

## 成功案例



## 利用 Informatica 建立的数据仓库可提供美国水质的重要信息

“之所以选择 Informatica 作为我们的数据集成解决方案，是因为它成熟、易用、功能齐全，并拥有业内领先的集成体系架构。”

— USGS 数据仓库实施负责人 Harry House

### 客户

美国地质调查局 (USGS)

### 挑战

集成四个主要遗留数据源系统中的数据以合并水质数据并提升决策能力

### 解决方案

- Informatica 数据集成平台

### 收益

- 节约成本，见效快
- 增强决策能力
- 提高性能
- 简单易用，实施迅速

### 要点

- 数据集成：Informatica
- 来源：文本文件、Oracle、SQL Server、IMS
- 目标：Oracle
- 平台：NT

1879 年，美国地质调查局 (USGS) 依据议会法案正式成立，在其后 120 年的发展过程中，其能力和知识也随着科学技术的进步而不断拓展。USGS 的年预算达 10 多亿美元，它运用广博的资源及专业知识与 2,000 多各州级、地方级和部落政府机构、学术团体、其它联邦联盟、非政府组织和私营机构保持着合作关系。

### 挑战

1991 年，美国地质调查局 (USGS) 设立了国家水质评估计划 (NAWQA)，旨在采集并分析全国主要地下水和地表水系统的水质数据。USGS 评估水质情况、造成这些情况的原因以及趋势，然后水资源管理者和科学家据此制定科学可行的应对方案。

要为所收集水质数据建立无比庞大的数据集，USGS 团队面临的主要难题就是数据集成。USGS 有四个主要的遗留数据源系统，由于每个系统的数据格式和参照系统都是唯一的，因此需要用四种不同的方法来操作和访问数据。例如，关于土地利用特征、人口及化学品使用的信息与源于 12 个支局的水质数据便存储在不同的系统中。而且像大多数自然现象一样，水质具有许多不同的指标来帮助确定其状态和未来趋势，为了进行有效评估，科学家需要综合而详细地了解所有相关信息。

以前，研究人员使用多种集成方法来分析水质，他们将数据子集导入 Microsoft Access 或 Microsoft Excel 中。但是在这样的环境中操作，通常要耗费一天甚至更长时间来编译待分析数据，而当系统移动数百万条记录时，还可能会因不堪重负而宕机。

## 解决方案

为了建立整合、集成的水质数据仓库，USGS 数据仓库团队最终选择了 Informatica 数据集成平台，因为该平台具备强大的功能、集成的体系架构和良好的易用性。USGS 目前拥有多个国家级数据仓库系统，其中包括 NAWQA 计划和几个针对地区合作伙伴的解决方案。

科学家能够通过 NAWQA 量化水质问题并解释可能的原因，同时为水质管理者和科学家作出决策提供可靠合理的机制。实际上，NAWQA 包含关于六百多万条水体和生物的调查结果，它变革了 USGS 研究人员发现并验证水质趋势和热点的方式，使研究人员可以检验以前无法检验的假说。

为了继续有效地分析水质并把握趋势，NAWQA 数据仓库需要定期从文本文件、Oracle 数据库、Microsoft Access 系统和 SQL Server 事务处理系统等多个数据源更新资料。数据要从 12 个不同的地方、地区及国家级机构集成到一个 Oracle 集结区。在该集结区，将通过多个并发源并大量使用键查找来转换数据，然后集中数据并重新格式化为非规范星型架构，供最终用户访问。

## 成果

### 节约成本，见效快

USGS 可以在 Informatica Designer GUI 界面中直观地定义映射和转换，而不用编写自定义代码，从而削减了开发成本。在执行项目的第一年，USGS 就节省了至少 30 天的开发时间和大量资源。现在，通过使用 Informatica 数据集成平台，每位开发人员的工作效率都得到最大限度提高。

### 加快决策，提高决策能力

通过水质数据的综合视图，USGS 科学家能够从 7 千万条化学、生物和物理水质数据中进行选择，而这些数据是由过去 10 年从约 50 片水域采集而得。例如，如今通过综合视图，科学家可以对比分析不同地区在特定污染物方面的差异。

### 提高了系统性能

与定制解决方案相比，通过使用 Informatica 的内存中服务器侧缓存处理来有效利用系统资源，USGS 已经能够将系统的复杂数据加载性能提升 30% 多。

### 简单易用，快速实施

Informatica 数据集成平台的易用性得到充分展现，USGS 只需最简单的培训就能开始开发映射。实际上，在授权 Informatica 数据集成平台的一周内，USGS 就进入了正常运作。



[www.informatica.com.cn](http://www.informatica.com.cn)

#### 北京办事处

地址：北京市朝阳区建国门外大街乙 12 号  
LG 双子座大厦东塔 19 层 1906 室  
邮编：100022  
电话：86-10-5879 3366  
传真：86-10-5879 3130

#### 上海办事处

地址：上海市浦东世纪大道 201 号  
渣打银行大厦 5 楼  
邮编：200120  
电话：86-21-6182 6825  
传真：86-21-6182 6755

#### 广州办事处

地址：广州市天河区体育东路 138 号  
金利来数码网络大厦 3 层 321 室  
邮编：510620  
电话：86-20-2886 0646  
传真：86-20-3878 1801