。著名思想家、文学家、哲学家。<sup>[1]</sup>徽州绩溪人，以倡导“白话文、领导新文化运动闻名于世。”

幼年就读于家乡私塾，19岁考取庚子赔款官费生，留学美国，师从哲学家约翰·杜威，1917年夏回国，受聘为北京大学教授。1918年加入《新青年》编辑部，大力提倡白话文，宣扬个性解放、思想自由，与陈独秀同为新文化运动的领袖。他的文章从创作理论的角度阐述新旧文学的区别，提倡新文学创作，翻译法国都德、莫泊桑、挪威易卜生的部分作品，又率先从事白话文学的创作。他于1917年发表的白话诗是现代文学史上的第一批新诗。胡适信奉实验主义哲学。“五四运动”后，同李大钊、陈独秀等接受马克思主义的知识分子分道扬镳，由“问题与主义之争”开其端，倡导改良，从此改变了他“20年不谈政治；20年不干政治”的态度。他于1920年代办《努力周报》，1930年代办《独立评论》，1940年代办“独立时论社”。1938～1942年出任国民政府驻美大使。1946～1948年任北京大学校长。1949年去美国。1952年返台湾任中央研究院院长。1962年在台北病逝。

胡适一生的学术活动主要在文学、哲学、史学、考据学、教育学、红学几个方面，主要著作有《中国哲学史大纲》（上）、《尝试集》、《白话文学史》（上）和《胡适文存》（四集）等。他在学术上影响最大的是提倡“大胆的假设、小心的求证”的治学方法。

人物关系

妻子 江冬秀 儿子 胡祖望 儿子 胡思杜 旧爱 曹诚英 旧爱 韦莲司 父亲 胡传 母亲 冯顺弟

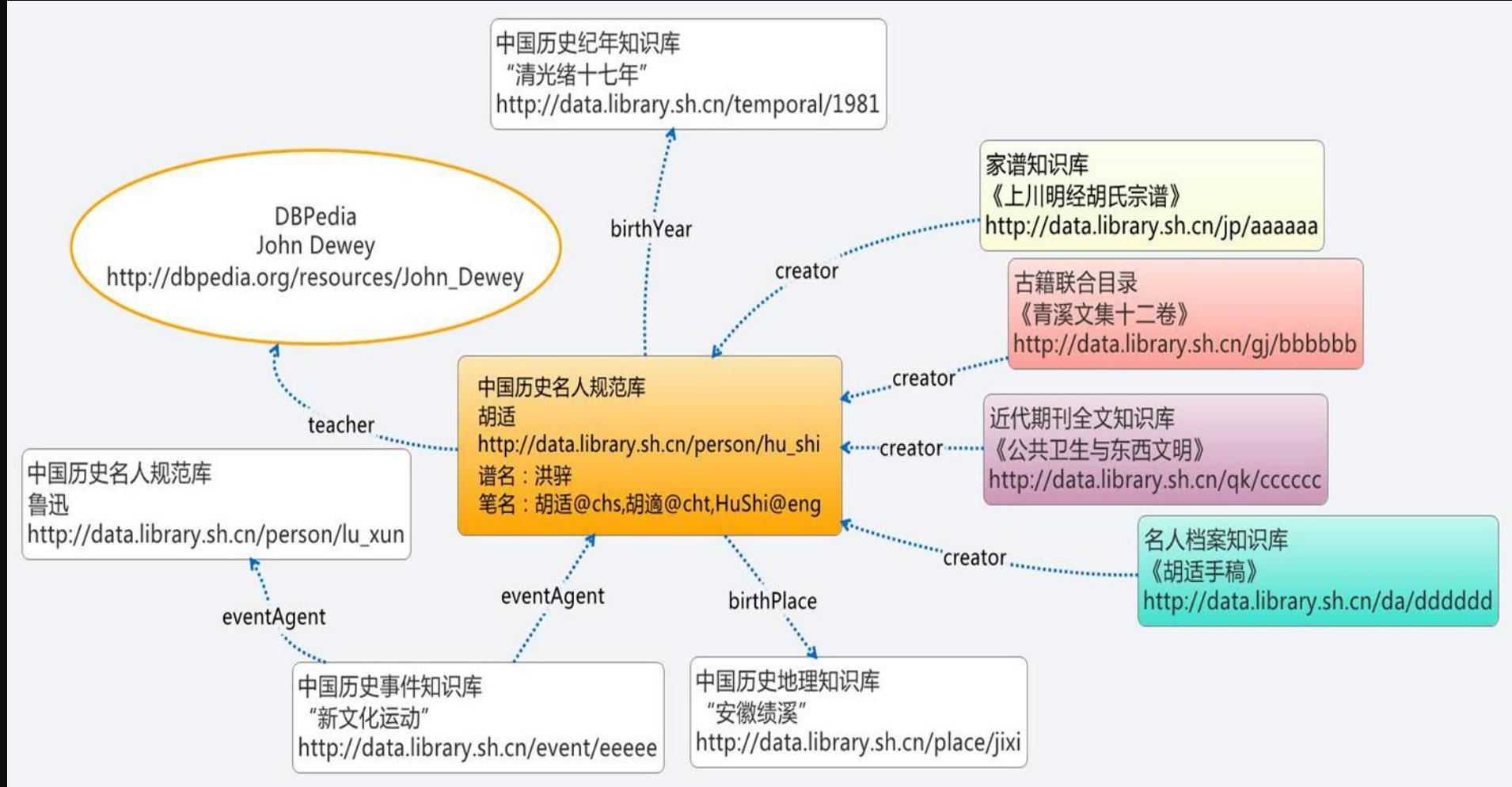
纠错

胡适图册

V百科 往期回顾

南京大屠杀

# 从文档的网络 web of documents 到实体的网络 (web of entities)



# 改变图书馆的生存环境 提升图书馆的核心竞争力



融入互联网

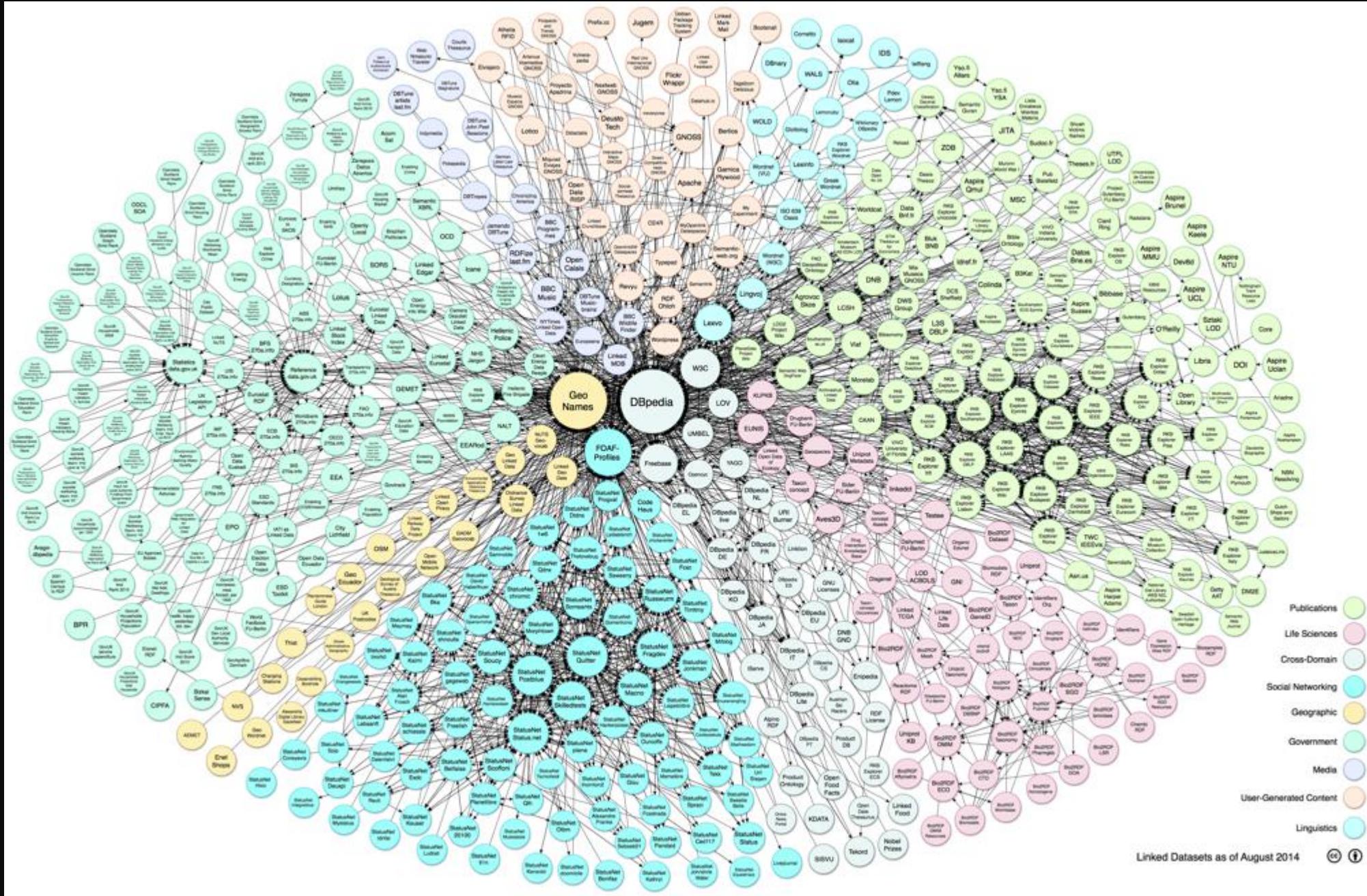
图档博&关联数据，  
让数据在网络上可见！

Of the Web

成为网络的一部分

Not On the Web

而非依附于网络之上



# 图书馆关联数据

## 关联的外部数据集



# 知识组织

# 关联数据用于知识组织

知识单元的细粒度化

知识组织的语义化 (机器可读、可计算、可理解、可推理)

知识呈现的可视化

# 以胡适先生家族（上川明经胡氏）迁徙路线为例



明经胡氏續修宗谱不分卷（江西省）  
婺源

【副题名】明經胡氏續修宗譜

【责任者】胡璘

【纂修时间】明嘉靖二十三年（1544）

【摘要】始祖昌翼，谱稱本唐昭宗幼子，母何后，天祐元年生，避禍來婺源考川，遂冒義父胡三公姓。後唐同光間，登明經進士。後世子孫散居新安諸邑、遂安、淳安、樂平、鄱陽、浮梁、興安、漢陽，乃遠至貴州、雲南者，世稱明經胡氏，以別安定胡氏。是為明經胡氏統宗譜。譜存譜序、墓圖及宋胡次焱撰《論始祖》、《論姓氏》、《論過房》、《論稱呼》等。书名据版心、卷端题。

【版本】刻本

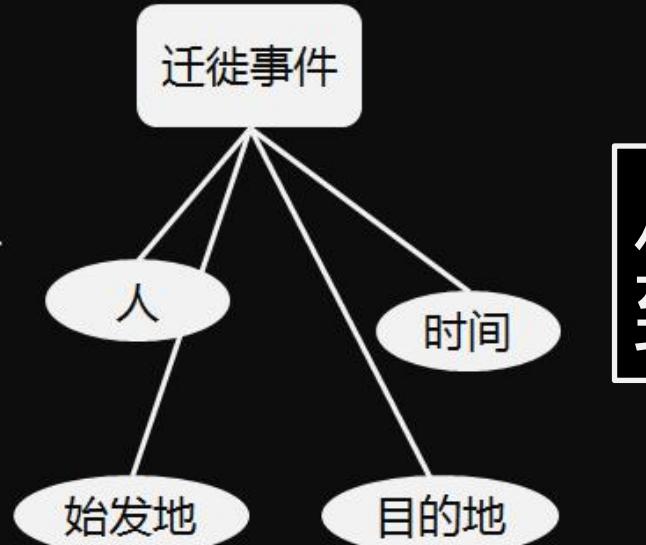
【数量】存三册

上图 山西家谱

馆藏地：上海图书馆  
地址：上海市 徐汇区 淮海中路1555号  
DOI为STJP011208：残本（索书号：911895-97）请至上海图书馆家谱阅览室借阅

面向文献，  
以元数据记  
录（MARC  
Record）为  
基本单位

# 基于本体的知识组织



从概念(Concept)  
到本体(Ontology)

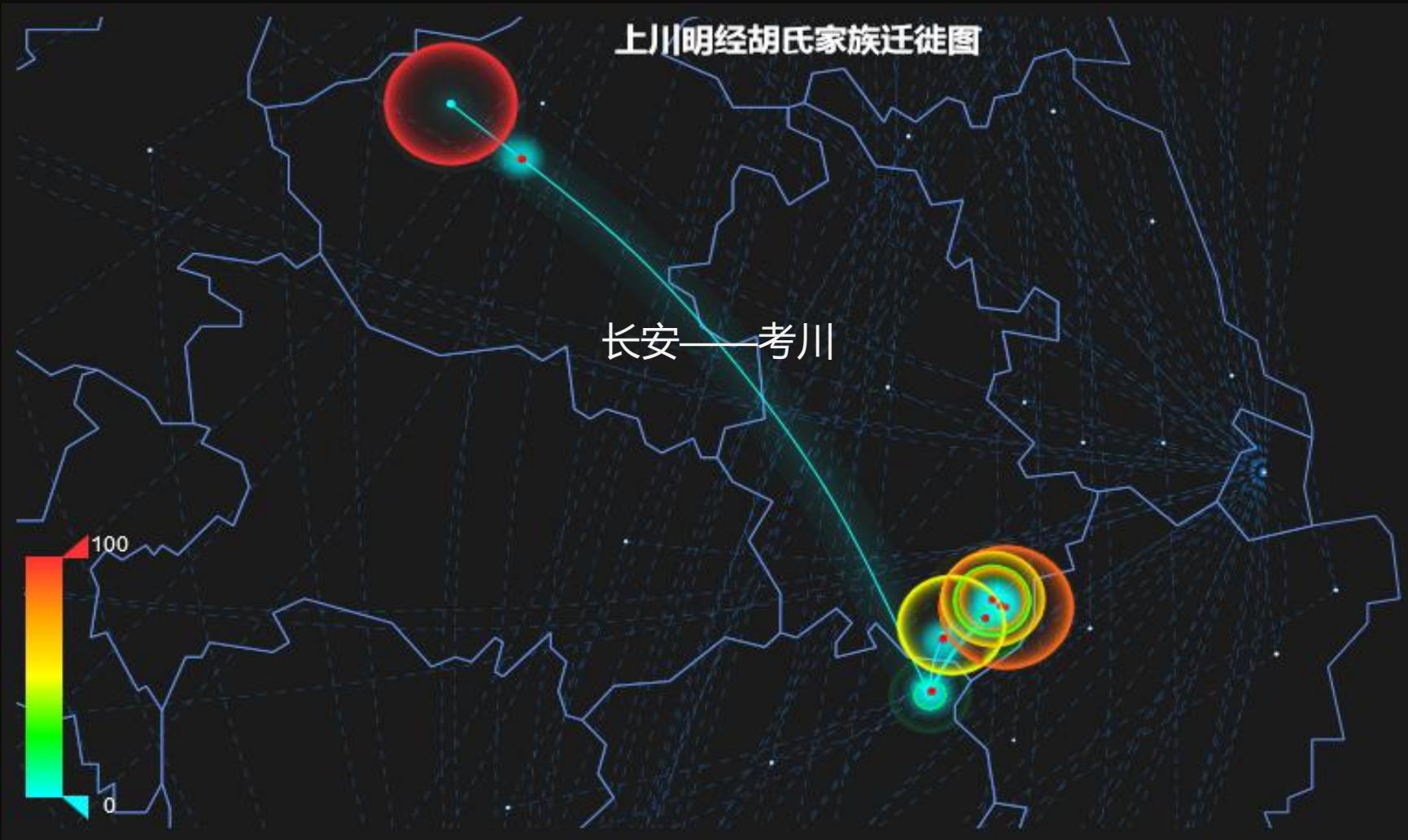
```
shl:Migration a rdfs:Class ;  
rdfs:label "迁徙" ;  
rdfs:subClassOf shl:Resource ;  
rdfs:comment "一次迁徙事件。" .  
shl:originalLocality a rdf:Property ;  
rdfs:label "原居地" ;  
rdfs:domain shl:Migration ;  
rdfs:range bf:Place ;  
rdfs:comment "迁徙之前的居地。" .  
shl:locality a rdf:Property ;  
rdfs:label "迁居地" ;  
rdfs:domain shl:Migration ;  
rdfs:range bf:Place ;  
rdfs:comment "迁徙完成后的居地。" .
```

# 知识组织的语义化

面向数据、事实和  
知识，以RDF  
Triples为基本单位

```
<http://gen.library.sh.cn/Migration/1> a shlgen:Migration ;
    bf:eventAgent <http://gen.library.sh.cn/Person/16607> ;
    shlgen:temporal <http://gen.library.sh.cn/Temporal/14> ;
    shlgen:locality <http://gen.library.sh.cn/Place/1129> ;
    shlgen:originalLocality <http://gen.library.sh.cn/Place/203>.
<http://gen.library.sh.cn/Person/16607> a shlgen:Person ;
    foaf:name "胡昌翼" ;
    shlgen:courtesyName "宏远";
    shlgen:pseudonym "眉轩" ;
    shlgen:description "唐昭宗幼子" .
<http://gen.library.sh.cn/Place/1129> a shlgen:Place ;
    bf:label "考川" ;
    shlgen:county "婺源县" ;
    shlgen:city "上饶市" ;
    shlgen:province "江西省" ;
    shlgen:town "紫阳镇" ;
    shlgen:village "考水村" ;
    geo:long "117.757787" ;
    geo:lat "29.27642" .
<http://gen.library.sh.cn/Temporal/14> a shlgen:Temporal ;
    shlgen:dynasty "唐" ;
    shlgen:reignTitle "天佑元年" ;
    shlgen:beginYear "904"^^xsd:int .
```

# 数据可视化



# 世系表——家谱文献中最重要的内容

兆賜公	次子	兆賜公	次子	兆賜公	次子	德振公	一德振公	錫紅公止
應祥公	——	天信公	止	娶某氏生一子				
應祥公	——	天輝公	——	德水公	——	錫清公	止	
兆思公	三子	兆思公	三子	兆思公	三子	德火公	止	
應成公	——	天輝公	——	根喜公	——	錫業公	——	貞開
生康熙庚子娶 某氏生二子以 次子天禮爲應 禎後		娶某氏生二子		娶某氏生二子		名觀業娶曹氏 生二子公葬上		
應環公	——	天仁公	——	根喜公	——	錫業公	——	貞開
字玉連生雍正 戊申五月二十 月初八娶某氏 生一子		生乾隆庚申六 月初八娶某氏 生一子		娶某氏生二子		庵腳氏葬社土		
兆恩公	子	應寶公	——	德高公	——	錫柏公	——	貞吉公止
兆蔭公	次子	字惟善生康熙 庚子十月十四 娶汪氏生康熙 己亥生一子		娶某氏生二子		貞爲公止		
應可公	——	天樂公	——	德高公	——	錫柏公	——	貞吉公止
應可公	——	德拱公	——	德高公	——	貞吉公止		
兆賜公	長子	應禎公	——	德拱公	——	錫順公	止	
兆蔭公	次子	應成公	——	德拱公	——	錫順公	止	
兆賜公	次子	遷休甯		德亮公	——	錫順公	止	
應成次子爲後				貞開公	止	貞奎公		

上川明經胡氏世系表扫描图片

# 上川明经胡氏世系表

## ——从扫描图片到结构化数据

Id	姓氏	谱名	字辈	字	号	排行	谥号	生于	卒于	父亲	母亲
5036	胡	三公	義	鑑之		0		唐會昌癸亥三月初五	後唐天成丙戌五月		
5037	胡	昌翼公	昌	宏遠		1		天祐甲子三月一日	宋咸平己亥十月初三		
5038	胡	延政公	延	克修、以禮	節庵	1		後唐天成己丑十二月十五	宋大中祥符己酉	5037	詹氏
8898	胡	延賓公	延	以敬		2		後唐長興壬辰十一月初七		5037	詹氏
8897	胡	延臻公	延	以福		3		唐清泰甲午九月十五	宋大中祥符甲寅十月	5037	詹氏
5039	胡	百彥公	百			1				5038	汪氏、詹氏、秦氏
5040	胡	百祿公	百	廷堅	禹焦	2				5038	汪氏、詹氏、秦氏
8899	胡	百英公	百			3				5038	汪氏、詹氏、秦氏
8900	胡	百忠公	百			4				5038	汪氏、詹氏、秦氏
8901	胡	百順公	百			5				5038	汪氏、詹氏、秦氏
8904	胡	文義公	文			1				8898	
8902	胡	文昊公	文			1				8897	程氏
8903	胡	文晟公	文			2				8897	程氏
5041	胡	簡公		居敬		1				5039	汪氏、戴氏、趙氏
5043	胡	笪公				2				5039	汪氏、戴氏、趙氏
5042	胡	官公				3				5039	汪氏、戴氏、趙氏
5044	胡	簫公				4				5039	汪氏、戴氏、趙氏
5045	胡	簧公				5				5039	汪氏、戴氏、趙氏

# 上川明经胡氏世系表

## ——从结构化数据到机器可读的知识



先祖名人共  
8915 人  
配偶 4733 人

```
<http://data.library.sh.cn/jp/entity/person/uqwn21pa7iah9p87> a shi:Person ;  
  shi:name "洪骍" ;  
  foaf:name "胡适" ;  
  shi:courtesyName "适之" ;  
  foaf:familyName  
<http://data.library.sh.cn/authority/familyname/rvmgfsec8os93mv> ;  
  shi:orderOfSeniority "3" ;  
  shi:birthday "光緒辛卯十一月十七未時" ;  
  shi:description "考派美國留學生名適字適之事迹見學了生光緒辛卯十一月十七  
未時娶江氏生光緒庚寅十一月初八辰時" ;  
  shi:relatedWork  
<http://data.library.sh.cn/jp/resource/work/jklhb5c3ga1rvxe3>;  
  shi:roleOfFamily <http://data.library.sh.cn/jp/vocab/ancestor/xian-zu>;  
  rel:spouseOf <http://data.library.sh.cn/jp/entity/person/kvop3qdif1okxhoa>;  
  rel:childOf <http://data.library.sh.cn/jp/entity/person/mc7khkqgql3rpceb> .
```

# 从小众人文研究到大众通俗文化

这些数据讲了什么样的故事？

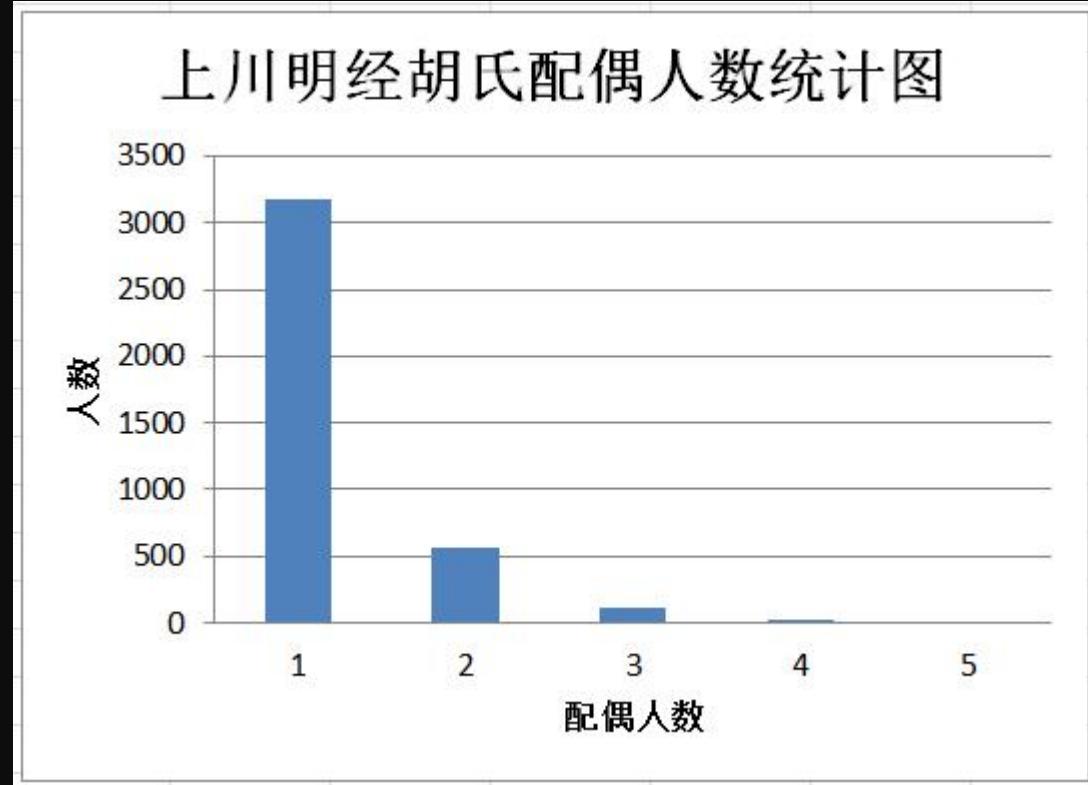


胡李不通婚的家规  
是真的吗？

从唐末至民初，真的没有一人娶过姓李的姑娘耶！

如果你是上川明经胡氏的后人，娶一位姓曹的女纸吧…

# 从小众人文研究到大众通俗文化



自唐末至民初的8915位男性成员中：  
娶了5位夫人的有4人；  
娶了4位夫人的有20人；  
娶了3位夫人的有109人；  
娶了2位夫人的有566人；  
娶了1位夫人的有3172人；  
有5043位孤单度过一生，占总人数的56.6%？？？

规范控制

# 从字符串 (String) 到东西 (Thing)

<http://data.library.sh.cn/entity/place/topq2nlfiugtg964>

```
a shl:Place;  
bf:label      "四川"@cht"四川"@chs;  
shl:abbreviateName    "川";  
shl:country     "中國"@cht"中国"@chs;  
owl:sameAs      <http://www.geonames.org/1815285>;  
gn:lat        30.67;  
gn:long       104.07.
```

<http://data.library.sh.cn/entity/place/1v1smmtsmqg29rlf>

```
a shl:Place;  
bf:label      "樂山"@cht"乐山"@chs;  
shl:province   "四川省";  
gn:parentADM2  <http://data.library.sh.cn/entity/place/topq2nlfiugtg964>;  
...
```

## 上海图书馆家谱文献列表 (953)

[高级检索](#)[简单检索](#)

标题

堂号

姓氏

先祖/名人

责任者

摘要

谱籍

馆藏信息

[查询](#)

## 上海图书馆家谱文献列表 (221)

[高级检索](#)[简单检索](#)

标题

堂号

姓氏

先祖/名人

责任者

摘要

谱籍

馆藏信息

[查询](#)

宣漢武氏族譜 (四川省達州市宣漢縣)

[责任者] 武澤宗

[摘要] 始祖國寧，明代人。始遷祖尚儒，明末自湖北麻城遷居四川宣漢。書名據封面題。



林氏家譜 (四川省達州市宣漢縣)

[责任者] 林因福

[摘要] 始遷祖長秀，清初由湖北省麻城縣移居四川省宣漢縣昆池壩林家壩。書名據書衣題。



岳氏族譜 (四川省南充市蓬安縣)

[责任者] 岳元虎

[摘要] 始祖同上。始遷祖之化，清初自湖北麻城縣移居四川南充縣北路會進橋（今屬四川蓬安）。書名據封面題。

**大山·水**  
2014.03.23 第3版

**巴山渝水**

## “湖广填四川” 重庆的移民记忆

“湖广填四川”起于元末均平初的洪武大移民，到了清代嘉庆、道光年间，“湖广填四川”达到高潮，湖广填四川”之说，至今已经600多年。

这一时期迁往四川、重庆的移民主要来自湖北、湖南、陕西、广东、福建、江西、广西、甘肃、江浙、浙江、鲁晋和云南等十余个省。尤以湖北、湖南为多，故有“江西填湖广，湖广填四川”之称。

在“湖广填四川”移民运动中，重庆有其较为特殊的地位——重庆是“湖广”移民避难归宿地。

入川后定居，繁衍、创业的需要地域，“湖广填四川”中移民的到来，使得渝湘两地农作物种植、饮食结构、语言等方面都有了改变。

至今，这些历史的印迹等都还依然存在。

### 五谷 麻婆豆腐、火锅都与移民有关

清朝乾隆时，重庆本地的农作物品种非常单一，除了水稻，几乎所有的主食都是从外省引进的。随着人口的增加，外来品种的传播也变得越来越快，品种也越来越丰富。而这时的重庆人，已经不再满足于本地的粮食作物，开始向更远的地方寻求新的粮源。于是，他们便开始大规模地向四川盆地移民，从而形成了“湖广填四川”的局面。

今天，我们所熟悉的麻婆豆腐、火锅等美食，其实都是从湖广地区传入重庆的。麻婆豆腐起源于四川省乐山市，而火锅则是从四川省自贡市传入重庆的。这两样美食的出现，都是因为湖广移民的到来，才使得重庆的饮食文化更加丰富多彩。

### 五谷 移民们带来 甘蔗南移至洋洋芋甘蔗

清朝末年，大量的移民从湖广等地迁入重庆，带来了先进的农业技术。其中，最为显著的就是甘蔗的引入。甘蔗原产于印度，明朝时传入中国，最初只在南方种植。但到了清朝末年，由于气候变暖，甘蔗开始向北方扩展。而在重庆，甘蔗的种植面积越来越大，逐渐成为了当地的主要经济作物之一。

甘蔗的引入，不仅改变了重庆的农业结构，也为当地居民提供了丰富的物质资源。甘蔗不仅可以作为糖料作物，还可以作为饲料、燃料等多种用途。因此，在重庆，甘蔗成为了人们生活中不可或缺的一部分。

### 五谷 “宝面”原指来自宝庆府的人

“宝面”一词最早出现在南宋时期，指的是来自湖南宝庆府（今邵阳）的移民。当时，宝庆府的移民大量涌入重庆，带来了先进的农业生产技术和丰富的劳动力。这些移民的到来，极大地促进了重庆的经济发展和人口增长。

“宝面”一词后来逐渐演变成了对重庆人的统称，特别是在重庆方言中，常常用来形容重庆人的外貌特征。例如，“宝面”就是指重庆人面部宽大、鼻梁低矮、皮肤黝黑等特点。

### 五谷 移民谋生催生了109个行帮

到了清末，重庆的商业活动非常繁荣，吸引了大量的外来商人。其中，来自湖广地区的商人占据了重要地位。他们带来了先进的商业理念和丰富的商业经验，促进了重庆商业市场的繁荣发展。

为了适应商业发展的需要，这些商人纷纷成立了各种行会组织。据统计，当时重庆共有109个行帮，涵盖了各行各业。这些行帮的成立，不仅促进了商业市场的规范管理，也为重庆的商业发展提供了有力的保障。

### 五谷 “巫山云雨”

“巫山云雨”是重庆著名的自然景观，也是“湖广填四川”移民的重要见证。据记载，早在唐朝时期，巫山就已经成为了著名的旅游胜地。到了宋朝，巫山的旅游业得到了进一步的发展，吸引了众多的游客前来观赏。

“巫山云雨”之所以闻名，是因为巫山的云雾缭绕，景色如画。每年春秋两季，巫山都会出现壮观的云海奇观，让人叹为观止。此外，巫山还有许多美丽的传说和故事，如“神女峰”、“巫山十二峰”等，更是增添了巫山的魅力。

# 为人文研究提供史料佐证 ——“家谱”与“正史”、“方志”一起构成人文研究的三大支柱。





上海图书馆

上海科学技术情报研究所

## 上海图书馆开放数据

### 地理名词表

国内

上海市

RDF | NT | XML | JSON x

在此输入朝代, 年号, 帝王姓名, 公元纪年检索

夏

商

周

西周

东汉

三国

晋

西晋

北宋

南宋

元

明

朝代    年号    帝王    帝王姓名

清    天命    太祖    爱新觉罗努尔哈赤

清    天聪    太宗    爱新觉罗皇太极

清    崇德    太宗    爱新觉罗皇太极



1880  
清



# 如何应用关联数据？



HTTP  
URI

用HTTP URI作为一切事物的名称

# 《上海图书馆URI设计规范》

设计模式： `http://{domain}/{type}/[concept][type]/{IRI}`

示例：

文献：<http://data.library.sh.cn/jp/resource/work/8y9p7s2euppwnerq>

人：<http://data.library.sh.cn/jp/entity/person/qxj915pkohm96unn>

姓氏：<http://data.library.sh.cn/authority/familyname/68n959cf8zdfkz3v>

地：<http://data.library.sh.cn/entity/place/tk5s4pej6linq9tr>

时：[sh.cn/authority/temporal/4alljneqivh5691](http://sh.cn/authority/temporal/4alljneqivh5691)



关系

# 知识本体(Ontology)的设计

# 家谱本体

定义文献、姓、人、地、时、事等概念及关联关系

The screenshot shows two tabs of an ontology editor:

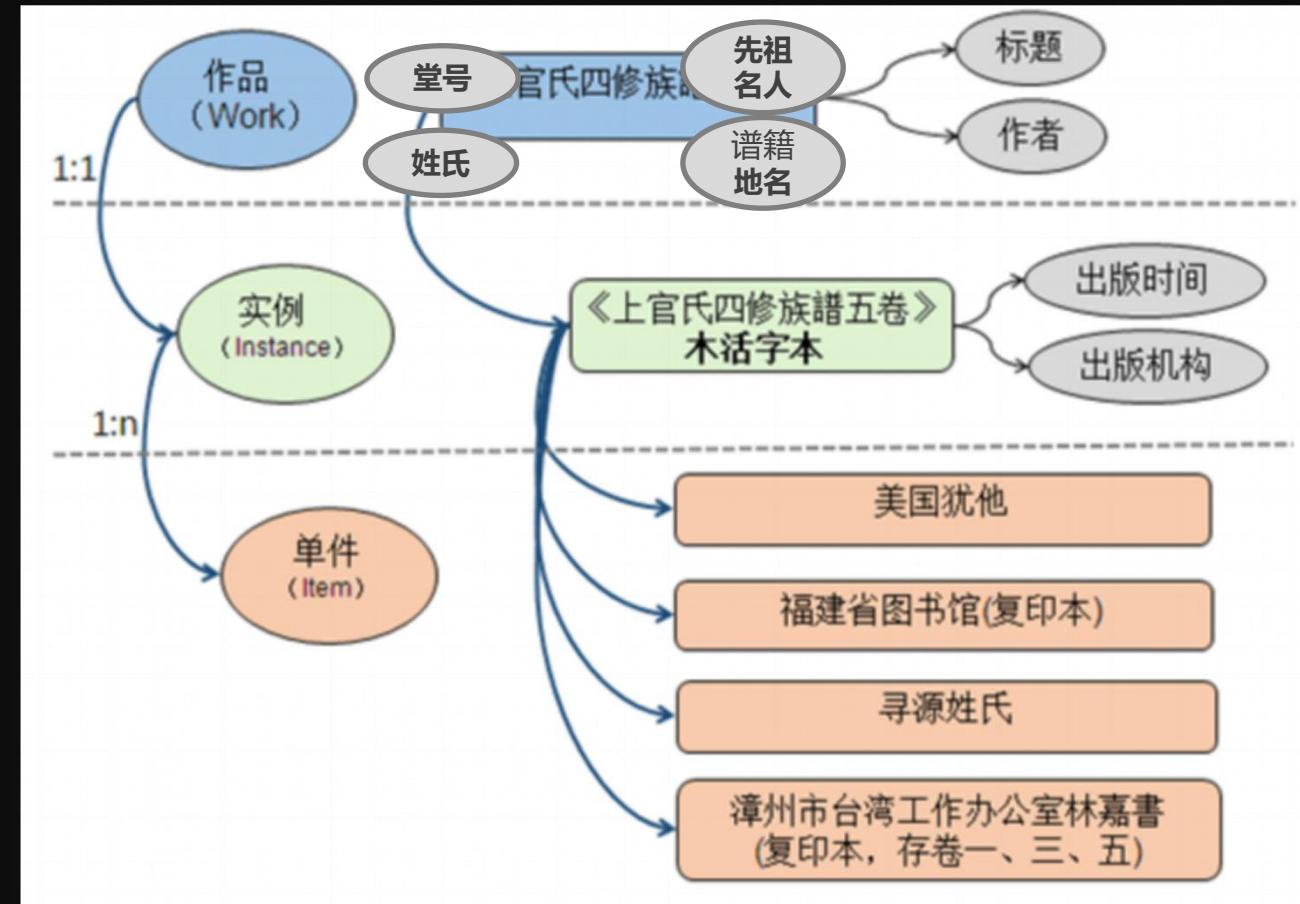
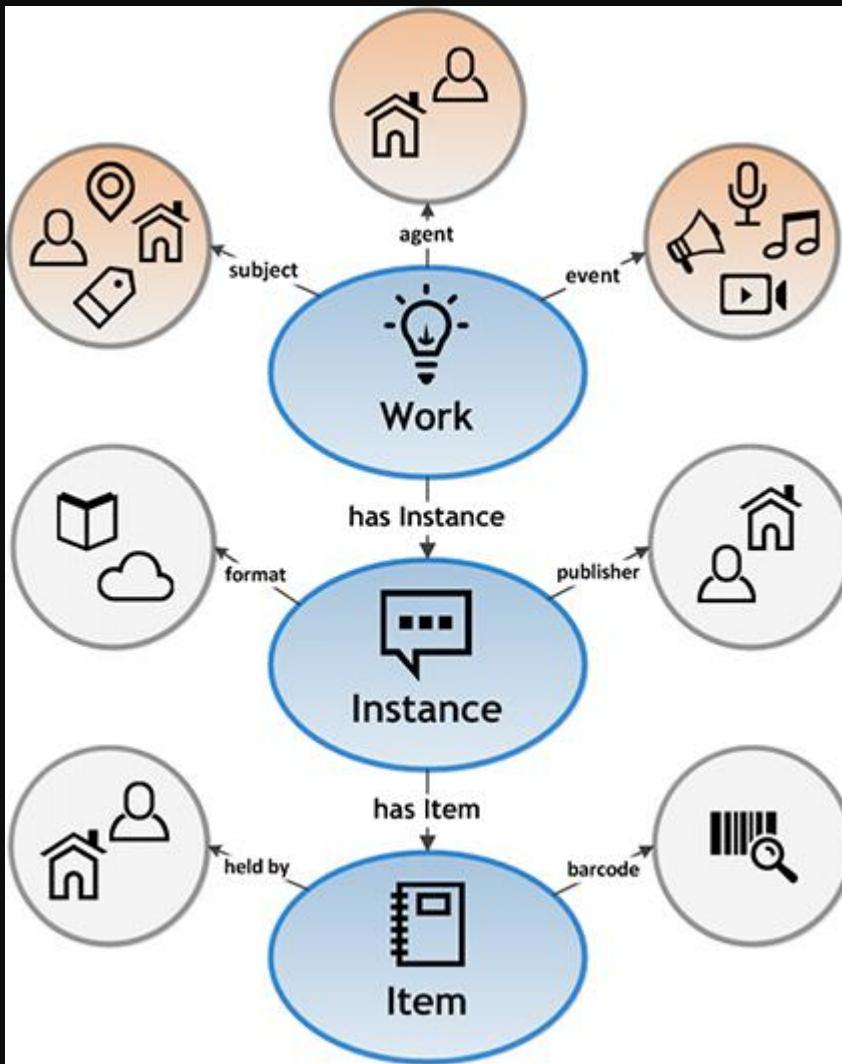
- Classes**: A tree view of classes. Rooted under "rdfs:Resource" are "bf:Resource" (marked with a plus sign), "shl:Resource" (marked with a minus sign), and several other namespaces like "foaf:Agent", "geo:SpatialThing", "shl:Agent", "shl:Event", etc.
- shl:Person**: A table listing properties for the "shl:Person" class. The columns are: Property, Subproperty of, Label, Used With, and Expected Value. The properties listed are:

Property	Subproperty of	Label	Used With	Expected Value
foaf:gender		性别	shl:Person	rdfs:Literal
rel:childOf		父	shl:Person	shl:Person
rel:spouseOf		配偶	shl:Person	shl:Person
shl:birthday		生于	shl:Person	rdfs:Literal
shl:courtesyName		字	shl:Person	rdfs:Literal
shl:deathday		卒于	shl:Person	rdfs:Literal
shl:genealogyName		谱名	shl:Person	rdfs:Literal
shl:generationCharacter		字辈	shl:Person	rdfs:Literal
shl:orderOfSeniority		行	shl:Person	rdfs:Literal
shl:posthumousTitle		谥号	shl:Person	rdfs:Literal
shl:pseudonym		号	shl:Person	rdfs:Literal
shl:roleOfFamily		角色	shl:Person	bf:Category

42

<http://gen.library.sh.cn:8080/ontology>

# 核心数据模型(BIBFRAME2.0)



## //Work

<<http://data.library.sh.cn/jp/resource/work/oyz2f36kouez9jdy>> a bf:Work;  
bf:creator <<http://data.library.sh.cn/jp/entity/person/y9fkgu3n87ccpa8m>>;  
shl:place<<http://data.library.sh.cn/entity/place/t3tec8y1oy2j3kjc>>;  
bf:subject <<http://data.library.sh.cn/authority/titleofancestraltemple/6biawgp5dbdm9hkm>>;  
bf:subject <<http://data.library.sh.cn/authority/familyname/4feibkhldtiroeu3>"

.....

## //Instance

<<http://data.library.sh.cn/jp/resource/instance/apz8aio37wj6y524>> a bf:Instance;  
bf:category <<http://data.library.sh.cn/vocab/binding/xian-zhuang>>;  
bf:edition <<http://data.library.sh.cn/vocab/edition/mu-zi-huo-ben>>;  
bf:extent "五册";  
shl:temporalValue 1936, "1936年" ;  
bf:instanceOf <<http://data.library.sh.cn/jp/resource/work/oyz2f36kouez9jdy>>;

.....

## //Item 1

<<http://data.library.sh.cn/jp/resource/item/fgxpi3bc8km3r672>> a bf:Item;  
bf:heldBy<<http://data.library.sh.cn/entity/organization/u0qz22aqnemd3idn>>;  
bf: itemOf<<http://data.library.sh.cn/jp/resource/instance/apz8aio37wj6y524>>;

## //Item 2

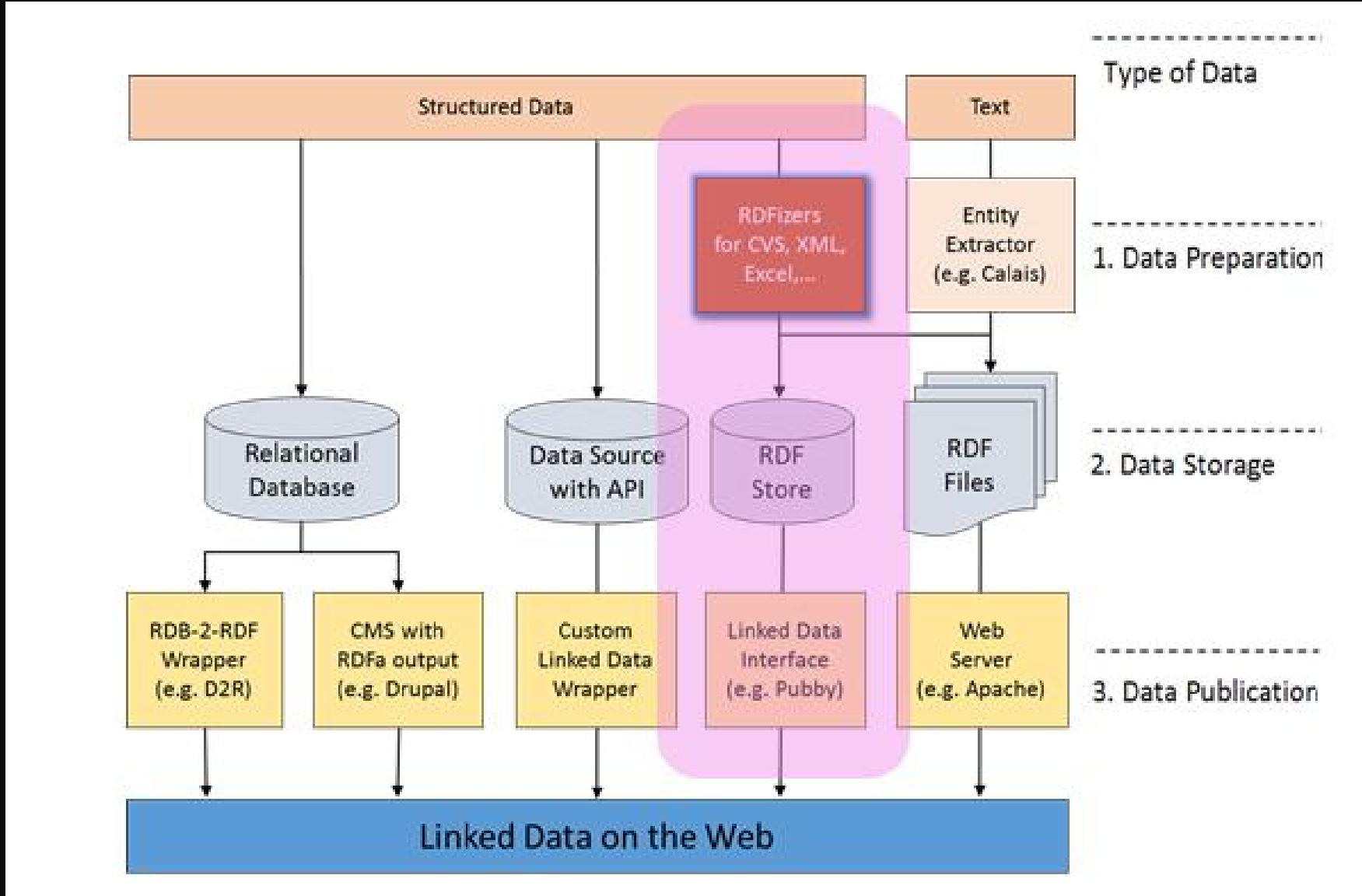
## //Item 3

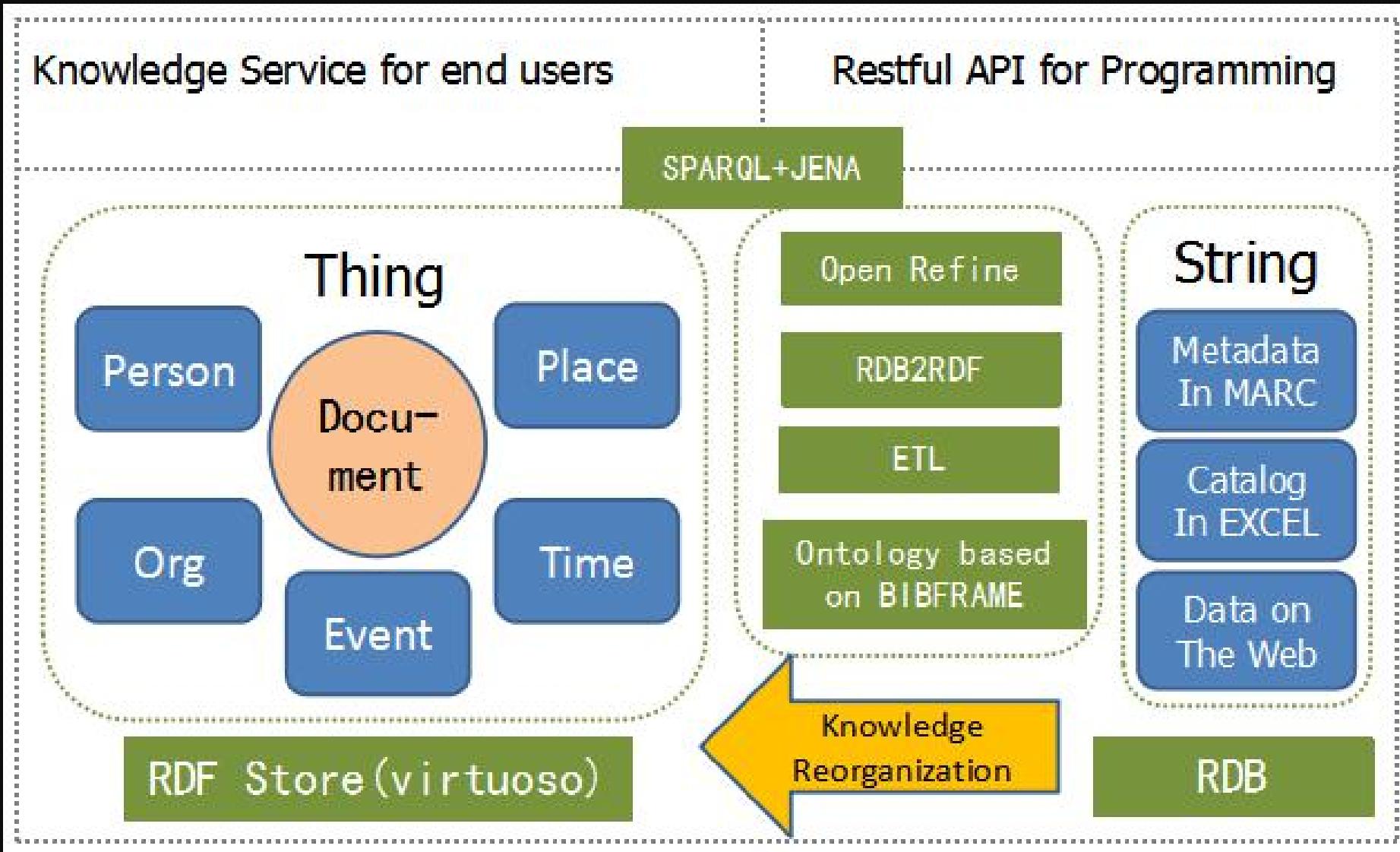
## //Item 4



当访问URI时提供RDF数据

# 生成并发布 RDF 数据

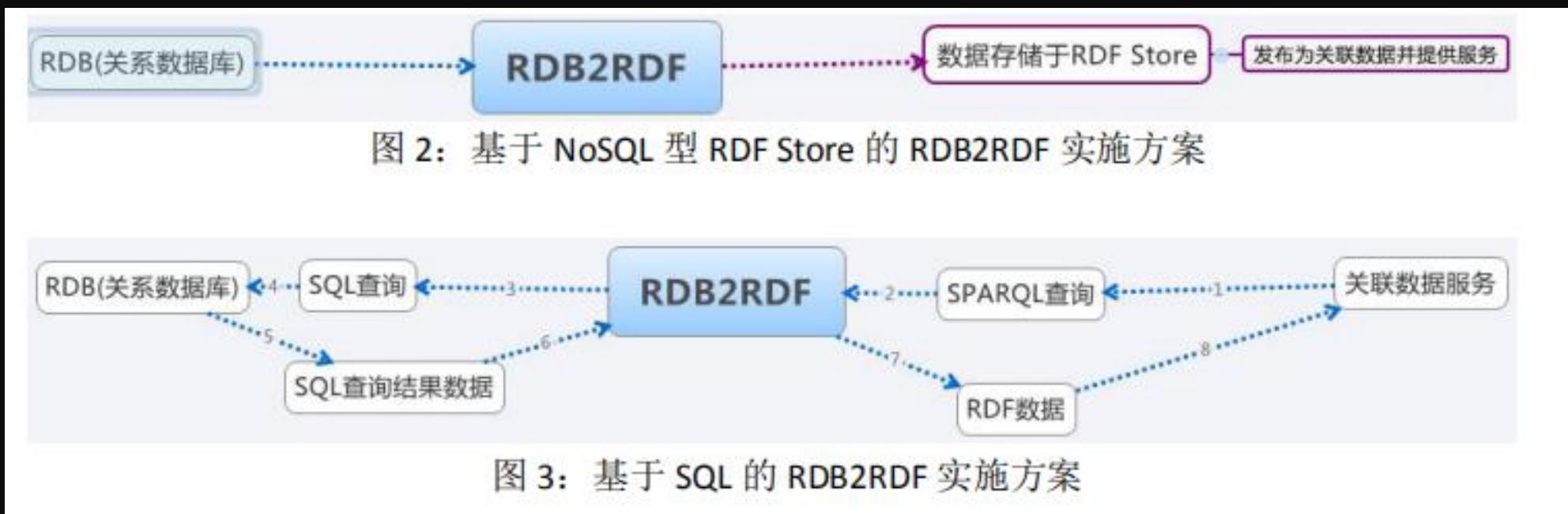




# RDB2RDF



图 1：关系数据库发布为关联数据并在 Web 上提供服务的示意图



# W3C的RDB2RDF推荐标准



R2RML: RDB to RDF Mapping Language

W3C Recommendation 27 September 2012

This version:

<http://www.w3.org/TR/2012/REC-r2rml-20120927/>

A Direct Mapping of Relational Data to RDF

W3C Recommendation 27 September 2012

This version:

<http://www.w3.org/TR/2012/REC-rdb-direct-mapping-20120927/>

R2RML 映射语言  
支持关系数据库表结  
构和自定义本体词表  
的映射

DM直接映射  
关系数据库表直接转  
换成RDF数据，表字  
段名即为RDF的属性  
名

# RDB2RDF工具

数据库初始化

数据库类型: Mysql

数据库地址: 127.0.0.1, 数据库端口: 3306, 数据库名称: Genealogy

数据库用户: root, 数据库密码: \*\*\*\*

连接数据库

配置

模式选择: R2RML

+ 添加命名前缀

映射配置

★ 名称:

是否需要查询构造表

1 SELECT ?s ?p ?o  
2 WHERE {?s ?p ?o}  
3 LIMIT 10

id: 166, 166, 166, 166, 166, 166, 166, 166, 166, 166

s: 234, 330, 330, 330, 330, 330, 330, 330, 330, 330

p: http://jp.library.sh.cn/ancestor/胡忠, http://jp.library.sh.cn/ontology/era, http://jp.library.sh.cn/ancestor/胡忠, http://jp.library.sh.cn/ontology/era, http://jp.library.sh.cn/ancestor/胡延政, http://jp.library.sh.cn/ontology/era, http://jp.library.sh.cn/ancestor/胡延政, http://jp.library.sh.cn/ontology/era, http://jp.library.sh.cn/ancestor/胡七二, http://jp.library.sh.cn/ontology/era

o: 宋, 宋, 宋, 明, 明, 明, 宋, 宋

查询

Ancestor

胡忠

胡宗文

胡麟豹

胡士良

胡彦诚

胡昌翼

胡七二

胡延政

图展示 内容展示 配置下载 RDF下载

The screenshot displays the RDB2RDF tool's user interface. On the left, there's a configuration panel for database initialization (Mysql, host 127.0.0.1, port 3306, database Genealogy) and mapping configuration (R2RML). In the center, a query editor shows a SPARQL query for retrieving ancestor data, and the results are displayed in a table. A graph on the right visualizes the family tree with nodes for individuals like 胡忠, 胡宗文, 胡麟豹, etc., and edges representing their relationships. At the bottom, there are tabs for '图展示' (Graph Display), '内容展示' (Content Display), '配置下载' (Configuration Download), and 'RDF下载' (RDF Download).

<http://data.library.sh.cn/tools/>

# Any2RDF ETL

Any2RDF 上海图书馆开放数据平台  
ETL 数据采集转换工具

任务

- 数据采集 >
- 数据加工 >
- 数据转换 >

模板

- 过滤模板
- 抽取模板
- RDF映射模板

资源

- Iri列表

任务列表

ID	任务名称	创建时间	结束时间	详情	采集源	状态	操作
32	6547	2017-03-23 09:38:19	2017-03-23 09:38:32	采集至2/2条 数据	百度百科	已完成	
31	0323	2017-03-23 09:27:02	2017-03-23 09:27:13	采集至2/2条 数据	百度百科	已完成	
30	20170322	2017-03-22 16:42:29	2017-03-22 16:42:35	采集至2/2条 数据	百度百科	已完成	

+ 新建任务

<http://data.library.sh.cn/tools/>

# OpenRefine

代	personID	father	谱名	名	字	号	排行	生	卒	迁徙	说明
42	16610	33134	洪骍	适	适之		三	光绪辛卯 上品上			考派美国留学生名适字适之事迹见学了生光绪辛卯十一月十七未时娶江氏生光绪庚寅十一月初八

Base URI: <http://gen.library.sh.cn/> edit

RDF Skeleton    RDF Preview

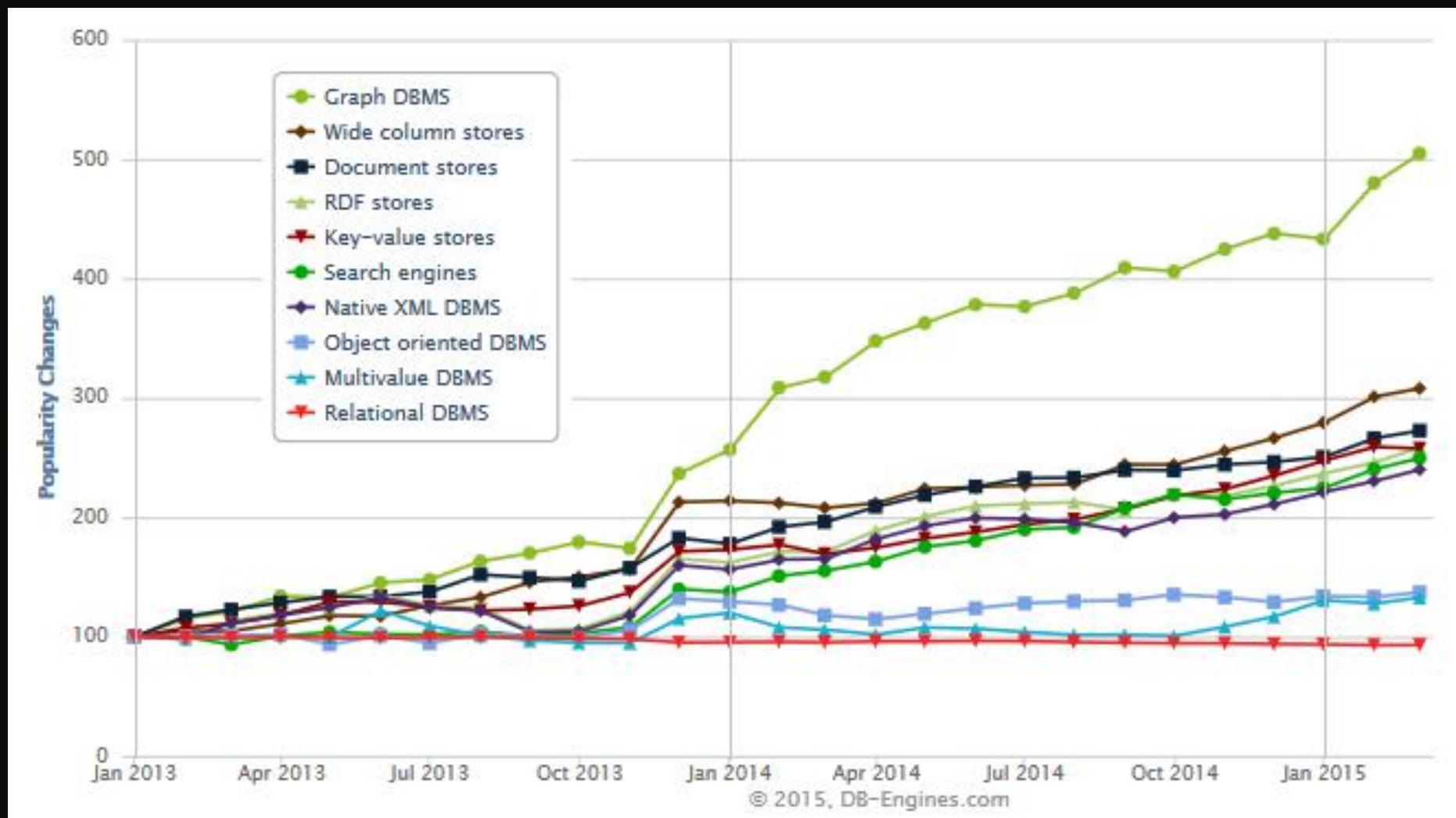
Available Prefixes: rdfs foaf xsd owl rdf rel shl + add prefix manage prefixes

personID URI	▷ shl:name →	▷ 谱名 cell
xshl:Person	▷ foaf:name →	▷ 名 cell
add rdf:type	▷ shl:courtesyName →	▷ 字 cell
	▷ shl:family →	▷ http://.../11268 + ... add rdf:type
	▷ shl:familyName →	▷ FamilyName/221 + ... add rdf:type
	▷ shl:orderOfSeniority →	▷ 排行 cell
	▷ shl:pseudonym →	▷ 号 cell
	▷ shl:birthday →	▷ 生 cell
	▷ shl:deathday →	▷ 卒 cell

# OpenRefine

```
<http://data.library.sh.cn/jp/entity/person/uqwn21pa7iah9p87> a shl:Person ;
  shl:name "洪骍" ;
  foaf:name "胡适" ;
  shl:courtesyName "适之" ;
  foaf:familyName <http://data.library.sh.cn/authority/familyname/rvmgjfsec8os93mv> ;
  shl:orderOfSeniority "3" ;
  shl:birthday "光緒辛卯十一月十七未時" ;
  shl:description "考派美國留學生名適字適之事迹見學了生光緒辛卯十一月十七未時娶江氏生光緒庚寅十一月初八辰時" ;
  shl:relatedWork <http://data.library.sh.cn/jp/resource/work/jklhb5c3ga1rvxe3>;
  shl:roleOfFamily<http://data.library.sh.cn/jp/vocab/ancestor/ming-ren>;
  rel:spouseOf <http://data.library.sh.cn/jp/entity/person/kvop3qdif1okxhoa>;
  rel:childOf <http://data.library.sh.cn/jp/entity/person/mc7khkqgql3rpceb> .
```

# RDF数据如何存储？



[http://db-engines.com/en/blog\\_post/43](http://db-engines.com/en/blog_post/43)

# GraphDB

# 还是 RDF Store?

强调数据与数据之间的关系

节点 (node) — 属性 (Property) — 节点 (node)

---

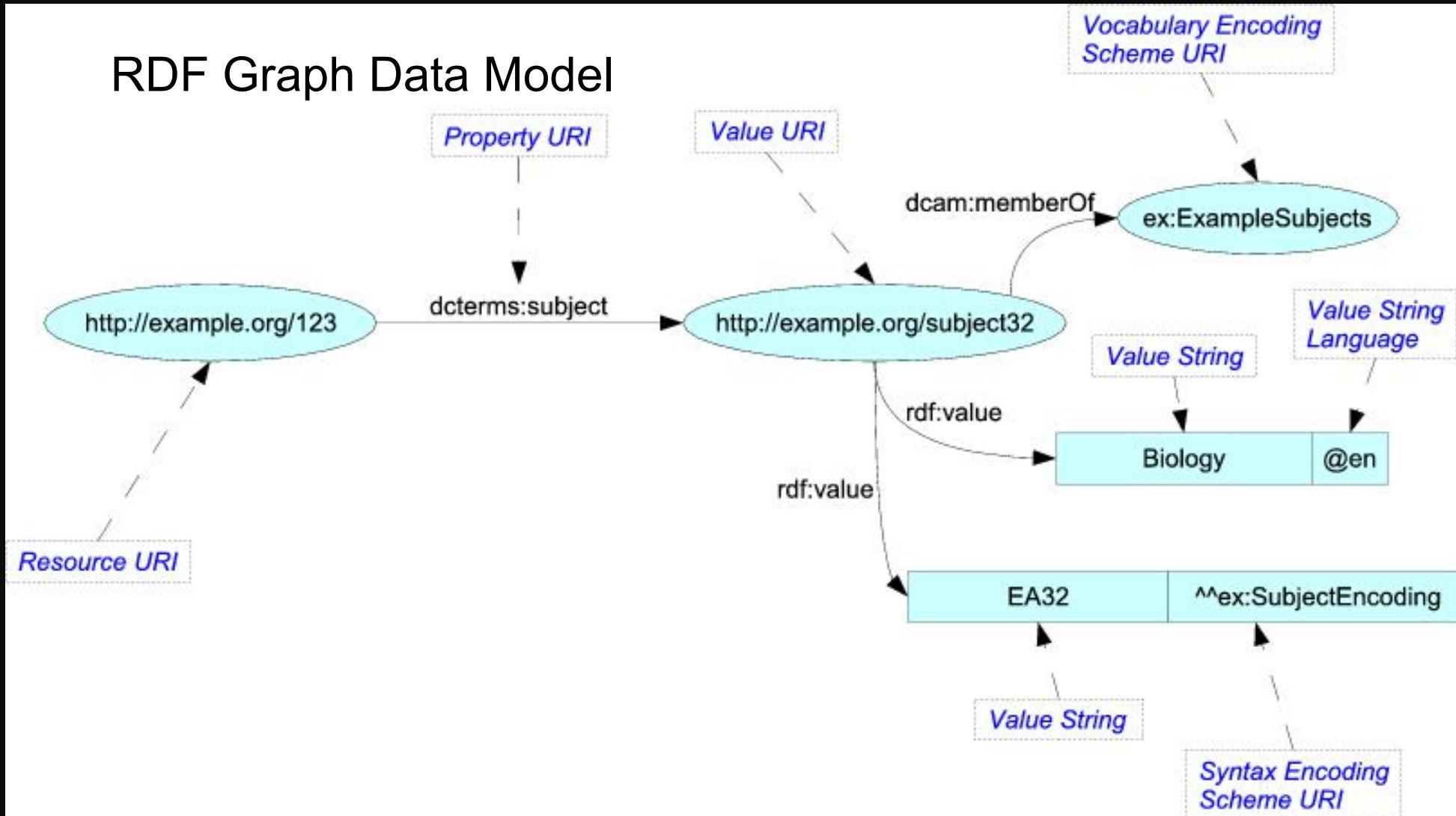
存储数据类型多：超图、无向图、权重图 (undirected graphs, weighted graphs, hypergraphs)

查询语言：

Cypher (neo4j) ,G,  
GraphLog, GOOD, SoSQL,  
BiQL, SNQL.....

存储数据类型仅为RDF三元组  
查询语言： SPARQL

# RDF Graph Data Model



# 如何提供RDF数据——SPARQL Endpoint

上海图书馆 Shanghai Library  
基于BibFrame的家谱本体 A Genealogy Ontology Based On BibFrame Vocabulary

上海图书馆 Home Model View Class View List View About (关于本站) Contact (联系我们) → RDF

## Class View

<http://gen.library.sh.cn:8080/ontology>

### Classes

- foaf:Agent
  - foaf:Person
    - shi:Person
  - + org:Organization
- + geo:SpatialThing
- rdfs:Resource
  - + bf:Resource
  - + shi:Resource
- + schema:Thing
- time:TemporalEntity
  - time:Instant
  - + time:Interval

### shi:Person

个人。

Property	Subproperty of	Label	Used With	Expected Value
foaf:gender		性别	shi:Person	rdfs:Literal
rel:childOf		父	shi:Person	shi:Person
rel:parentOf		子女	shi:Person	shi:Person
rel:spouseOf		配偶	shi:Person	shi:Person
shi:birthday		生于	shi:Person	rdfs:Literal
shi:courtesyName		字	shi:Person	rdfs:Literal
shi:deathday		卒于	shi:Person	rdfs:Literal
shi:genealogyName		谱名	shi:Person	rdfs:Literal
shi:generationCharacter		字辈	shi:Person	rdfs:Literal
shi:migration		迁徙	shi:Person	shi:Migration
shi:orderOfSeniority		行	shi:Person	rdfs:Literal
shi:posthumousTitle		谥号	shi:Person	rdfs:Literal
shi:pseudonym		号	shi:Person	rdfs:Literal
shi:roleOfFamily		角色	shi:Person	bf:Category

Virtuoso SPARQL Query Editor

About | Namespace Prefixes | Inference rules

Default Data Set Name (Graph IRI)

Query Text

```
select distinct ?Concept where {[] a ?Concept} LIMIT 100
```

(Security restrictions of this server do not allow you to retrieve remote RDF data, see [details](#).)

Results Format:

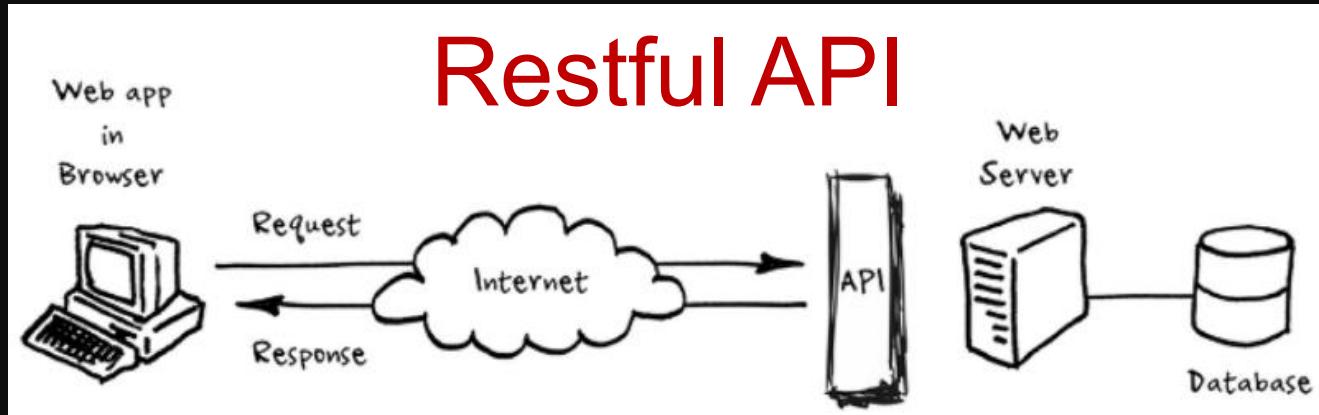
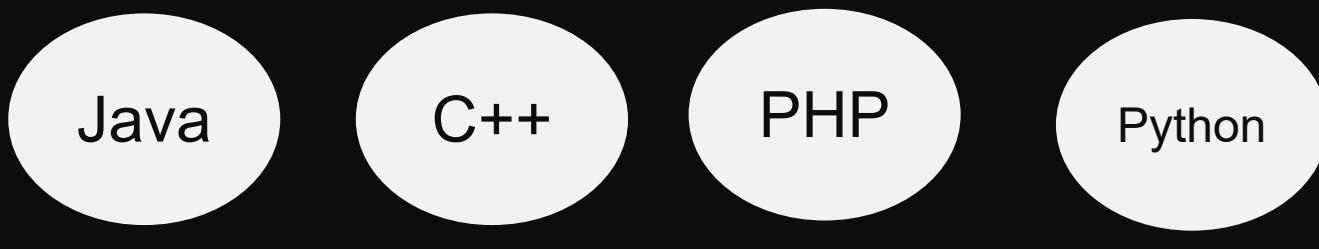
Execution timeout:  milliseconds (values less than 1000 are ignored)

Options:  Strict checking of void variables

(The result can only be sent back to browser, not saved on the server, see [details](#))

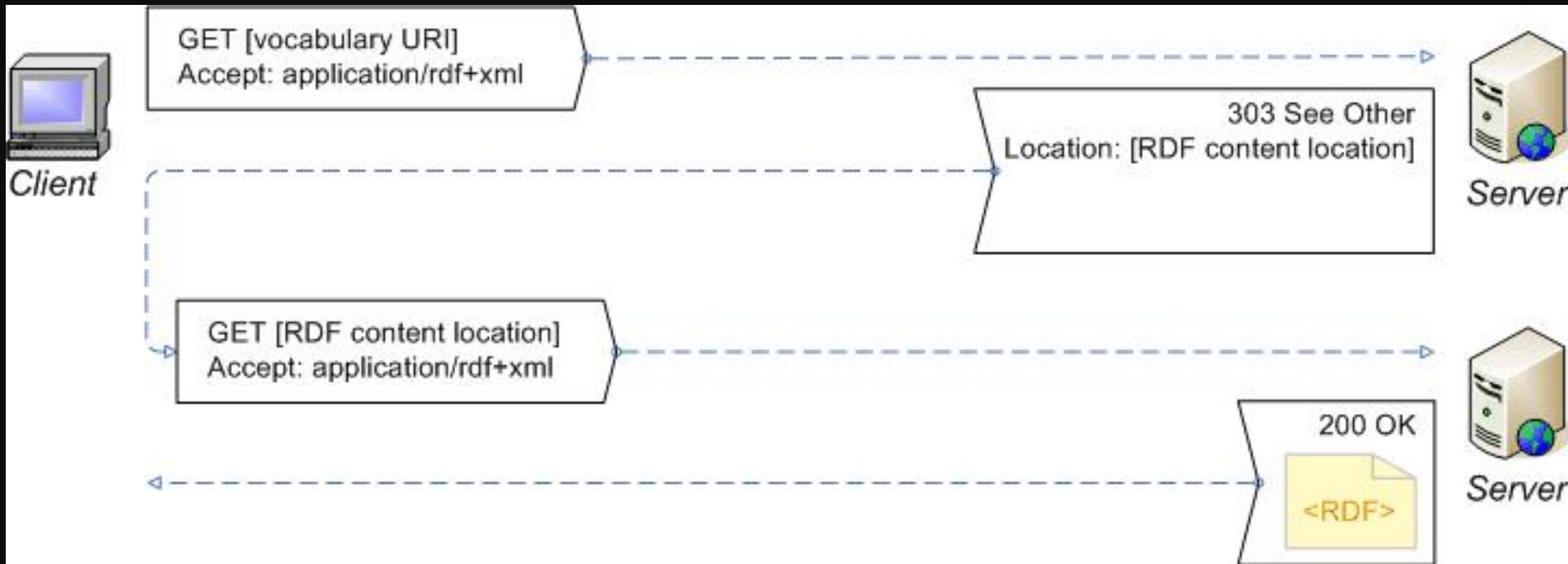
<http://data.library.sh.cn:8890/sparql>

# 如何提供RDF数据——Restful API



[http://data.library.sh.cn/jp/person/data?\[参数1\]&\[参数2\]&\[参数3\]?key=YourAPIKey](http://data.library.sh.cn/jp/person/data?[参数1]&[参数2]&[参数3]?key=YourAPIKey)

# 如何提供RDF数据——内容协商



一个URI标识的资源可以有不同的表示形式，如HTML、RDF、PDF或不同的语种等。因此，当客户端参引URI时，服务器端需要根据其发送请求的HTTP报头返回合适的表示形式。

开放数据服务  
让数据在流动中增值



SPARQL  
Endpoint

Restful  
API

HTTP  
URI

JSON  
-LD

返回数据

数据消费接口

# 关联开放数据 (Linked Open Data) ——从本地数据到公共知识

<http://pcrc.library.sh.cn/zt/opendata/>

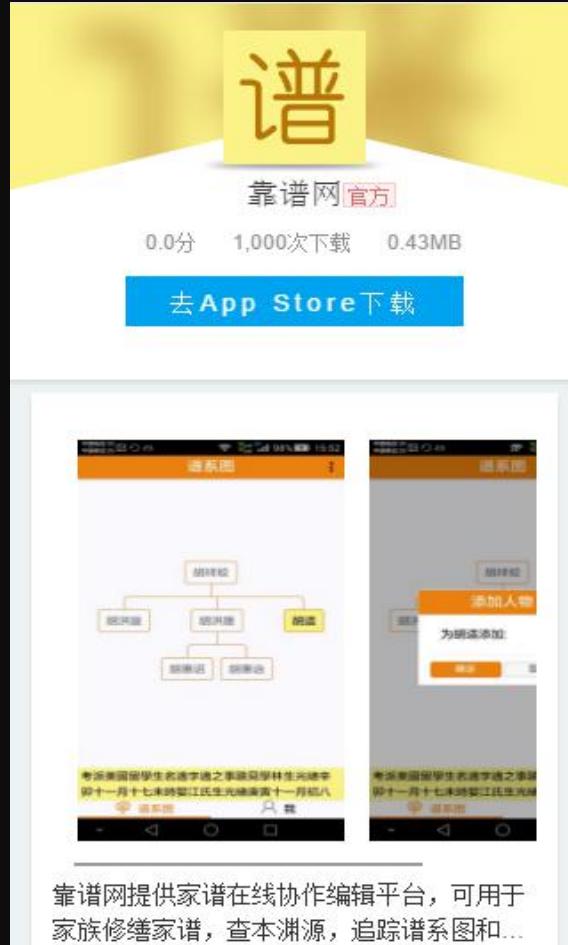
SHANGHAI LIBRARY

## 开放数据 应用开发竞赛

作品提交要求文档 (更新)

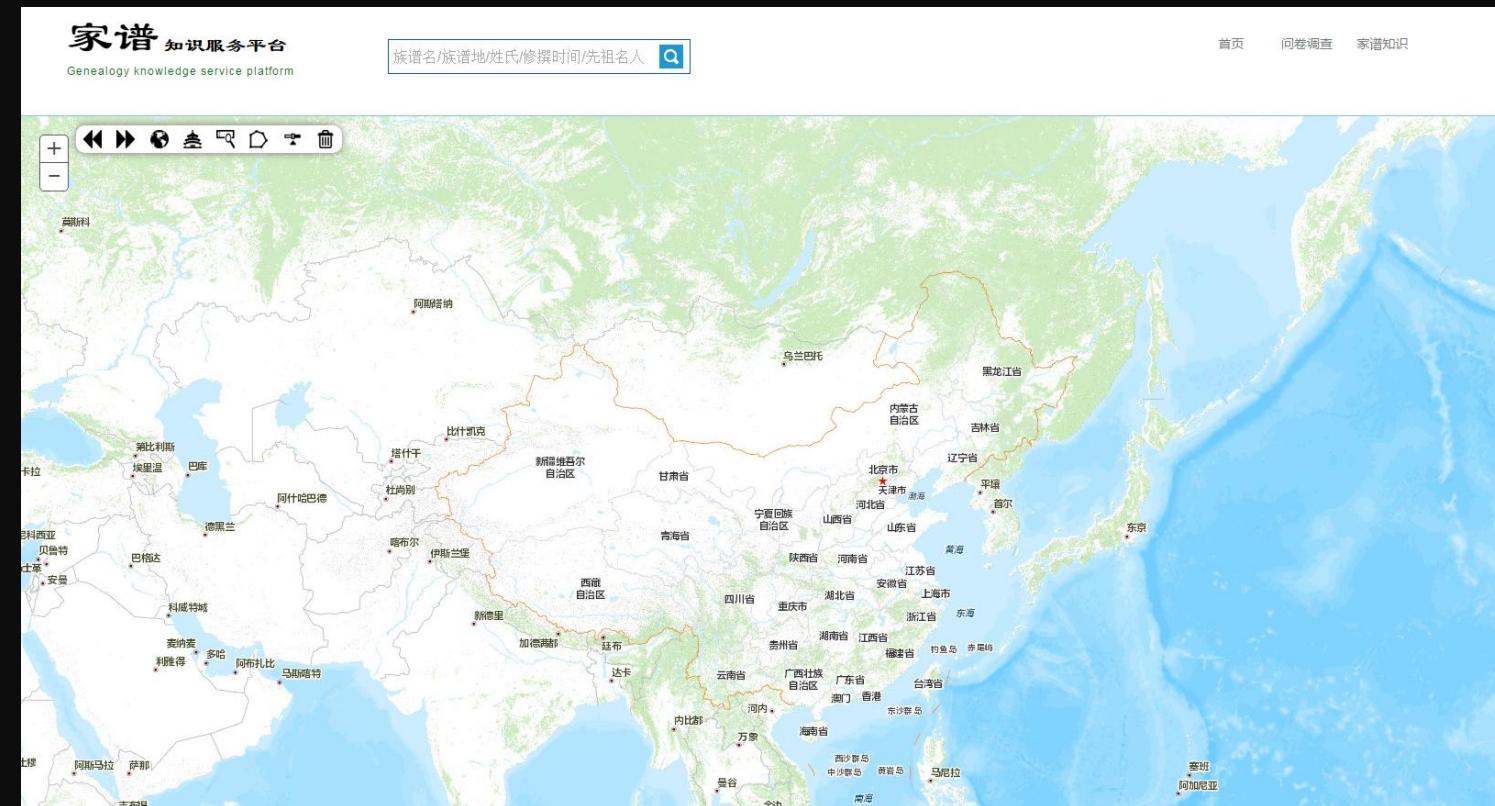
N

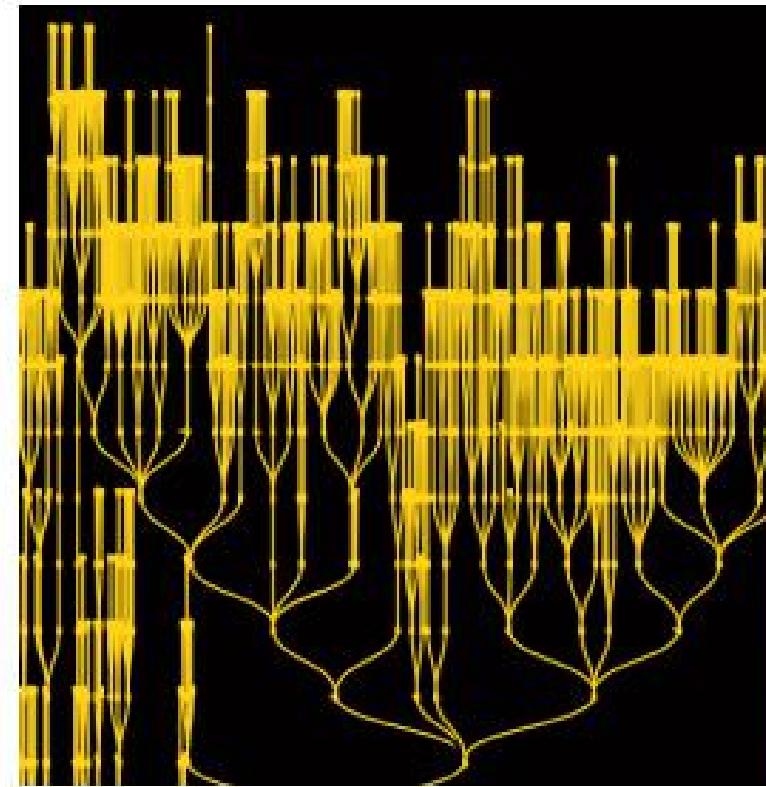
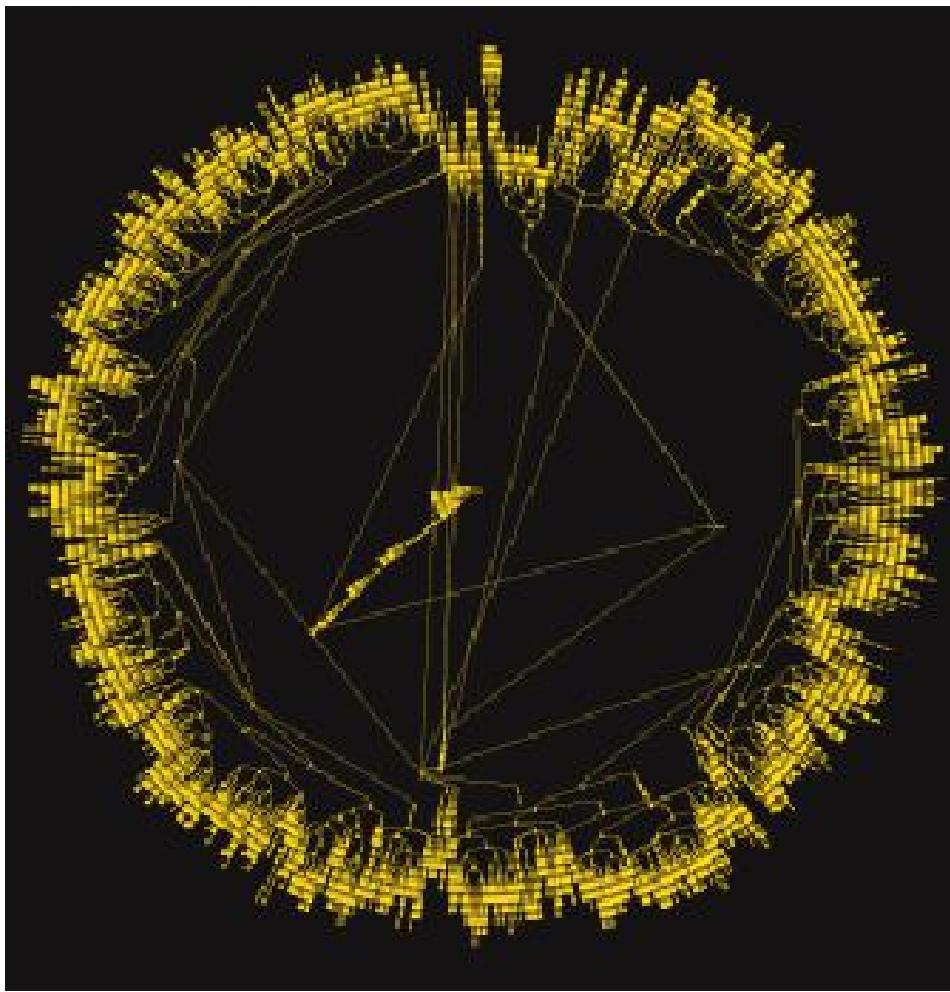
# 获奖作品



# Learn Chinese Surnames

上海图书馆开放数据应用开发竞赛作品





謝謝！  
Q&A！

cjxia@libnet.sh.cn