
《环境信息数据仓库系统技术规定》

(征求意见稿)

编制说明

《环境信息数据仓库系统技术规定》编制组

二〇一〇年十月

目 录

1	项目背景.....	1
1.1	任务来源.....	1
1.2	工作过程.....	1
2	制（修）订的必要性分析.....	1
2.1	国家及环保主管部门的相关要求.....	1
2.2	现行环保标准存在的主要问题.....	2
3	编制的依据与原则.....	2
3.1	编制的依据.....	2
3.2	编制的原则.....	2
4	主要技术内容.....	2
4.1	适用范围.....	2
4.2	技术规定的结构框架.....	2
5	对实施本技术规定的建议.....	3

《环境信息数据仓库系统技术规定》编制说明

1 项目背景

1.1 任务来源

我国环境保护形势日益严峻，污染减排工作得到了党中央和国务院的高度重视。为此，国务院提出建立和完善污染减排“三大体系”，包括科学的减排指标体系、准确的减排监测体系和严格的减排考核体系。环境保护部针对污染减排“三大体系”的要求，进行国控重点污染源自动监控项目建设、污染源监督性监测项目建设、环境监察执法项目建设、环境信息与统计能力项目建设。为了保障“环境信息与统计能力”项目顺利实施，需要建立相应的减排能力建设类技术规范，保障系统建设上下级之间标准一致，为环境信息管理和资源整合与协同共享奠定基础。

2009年10月30日，环境保护部信息中心在北京召开了“国家环境信息与统计能力建设项目”技术标准规范申报评议会，面向地方环保系统和社会公开征集标准规范研究协作单位。经专家评议与审核，2009年11月，信息中心以环信发[2009]11号文“关于确定‘国家环境信息与统计能力项目’技术标准规范协作单位的通知”，确定了各标准规范的承担单位，下达了标准规范编制任务。福建省环境信息中心承担《环境信息数据仓库系统技术规定》标准规范的编制工作。

1.2 工作过程

（1）编制任务正式启动

2009年11月26日，标准编制组在福州召开了标准编制工作启动会，正式确定了《环境信息数据仓库系统技术规定》标准编制组成员、组织形式、工作机制，项目具体研究内容、研究范围，项目工作计划、时间安排，工作任务分工。

（2）开题论证会情况

2009年12月14日，环保部在北京组织召开了“国家环境信息与统计能力建设项目”标准规范“环境信息数据仓库系统技术规定”开题论证会。论证专家委员会由中部信息中心领导、各单位专家组成。

专家组听取了编制组对标准的编制思路、前期调研、大纲内容、任务分工、经费预算及进度计划的汇报，审阅了相关文档，并进行了认真讨论，认为标准编制大纲和开题论证报告符合“环境信息与统计能力建设项目”标准编制需求，标准编制的基础环境基本成熟，技术路线可行，具有可操作性，标准编制进度及人员安排基本合理，同意编制组按计划开展下一步工作。

2 制（修）订的必要性分析

2.1 国家及环保主管部门的相关要求

本技术规定的制订是“环境信息与统计能力建设项目”建设的需要。

“环境信息与统计能力建设项目”以贯彻落实党中央、国务院关于节能减排工作部署为指导，以实现“十一五”期间重点污染物减排的目标指导为紧要任务，围绕建立与完善“科学的减排指标体系、准确的减排监测体系、严格的减排考核体系”的要求，加强数据传输、共享和应用能力、业务应用支撑能力、统计基础能力等环保信息化能力的建设，为实现“十一五”节能减排和换将保护工作目标奠定基础。

《环境信息数据仓库系统技术规定》是环境信息化服务保障机制的主要组成部分，是保证环境数据共享和服务的重要手段。

2.2 现行环保标准存在的主要问题

在环境信息数据仓库系统技术规定方面，目前国内还是空白。

3 编制的依据与原则

3.1 编制的依据

(1) 国外相关标准和应用

国外在数据仓库的建设标准上进行了广泛的研究，并在数据仓库元数据领域、数据的ETL领域、主题数据的多维分析领域建立了相应的标准。

在数据仓库元数据管理方面有CWM（Common Warehouse Metamodel，公共仓库元模型规范），DC（Dublin Core Metadata，都柏林核心元数据）等规范，其经过不断的发展和完善已经广泛应用于数据库、数据仓库、数据挖掘、商业智能等应用领域，并指导相关应用的建设，是各数据库、数据仓库等软件系统厂商共同遵循的标准。

(2) 国内相关标准和应用

国内信息化工作起步相对较晚，特别是数据仓库在环境领域的建设案例几乎是空白，但国内也在数据仓库及地理信息系统方面的标准研究方面做了大量的工作，并在地理信息系统元数据方面建立了相应的标准，但应用领域相对还比较小，并未在环境数据仓库领域有比较全面的标准。

3.2 编制的原则

在标准制定的过程中，遵循以下基本原则：

(1) 完整性原则

标准应覆盖数据仓库体系主要内容和步骤，有效支持环境信息数据仓库系统建设。

(2) 可操作性原则

标准应采用现有主流数据仓库技术及概念，保证环境信息数据仓库系统可操作。

(3) 针对性原则

标准应体现出环境信息数据仓库系统的特点。

4 主要技术内容

4.1 适用范围

本技术规定对环境信息数据仓库系统体系架构及主要建设环节的工作方法和建设内容进行了规范。

本技术规定适用于指导国家环境信息与统计能力建设项目数据仓库系统的设计、建设和运行维护。

4.2 技术规定的结构框架

《环境信息数据仓库系统技术规定》共有13章组成，主要内容有：

第一章为适用范围：概述了本技术规定的编制目的和适用范围。

第二章为术语和定义：列出了在本技术规定中出现的相关术语及其定义。

第三章为缩略语：列出了主要缩写的定义。

第四章为环境信息数据仓库系统架构：定义了环境信息数据仓库系统的组成内容。

第五章为环境信息数据仓库系统数据源规定：规定了环境信息数据仓库系统的数据来源。

第六章为环境信息数据仓库系统数据采集层（ETL）技术要求：定义了采集层的位置和作用，定义了ETL主要功能。

第七章为环境信息数据仓库系统数据层技术要求：确定了环境信息数据仓库逻辑建模策略、物理建模策略。

第八章为BI层技术要求：定义了BI层主要应用模式。

第九章为环境信息数据仓库系统元数据技术要求：定义了环境信息数据仓库系统元数据管理需遵守的规则。

第十章环境信息数据仓库系统标准规范体系要求、第十一章环境信息数据仓库系统基础设施要求、第十二章境信息数据仓库系统数据安全要求和第十三章环境信息数据仓库系统运行维护要求对环境信息数据仓库系统其他组成部分进行了定义和要求。

5 对实施本技术规定的建议

由于数据仓库技术规定在国内尚属空白，本技术规定还需补充、完善和细化，建议在实施本技术规定时多交流、多学习同行及其他部委机构的成功经验。

6 附录 征求意见情况汇总处理表

序号	标准条款编号	意见内容	提出单位	处理意见及理由	备注
	5.3	<p>建议在数据层中增加操作数据存储区，操作数据存储区主要用于存放保留业务发生的原始信息及历史变更信息，以支撑相关业务查询等基本信息服务。设立操作数据存储区，对于数据从数据源到数据仓库是十分必要的，操作数据两大应用特点是，以获取和提供全局一致的综合信息为目标，但并不需要考虑太多的历史数据，并且要求有较快的响应速度。它的作用主要体现在以下几个方面：</p> <p>1) 在业务系统和数据仓库之间形成一个隔离层。</p> <p>2) 对采集的数据进行必要的校验和比对。</p> <p>3) 转移一部分业务系统细节查询的功能。</p>	总集成单位	未采纳。由于采用的基础数据库已经起到了业务系统和数据仓库之间隔离层的作用，现有的数据仓库系统性能支持设计到细节级数据，设计操作数据存储区还会大幅增加系统的复杂性不利于后期维护，因此不需要增加操作数据存储区。	
	4	在体系架构中增加标准体系规范，因为标准体系规范标准体系架构是规范数据仓库各个部分，包括运维、应用、数据、基础设施等各方面建设内容的标准与规范的综合，用来指导、协调和优化数据仓库的建设与运维。	总集成单位	采纳。见图5环境信息数据仓库系统架构图。增加章节10进行要求。	
	5	在体系架构中，缺少数据仓库基础设施层，也就是缺少系统运行所需的软硬件及网络环境。	总集成单位	采纳。见图5环境信息数据仓库系统架构图。增加章节11进行要求。	
	第二章 规范性引用文件	整个章节缺失	工程监理	未采纳，同时修改编制说明中关于规范性引用文件部分。	
	第五章以后内容	征求意见稿与编制说明描述的内容不完全一致	工程监理	采纳。按照征求意见稿对编制说明第4节进行调整	
	第11章 环境信息数据仓库系统运行维护要求	缺少对于运维体系建立培训的要求	工程监理	采纳。见13章环境信息数据仓库系统运行维护要求。	
	2	术语定义的出处	环保部信息中心	采纳。在引用的定义增加出处。	

序号	标准条款编号	意见内容	提出单位	处理意见及理由	备注
	1	应定义一下环境数据仓库建设的内容和主要目的	环保部信息中心	采纳。对适用范围进行明确。	
	图 5	环境数据仓库架构图有问题	环保部信息中心	采纳。对图 5 进行设计调整。	
	5	数据源范围	环保部信息中心	采纳。调整了章节 5 环境信息数据仓库系统数据源规定。	
	7.1.4 度量值设计	度量值 根据业务需要选取数值型指标。	环保部信息中心	采纳。调整了 7.1.4 度量值设计。	
	7.1.2 粒度设计	粒度设计针对具体的环境业务来说	环保部信息中心	采纳。调整了 7.1.2 粒度设计。	
	9 元数据要求	试图描述全，但感觉凌乱，不系统。	环保部信息中心	采纳。调整了章节 9 元数据技术要求。	
	2	部分定义存在问题	环保部信息中心	采纳。调整了 2.3 数据集市定义，2.4 决策支持系统定义，2.14 维度建模，2.14 事实表，2.21 表分区，2.22 商业智能，2.23 数据挖掘，2.24 聚集。	
	图 1	ETL 示意图有问题	环保部信息中心	采纳。根据抽取、清洗、转换、加载步骤调整了示意图。	
		数据源定义	环保部信息中心	待定	
	4.2.2	数据抽取清洗、加载部分需要调整		采纳。将抽取、清洗、加载分开进行描述。	
		其他格式及小细节问题。	环保部信息中心	采纳，已修改。	