

一体化政务信息传播平台建设实践^①



陈 雄, 杜义华, 王闰强

(中国科学院 计算机网络信息中心, 北京 100190)
通讯作者: 陈 雄, E-mail: chenxiong@cashq.ac.cn

摘 要: 结合政务网站集约化建设要求, 针对网站规模不断扩大提出一种基于微服务架构的一体化信息传播及管理解决方案, 以统一组织机构、共享稿件、资源标识、系统接口为平台公共基础服务, 以数据资源全集为平台公共数据服务, 实现信息从生产、编辑加工到网络传播、绩效统计的全流程管理, 形成以组织机构为基本管理单元的一体化政务信息传播管理平台。

关键词: 政务信息传播; 一体化; 内容管理

引用格式: 陈雄, 杜义华, 王闰强. 一体化政务信息传播平台建设实践. 计算机系统应用, 2021, 30(4): 82-87. <http://www.c-s-a.org.cn/1003-3254/7863.html>

Construction Practice of Integrated Government Information Communication Platform

CHEN Xiong, DU Yi-Hua, WANG Run-Qiang

(Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: As required by intensive construction of government websites, an integrated information communication and management solution is proposed, which is based on the micro-service architecture against continuous expansion of websites. It serves for the public base of the platform with unified organization, shared manuscripts, resource identifiers and system interfaces and supports the public data with a complete set of data resources. Thus the whole process management of information is realized, from producing, editing, and processing to network communication and performance statistics. Then an integrated government information communication platform is formed, with an institutional framework as the basic management unit.

Key words: government information communication; integrated; content management

1 引言

随着 Web 3.0 时代到来, 基于移动互联网的微博、微信已大规模应用, 政府机构纷纷开通官方微博和微信, 同时抖音、快手等短视频应用也迅猛发展, 部分官方机构已陆续入驻开通账号. 新时代政务信息门户发布已不局限于传统网站, 还包括微博、微信及短视频等各类第三方自媒体渠道.

2017 年, 《政府网站发展指引》^[1] 发布, 国家对于政务网站集约化要求更加清晰. 按照统一标准体系、

统一技术平台、统一安全防护、统一运维监管原则, 集中管理信息数据, 集中提供内容服务, 实现政府网站资源优化融合、平台整合安全、数据互认共享、管理统筹规范、服务便捷高效, 以此实现集约化是解决政府网站“信息孤岛”等问题的有效途径.

对于包含多级组织机构、需运行多个网站和新媒体账号的政务机构, 按照集约化原则建设信息传播平台, 既是机遇也是挑战: 一方面, 可以优化重组各类信息系统, 进一步梳理优化政务信息传播环境; 另一方面,

① 基金项目: 智慧中科院建设推进工程 (二期)(XXH13504-04)

Foundation item: Smart CAS Construction Promotion Project (Phase II)(XXH13504-04)

收稿时间: 2020-08-12; 修改时间: 2020-09-03; 采用时间: 2020-09-11; csa 在线出版时间: 2021-03-30

面临信息数据规模增大、多组织信息快速共享及网络安全防护等压力。

2 现状分析及问题

按照集约化建设原则,部分机构已建设新媒体集成发布平台^[2],实现系统的统一认证及各媒体间内容关联和分享,初步解决多渠道信息发布和统一集成管理的问题。对于存在多层级隶属关系的组织机构体系,各机构按要求建设网站,结合自身需求已开设中英文网站、移动APP、微博、微信公众号、手机报、电子杂志等。但随着组织机构发展,网站及新媒体各类账号规模不断增大,各业务系统间接口调用频繁、耦合度高,在实际应用中存在各种问题。

(1) 业务系统间接口耦合

各业务系统间数据推送等交互功能为系统间接口直接调用,不能控制系统间接口调用权限,监测各系统接口状态。当业务系统调整接口参数或进行业务升级时,其他相关业务系统需要相应调整调用方式,系统间耦合度高,需提供相对独立的接口网关服务,统一管理接口调用、状态及权限等。

(2) 各业务系统中组织机构关系体现

由于网站、微博、微信公众号等内容管理系统相对独立运行,机构在同时开设中英文网站、新浪官方微博官方账号、微信订阅号的情况下,各业务系统间对同一机构网站或新媒体无法关联体现属于同一组织机构,无法方便快捷统计机构下站点、新媒体账号等运行情况;多级组织机构下各机构网站或新媒体,在业务系统中无法体现隶属层级关系。因此,需在平台中提供统一的组织机构管理和开放接口,供各业务系统关联调用。

(3) 发布信息多渠道流转查询

信息经过二次编辑以不同形式在各渠道发布后,在各业务系统中只能查看信息最终处理结果,无法查看信息在各业务系统的具体流转情况,包括信息的相关操作用户、所属业务系统以及相关处理过程等。因此,需在各业务系统中提供稿件统一跟踪查询功能,查询展示稿件在平台中详细流转过程。

(4) 多渠道传播效果综合统计

传播效果统计包括信息浏览量、访客数、点赞数、评论数等常用数据,以及机构信息更新量和机构间信息推送相关统计数据。各渠道传播数据经过汇总后只能分开独立查看,不能查询信息在各渠道传播效

果的综合结果,也不能基于机构整合所有渠道的传播效果进行综合排名。因此,需在系统中提供按信息、机构等维度统计的各个渠道综合传播效果统计功能。

(5) 传播信息类别统一标识

数据在各业务系统中有不同信息类别标识,每个业务系统根据自身功能需求有不同分类标准,网站通过栏目类别标识,微信、微博等数据通过信息分类标识,通过信息资源池汇聚后,不能对汇总数据按照统一标准进行类别标识,不利于进一步数据分类统计和分析。因此,需在系统中提供统一规范的资源标识功能,供各业务系统按标准对信息资源标记分类。

3 目标与整体设计

3.1 工作目标

针对现有平台运行中的各类问题,需基于现有平台调整平台运行业务流程,以统一组织机构、共享稿件、资源标识、系统接口为平台公共基础服务,以数据资源全集为平台公共数据服务,实现信息从生产到传播、发布再到分析的全流程管理,形成一体化的政务信息传播管理平台。

平台建设引入松耦合的微服务架构,构建平台公共基础服务,建设统一组织机构库,实现平台组织机构信息统一管理和接口开放;建设共享稿件库,为业务系统间数据推送提供统一消息队列服务;建设资源标识库,为业务系统提供一致数据分类标准;建设平台公共接口网关,作为业务系统接口调用的统一入口和分发路由,实现接口统一监控、权限校验和审计;建设数据资源全集,汇聚包括文档、图片、视频、文件、站点、栏目、用户、新媒体账号、日志等各类业务数据,实现海量数据统一查询分析。在不改变业务系统已有功能前提下,进行系统功能接口改造,统一与接口网关进行对接,实现系统间的松耦合和高可用。

3.2 整体架构

在业务系统上层构建接口网关,统一提供接口调用的路由服务、访问控制服务和状态监控服务,各业务系统接口调用统一由接口网关进行封装整合、解析调度和路由分发,确保可实时监控各接口运行状态;组织机构库、共享稿件库、资源标识库作为平台公共基础服务,为各项业务系统开放接口,提供包括机构查询检索、稿件推送和跟踪、资源标识查询等远程调用服务;数据资源全集作为平台公共数据服务,定时抽取各

业务系统及公共基础服务的运行数据。

如图 1, 用户统一认证、各类业务系统和公共基础服务, 由原来系统间相互直接调用方式, 调整为通过公共接口网关进行相互调用. 对于接口调用方来说, 公共接口网关隐藏各个接口具体细节, 聚合成粗粒度业务接口; 对于接口被调用方来说, 公共接口网关集成接口

调用的负载均衡、权限验证、参数过滤及安全防护等通用功能, 使接口实现专注于业务本身, 并为业务功能升级或接口改造时不停止服务提供可能. 通过引入公共接口网关, 对于系统间的相互调用, 形成了统一调度方式、统一安全策略、统一监控和日志记录, 进一步降低系统间耦合度.

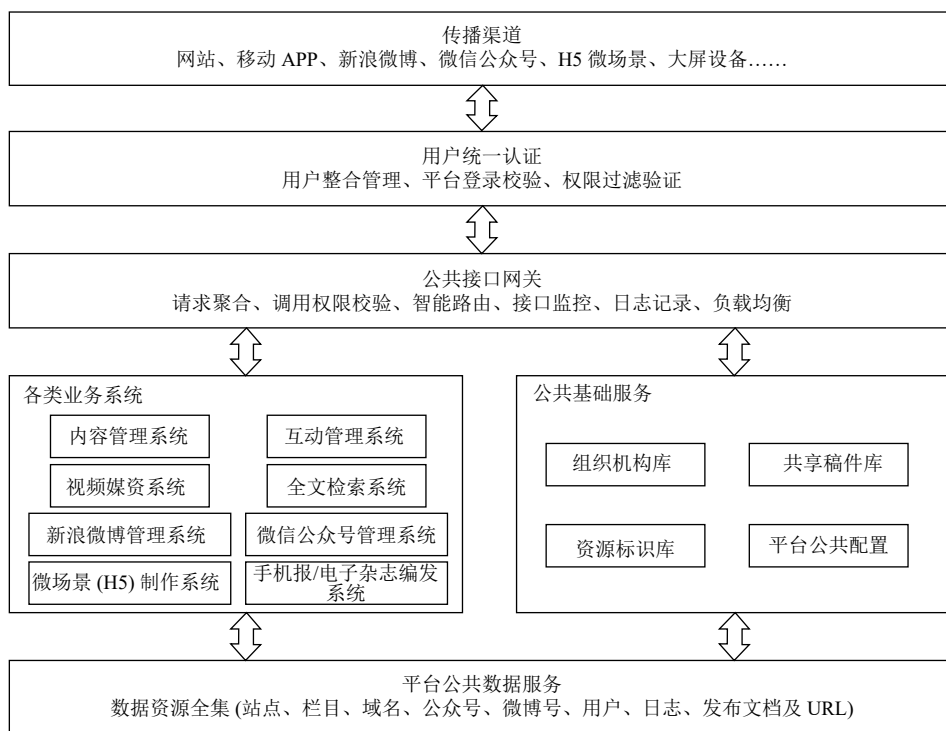


图 1 系统架构图

4 功能设计实现

通过新建公共接口网关、组织机构库、共享稿件库、资源标识库、数据资源全集等公共服务系统, 并对已有业务系统进行局部功能调整, 建设实现一体化政务信息传播平台。

4.1 公共接口网关

系统采用 Spring Cloud^[3] 开源项目实现公共接口网关对接口的统一管理调用, 各系统间功能调用均通过接口网关进行调度. 主要包括请求聚合、权限校验、智能路由、接口监控、日志记录等功能。

请求聚合指在各业务系统接口服务基础上, 按照各项业务功能对业务系统最细颗粒度接口进行功能聚合, 确保一次接口调用完成一次业务功能, 提升网关的响应能力和速度; 权限校验指在通过网关调用接口时, 会统一过滤验证调用方的权限和参数合法性, 并在向被调用方验证通过后放行; 智能路由指在调用各业务

接口时, 隐藏接口详细信息, 统一由网关将接收到的请求按照注册中心配置进行转发; 接口监控指网关会对所有接口进行实时监控, 包括接口的状态、实时流量、调用情况等, 对于恶意调用可及时进行接口限流或临时熔断; 日志记录会详细记录每次接口调用详细情况, 包括调用方的详细请求信息、调用结果等, 用于异常监控和优化分析。

结合新媒体集成发布平台运行情况, 以不改变业务系统功能架构为前提, 仅对各业务系统开放接口部分进行功能改造, 如图 2 所示, 将各业务系统开放接口作为独立微服务, 所有接口注册后统一由接口网关进行代理和路由, 既提升了平台扩展性和稳定性, 也减少了运维过程中的复杂度和工作量。

4.2 组织机构库

组织机构库是平台中统一管理机构下所有组织信息的公共基础系统, 提供对组织机构的管理维护、查

询检索、查询接口等功能,平台中网站、微博账号、微信公众号、手机报等业务系统均以组织机构为单位进行关联管理,是平台的核心功能系统。

组织机构库管理维护机构信息主要包括机构 ID、机构名称、机构全称、机构代码、机构类别、状态、父机构代码、机构图标、所属区域、机构地址、联系人相关信息、备注说明等字段;系统本身提供组织机

构树形结构的相关管理功能,包括数据增删改查、图片上传、顺序调整等常用功能;系统同时提供对组织机构查询的开放接口,接受组织机构 ID、关键字、区域等参数进行远程调用查询,所有业务系统都和组织机构库的组织机构进行关联,每个网站、新浪微博、微信公众号、手机报及其它业务均以组织机构为基本管理单元。

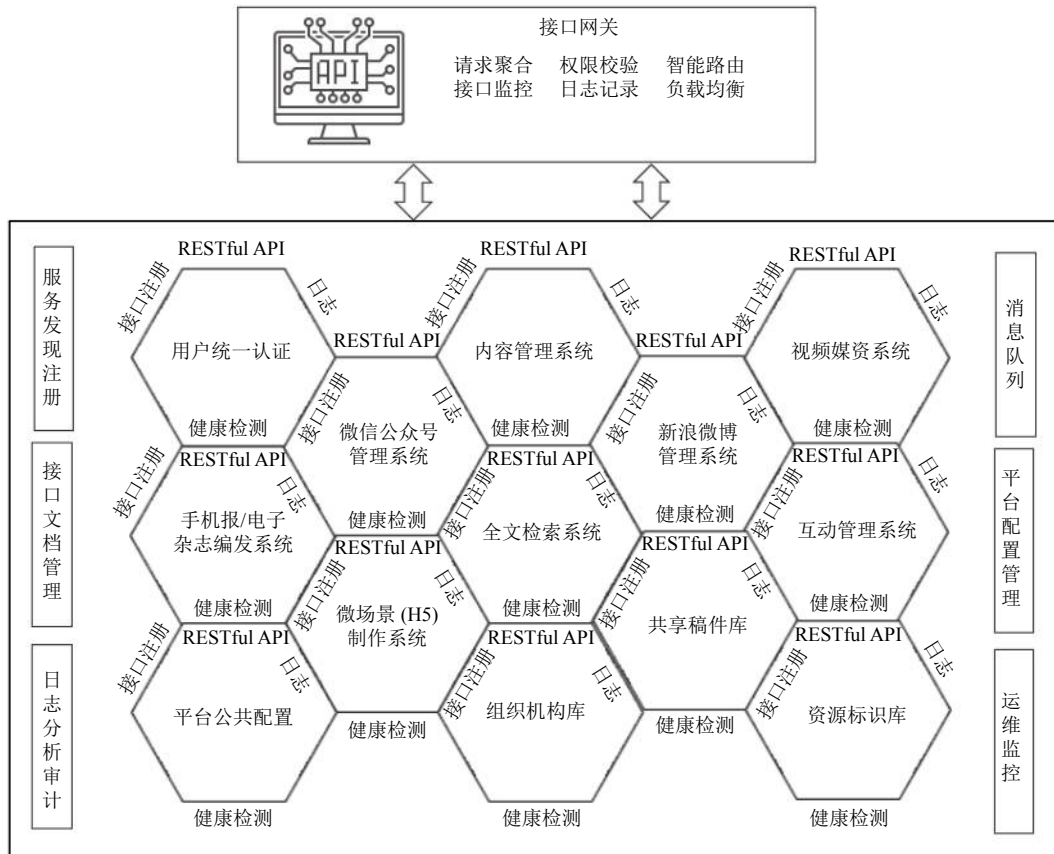


图2 微服务架构示意图

4.3 共享稿件库

共享稿件库是业务系统间进行数据推送的消息队列系统,采用点对点通信模式实现,主要为各业务系统提供数据相互推送的中转、权限控制和共享稿件查询等服务,包括稿件推送、推送权限配置、稿件跟踪、推送稿件查询、开放接口管理等功能。

稿件推送指业务系统间数据共享,实现各系统间双向信息推送.例如信息在内容管理系统中发布到网站后,可将信息推送到本机构或上级机构的微信管理系统中,微信可对稿件进行二次编辑后进行发布.推送权限配置指不同层级机构间的推送关系配置,平台中

数据缺省可在同一机构下不同业务系统间相互推送,对于不同机构间数据推送,需要配置机构到机构的推送权限;稿件跟踪指对于已推送稿件可通过共享稿件库查询稿件的详细流转情况,包括来源系统、推送用户及时间、目的系统、稿件处理情况等;推送稿件查询指共享稿件库本身提供按机构查询展示所有推送稿件,并提供基本的统计分析;共享稿件库提供稿件推送、附件上传、稿件跟踪、稿件查询等开放接口。

如图3所示,网站、新浪微博、微信公众号都以组织机构为基本管理单元.共享稿件库中各组织机构下的稿件库逻辑上相互独立,通过共享稿件库进行机

构到机构的推送权限配置,推送稿件前先进行权限过滤验证,通过后将稿件推送到对应的消息推送队列,业务系统中的接收机构从消息推送队列中获取稿件信息,并在共享稿件库中存储稿件进行记录和查询。

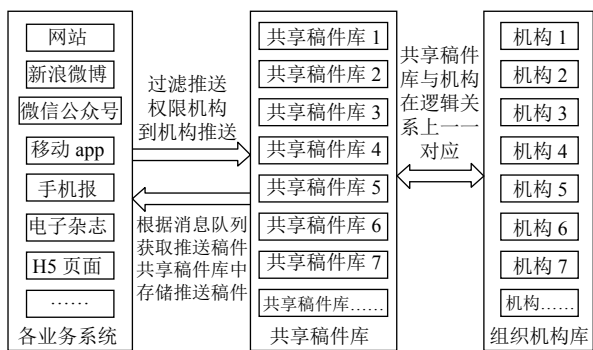


图3 推送逻辑示意图

4.4 资源标识库

资源标识库是对平台所有业务信息进行统一标识分类的公共基础服务,主要提供资源标识分类管理和开放接口功能。平台中包含各种结构的业务信息,通过统一数据分类标准进行规范,便于数据汇总后综合分析。

资源标识管理指可对平台各类资源目录进行标识管理,制定统一标准的资源目录,包括站点、栏目、文档、图片、视频等各类信息,均制定平台公共资源分类,实现常用的增删改查等管理功能;同时提供标识查询开放接口,接受标识ID、关键字等参数进行远程调用查询。

4.5 数据资源全集

数据资源全集是对平台各类业务信息进行统一过滤汇聚的全文检索系统,汇聚分散在各业务系统的站点、栏目、域名、公众号、微博号、用户、日志、发布文档及URL等各类数据,主要功能包括各类数据查询、数据流转分析、传播效果统计等功能。数据资源全集采用开源ETL工具Kettle^[4]实现数据的抽取和汇聚,采用开源搜索分析引擎Elasticsearch^[5]建设分布式的数据资源存储和索引。

各类数据查询指按组织机构对站点、栏目、文档、用户、图片、视频、日志等各类结构化数据的综合查询,包括各类数据的详细信息;数据流转分析指可对文档信息在平台中的所有流转情况进行关联查询分析,通过汇聚平台所有文档信息,可查看文档在各个渠

道的发布情况、编辑版本、推送采纳情况等,对后续的内容运维具有指导作用;传播效果统计指按照组织机构对各个发布渠道提供包括文档浏览量、访问数、点赞数、评论数、更新量、收录量等各类传播效果的综合统计分析,为各机构的信息传播绩效统计提供切实有效的数据参考。

如图4,通过Kettle将各业务系统已发布数据经过清洗过滤后存储至ES集群中,其中图片、视频只抽取文件相关说明信息,不抽取文件内容,由于数据资源全集对数据实时性要求不高,针对各类数据在kettle上设置定时同步任务,每天凌晨增量更新。

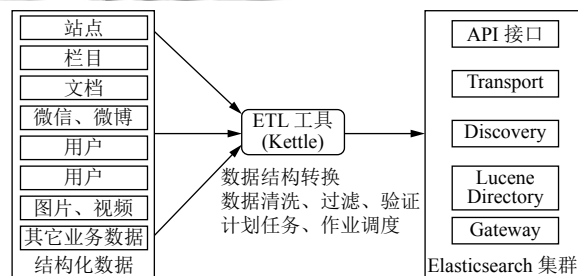


图4 数据抽取逻辑示意图

5 实践与分析

该平台已在中国科学院新媒体环境下的科学传播平台建设课题中实际应用,初步形成以“中科院之声”品牌为核心的科学传播矩阵,为中科院百余个科研机构网络宣传工作提供技术支撑。

5.1 业务系统改造

中科院前期已建成以网站群为核心的集约化新媒体发布集成环境,在此基础上进行系统改造,实现网站、新媒体等多渠道融合传播管理,主要改造工作包括:(1)改造业务系统功能接口,与公共接口网关进行对接,系统调用接口在接口网关进行注册和调用,实现各功能接口的统一管理和监控;(2)针对组织机构库和资源标识库,业务系统需要修改与组织机构和资源标识关联的业务表,相应添加机构代码和标识编码字段,用于系统间的数据展示;(3)针对数据资源全集,各系统需要相应开放数据抽取的数据库查询权限,提供结合自身业务实现的各类业务视图,供Kettle批量抽取数据并存放到ES中。

在系统实际改造中,还需注意:(1)组织机构库需根据机构调整及时维护完善机构信息和层级关系,平

台所有业务系统均基于组织机构为管理单元, 组织机构信息的准确性直接影响各类信息的传播监控和数据综合统计; (2) 共享稿件库实现了信息发布至各个渠道不同版本的跟踪查询, 实际运行过程中需结合机构自身情况考虑稿件的反复修改和多版本对照需求, 可增加稿件处理备注功能和同一稿件多版本比较功能, 方便用户实际操作; (3) 数据资源全集汇聚了平台包括文档在内的各类信息, 对于已同步至资源全集又有修改或临时撤销的信息, 这类数据需及时从资源全集中同步修改或删除, 具体实现采用数据库增加触发器和建临时表存放撤销、修改数据, 再通过 Kettle 定时同步更新资源全集, 也可结合业务实际考虑其他实现方式。

5.2 实践分析

平台为中科院主站、院机关、研究单位及下级机构的门户网站及新媒体运营等提供了运行支持。公共接口网关日均调用近万次, 组织机构库已有数据 1300 余条, 共享稿件库日均接收稿件 500 余篇, 数据资源全集已汇总用户信息近 3000 条, 站点信息 1000 余条, 栏目信息 20 余万条, 文档信息 500 多万条, 微信号 600 余个, 微博号 70 余个。

平台还将结合需求继续改进完善: (1) 平台安全防护还需提升, 包括接口调用数据加密、身份信息动态标识验证、恶意访问监控预警等; (2) 微博、微信等第

三方平台接口持续变化的处理, 第三方接口变化会直接影响平台业务功能, 需制定合理可行的技术处理流程及时调整; (3) 共享稿件库稿件版本管理, 稿件实际发布到微信等渠道后通常会进行二次编辑, 需考虑增加稿件版本对照功能方便用户实际使用。

6 结语

该平台在实际应用中进一步提升了中科院网络宣传工作的效率和质量, 平台建设思路对其他机构政务信息传播工作有一定借鉴作用。下一步将基于数据资源全集进行大数据分析, 并在此基础上探索科学传播工作的态势分析和智能化应用。

参考文献

- 1 国务院办公厅. 国务院办公厅关于印发政府网站发展指引的通知. 北京: 国务院办公厅. http://www.gov.cn/zhengce/content/2017-06/08/content_5200760.htm. [2017-06-08].
- 2 陈雄, 杜义华, 郭小龙. 政务网站新媒体发布集成环境建设探讨. 计算机系统应用, 2016, 25(8): 86-90. [doi: 10.15888/j.cnki.csa.005311]
- 3 Spring cloud. <https://spring.io/projects/spring-cloud>. [2020-08-10].
- 4 Kettle. <http://www.kettle.net.cn/>. [2020-08-10].
- 5 Elasticsearch. <https://www.elastic.co/guide/index.html>. [2020-08-10].