引文格式：曹千红,沈琼,陆海霞. 自然资源指标体系探索与实践[J]. 地理空间信息, 2023, 21(7): 86-89. doi:10. 3969/j.issn. 1672-4623.2023.07.021

自然资源指标体系探索与实践

曹千红[1]，沈　琼[1]，陆海霞[2]
（1. 湖北省空间规划研究院，湖北 武汉 430071；2. 湖北省洪湖市府场镇自然资源和规划所，湖北 洪湖 434000）

摘　要：为加快适应数字政府建设的新形式，针对自然资源行业数字化转型过程中由于底层数据库标准各异而导致的统计口径不统一、数据共享难、数据资源无法有效利用等问题，研究设计了一套基于自然资源指标体系的混合模式数据仓库，将范式建模和维度建模两种方式进行有机结合，解决了多源异构自然资源数据的统计、分析、更新等问题。基于此，建立了自然资源综合统计系统，为建设用地统计、分析、管理以及相关业务提供了支撑服务。

关键词：自然资源指标体系；综合统计；数据仓库
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2023）07-0086-04

Exploration and Practice of Natural Resource Indicator System

CAO Qianhong[1], SHEN Qiong[1], LU Haixia[2]
(1. Hubei Spatial Planning Research Institute, Wuhan 430071, China;
2. Hubei Province Honghu Fuchang Town Institute of Natural Resources and Planning, Honghu 434000, China)

Abstract: In order to accelerate the adaptation to the new form of digital government construction, solve the problems such as inconsistent statistical caliber, difficult data sharing and ineffective use of data resources due to the different underlying database standards in the process of digital transformation of natural resource industry, we designed a set of hybrid data warehouse based on the natural resource indicator system which combined normal form modeling and dimension modeling, to solve the problem of statistics, analysis and update of multi-source heterogeneous natural resource data. On the basis of this, we established a comprehensive statistical system of natural resources to provide support services for the statistics, analysis, management and related businesses of construction land.

Key words: natural resource indicator system, comprehensive statistics, data warehouse

近年来，党中央高度重视发展数字经济，“数据即资产”成为核心的产业趋势。2022年6月6日国务院发布《国务院关于加强数字政府建设的指导意见》，明确加强数字政府建设是适应新一轮科技革命和产业变革趋势、引领驱动数字经济发展和数字社会建设、营造良好数字生态、加快数字化发展的必然要求，是建设网络强国、数字中国的基础性和先导性工程，是创新政府治理理念和方式、形成数字治理新格局、推进国家治理体系和治理能力现代化的重要举措，对加快转变政府职能，建设法治政府、廉洁政府和服务型政府意义重大。为加快适应数字政府建设的新形式，自然资源部门数据统计工作在完善统计指标、创新统计方式、推进统计信息化等方面开展了积极探索，在数据共享和开发利用方面取得了积极进展；但统计系统建设仍存在底层数据库标准不够完善、数据壁垒、统计口径亟需统一等问题，距离自动化预警预测、智能化辅助决策的统计分析需求还存在较大差距[1]。本文基于《湖北省自然资源信息化建设总体实施方案（2020—2025年）》中自然资源综合统计与辅助决策系统的建设成果展开分析。该系统的建设目标是打造湖北省的自然资源“决策大脑”，成为领导决策的抓手、业务人员办公的帮手以及省厅数据统计的唯一出口。为实现该目标，探讨了基于指标体系的数据仓库设计与应用实践，并以批准建设用地指标应用为例展开应用研究，为实现自然资源数字化转型提供实践基础。

1　需求分析

随着自然资源管理信息化建设的发展，收集了庞大的自然资源数据。如何“让数据持续用起来”、实时展示最新的自然资源概况、应对各种业务维度的统计分析，是自然资源管理数字化转型需要面对的问题[2]。传统的面向单一业务的系统架构模式已不能适应海量数据的分析需求，数据仓库建设迫在眉睫。其

收稿日期：2023-03-03；修回日期：2023-04-10。
第一作者简介：曹千红（1978—），高级工程师，研究方向为土地资源管理，E-mail：1004011352@qq.com。

本质是对多个异构数据源的集成，既是对事务型信息系统的数据集成，又是对分析型数据的统计，使沉淀的数据发挥最大价值，以满足业务统计、分析决策等各方面对信息化应用的期望。数据仓库包括基于Inmon的三范式关系建模和基于Kimball的维度建模两种方式。Inmon[3]的核心是要求规范化数据库表结构，优点为没有数据冗余、保证了数据的一致性，解耦数据、方便数据维护，常用于金融和电信等业务模式稳定的行业；缺点为开发周期慢、需耗费大量时间梳理和设计数据表结构，表的数量多、降低了查询性能。Kimball则是从需求端出发，根据具体业务场景构建数据仓库分析模型，优点为数据表结构可根据需求灵活变换、开发周期短，较常用于业务需求快速迭代的互联网行业；缺点为存在大量数据冗余，后期维护成本高。数据仓库中的指标是指将业务单元细分后量化的度量值，使得业务目标可描述、可度量、可拆解。它是业务与数据的结合，是统计的基础，也是量化效果的重要依据。自然资源指标体系是以“框架统一、需求拓展”为宗旨，在部、省级自然资源综合统计指标体系的基础上进行细化与拓展，最终构建的满足省、市、县、乡、村5级指标体系。

自然资源业务复杂、系统繁多，采取任何单一数据仓库建模方式均不能满足行业对新型数据的分析需求。因此，本文探索将结合两种建模方式的自然资源指标体系应用于数据仓库设计。设计思路为：将数据仓库分为贴源数据（ODS）层、统一数仓（DW）层和应用数据（ADS）层，由于ODS层和DW层规范化程度高、数据格式相对稳定，采用范式建模方法（Inmon模式）；而ADS层业务需求多变，特异性、定制化程度高，则采用维度建模方法（Kimball模式）。基于此，按照“松耦合、层次化”的基本原则建设符合自然资源行业的混合模式数据仓库。

2　数据仓库设计

2.1　基于Inmon模式的ODS层

ODS层与源数据是保持一致的，主要收集各业务系统生产的原始数据[4]。由于各业务系统的数据库架构均满足范式要求，因此ODS层数据也符合范式设计规范。

2.2　基于Inmon模式的DW层

DW层为数据处理的核心[5]。Immon模式主张自上而下的架构，需系统梳理自然资源业务相关内容并形成指标体系，以满足多层级、多维度的数据综合统计需求。自然资源领域涉及的业务繁多，建设的业务系统相互独立，没有通用的数据标准，无法统一管理指标数据，导致数据间相互矛盾，失去了数据统计分析的参考价值。因此，自然资源综合统计系统需要单独设计一组统一的、有逻辑的定义规范数据指标，即指标体系。

在全面分析自然资源数据统计信息化建设现状的基础上，按照“统一组织开展、统一法规依据、统一调查体系、统一分类标准、统一技术规范、统一数据平台”的要求，为现有各种自然资源综合统计调查的技术、方法、措施、成果设计了一套省、市、县3级统一的自然资源综合统计指标体系[6]，包括土地指标、国土空间规划指标、不动产登记指标等8类单项指标以及综合指标、复合指标和社会经济指标，详见图1。

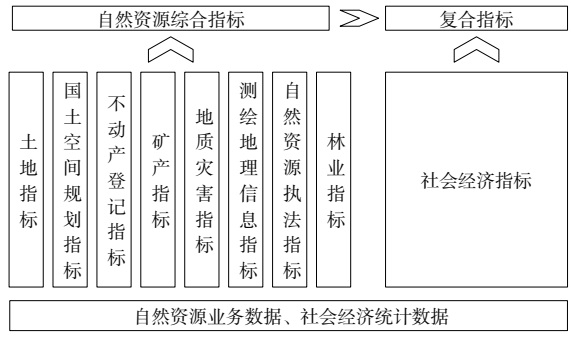


图1　自然资源综合统计指标体系

2.3　基于Kimball模式的ADS层

Kimball模式主张自下而上的架构，即通过业务需求建立指标和维度，进而构建数据仓库的分析模型。分析模型是指星型模型或雪花模型。通过数据仓库可从各业务系统的多个事务性型数据库中收集信息，以统一的数据模式实现OLTP与OLAP的业务数据分离[7]。星型模型是一个类似于星型的实现结构，解决了操作型数据之间的复杂关系，适用于创建简单且迅速的数据分析及其报告。雪花模型则是对星型模型的扩展，当一个或多个维度表没有直接链接到事实表上，而是通过其他维度表链接到事实表上时，其图解就像多个雪花连接在一起，故称雪花模型。星座模型是由星型模型延伸而来的，星型模型和雪花模型是基于一张事实表，而星座模型是基于多张事实表，且共享维度表信息，这种模型往往应用于数据关系更复杂的场合。自然资源业务复杂，每个业务都对应一张事实表，因此符合星座模型的特点。

本文以土地批供流转中的数据模型为例，利用星座模型对批地项目和供地项目数据建模（图2）。通过建立批地项目、供地项目事实表，配合案卷名称、审

批级别、用途、行政区划等维度，即可在系统中进行查询，如可直接查询审批级别为“国务院审批”、用途为“水域及水利设施”、批复时间为“批复时间年”的项目信息。

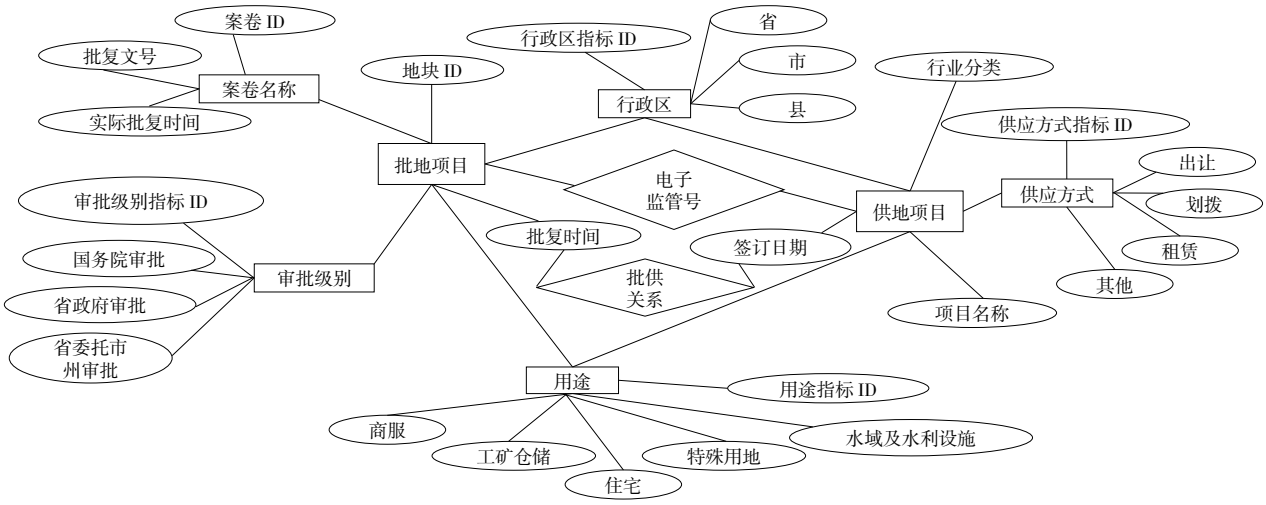


图2　批供关系星座模型

2.4　混合模式数仓指标

基于上述自然资源综合统计指标体系设计和数仓建模方法，可以很方便地进行数仓指标建设。指标体系的树形结构见图3，其中的指标节点可直接对应数据仓库中的一个分析指标，并可向上汇总成为更高维度的指标。

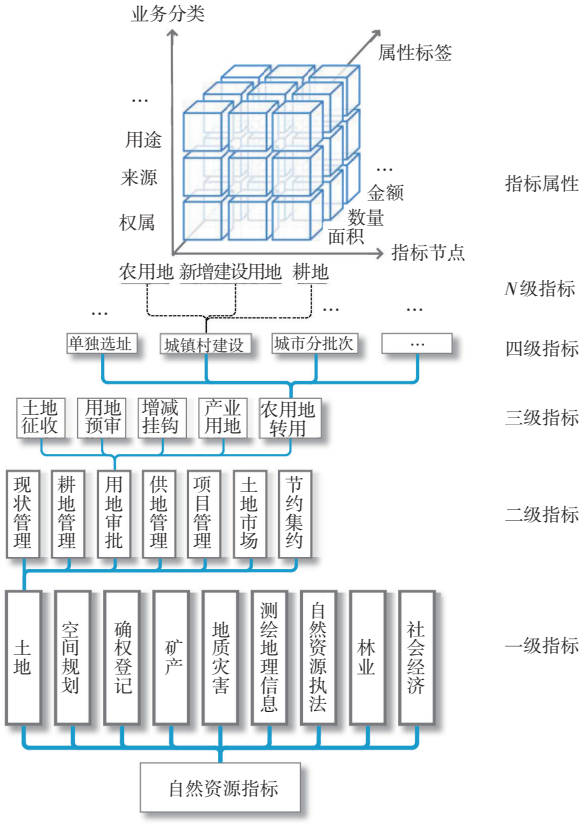


图3　混合模式数仓指标

2.5　数据编码方案

在整个数据仓库中，采用统一的方式维护各类定义信息是非常必要的，也是统一数据统计口径的重要手段[8]。例如，在“建设用地批准”中，“省政府审批”的案卷在综合统计口径下包含了“省委托市州审批”的案卷，而在用途管制统计业务中则单独划分为“省政府审批”和“省委托市州审批”两个统计项，若没有统一的定义标准，容易造成数据仓库数据访问获取困难、数据质量不高等问题。基于此，系统配合数仓采用数字化建模方式，对指标体系进行数字编码。由于自然资源综合统计指标体系将指标层级划分为7级，因此指标编码采用7级分类，预留一级作为顺序码，方便后续扩展。整个指标编码采用8层16位层次码结构，由指标类型特征码和顺序号构成，指标类型特征码共7层，每层两位编码，顺序号为两位编码（图4）。在考虑节约存储空间的情况下，最多可支持10^{16}个指标进行编码。同时，严格规定了指标不同类型对应的编号，以保证系统生成的编码符合规则且不重复。“一级类”对应指标树的“一级指标”，“二级类”对应 “二级指标”，以此类推。通过该指标编码可清晰划分对应的指标所属层级。在上述问题中，将“省政府审批”包含“省委托市州审批”的指标用4级编码“0202020103000000”表示，将不包含“省委托市州审批”的指标用5级编码“0202020103010000”表示，即可快速区分对应的不同指标含义。指标体系实现了指标的规范统一，且指标编码的分层设计很好地支持了数据的递归统计分析。

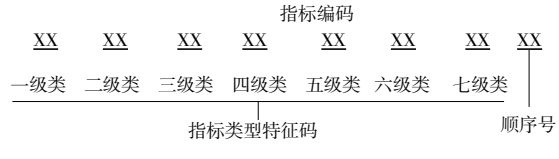


图4　指标体系编码规则

3　基于指标体系的数据仓库应用实践

3.1　建设用地全流程年度流量分析

传统业务系统通常只做到了单个项目的建设用地全流程监管，自然资源管理部门领导不能实时掌握建设用地的整体情况。例如，领导想知道当年批了多少地，有多少已经供出去了，或者当年供了多少地，是哪些年份批复的？传统数据库结构由于不能提前预知组合条件，需要开发人员进行条件定制开发，成本高、效率低，且随着数据量和组合条件的增加，数据库响应速度也会大幅降低，严重影响用户使用。

系统采用基于指标体系的数据仓库技术，提前通过指标体系对数据进行预计算，使数据响应时间由超过10 s缩减至1 s以内。如图5所示，可分别获取批地、供地、用地等数据的年度汇总情况；通过指标体系中的“电子监管号”、“政府文号”、“项目名称”等关键字段建立单个项目在批、供、用、补、查各环节的联系；再根据各环节的审核时间分别进行统计汇总，可获取对应环节的历年数据来源与数据去向；最后利用能量分流图，以流量的形式直观展示统计结果，达到“实时变换一图尽览”的效果。

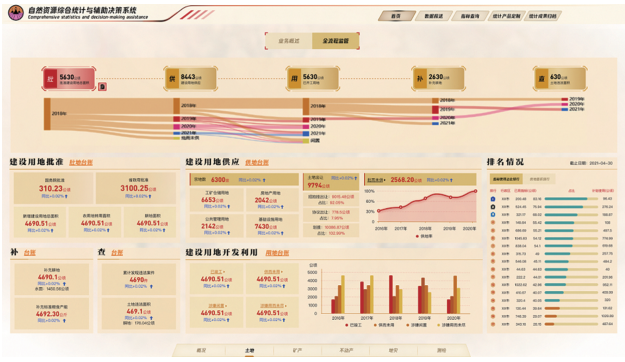


图5　土地全流程年度流量分析图

3.2　自定义组合维度指标查询

自然资源管理部门需要每月制作月度简报，向领导反映自然资源整体状况。传统统计方法存在统计维度固定、数据更新滞后等问题。通过自然资源综合统计系统开发指标查询功能，借助数据仓库的OLAP技术，全面支持任意组合指标维度的统计与分析，原业务需要2 d整理完成的数据，系统2 min即可统计完成，极大地提升了数据统计与分析的工作效率。用户点击“指标查询”标签页，选择“建设用地审批情况”目录树，系统可提供初始查询结果；勾选业务分类、统计来源、业务类型指标属性，选择时间、区域等维度信息，即可查询当前最新的批地数据情况；勾选同比、占比等选项，可反馈数据变化趋势；选择单位切换，可根据公式进行快速换算（图6）。



图6　指标组合查询图

4　结　语

本文对自然资源指标体系的构建进行了探索，并在自然资源综合统计系统上进行了应用，实现了自然资源统计指标的统一和标准化，为自然资源数据分析与应用打下了坚实基础，为信息技术创新支撑政府数据治理提供了新思路。然而，随着统计需求的日新月异，指标的不断更新变化，如何融合新旧指标、兼容不同时期的数据分析需求还需进一步研究探索。

参考文献

[1]　中共中央国务院. 国务院关于加强数字政府建设的指导意见[Z].

[2]　付登坡,江敏,任寅姿,等. 数据中台:让数据用起来[M]. 北京:机械工业出版社,2020

[3]　黎昊. 关于业务系统中应用数据中台技术的研究[J]. 数字化用户,2019(46):82

[4]　钱文杰,王强,王新洲,等. 苏州工业园区污染源企业环境行为大数据评价系统研究[J]. 环境生态学,2019,1(3):46-52

[5]　董星星,段修亭,陈肖磊,等. Web分布式空间数据仓库体系结构设计[J]. 地理空间信息,2011,9(4):91-93

[6]　自然资源部办公厅. 自然资源标准体系[S].

[7]　刘耀林,何建华. 地理国情多层次统计分析指标体系设计[J]. 地理空间信息,2014,12(3):1-5

[8]　何丽华,於新国,于胜杰. 城镇空间适宜性评价方法与应用[J]. 地理空间信息,2021,19(8):9-12