

The First ICSU World Data System Conference

Managing and Linking Scientific Data on the Web via VisualDB

Zhihong SHEN, Jianhui LI, Chengzan LI, Xing HE, Xianming SU

Scientific Data Center
Computer Network Information Center, Chinese Academy of Sciences



Outline

1. Background and Goals of VisualDB
2. Software architecture and features
3. Developing user applications based on VisualDB
4. Application cases
5. Future work

Background and Goals of VisualDB

Managing and Linking Scientific Data on the Web via VisualDB

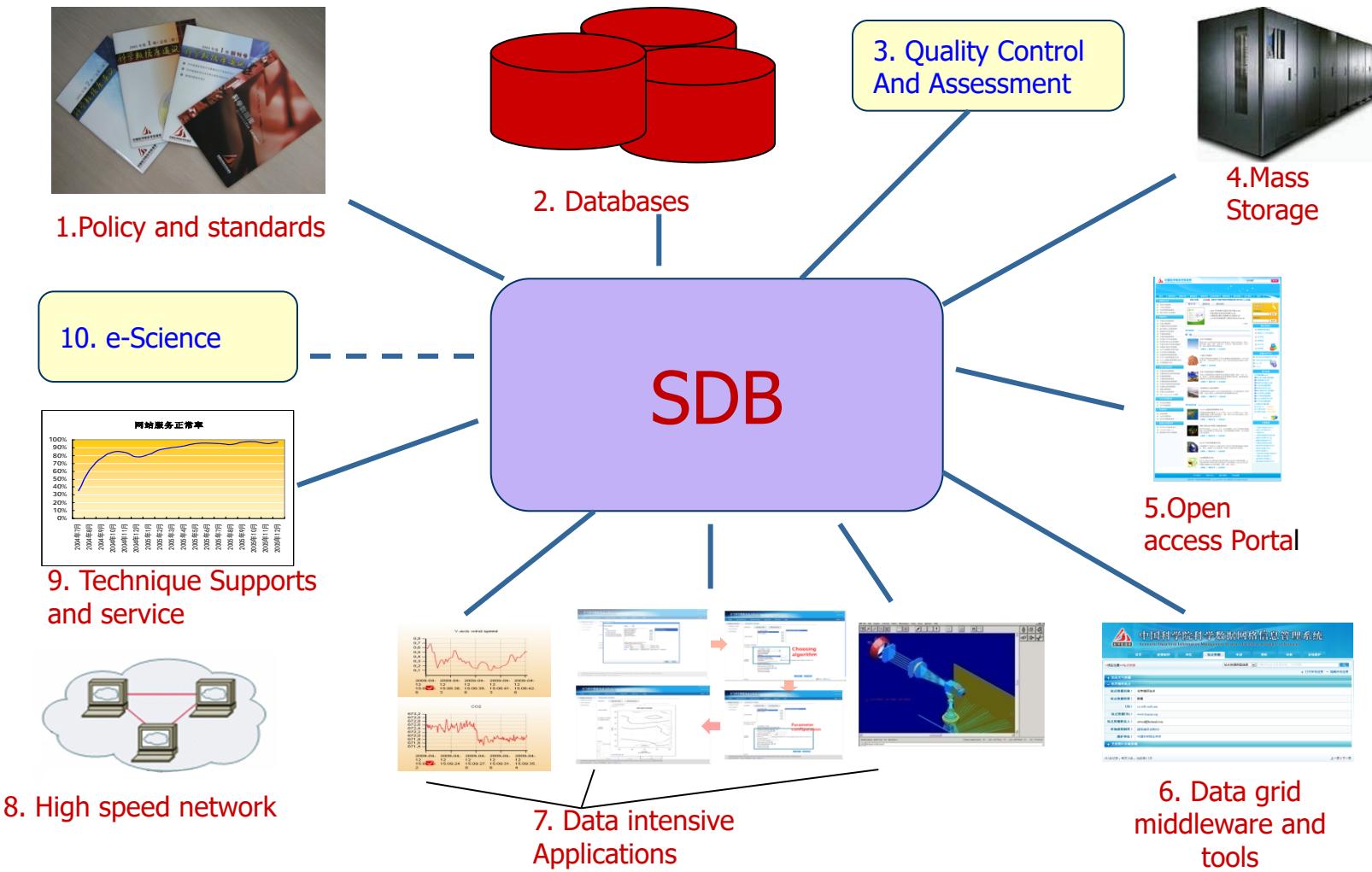
Background - Scientific DataBase project

□ CAS(Chinese Academy of Sciences) Scientific DataBase (SDB) project

- A Long-term mission started in 1986 which was funded by CAS
 - data from research, for research
- Collecting multi-discipline research data and promoting data sharing (2006-2010)
 - More than **500** research databases/datasets by 61 institutes, CAS
 - Over **200TB** data available to open access and download

Background - Scientific DataBase project

SDB architecture



Background- Scientific Databases

- 8 Resource databases
 - Geo-Science
 - Biodiversity
 - Chemistry
 - Astronomy
 - Space Science
 - Micro biology and virus
 - Material science
 - Environment
- 37 discipline databases

- 2 Reference databases
 - China Species
 - compound
- 4 Application-Oriented databases
 - High Energy (ITER)
 - Western Environment Research
 - Ecology research
 - Qinghai Lake Research

人地系统主题数据库

Heterogeneous Database for Human-earth System

分类查询 分类浏览 选择查询 选择浏览 关键字查询 分省数据查询 分县数据查询 交界数据查询
资源图集 经济图集 航片目录 生成地图 生成统计图 数据下载 模板数据下载 欢迎您的留言

人地系统主题数据库简介

人地系统主题数据库是面向人地系统基础研究、国家经济建设和社会战略需求，以人口、资源、环境和发展（PERD）为核心的数据库服务体系。它由中科院信息中心提供支持、中科院地理科学与资源研究所承建。人地系统主题数据库是在完善原中国自然资源数据库的基础上，通过整合东北黑土区、黄土高原、西南山地地区人地关系模型区域以及中国周边与全球主要国家（地区）的人地系统数据形成。数据库的数据涉及自然资源、环境、人口、社会经济、生态等多个方面。…… [详细]

使用指南

■ 数据库使用方法
人地系统主题数据库服务系统功能丰富，使用简单、方便。
具有数据查询浏览、数据下载、图形浏览和生成以及统计数据可视化与在线分析功能。…… [详细]

服务案例

■ 国区域可持续发展模拟与决策支撑平台
■ 京津冀都市圈区域规划

数据检索

库推荐

自然资源及其开发利用



灾害及其治理专题库

中国科学院天文学科学数据库

首页

查找数据库

数据服务

使用指南

服务案例



数据库简介

天文科学数据主题库将国家天文台科学数据和世界数据中心中心天文台，以及联合申请单位上海天文台和紫金山天文台，多年来的积累基础上，继续集成综合申请单位，本学科的科学数据，丰富数据源体系；构建网络化的科学数据共享环境，努力成为天文信息化的数据服务平台；建立科学数据服务体系，推动天文数据在科研中的应用；推动天文科学数据共享政策的制订和实施。…… [详细]



推荐数据库

南极CSTAR射电数据库

CSTAR (Chinese Small Telescope ARray) 是中国制造的第一台南极天文望远镜。CSTAR由4台相同的探测器组成。…… [更多]

热数据简介 | 元数据 | 共享声明 | 数据检索



服务案例

为国家植物西南种质资源库提供技术服务

更多>>

咨询台

QUESTION&ANSWER

Established databases

大气科学数据库

Atmospheric Scientific Database

首頁 檢索數據庫 數據檢索 數據服務 標準/軟件 作品展示 常見問題

大氣科學數據庫

大氣科學數據庫主要收集整理大氣科學研究常用數據，並提供進一步數據開發所需的網絡化研環境。目前包括四個子庫：大氣科學數據庫、大氣科學文獻庫、圖形可視化軟件庫、大氣科學算法和模式庫。系統的總量已接近1390GB，在線數據達到600GB。

站內檢索

標題 作者 檢索

中国植物物种信息数据库

Scientific Database of China Plant Species

首頁 名錄索引 志書式檢索 分等級檢索 數據查詢 云南植物志

数据库简介

数据库目录

植物物种信息数据库 云南高等植物数据库 云南经济植物数据库 有花植物数据库 物种行政区分布数据库

数据库推荐

中国植物物种信息数据库

该数据库由植物学学科积累深厚和资源丰富的中科院昆明植物研究所、中科院植物所、中科院武汉植物园和中科院华南植物园联合建设，面对国家重大资源和重大领航前沿需求，紧密围绕物种独特性、稀有的、或成的植物数据库，充分运用植物学、植物资源学、区域地理学等有关理论、方法和技术，在顶层设计的基础上，植物学专家（参考型数据库）。…… [詳細]

冰雪冻土环境本底与可持续发展专题数据库

The database of environmental background and sustainable development of glacial and frozen regions

首頁 查找数据库 數據檢索 應用服務 重要服務 使用指南 服務案例 建設隊伍

冰雪冻土环境本底与可持续发展专题数据库

冰雪冻土环境本底与可持续发展专题数据库主要收集冰川冻土、多年冻土、冻土退化与恢复、冻土与气候变化、冻土与植被、冻土与水文、冻土与土壤、冻土与生物、冻土与工程、冻土与灾害、冻土与生态、冻土与气候、冻土与区域水文、冻土与区域生态等。…… [詳細]

数据服务

冰雪冻土环境本底与可持续发展专题数据库

冰雪冻土环境本底与可持续发展专题数据库主要服务于冰川冻土、多年冻土、冻土退化与恢复、冻土与气候变化、冻土与植被、冻土与水文、冻土与土壤、冻土与生物、冻土与工程、冻土与灾害、冻土与气候、冻土与区域水文、冻土与区域生态等。…… [詳細]

灾害服务

冰雪冻土环境本底与可持续发展专题数据库

冰雪冻土环境本底与可持续发展专题数据库主要服务于冰川冻土、多年冻土、冻土退化与恢复、冻土与气候变化、冻土与植被、冻土与水文、冻土与土壤、冻土与生物、冻土与工程、冻土与灾害、冻土与气候、冻土与区域水文、冻土与区域生态等。…… [詳細]

聚变数据库

Fusion Database

首頁 查找数据库 數據檢索 標準/軟件 數據服務 使用指南 服務案例

聚变数据库

聚变数据库在收集丰富聚变数据的基础上，辅以分析、使用聚变数据库所收集数据。聚变数据库包括聚变材料数据库、聚变部件数据库和等离子体物理实验数据中聚变数据库，采用聚变中物理设计计算所需各种核聚变材料的数据。…… [詳細]

数据服务

聚变数据库

聚变数据库在收集丰富聚变数据的基础上，辅以分析、使用聚变数据库所收集数据。聚变数据库包括聚变材料数据库、聚变部件数据库和等离子体物理实验数据中聚变数据库，采用聚变中物理设计计算所需各种核聚变材料的数据。…… [詳細]

联系方式

联系人：胡丽琴
电 话：0511-6370348
传 真：0511-6370348

Background- Scientific Databases - toolsets

Vooyle



<http://vdb.csdb.cn>



SDB Single-sign-on Service



<http://resstat.csdb.cn>
Resource statistic system



<http://msis.csdb.cn>
Service monitoring system



Service monitoring system



Goals of VisualDB

□ VisualDB^[1]

- Developed by SDC, CNIC, since 2002
- Now VisualDB v2.0 is available

□ Aims at

- Providing web-based, easy-to-use, customizable tools
 - **data managers know little about SQL, DBMS, HTML, Java...**
- to help data managers manage and publish relational data or metadata and files
- under strict access control
- And also providing web services for computer programs to access data remotely
- Enabling link all databases together to construct a whole **database** (or a web of data)

[1]VisualDB homepage. <http://vdb.csdb.cn>

Software architecture and features

Managing and Linking Scientific Data on the Web via VisualDB

Software architecture and features

□ VisualDB consists of

VdbEngine

- An invisible **VdbEngine** which organizes all data into a whole database

VdbTools

- A set of **VdbTools** which help data managers manage and publish their data

VdbServer

- A **VdbServer** which serves local database as an accessible data source on the Web

VdbFramework

- A **VdbFramework** which help developers build applications (VdbApps) based on VisualDB

Software architecture and features

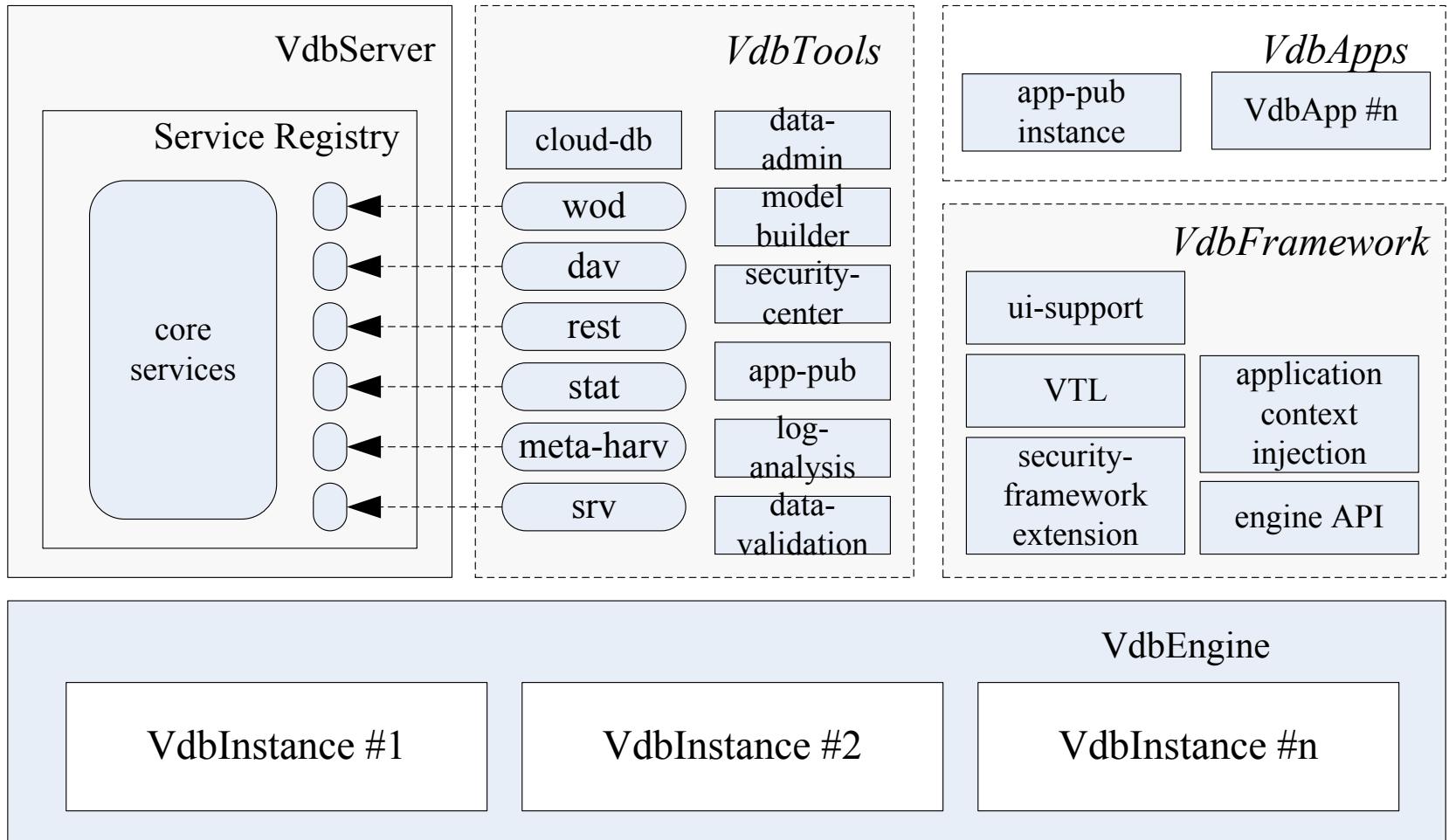
□ VdbEngine's work

- 1 • connects relational records and files distributed in multiple repositories
- 2 • maps physical records and files into logical entities, according to entity models
- 3 • links records, records and files
• extracts metadata of files[1][2]
- 4 • maps records with fields, or files with merged metadata into resources
- 5 • provides a language to query over all resource descriptions

[1] NIE (Nepomuk Information Element) Ontology Framework. <http://www.semanticdesktop.org/ontologies/nie/>

[2] <http://aperture.sourceforge.net>

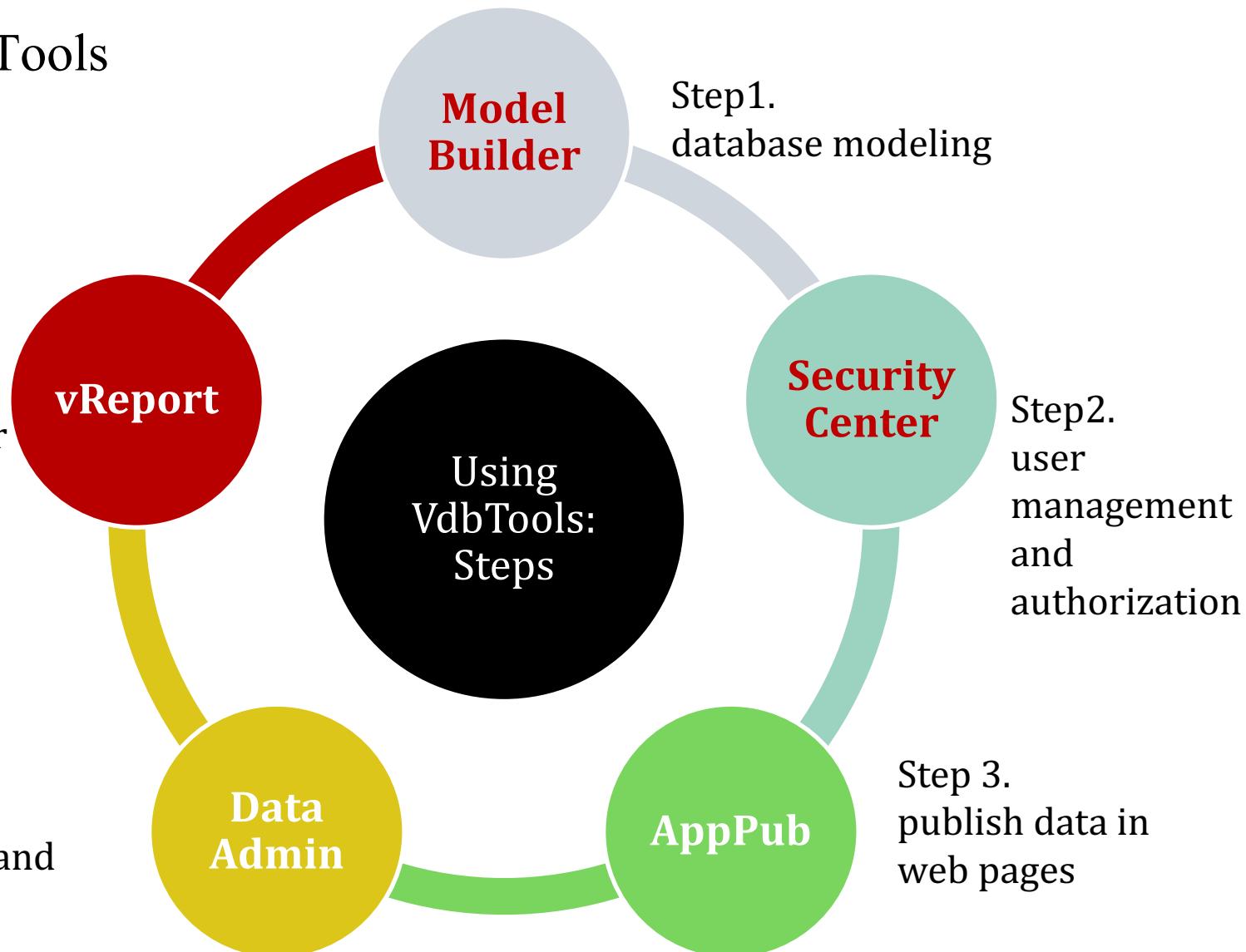
Software architecture and features



Software architecture and features

□ VdbTools

Step 5.
statistics on
data quantity
and quality,
with
graphical user
interfaces
(charts and
tables)

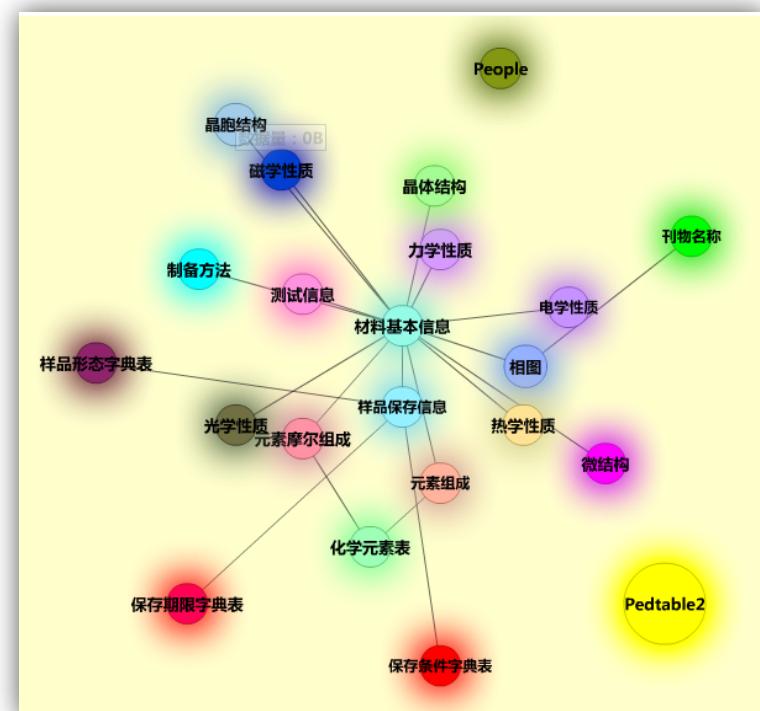


Software architecture and features

□ ModelBuilder

- LEM: A linked entities model for records and files
- Each record and file (with its metadata) is represented as an entity in LEM
- Describes links between files, records, and external services

- The right figure shows a diagram of a model in a material database, in which 22 entities (shown as circles) and 20 relationships (shown as lines) are described.

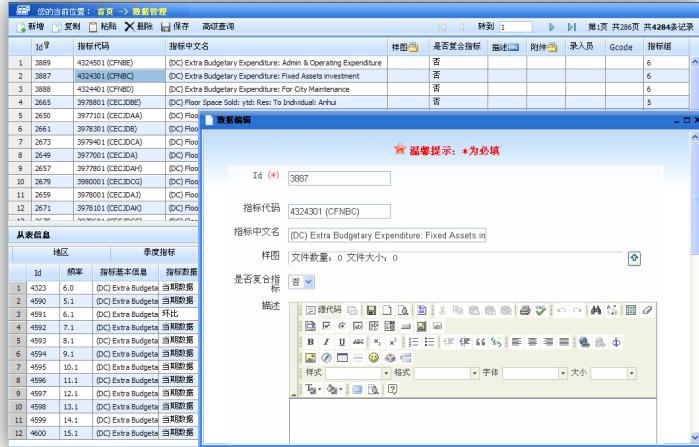


Software architecture and features

□ VdbTools - DataAdmin

■ SqlExplorer

- Microsoft Excel-like interface
- Browse, edit data
- Import, export data following by wizard
 - EXCEL
 - SQL scripts



■ FileExplorer

- upload/download files
- VisualDB will show an “Add/Update Metadata” option in the context menu when users select a file
- Classification & Tagging



Software architecture and features

□ VdbTools - AppPub

- Publishing web pages for data browse, search and view
- Cross repositories search is provided in the published application
- Customize webapp
 - 6 layout themes
 - 40+ web components
 - Configurable
 - drag and drop
 - 15 editor/view styles for different value types



Software architecture and features

□ Features

- Full java, independent of OSs
- Full support for most RDBMS including Oracle、SQLServer、MySQL、Access、PostgreSQL、HSQL、SQLite、derby
- Fine-grained access control, especially focuses on each field of each record
- Extensible component-lib (library of web components)
- Extensible typelib (library of wrappers and renderers for different types of data, such as images, videos, dates etc.)



Software architecture and features

- editors/view controls for different value types

The collage illustrates several software components:

- Date picker:** A calendar interface showing November 1999, with a specific date highlighted.
- Table editor:** A window titled "从表添加" (Add from table) displaying a list of teacher course assignments. It includes a search bar, navigation buttons, and a modal for adding new entries.
- File uploader:** A window titled "上传文件" (Upload file) showing two files: "vdb改进意见和建议.doc" and "近期工作汇报.ppt".
- Rich Editor for Text:** A text editor with a toolbar featuring icons for source code, cut, copy, paste, bold, italic, underline, superscript, and various styling options.
- Chemical Structures:** A software interface for drawing chemical structures, showing a complex polycyclic aromatic hydrocarbon.
- Map:** A Google Maps interface showing a location in Beijing, China, with labels for "中关村东" and "清华园站". It includes zoom controls and a legend.
- FASTA format of genes:** A sequence of DNA bases: acc=AB000263 | descr=Homo sapiens mRNA for prepro cortistatin | TTGTCCCCCGGCCTCCTGCTGCTGCTCCGGGCCACGGCCACCGCTGCCCTGCC | GCCCCCACCGGCCGAGACAGCGAGCATATGCAGGAAGCGGCAGGAATAAGGAAAAGCAGC | CCTCGCTTGGTGGTTGAGTGGACCTCCAGGCCAGTGCCGGCCCTCATAGGAGAGG | GTGGCCAGGCCAGGAAGGCCACCCCCCCCAGCAATCCCGCGCGCCGGGACAGAATGCC | TCTTCTGGAAGACCTCTCCTGCAAATAAAACCTCACCATGAATGCTCACGCAAG | JACCTGAA

Developing user applications based on VisualDB

Managing and Linking Scientific Data on the Web via VisualDB

Developing user applications based on VisualDB

□ Developing applications

- VdbTools, as generic tools, may not satisfy all users' requirements, so data managers need to develop their applications upon VisualDB

building
native
applications

- Developing a VdbApp based on the **VdbFramework**

developing
remote
clients

- Connecting to a **VdbServer** and consuming data remotely

Developing user applications based on VisualDB

□ VdbFramework: developing native applications

Engine APIs

- Access the running VdbEngine
- e.g. add or delete a file into/from a repository

application context injection

- override default configurations
- e.g. add a custom caching mechanism into the persistence layer

security-framework extension

- control user's access on resources such as Web pages or services

VTL

- VisualDB Template Language

ui-support

- page layout template
- JSP tags in Web pages

Developing user applications based on VisualDB

- VdbFramework: developing native applications
 - Writing web pages in VTL(VisualDB Template Language)
 - Powered by Velocity^[1]
 - Easy to learn, easy to read
 - Easy to reference java objects

For example:

hello, \$user.name! welcome to \$database.title.

→

hello, ShenZhihong! welcome to China lakes databse.

[1] The Apache Velocity Project. <http://velocity.apache.org/>

Developing user applications based on VisualDB

□ VTL(VisualDB Template Language)

- Defining a virtual field which value comes from an expression

```
person.age=  
 ${today.year-person.birthday.year}
```

- Customizing query and report templates

```
query1=  
 select * from persons where age>$ageVar
```

The value of ageVar may be passed in as a parameter in a HTTP request

Developing user applications based on VisualDB

□ Developing remote clients

- VdbServer hosts multiple pluggable services
 - Core services
 - SRV services: **Search** and **Retrieval** over **VisualDB**



- Message formats: binary/JSON/RDF
 - Hessian binary web service protocol^[1] is chosen for its high performance.

[1] <http://hessian.caucho.com/>

[2] <http://www.w3.org/DesignIssues/LinkedData.html>

Developing user applications based on VisualDB

- Developing remote clients
 - Especially, VDB-WOD services are provided
 - implementation of Linked Data^[2] interface
 - each record or file has an URI
 - responses in RDF/XML on request

[1] <http://hessian.caucho.com/>

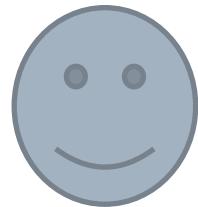
[2] <http://www.w3.org/DesignIssues/LinkedData.html>

Developing user applications based on VisualDB

□ Linked Data interface



For each entity



When user access the URI

Such an URI is assigned
`http://qinghailake.csdb.cn/birds/2`



HTML representation

The screenshot shows a Windows Internet Explorer window displaying the RDF representation of a bird entity. The page lists various properties and their values, such as '物种' (Species) set to 'true', '国家一级保护动物' (National First-class Protected Animal), '青海湖主要觅食鸟' (Main feeding birds in Qinghai Lake), and 'IUCN世界自然保护联盟' (IUCN World Conservation Union). The page also includes a table of '青海湖鸟类调查表' (Qinghai Lake Bird Survey Form) with columns for '物种' (Species), '分布状况' (Distribution Status), and '分布范围' (Distribution Range).

Application cases

Managing and Linking Scientific Data on the Web via VisualDB

Application cases

- VisualDB has deployed in more than 30 institutes, and helped data managers establish their data management and sharing systems^[1] without programming
 - China soil database
 - China animal database
 - China lake database
 - China plant database
 - China coastal environment remote sensing database
 - Plant introduction and conservation database
 - More cases are listed in <http://vdb.csdb.cn/cases>



[1] <http://www.vdb.csdb.cn/cases>

Application cases

- more than 5 data management systems are developed based on VdbFramework
 - China northwest cultural database
 - China Brainbank database
 - Qinghai project management system
 - Beijing food safety standards database
 - Research network for applied microbes(**RNAM-DB**)
 - <http://www.rnam.org/>
 - Managing data such as strain information
 - Organization: headquarter → centers → labs
 - Users in different level are granted different privileges

Application cases

□ Authorization in RNAM-DB

headquarter

centers

labs

The screenshot shows the RNAM-DB Security Center interface. On the left is a tree view of research centers and laboratories. Three lines from the text labels 'headquarter', 'centers', and 'labs' point to the 'Industrial Microbiology Research Center (North)', 'Industrial Microbiology Research Center (South)', and 'Environmental Microbiology Research Center (Chengdu)' respectively. The main panel displays a 'User List' table with 6 entries, all assigned to the 'Food Microbiology Network Laboratory (North)'. Below it is a 'Grant Privileges to Current User' section for the user 'hexiuping', which includes a table of roles and checkboxes for system management, data management, and Web Service users. A large red watermark 'grant privileges to users' is overlaid across the bottom right of the interface.

序号	用户名	用户组
1	hexiuping	食品微生物网络实验室(北方)
2	sunyuanxia	食品微生物网络实验室(北方)
3	liyin	食品微生物网络实验室(北方)
4	guoxuena	食品微生物网络实验室(北方)
5	zhutaicheng	食品微生物网络实验室(北方)
6	wangzhaoyue	食品微生物网络实验室(北方)

综合信息	页面
前台列表页面	个人数据
前台细览页面	默认前台列表页面 序列信息列表 化学分子信息
编辑列表页面	个人数据 实验室组 中心组
添加记录页面	默认添加记录页面
修改记录页面	默认修改记录页面

培养基信息	页面
前台列表页面	个人数据
前台细览页面	默认前台细览页面
编辑列表页面	个人数据
添加记录页面	默认添加记录页面
修改记录页面	默认修改记录页面

Application cases

- Furthermore, three integrated databases in SDB project are (being) developed via linking VdbServers

CAS chemical subject database

Properties of the record come from different VdbServers Beijing, Shanghai, Changchun

CAS material subject database

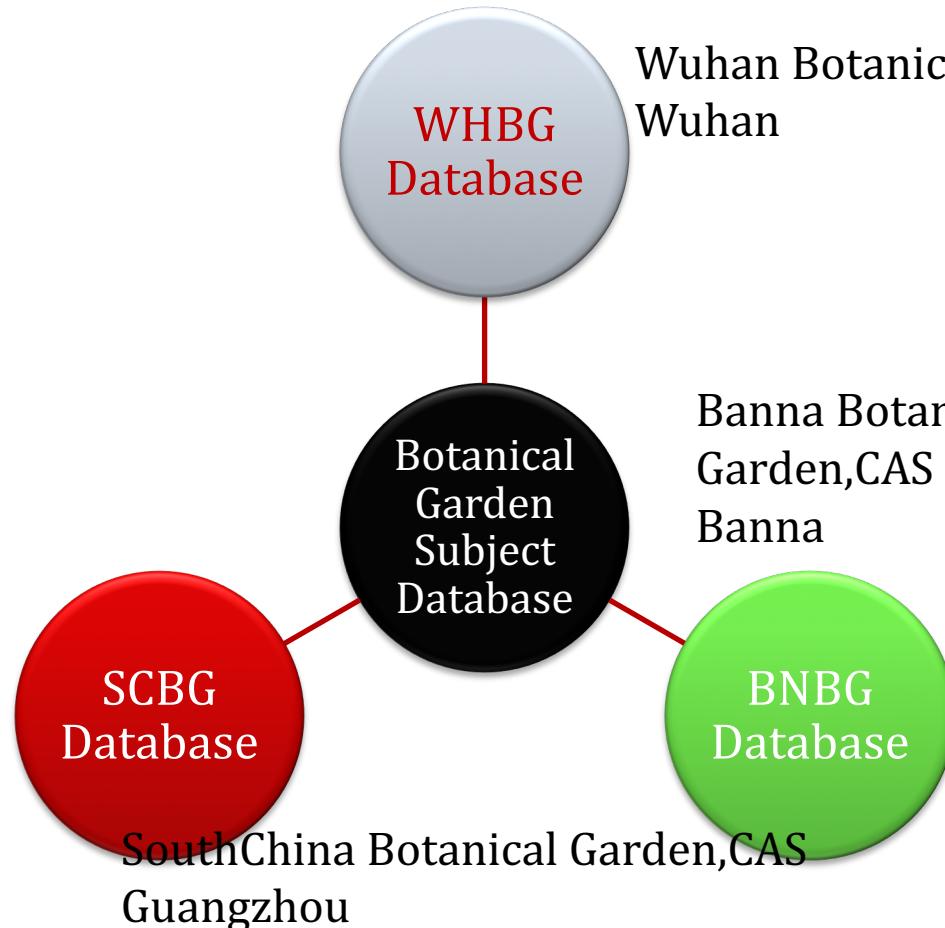
The entities come from different VdbServers
Shanghai, Shenyang

金属材料

- + 基础信息
- + 加工测试信息
- + 其它信息
- + 应用信息
 - 无机非金属材料
 - 无机非金属材料数据库
 - 无机非金属材料
 - 测试信息
 - 晶体结构
 - 晶胞结构
 - 光学性质
 - 热学性质
 - 力学性质
 - 电学性质
 - 磁学性质
 - 相图

Application cases

- The 3rd integrated database
 - CAS botanical garden subject database
 - integrates 3 botanical garden databases



The screenshot shows the homepage of the "植物园主题数据库" (Plant Theme Database) from the Wuhan Botanical Garden, CAS. The page features a banner image of a garden, navigation links like "首页", "资源数据库", "高级搜索", etc., and a search bar. Below the search bar is a table titled "检索结果" (Search Results) showing 15 entries for the Camellia family (山茶科).

中文学名	英文名	引种	属	种	属	种	属	种	属	种	属	种
1 文毛冬青	Brynnearia	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
2 山茶科	Bympelaceae	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
3 文毛冬青	Agavaceae	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
4 文毛冬青	Brynnearia	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
5 山茶科	Bympelaceae	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
6 文毛冬青	Agavaceae	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
7 文毛冬青	Brynnearia	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
8 山茶科	Bympelaceae	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
9 文毛冬青	Agavaceae	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
10 文毛冬青	Brynnearia	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
11 山茶科	Bympelaceae	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
12 文毛冬青	Agavaceae	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
13 文毛冬青	Brynnearia	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
14 山茶科	Bympelaceae	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200
15 文毛冬青	Agavaceae	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200

Thank you for your attention!

<http://vdb.csdb.cn>

mailto: bluejoe@cnic.cn



The screenshot shows the homepage of the VisualDB website. At the top, there's a navigation bar with links for '产品介绍' (Product Introduction), '服务案例' (Service Case Studies), '问题答疑' (FAQ), '支持下载' (Support Download), and '在线试用' (Online Trial). Below the navigation, a large banner features the text 'VisualDB免费下载试用，助力数据应用，同创企业未来。[赶快下载体验吧]' (VisualDB free download trial, assist data application, co-create enterprise future. [Download and experience now]). A prominent '立即下载' (Download Now) button is located below the banner. To the right, there's a section titled 'VisualDB 简介' (Introduction) with a brief description of the product and a '查看详情' (View Details) link. On the far right, there's a sidebar titled '产品动态' (Product News) listing recent news items with dates like '2010-8-23'.

版权所有：中国科学院