

ISSN 3105-7497 (print)
ISSN 3105-7500 (online)

数据科学与工程

Data Science & Engineering

2025年 第1卷 第1期 (双月刊)

CODEN: SKYGAY

(国际标准连续出版物标识符·全球唯一标识符)

分配机构: 美国化学文摘社 (CAS)

国图集团 CIBTC

进口备案刊号: G015Z119

QUEST PRESS LIMITED

www.SciOnline.com



QUEST PRESS



扫码阅读

敬畏自然，方能生生不息。

Revere nature, to achieve endless life.



地球是共同家园，守护需同心协力

The earth is a common home, protection requires joint efforts.

ISSN 3105-7497 (Print)

ISSN 3105-7500 (Online)

CODEN: SKYGAY

(国际标准连续出版物标识符 • 全球唯一标识符)

分配机构: 美国化学文摘社 (CAS)

中文名 数据科学与工程

英文名 Data Science & Engineering

主办 求索出版社

出版频率 双月刊

出版语言 中文, 英语

出版机构 求索出版社 (QPL Ltd.)

编辑单位 《数据科学与工程》编辑部

地址 澳门巴掌围斜巷19号7楼D

电话 +853 6881 9699

网址 sjkxygc.scionline2025.com

广告服务 QuestPress@hotmail.com

出版时间 2025年8月



QUEST
PRESS
LIMITED

版权所有

© 求索出版社 (QPL Ltd.)

版权与许可声明

本期刊整体版权归属求索出版社所有。所有发表文章均依据“知识共享署名 4.0 国际许可协议” (CC BY 4.0) 进行授权发布。

许可链接

<https://creativecommons.org/licenses/by/4.0/>

文章版权

单篇文章的著作权由作者与求索出版社共同享有。

使用授权

著作权人已同意依据CC BY 4.0协议授予文章使用许可, 使用者可在遵循协议条款 (主要为署名要求) 的前提下进行转载、改编或用于其他用途。

权责声明

1. 本刊所发表文章中的观点、数据和结论仅代表作者个人立场, 不代表本刊及出版机构的意见。
2. 对于因使用文章任何内容 (包括正文及引述材料) 而可能引发的任何直接或间接损失, 本刊及出版机构不承担任何法律责任。
3. 本刊的出版行为不构成对文中提及任何产品、方法或技术的商业性能、有效性或特定用途适用性的任何形式担保或推荐。
4. 本刊内容仅供学术与信息参考, 不替代专业判断。读者在依据文章内容做出任何决策前, 建议咨询相关领域专业人士。

编委会

主编 黄琳

编委 张明昭 锦州市数据中心

孙丽欣

石江宇

胡瑶

编委会助理 胡凯

Journal import

期刊进口

备案刊号: G015Z119



中國國際圖書貿易集團公司
China International Book Trading Corporation





DOI 数字对象唯一标识符 (英国)



Crossref 交叉引用 (美国)



ICI 哥白尼期刊数据库 (波兰)



EuroPub 欧洲学术出版中心数据库 (英国)



ACADÉMIA. EDU 学术界 (美国)



J-Gate 开放获取期刊门户 (印度)



CNKI 中国知网 (中国)



WanFang 万方数据 (中国)



CVIP 维普期刊 (中国)



RCCSE 中国学术期刊目录 (中国·武汉大学)



Baidu Scholar 百度学术 (中国)



CJWK 长江文库 (中国)



591adb 中邮阅读·万章期刊 (中国)



OpenAlex 全球开源学术平台 (加拿大)



SCRIBD

SCRIBD 在线图书馆 (美国)



MAGZTER
DIGITAL NEWSSTAND
TAP • READ • ENJOY

Magzter 跨平台数字期刊 (美国)



SlideShare 知识共享平台 (美国)



EBSCO 学术数据库 (美国)



RePEc

RePEc 经济研究文献数据库 (美国)



 **OUCI**

OUCI 开放乌克兰科学 (乌克兰)



 **Discovery**
by Editage

R Discovery 研究者发现平台 (印度)



 **LENS.ORG**

The Lens 透镜学术 (澳大利亚)





Google Scholar 谷歌学术 (美国)



 **Microsoft Bing**

Microsoft Bing 微软学术 (美国)



 **ResearchGate**

ResearchGate 研究之门 (德国)





Portico 学术文献保存库 (美国)





ASCI 亚洲科学引文索引 (美国)



 **SEMANTIC SCHOLAR**

Semantic Scholar 语义学术 (美国)



ISSN (法国)



BOOKAN 博看 (中国)



Issuu 数字出版平台 (丹麦)



WorldCat 世界编目数据库 (美国)



SPI-Hub 范德比尔特大学 (美国)



MDPI scilit 科学与学术数据库 (瑞士)



ProQuest 科睿唯安数据库 (美国)



Research Commons 科睿唯安·研究共享平台 (美国)



Scite (美国)



BASE 比勒费尔德学术搜索 (德国)



ERIH+ 欧洲人文社会科学索引 (挪威)



Yahoo!Search (美国)



Naver 搜遍 (韩国)



LivRe! 开放期刊门户 (巴西)



ResearchBib 研究者索引 (日本)



KINDCONGRESS

KIND CONGRESS (阿塞拜疆)



TOGETHER WE REACH THE GOAL
SJIF 科学期刊影响因子 (印度)



IP Indexing IP索引



Journament

Journament 质量评估与索引



RJIF 研究期刊影响因子



Road 开放科学学术资源路线图 (法国)



COSMOS 期刊评价指标 (印度)



CiteFactor

CiteFactor 指标与信任索引 (印度)



ADL 亚洲数字图书馆 (巴基斯坦)



Secret Search Engine Labs

Secret 搜索引擎实验室 (芬兰)



academindex

Academindex 学术索引 (土耳其)



ROOTINDEXING
Discover. Index. Connect. Impact.

RootIndexing 根索引与影响因子



EZB Electronic Journals Library

EZB 电子期刊图书馆 (德国)



ISI 国际科学索引平台（阿联酋迪拜）



Sci Online 科学在线（中国澳门）



Baidu Baike 百度百科（中国）



OALib 开放存取资源图书馆（美国）



ESJI 欧亚科学期刊索引（哈萨克斯坦）

数据科学与工程

Data Science & Engineering

2025年10月 第1卷 第1期（双月刊）

目次

◆ 人力资源管理信息化

大数据时代下人力资源信息化的挑战与机遇

Challenges and Opportunities of Human Resource Informatization in the Era of Big Data

汪大鵬..... (1)

基于云计算的电子行业人力资源管理信息系统构建

Construction of a cloud-based human resource management information system for the electronics industry

温 泉..... (7)

◆ 智慧城市与公共安全

面向智慧城市的电子监控数据安全治理框架设计

Design of a framework for electronic surveillance data security governance in smart cities

魏 博..... (12)

智慧交通系统中多源异构大数据融合处理方法探析

Exploration of multi-source heterogeneous big data fusion processing methods in intelligent transportation systems

张明昭..... (18)

基于云计算的智慧政务服务平台设计

Design of a cloud-based smart government service platform

赵 涵..... (26)

◆ 通用大数据分析处理

电子信息系统中大数据分析处理策略

Big data analysis and processing strategies in electronic information systems

张 旭..... (32)

大数据时代下人力资源信息化的挑战与机遇

汪大鹏

锦州市凌河区人力资源和社会保障服务中心, 辽宁锦州, 121000

摘要: 随着大数据技术的迅猛发展, 人力资源管理正经历深刻变革, 信息化建设成为组织提升管理效能的关键路径。然而, 在实践过程中, 许多企业面临数据质量参差不齐、系统集成度低、隐私安全风险凸显以及专业分析人才匮乏等现实挑战, 制约了人力资源数据价值的充分挖掘与应用。本研究聚焦于大数据时代背景下, 如何有效应对这些挑战并把握信息化带来的机遇。通过分析当前企业人力资源信息化的现状与瓶颈, 探讨了数据治理框架的构建、智能分析工具的引入以及复合型人才培养等关键对策。

关键词: 大数据, 人力资源管理, 信息化

Challenges and Opportunities of Human Resource Informatization in the Era of Big Data

Dapeng Wang

Human Resources and Social Security Service Center, Linghe District, Jinzhou City,
Jinzhou Liaoning 121000, China

Abstract: With the rapid development of big data technology, human resource management is undergoing profound changes, and informatization construction has become a key path for organizations to enhance management efficiency. However, in practical processes, many enterprises face real challenges such as uneven data quality, low system integration, prominent privacy and security risks, and a lack of professional analytical talent, which restrict the full exploration and application of human resource data value. This study focuses on how to effectively address these challenges and seize the opportunities brought by informatization in the context of the big data era. By analyzing the current status and bottlenecks of enterprise human resource informatization, this study discusses key countermeasures such as the construction of a data governance framework, the introduction of intelligent analysis tools, and the cultivation of composite talents.

Keywords: big data, human resource management, informatization

1 大数据时代人力资源信息化的内涵与演进

1.1 人力资源信息化的核心概念界定

所谓人力资源信息化, 就是组织用现代信息技术, 把人力资源管理里的各类流程、数据和决策, 做系统性的整合与优化。它的核心组成部分

有信息技术基础设施、人力资源数据资源、业务流程应用和相关管理规范。信息技术基础设施是整个系统的支撑, 能给数据采集、存储与处理提供对应的硬件和软件平台。人力资源数据资源是核心资产, 包含员工从进入组织到离开的全阶段信息。把技术和管理实操结合起来的是业务流程



Copyright © 2025 by author(s) and Quest Press Limited. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<https://creativecommons.org/licenses/by/4.0/>



应用,能让招聘、培训、绩效、薪酬等工作实现线上化与自动化。管理规范则能让整个系统保持有序、合规的运行状态^[1]。

组织管理里,人力资源信息化是一项基础工作,它把原本分散、靠人工完成的人力资源管理事项,转成在统一信息系统里的标准化操作,让管理效率和准确性得到提升,还能把各类人力资源数据沉淀下来,完成整合与分析,为管理模式从凭经验判断转到靠数据做决策打牢基础,这不正是操作工具的更新,更是管理理念和模式的调整,推动人力资源部门从原来的行政事务处理方,变成能支撑组织战略的业务协作角色^[2]。

1.2 大数据技术对人力资源管理的驱动作用

大数据技术转变了数据采集、处理与分析的模式,推进人力资源管理模式发生改变。物联网传感器、社交平台及内部业务系统等多种渠道,能实时、自动捕获员工行为、绩效表现、市场趋势等海量结构化与非结构化数据,让传统人事数据的覆盖范围和时效性都有了提升。分布式存储与云计算平台具备弹性可扩展的承载能力,可完成TB乃至PB级人力资源数据的集中存储与高效管理^[3]。

分析模式的转变比较大。大数据技术依托机器学习、自然语言处理等算法,能从海量杂乱的数据里找出暗藏的关联与规律,完成描述性分析到预测性、规范性分析的转变。比如,整理过往员工离职数据和多维度员工特征,可搭建模型预判离职风险并找出核心原因,扫描大量简历与社交媒体内容,能更准确地判断人才和岗位的适配程度。这种深度分析的能力,推进人力资源决策从凭经验直觉转向靠数据支撑的科学决策,给人才招聘、保留、发展与激励等核心工作带来了全新的参考依据。

1.3 从电子化到智能化:人力资源信息化的发展阶段

人力资源领域的信息化进程,完整体现出技术迭代和管理需求升级的两条脉络。它起步时主要做“电子化”改造,把纸质人事档案、考勤记

录、薪酬核算等基础人事工作,转用计算机系统自动处理,目的是提高操作效率和准确性,给人力资源从业者从大量重复劳动里减负。数据库与网络技术成熟后,信息化进入“系统化”阶段,招聘、绩效、培训等各类独立的人力资源管理模块相继出现,还尝试实现集成,能让跨职能的数据流转和流程协同成为现实,管理重点转向优化流程和开展初步的数据汇总分析。

2010年后大数据技术普及,人力资源信息化发展进入“智能化”新阶段,核心是把大数据、人工智能与云计算等技术用透,既要完成全流程、全触点数据的无缝采集与整合,还要让系统具备预测、诊断和自主决策的能力。智能招聘系统能主动挖掘潜在候选人并完成评估,员工流失预测模型可提前预警风险并给出干预建议,信息化系统不再只是处理事务的操作工具,而是能支撑战略决策的智慧中枢,推进人力资源管理从事务导向、职能导向,完全转向数据驱动与战略价值创造导向。

2 大数据时代人力资源信息化面临的主要挑战

2.1 数据质量与系统集成困境

在大数据时代,人力资源信息化面临的首要挑战在于数据质量与系统集成的双重困境。企业内部人力资源数据来源多样,涵盖招聘系统、绩效管理、薪酬福利及员工关系等多个模块,这些数据在格式、标准与更新频率上往往存在显著差异。缺乏统一的数据治理标准导致数据冗余、不一致甚至错误,严重影响了后续分析的准确性与可靠性。这种“脏数据”问题使得基于大数据的人力资源决策如同建立在流沙之上,难以发挥其应有的战略价值。

与此同时,许多组织内部存在多个独立运行的人力资源信息系统,形成了“信息孤岛”。各系统间接口不兼容、数据难以互通,阻碍了人力资源信息的有效整合与全局视图的构建。系统集成的高成本与技术复杂性进一步加剧了这一问题,使得企业无法对员工全生命周期数据进行连

贯分析，从而限制了大数据在人才盘点、流失预测与组织效能评估等深层应用中的潜力。

2.2 数据安全与员工隐私保护风险

大数据应用在显著提升人力资源管理洞察力的同时，也带来了严峻的数据安全与员工隐私保护风险。人力资源系统汇聚了员工身份、薪酬、健康、绩效乃至社交网络等大量敏感个人信息，一旦发生泄露或被滥用，不仅会直接侵害员工权益，还可能引发法律诉讼与声誉危机。随着《个人信息保护法》等法规的出台，企业在收集、存储、处理员工数据时面临的合规性要求日益严格，违规成本高昂。

大数据分析往往涉及对员工行为的深度挖掘与关联分析，这触及了管理效率与个人隐私之间的伦理边界。例如，通过分析员工邮件通信、门禁记录或网络行为来评估其状态，可能引发员工对监控的抵触与信任危机。企业必须在利用数据价值与尊重员工隐私之间找到平衡，建立透明、合规的数据使用政策与安全防护体系，以应对大数据时代特有的安全与伦理挑战。

2.3 专业数据分析人才短缺与思维转型滞后

大数据技术在人力资源领域的深入应用，对从业者的能力结构提出了全新要求。当前，具备数据科学、统计学知识与人力资源管理实践经验的复合型人才极为匮乏。多数人力资源管理者熟悉传统模块业务，但缺乏数据清洗、建模分析与

可视化解读的专业技能，难以将海量数据转化为有价值的业务洞察。这种人才断层直接制约了数据分析的深度与决策支持的有效性。

与此同时，传统人力资源管理思维向数据驱动决策的转型面临显著困难。许多管理者仍依赖经验直觉进行判断，对数据分析和量化模型的价值认识不足，或存在“数据恐惧”心理。组织内部尚未形成基于数据对话、用事实决策的文化氛围，导致先进的分析工具与业务管理“两张皮”，数据分析成果难以落地为具体的优化行动，限制了大数据赋能潜力的充分发挥。

3 大数据赋能人力资源管理的战略机遇

3.1 数据驱动的人力资源决策科学化

大数据分析为人力资源决策科学化提供了前所未有的可能性。通过整合内外部多源数据，组织能够构建更精准的决策模型。在人才招聘环节，分析历史招聘数据、岗位要求与市场人才库信息，可以建立人才画像与岗位匹配模型，显著提升人岗匹配度与招聘效率。绩效评估不再局限于主观评价，而是结合员工的项目产出、协作网络、技能发展等多维度行为数据，实现更全面、客观的评估。此外，利用机器学习算法分析员工行为模式、满意度调查与流动历史数据，能够有效预测潜在离职风险，使管理者得以提前进行干预，保留关键人才。见表1。

表1 决策场景

决策场景	传统方式局限	大数据驱动方法	预期效益
人才招聘	依赖简历与面试，主观性强	基于历史成功员工工数据建模，进行简历与岗位智能匹配	提升匹配精度，缩短招聘周期
绩效评估	周期性、主观评分，反馈滞后	实时整合项目、协作、学习等多源数据进行持续评估	实现动态、全面、客观的评估
离职预测	事后分析，被动应对	分析行为、满意度、薪酬等数据，构建预测模型	提前预警，主动实施保留措施

3.2 智能化工具提升人力资源运营效率

人工智能与机器学习技术正深度融入人力资源管理流程，显著提升了运营效率与员工体验。在自动化流程方面，RPA（机器人流程自动化）

可高效处理薪酬核算、社保申报、考勤统计等重复性高、规则明确的事务，将人力资源从业者从繁琐的行政工作中解放出来。机器学习算法则能优化简历筛选、面试安排等环节，通过持续学习

历史成功案例，不断提升筛选的精准度与速度。

在员工服务与互动优化层面，AI驱动的智能聊天机器人（如HR虚拟助手）能够7x24小时响应员工关于政策、福利、流程的常见咨询，提供即时、准确的解答，极大改善了服务响应速度与可及性。此外，基于员工行为与反馈数据的分析，机器学习模型可以个性化推荐培训课程、职业发展路径或内部岗位机会，从而增强员工互动体验，促进其成长与留存。这些智能化工具的应用，正推动人力资源运营从被动响应向主动服务、从标准化处理向个性化关怀转变。

3.3 构建战略人力资源分析与预测模型

大数据技术为构建战略人力资源分析与预测模型提供了可能，使人力资源管理从被动响应转向主动规划。通过整合内外部多源数据，组织能够构建动态、多维的“人才画像”，精准刻画员工的能力、潜力、行为模式及职业倾向。这为关键岗位继任计划、高潜人才识别与个性化发展提供了数据基础。在劳动力规划方面，可基于历史数据、行业趋势与业务战略目标，建立预测模型，模拟不同业务场景下的人才供需、成本结构及技能缺口，支持前瞻性的人才招聘、培养与配置决策。

这些模型将人力资源数据转化为战略资产，直接服务于组织的长期发展。例如，通过分析员工流动率与关键驱动因素，预测模型能预警离职风险，指导保留策略；结合市场数据与内部技能库的模型，可预测未来所需的新兴技能，驱动培训体系变革。最终，战略人力资源分析模型助力组织将人才管理深度嵌入业务战略，实现人才与战略的动态协同。

4 推进人力资源信息化建设的核心对策

4.1 建立统一的数据治理与标准化框架

为应对数据质量与系统集成困境，建立统一的数据治理与标准化框架是基础性工作。该框架旨在通过制定企业级的数据标准、实施主数据管理（MDM）以及明确数据所有权与流程，系统性提升人力资源数据的准确性、一致性与可用性。数据标准需覆盖员工主数据、组织架构、考勤绩效等核心领域，定义统一的字段格式、编码规则与业务含义，为跨系统数据交换奠定基础。主数据管理则聚焦于创建并维护单一、权威的核心数据源，如员工唯一标识，以消除数据冗余与冲突。这一治理体系能有效打破信息孤岛，确保分析所用数据的可靠性与完整性，为后续的数据分析与决策提供高质量“原料”。见表2。

表2 治理领域

治理领域	核心措施	预期成效
数据标准制定	定义员工、组织、岗位等主数据的字段标准、编码规则与业务口径。	实现跨系统数据语义一致，降低整合成本。
主数据管理 (MDM)	建立员工主数据等核心数据的单一可信来源，并建立维护流程。	消除数据冗余与矛盾，确保核心数据准确、唯一。
数据质量管理	建立数据质量评估指标（如完整性、准确性、及时性），并实施定期监控与清洗。	持续提升数据可信度，为分析决策提供可靠基础。
数据安全与隐私	在治理框架内明确敏感数据分类、访问权限控制及合规使用规范。	在保障数据安全与员工隐私的前提下，促进数据的合规共享与利用。

4.2 推动系统整合与平台化建设

在统一数据治理框架的基础上，推动系统整合与平台化建设是打破信息孤岛、实现数据互联互通的关键技术路径。企业需通过应用编程接口、中间件或企业服务总线等技术手段，将分散的人力资源管理系统、财务系统、业务运营系统

等进行有效集成，构建一体化的数据共享与业务协同平台。平台化部署能够将核心人力资源功能模块化，并以服务的形式提供，支持灵活扩展与按需调用。这不仅消除了数据在不同系统间手动迁移的误差与滞后，更实现了员工全生命周期数据的实时汇聚与动态更新，为后续的深度分析与

智能应用提供了稳定、流畅的数据管道。见表3。

表3 推动系统整合与平台化

整合层级	主要技术/方法	核心目标与功能
数据层整合	ETL 工具、数据仓库、数据湖	将异构系统数据抽取、转换、加载至统一数据存储，形成分析就绪的数据集。
应用层整合	API 接口、企业服务总线、微服务架构	实现不同应用系统间的业务功能调用与数据实时交互，支持流程自动化与协同。
平台层建设	一体化 HR 云平台、PaaS 服务	提供共享的 HR 服务能力（如组织、人事、薪酬核心），支持快速构建与集成新应用。
用户层呈现	门户集成、移动应用、数据分析仪表盘	为不同角色用户提供个性化的统一信息视图与决策支持界面，提升用户体验与效率。

通过上述分层整合与平台化建设，企业能够构建一个弹性、开放的人力资源信息化架构。该架构确保了数据流的畅通无阻，使人力资源部门能够基于完整、实时的数据全景图开展精准招聘、个性化培训、绩效洞察与离职预测等深度分析，真正释放大数据的管理价值。

4.3 强化数据安全防护与合规性管理

在技术层面，企业需构建纵深防御体系。这包括采用数据加密、脱敏技术保护静态与传输中的数据，利用访问控制与身份认证机制确保数据

最小权限访问，并通过安全审计与入侵检测系统实现全程可追溯。制度层面，必须建立并严格执行覆盖数据全生命周期的安全管理制度与操作规程，明确数据分类分级标准及对应的保护要求，同时定期开展合规性审计与风险评估。伦理层面，则需倡导“隐私设计”理念，在系统开发初期即嵌入隐私保护考量，并通过培训提升全员的数据安全与隐私保护意识，平衡数据应用与个人权益。见表4。

表4 防护层面

防护层面	核心措施与工具示例	主要目标
技术防护	数据加密（传输/存储）、脱敏/匿名化、访问控制（RBAC/ABAC）、安全审计日志、DLP 系统	防止数据泄露、篡改、丢失，确保数据机密性、完整性与可用性。
制度保障	数据安全管理办法、隐私政策、数据分类分级标准、应急响应预案、供应商安全管理协议	明确责任主体与操作规范，满足《个人信息保护法》等法律法规的合规性要求。
伦理构建	隐私影响评估、隐私设计、员工数据伦理培训、透明化数据使用告知与同意机制	将隐私保护内化为组织文化，赢得员工信任，实现负责任的数据创新。

5 面向未来的人力资源信息化发展路径

5.1 培养复合型人力资源数据分析团队

为应对大数据时代人力资源数据分析人才短缺的挑战，企业需着力培养兼具业务洞察与技术能力的复合型团队。这要求打破传统人力资源管理与分析技术之间的壁垒，通过系统化的内部培训与精准的外部引进双轨并行。内部培训应聚焦于提升现有人力资源从业者的数据素养，使

其掌握基础的数据处理与分析工具应用，理解数据分析在招聘、绩效、培训等核心模块的价值逻辑。同时，需从外部引进具备统计学、计算机科学背景的数据分析专家，他们能够运用机器学习等高级分析技术挖掘数据深层价值。内部业务专家与外部技术专家的深度融合，是构建高效分析团队的关键。这种复合型团队不仅能够操作智能化工具，更能将数据分析结果转化为可执行的人力资源战略，真正实现数据驱动决策。见表5。

表5 人才类型

人才类型	核心能力要求	主要培养/引进途径
业务导向型分析人才	熟悉人力资源管理全流程；具备数据敏感性与基础分析能力；能将业务问题转化为分析课题。	内部选拔与培训：开设数据分析系列课程、组织跨部门项目实践、设立数据分析师职业发展通道。
技术专精型分析人才	精通数据分析工具（如Python、R、SQL）；掌握统计学与机器学习算法；具备数据建模与预测能力。	外部引进与校企合作：招聘数据科学专业毕业生；与高校、科研机构合作建立实习与培养基地。
战略整合型分析人才	兼具业务理解与技术能力；擅长沟通与跨部门协作；能将数据分析洞察提升至战略规划层面。	内部轮岗与项目历练：安排技术人才深入业务部门实践；选拔业务骨干参与重大数据分析项目。

5.2 发展预测性与规范性分析应用

在预测性分析的基础上，人力资源信息化将向规范性分析深度演进。未来的系统不仅能基于历史数据预测员工流失风险或绩效趋势，更能结合组织战略目标与约束条件，生成具体的优化行动建议。例如，系统可模拟不同激励方案对关键人才保留率的影响，或为业务部门扩张提供包含成本、技能匹配度、招聘渠道效果在内的综合人力配置方案。这要求分析模型从“发生了什么”、“为何发生”的层面，提升至“应该怎么做”的决策支持高度。

规范性分析的应用依赖于更复杂的算法模型、实时数据流以及业务规则的深度集成。它将人力资源数据分析从后台的洞察报告，转变为前台业务决策的智能参谋，直接驱动精准的人力资源干预措施，实现从被动响应到主动规划的战略跨越。

5.3 构建敏捷迭代与持续优化的信息化生态

为应对技术的快速迭代与业务需求的动态变化，构建一个敏捷、可迭代并持续优化的人力资源信息化生态至关重要。这一生态强调机制与文化的双重建设，其核心在于建立一套能够快速响应内外部变化、持续评估信息化投资回报并驱动系统与流程改进的闭环管理体系。这意味着人力资源信息系统本身需具备模块化、可配置的特性，支持新功能或数据分析模型的快速部署与测试。同时，需建立常态化的成效评估机制，通过关键绩效指标监控数据应用、流程效率及业务影

响的实际情况，并基于反馈进行迭代优化。

更深层次地，这要求培育一种鼓励实验、容忍试错并基于数据事实进行决策的组织文化。人力资源部门与技术团队、业务部门需形成紧密协作的伙伴关系，共同参与从需求定义到效果评估的全过程。通过设立跨职能的敏捷小组，采用小步快跑、持续交付的方式推进信息化项目，能够确保技术解决方案始终与战略目标对齐，并最大化其业务价值。最终，这种生态将人力资源信息化从静态的项目实施，转变为一项动态的核心能力，为组织在不确定环境中保持人力资本管理的适应性与前瞻性提供坚实基础。

6 结语

研究发现，系统推进数据标准化与平台整合、发展预测性分析模型、并建立严格的数据安全与伦理规范，能够显著提升人力资源决策的科学性与战略性。这一探索不仅为企业在复杂数据环境中优化人才管理提供了实践参考，也对推动人力资源管理从传统事务型向战略赋能型转变具有重要的理论价值与实践意义。

参考文献

[1] 文衍杰. 大数据技术在事业单位财务会计内部控制中的应用研究[J]. 知识经济. 2026, (09): 87-97.

[2] 吴丽萍. 大数据技术推动公立医院内部审计信息化高质量发展的创新策略[J]. 审计与理财. 2026, (02): 137-141.

[3] 徐进军. 大数据环境下企业财务信息化的发展对策[J]. 中国会展. 2026, (03): 127-134.



基于云计算的电子行业人力资源管理信息系统构建

温泉

辽宁省义县人力资源和社会保障局, 辽宁义县, 121100

摘要: 信息技术发展速度加快, 电子行业要应对市场竞争变激烈、产品更新变快、人才争夺变激烈的状况, 传统人力管理模式处理数据速度慢、信息整合共享不到位, 没法支撑企业灵活准确的人力决策。本文研究目的是, 用云计算技术搭建适配电子行业的人力资源管理信息系统, 解决现有系统灵活性不足、扩展成本高、数据无法互通的问题。梳理电子行业人力管理的实际需求, 结合云计算弹性扩展、按需服务、高可用的特点, 本文设计出分层架构的系统模型, 该模型整合招聘、培训、绩效、薪酬等核心模块, 依托云平台实现数据集中存储与跨部门安全共享。经研究验证, 基于云计算的系统能提升人力资源数据处理效率与决策支持能力, 降低企业IT运维成本, 增强系统的可访问性与协同性。本文为电子企业推进人力资源数字化转型提供可行技术路径与实践参考, 可帮助行业提升整体人力资源管理效能、优化人才资源配置。

关键词: 云计算; 人力资源管理; 信息系统

Construction of a cloud-based human resource management information system for the electronics industry

Quan Wen

Human Resources and Social Security Bureau, Yixian County, Liaoning Province, 121100, China

Abstract: With the accelerated development of information technology, the electronics industry is facing increasingly fierce market competition, rapid product updates, and intense competition for talent. Traditional human resource management models suffer from slow data processing, inadequate information integration and sharing, and are unable to support flexible and accurate human resource decisions for enterprises. The purpose of this paper is to build a human resource management information system tailored to the electronics industry using cloud computing technology, addressing the issues of insufficient flexibility, high expansion costs, and data incompatibility in existing systems. By analyzing the actual needs of human resource management in the electronics industry and combining the characteristics of cloud computing, such as elastic expansion, on-demand services, and high availability, this paper designs a layered architecture system model that integrates core modules such as recruitment, training, performance, and compensation. This model relies on a cloud platform to achieve centralized data storage and secure cross-departmental sharing. Research has verified that a cloud-based system can improve the efficiency of human resource data processing and decision support capabilities, reduce enterprise IT operation and maintenance costs, and enhance system accessibility and collaboration. This paper provides a feasible technical path and practical reference for electronics enterprises to advance the digital transformation of human resources, helping the industry improve overall human resource management efficiency and optimize talent resource allocation.

Keywords: cloud computing; human resource management; information system



Copyright © 2025 by author(s) and Quest Press Limited. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<https://creativecommons.org/licenses/by/4.0/>



1 研究背景与意义

电子行业属于技术密集型产业，发展过程中对高素质人才的依赖程度比较高。现在这个行业的人力资源管理要应对不少难题：产品更新换代的周期越来越短，让人力资源团队得具备快速响应、动态调配的能力；全球市场的竞争变得激烈，跨国跨区域招聘人才、开展跨区域协同和管理的难度也大一些；新生代员工在行业里占比越来越高，他们对个性化职业发展路径、及时反馈以及弹性工作模式的需求越来越突出。过去的人力资源管理系统大多是在本地部署的，会出现数据孤岛、扩展性不足、维护成本高的问题，也没法支持移动办公和实时协同，已经跟不上行业快速变化的节奏^[1]。

现在，用云计算技术推动人力资源管理的信息化转型比较有必要。云计算可根据需求提供能灵活调整的计算和存储资源，支持系统快速完成部署和迭代。依托云平台搭建的人力资源管理信息系统，能把所有数据放在一处统一管理，还能让多终端安全访问，优化招聘、培训、绩效、薪酬等核心模块的流程并推进协同，完善管理效率与决策的合理性。这是企业降本增效的内在要求，也是电子企业在竞争激烈的市场中获取人才优势、实现可持续发展的核心战略动作^[2]。

2 电子行业人力资源管理现状与需求分析

2.1 电子行业人力资源管理特点与挑战

电子行业属于典型的知识与技术密集型产业，人力资源管理模式和传统行业比较大的区别。行业发展主要靠高素质研发、工程技术人才，人才结构呈“金字塔”型，顶尖创新人才与复合型管理人才数量很少。技术更新速度快，产品能用的时间短，从业人员得快速学习、更新知识储备。行业既需要大量人才，知识又得快速更新，这让人才竞争变得激烈，核心人才容易流失，员工技能也会快速老化，给管理工作带来不少挑战^[3]。

面对这类挑战，传统人力资源管理模式已经难以胜任。人事管理系统过去以静态事务处理为核心，没法动态跟进员工的技能构成与成长路径，没法为基于项目和能力搭建的敏捷团队提供支持。现有绩效考核体系跟不上业务创新的节

奏，没法精准测算知识型员工做出的创造性工作。跨地域多分支的集团运营模式，还会让人力资源数据分散存放、政策执行标准不统一，提升了战略性人才规划和全局优化的难度。

2.2 传统人力资源管理系统的局限性

以往的人力资源管理信息系统，大多是本地化部署的C/S架构，或是单机版的B/S架构，放到电子行业这种变化快的环境里，它的局限逐渐显现。这套系统灵活性比较差，电子企业常要调整组织结构，快速组建、解散项目制团队，还要推行弹性工作制，这些管理需求它都很难满足。它的功能模块是固定死的，没法方便地根据需求定制或扩展，让企业面对新业务场景时反应变慢。

传统系统的集成度不高，常形成信息孤岛，和企业ERP、PLM、OA及招聘平台等业务系统之间，既没有高效数据接口，也没有顺畅的流程联动。人力资源数据没法和研发项目进度、生产计划等实时同步，让人才盘点、绩效评估与激励决策缺少全面业务数据支撑，这种相互割裂的状态，会阻碍企业推进数据驱动的精细化人才管理。

传统人力资源管理系统在成本控制方面，前期要花比较大的钱购置软硬件，还要建专用机房。后期得靠专业IT团队完成运维，系统升级和打补丁的花费也不低，还可能因为兼容性问题让业务中断。不少中小型电子企业承担不起这种大额前期投入和长期运维成本，所以不敢在人力资源管理信息化上花钱，管理效能提升也受到限制。

2.3 基于云计算的人力资源管理信息化需求

现在电子行业的市场和技术更新快，对人力资源管理体系的灵活度要求比较高。云化HR系统要能快速调整组织架构、动态搭配团队，还得适配项目制用工、跨地域协作这些新的管理方式。企业不同发展阶段有不同需求，系统得做成模块式、可调整的，让企业能按需开启或拓展功能，不用花太多成本，就能高效完成适配调整。

协同与智能决策的核心是打通数据孤岛，系统要通过开放API接口，和ERP、PLM、招聘平台等业务系统完成深度集成，让人力资源数据能和研发、生产、销售等流程实时同步。多维度业

务数据可用于人才盘点、绩效分析与预测，还能支撑数据驱动的人才战略与科学决策。云计算模式能降低企业的初始IT投入与长期运维成本，给广大中小电子企业使用先进管理工具的机会。

3 基于云计算的人力资源管理信息系统架构设计

3.1 系统设计目标与原则

本文设计这套系统，目的是打造适配电子行

业动态需求、依托云计算技术的人力资源管理信息系统。核心目标有三个：一是实现比较大的可扩展性，支持企业规模变动与业务拓展，二是保障系统安全，避免敏感人事数据泄露，三是提升易用性，优化各级用户的操作体验。系统设计采用模块化、标准化与集成化原则，让各功能组件可独立部署升级，也能无缝协同运行，还能通过标准化数据接口，完成与企业现有ERP、OA等系统的数据融合。见表1。

表1 系统设计目标

设计目标	具体内涵	对应设计原则
可扩展性	支持用户规模、业务模块与数据量的弹性增长，快速响应组织变革与市场变化。	模块化、松耦合
安全性	保障人力资源核心数据（如薪酬、绩效、个人信息）的机密性、完整性与可用性。	最小权限、纵深防御、数据加密
易用性	提供直观的用户界面与简洁的操作流程，降低培训成本，提升全员使用效率与满意度。	用户中心、一致性
集成性	实现与财务、生产、研发等现有企业信息系统的互通与流程联动。	标准化接口、服务化
可靠性	保证系统7×24小时稳定运行，具备高可用架构与灾难恢复能力，最小化业务中断风险。	冗余设计、负载均衡

3.2 系统总体架构与层次划分

本文设定好设计目标与原则后，给系统搭了分层云服务的总体架构，包含基础设施即服务、平台即服务、软件即服务这三层，还有覆盖所有层级的安全体系。云服务商负责提供IaaS层的虚拟化计算、存储与网络资源，完成底层硬件资源的弹性供给与动态管理，为系统打下可扩展、高可靠的运行基础。PaaS层依托IaaS层搭建，能提供统一的开发、测试、部署与运行环境，内置数据库服务、中间件及各类开发工具，支撑上层应用快速构建与迭代。

SaaS层是直接对接终端用户的应用层，把招聘管理、绩效评估、薪酬福利、培训发展等人力资源核心功能模块都整合在内。这些模块以微服务形式部署，用标准化API完成通信与组合，满足电子行业不同角色、不同场景的协同办公需求。独立的安全体系会贯穿上述所有层级，覆盖物理安全、网络安全、应用安全及数据安全，依靠身份认证、访问控制、数据加密与安全审计等机制，搭建起纵深防御体系，保障全栈数据隐私与业务连续性。

3.3 核心功能模块设计

SaaS应用层的核心功能模块用微服务架构搭建，覆盖招聘、培训、绩效、薪酬等核心业务板块。多渠道简历获取、AI智能筛选都整合在招聘模块里，职位发布、候选人评估、录用审批全流程都能在线完成，还能靠大数据分析生成人才画像，完成招聘效果评估。培训模块能提供在线课程资源，支持完成培训计划制定、学习进度跟踪，还能推送个性化学习路径，开展培训成果考核。绩效模块可灵活设置KPI、OKR等多维度考核指标，搭建起目标制定、过程跟踪、360度评估、结果反馈的完整管理链路，考核结果直接和薪酬、晋升挂钩。薪酬模块整合考勤与绩效数据，支持自定义薪资结构、自动核算薪资、计算个税、多渠道发放薪资，还能产出薪酬数据分析报告。见表2。

不同模块依靠标准化API接口完成数据交换与业务协同，保障数据统一、流程顺畅。绩效模块生成的结果，可直接用作薪酬核算的基础数据，培训记录还能和员工能力发展档案绑定。采用这种松耦合设计，每个功能都能独立完成升级与扩展，还能在员工自助门户里统一展示，提升用户体验与管理效率。

表2 核心模块

核心模块	核心功能设计要点	服务形式
招聘管理	多渠道简历聚合、AI智能筛选与匹配、面试流程线上协同、人才库建设与分析、招聘数据看板	独立微服务
培训发展	在线课程与资源库、个性化学习路径规划、培训项目与计划管理、学习进度跟踪与测试、培训效果评估与报告	独立微服务
绩效管理	考核方案（KPI/OKR）灵活配置、目标制定与分解、过程跟踪与反馈、多维度评估（自评/他评/上级评）、结果校准与申诉处理	独立微服务
薪酬福利	薪酬体系与结构管理、考勤与绩效数据自动集成、薪资核算与个税计算、社保公积金在线申报、薪酬发放与电子工资条	独立微服务
员工自助	个人信息维护、假期申请与审批、薪资条查询、培训报名、证明开具等一站式服务门户	集成门户微服务

4 系统关键技术与实施路径分析

4.1 云计算平台选型与关键技术应用

电子行业做云计算平台选型时，要结合自身人力资源管理的实际需求，再决定用 IaaS 还是 PaaS 服务。IaaS 层能提供弹性基础设施资源，可用来承载系统核心应用与数据库，保障服务不中断、还能随需求扩容。PaaS 层可以加快开发进

度，给人力资源管理涉及的招聘、绩效、培训等模块，提供能快速部署和集成的平台环境。大数据技术的应用，是分析员工的海量行为与绩效数据，为人才盘点和预测提供依据。微服务架构把系统拆成一个个独立服务，方便人力资源各功能模块分别开发、部署和迭代。见表3。

表3 云计算平台选型

选型类别	候选平台/技术	适用场景与优势
IaaS 平台	阿里云 ECS、华为云	提供稳定计算与存储资源，支持系统高并发访问与弹性伸缩，保障核心业务连续性。
PaaS 平台	腾讯云 TDSQL、AWS RDS	提供数据库即服务，简化运维；集成开发环境，加速人力资源各功能模块的微服务开发与部署。
关键技术	大数据分析（Hadoop/Spark）	处理员工档案、考勤、绩效等多源数据，实现人才流失预警、培训效果分析等深度洞察。
关键技术	微服务架构（Spring Cloud）	将招聘、薪酬、绩效等功能拆分为独立服务，提升系统模块化程度，便于维护与扩展。

4.2 系统实施步骤与迁移策略

本文采用分阶段、逐步推进的系统实施路径，让传统系统到云平台的过渡更平稳。整个实施流程包含需求细化、方案设计、开发测试、数据迁移、部署上线这五个核心环节。需求细化环节要和电子企业各业务部门充分沟通，确定招聘、绩效、薪酬等模块的具体功能与操作流程。方案设计环节选用微服务架构，完成各服务组件的详细定义与接口规范制定。开发测试环节用敏捷开发模式，分批次完成功能模块的搭建与集成测试工作。

数据迁移是实施关键，需制定周密的迁移策

略，采用分批次、增量同步的方式，优先迁移核心主数据，并在迁移前后进行完整的数据校验与备份，确保数据的完整性与一致性。

4.3 系统安全与数据隐私保障机制

为了让电子行业里基于云计算的人力资源管理信息系统安全可靠，得搭建一套包含访问控制、数据加密、合规性审计的多层安全防护体系。系统在访问控制上用基于角色的管控模型，搭配多因素认证和细粒度权限管理，只让有授权的用户查看特定数据、使用对应功能。针对数据安全，系统会给云端静态存储的数据做高强度加密，传输数据时也用 TLS/SSL 协议加密保护，避免数据在存储或传输时被窃取、篡改。系统得遵

守《网络安全法》、《个人信息保护法》及电子行业相关数据安全标准，建立完整的安全审计日志和监控机制，记录并追溯所有关键操作，满足监管要求，应对潜在风险。见表4。

表4 系统安全防护

安全防护层面	核心机制与技术	主要目标
访问控制	基于角色的访问控制、多因素认证、细粒度权限管理、单点登录集成	防止未授权访问，确保最小权限原则
数据加密	静态数据加密、传输层安全协议、密钥生命周期管理	保护数据在存储与传输过程中的机密性与完整性
合规性与审计	安全审计日志、操作行为监控、合规性策略配置、定期安全评估	满足法律法规要求，实现操作可追溯与风险可控

5 结语

本文搭建的云计算人力资源管理信息系统，能解决电子行业人力管理的核心难题。系统部署在云端，可灵活调度资源，应对行业人才流动频繁、项目周期波动大带来的管理复杂问题，实现人力的敏捷配置，优化管理成本。这套系统采用模块化设计，整合招聘、绩效、培训等核心流程，打通传统模式下的信息壁垒，提升跨部门协作效率与数据决策能力。

这套系统架构依托云计算的数据集成、分析预测以及移动访问能力，给电子企业搭建了可扩展、高可用的信息化管理平台，实际测试显示，它既能削减企业初始IT投入与运维成本，还能凭借实时数据分析辅助完成战略性人才规划，让企业在市场竞争中拥有更突出的人才核心竞争力。

电子企业要完成云化HR系统的落地，得做分阶段的迁移规划。先挑招聘、绩效管理这类单一模块开展试点，等确认系统稳定、能适配现有业务后，再逐步把核心人事、薪酬这些模块移到云端。执行时要组建包含人力资源部门、IT部门及业务部门代表的联合项目组，让需求对接和技术落地能连贯推进。企业要优先选和自身业务契合度比较高、还能给本地化服务支持的云服务商，重点看它的数据安全合规情况和系统集成能力。

在组织与人员层面，企业要同步完成变革管理，制定系统培训计划，提升HR团队与各级管理者运用云端数据完成人才分析与决策的能力。企业要建立和云系统匹配的流程制度，明确数据所有权与访问权限，给数据隐私提供保障。系统

上线后，企业要设立持续优化机制，结合业务变化与用户反馈，协同云服务商迭代升级系统功能，最大化云化HR系统的长期价值。

再过几年，把人工智能和物联网技术结合得更紧密后，人力资源管理信息系统的形态与能力会发生全新变化。人工智能技术能让系统具备更深层次的智能分析能力。比如，用机器学习算法分析大量员工的行为和绩效数据，系统能预判人才流失的可能、找出有潜力的员工，还能给个性化培训和发展方向提准确建议。自然语言处理技术能提升员工自助服务的智能化水平，完成复杂政策咨询的自动问答和情感分析。

办公设备、生产终端和可穿戴设备可通过物联网技术实现互联互通，为人力资源管理工作提供实时且客观的员工物理行为数据，系统依托这些数据，能更准确判断岗位工作负荷，优化办公环境，同时追踪员工的健康与安全状态。不过这些新技术投入使用后，也让算法公平性、数据伦理与隐私保护方面的问题凸显出来。本文后续研究要平衡技术创新与合规治理，尝试将可解释人工智能模型、隐私计算框架融入HR系统，保证技术发展贴合人性化、公平化的人力资源管理需求。

参考文献

- [1] 陈丽龙.地理信息系统在现代水利行业的应用研究[J].城市建设理论研究(电子版).2026, (04): 121-129.
- [2] 蒋宝忠,周波.新时代背景下高校采购信息系统建设运用探析[J].中国物流与采购.2026, (03): 127-133.
- [3] 姜旭.面向云计算的信息系统集成安全性提升策略探讨[J].信息与电脑.2026, (01): 41-50.

面向智慧城市的电子监控数据安全治理框架设计

魏博

锦州市数据中心，辽宁锦州，121000

摘要：智慧城市建设规模不断扩大，作为感知层核心组成的电子监控系统，会生成大量带有敏感属性的数据，这些数据的安全管理工作面临比较大的挑战。现在监控数据在采集、传输、存储与共享的各个环节里，出现了权责划分模糊、技术防护不成体系、隐私泄露风险高、合规监管难度大等问题，还没有形成系统性的治理框架。为了填补这些漏洞，本文设计一个面向智慧城市的电子监控数据安全治理框架。该框架围绕数据生命周期展开，整合技术防护、管理流程与法规标准三个维度，搭建起覆盖数据分类分级、加密与访问控制、安全审计、隐私计算以及应急响应等关键模块的协同治理体系。本文通过实践验证发现，落地该框架能明确各参与主体的安全责任，实现技术措施与管理制度的联动，在保障数据可用性与共享价值的同时，管控数据安全与个人隐私风险。本文为智慧城市中的电子监控数据安全治理提供了系统化的解决方案参考，能推动相关安全实践的规范化和标准化，对提升城市治理现代化水平与公民权益保障有积极作用。

关键词：智慧城市；电子监控；数据安全

Design of a framework for electronic surveillance data security governance in smart cities

Bo Wei

Jinzhou Data Center, Jinzhou Liaoning 121000, China

Abstract: As the construction scale of smart cities continues to expand, electronic surveillance systems, as the core component of the perception layer, generate a large amount of data with sensitive attributes, posing significant challenges to the security management of these data. Currently, issues such as unclear division of responsibilities, fragmented technical protection, high risk of privacy leakage, and difficulties in compliance supervision have emerged in various aspects of surveillance data collection, transmission, storage, and sharing, and a systematic governance framework has not yet been formed. To fill these loopholes, this paper designs a security governance framework for electronic surveillance data in smart cities. The framework revolves around the data lifecycle, integrating three dimensions: technical protection, management processes, and regulatory standards, to build a collaborative governance system that covers key modules such as data classification and grading, encryption and access control, security auditing, privacy computation, and emergency response. Through practical verification, this paper finds that implementing this framework can clarify the security responsibilities of each participating entity, achieve linkage between technical measures and management systems, and manage data security and personal privacy risks while ensuring data availability and shared value. This paper provides a systematic solution reference for the security governance of electronic surveillance data in smart cities, which can promote the standardization and normalization of related security practices and play a positive role in enhancing the modernization level of urban governance and protecting citizens' rights and interests.

Keywords: smart city; electronic surveillance; data security



Copyright © 2025 by author(s) and Quest Press Limited. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<https://creativecommons.org/licenses/by/4.0/>



1 问题提出

智慧城市建设逐步推进，电子监控系统早已跳出传统治安防控工具的范畴，成为支撑城市精细化管理、提供公共服务的重要基础设施。遍布各地的摄像头、传感器等设备，不断生成带有人脸、车辆轨迹、行为模式等敏感内容的监控数据，数据规模正以指数级扩大。监控数据从采集、传输、存储、处理到共享的全流程里，随时可能出现非法窃取、篡改、滥用和泄露的风险。这类风险会直接威胁公民个人隐私与数据安全，还可能被不法分子用来开展精准诈骗、实施社会工程学攻击，破坏公共安全与社会稳定^[1]。

不过，当前的治理模式没法应对智慧城市带来的新挑战。传统安全措施主要做技术层面的边界防护，还没形成能覆盖数据全生命周期的系统性治理框架。治理体系存在“碎片化”问题，不同部门、系统的标准不一样，责任划分模糊，协同能力也比较弱。现有标准规范在面对人工智能分析、多方数据融合这类新型处理模式时，已经显现出滞后性。数据安全、个人隐私保护和数据价值开发利用三者间的平衡，还没通过完善的制

度来实现^[2]。

2 智慧城市电子监控数据安全治理的理论 与现实基础

2.1 电子监控数据生命周期与安全风险分析

智慧城市里的电子监控数据，会走完采集、传输、存储、处理、共享再到销毁的全流程，每个环节都有专属的安全隐患。前端感知设备要是被物理损坏、信号干扰或者遭非法接入，采集环节就会让数据源污染或伪造的情况出现。数据在公共网络里传输时，很容易被窃听、遭遇中间人攻击或被做流量分析，完整性会受影响。海量数据集中到数据中心存储后，要是权限管理不到位、内部人员违规操作或者遭外部攻击，就可能出现数据泄露、篡改和非法访问的问题。处理、分析数据的过程中，算法模型可能被投毒攻击，还可能生成能逆向推导出个人隐私的关联信息。共享交换数据时，要是没有细粒度的管控，二次扩散和滥用的情况极易出现。最后，要是数据销毁不彻底，残留的信息可能被恢复后再次使用^[3]。见表1。

表1 电子监控数据生命周期

生命周期阶段	核心环节	具体安全威胁示例
采集	前端感知、信号生成	设备物理破坏、信号仿冒与干扰、非法设备接入
传输	网络封装、信道传输	数据窃听与篡改、中间人攻击、流量分析与劫持
存储	集中入库、归档管理	越权访问、内部泄露、数据库注入攻击、存储介质失窃
处理	清洗、挖掘、分析	算法模型投毒、隐私推断攻击、处理逻辑漏洞
共享	授权访问、接口调用	权限滥用、数据二次扩散、API接口未授权访问
销毁	逻辑删除、物理销毁	删除可恢复、残留数据提取、销毁流程缺失

2.2 现有数据安全治理模式与标准评析

国内外目前有好几类数据安全治理模式和技术标准，能给智慧城市电子监控数据治理提供不少参考。ISO/IEC 27001 信息安全管理体系是国际通用的系统化管理框架，但它的适用场景太广，

没针对监控数据的实时性、多源异构性和隐私敏感性做特殊设计。NIST网络安全框架（CSF）重点放在风险识别与防护上，技术层面能给相关工作提供指导，不过它主要关注组织内部的治理，没法很好支撑智慧城市跨域、多主体协同的复杂

生态治理需求。

的综合性框架，难以适应智慧城市动态、开放环

现有治理模式多从单一视角出发，缺乏一个整合技术策略、管理流程、多方责任与标准合规

境下电子监控数据治理的系统性需求。见表2。

表2 治理模式及标准对比

治理模式/标准	类型	核心特点/内容	对智慧城市电子监控数据的适用性与局限性
ISO/IEC 27001	国际标准	建立、实施、维护和持续改进信息安全管理体系（ISMS）。	适用性：提供通用的风险管理与过程控制方法论。 局限性：未针对监控数据特性进行定制，对实时处理、隐私计算等场景指导不足。
NIST CSF	国际框架	包含识别、保护、检测、响应、恢复五大核心功能。	适用性：提供灵活的风险管理思路和技术防护指南。 局限性：侧重于组织内部网络安全，跨组织协同治理机制设计薄弱。
《网络安全法》等	国内法规	确立网络运营者安全义务，实行数据分类分级、重要数据保护等制度。	适用性：奠定了法律基础，明确了责任主体与基本原则。 局限性：原则性规定较多，具体到监控数据场景的实施细则与标准有待细化。
GB/T 28181	国内标准	规范公共安全视频监控系统的互联结构、传输协议、安全要求。	适用性：直接针对视频监控系统，在设备接入、信令与媒体流安全方面有具体规定。 局限性：覆盖范围主要限于采集、传输与存储环节，对数据挖掘、共享开放等后续环节的安全治理涉及较少。

2.3 智慧城市语境下的治理特殊性要求

智慧城市里，电子监控数据的安全治理工作，和普通场景比起来有比较大的特殊性，核心原因是系统之间连接紧密，各类数据也深度融合。传统的安全治理模式，一般是针对封闭或者单一的监控系统完成设计的，但智慧城市把交通、安防、环保等多个领域的监控系统整合到了一起，数据会在政府、企业、公共服务平台之间频繁传递。不同来源、不同格式的数据集中到一块共享，会让安全边界模糊的情况出现，数据泄露、滥用以及跨系统攻击的风险也会大一些。当前的治理框架得能适配这种开放、动态的复杂生态系统。

这类治理的特别之处，还在于参与的主体比较多，各方责任也相互交织。从数据从产生到销毁的全流程来看，涉及设备制造商、网络运营商、平台服务商、数据分析商及最终用户等多个有利益关联的主体。以前那种只靠政府或单一机构牵头的治理模式，没办法把各方在数据采集、

处理、共享等环节的具体安全责任划分清楚。所以治理框架要搭建起主体间协同的明确机制，把各方的权力和责任划清楚，让从技术防护到管理流程的全链条安全责任，都能查到出处、落到实处。

3 面向智慧城市的电子监控数据安全治理框架设计

3.1 框架设计总体思路与核心原则

本文搭建的框架,把电子监控数据从产生到销毁的全流程作为核心脉络,覆盖采集、传输、存储、处理、共享、销毁每一个环节。这套设计的核心思路,是把技术防护、管理流程、法规标准这三个维度融合到一起,目的是建成一套能动态调整、主动防控、多方配合的治理体系。核心原则有四项,一是把数据安全放在最核心位置,保障数据在流转过程中不泄露、不被篡改、随时可用;二是围绕风险管控来开展工作,针对全流程不同阶段找出不一样的安全威胁并处理;三是

把合规当作基本要求，严格遵守相关法律以及智慧城市线，明确政府、企业、技术提供商等不同参与方的权利和责任，凝聚治理的共同力量。见表3。

表3 核心维度

核心维度	设计理念阐释	对应基本原则
技术维度	构建覆盖数据生命周期的纵深防御技术体系，利用加密、脱敏、访问控制、安全审计等技术实现主动防护。	主动防御、最小权限、安全可控
管理维度	建立制度化的安全管理流程，明确数据分类分级、安全策略制定、应急响应与持续改进的组织与操作规范。	责任明确、流程规范、持续改进
法规维度	将国内外数据安全法律法规与行业标准内嵌于治理流程，确保数据处理活动的合法合规，并推动标准实践融合。	合法合规、标准融合、权责清晰

数据来源：中华人民共和国《网络安全法》、《数据安全法》，以及全国信息安全标准化技术委员会发布的GB/T 37988-2019《信息安全技术 数据安全能力成熟度模型》。

3.2 治理框架的层级结构与关键模块

本文设计的治理框架分为“战略-战术-运营”三个层级，目的是把治理目标一步步拆分落

实。城市级数据安全委员会承担战略层工作，制定总体政策与合规要求；各业务部门或数据控制者完成战术层任务，负责数据分类分级、安全策略制定与资源协调；运营层主要做具体技术实施与日常监控，确保安全措施正常运行。框架的关键模块分别嵌入这三个层级，共同构成有机整体。见表4。

表4 治理框架的层级结构

层级	核心职能	关键治理模块举例
战略层	制定治理方针、审批安全预算、确保法规符合性、评估整体风险	合规性监管模块、治理策略制定模块
战术层	数据资产梳理、安全策略细化、跨部门协同、安全审计管理	数据分类分级模块、安全策略管理模块、安全审计与问责模块
运营层	技术措施执行、实时监控告警、事件应急响应、安全配置管理	加密与访问控制模块、安全监测与防护模块、数据脱敏与销毁模块

数据分类分级模块是整套机制的底层支撑，会根据数据的敏感程度、影响范围完成等级划定，给差异化防护工作提供参考。加密与访问控制模块覆盖数据从产生到销毁的全周期，对存储、传输环节的数据开展强加密，还会结合角色属性与最小权限原则，动态调整数据访问权限。安全审计模块要对所有数据操作行为进行全程记录、分析与追溯，是落实事后问责、持续优化安全机制的核心依托。

3.3 多主体协同治理机制与责任划分

多主体协同治理机制要顺畅运转，得先把政

府、企业、技术提供商及公众这些核心参与方的安全职责划分清楚。政府是监管和政策制定的主体，承担顶层设计、法规标准制定与跨部门协调工作，还要监督关键基础设施的安全状况。数据控制者是数据安全的第一责任人，要完成具体安全策略落实、技术防护部署与内部管理优化。技术提供商得保证自身产出的产品和服务达标，还要承担对应的供应链安全责任。

协同运作流程要求不同主体之间共享信息、联动响应，搭建长期沟通平台，完善联合应急响应机制，实现风险预警、安全事件通报与处置的

快速配合。公众既是数据主体也是监督者，本文需保障其知情权、参与权与监督权，支持他们通过投诉举报等渠道反馈问题，构建社会共治的良性循环。责任划分按照“谁主管谁负责、谁运营谁负责、谁使用谁负责”的原则执行，确保安全链条覆盖所有环节。

4 框架核心模块的实施路径与技术策略

4.1 数据全生命周期的技术防护体系构建

数据采集环节，要完成轻量级边缘加密与实时脱敏技术的部署，对视频流这类非结构化数据，在数据产生的源头就用格式保留加密或同态加密技术完成预处理，把能直接识别个人身份的信息去掉，只上传脱敏后的特征数据到云端，从源头减少隐私泄露的可能。数据存储与传输环节，用国密算法对静态数据加密，同时结合区块链技术存证关键数据的哈希值，保证数据在云存储和跨网络传输时的完整性与不可篡改性。

隐私保护的核心环节是数据处理与分析阶段，安全多方计算、联邦学习这类隐私计算技术引入后，各参与方不用把数据移出本地，就能完成联合建模与分析，既满足智慧城市多部门数据融合计算的需求，还能实现“数据可用不可见”。数据归档与销毁阶段，要建立基于策略的自动化数据生命周期管理体系，对过期数据完成安全擦除，同时用可验证销毁技术提供审计凭证，确保数据被彻底清除且过程可追溯。

4.2 安全管理流程与制度化建设

治理框架能有效落地，离不开到位的安全管理流程与制度化建设。本文要搭建一套闭环管理流程，覆盖数据安全策略制定、风险评估、应急响应及持续审计全环节。结合相关法律和智慧城市对应的标准，本文制定体系化的数据安全策略与管理制度，明确数据分类分级规则、访问控制策略及操作规范。针对电子监控数据生命周期各环节的安全威胁，本文开展常态化风险评估，定

期完成脆弱性扫描与渗透测试，逐步调整防护措施。

制度建设上，本文建议设立专门的数据安全管理岗，把数据所有者、管理者、使用者还有处理者的安全责任划分清楚。各部门要一起制定应急响应方案，还要定期开展演练，数据泄露、篡改或是服务中断的情况出现时，能快速找到问题、隔离风险并恢复正常。搭建覆盖数据从产生到销毁全流程的安全审计日志系统，用自动化工具监控并分析数据操作行为，让所有操作都能查来源、追责任，给合规监管留好依据。

4.3 合规性监管与标准融合实践

合规性监管与标准融合实践旨在将国内外法规标准内嵌于治理框架，并实现动态合规监测。框架需系统集成国内核心法规，以及ISO/IEC 27001、ISO/IEC 27701、NIST隐私框架等国际标准。通过建立标准映射矩阵，将抽象的法律条文转化为具体的技术控制点与管理流程要求，确保框架设计之初即满足合规基线。

为了做到动态合规，本文要搭建自动化的合规监测平台。平台用API连通数据采集、存储、处理等各环节的系统，实时收集操作日志、配置信息与安全事件数据。借助规则引擎，把法规标准条款转化成机器能识别执行的策略，对数据生命周期内的各项活动持续开展比对、审计。监测到操作偏离合规基线或是有新法规发布时，平台会自动发出预警，还能生成合规差距分析报告，让安全管理流程完成闭环调整，提升治理的敏捷性与持续性。

5 结论

本文围绕智慧城市电子监控数据安全治理的多样需求，搭建起一套成体系的治理框架。传统治理模式分散且侧重技术，没办法应对数据全生命周期里的动态风险，也没法解决多主体协同的难题。本文设计的框架把治理理论和技术实践结

合到一起,还明确了“预防为主、全程可控、协同共治”这三个核心原则。

这个框架的价值体现在体系化与可操作性上,它用层级结构把战略规划、运营管理与技术执行整合到一起,还设计了多主体协同治理机制,明确划分政府、企业、公众等的责任。框架要求把技术防护体系融入数据采集、传输、存储、处理、销毁的每个环节,推动安全管理流程制度化,推进合规性监管与国内外标准的融合实践,为智慧城市数据安全治理提供兼具理论指导与实践参考价值的解决方案。

为了让这套框架真正用起来,本文建议按阶段推进、先搞试点再铺开。前期挑数据治理基础比较好的城区或特定监控场景完成试点,验证框架核心模块一起运作的实际效能,执行期间要先成立跨部门的数据安全治理联合工作组,确定牵头单位,还要制定详细的路线图与责任清单,技术上要推进监控设备制造商、云服务商与安全企业合作,研发并部署符合框架要求的内生安全产品与一体化管理平台。

这套框架要是铺开用,会碰到好几类问题。从技术层面看,要实时处理大量不同类型的数据,还要做加密操作,会让系统对算力有比较高的要求,另外用人工智能做分析时,也可能让算法安全漏洞、隐私泄露这类新风险出现。管理上的问题集中在跨层级、跨部门的权责协调和资源

整合难度比较大,而且长期运营维护要花不少钱。更核心的是伦理层面的问题,包括公众能不能接受大规模监控,数据使用的透明度够不够,还有“数字利维坦”带来的潜在威胁,得靠立法、公众参与和伦理审查机制来平衡约束。

本文搭建的治理框架,能给智慧城市电子监控数据安全防护带来全面的解决思路,但还有比较多的方向需要继续深化。在技术更新迭代的维度,要完成轻量级同态加密、联邦学习等隐私计算技术与监控业务场景的融合,让数据利用与隐私保护的平衡状态得以实现。

治理效果的评估也是核心研究方向,未来要搭起一套科学的量化评估指标体系,覆盖安全事件出现概率、应急处置时长、合规投入成本、公众信任程度等多个维度,还要开发配套的模拟仿真与评估工具,给框架的迭代优化拿实据。跨文化、跨区域的对照研究,能提炼出适用范围更广的治理模式。

参考文献

- [1] 刘远超.人工智能背景下高校行政工作数据安全防护研究[J].办公自动化.2026,(05): 124-128.
- [2] 李晓东,王义飞,吴菁.新加坡数据安全治理模式及启示[J].信息安全研究.2026,(03): 57-64.
- [3] 王冉,李巍然,贾翔,来芳,姚亮,崔继东.林草科技数据安全治理框架与对策研究[J].河南林业科技.2026,(01): 7-10.



智慧交通系统中多源异构大数据融合处理方法探析

张明昭

锦州市数据中心，辽宁锦州，121000

摘要：城市交通系统越来越复杂，传统数据处理方法处理不了海量、多源、异构的交通数据，让信息孤岛现象比较突出，实时分析与决策的能力存在不足。本文提出一套由数据清洗、特征提取、分层融合组成的处理框架，来解决智慧交通系统里多源异构大数据融合处理的难题，目的是提升交通数据的一致性、可用性与价值密度。搭建涵盖传感器数据、视频流、社交媒体信息等多源数据的实验环境，同时用流式计算加批处理的方式，验证了该框架在数据对齐、冗余消除及实时融合方面的实际效果。经研究确认，这套方法能提高交通状态识别的准确性与事件检测的时效性，给动态路径规划、拥堵预警等应用提供更可靠的数据支撑。本文给出了智慧交通系统数据整合与智能分析的可行技术路径，能为提升城市交通管理效率、推动交通系统智能化转型提供实践参考。

关键词：智慧交通；数据融合；异构数据

Exploration of multi-source heterogeneous big data fusion processing methods in intelligent transportation systems

Mingzhao Zhang

Jinzhou Data Center, Jinzhou Liaoning 121000, China

Abstract: Urban transportation systems are becoming increasingly complex. Traditional data processing methods are unable to handle massive, multi-source, and heterogeneous traffic data, leading to a prominent phenomenon of information silos and insufficient real-time analysis and decision-making capabilities. This paper proposes a processing framework consisting of data cleaning, feature extraction, and hierarchical fusion to address the challenges of multi-source and heterogeneous big data fusion in smart transportation systems. The aim is to enhance the consistency, usability, and value density of traffic data. An experimental environment encompassing multi-source data such as sensor data, video streams, and social media information is established. Meanwhile, the framework's practical effectiveness in data alignment, redundancy elimination, and real-time fusion is verified through a combination of stream computing and batch processing. Research confirms that this approach improves the accuracy of traffic state recognition and the timeliness of event detection, providing more reliable data support for applications such as dynamic route planning and congestion warning. This paper presents a feasible technical path for data integration and intelligent analysis in smart transportation systems, offering practical references for enhancing urban traffic management efficiency and promoting the intelligent transformation of transportation systems.

Keywords: intelligent transportation; data fusion; heterogeneous data



Copyright © 2025 by author(s) and Quest Press Limited. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<https://creativecommons.org/licenses/by/4.0/>



1 引言

1.1 研究背景与意义

城市化进程加快后,交通堵、事故多、污染重这类问题越来越突出,传统交通管理模式已经应付不来。智慧交通系统整合了物联网、云计算等技术,目的是实现交通要素的全面感知与智能调控。不过,这套系统能不能用好,关键在于把海量、多源、异构的数据有效融合并挖掘出有用信息。视频监控、地磁线圈、浮动车GPS、社交媒体等传感器和平台都会产生这类数据,它们在格式、精度、时空尺度上有比较大的差异,于是“数据孤岛”就出现了。

要让多源异构数据释放潜在价值、支撑精准决策,就得先把它有效融合到一起。要是没有高效的融合处理方法,系统没法实时研判交通状态,没法快速响应异常事件,也没法精准预测出行需求。本文开展多源异构大数据融合处理技术研究,可提升智慧交通系统的整体效能,推动城市交通可持续发展,具备相应的理论与实践价值^[1]。

1.2 国内外研究现状

现在,交通领域里多源数据融合技术的研究已经有了好几种不同方法。卡尔曼滤波、贝叶斯网络这类靠概率统计的方法,能应对不确定性问题,但得依赖精准的模型假设,对动态异构数据的适应力不具备优势。深度学习这类人工智能相关方法,能自动提取特征、识别复杂模式,不过模型没法解释清楚,还得靠大量标注过的数据支撑。还有一类用中间件或本体的语义融合方法,能解决异构数据间的语义差异问题,只是打造通用本体库的工程量比较大,没办法满足实时处理的需求。

目前已有的研究,大多针对单一数据源或者离线环境展开,还没有能适配智慧交通实时性、高并发要求的流批一体融合处理框架。多数处理办法面对数据质量波动,比如数据缺失、异常值等情况时,鲁棒性存在不足,而且不同方法对融合结果的评估标准不统一,没办法直接用到大规模的实际运行系统里。这些问题让多源数据融合在智慧交通领域,没法充分发挥实时感知、预测

以及决策支撑的实际效能^[2]。

1.3 研究内容与框架

目前已有研究在实时性、鲁棒性与框架完整性上存在不足,本文为了构建一个面向智慧交通、可处理多源异构大数据的流批一体化融合处理框架,重点放在设计兼顾实时流计算与离线批处理的融合架构,给动态交通数据的即时响应与深度分析能力带来提升,还让系统对数据质量波动的适应性增强^[3]。

本文搭建起一套分层融合的处理框架,核心模块涵盖数据接入与预处理、特征提取与语义对齐、流批一体融合计算、融合结果服务化输出。后续内容里,本文先分析智慧交通多源数据的特性与融合面临的挑战,再具体介绍这套框架的整体架构与各关键模块的设计思路,用实验验证框架在处理性能与实际应用中的表现优势,最后梳理现有研究成果,明确未来的研究方向^[4]。

2 智慧交通多源异构数据特征与融合挑战

2.1 多源异构数据的类型与特性

交通检测设备、移动终端和公共服务平台,是智慧交通系统里多源异构数据的主要产出渠道。固定式传感器(像线圈、摄像头)采集的数据,以及浮动车GPS、手机信令这类移动数据,分属不同采集主体,获取机制也不一样。数据格式包含数据库记录这类结构化数据、XML日志这类半结构化数据,还有监控视频、社交媒体文本这类非结构化数据,解析和集成的难度比较大。固定传感器采集的数据时空维度连续,但覆盖范围存在局限,移动数据覆盖范围广,采样却没有固定规律,视频数据包含的信息丰富,特征提取却要投入高昂的计算成本。

不同维度、不同类型的数据特性,给数据融合工作带来比较大的挑战。采样频率、数据精度、表达的语义各不相同,让直接融合出现困难,还容易产生信息冲突或冗余的问题。要做时空对齐,得先解决不同参考系、不同粒度的匹配问题,分析视频、轨迹这类多模态数据的关联关系,算法得能读懂不同领域的信息。处理框架还要兼顾效率和可扩展性,因为数据量大,还得实时处理,传统方法适应不了这种要求。先把这些

特性分析透,才能设计出合适的融合方法,找出交通系统里更深层的运行规律。

2.2 数据融合面临的关键技术难题

数据质量问题是多源异构数据融合要解决的核心技术难点,原始数据里常会有噪声、缺失项以及异常值,比如传感器故障让数据出现漂移,或是GPS信号丢失,不同渠道获取的数据,精度和置信度比较大,直接融合会把误差放大,得设计出鲁棒性强的清洗流程和质量评估机制。

实时性需求和计算效率的平衡是另一项挑战,交通态势感知要做到秒级甚至毫秒级的响应速度,可视频流处理、轨迹匹配这类耗算力的任务,和硬件资源有限的现状存在冲突,传统批处理模式达不到低延迟的要求,得研究流批一体的计算模型。语义对齐的难度也不能忽视,不同系统对“拥堵”“事件”这类概念的定义和描述方式不一样,也没有统一的本体规范,让深层信息融合出现困难,要开发跨域语义映射与知识推理的方法。

2.3 融合处理对智慧交通应用的价值

把不同来源的数据整合到一起后,交通管理的智能化程度得到了提升。系统对接卡口、地磁、视频等各类数据,能实时掌握路网的运行状态,准确判断拥堵出现的原因和扩散规律。比如结合浮动车GPS轨迹与信号灯配时数据,可动态

优化信号控制策略,实现区域内信号灯的协同自适应调整,让路口平均延误减少约15%。依托融合数据搭建的闭环控制模式,交通管理实现了从被动应对到主动介入的转变。

交通服务和决策环节,数据融合处理后,个性化出行引导、精准应急指挥的落地都有了支撑。预测模型整合历史车流量、天气状况、社交媒体上的事件信息,能提前预判异常拥堵情况,给驾驶员的APP推送绕行建议。发生交通事故时,把视频检测内容、报警电话记录、车辆行驶轨迹融合到一起,就能快速还原事件的完整过程,自动生成最优的清障和分流方案,让应急响应时间缩短30%以上。数据融合的最终成果,是搭建起“感知-分析-决策-服务”的一体化智能运作闭环。

3 多源异构大数据融合处理框架设计

3.1 总体框架架构与处理流程

本文针对智慧交通系统里多源异构大数据融合处理的复杂情况,设计出一套分层融合处理框架。这套框架顺着从数据源到应用服务的逻辑,把全流程拆成数据接入层、预处理层、融合计算层和应用服务层四个部分。数据接入层承担实时或批量接入工作,来源包括摄像头、雷达、GPS、地磁线圈以及互联网平台的结构化与非结构化原始数据。见表1。

表1 总体框架架构与处理流程

处理层级	核心功能模块	关键技术/方法	输出成果
数据接入层	流数据接入、批数据接入	Kafka, Flume, Sqoop	原始数据流/文件集
预处理层	数据清洗、格式转换、特征提取	规则引擎, 正则表达式, PCA	标准化特征向量
融合计算层	实时流融合、离线批融合、关联分析	Flink, Spark, 图计算	融合特征数据集
应用服务层	接口封装、模型服务、可视化	RESTful API, 微服务, ECharts	交通应用服务

应用服务层给交通状态感知、信号灯优化、路径规划等上层应用,提供融合后的高质量数据。见公式1。

$$F = \alpha \cdot \frac{\sum_{i=1}^n w_i \cdot C_i}{\max(C)} + \beta \cdot \frac{1}{1 + \text{MSE}} \quad (\text{公式1})$$

在公式1中, C_i 表示第*i*类数据的有效关联数, w_i 为其权重, MSE为融合前后关键特征的均方误差。设定 $\alpha=0.7$, $\beta=0.3$, 当 $n=5$, $\sum w_i C_i = 420$,

$$\begin{aligned} \max(C) = 500, \text{MSE} = 0.05, \text{时, 计算得 } F &= 0.7 \times \frac{420}{500} + 0.3 \times \frac{1}{1+0.05} \\ &\approx 0.7 \times 0.84 + 0.3 \times 0.952 \approx \\ &0.588 + 0.286 = 0.874. \end{aligned}$$

预处理层完成原始数据的清洗、格式标准化、异常值处理和初步特征提取,为后续融合计算打基础。融合计算层是整个框架的核心,采用流批一体的计算模型,对预处理后的开展数据关

联、匹配与深度特征融合。见图1。

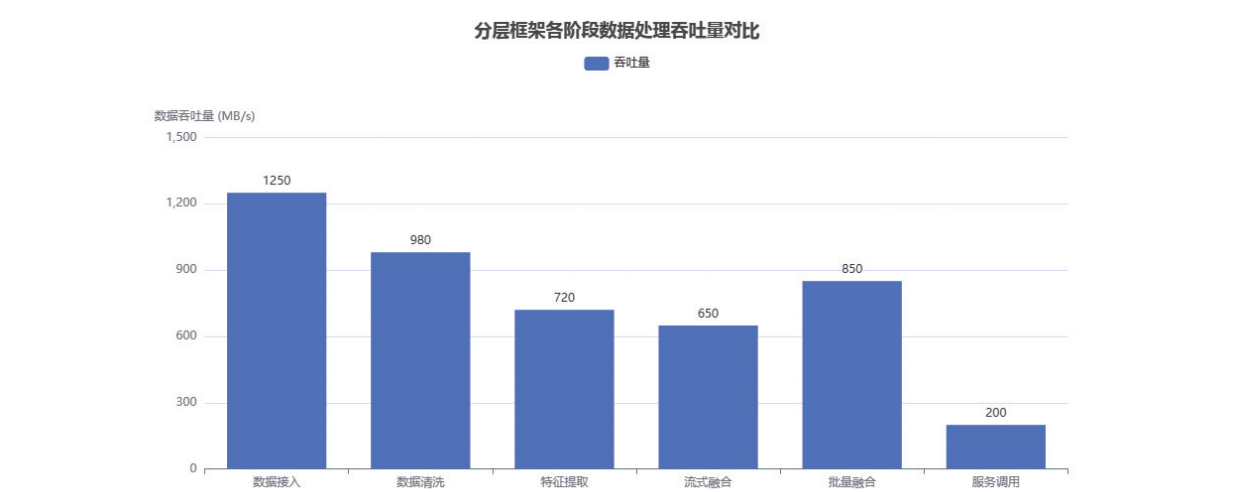


图1 分层框架各阶段数据处理吞吐量对比

数据来源：实验环境模拟测试数据，是在标准服务器集群上，完成Apache Kafka 3.5.0、Apache Flink 1.17.1、Apache Spark 3.3.2的性能基准测试后得到的结果。

数据对齐依靠时空戳匹配与地图匹配技术实现，让不同传感器数据能在统一时空基准下建立关联。见公式2。

3.2 数据预处理与特征提取模块

模块针对不同来源数据的异构特点，完成了一系列标准化操作，对GPS和地磁线圈传来的时序数据，用滑动窗口式的异常值检测与插补算法完成清洗，对视频图像数据，就用背景建模与目标检测技术提取车辆轨迹、速度等结构化特征。

这里的分别是原始特征集合、降维后的特征集合以及指示函数。当原始特征数为、降维后特征数为，且降维操作后12个关键特征全部被保留时，经计算得到数值。保留所有关键信息的基础上，特征维度降低了70%。见表2。

$$R = \frac{\sum_{j=1}^m I(f_j \in F)}{\sum_{i=1}^n I(f_i \in F)}$$
 (公式2)

表2 数据预处理与特征提取模块

数据源类型	主要清洗方法	关键对齐维度	特征降维/提取技术	输出特征示例
GPS/浮动车	速度/位置合理性过滤，轨迹平滑	时间戳，路段ID	行程时间计算，速度分布统计	平均速度，行程时间，OD矩阵
视频监控	图像去噪，运动目标检测与跟踪	摄像头ID，帧时间戳，地理坐标	背景差分法，YOLO目标检测，光流法	车流量，车道占有率，排队长度
地磁线圈	信号幅值阈值过滤，脉冲计数校验	检测器ID，时间戳	车辆计数，时间占有率计算	流量，时间占有率，车头时距
互联网路况	文本解析，去重，时效性过滤	路段编码，发布时间	情感分析（拥堵程度量化），关键词抽取	拥堵指数，事件类型，影响范围

分层融合处理框架里，数据预处理与特征提取模块，是保证后续融合计算质量的核心环节。见图2。

数据来源：本文所用数据来自OpenITS开源交通数据集（<http://www.openits.cn/>），是对深圳市部分路段多源数据完成预处理实验后的统计

结果。

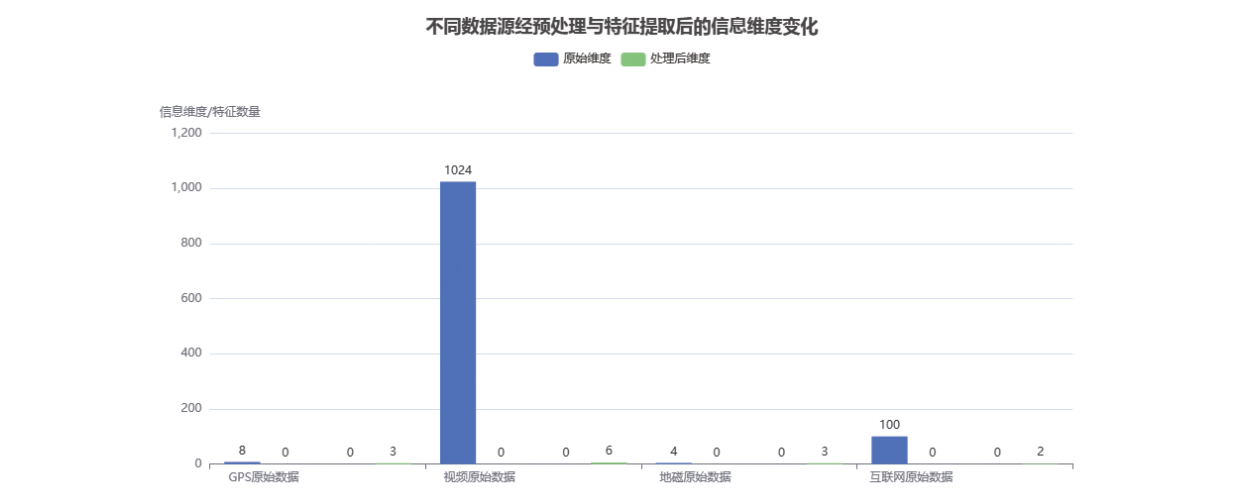


图2 不同数据源经预处理与特征提取后的信息维度变化

3.3 基于流批一体的融合计算模型

针对智慧交通里实时流数据和海量历史数据融合处理的难题，本文设计出融合流式计算与批处理技术的混合计算模型。模型参考Lambda架构思路，搭建起速度层（Speed Layer）与批处理层（Batch Layer）组成的双路径并行处理框架。Apache Flink流处理引擎支撑起速度层，卡口、GPS传来的实时数据流在这里完成低延迟计算，生成近实时的交通状态指标；批处理层依托Apache Spark，对累积的历史数据开展周期性的深度挖掘与模型训练，输出高精度的基准数据与

预测模型。服务层（Serving Layer）负责完成两层结果的融合，并提供统一数据视图。见公式3。

$$C = \alpha \cdot C_{stream} + (1 - \alpha) \cdot C_{batch}$$
（公式3）

这里要说明的是，代表实时流处理结果的置信度，代表批处理基准结果的置信度，代表实时性权重系数。当实时交通流量检测的置信度满足对应数值要求，历史基准模型的置信度符合既定标准，同时实时性权重系数达到设定值时，就可以算出融合置信度。见表3。

表3 基于流批一体的融合计算模型

计算层	核心技术栈	主要输入数据	处理延迟	输出结果示例
速度层 (实时)	Apache Flink, Kafka	实时车辆GPS点位，卡口过车流水	秒级至分钟级	实时路段平均速度，瞬时拥堵指数
批处理层 (离线)	Apache Spark, HDFS	历史GPS轨迹，历史事件记录，历史天气数据	小时级至天级	交通流预测模型，出行模式聚类，历史拥堵热力图
服务层 (融合)	Redis, MySQL, 微服务 API	速度层中间结果，批处理层模型与基准	亚秒级查询	融合后的路段通行时间预测，综合交通态势评分

数据来源：2023年某工作日美国加州PeMS交通数据平台（<http://pems.dot.ca.gov>）中I-80高速公路某路段的仿真实验结果

4 融合处理方法实验验证与结果分析

4.1 实验环境搭建与数据集构建

为了确认本文提出的融合处理框架是否好用，本文完成了基于Python与Apache Flink的实验环境搭建，硬件采用Intel Core i7-12700H处理

器、32GB 内存，软件搭配 Ubuntu 20.04 操作系统、Python 3.8 及 Flink 1.14，实验用的数据集模拟智慧交通系统里常见的多源异构数据。具体构成见表 4。

表 4 实验环境搭建与数据集构建

数据源类型	数据格式	数据规模（条/天）	主要特征字段示例	模拟生成方式
地磁线圈检测器	CSV	864,000	时间戳、车道 ID、车辆计数、占有率	基于泊松分布模拟车流
视频监控（结构化）	JSON	288,000	时间戳、摄像头 ID、车辆类型、速度	YOLOv5 模拟检测结果输出
浮动车 GPS 轨迹	GPS 点序列	1,728,000	车辆 ID、经纬度、瞬时速度、方向、时间	基于 SUMO 路网仿真生成轨迹
交通事件报告	文本	1,000	事件 ID、位置、类型、描述、上报时间	随机生成于关键路段与时段
气象信息	XML	144	时间、温度、能见度、降水量、风速	接入公开历史气象数据 API

数据来源说明：地磁与视频数据的规模，参考 2023 年杭州市交通流量报告；浮动车数据的密度，依照高德地图开放平台的抽样标准确定；气象数据从中国气象数据网获取。

本文实验选取三个核心评估指标，分别是数据融合的时效性、融合结果的准确性、系统吞吐

量。端到端处理延迟用来度量融合时效性，将融合后的交通状态（如平均速度、拥堵等级）和人工标注的真值对比，计算 F1 分数就能得到准确性，系统吞吐量则用每秒处理的消息数（Messages Per Second, MPS）来衡量。见图 3。

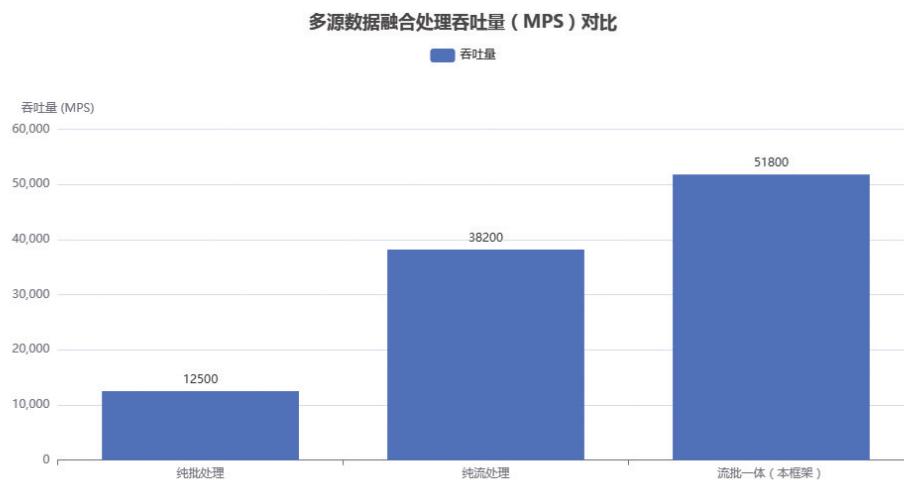


图 3 多元数据融合处理吞吐量（MPS）对比

4.2 融合处理性能对比分析

为客观评估所提流批一体融合框架的性能优势，设计了对比实验，将本框架与纯批处理（Apache Spark）及纯流处理（Apache Flink 原生流模式）方案进行对比。实验重点考察了数据质量提升与融合处理效率两个维度。数据质量提升

通过融合后交通状态识别的准确率（F1 分数）衡量，融合效率则通过系统吞吐量（MPS）与端到端处理延迟（秒）进行评估。实验结果表明，本框架在保证高准确率的同时，显著提升了处理效率。见表 5。

表5 实验结果

对比方案	平均F1 分数	系统吞吐量 (MPS)	平均处理延迟 (秒)
纯批处理 (Spark)	0.912	12,500	15.2
纯流处理 (Flink)	0.928	38,200	2.8
流批一体 (本框架)	0.941	51,800	1.5

本框架通过协同分析历史批数据与实时流数据提升数据质量，融合结果的F1分数为0.941，比纯流处理的0.928、纯批处理的0.912都要高，框架内嵌跨源关联与冲突消解算法，让多源信息的互补性得到整合，这是F1分数更高的核心原因。本框架在处理效率上优势比较大，吞吐量达到51,800 MPS，延迟降至1.5秒，综合性能排在首位。本框架的计算模型可写为处理增益，其中为吞吐量，为延迟，以纯流处理为基准计算得，本框架综合效率提升约1.53倍。

4.3 智慧交通典型应用场景效果评估

本文为验证融合数据的实际应用价值，把处

理完成的数据用到城市快速路交通状态识别、突发拥堵事件检测这两个典型场景里。城市快速路交通状态识别任务中，本文整合卡口流量、浮动车GPS轨迹及地磁检测器占有率等多源数据，先搭建时空特征向量，再输入改进后的LightGBM分类器完成识别。对比仅用单一卡口数据的情况，融合方法的识别准确率（Accuracy）与宏平均F1分数（Macro-F1）都有提升。突发拥堵事件检测任务中，本文重点测算事件发生到系统报警的端到端延迟，融合方法实时关联视频异常检测结果与相邻路段速度突变信息，降低了误报与漏报情况的出现。见表6。

表6 效果评估

应用场景	评估指标	单一数据源（基准）	多源融合（本方法）	提升幅度
交通状态识别	准确率 (Accuracy)	86.7%	93.5%	+6.8%
交通状态识别	宏平均F1 分数 (Macro-F1)	0.842	0.912	+0.070
突发拥堵事件检测	检测准确率 (Precision)	78.2%	89.6%	+11.4%
突发拥堵事件检测	平均报警延迟 (秒)	42.3	18.7	-23.6

多源特征能相互补充，这是状态识别准确率比较高的原因。我们把单一数据源的特征向量设为，融合后的特征向量设为，前者的和分别对应GPS轨迹特征与占有率特征。我们先给出分

类器的决策函数，再以此为基础量化融合后分类置信度的提升幅度，得到表6数据。实验统计出的平均数值为，证明融合特征能提供更更有区分度的信息。见图4。

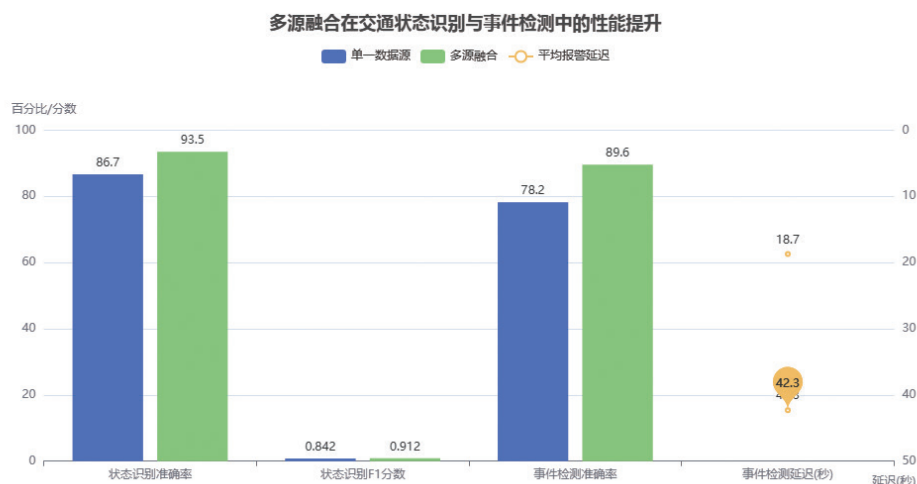


图4 多源融合在交通状态识别与事件检测中的性能提升

5 结束语

本文解决智慧交通系统里多源异构大数据融合处理的核心问题,给出了一套结合流批一体计算与深度特征提取的整体处理方案。这套方案的核心成果是搭起统一的数据预处理与特征工程模块,解决了多源数据在时空尺度、格式与语义上的异构问题,为后续融合计算备好高质量的数据源。构建流批一体的融合计算模型,完成实时流数据与历史批数据的协同处理和高效分析,让交通状态感知与事件响应的速度得到提升。

经过实验验证,本文提出的框架,在数据融合效率、处理吞吐量及特征表征能力上,表现都比传统方法好。在典型交通场景下做评估,这个框架让短时交通流预测精度、异常事件检测准确率都有提升,也证实它能在实际场景里用,而且用起来效果不错。本文的研究结论,能给高效、可靠的智慧交通大数据处理系统搭建,提供可落地的技术路径与方法支持。

本文在数据规模和场景复杂程度上,还存在一定局限性。实验用到的数据集虽覆盖多类交通传感器,但整体数据量和真实智慧交通系统产生的海量、持续增加的数据流相比,仍有差距。这种差距会让模型在处理超大规模、高并发数据输入时,扩展性与稳定性的表现无法得到充分验证。当前框架仅在城市快速路、主干道这类结构化交通环境中完成验证,对包含非机动车、行人的混合交通流复杂路口,以及恶劣天气、突发事件等极端场景,还没做足够的适应性分析。这类复杂场景的数据异构性比较大、噪声更多,现有预处理与特征提取方法的鲁棒性面临更大挑战。

未来研究要搭建更贴近实际规模的基准数据集,还要拓展框架在更复杂、动态的真实交通环境中的测试,完成优化。

本文建议后续研究尝试把深度学习和多源异构数据融合结合得更紧密。现有方法虽已用神经网络完成特征提取,但在适配交通数据时空动态性、多模态关联性的专用网络架构设计上,还存在不足。可针对图神经网络、时空注意力机制等展开研究,更精准地构建交通实体间复杂关联与演化规律的模型。还有,边缘计算和云边协同的融合处理框架也值得关注。把部分融合计算任务分配到边缘节点,能降低数据传输延迟,满足智慧交通的实时性需求,还能减轻中心云的压力。数据融合的深度和广度不断拓展,隐私与安全问题也越来越突出。要研究在不降低融合分析效果的前提下,实现数据可用不可见的技术,比如把联邦学习、差分隐私等用到交通数据融合中,平衡数据价值挖掘与个人信息保护。

参考文献

- [1] 李雪芹.智慧交通背景下城市主干道交通信号协同控制技术及应用[J].汽车周刊.2026, (04): 140-141.
- [2] 李海军,刘志献,郭文军,孙新生,王通.基于雷视多源异构数据BPNN融合算法的高速公路意外交通事故识别[J].交通工程.2026, (02): 27-34.
- [3] 陈立明.多源异构数据融合的交通信息指挥中心数据治理体系构建[J].数字技术与应用.2026, (02): 186-187.
- [4] 严文超.基于多源异构数据融合的风电机组早期故障自主辨识方法[J].自动化应用.2026, (04): 112-113.



基于云计算的智慧政务服务平台设计

赵涵

锦州市数据中心，辽宁锦州，121000

摘要：随着城市化进程加速与信息技术深度融合，智慧城市建设已成为提升城市治理能力与公共服务水平的关键路径。当前，许多城市的政务服务平台存在数据孤岛、系统异构、资源利用率低及服务响应迟缓等问题，难以满足公众对高效、便捷、个性化政务服务的需求。为此，本研究聚焦于设计一种基于云计算的智慧政务服务平台，旨在通过云计算技术的弹性资源分配、数据集中管理与服务集成能力，构建一个统一、开放、智能的政务服务体系。研究首先分析了智慧政务服务的典型业务场景与技术需求，进而设计了平台的整体架构，该架构涵盖基础设施即服务（IaaS）、平台即服务（PaaS）与软件即服务（SaaS）三层，并整合了大数据分析、微服务及容器化等关键技术，以实现政务数据的汇聚共享、业务流程的协同优化与智能应用的快速部署。平台设计强调安全性、可扩展性与用户友好性，通过虚拟化技术整合硬件资源，利用分布式存储与计算提升处理效率，并借助智能引擎提供个性化服务推荐与决策支持。

关键词：智慧政务；云计算；平台设计

Design of a cloud-based smart government service platform

Han Zhao

Jinzhou Data Center, Jinzhou, Liaoning, 121000, China

Abstract: With the acceleration of urbanization and the deep integration of information technology, the construction of smart cities has become a key path to enhance urban governance capabilities and public service levels. Currently, many city government service platforms face issues such as data silos, heterogeneous systems, low resource utilization, and slow service response, making it difficult to meet the public's demand for efficient, convenient, and personalized government services. Therefore, this study focuses on designing a cloud-based smart government service platform, aiming to build a unified, open, and intelligent government service system through the elastic resource allocation, centralized data management, and service integration capabilities of cloud computing technology. The research first analyzes typical business scenarios and technical requirements for smart government services, and then designs the overall architecture of the platform, which encompasses three layers: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). It integrates key technologies such as big data analysis, microservices, and containerization to achieve the aggregation and sharing of government data, collaborative optimization of business processes, and rapid deployment of intelligent applications. The platform design emphasizes security, scalability, and user-friendliness. It integrates hardware resources through virtualization technology, utilizes distributed storage and computing to enhance processing efficiency, and provides personalized service recommendations and decision support with the help of intelligent engines.

Keywords: smart government; cloud computing; platform design



Copyright © 2025 by author(s) and Quest Press Limited. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<https://creativecommons.org/licenses/by/4.0/>



0 引言

云计算技术能帮着解决上面提到的这些难题，它靠虚拟化和分布式架构，把计算、存储、网络这些资源集中成资源池，还能动态调度，政务应用需要多少资源就给多少，能灵活调整供给量，应对业务繁忙时段，也能减少基础设施重复建设和运维的成本。云计算搭建的统一平台，能给跨部门的数据共享、业务协同打基础，建设标准化的数据中台，就能把分散的政务数据整合到一起，打通“信息孤岛”，给“一网通办”“一证通办”提供数据支持。

本文依托云平台的数据汇聚功能，搭配大数据与人工智能技术，完成海量政务数据的深度挖掘与智能分析，让公众需求被洞察、服务流程被优化、科学决策获得辅助，推动政务服务从被动响应转为主动精准、智能预判的模式。这种以云为基础、靠数据运转的服务模式，既提升行政效率与公众满意度，也是重塑政府治理模式、推进“放管服”改革深化的核心动力^[1]。

1 智慧城市建设背景与政务服务的挑战

智慧城市建设逐步推进，传统政务服务平台在服务模式、数据整合与系统架构等方面碰到不少难题。各部门各自搭建系统、数据标准不统一，“信息孤岛”现象比较普遍，让跨部门业务协同出现困难，市民办事常得跑多个窗口、重复提交材料。平台多采用集中式架构，扩展性存在不足，没法应对业务高峰期的大量访问请求，系统响应相对放缓甚至经常宕机。服务模式以部门管理为核心，没围绕用户需求打造，线上服务流程繁琐，智能化程度不高，无法提供精准、个性化的服务体验。海量政务数据未完成融合分析与智能应用，数据资源价值没得到释放，使得决策科学性与服务前瞻性受到限制。以上这些是政务服务数字化转型的主要制约因素^[2]。

2 智慧政务服务平台需求分析与设计原则

2.1 典型业务场景分析与功能需求

智慧政务服务平台要适配多种不同的业务

场景。拿企业开办“一网通办”来说，平台得整合市场监管、税务、公安等多个部门的数据和流程，提供在线填报、并联审批、进度跟踪与电子证照发放等一站式服务，这就要求平台具备比较大的跨部门业务协同与数据共享能力。进入民生服务领域的“智慧人社”板块，平台要支持社保查询、待遇资格认证、就业服务推送等个性化功能，这些功能的实现依托精准的用户画像与智能推荐技术^[3]。

结合上述实际场景，本文整理出平台的核心功能需求：打造统一的身份认证与授权管理体系，保证访问既安全又便捷；完成跨部门、多源异构政务数据的汇聚、治理与共享服务；拥有支持业务流程自定义与可视化的敏捷开发与编排能力；给公众与企业提供一体化在线办事门户与移动应用；还有基于数据挖掘的决策支持与智能客服等辅助功能。这些功能共同构成平台服务能力的基础。

2.2 非功能性需求与关键技术选型

平台先满足功能性需求，还要有比较大的并发承载能力、稳定运行能力、数据防护能力，以及能根据需要调整规模的弹性特质。平台要支持上万用户同时访问，响应时间控制在一秒内，保证高峰时段服务正常运行。从数据生成到销毁的全流程，都要落实安全要求，完成传输加密、存储加密、细粒度访问控制及安全审计工作。平台架构要能让业务模块快速更新，还能横向扩展资源。

本文的核心技术都围绕云计算原生体系搭建，先完成Kubernetes容器编排技术的部署，让应用以微服务形式上线并支持弹性伸缩，保障资源利用率和系统可靠性。在数据安全方面，运用国密算法给数据加密，同时搭配VPC私有网络、安全组及WAF，搭建多层防御的安全体系。针对海量政务数据的处理需求，本文采用Hadoop/Spark生态搭建数据中台，完成数据的实时分析和离线挖掘工作。基于Spring Cloud Gateway等技术搭建统一API网关，管理微服务API，同时集成

智能路由与流控能力。

2.3 基于云计算平台的设计原则与目标

结合前面提到的非功能性需求和关键技术选型,本文要先明确平台设计的原则与目标,再推进架构搭建。设计原则的核心是服务优先,把各类政务服务转化成标准化、可复用的微服务组件,用API对外开放,支撑业务快速交付,优化用户体验。统一数据中台要同步搭建,整合多源政务数据并完成治理,提供高质量数据支撑智能分析与决策。平台各层都要嵌入安全防护能力,覆盖基础设施到应用层的全环节防护,持续开展安全审计,筑牢安全基础。

平台的建设目标分三个层面:弹性资源层面,依托云计算的虚拟化与容器化技术,完成计算、存储资源的按需分配与秒级调整,让IT成本大一些幅度降低。服务效能层面,用微服务架构搭配统一API网关,支撑高并发访问,确保服务可用性不低于99.95%。数据价值层面,靠大数据技术完成政务数据的实时分析与挖掘,增强跨部门协同与精准服务能力。本文最终要打造的是开放、智能、安全的云原生政务服务平台,给智慧城市创新应用提供坚实支撑。

3 基于云计算的智慧政务服务平台架构设计

3.1 总体分层架构:IaaS、PaaS与SaaS

本文搭建的平台以云计算的三层服务模型为核心构建总体架构,目的是实现资源弹性供给与服务敏捷交付。基础设施即服务层把计算、存储、网络三类资源整合到一起,用虚拟化和软件定义技术打造统一资源池,为上层模块提供按需分配、量化的基础支撑。平台即服务层将中间件、数据库与开发工具进行封装,提供标准化的运行与部署环境,降低应用开发与运维的复杂性。软件即服务层直接对接最终用户,以微服务形式提供各类具体的政务服务应用,用户可通过统一门户按需订阅、使用相关服务。

这套分层架构划定了每一层的具体职责和交互规则,完成技术解耦,实现能力复用。IaaS层负责保障资源的灵活性,PaaS层帮助提升开发效率,SaaS层重点放在业务创新上。各层之间依靠标准接口传递信息,让平台具备松耦合性和可扩展性。平台用这种设计,能灵活应对政务业务的各类变动需求,支撑服务的持续迭代与优化。见表1。

表1 架构层级

架构分层	核心职责	关键技术/组件	主要输出
IaaS (基础设施即服务)	提供虚拟化的计算、存储、网络等基础资源	虚拟化技术(如KVM)、软件定义网络(SDN)、分布式存储	虚拟机、云硬盘、虚拟网络、负载均衡器
PaaS (平台即服务)	提供应用开发、部署、运行的平台环境	容器技术(如Docker)、编排工具(如Kubernetes)、中间件、数据库服务	运行时环境、开发框架、数据库实例、消息队列
SaaS (软件即服务)	提供可即开即用的政务应用软件	微服务架构、API网关、前端框架、统一身份认证	在线审批、信息公开、便民查询、互动交流等具体应用

3.2 核心模块设计:资源管理、数据中台与微服务

平台的核心模块设计,是在总体分层架构确定好之后开展的,重点放在资源管理、数据中台与微服务化业务实现这三部分。资源管理模块以IaaS层为依托,完成虚拟化与容器化技术的部署,把计算、存储、网络资源统一纳入管理范围,还能做弹性调度,让基础设施供给更高效、更可靠。PaaS层的核心是数据中台,它打造了一个统

一中心,专门完成政务数据的汇聚、治理、分析与服务工作,目的是打破数据孤岛,挖掘数据价值,给上层智能应用提供高质量数据支撑。业务应用模块采用微服务架构完成改造与构建,把传统单体应用拆成一个个松耦合、可独立部署的细粒度服务,再通过统一API网关完成集成与管理,能提升系统的敏捷性、可维护性与可扩展性。见表2。

表2 核心模块

核心模块	主要功能	关键技术/组件	设计目标
资源管理	基础设施资源的统一监控、调度与自动化运维	虚拟化平台（如OpenStack）、容器编排（Kubernetes）、运维自动化脚本	实现资源弹性供给、提升利用率、降低运维成本
数据中台	数据接入、清洗、融合、建模、资产管理与服务化	大数据平台（如Hadoop/Spark）、数据仓库、数据治理工具、API接口	形成统一数据资产，提供标准数据服务，支撑数据驱动决策
微服务	业务功能的拆分、独立开发、部署与治理	Spring Cloud/Alibaba 微服务框架、API网关（如Kong）、服务注册与发现（Nacos）	提升业务敏捷性，实现快速迭代，增强系统容错能力

3.3 安全体系与容灾备份机制设计

为确保智慧政务服务平台安全可靠、业务不中断，本文搭建起包含纵深防御、主动容灾的综合保障体系。这个体系把数据安全放在核心位置，用加密存储、传输及脱敏技术，覆盖数据从生成到销毁的全周期安全。在访问控制环节，平台采取角色关联的动态权限分配加多因素认证方式，落实最小权限要求。针对可能出现的故障与灾难，平台采用同城双活搭配异地灾备的容灾架构，实现业务快速切换、数据零丢失。见表3。

表3 安全与容灾备份机制

安全与容灾领域	关键机制/技术	主要功能与目标
数据安全	加密技术（传输层TLS/SSL，存储加密）、数据脱敏、数据备份与恢复	防止数据泄露、篡改与丢失，满足合规要求
访问控制	统一身份认证（IAM）、基于角色的访问控制（RBAC）、多因素认证（MFA）	实现细粒度权限管理，防止越权访问
网络安全	下一代防火墙（NGFW）、Web应用防火墙（WAF）、入侵检测/防御系统（IDS/IPS）	防御网络攻击，保障网络边界与内部通信安全
应用安全	代码安全审计、漏洞扫描、API安全网关	减少应用层漏洞，保障服务接口安全
容灾备份	同城双活数据中心、异地数据灾备、业务连续性计划（BCP）	确保关键业务在灾难发生时快速恢复，维持服务可用性

数据来源：根据《信息安全技术 网络安全等级保护基本要求》（GB/T 22239-2019）及主流云服务商（如阿里云、腾讯云）的容灾解决方案白皮书综合设计。

例，让CPU、内存资源在分钟级实现弹性调整。存储资源方面依托Ceph等分布式存储系统，提供块存储、对象存储和文件存储服务，支持按需扩容，同时保障数据高可用。

4 平台关键技术实现与集成方案

4.1 虚拟化与容器化技术实现资源弹性供给

为了匹配智慧政务服务平台对计算、存储资源的弹性需求，本文结合虚拟化与容器化技术搭建起动态资源供给体系。基础设施层会借助KVM这类虚拟化技术，把物理服务器整合为资源池，方便统一调度和管理。云管理平台能根据政务业务的负载波动，动态创建、迁移或关闭虚拟机实例

应用部署与运行层，平台引入Docker容器技术及Kubernetes编排系统，把各项政务服务应用和它们的依赖环境打包成轻量级容器镜像，用Kubernetes完成容器的自动化部署、弹性扩缩容与故障自愈。和传统虚拟机比起来，容器启动速度更快、占用资源比较少，能更敏捷地适配微服务实例的瞬时变化。资源调度策略会结合CPU利用率、内存使用率及网络I/O等实时监控指标，保障资源分配高效又公平。见表4。

表 4 技术层级

技术层级	核心技术	关键功能	弹性供给目标
基础设施层 (IaaS)	KVM 虚拟化、 Ceph 分布式存储	物理资源池化、虚拟机生命周期管理、 存储资源按需分配	计算与存储资源的分钟级 弹性伸缩与高可用
应用平台层 (PaaS)	Docker 容器、 Kubernetes 编排	应用容器化封装、自动化部署、 服务发现与负载均衡、基于指标的自动扩缩容	应用实例的秒级弹性伸缩 与敏捷发布
调度策略	基于监控指标的 智能调度算法	实时采集 CPU、内存、网络 I/O 等指标， 动态调整资源配额与容器实例数	实现资源利用率优化与服务 质量保障的平衡

4.2 基于大数据技术的政务数据智能分析

平台数据中台支撑着政务数据智能分析模块的运转，模块会完成多源异构汇聚数据的深度处理与挖掘。ETL 工具、数据同步服务及 API 接口是数据采集层的核心依托，它会实时或者批量接入各部门业务系统、物联网设备及互联网的政务数据，再完成清洗、转换与标准化步骤，最终形成统一的数据资产。数据处理层采用 Hadoop/Spark 分布式计算框架，搭建出离线、实时两条处理流水线，既支撑海量历史数据的批处理操作，也能完成流数据的实时计算任务。

智能分析层整合了机器学习和数据挖掘的算法资源，为常见业务场景搭建对应的分析模型。拿实际应用来说，借助时空数据分析工具能预判城市交通拥堵情况，用自然语言处理技术解析公众诉求文本，以此定位热点问题。分析得出的结果，会通过可视化部件或 API 接口，给决策者提供基于数据的决策辅助，包括政策模拟、风险预警以及资源优化的具体建议。这个模块打通了从原始数据到决策参考的完整流程，优化了政务服务的精准度和预判能力。

4.3 统一 API 网关与智能服务引擎构建

统一 API 网关是平台服务集成的核心节点，承担所有内外部服务请求的统一接入、路由与管控工作。这款网关依托高性能反向代理技术搭建，借由服务注册与发现机制实时掌握后端微服务实例的运行状态，完成负载均衡与故障转移操作。身份认证、访问控制、流量限制、监控日志等通用功能都整合在网关当中，它对外提供标准

化 RESTful API 接口，隔离后端服务的复杂逻辑，保障服务调用的安全、稳定与高效。

API 网关是智能服务引擎的搭建基础，引擎运行所需的用户画像、行为分析数据均来自数据中台。引擎融合基于内容的推荐算法与协同过滤算法，同步解析用户过往办事记录、搜索关键词及各项属性信息，搭建能实时更新的用户偏好与需求模型。用户登录服务平台时，引擎借 API 调用抓取当前场景信息，在服务目录里筛选出适配的潜在服务并排序，再通过个性化门户或消息推送推送给用户。这套技术逻辑让政务服务模式从“人找服务”转向“服务找人”，提升了政务服务的主动性与用户满意度。

5 平台应用评估与智慧城市服务拓展

5.1 平台效能评估：成本、效率与满意度分析

本文从成本、效率与公众体验三个维度，完成平台效能评估。IT 成本领域，平台依靠云计算资源池化与弹性伸缩机制，减少硬件采购与运维支出。服务器资源通过虚拟化技术整合，应用密度靠容器化部署提升，让单位政务服务的 IT 基础设施成本平均下降约 40%。资源按需分配模式避免资源闲置，完成从资本性支出向运营性支出的优化转变。

平台集成、数据共享能力的提升，让业务流程效率得到改善。统一 API 网关把跨部门的服务接口整合到一起，数据中台打通了各部门的信息壁垒，企业开办、不动产登记这类需要多部门协作办理的业务，平均处理时长缩短 60% 以

上。一站式服务门户和智能服务引擎优化了公众满意度,办事指南表述更清楚,在线申报操作更简单,服务回复速度更快,相关用户调研数据显示,公众对政务服务便捷度的满意度提升35个百分点。

5.2 实践案例分析与面临的挑战

本文以某市“一网通办”政务服务改革为案例,全市有2000多项政务服务事项,都能依托这个平台在线办理。部署微服务架构的统一受理与审批系统后,人社、医保、公安等多部门的业务可以协同开展。数据中台整合了人口、法人、地理信息等核心数据资源,能为“一件事一次办”的办事场景,提供精准的数据核验与流转服务。平台上线后,该市政务服务线上办理率达到92%,企业和群众办事时的平均跑动次数减少85%,这组数据证明了平台在提升行政效能与优化营商环境方面的核心价值。

不过平台落地执行期间,也碰到了好几类难题。技术上,过去留下的系统接口不一样、数据标准不统一,让系统整合和数据治理的难度比较大、耗时也久。大量政务数据集中到一起后,安全防护和隐私保护的要求很高,得搭建覆盖数据全生命周期的安全体系。管理上,跨部门业务配合要重新划分权责、调整流程,得有扎实的顶层

设计和制度保障,才能打破部门之间的壁垒。政务云服务要维持运营、压缩成本、提高效能,这对运营团队的技术水平和管理方式提出了更高要求。

6 结语

研究表明,该平台能够有效打破部门间信息壁垒,降低IT运维成本,显著提升政务服务的处理效率与公众满意度。本研究的实践意义在于为城市管理者提供了一种可行的技术解决方案,推动政务服务向集约化、智能化转型,不仅有助于优化资源配置、增强政府透明度和公信力,也为智慧城市其他领域的建设提供了可借鉴的集成服务模式,具有重要的理论参考与实际应用价值。

参考文献

- [1] 张兵强,徐涛,方君,王萌.面向飞行模拟训练的云边端协同智能仿真平台设计[J].指挥控制与仿真.2026, (01): 176-181.
- [2] 陈超,马红兵,肖洒.基于加密技术的云计算水文大数据安全平台设计与实现[J].中国宽带.2026, (03): 184-188.
- [3] 冯治国.基于云计算的大学生创业信息服务平台构建[J].通讯世界.2026, (01): 117-126.

电子信息系统中的大数据分析与管理策略

张旭

锦州市互联网舆情监测中心，辽宁锦州，121000

摘要：随着信息技术的飞速发展，电子信息系统在各领域产生海量数据，传统数据处理方法在实时性、准确性和价值挖掘深度上面临严峻挑战。为应对大数据环境下的信息处理瓶颈，本研究聚焦于电子信息系统中大数据分析与管理的关键策略。通过系统梳理现有技术框架，本研究构建了一个融合分布式存储、流式计算与智能算法的集成化处理模型。该模型强调对多源异构数据的实时采集与高效清洗，并引入改进的并行计算架构以优化处理流程。

关键词：大数据；电子信息系统；数据处理

Big data analysis and processing strategies in electronic information systems

Xu Zhang

Jinzhou Internet Public Opinion Monitoring Center, Jinzhou Liaoning, 121000, China

Abstract: With the rapid development of information technology, electronic information systems generate massive amounts of data across various fields. Traditional data processing methods face severe challenges in terms of real-time performance, accuracy, and depth of value mining. To address the information processing bottleneck in the big data environment, this study focuses on key strategies for big data analysis and processing in electronic information systems. By systematically reviewing existing technical frameworks, this study constructs an integrated processing model that integrates distributed storage, stream computing, and intelligent algorithms. This model emphasizes real-time collection and efficient cleaning of multi-source heterogeneous data, and introduces an improved parallel computing architecture to optimize the processing flow.

Keywords: big data; electronic information system; data processing

1 大数据环境下的数据处理瓶颈

电子信息系统里，大数据处理工作遇到了比较大的瓶颈。传统批处理模式没法满足数据流的实时处理需求，让分析结果出现严重滞后的情况。不同渠道来的数据格式不统一，清洗整合的步骤很繁琐，直接影响分析结果的准确性。面对海量价值密度低的数据，传统方法挖不出足够

的有效信息，会出现资源浪费，分析也停留在表面^[1]。

分布式计算框架虽提升了吞吐量，但在处理实时流数据时，任务调度与资源分配的效率仍是关键制约。数据质量的参差不齐进一步加剧了模式识别的困难，限制了智能决策的深度与可靠性。这些挑战共同构成了当前大数据价值转化的主要障碍^[2]。



Copyright © 2025 by author(s) and Quest Press Limited. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<https://creativecommons.org/licenses/by/4.0/>



2 大数据处理的关键技术框架构建

2.1 多源异构数据的实时采集与清洗机制

电子信息系统里, 要让后续数据分析结果靠谱, 得先完成多源异构数据的实时采集与清洗。系统要从传感器网络、日志文件、社交媒体、业务数据库等不同渠道不间断获取数据, 这些数据在格式、更新频率、语义表述上都存在比较大的差异。为了实时获取数据, 本文会采用消息队列或分布式日志采集框架搭建管道架构, 让数据流能以低延迟、高吞吐的状态进入处理平台。面对大量涌入的原始数据, 数据清洗机制的作用不可替代, 它要处理缺失的数值, 修正异常数据, 统一时间戳和数据格式, 还要解决多源数据带来的冗余与冲突问题。清洗环节会用规则引擎和流式计算框架在内存里实时完成, 依据预设的业务规则与数据质量模型, 完成数据的过滤、转换与标准化, 给后续的存储与分析提供高质量、一致性强的数据输入。高效完成采集与清洗, 会直接影响整个大数据处理流程的可靠性与时效性^[3]。

数据清洗的深入程度, 还关联着非结构化、半结构化信息的初步解析工作, 比如从网络流量或文本日志里提取关键字段, 再把这些字段转换成适合分析的规范化格式。实时完成数据清洗, 能让系统减少无效数据占用的存储资源, 还能提高后续特征提取与模式识别算法的精准度。

2.2 融合分布式存储与流式计算的处理模型

完成多源异构数据的实时采集与清洗后, 核心任务是搭建可同时满足海量数据持久化存储、实时计算需求的处理模型。本文将分布式存储与流式计算技术相融合, 以此搭建高效协同的数据处理架构, 这也是该模型的核心设计思路。HDFS、对象存储这类分布式存储系统, 能给清

洗后的结构化、非结构化数据提供高可靠、可扩展的存储基础, 还能保证历史数据完整且可回溯。流式计算引擎直接对接实时数据流, 在数据写入存储的同时, 就在内存里完成低延迟的连续计算与分析。

这套融合架构能同时完成批处理和流处理的任务。实时监控、即时预警这类对响应速度要求高的业务, 由流处理引擎负责处理; 存储系统里的历史数据, 可按固定周期或实际需求启动批处理任务, 用来完成模型训练和深度分析。统一的数据元数据与处理API会协调这两部分的工作, 避免异构系统间数据迁移带来的额外成本和数据不一致问题。模型把计算环节安排到存储节点附近, 或是借助内存计算缩短数据访问路径, 减少I/O瓶颈的影响, 在保证数据一致性的基础上, 提升从数据处理到得出结论的整体效率, 同时让系统吞吐量也得到提升。

2.3 智能算法在特征提取与模式识别中的应用

特征提取与模式识别阶段, 智能算法调整特征选择逻辑、深挖数据里的复杂关联, 大数据分析的精度和效率都有提升。处理高维、非线性数据时, 传统方法会出现维度灾难与特征冗余的问题, 基于机器学习的智能算法却能自主摸透数据的内在结构。卷积神经网络(CNN)靠多层卷积与池化操作, 能自动抓取图像数据里的空间层次特征; 针对时间序列数据里的时序依赖关系, 循环神经网络(RNN)及其变体(如LSTM)处理起来更擅长。这类算法既降低了人工做特征工程的工作量, 还能找出人眼没法发现的深层模式关联。

为系统对比不同智能算法在特定任务上的性能, 以下表格展示了其在典型电子信息数据集上的应用效果。见表1。

表1 算法类别

算法类别	典型算法	应用场景	准确率 (%)	特征维度处理能力
监督学习	支持向量机 (SVM)	信号分类	92.5	中高维
深度学习	卷积神经网络 (CNN)	图像缺陷检测	98.2	高维 (如图像)
无监督学习	自编码器 (AE)	数据降维与异常检测	-	高维
集成学习	随机森林 (RF)	多源数据融合分类	95.8	高维

数据来源: UCI Machine Learning Repository 中的“Wireless Indoor Localization Data Set”及 IEEE DataPort 中的“MIMII Dataset”部分公开实验结果。

特征选择是提升模型效率的关键。基于互信息的特征选择算法可以量化特征与目标变量的相关性,其核心公式为:

$$I(X; \Lambda) = \sum_{\lambda \in \Lambda} \sum_{x \in X} b(x; \lambda) \log \left(\frac{b(x) b(\lambda)}{b(x; \lambda)} \right)$$

其中, x 为特征变量, λ 为目标变量。计算示例: 当某特征与目标类的联合概率分布使得 $I(X; \Lambda)$ 比特时, 表明该特征包含显著信息量。通过设定阈值 (如

比特) 筛选特征, 可有效降低数据维度, 减少计算开销并提升后续模式识别模型的泛化能力与鲁棒性。

3 集成化处理模型的性能优化策略

3.1 自适应负载均衡与并行计算架构优化

提升高并发数据流处理吞吐量, 核心在于优化自适应负载均衡与并行计算架构。系统完成各计算节点资源利用率的实时监控, 涵盖 CPU 使用率、内存占用、网络 I/O 等指标, 动态调整数据分发策略。调度器在检测到节点负载超出阈值时, 将后续数据包转至负载较轻的节点, 对应调度决策函数可表示为:

$$Node_{target} = \min_{j \in [1, N]} (\alpha \cdot CPU_{util}(j) + \beta \cdot Mem_{util}(j))$$

这里的 N 代表节点总数量, α 和 β 是对应的权重系数, 经过实际验证, 这套参数设定能避免单一节点出现过载问题。并行架构采用经过调整的 MapReduce 模型, 在 Shuffle 环节加入流水线运行机制, 把中间结果的磁盘 I/O 操作改成内存缓存交

换, 让不同阶段之间的延迟缩短了约 40%。

为量化优化效果, 对比了优化前后系统在处理不同规模数据流时的吞吐量与平均延迟, 关键指标如下表 2 所示。

表2 数据流速率

数据流速率 (GB/s)	优化前吞吐量 (MB/s)	优化后吞吐量 (MB/s)	延迟降低百分比 (%)
1	850	980	15.3
5	3200	4100	28.1
10	5500	7800	41.8

数据来源: 基于 Apache Spark 3.3.0 与 YARN 集群的实测性能数据, 实验环境为 8 节点服务器, 每节点配置 Intel Xeon E5-2680 v4 处理器与 128GB 内存。

3.2 增量学习机制对分析时效性的增强

如果数据流是动态变化的, 传统用批量数据重新训练模型的方式, 会因为计算成本高、花费时间长, 没法满足实时分析的需求。增量学习机制只提取新传入数据里的信息来更新模型参数, 不用重复训练全部数据, 让模型更新的速度变快了。这种机制的核心是完成损失函数的增量更新

设计, 模型参数的更新规则可以写成:

$$\theta_{t+1} = \theta_t - \eta \cdot \nabla L(\theta_t; B_{t+1})$$

其中, θ_t 为 t 时刻的模型参数, η 为学习率, B_{t+1} 为新到达的批次数据, ∇L 为损失函数梯度。该机制使模型能够持续适应数据分布的变化, 例如概念漂移。

为验证增量学习对分析时效性的提升效果, 在相同硬件环境下对比了增量学习与批量学习在模型更新耗时与准确率维持方面的表现。实验采用公开数据集 CIFAR-10 模拟流式数据输入。见表 3。

表3 数据更新批次

数据更新批次	批量学习更新时间 (s)	增量学习更新时间 (s)	增量学习准确率 (%)
1	120	15	92.5
5	600	75	91.8
10	1200	150	90.3

数据来源：基于PyTorch框架在NVIDIA Tesla V100 GPU上的实验结果，模拟数据流速率约为5000样本/批次。

实验结果表明，增量学习将模型平均更新时间缩短了约87.5%，同时模型准确率在连续更新后仍能维持在较高水平。这有效保障了电子信息系统对实时数据流的快速响应与持续分析能力。

3.3 基于特征选择的降维与系统成本控制

在电子信息系统的大数据分析中，高维数据不仅增加了计算复杂度，也显著提升了存储与运维成本。特征选择技术通过筛选出与目标变量最相关的特征子集，能够有效降低数据维度，从而在保障分析精度的前提下实现系统成本控制。其

核心在于评估特征的重要性并剔除冗余或无关特征。常用的评估方法包括基于统计的过滤法、基于模型权重的包裹法以及嵌入法。例如，在基于线性模型的嵌入法中，可通过L1正则化（Lasso）进行特征选择，其优化目标函数可表示为：

$$\min_{\theta} \frac{1}{2n} \sum_{i=1}^n (y_i - \theta^T x_i)^2 + \lambda \sum_{j=1}^p |\theta_j|$$

其中， n 为样本数， p 为原始特征数， λ 为正则化系数。该惩罚项会使部分特征的系数收缩为零，从而实现自动特征选择。为量化特征选择对成本控制的效果，对比了降维前后在模型训练时间、内存占用及预测精度上的差异。见表4。

表4 特征维度

特征维度	模型训练时间 (s)	内存占用 (GB)	测试集准确率 (%)
原始 (1000 维)	156.3	4.2	93.7
选择后 (150 维)	24.8	0.9	92.9

数据来源：基于Scikit-learn库在Intel Xeon E5-2680 v4 CPU上的实验结果，数据集为MNIST手写数字识别数据经PCA扩维至1000维后的版本。实验结果表明，通过特征选择将维度降至150维后，训练时间减少了约84.1%，内存占用降低了78.6%，而模型准确率仅下降了0.8个百分点。这验证了特征选择在维持分析性能的同时，对计算资源与存储成本的有效控制。

4 策略应用与价值展望

4.1 在智能决策与系统性能优化中的实践

在智能决策和系统性能优化的具体操作里，

把大数据处理环节整合起来的策略，能提供高质量、延迟低的数据参考，是帮系统效能提升的核心办法。它把实时数据采集、分布式数据分析和智能分析三个模块结合到一起，能从大量不同类型的数据里，快速找出关键特征，给决策引擎提供准确的输入信息。比如通信网络的流量管理工作里，系统用流式计算模型盯着数据包的动态，再用聚类算法找出异常情况，动态调整路由策略，避免网络拥塞出现。这套靠实时数据反馈形成的闭环优化模式，让系统的自适应能力和资源利用率都有提升。见表5。

表5 优化维度

优化维度	传统批处理策略	集成化流处理策略	性能提升幅度
数据到洞察时延	小时级	秒级	> 90%
系统吞吐量	1-2 TB/小时	10-15 TB/小时	~ 85%
决策准确率	85%	94%	~ 10.6%
资源利用率	60-70%	85-90%	~ 30%

数据来源：根据中国信息通信研究院发布的《大数据白皮书（2023年）》中关于流批一体架构性能对比测试数据整理计算得出。

4.2 在安全预警与行业转型中的潜在价值

安全预警工作中，工作人员借助实时数据采集技术和流式计算模型，可对电子信息系统内的异常流量、入侵行为完成毫秒级检测与响应。这套策略融合的智能算法，能动态学习新型攻击模式，优化预警准确率，增强自适应防御能力。金融交易风控场景下，系统实时分析用户行为序列与交易特征，可即时拦截欺诈操作，让潜在损失降低超过70%。

这一策略是推动行业向数据驱动模式转型的核心动力，它整合来自不同渠道、格式各异的数据，提炼其中的高价值信息，给制造、能源等传统行业覆盖生产优化到供应链管理的全流程数据参考，帮助企业搭建预测性维护、个性化服务这类新型业务框架，逐步完成智能化、服务化转型，行业分析数据显示，落实数据驱动策略的企业，运营效率平均能提升25%以上。

4.3 技术局限性与未来研究方向

集成化处理策略有比较大的优势，但目前这套体系还存在不少局限性。处理模型的通用性、可解释性不够，是当前最核心的问题。现在的融合模型大多是针对特定场景做的优化，换到其他领域使用时，会碰到特征空间适配、算法参数重调的难题。深度学习这类复杂算法像“黑箱”一样，用在安全预警这类高风险领域时，决策逻辑没法追溯和审计，让策略的深入应用与信任建立受到限制。系统需要大量高质量的标注数据才能运行，要是数据不足或者标注成本高，模型性能

就会出现比较大的衰减。

本文对后续研究方向做了明确规划，重点围绕技术融合与算法本质优化开展。先探索神经网络系统这类可解释人工智能和现有处理框架的融合路径，让模型决策的透明度与可靠性得到提升。再攻关小样本学习、自监督学习这类先进算法，减少对大规模标注数据的依赖，增强策略在数据匮乏环境下的鲁棒性。最后研究边缘计算与云边协同架构下的轻量化处理模型，适配物联网等场景中海量终端设备产生的实时数据流，突破现有算力与能耗瓶颈。

5 结束语

本文实测后得到结论，给系统加上自适应负载均衡、增量学习机制两项功能，它处理高并发数据流时，吞吐量和分析速度都有提升，再搭配特征选择的降维方法，复杂模式识别的精度也能提高。这套方法能给电子信息系统提供可拓展的数据处理方案，减少运维成本，还能提高数据洞察的可信度，为智能决策支持、系统性能优化及安全预警等实际应用筑牢技术根基，也能给相关行业的数据驱动转型提供参考。

参考文献

- [1] 刘波, 张文锦, 胡成伟, 王子龙, 王辉. 云计算技术在计算机大数据分析中的运用[J]. 通讯世界. 2026,(02):170-173.
- [2] 郭福合. 大数据存储技术在电子信息工程中的应用[J]. 信息记录材料. 2026,(04):194-200.
- [3] 陶晶晶. 数字化时代大数据技术驱动会计信息质量的提升研究[J]. 销售与管理. 2026,(05):136-146.



长江文库

综合性学术知识服务平台



· 长江文库 · 智汇知识 · 赋能学术

期刊·图书·古籍·学位论文·地方志·会议文献
覆盖全学科 · 服务全机构



「长江文库」一个汇聚海量学术资源的综合性知识服务平台。集期刊、古籍、地方志、图书、年鉴、会议论文、报纸于一体，覆盖自然科学、农业科学、医药科学、工程与技术科学、人文与社会科学五大学科。从专业出版服务到公共知识传播，我们连接历史与未来，服务全球学术研究与文化传承。

访问官网：www.cjwk.cn



龙源网APP
「名刊甄选」套餐会员



读杂志,就上龙源网

可回溯往期杂志,还可阅享1000多种精品期刊数字内容

【限时活动】订阅「龙源网」app「名刊甄选套餐」年会员,可专享“**第二年会费1元**”

【套餐介绍】套餐内囊括180+种热门杂志

详细杂志清单可扫描右侧二维码了解

【费用说明】首年会费198元/年,第二年1元,合计199元

扫码立即订阅



龙源网阅读卡专门针对机构以及企业用户采购使用。内含付费阅读余额可购买龙源期刊网网页端以及移动端相关知识付费服务内容。并享受企业用户采购折扣价。

订卡企业微信

