



# Public Health Scientific Data Sharing Service

**Han Hong**  
**2006.10.24**



# Outline

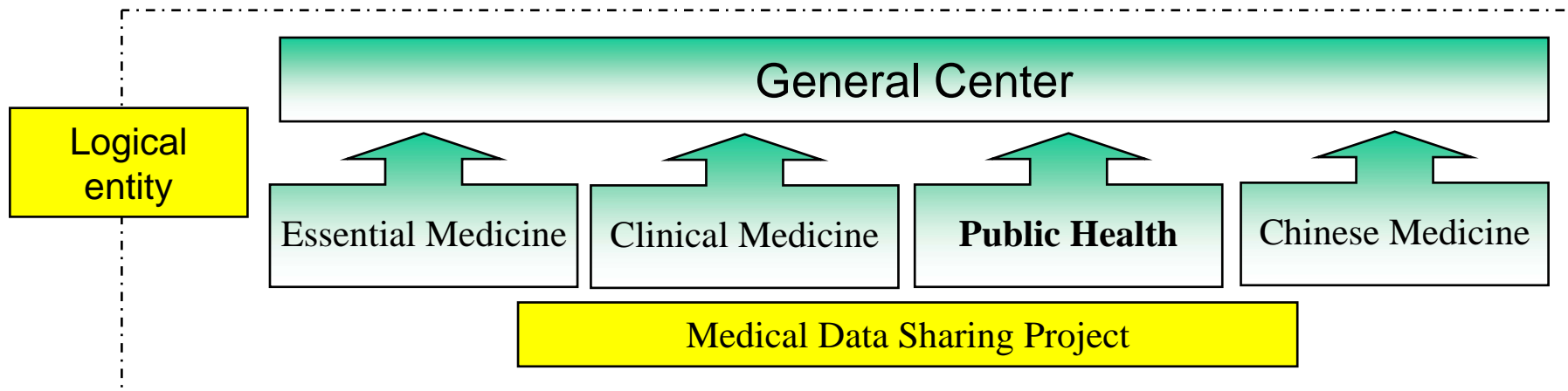
- Introduction
- Data Submitting & Processing
- Data Service
- Chance & challenge



# Outline

- Introduction
  - General Introduction
  - Characteristics of Public Health Data & Users
  - General Sharing Process
- Data Submitting & Processing
- Data Service
- Chance & challenge

# General Introduction



- General center & 4 data centers
- Distributed in physics but unified in logic
- Share common entry & metadata service
- Share common standards
- Different emphasis

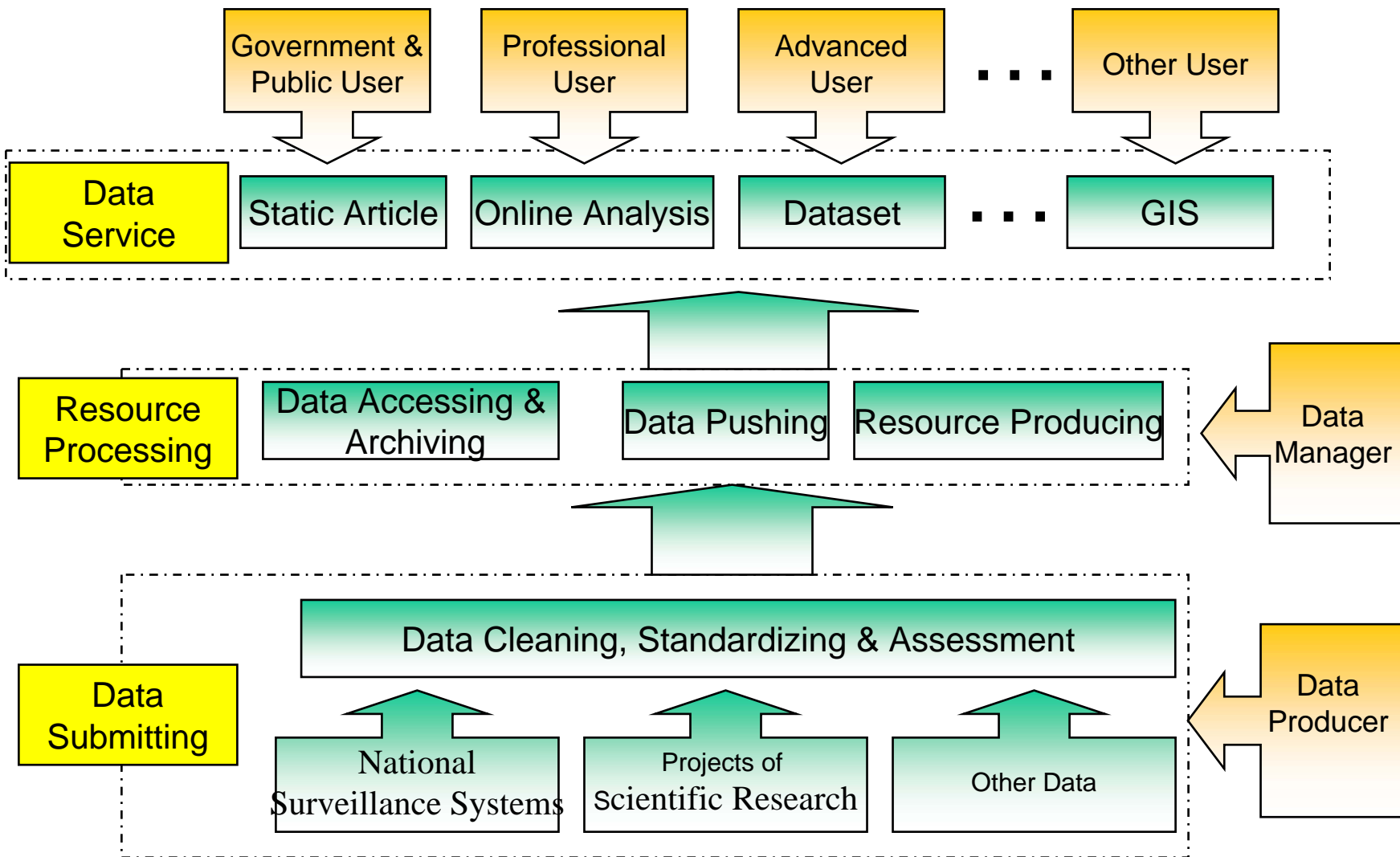
# Characteristics of Public Health Data & Users



- Data -> the importance of quality control
  - Source: Surveillance & Report
  - Content: Epidemiology
  - Data capacity : Big or great, often millions
  - Type: Enumeration data
- Users -> the modes of data sharing
  - Government & public
  - Professional
  - Advanced user



# General Sharing Process





# Outline

- Introduction
- Data Submitting & Processing
  - Resource Investigation
  - Data Cleaning & Standardizing
  - Resource Processing
  - Quality Control
- Data Service
- Chance & challenge

# Resource Investigation



- Methods
  - Survey
  - Web

表 1 公共卫生信息资源调查表

数据库名称	
数据库状态	1. <input type="checkbox"/> 已经建成 2. <input type="checkbox"/> 正在建中
所属资助项目名称	
项目类型	1. <input type="checkbox"/> 科技部科技基础性工作转向 2. <input type="checkbox"/> 社会公益研究专项 3. <input type="checkbox"/> 科技攻关 4. <input type="checkbox"/> 863 5. <input type="checkbox"/> 973 6. <input type="checkbox"/> 其它科技部项目 7. <input type="checkbox"/> 国家自然科学基金资助项目 8. <input type="checkbox"/> 中科院项目 9. <input type="checkbox"/> 教育部项目 10. <input type="checkbox"/> 其它部委项目 11. <input type="checkbox"/> 本单位自筹资金项目 12. <input type="checkbox"/> 其它 (请注明)
所属学科领域	
项目承担单位	
项目负责人	
建库单位	
建库/完成时间	建库: 年 月 完成: 年 月
数据更新	1
数据类型	1 6
数据来源	1 6 您的位置: 首页 - > 资源调查 - > 公共卫生信息资源现场在线调查表
使用限制	1
收费方式	1
服务方式	1
网址	1 6
主要用户	1 6
数据库操作系统	1 5
数据库类型	1 6
数据管理系统	1 6 9
数据量	1 6
记录数	1 6
数据年限	1 6



国家医药卫生科学数据共享网

中国公共卫生科学数据中心

The Data-Center of China Public Health Science

数据库信息(\*为必填项)

数据库名称: \*

数据库状态: 请选择 \*

所属资助项目名称: \*

项目类型: 请选择 其他(请注明) \*

所属学科领域:  营养学  毒理学  消毒学  流行病学  传染病学  媒介生物控制学  环境医学  职业病学  地方病学  社会医学  卫生检验学  食品卫生学  儿少卫生学  妇幼卫生学  环境卫生学  劳动卫生学  放射卫生学  卫生工程学  卫生经济学  卫生统计学  计划生育学  优生学  健康教育学  卫生管理学  预防医学与卫生其他学科 \*

项目承担单位: \*

项目负责人: \*

建库单位: \*

建库/完成时间: 建库时间: 完成时间: \*

数据更新: 请选择 \*

更新周期: 请选择 其他(请注明) \*

数据来源:  数值  文本  图形  图像  声音  影像  其它(请注明)  检测  检验  其它(请注明)  其它(请注明)  统计  研究  监测

使用限制: 请选择 \*

收费方式: 请选择 \*

服务方式: 请选择 \*

网址: \*

主要用户: 请选择 其他(请注明) \*

数据库操作系统: 请选择 其他(请注明) \*

数据库类型: 请选择 其他(请注明) \*

数据管理系统: 请选择 其他(请注明) \*

数据量: (GB) \*

数据年限: (GB) /年 \*

记录数: (条) \*

数据年限: 年~ 年

数据库内容简要描述: \*

应用范围: \*

联系人姓名: \*

联系电话: \*

传真: \*

联系单位: \*

联系地址: \*

邮编: \*

提交 重置



# Data Cleaning & Standardizing



- Data Cleaning Regulation
  - Familiar mistakes: general & logical
  - Cleaning methods
  - Reports writing
- Data Standardizing Regulation
  - Applicable Standards: coding standards
  - Standardizing methods
  - Reports writing

# Resource Processing



- Metadata Class
  - Data background
  - Metadata
  - Data element
- Sharing Data Class
  - Static articles
  - Query analysis
  - OLAP analysis
  - Downloadable dataset
  - GIS analysis

# Quality Control



- By Means of
  - Proper regulations
  - Quality assessment
- Throughout
  - Data Submitting
  - Resource Processing



# Outline

- Introduction
- Data Submitting & Processing
- Data Service
  - Supporting Web Platform
  - Shared Data Resource
  - Service Mode
  - Authority Management
- Chance & challenge

# Supporting Web Platform



- Sharing platform
  - Data finding
  - Data obtaining
- Working platform
  - Data manufacturing
  - System management

国家医药卫生科学数据共享网——  
中国公共卫生科学数据中心  
The Data-Center of  
China Public Health Science

共享政策 | 数据更新动态 | 工作进展 | 数据共享产品

国家医药卫生科学数据共享网——  
中国公共卫生科学数据中心  
The Data-Center of  
China Public Health Science

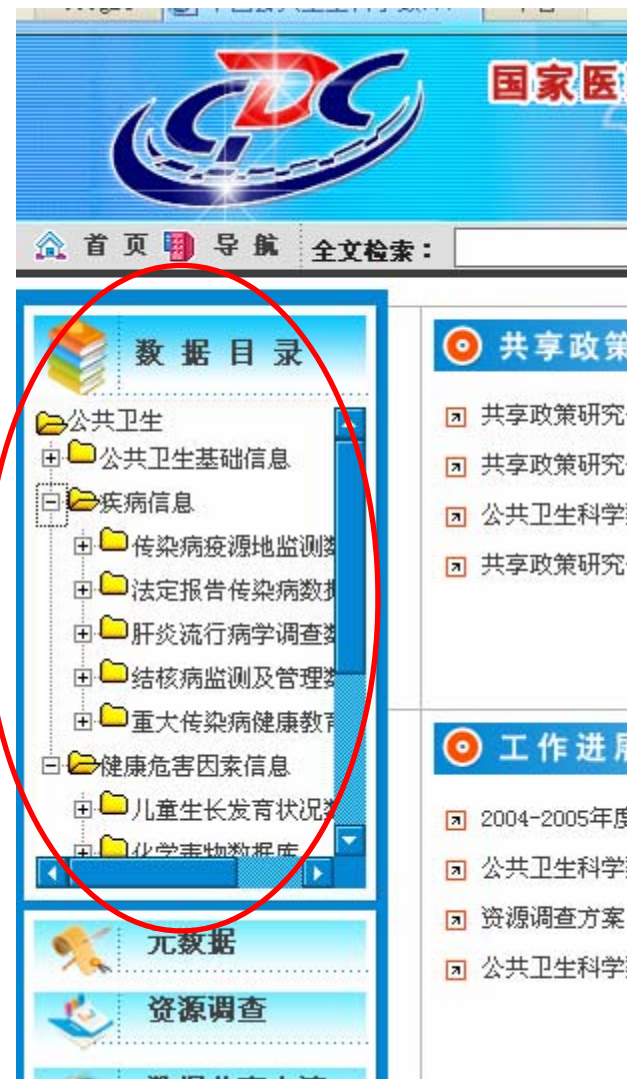
静态资源

标题	上传用户	上传日期	发布用户	发布日期	标志	调度标志	使用角色	分配角色	撤发	编辑	删除
全国疾病监测系统人群缺血性心脏病城乡分年龄别死亡率(2000)	admin	2006-09-14	admin	2006-08-30	发布	已调度	注册用户				
全国疾病监测系统不同类别农村人群结核死亡率变化趋势(1991-2000)	admin	2006-09-14	admin	2006-08-30	发布	已调度	注册用户				
全国疾病监测系统结核分枝杆菌分年龄别死亡率(2000)	admin	2006-09-14	admin	2006-08-30	发布	已调度	注册用户				
全国疾病监测系统人群肝癌城乡分年龄别死亡率(2000)	admin	2006-08-16 15:43:08:696	admin	2006-08-30	发布	已调度	注册用户				
中国女性产科疾病估计死亡率变化趋势(1950-2000)	admin	2006-08-16 15:43:08:55	admin	2006-08-30	发布	已调度	注册用户				
全国疾病监测系统人群肺炎城乡分年龄别死亡率(2000)	admin	2006-08-16 15:43:08:526	admin	2006-08-30	发布	已调度	注册用户				
中国围生期疾病估计死亡率变化趋势(1950-2000)	admin	2006-08-16 15:43:08:345	admin	2006-08-30	发布	已调度	注册用户				
中国人群婴儿死亡率变化趋势(1950-2000)	admin	2006-08-16 15:43:08:205	admin	2006-08-30	发布	已调度	注册用户				
中国城市和农村人群糖尿病死亡率变化趋势(1974-2000)	admin	2006-08-16 15:43:07:695	admin	2006-08-30	发布	已调度	注册用户				
中国城市和农村人群死亡率变	admin	2006-08-16	admin	2006-08-30	发布	已调度	注册用户				

# Shared Data Resource



- 13 Databases (2T)
  - Basic information
    - Death Surveillance
  - Diseases information
    - Infective diseases legal report
  - Health hazardous factor information
    - Smoking survey



# Service Mode



- Data finding service
  - Metadata
  - Full text search



# Service Mode



- Data obtaining service
  - Online & off-line service
  - Free of charge
  - Registry is needed
    - For government & public role: static articles
    - For professional role: query & OLAP analysis
    - For advanced role: downloadable dataset



# Service Mode



- static articles
  - Tree view
  - PDF

The screenshot shows the 'China CDC' website interface. On the left, a '数据目录' (Data Directory) tree view is expanded to show '2004年10月全国法定报告甲、乙类传染病发病、死亡情况'. The main content area displays a list of static articles under the heading '静态查询'. The selected article is a PDF document titled '2004年10月全国法定报告甲、乙类传染病发病、死亡情况'. The PDF content includes a summary of 2004 national infectious disease statistics, a list of top 5 diseases by incidence and mortality, and a detailed table of 2004 national reported infectious disease statistics.

**表 2004年10月全国法定甲、乙类报告传染病发病、死亡统计表**

病名	发病数	死亡数
合计	272151	522
鼠疫	19	8
霍乱	37	0
病毒性肝炎	99292	84
甲肝	7451	9
乙肝	80007	59

# Service Mode



- Query analysis

The screenshot displays the website interface for the China Public Health Science Data Center. The top banner includes the logo and the text "国家医药卫生科学数据共享网" and "中国公共卫生科学数据中心". The main content area is divided into a left sidebar with a "数据目录" (Data Directory) and a main "动态查询" (Dynamic Query) section. The "动态查询" section contains a table of query results.

标题	发布用户	发布日期
有关性病/HIV的知识、态度和行为情况查询	admin	2006-07-10
<b>高血压、高血脂意识情况查询</b>	admin	2006-07-10
血糖检查和高血糖及糖尿病的检测情况查询	admin	2006-07-10
肿瘤筛查情况查询	admin	2006-07-10
吸烟流行水平	admin	2006-07-11
医疗保险情况查询	admin	2006-07-10
健康体检情况查询	admin	2006-07-10
被动吸烟情况查询	admin	2006-07-10
烟草控制活动的执行和效果查询	admin	2006-07-10
饮酒情况查询	admin	2006-07-10
体力活动情况查询	admin	2006-07-10
体重评价和体重控制情况查询	admin	2006-07-10
对肝炎的知识、态度和行为情况查询	admin	2006-07-10

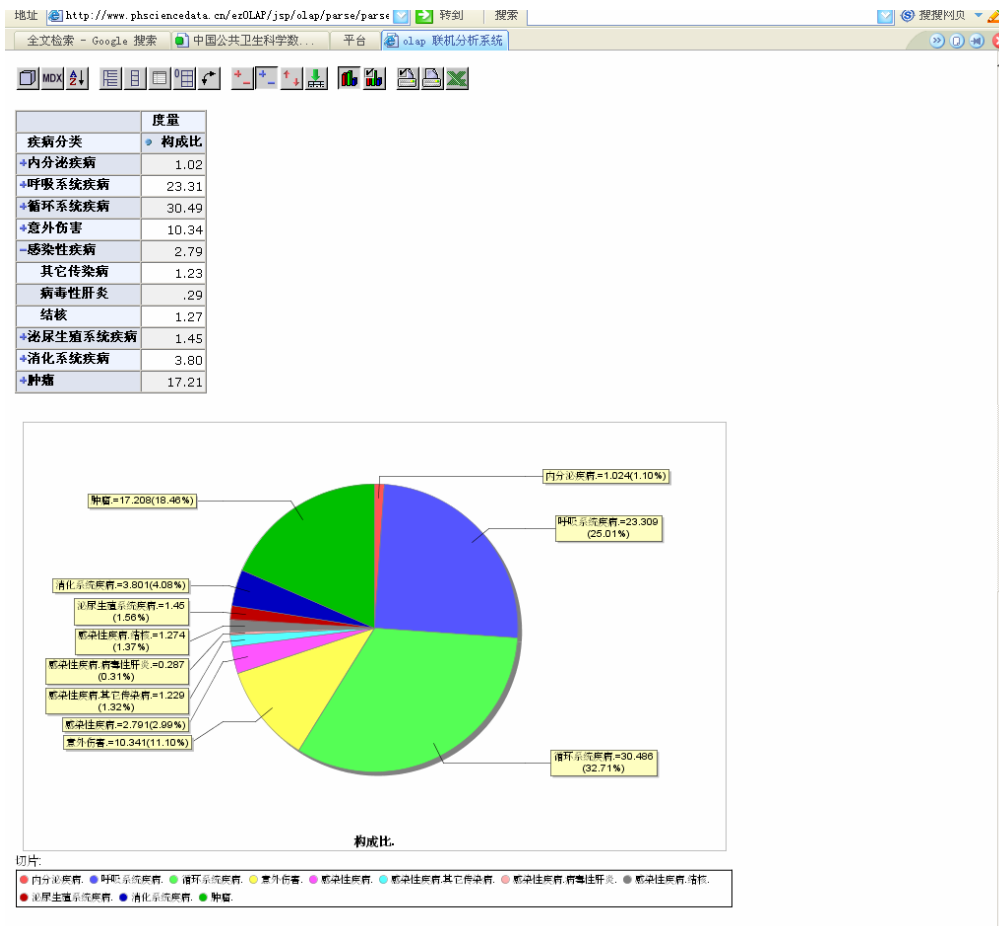
The bottom part of the screenshot shows a detailed query for "高血压、高血脂意识情况查询". It includes a form with filters for "地域" (Region: 东部), "城市" (City), "性别" (Gender: 未选择), "年龄段" (Age Group: 未选择), and "文化程度" (Education Level: 未选择). Below the form is a table of results:

对高血压危害认识正确的比例 (%)	25.21
对高血脂危害的认识正确的比例 (%)	7.41
需定期测血压的认识率 (%)	74.67
需定期测血脂的认识率 (%)	73.6
35岁及以上定期测血压的认识率 (%)	42.43
40岁及以上定期测血脂的认识率 (%)	35.47

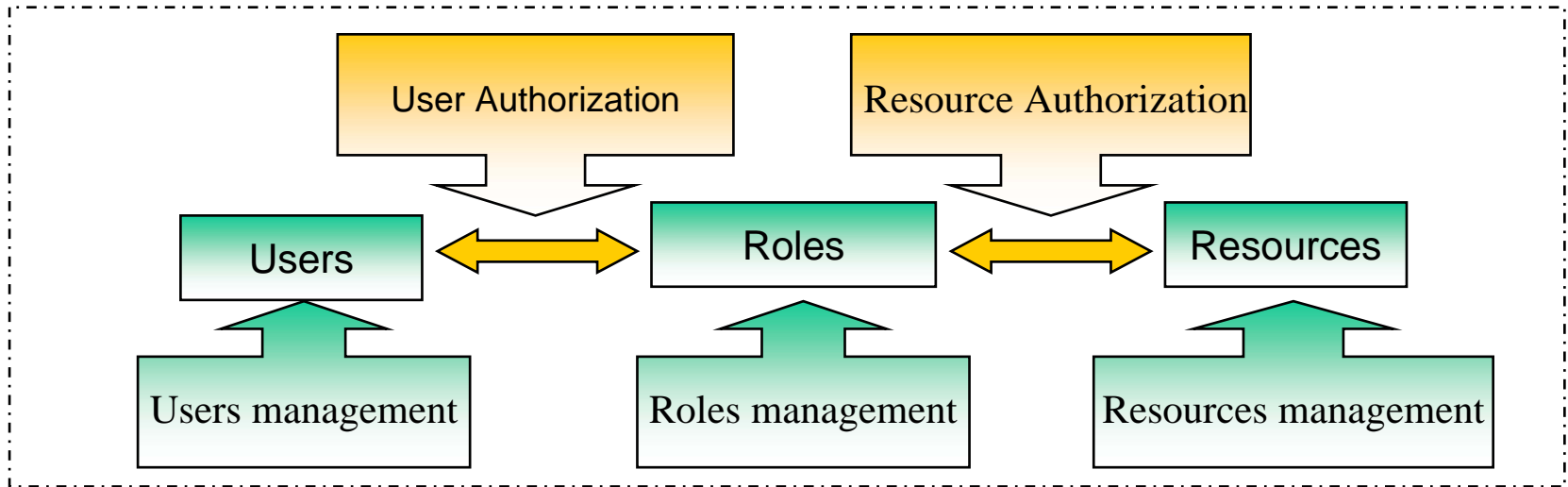
# Service Mode



- OLAP analysis



# Authority Management



- Mode
  - User-Role-Resource
- Users, role & resources management
  - Add, Modify, &Delete
- Authorization
  - Resource: to specify a resource with a role
  - User: to specify a user with a role



# Outline

- Introduction
- Data Submitting & Processing
- Data Service
- Chance & challenge
  - Chance
  - challenge

# Chance



- Chance
  - Plenty of Data: new data & historical data
  - Attach more importance: the Government
- Challenge
  - Quality control: historical data
  - People's awareness to share



# Thanks

Chinese Center for Disease Control  
and Prevention

<http://www.phsciencedata.cn>

Han\_cdc@126.com