

中华人民共和国新闻出版行业标准

CY/T 231—2020

图书出版发行物联网平台总体结构

Structure of book publishing IoT platform

行业标准信息服务平台

2020-11-16 发布

2021-2-1 实施

国家新闻出版署 发布

目 次

目次	I
前言	II
图书出版发行物联网平台总体结构	1
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 平台架构	2
4.1 总体架构	2
4.2 感知层	3
4.3 数据源层	3
4.4 传输层	3
4.5 数据层	3
4.6 应用层	3
4.7 表示层	3
4.8 安全保障	3
4.9 标准规范	4
5 数据采集	4
5.1 数据上传节点	4
5.2 数据上传方式	4
5.3 上传文件格式	4
5.4 数据采集内容	4
6 数据应用	5
6.1 数据流向	5
6.2 图书出版辅助决策	5
6.3 图书市场营销支持	5
6.4 图书产品信息资讯	5
7 平台性能	5
7.1 基本性能要求	5
7.2 存储性能要求	6
7.3 网络性能要求	6
7.4 数据交换要求	6

前 言

本标准按照GB/T 1.1给出的规则起草。

本标准由北京印刷学院提出。

本标准由全国出版物发行标准化技术委员会（SAC/TC 505）归口。

本标准主要起草单位：北京印刷学院。

本标准主要起草人：杜艳平，艾立民，唐贾军，王兆华，刘冠一，窦水海，宁毅，高振清，张媛，李宏峰。

行业标准信息服务平台

图书出版发行物联网平台总体结构

1 范围

本标准规定了图书出版发行物联网平台架构、数据采集、数据应用、平台性能。
本标准适用于图书出版发行物联网平台的设计、开发和使用。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 12451 图书在版编目数据

GB/T 30330 中国出版物在线信息交换 图书产品信息格式规范

GB/T 33745 物联网 术语

GB/T 36626 信息安全技术 信息系统安全运维管理指南

GB/T 36958 信息安全技术 网络安全等级保护安全管理中心技术要求

3 术语和定义

下列术语和定义适用于本文件

3.1

感知 sensing

通过感知设备获得对象的信息的过程。

[GB/T 33745, 术语和定义2.1.8]

3.2

感知设备 sensing device

能够获取对象信息的设备，并提供接入网络的能力。

注：常见的感知设备有传感结点、RFID读写器等。

[GB/T 33745, 术语和定义2.1.9]

3.3

数据 data

信息的可再解释的形式化标识，以适用于通信、解释或处理。

[GB/T 33745, 术语和定义2.5.1]

3.4

信息 information

关于客体（如事实、事件、事物、过程或思想，包括概念）的知识，在一定的场合中具有特定的意义。

[GB/T 33745，术语和定义2.5.2]

3.5

数据挖掘 data mining

从大量的数据中通过算法搜索隐藏于其中信息的过程。

注：一般通过包括统计、在线分析处理、情报检索、机器学习、专家系统（依靠过去的经验法则）和模式识别等方法来实现。

[GB/T 33745，术语和定义2.5.3]

4 平台架构

4.1 总体架构

图书出版发行物联网应用平台主要由感知层、数据源层、传输层、数据层、应用层、表示层、安全保障体系和标准规范体系八部分组成。如图1所示。

平台的使用者为图书出版发行产业链相关用户，包括行政管理部门、出版商、中盘商、零售商、消费者等。

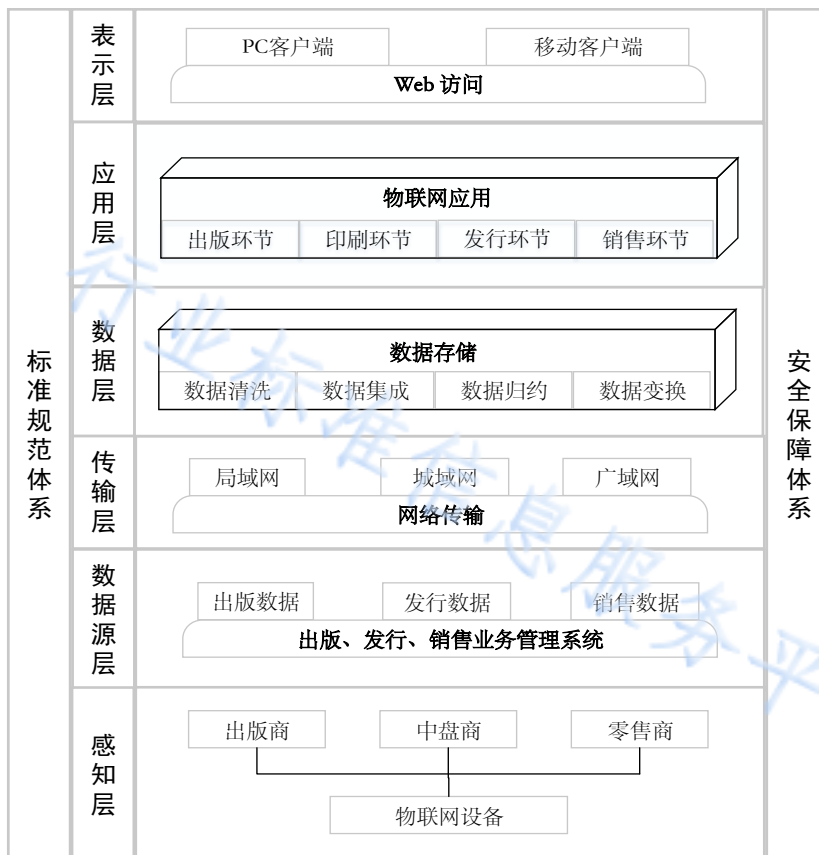


图1 图书出版发行物联网应用平台总体架构

4.2 感知层

出版商、中盘商、零售商通过物联网设备感知数据并上传至业务管理系统。

4.3 数据源层

数据源层分为出版数据、发行数据、销售数据三类数据。

4.4 传输层

数据源层通过网络将数据上传至数据层，网络接入可采用多种方式，包括但不限于专线、VPN、ATM、PPP等。

4.5 数据层

4.5.1 数据清洗

删除数据层上传数据中的无关数据、重复数据，平滑噪声数据，筛选与挖掘主题无关的数据，处理缺失值、异常值。

4.5.2 数据集成

将不同来源、格式、特点的数据源集成为适应平台应用开发的数据，并整合成一致的数据存储。

4.5.3 数据归约

在尽可能保持数据原貌的前提下，用替代的、较小的数据表示形式替换原始数据，最大限度地精简数据量。

4.5.4 数据变换

将数据转换成适合数据挖掘的形式。

4.5.5 数据存储

将整合成一致的数据建立相互关联的数据结构，统一存储于数据库。

4.6 应用层

通过物联网技术对数据层存储的数据进行数据挖掘，为出版商、中盘商、零售商等在出版、印刷、发行、销售环节，提供出版印刷辅助决策、市场营销支持、图书产品信息资讯等服务。

4.7 表示层

通过 Web 访问方式，为用户提供应用层输出的信息服务。支持 PC 端和移动客户端两种访问方式。

4.8 安全保障

4.8.1 物理安全

对平台环境、系统设备、存储介质等所有支持平台运行的硬件设施应采取安全措施。

4.8.2 网络与系统安全

网络与系统安全包括系统内部稳定运行安全以及系统外部防范病毒、入侵两方面内容,应遵循 GB/T 36958 的规定。

4.8.3 应用安全

应用安全包括用户身份认证安全,用户权限管理安全两方面内容。应遵循 GB/T 36626 的规定。

4.9 标准规范

平台的建设和运行须遵从国家和行业相关标准,全面综合考虑各环节数据特点,建立兼容性强、简洁便利的数据感知标准、数据传输标准、数据存储标准、安全保障标准等。

5 数据采集

5.1 数据上传节点

数据源层分别在图书出版、图书发行、图书销售三个环节通过传输层上传数据。

出版数据上传节点: 出版商上传图书基本信息、单册图书唯一标识和图书库存等数据。

发行数据上传节点: 中盘商上传图书发行物流信息和图书库存等数据。

销售数据上传节点: 出版商、中盘商和零售商上传图书零售数据; 零售商上传图书关注度和图书库存等数据。

5.2 数据上传方式

数据上传可采用标准API接口方式,接入平台的API接口采用RESTful API或web service方式。

5.3 上传文件格式

上传文件须符合国家标准、行业标准和行业工程标准的格式要求。

5.4 数据采集内容

5.4.1 图书基本信息

由出版商将图书基本信息上传至平台,所包含的内容应符合《GB/T 12451 图书在版编目数据》和《GB/T 30330 中国出版物在线信息交换 图书产品信息格式规范》的相关要求。

5.4.2 单册图书唯一标识

由出版商将写入图书RFID标签的单册图书唯一标识上传至平台,单册图书唯一标识应与图书基本信息建立关联。

5.4.3 图书库存数据

由出版商、中盘商和零售商将物联网设备感知的库存单册图书唯一标识上传至平台。

5.4.4 图书销售数据

由出版商、中盘商和零售商将物联网设备感知的零售单册图书唯一标识上传至平台。

5.4.5 图书关注度数据

由实体书店将物联网设备感知的读者阅览次数和单次阅览时长,以及相应的单册图书唯一标识上传至平台。

5.4.6 图书流转数据

由中盘商将物联网设备感知的图书物流信息上传至平台。

6 数据应用

6.1 数据流向

如图2所示,出版商、中盘商及零售商将物联网感知数据通过网络传输方式传输至图书出版发行物联网应用平台。图书出版发行物联网应用平台将海量多元异构数据进行清洗、集成、归约、变换等预处理,对预处理后的数据进行深度挖掘,将深度挖掘分析结果反馈至出版商、中盘商、零售商。

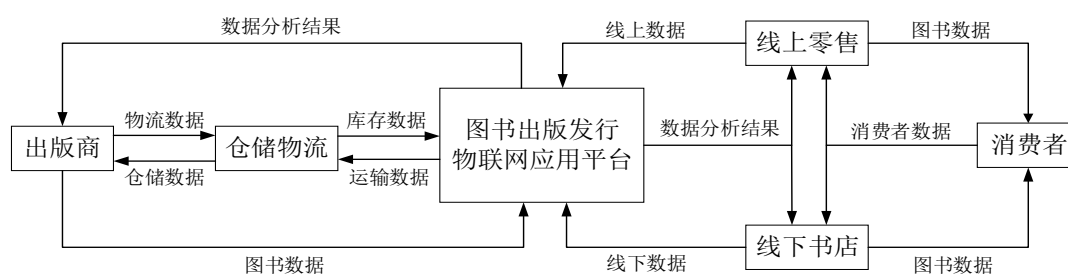


图2 图书出版发行物联网应用平台数据流向

6.2 图书出版辅助决策

提供各类图书销售类别、各类图书销售热度、各类图书读者群体、各发行企业库存等信息服务,为出版企业选题策划、确定印刷量等工作提供辅助决策支持。

6.3 图书市场营销支持

提供各类图书的销量、热销排行、读者关注度等信息服务,为出版发行企业图书市场预测、制订营销策略等工作提供支持。同时提供出版商和发行商的图书库存分布。

6.4 图书产品信息资讯

提供各类图书的防伪查询、热度分析、热销品种、关联产品等服务,为消费者提供图书查询等信息服务。

7 平台性能

7.1 基本性能要求

- 在规定的并发用户数范围内,对于平台的请求响应时间,简单应用不超过3s,一般应用不超过10s,复杂应用不超过30s;
- 并发用户数应不低于设计用户数的10%。

7.2 存储性能要求

- a) 应支持对异构存储系统虚拟化后对原有数据的管理，包括数据导入、导出、整合、迁移等；
- b) 应在存储数据传输过程中保持数据的完整性、一致性、高可用性；
- c) 应支持数据保护，包括快照、恢复、备份、复制、镜像、克隆、冗余等数据保护功能；
- d) 应支持国家密码管理机构允许的密码算法的硬件存储系统，支持数据加密存储和签名。

7.3 网络性能要求

- a) 外部接入网络带宽速率应达到1000M以上；
- b) 内部网络带宽速率应达到10000M以上；
- c) 虚拟网络资源的服务准备时间不超过20min；
- d) 内部网络最大端到端延迟时间小于50ms；
- e) 内部网络丢包率小于0.5%。

7.4 数据交换要求

- a) 数据交换服务适配器运行环境一次至少加载100个适配器实例；
- b) 每次交换至少支持2000条记录；
- c) 每次交换支持20M以内的公文；
- d) 数据交换服务适配器运行环境应至少支持50个并发数据交换服务；
- e) 数据交换服务最少支持500个并发数据交换；
- f) 交换服务、适配器运行环境的运行状态响应时间不超过15s。

行业标准信息服务平台