

徐晓庆,张智,卓凤艳,等. 宁夏气象资料自动归档管理系统设计与实现[J]. 中低纬山地气象,2023,47(5): 95 - 101.

宁夏气象资料自动归档管理系统设计与实现

徐晓庆^{1,2},张 智^{1,2},卓凤艳^{1,2},王添男^{1,2}

(1. 中国气象局旱区特色农业气象灾害监测预警与风险管理重点实验室,宁夏 银川 750002;
2. 宁夏气象防灾减灾重点实验室,宁夏 银川 750002)

摘 要:气象探测资料种类多、数据量大,早期宁夏实时气象探测资料主要存储在各类服务器中,如何便捷高效地使用气象探测资料是困扰业务人员的难题。该文分析了宁夏气象资料归档业务需求,制定气象资料自动归档策略;介绍了系统的总体架构和各功能模块的设计思路;采用 MySQL、NAS 和 Bootstrap 等关键技术,设计开发“宁夏气象资料自动归档管理系统”,并统计宁夏气象资料的自动归档情况,展示资料查询下载效果。经试运行,该系统可以实时自动归档宁夏气象探测资料和基础预报资料,能够查询下载数值模式产品、雷达数据和天气图等归档资料。

关键词:气象资料;MySQL;NAS;自动归档;管理

中图分类号:TP311.521 **文献标识码:**B

Design and Implementation of Ningxia Meteorological Data Automatic Archiving Management System

XU Xiaoqing^{1,2},ZHANG Zhi^{1,2},ZHUO Fengyan^{1,2},WANG Tiannan^{1,2}

(1. Key Laboratory for Meteorological Disaster Monitoring and Early Warning and Risk Management of Characteristic Agriculture in Arid Regions, CMA, Yinchuan 750002, China;
2. Ningxia Key Lab of Meteorological Disaster Prevention and Reduction, Yinchuan 750002, China)

Abstract: There are many kinds and a large amount of meteorological observation data. The early real - time meteorological observation data were mainly stored in various servers. How to use meteorological observation data conveniently and efficiently is a difficult problem for business personnel. This paper analyze the requirements of archiving meteorological data, formulate the automatic archiving strategy of meteorological data expound the overall structure of system and explain the design ideas of various functional modules. Use the key technologies such as MySQL, NAS and Bootstrap, and designed and developed the "Ningxia meteorological data automatic archiving management system". The automatic archiving of meteorological observation data is counted, and the effect of data query and download is shown. After trial operation, the system can automatically archive Ningxia meteorological observation data and basic forecast data in real time, query and download numerical model products, radar data, weather charts and other archived data.

Key words: meteorological data; MySQL; NAS; automatic archiving; management

收稿日期:2022 - 09 - 06

第一作者简介:徐晓庆(1987—),女,硕士,工程师,主要从事气象信息处理技术研究,E - mail:519217609@qq.com。

通讯作者简介:张智(1967—),男,正高,主要从事气象信息处理技术研究,E - mail:ZZ_tt@126.com。

资助项目:中国气象局旱区特色农业气象灾害监测预警与风险管理重点实验室项目(CAMP - 201913)(CAMP - 202102)。

0 引言

气象资料是天气气候预测和规律研究的重要依据。随着我国气象探测设备现代化程度的不断提高,气象探测资料的种类和数量快速增长,信息化气象资料的存储及归档工作也越来越受到各级气象部门的重视。沈文海等^[1]设计了国家级气象资料存储检索系统,实现了结构化和非结构化气象资料的收集、存储、检索和备份等多项功能;王恕等^[2]、闫莽莽等^[3]设计了省级气象资料归档管理检索系统,以磁盘阵列、移动硬盘和 DVD 光盘为主要存储介质,实现了结构化和非结构化气象资料的自动存储与归档;滕水昌等^[4]设计了市级地面资料数据管理系统,实现了地面结构化气象资料的存储与管理;虎文珺等^[5]利用 VB 编程实现了天气雷达观测数据的自动归档,免除了人工整理、压缩、存储各类雷达观测资料的繁重低效工作;张恩红等^[6]建立了气象数据光盘归档系统,以蓝光光盘为存储介质,实现了省级气象资料的不定时归档。

受存储条件的限制,早期宁夏实时气象探测资料主要存储在各类服务器中,其中非结构化气象资

料只存储 3—6 个月。宁夏气象档案馆仅能归档存储各类月、年报表等结构化数据文件,无法考虑实时报文等结构化、雷达等非结构化气象资料的归档存储问题。2020 年宁夏气象信息中心建成 1PB 容量的存储服务器(简称 NAS),为宁夏非结构化气象资料的归档提供了存储条件。

本文在分析宁夏气象资料归档流程的基础上,设计开发了“宁夏气象资料自动归档管理系统”(简称“归档系统”),能够实现气象资料自动归档、网页查询下载、归档率统计等功能,提高了气象资料业务的工作效率。

1 归档资料范围

按照《气象资料分类与编码》^[7]中气象资料分类规定,归档系统将宁夏行政区域内地面、高空、辐射、农气、大气成分、雷达等气象观测数据列入自动归档范围;经征询省级天气气候业务单位需求,选定天气预报核心业务系统所需基础数值模式产品、天气图等资料也列入自动归档范围,共计 7 个大类、22 种(表 1)。

表 1 自动归档气象资料种类表
Tab.1 Automatic archiving of meteorological data type table

大类	归档资料名称	文件特征	存储频次
地面	国家级站逐小时观测资料(BUFR 格式)	结构化	月
	区域站逐小时观测资料(BUFR 格式)	结构化	日
	公路交通站逐小时观测资料	结构化	月
	地面天气图	非结构化	日
高空	高空观测数据文件(BUFR 格式)(原始地面—500 hPa)	结构化	月
	高空观测数据文件(BUFR 格式)(原始地面—100 hPa)	结构化	月
	高空观测数据文件(BUFR 格式)(原始全部)	结构化	月
	高空观测数据文件(BUFR 格式)(质控地面—500 hPa)	结构化	月
	高空观测数据文件(BUFR 格式)(质控地面—100 hPa)	结构化	月
	高空观测数据文件(BUFR 格式)(质控全部)	结构化	月
	高空探测和状态参数文件(XML 格式)	结构化	月
	高空观测基数据文件(S 文件)	结构化	月
	高空天气图	非结构化	日
辐射	逐小时观测资料(BUFR 格式)	结构化	月
农气	自动土壤水分站逐小时观测资料	结构化	月
数值预报	欧洲中心(ECMWF)模式产品	非结构化	月
	宁夏 WRF 产品	非结构化	月
大气成分	气溶胶数浓度(NSD)观测资料	结构化	月
	气溶胶质量浓度(PMMUL)观测资料	结构化	月
雷达	标准格式基数据	非结构化	日
	逐仰角标准格式基数据	非结构化	日
	PUP 产品	非结构化	日

2 系统设计

2.1 总体架构

归档系统以数据收发业务监控系统(简称 CTS)和气象信息综合分析处理系统(简称 MICAPS)为数据源,以虚拟化服务器为硬件支撑,采用 B/S 体系架构,Web 网站采用 Windows + IIS + ASP.NET 架构,资料归档程序采用 Linux + Shell 的结构,数据库采用 Linux + MySQL 数据库语言,服务器端包括 Web 服务器、归档服务器和数据库服务器,客户端

基于 Web 可视化技术采用浏览器方式实现归档气象资料查询下载和资料归档率统计结果展示^[8,9]。

归档系统总体架构在数据资源标准化体系的规范下和基础设施的支持下设计,采用 3 层体系架构:服务层、处理层和数据层^[10,11](图 1)。其中服务层为标准的 Web 客户端,负责与用户交互,并通过 Bootstrap 表格方式实现各类归档气象资料的查询展示;处理层通过定时任务、节目表实现气象资料的自动归档和日志信息入本地数据库;数据层是数据的来源,包括 CTS 和 MICAPS。

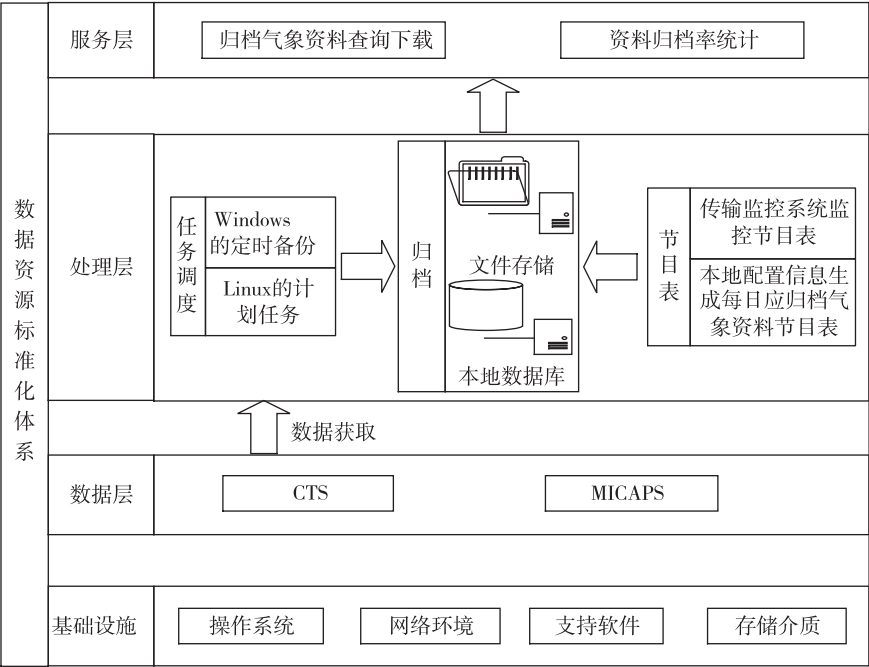


图 1 总体架构图

Fig. 1 Overall structure picture

2.2 功能设计

归档系统功能设计主要包括处理层的气象资料自动归档和服务层的归档气象资料查询 2 大模块,其中气象资料自动归档模块包括实时气象资料归档、预报业务资料归档、归档资料日志入库、归档资料日志信息统计等 4 个子模块;归档气象资料查询模块包括归档气象资料查询下载、资料归档率统计等 2 个子模块。

2.2.1 自动归档模块设计 自动归档策略(图 2):实时气象资料的归档程序根据配置信息每日定时从 CTS 下载前 1 d 的气象资料存入 NAS 盘阵;预报业务资料的归档程序根据配置信息每日定时从 MICAPS 数据库下载前 1 d 的预报资料存入 NAS 盘阵;归档资料日志入库程序遍历 NAS 盘阵的归档目录,并将归档信息存入归档文件日志表。归档资料

日志信息统计程序每日定时生成当日节目表,根据归档文件日志表统计前 1 d 归档资料入库信息,并存入归档文件信息统计表。具体功能设计如下:

(1)实时气象资料归档。归档系统从 CTS 系统获取各类实时气象资料。CTS 系统具备资料推送功能,但推送失败并达到一定程度会影响业务运行。为减少对 CTS 系统的影响,归档系统采用 Shell 脚本编写自动归档程序,定时从 CTS 系统通过 FTP 方式采集气象资料。自动归档程序部署在 Linux 系统中,该计算机挂载气象资料归档存储的 NAS 盘阵,通过基于 Linux 计划任务(crontab)实现数据的自动下载。

(2)预报业务资料归档。用于 MICAPS^[12]产品数据文件的冷备份业务,也可用于近实时预报业务资料的备份业务。预报业务资料归档程序部署在

Windows 计算机中。该计算机挂载气象资料归档存储的 NAS 盘阵,采用每日定时备份方式从 MICAPS 分布式数据库系统中获取欧洲中心模式产品 (ECMWF)、地面天气图和高空天气图数据存储至 NAS 盘阵。

(3) 归档资料日志入库。用于实现从归档气象资料存储 NAS 盘阵中列出归档资料,生成归档文件日志并写入归档文件日志表。归档气象资料日志

入库程序遍历归档气象资料存储 NAS 盘阵,列出已归档文件,并根据归档资料配置信息表获取归档文件日志信息,写入归档文件日志表。归档气象资料日志入库程序采用 Shell 脚本编写,部署在 Linux 系统中。该计算机挂载气象资料归档存储的 NAS 盘阵,通过计划任务 (crontab) 实现每日定时归档资料日志入库。

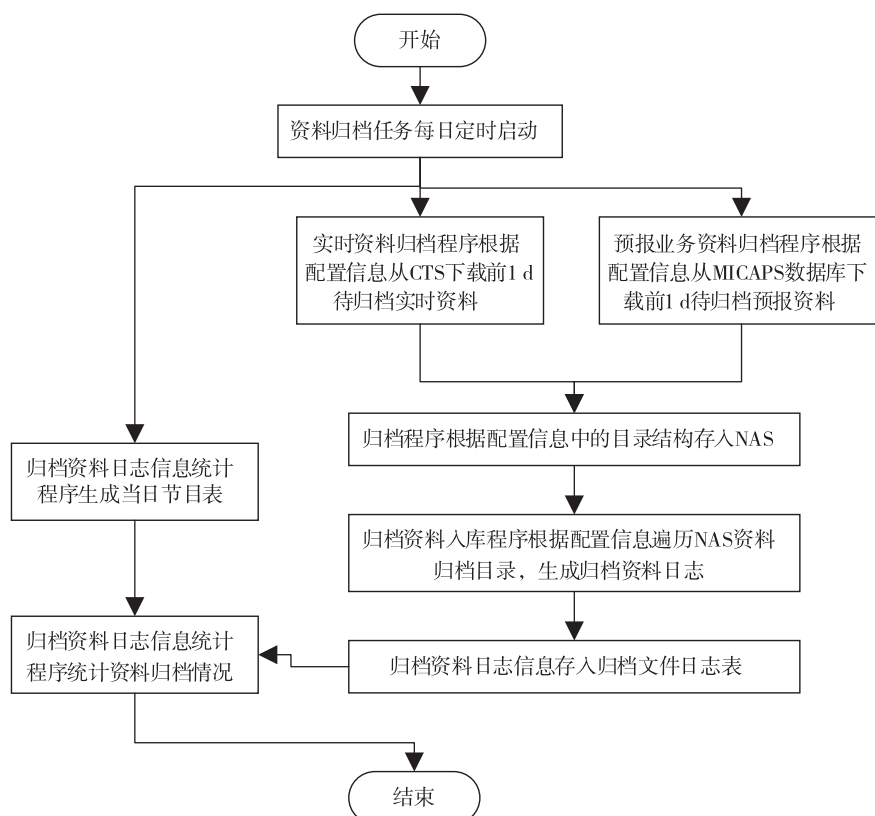


图 2 归档策略

Fig. 2 Archive policy

(4) 归档资料日志信息统计。用于实现生成每日应归档气象资料数量,统计每日实际归档气象资料数量。归档日志信息统计程序根据传输监控系统监控节目表或本地配置信息生成每日应归档气象资料节目表,并从归档文件日志表中统计实际归档气象资料数量,写入归档文件信息统计表。

2.2.2 查询下载模块设计 目前宁夏气象部门结构化数据在大数据云平台入库后可长时间存储并提供数据服务,而非结构化数据在大数据云平台一般存储 3~6 个月。因此归档系统的归档气象资料查询下载只针对非结构化数据,主要包括模式产品 (ECMWF、WRF)、雷达 (全仰角、逐仰角、PUP 产品) 和天气图 (地面、高空) 等 3 大类。模式预报和天气图资料按时次查询,雷达资料按时次、站号查询,查询后以表格形式显示,并可从远程服务器实现归档

气象资料下载。

资料查询的设计:选择查询条件后,利用 ajax 方法将查询条件传入后台,从 MySQL 数据库归档文件日志表中查询资料的文件名、路径、文件大小、观测时间等信息,通过 json 串将查询的数据传入网页,并通过 Bootstrap 表格的形式展示。

资料下载的设计:根据从 MySQL 数据库归档文件日志表中查询资料的文件名、路径,通过 ftp 方式从远程服务器上逐个下载文件,判断远程服务器上是否存在下载的文件,若存在则将文件依次下载至程序所在共享目录 share 下,若不存在则继续循环判断下个文件是否存在。判断下载的资料是否为 ECMWF 资料,若是则统计下载的文件大小 sum 值,判断 sum 值是否大于 2G,若是对下载的文件进行打包压缩;若下载的资料非 ECMWF 资料,则直接打包

压缩。将下载的文件名和路径组成 json 串传回至前台,前台根据文件名和路径从共享目录下载资

料。下载归档资料流程图见图 3。

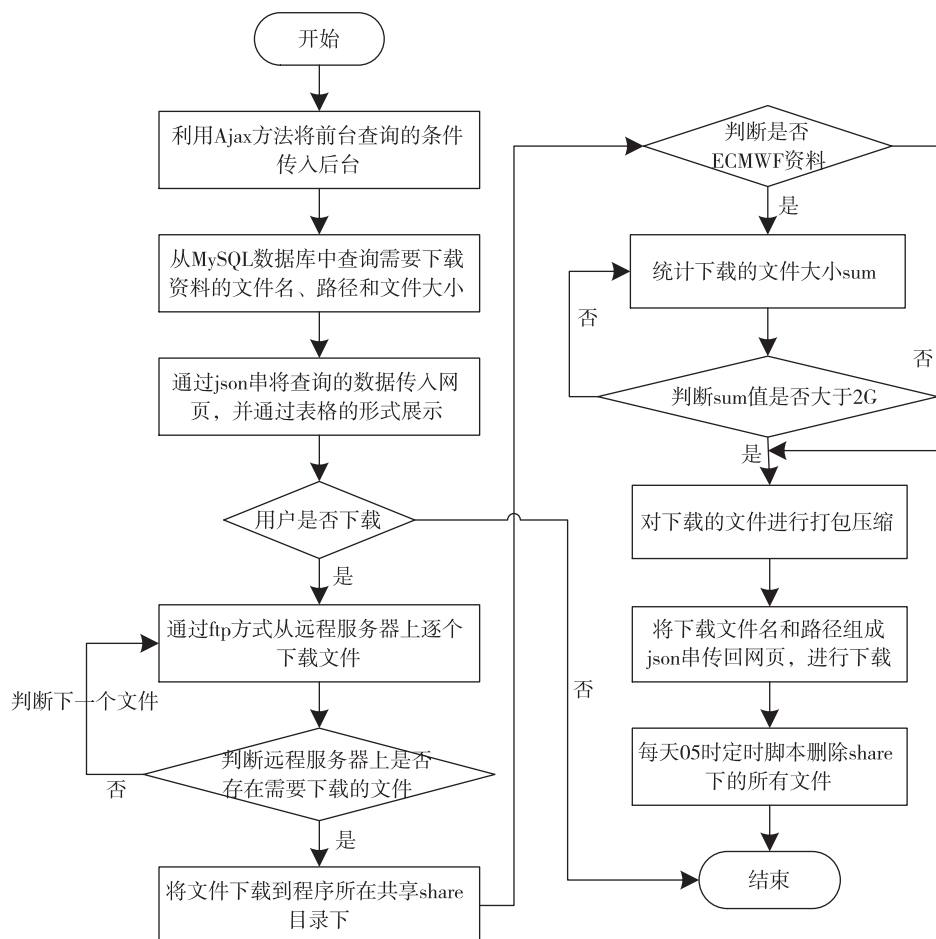


图3 下载归档资料流程图

Fig. 3 Download archived data flow chart

资料归档率统计。根据实际归档业务需求,实现 7 大类、22 种归档气象资料按不同时间段进行归档率统计。查询条件中的资料类型,资料种类,市、县、站信息均来自 MySQL 数据库归档文件信息统计表,根据该数据库的表信息的变动可以自动变动查询条件,并以表格的形式显示归档率统计情况。

3 关键技术

3.1 MySQL 数据库技术

MySQL 作为一种关系型数据库管理系统,所使用的 SQL 语言是用于访问数据库的最常用标准化语言,支持多个开发语言,特别是开放源码,被许多开发选择^[13,14]。本归档系统亦选用 MySQL 作为数据库管理软件。

归档系统数据库数据表包括:归档文件配置信息表、归档文件日志表、归档文件信息统计表、站号与行政编码信息对照表。其中归档文件配置信息

表(ARCHIVE_TYPE_INFO)用于存储归档系统配置信息,包括资料类型、是否下载、下载路径、归档路径等;归档文件日志表(ARCHIVE_FILE_LOG_YYYY)用于存储归档文件信息,包括文件名、文件大小、文件路径等;归档文件信息统计表(ARCHIVE_FILE_STAT)用于存储归档文件统计信息,包括资料类型、应归档文件/站点数、实际归档文件/站点数等;站号与行政编码信息对照表(ARCHIVE_FILE_PROGRAM)用于存储归档系统站号、行政编码对应关系,包括资料类型、应归档文件/站点数、实际归档文件/站点数等。

3.2 Shell 脚本技术

Shell 是一个命令行解释器,采用 C 语言编写,遵循一定的语法将输入命令加以解释并传给系统,是用户使用 Linux 的桥梁。Linux 服务器部署 Shell 脚本,结合 crontab 可实现脚本的定时高效执行,提高了运维工作效率^[15]。

归档系统实时气象资料归档程序和归档资料日志入库程序均采用 Shell 脚本编写,通过 crontab 定时执行脚本,实现资料定时下载和归档资料写入数据库。

3.3 NAS 网络存储技术

NAS 全称为 Network Attached Storage(网络附属存储),是通过网络访问一台存储服务器,其最大的特点是存储空间共享,并可根据存储容量需求进行扩展。

2020 年宁夏气象信息中心建成 1PB 容量的 NAS 存储,目前 NAS 资料存储情况(表 2),还有约 910 TB 空间可使用。归档气象资料存储空间需求大,不仅有结构化资料,还有非结构化资料,如模式预报 1 a 存储量约 10 TB。基于 NAS 的大容量和存储空间共享的优点,归档系统选用 NAS 作为存储设备,并可随时扩展归档存储空间。

表 2 NAS 资料存储情况表
Tab. 2 NAS data storage table

工作节点	存储容量/TB	已用容量/ TB	占比/%
store01	1024	110	10.74
store02	50	3.3	6.6
store03	20	4.2	21

表 3 2—4 月气象资料自动归档情况统计表

Tab. 3 Statistical table of automatic archiving of meteorological data from February to April

资料类型	应归档数量/个	实际归档数量/个	实际存储量/MB	归档率/%
地面	2 113 752	2 109 650	9 799	99.81
高空	1 780	1 780	212	100.00
辐射	2 136	1 068	1.45	50.00
农气	79 032	79 032	70	100.00
模式预报	10 858	11 837	2 987 008	100.00
大气成分	6 408	6 408	40.3	100.00
雷达	524 350	5 787 798	66 464.2	100.00
合计	2 738 316	7 997 573	3 063 594.95	—

4.2 查询下载

以前气象资料一般利用 FTP 从服务器共享目录内获取,极大地限制了气象资料的共享性。归档系统基于 C#、jQuery、Bootstrap 和 json 等技术^[17-19]开发归档气象资料查询下载页面,通过浏览器即可在线查询下载归档气象资料。

根据界面选定的查询条件,可将气象资料查询、资料归档率统计以 Bootstrap 表格格式显示,使用表格右上角的搜索框可以快速查找用户需要的资料。如需下载,可自动将查询到的归档气象资料下载到本地谷歌浏览器的默认路径。

3.4 基于 Bootstrap 气象数据可视化技术

Bootstrap 是基于 HTML、CSS、JavaScript 开发的前端开发框架,提供了窗体、字体排印、导航、按钮等其它组件,使得 Web 开发更加快捷^[16]。

基于 Bootstrap 功能,归档系统采用基于 Bootstrap 的 B/S 开发技术,通过浏览器实现基于应用层 http 协议提供的 Web 服务,用户通过浏览器即可查看 Bootstrap 表格形式的归档气象资料。

4 应用实现

归档系统于 2022 年 1 月开始进行业务试运行,各模块均运行正常。

4.1 自动归档

选取 2022 年 2—4 月,统计自动归档程序归档气象资料的运行情况(表 3)。气象资料自动归档约 800 万个数据文件、2.9 TB 存储量,其中雷达资料数量最多约 580 万个,模式预报资料存储量最大约 2.8 TB。只有地面资料和辐射资料的归档率未达到 100%,其中地面资料是因逐小时资料中有个别站上传报文缺报导致,辐射资料是因 1 个站软件故障一直未上传报文导致。从归档率整体情况看,自动归档模块能够及时、完整地自动归档气象资料。

5 结论

根据宁夏现有气象探测资料的种类和天气预报核心业务系统所需基础预报资料,制定符合气象行业标准和省级气象资料存储管理规定的自动归档策略,设计开发“宁夏气象资料自动归档管理系统”,实现了宁夏气象资料自动归档、Web 查询下载和自动归档资料统计等功能,不仅为天气预报业务提供基础预报资料,还可为各级业务管理部门了解掌握属地气象探测资料的归档情况提供查询工具。同时系统具有可扩展性和可维护性,能适应今后新

增气象资料的自动归档业务需求。

参考文献

- [1] 沈文海,赵芳,高毕云,等. 国家级气象资料存储检索系统的建立[J]. 应用气象学报,2004,15(6):727-736.
- [2] 王恕,戴萍,刘晓梅,等. 实时气象资料归档与检索系统[J]. 气象与环境学报,2007,23(1):56-59.
- [3] 闫莽莽,刘莹,刘园园,等. 省级气象资料归档管理系统设计与实现[J]. 气象水文海洋仪器,2017,34(1):50-53.
- [4] 滕水昌,王生元,王荣喆. 基层台站地面气象资料数据管理检索系统[J]. 气象科技,2012,40(1):62-64.
- [5] 虎文琨,张芳. 利用VB编程实现天气雷达观测数据的自动归档[J]. 气象水文海洋仪器,2013,28(4):67-68.
- [6] 张恩红,张金标,李高洁,等. 气象数据光盘归档系统设计与实现[J]. 广东气象,2018,40(2):73-76,80.
- [7] 中国气象局. GB/T40153-2021 气象资料分类与编码[S]. 北京:气象出版社,2021.
- [8] 王文闻,王光明,黄祯珍,等. 怀化市 CIMISS 资料查询统计系统的设计与实现[J]. 中低纬山地气象,2021,45(2):77-81.
- [9] 陈关清,李逸,杨群,等. 贵州省气象科技信息共享系统的设计与实现[J]. 中低纬山地气象,2019,43(1):64-68.
- [10] 杨礼敏,胡平,王亚东. 上海气象信息全流程可视化监控系统的
- 设计与实现[J]. 气象科技进展,2019,9(3):29-35.
- [11] 孙钢灿,张宁宁,赵传勇. 智能井盖锁云端监测管理系统设计与实现[J]. 计算机应用与软件,2022,39(2):16-20.
- [12] 王若瞳,王建民,黄向东,等. MICAPS4 服务端系统架构设计[J]. 应用气象学报,2018,29(1):1-12.
- [13] 邱忠洋,雷正翠,刘文伟. 基于Web的气象项目管理系统的设计与实现[J]. 计算机技术与发展,2020,30(7):204-209.
- [14] BOUAMRANE K, MATALLAH H, BELALEM G. Comparative study between the MySQL relational database and the MongoDB NoSQL database [J]. International Journal of Software Science and Computational Intelligence(IJSSCI), 2021, 13(3):38-63.
- [15] 白铁男,唐维尧,谭海波,等. 贵州天气雷达集约监控系统的研究与实现[J]. 中低纬山地气象,2022,46(3):126-128.
- [16] 刘荣英. Bootstrap 前端开发(全案例微课版)[M]. 北京:清华大学出版社,2021.
- [17] 叶广仔,李卫华,刘晓蔚. 可拓传导知识挖掘系统的设计与实现[J]. 计算机应用与软件,2018,35(4):129-136.
- [18] 李新庆,卫建国,单新兰,等. 宁夏气象业务内网的功能与实现[J]. 干旱气象,2017,35(6):1077-1083.
- [19] 李艳萍,陶伟,陈为超,等. 国家级地面气象观测站设备监控系统设计与实现[J]. 气象水文海洋仪器,2020,37(4):62-65.