引文格式：侯璨, 李志飞, 李江河, 等. 不动产数据清理系统的设计与实现——以长沙市为例[J]. 地理空间信息, 2023, 21(10): 108-111. doi:10.3969/j.issn. 1672-4623.2023.10.027

不动产数据清理系统的设计与实现——以长沙市为例

侯　璨[1,2]，李志飞[3]，李江河[1]，吴希欢[1]
（1. 广东南方数码科技股份有限公司，广东 广州 510630；
2. 上海飞未信息技术有限公司湖南分公司，湖南 长沙 410005；
3. 湖南星图空间信息技术有限公司，湖南 长沙 410004）

摘　要：针对长沙市存量不动产登记数据质量不均、数据标准不一、图形未落宗关联等问题，结合长沙市不动产登记业务数据现状，设计开发了一套“在线”+“离线”结合模式的不动产数据清理系统，实现长沙市存量不动产登记数据的流程化深度清理。在保证数据清理成果与不动产登记业务系统实时交互的同时，实现清理过程全留痕、修改记录可回溯。基于长沙市2021年存量不动产登记数据深度清理的统计数据表明，随着存量不动产数据深度清理的逐步推进，不动产日常业务数据修改和数据核查的工作量均逐渐减少，说明存量不动产数据清理取得了良好的效果。

关键词：不动产；数据清理；数据库；系统
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2023）10-0108-04

Design and Implementation of Real Estate Data Cleaning System: A Case Study of Changsha City

HOU Can[1,2], LI Zhifei[3], LI Jianghe[1], WU Xihuan[1]

(1. South Digital Technology Co., Ltd., Guangzhou 510630, China;
2. Hunan Branch, Shanghai Feiwei Information Technology Co., Ltd., Changsha 410005, China;
3. Hunan Satellite Image Information Technology Co., Ltd., Changsha 410004, China)

Abstract: Considering uneven data quality, unified data standards and the unresolved graphs of real estate registration stock of Changsha, we designed a set of “online+offline” system for real estate data based on the current situation of real estate registration data in Changsha, realizing in-depth process cleaning of the existing data. While ensuring real time interaction between the data cleaning results and the real estate registration business system, the system can remember and trace all traces of the cleaning process and modification records. The statistical data based on the deep cleaning of real estate registration stock of Changsha in 2021 shows that the workload of daily real estate business data modification and data verification has gradually decreased with the advancement of the deep cleaning of stock real estate data, indicating the data cleaning system has achieved good results.

Key words: real estate, data cleaning, database, system

近年来，自然资源部、各省不动产登记主管部门不断发文倡导进行不动产登记数据整合。将“存量不动产登记数据”按照不动产登记数据库标准进行整合、理清登记数据业务历史关系、建立“图形、属性、档案”联动的不动产登记库，是开展不动产数据整合工作的目标，也是不动产登记工作顺利、高效实施的有利保障。不动产登记存量数据深度清理在不动产统一登记工作中具有十分重要的意义[1-5]。

本文以长沙市存量不动产登记数据整合为例，在满足长沙市日常不动产登记业务的基础上，设计开发了一套基于完整数据整合流程的、“在线”+“离线”结合模式的不动产数据清理系统，为存量不动产登记数据整合提供可视化平台和技术支撑，实现不动产数据深度清理即清即用。该系统与其他不动产数据清理系统相比，能够充分利用数据库管理工具和GIS数据处理软件，实现不动产数据在离线系统中批量清理，极大地提高了数据清理效率。且该系统与不动产登记业务系统实时交互，避免不动产数据清理与不动产日常业务办理冲突产生新的数据问题。

收稿日期：2022-06-20。
第一作者简介：侯璨（1990—），硕士，工程师，注册测绘师，主要从事地理信息应用、自然资源信息化建设等工作，E-mail：canning000@126.com。

1　长沙市不动产登记业务数据现状分析

长沙市不动产登记中心2021年业务受理量为494 901笔，每日平均业务受理1 980笔，业务办理过程中发现问题而进行数据修改共2 324笔，每日平均修改近10笔。日常业务办理量大、数据修改量较为常见。同时存量不动产登记数据质量不均，部分不动产登记业务因受制于数据质量和历史遗留问题，办理时间长达一个月。

设计不动产数据清理系统在满足不动产数据安全性、规范性、完整性清理的基本要求下[6]，能够衔接不动产登记业务系统，实现边清理边应用，不仅是不动产数据整合的要求，也是支撑日常不动产登记工作顺利开展、降低业务风险的关键[7]。

2　系统设计

2.1　系统架构设计

系统采用C/S体系架构，结合各类标准规范体系和安全管理体系，自下而上分为基础设施层、数据服务层、公共支撑层和应用层4个部分。

1）基础设施层是对数据清理系统底层支撑设施的表现，包括支撑系统运行所需的服务器、存储设备、安全设备、网络设施、操作系统，以及组织系统数据存储与管理的Oracle数据库，提供空间数据支撑的ArcSDE等。

2）数据服务层包括支撑数据清理系统运行所需要的各类数据，由不动产登记业务库数据、不动产空间数据库数据、不动产档案数据、高时效遥感影像、不动产清理在线库、不动产清理离线库等组成。

3）公共支撑层是数据清理系统开发所用的工具和软件环境。不动产登记业务数据和楼盘表管理基于微软.NET Framework4.0框架，选用C#开发语言，保证了系统技术框架的稳定性、兼容性、可扩展性。空间数据管理、分析与展示基于Esri公司的ArcEngine框架进行开发。

4）应用层是不动产清理系统的需求和功能的表现，提供可视化的操作界面，实现不动产数据在线抽取、不动产数据在线清理、不动产数据离线清理、不动产数据检查、不动产数据抽查、日常业务与清理业务关联和不动产数据回库等功能。

2.2　数据库设计

数据清理系统数据库采用大型关系型数据库Oracle 11g进行存储，管理不动产登记业务库数据、不动产档案数据、不动产清理在线库和不动产清理离线库。

同时使用ArcSDE作为空间数据引擎，管理存储于Oracle数据库中的空间数据，确保数据清理过程中，属性数据、档案数据与空间图形数据可同步关联，同步整合[8]。

清理系统库表字段设计，以长沙市不动产登记库数据标准为基础，结合《不动产登记暂行条例》[9]《不动产数据整合建库技术规范（试行）》[10]等标准规范以及长沙市不动产数据深度清理的实际需求，建立登记业务库与清理库之间的对应关系，形成满足存量数据整合与汇交要求的清理库库表结构。

3　系统关键技术实现

不动产数据清理系统进行存量不动产数据整合的流程为：不动产数据抽取–>在线数据清理–>离线数据清理–>数据质检–>业务科室抽检–>数据入库。

清理库：不动产数据整合工作人员从不动产登记业务系统登记库抽取，待清理的存量不动产登记数据。每次抽取数据，系统会自动记录数据抽取人员的ID、抽取时间等信息，以作留痕管理；同时将数据抽取时的不动产登记信息作备份，备份的数据在不动产数据深度清理过程中属性一直不变，用于数据整合各环节进行数据比对。清理库与不动产登记库部署于同一网络环境、不同数据库中，可进行数据交互。

在线库：从不动产登记业务系统中抽取不动产登记数据，随着不动产数据深度清理的进行，不断地修改、完善的清理数据。在线库与清理库存储于同一个数据库中，不动产在线清理系统数据来源于此数据库。

离线库：单独部署在不动产数据深度清理服务器上，与在线库数据库结构完全一致的离线库。离线库中数据依据不动产数据整合需要，由不动产在线清理系统通过数据同步分发至离线库。不动产离线清理系统数据来源于此数据库。

不动产在线清理系统：与不动产登记业务系统部署于同一网络环境，可实现不动产数据抽取、可视化界面数据整合的流程版数据清理系统。为避免数据整合过程中，因操作失误或核查不到位而引起数据批量错误，数据库权限不对数据整合工作人员开放。可依据数据整合情况，将待清理数据分发至不动产离线清理系统进行数据深度清理。

不动产离线清理系统：单独部署在不动产数据整合服务器上的非流程版清理系统，拥有与在线清理系统相同的数据整合可视化界面，还可将数据库权限对

不动产数据整合工作人员开放，可通过数据库管理工具实现不动产数据增加、删除、修改、查询。

3.1 系统流程实现

不动产数据清理系统以数据清理流程为导向，在线清理系统，数据清理人员通过宗地、自然幢、房屋等查询条件，查询出需要清理的业务。选择需要清理的业务，提交至在线清理系统进行数据在线清理。系统后台自动记录需要清理的业务的登记编号。系统支持以宗地及其附着房屋为单位进行整合，简单业务数据整合在不动产在线清理系统中进行，通过数据规范整理或数据新增、补录，确保数据关联和图形挂接正确后经质检合格后完成数据清理；复杂业务数据可应用“在线”+“离线”数据同步技术，将数据从不动产在线清理系统中同步至不动产离线清理系统中进行深度清理，离线清理整理完毕后，数据同步到在线清理系统，继续后续的流程。

数据离线阶段，需要对房地数据进行现状核对、图形清理和业务历史梳理。在离线清理系统中，可在系统界面进行可视化数据清理，也可在数据库后台，利用数据库管理工具批量修改数据。可充分利用成熟的GIS数据处理软件对图形数据进行编辑处理。梳理业务历史过程中，需要对查阅相关档案进行核实、补录。部分缺失宗地红线或户室分层分户图的数据，还需要开展外业测绘与调查，形成满足不动产发证需要的不动产测绘数据。离线整理完成的存量不动产数据需要进行自检，自检合格后，应用“在线”+“离线”数据同步技术，将清理后的不动产数据由不动产离线清理系统同步至不动产在线清理系统中。

数据同步到不动产在线清理系统后，将进行数据质检和业务科室抽检，质检无误后，回写到登记业务库，完成数据深度清理。回写登记业务库时，系统将自动检测清理过程中，不动产登记业务系统是否有被清理土地和户室的新办业务。若清理过程中无新办业务，则依据清理结果更新登记业务库，否则依据新办理的业务更新业务状态。

3.2 “在线”+“离线”数据同步技术

长沙市不动产数据清理系统包括部署在自然资源专网的“在线”清理系统，和部署在不动产数据整合服务器上的“离线”清理系统，两者数据库表结构一致、界面一致。“在线”清理系统中，可以选择把业务数据同步分发到“离线”清理系统进行深度清理。系统通过数据库定时任务进行数据同步管理调度，通过系统后台同步服务，定时读取日志表，获取需要分发到“离线”清理系统的业务，分批次自动同步分发到“离线”清理系统。每一笔登记业务会同步分发到“离线”清理系统的时间、状态（成功或失败）。对于同步分发失败的数据，下一次任务开始时，会重新读取，进行同步分发操作。“离线”清理系统中，数据深度清理完成后，点击“整理完成”按钮，系统自动记录清理状态。对于标识“整理完成”状态的业务，自动同步回写到“在线”清理系统。离线数据回写到“在线”清理系统库时，采用表对表、字段对字段的全覆盖模式。同步回写后会更新对应的同步回写时间、回写状态（成功或失败）。对于同步回写失败的，下一次任务开始的时候，会重新读取，进行同步回写操作。同步回写时，系统自动记录回写时间、数据在“离线”清理系统中修改内容以及数据清理人员等信息，在提高数据整合效率的同时，可实现数据清理留痕管理，降低数据整合安全风险。

3.3 楼盘表管理

系统提供楼盘表快速查询功能，可查看楼盘表的数据分布和业务状态，楼盘表对应的数据与不动产登记业务系统保持一致。在楼盘表管理模块，可进行层信息和户室信息的新增、修改、删除等管理。

3.4 清理数据质量检查

不动产清理系统提供不动产整合数据的自动质检功能，包括对图形的拓扑关系检查和属性的完整性、必填项、值域检查，系统支持输出质检报告。

3.5 清理系统与不动产登记业务系统数据交互

不动产在线清理系统保持与不动产业务系统的实时数据交互，正在办理不动产业务的数据不允许进行数据整合；同时，整合完成的不动产数据在更新登记业务系统之前，清理系统将与不动产登记业务系统进行数据共享，对比数据深度清理期间，被整理的不动产单元是否有新办理的业务，若无新办业务，则依据清理结果更新登记系统业务库；若有新办业务，则在数据回写至登记业务系统时依据新办理的业务，更新不动产数据的业务状态。如此，保证数据的安全，避免不动产登记业务与不动产数据整合有冲突。经整合的数据，回写到登记库之后，清理系统可显示数据整合前后的变化情况。

4 系统应用与结果分析

不动产日常业务数据修改，是指在办理日常不动产登记业务过程中，发现存量不动产数据质量问题，

经查看登记档案、登记证书或证明等材料，核实数据情况后，对存量不动产登记数据进行修改的工作。不动产数据核查，是指在在办理日常不动产登记业务过程中，发现存量不动产数据质量问题，经查阅登记档案和相关材料后，无法确认不动产登记数据，需通过开展外业调查或外业测绘的工作来核实或添加不动产登记材料的工作。不动产登记日常业务能否顺利开展，不动产日常业务数据修改和数据核查业务量有无减少，是检验存量不动产数据整合是否取得良好效果的重要指标。本文以长沙市不动产存量数据整合为例，统计2021年1月到12月，随着存量不动产数据深度清理的开展，不动产日常业务数据修改和数据核查的工作量变化情况。

图1为长沙市2021年业务数据修改工作量变化趋势图，根据每月统计数据，长沙市2021年业务数据修改总量为2 324笔，其中数据修改业务量在2月份由于春节期间，业务办理量相对较少；数据修改业务量在3月份最大，为389笔，之后数据呈逐月减少的趋势。在第四季度，数据修改业务量每月均下降为一百笔以内。可见，随着不动产存量数据深度清理的不断推进，业务数据质量提升明显。

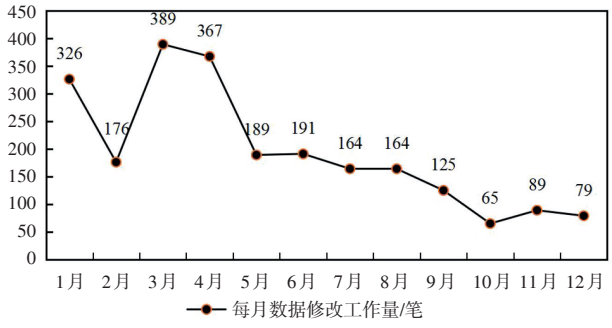


图1　2021年业务数据修改工作量变化趋势图

图2为长沙市2021年业务数据核查工作量变化趋势图，根据每月统计数据，长沙市2021年业务数据核查总量为1 958笔，相比2020年业务数据核查总量（2 911笔），总量下降了32.7%。其中数据核查业务量在2月份由于春节期间，业务办理量相对较少；数据核查业务量在3月份最大，为362笔，之后数据呈逐月减少的趋势。在第四季度，数据核查业务量每月均下降为一百笔以内。可见，随着不动产存量数据深度清理的不断推进，业务数据质量提升明显。

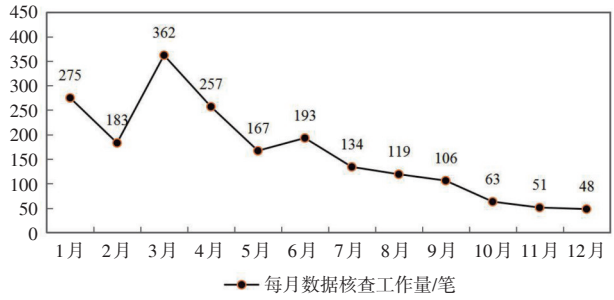


图2　2021年业务数据核查工作量变化趋势图

5　结　语

存量不动产数据因历史问题，存在字段缺失、属性错误、未落宗关联等问题，是制约不动产日常业务高效办理的主要因素。本文以长沙市存量不动产登记数据深度清理为目标，设计开发了一套“在线”+“离线”结合模式的不动产数据清理系统，实现长沙市存量不动产登记数据的流程化清理。系统可实现不动产数据整合的界面化修改，可分发数据至本机利用数据库管理工具对属性数据进行批量处理、利用GIS软件对图形数据进行编辑管理。系统对数据深度过程进行留痕管理，可对数据清理成果进行质量检查。不动产数据清理系统与不动产登记业务系统实时交互，避免数据整合与日常业务办理冲突产生新的数据问题。

参考文献

[1] 吴明辉. 不动产登记存量数据整合研究[J]. 国土资源信息化,2018(5):39-42

[2] 焦汉科,易家莉,刘振荣. 不动产登记房地存量数据整合技术路线研究:以广州市增城区为例[J]. 地理信息世界,2019,26(1):132-138

[3] 张彤蕴. 不动产存量数据整合系统的设计与实现[J]. 地理空间信息,2021,19(12):100-102

[4] 王江涛,李政,崔帅,等. 基于FME的榆林市不动产数据整合系统建设探讨[J]. 测绘标准化,2018,34(1):43-46

[5] 程丽丽. 基于SuperMap的不动产登记数据整合建库研究[J]. 测绘通报,2015(增刊):7-10

[6] 范高兴. 不动产数据整合系统的设计与实现[J]. 中小企业管理与科技(下旬刊),2020(1):94-95

[7] 王化娟. 不动产登记存量数据整合系统的设计与实现[J]. 测绘与空间地理信息,2019,42(8):88-90

[8] 贾胜韬,李涛,李治明,等. 不动产登记存量数据整合系统设计与实现:以金昌市为例[J]. 测绘,2020,43(5):228-233

[9] 中华人民共和国国土资源部. 不动产登记数据整合建库技术规范(试行)[S].