

DreamBI 大数据分析平台
技术白皮书

目录

第一章 产品简介.....	4
一、 产品说明.....	4
二、 产品特点.....	4
三、 系统架构.....	4
四、 基础架构.....	7
五、 平台架构.....	7
第二章 功能介绍.....	7
2.1. 元数据管理平台.....	7
2.1.1. 业务元数据管理.....	8
2.1.2. 指标元数据管理.....	10
2.1.3. 技术元数据管理.....	14
2.1.4. 血统管理.....	15
2.1.5. 分析与扩展应用.....	16
2.2. 信息报送平台.....	17
2.2.1. 填报制度管理.....	17
2.2.2. 填报业务管理.....	33
2.3. 数据交换平台.....	54
2.3.1. ETL 概述.....	55
2.3.2. 数据抽取.....	56
2.3.3. 数据转换.....	56
2.3.4. 数据装载.....	57
2.3.5. 规则维护.....	58
2.3.6. 数据梳理和加载.....	65
2.4. 统计分析平台.....	67
2.4.1. 多维在线分析.....	67
2.4.2. 即席查询.....	68
2.4.3. 智能报表.....	70
2.4.4. 驾驶舱.....	74

2.4.5.	图表分析与监测预警.....	75
2.4.6.	决策分析.....	79
2.5.	智能搜索平台.....	83
2.5.1.	实现方式.....	84
2.5.2.	SolrCloud.....	85
2.6.	应用支撑平台.....	87
2.6.1.	用户及权限管理.....	87
2.6.2.	统一工作门户.....	94
2.6.3.	统一消息管理.....	100
2.6.4.	统一日志管理.....	103
第三章	典型用户.....	106
第四章	案例介绍.....	108
一、	高速公路大数据与公路货运统计.....	108
二、	工信部-数据决策支撑系统.....	110
三、	企业诚信指数分析.....	111
四、	风险定价分析平台.....	112
五、	基于斯诺模型的增长率测算.....	113
六、	上交所-历史数据回放引擎.....	114
七、	浦东新区能耗监控.....	115

第一章 产品简介

一、 产品说明

DreamBI 摒弃了原有数据采集及开发应用系统建设时间长、花费高的传统建设模式，提供了一套方便易用的元数据管理平台，并以此为基础，采用独有的拖拽式可视化操作，为用户提供了一整套简单快捷的数据采集、查询统计、在线分析、资料制作、信息发布等功能模块。

二、 产品特点

让每个人都会用的可视化展现及分析系统

- ◆ 拖拽即成，人人会用
- ◆ 跨平台、跨设备轻松展现
- ◆ 全程元数据管理
- ◆ 完美中国式报表，符合国情
- ◆ OLAP 分析、即席查询、交互式图表分析、数据挖掘四项全能

三、 系统架构

总体架构设计采用云服务的架构模式系统建设总体架构图如下：



如上图所示，项目功能整体设计以服务为理念，自下而上分别提供：

一、IaaS（设备即服务）层

这一层，利用服务器、存储设备、网络设备等硬件设施，同时配合操作系统、虚拟化管理软件、数据库、网络管理系统为整个项目的部署、运行提供高效可控的硬件设施服务。

二、Daas（数据即服务）层

数据层，作为整个系统的信息存储层，这里主要存储信息包括：元数据、企业名录库、共享数据、私有数据以及应用信息库（用户权限信息、系统配置信息、日志信息等）。

三、PaaS（平台即服务）层

作为开放、可扩展、成熟的应用系统，大部分的应用模块都应该都是通过完善可靠的应用开发平台，结合实际的业务需要，通过二次开发或配置完成，配合本期项目的实际需要以及未来可扩展性的考虑，我们将提供元数据管理平台、智能搜索平台、信息交换平台、信息报送平台、统计分析平台、应用支撑平台这六大基础服务平台，大部分的功能模块将通过这六大平台二次开发或配置完成，同时保证业务模块的快速调整或功能扩展。

四、SaaS（软件即服务）层

软件既服务这一层，将根据用户实际的业务需求，结合基础服务平台，提供具体的功能模块，包括：

(一)、信息报送子系统

- 任务发布
- 信息报送
- 数据审核
- 信息归档
- 任务监控
- 报送情况统计

(二)、数据处理子系统

- 企业名录库
- 数据加载
- 查询、统计、计算、报表设计、图表设计
- 信息发布
- 批处理脚本

(三)、信息共享门户

- 信息发布审核
- 共享目录
- 智能检索
- 统计出版物
- 专题分析
- 分析报告
- 数据解读

(四)、微信服务号及移动应用

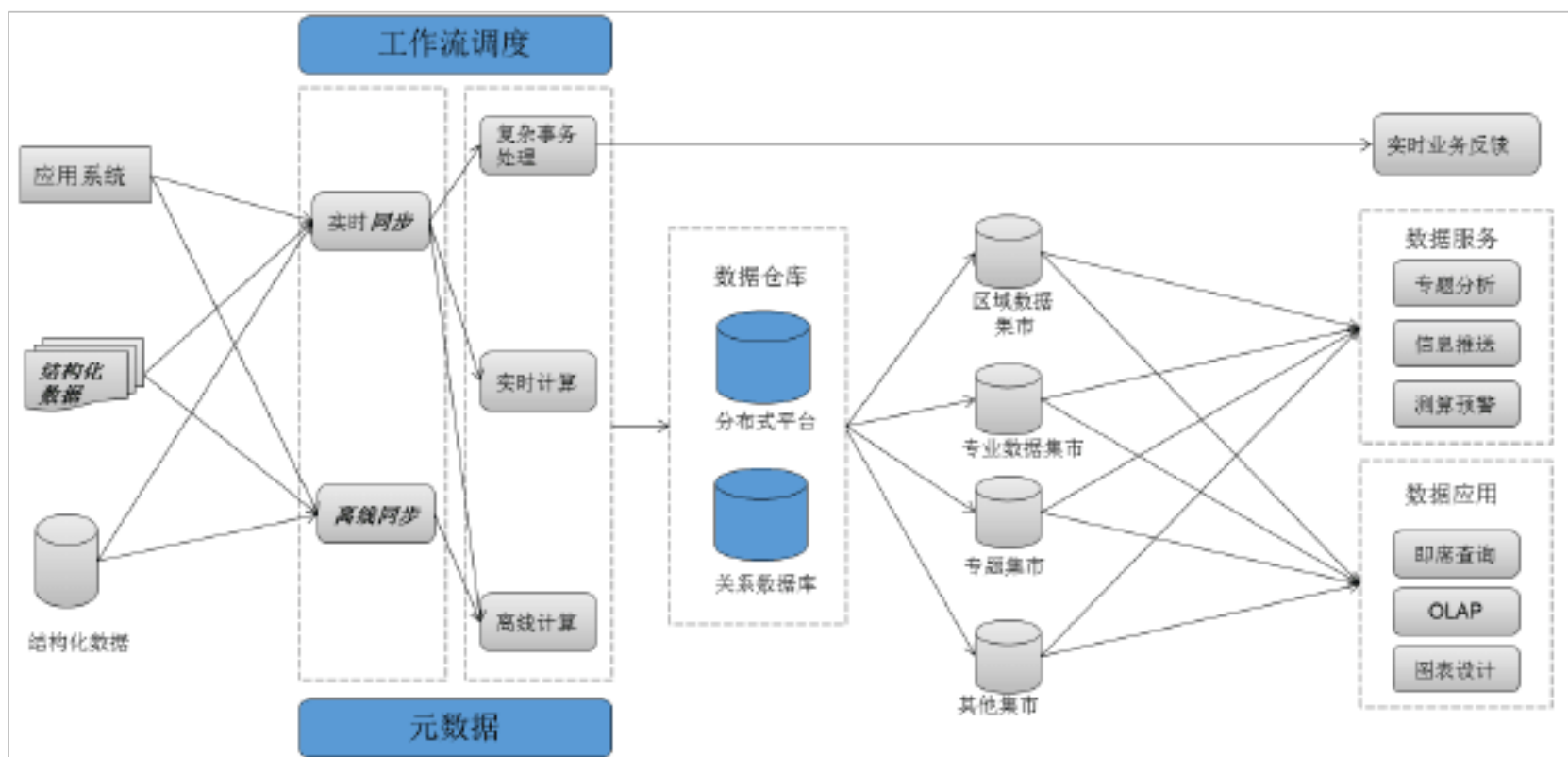
- 用户绑定
- 智能应答
- 信息推送
- 报送审核
- 辅助功能

- 管理功能

四、基础架构



五、平台架构



第二章功能介绍

2.1. 元数据管理平台

数据共享平台，作为一个以资源信息库为核心的建设项目，数据标准的制定

和管理是这个项目是否成功的主要因素，所以，元数据管理平台是支撑数据共享平台的主要组成部分。通过该平台来规范管理全系统数据资源的规范定义、命名、分类等。同时，元数据管理平台也将帮助我们技术的角度梳理所有的信息系统，了解所有数据资源的生产、存储、转换和同步等所有相关活动，从而建立数据元的血统世系，帮助我们理解每一个数据的来龙去脉。元数据管理系统将提供如下一些主要的功能：

2.1.1. 业务元数据管理

根据全系统信息资源规划的结果和行业信息资源标准体系,从业务的视野出发，建立的关键数据元、指标等的分类、规范定义体系。具体功能如下：

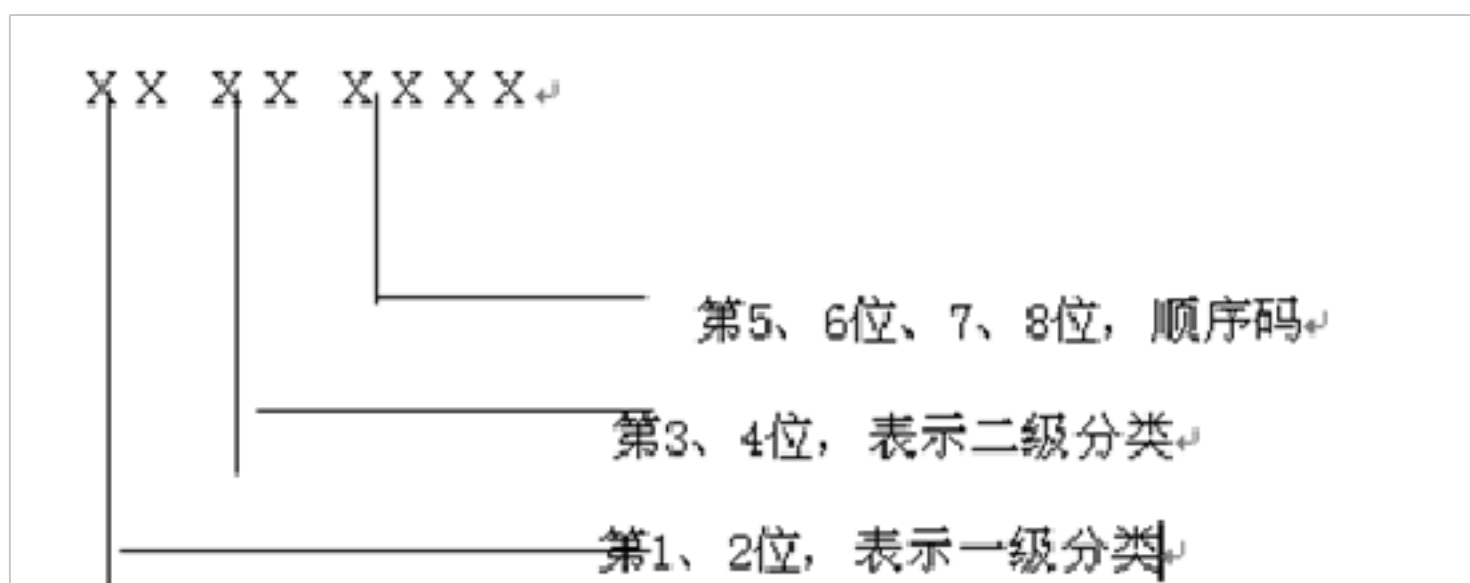
2.1.1.1. 元模型定义

元模型定义用于定义元模型结构，一般包括编号、名称、说明、同义词、语境、标注字段命名、值域、备注等要素。为建立与行业标准的对照，建议元模型应包含与行业标准的对照项。一个推荐的业务数据元元模型如下（最后一栏为国标行政区划的案例）：

序号	项目	案例（行政区划）
1	编号	SDT0001
2	名称	行政区划
3	说明	设有国家政权机关的各级地区的名称，按层级顺序表示：省（自治区、直辖市、特别行政区）、市（地区、自治州、盟）、县（自治县、市、市辖区、旗、自治旗）。
4	数据表示	C..40
5	计量单位	
6	同义词	
7	标准字段名	DIV_NAME

8	值域	采用 GB/T 2260 《中华人民共和国行政区划代码》中的名称
9	语境	通用
10	国标对照编号	1110
11	备注	

数据元编号建议采用特征组合编码方式进行编码,一个建议的编号模式如下:



2.1.1.2. 元数据维护

根据业务需求,制定统一的技术元数据和业务元数据标准,覆盖多种来源统计数据采集、加工、清洗、加载、多维生成、分析利用、发布、归档等各个环节,建立相应的管理维护机制,梳理并加载各种元数据。根据统计业务特点,制定元数据标准,要满足统计元数据在口径、分类等方面的历史变化。支持对派生元数据的管理,如派生指标、代码重新组合等,对元数据管理实行权限控制。

2.1.1.3. 统一管理和利用

通过元数据,实现对各类数据的统一管理和利用,包括:

- 填报数据: 与各级业务数据系统的无缝集成。
- 基础数据管理: 建立各类基础数据与元数据的映射关系,实现统一的数据查询、处理、报表管理。
- ETL: 通过元数据获取 ETL 规则的描述信息,包括字段映射、数据转换、

数据转换、数据清洗、数据加载规则以及错误处理等。

- 数据库：利用元数据实现对数据库结构的描述，包括仓库模式、视图、维、层次结构维度描述、多维查询的描述、立方体（CUBE）的结构等。
- 版本控制及追溯
- 对于元数据的维护版本进行控制和追溯。

2.1.2. 指标元数据管理

作为一项特殊的业务元数据，指标元数据在业务应用中具有非常重要的地位，在元数据管理系统的基础上，实现对信息资源的指标元数据的管理与维护功能：

- 对指标元数据的分类管理



The screenshot displays a software interface for managing indicator metadata. On the left, a tree view shows a hierarchy starting with '指标元数据' (Indicator Metadata), which includes '公共指标' (Public Indicators), '单位基本情况' (Unit Basic Information), '公用分类' (Public Classification), and '专业指标' (Professional Indicators). Under '专业指标', there are sub-categories like '房地产业' (Real Estate), '劳动' (Labor), '服务业' (Service Industry), '工业' (Industry), and '投资' (Investment). The right pane, titled '指标元数据专业指标' (Professional Indicator Metadata), contains a table with the following data:

	名称	备注	创建时间	操作
1	房地产		2010-08-19	删除
2	能源	能源	2011-10-11	删除
3	劳动		2011-10-11	删除
4	服务业		2011-10-11	删除
5	工业		2011-10-11	删除
6	投资		2011-12-05	删除

- 基本信息维护

指标详细信息 -- 网页对话框

指标路径: 数据仓库/年、订报/2009/房地产/企业财务状况【X103】

指标名称: <input type="text" value="投资性房地产"/>	指标类型: <input type="text" value="统计指标"/>
显示名称: <input type="text" value="投资性房地产"/>	运算规则: <input type="text" value="汇总求和"/>
指标状态: <input checked="" type="radio"/> 生效 <input type="radio"/> 失效	计量单位: <input type="text" value="货币"/> <input type="text" value="千元"/>
派生指标: <input checked="" type="radio"/> 非派生 <input type="radio"/> 派生	审核状态: <input type="text" value="待审核"/>
有效开始时间: <input type="text" value="2000-09-02"/>	指标最小值: <input type="text"/> 只能填写数字!
有效结束时间: <input type="text" value="2020-09-02"/>	指标最大值: <input type="text"/> 只能填写数字!
绑定字段: <input type="text" value="X103.B30"/>	
指标描述: <input type="text"/>	

➤ 使用情况监管

指标详细信息 -- 网页对话框

元数据: 投资性房地产【投资性房地产】

路 径: 指标元数据/专业指标/房地产企业财务状况/投资性房地产

- 数据仓库
 - 年、订报
 - 2009
 - 房地产
 - 企业财务状况【X103】
 - 投资性房地产【投资性房地产】 单位: 货币 - 千元
 - 2008
 - 房地产
 - 企业财务状况【X103】
 - 投资性房地产【投资性房地产】 单位: 货币 - 千元
 - 2010
 - 房地产
 - 企业财务状况【X103】
 - 投资性房地产【投资性房地产】 单位: 货币 - 千元

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/495304230234011132>