

ICS 35.240.01

CCS L67

SF

中华人民共和国司法行政行业标准

SF/T 0089—2021

司法行政信息资源中心建设规范

Specification for construction of judicial administrative information resource center

2021 - 08 - 31 发布

2021 - 08 - 31 实施

中华人民共和国司法部 发布

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 功能要求	1
6 分级建设要求	2
7 总体架构	2
8 技术支撑层	4
9 数据层	4
10 平台层	4

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由北京市司法局提出。

本文件由司法部信息中心归口。

本文件起草单位：北京市司法局、中国电子科技集团公司第十五研究所、北京北大软件工程股份有限公司、上海同道信息技术有限公司。

本文件主要起草人：田野、李洋、王宏业、吴卫荣、梅金芳、吴巍、刘颖、王树良、魏述强、孙苗苗、张靖杰。

司法行政信息资源中心建设规范

1 范围

本文件规定了司法行政信息资源中心的功能要求、分级建设要求、总体架构以及技术支撑层、数据层和平台层的要求。

本文件适用于司法行政信息资源中心的总体规划、设计、建设与应用。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；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。

- SF/T 0011 全国司法行政信息资源交换规范
- SF/T 0014 全国监狱信息化应用技术规范
- SF/T 0015 全国社区矫正管理信息系统技术规范
- SF/T 0016 全国社区矫正人员定位系统技术规范
- SF/T 0017 全国司法所管理信息系统技术规范
- SF/T 0018 全国人民调解管理信息系统技术规范
- SF/T 0019 全国安置帮教信息采集及数据交换技术规范
- SF/T 0020 全国司法行政信访管理信息系统技术规范
- SF/T 0021 全国司法行政戒毒信息化应用技术规范
- SF/T 0022 全国律师综合管理信息系统技术规范
- SF/T 0023 全国公证综合管理信息系统技术规范
- SF/T 0025 全国司法鉴定信息化技术规范
- SF/T 0026 全国人民监督员管理信息系统技术规范
- SF/T 0050—2019 司法数据资源平台和司法共享服务平台技术规范
- SF/T 0052 行政执法综合管理监督信息系统数据元和代码集

3 术语和定义

SF/T 0050—2019界定的以及下列术语和定义适用于本文件。

3.1

司法行政信息资源中心 **judicial administrative information resource center**

基于大数据相关平台或工具的功能支撑，利用大数据相关技术对司法行政信息资源进行统一归集、治理和分析，形成的司法行政信息资源池。

注：司法行政信息资源中心为司法行政系统使用和共享交换提供数据支持。

4 缩略语

下列缩略语适用于本文件。

- AI 人工智能 (Artificial Intelligence)
- API 应用程序接口 (Application Programming Interface)
- BI 商业智能 (Business Intelligence)

5 功能要求

司法行政信息资源中心建设应实现司法行政信息资源的统一管理，实现信息资源的汇聚、治理、分析和展示功能，同时也为外部单位和司法行政系统提供数据和应用支撑。

6 分级建设要求

司法行政信息资源建设应按照第7章的要求分级建设部级和省级司法行政信息资源中心，并进行层级递进的数据上报与交互，要求如下：

- a) 司法部建设部级司法行政信息资源中心；
- b) 各省建设省级司法行政信息资源中心；
- c) 市级司法机关宜使用本省司法行政信息资源中心或建设本省司法行政信息资源地市级分中心。

7 总体架构

图1规定了司法行政信息资源中心总体架构，包括技术支撑层、数据层和平台层，以及标准规范体系、运维保障体系和安全保障体系。各层内容如下：

- a) 技术支撑层：包括机房/互联网数据中心（IDC）、云资源（电子政务云和公有云）、电子政务外网和互联网；
- b) 数据层：包括依法治国（省、市）类数据、行政立法类数据、行政执法类数据、刑事执行类数据、公共法律服务类数据、政务管理类数据和支撑数据分析的非司法行政类数据；
- c) 平台层：包括数据分析管理、数据资源子平台、数据共享和运行监控；
- d) 标准规范体系：包括各层所需的标准和规范；
- e) 运维保障体系：为各层稳定可靠运行提供管理与服务支撑，包括但不限于运维监控、运维审计和应急预案；
- f) 安全保障体系：为各层提供安全技术保障。

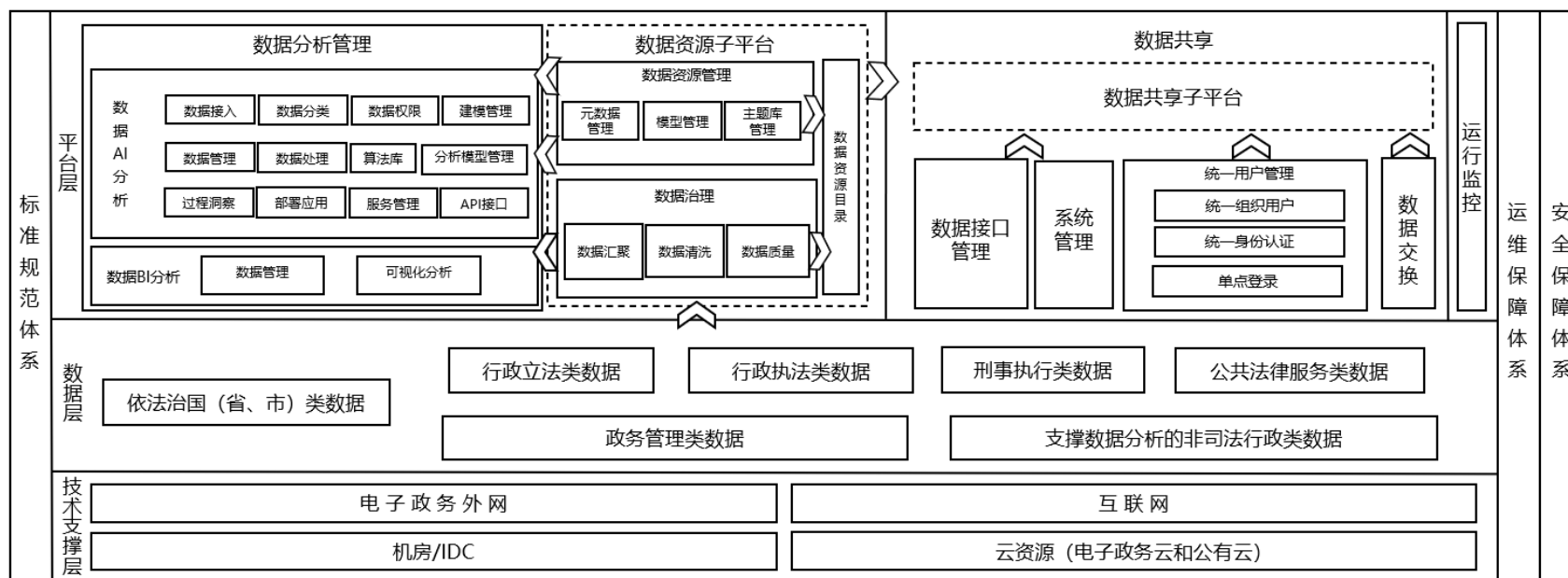


图1 司法行政信息资源中心总体架构

8 技术支撑层

司法行政信息资源中心应依托技术支撑层构建对内对外的计算环境。技术支撑层应符合以下要求：

- a) 具备承载大数据分析算力，提供大数据分析服务和海量数据共享交换的能力；
- b) 利用虚拟机灵活分配计算资源和存储资源，其中：
 - 1) 计算资源包括但不限于离线计算、实时计算、流式计算和内存计算；
 - 2) 存储资源包括但不限于关系型数据库、分布式数据库、列式数据库和图数据库。

9 数据层

数据层应符合以下要求：

- a) 数据层集中司法行政信息化已建成和在建的司法行政业务应用的数据；
- b) 数据层作为数据源提供方，能够为司法行政信息资源中心提供数据支撑，供数据分析管理和数据资源子平台进行数据分析和数据治理。

10 平台层

10.1 数据分析管理

10.1.1 基本要求

数据分析管理应包括但不限于数据BI分析和数据AI分析功能。利用数据资源子平台提供的元数据、经过数据治理的业务数据和主题库数据进行BI分析和AI分析。

10.1.2 数据 BI 分析

数据BI分析应支持可视化数据分析功能，可通过拖拽的形式简化用户操作，并支持基于权限控制，共享BI分析能力和分析结果。包括但不限于数据管理和可视化分析功能。应符合以下要求：

- a) 数据管理用于维护 BI 分析所用的数据源、数据模型和视图，并进行数据权限分配与管理，实现分析数据的统一接入与按需分发，包括但不限于数据接入、数据模型管理和数据权限；
- b) 可视化分析将数据进行预处理后通过图形化和拖拽式的设计模式，设计不同的应用场景并能进行成果管理，包括但不限于数据准备、可视化设计、自由式报表、场景设计和成果管理功能。

10.1.3 数据 AI 分析

数据AI分析应支持可视化AI分析功能，可通过拖拽的形式简化用户操作，支持基于权限控制，共享AI分析能力和分析结果。数据AI分析包括但不限于数据接入、数据分类、数据权限、建模管理、数据管理、数据处理、算法库、分析模型管理、过程洞察、部署应用、服务管理和API接口。应符合以下要求：

- a) 数据接入支持多种数据源的接入，包括但不限于关系型数据库、大数据引擎、文本数据和第三方接口数据；
- b) 数据分类提供轻量级的元数据管理模块，包括但不限于数据主题域划分和数据模型创建维护功能；
- c) 数据权限具备严密的数据权限管理机制，确保数据的安全、可控、可跟踪和可溯源，包括但不限于数据权限控制、数据权限分配日志和数据下载权限控制；
- d) 建模管理支持以图形化和流程式的方式进行建模；
- e) 数据管理具备简易和自助式数据准备能力，包括但不限于数据的关联、追加、输入和输出。其中：
 - 1) 数据输入包括但不限于文件输入、关系型数据库输入、大数据输入和 API 输入；
 - 2) 数据输出包括但不限于关系型数据库输出、大数据输出和文件输出。
- f) 数据处理提供数据预处理，对数据进行规范化、降维和特征提取操作。功能包括但不限于行处理、列处理、数据融合、特征工程和流数据处理；

- g) 算法库提供基础算法满足数据建模，同时支持文本分析相关算法和多种扩展能力。具体如下：
 - 1) 基础算法包括但不限于回归、分类、聚类、时间序列、综合评价和关联规则；
 - 2) 文本分析相关算法包括但不限于分词、信息抽取、文本过滤和关键词提取；
 - 3) 多种扩展能力包括但不限于 R/PYTHON/JAVA 编程语言的扩展。
- h) 分析模型管理包括但不限于模型评估、模型比较、模型读取、模型输出、模型利用和多版本保存及利用；
- i) 过程洞察提供对建模过程进行完整流程跟踪的能力，包括但不限于流程综述、节点洞察和报告导出功能；
- j) 部署应用提供将成果模型部署上线并发布为服务的能力；
- k) 服务管理支持 AI 分析服务的统一管理。包括但不限于服务列表、服务授权、服务运行信息统计、服务调度和任务管理；
- l) API 接口提供核心功能 API 接口，支持基于 API 接口快速进行自定义二次开发。

10.2 数据资源子平台

10.2.1 基本要求

数据资源子平台应包括但不限于数据治理、数据资源管理和数据资源目录，对数据层数据进行数据汇聚、数据清洗和数据质量检查，形成元数据、模型、主题库和数据资源目录。

数据资源子平台将元数据和经过数据治理的业务数据形成数据资源目录，供数据共享使用。

10.2.2 数据治理

10.2.2.1 数据汇聚

数据提供者提供的数据，应包括数据库数据、数据汇聚编目、数据字典和表关联关系说明，并定期维护，数据汇聚要求如下：

- a) 支持多样数据汇聚方式：应包括但不限于提取转换加载（ETL）和接口方式；
- b) 支持多种数据源：应包括但不限于关系型数据库、非关系型数据库和分布式文件系统；
- c) 支持调度任务管理：应包括但不限于支持数据汇聚任务调度管理，创建、编辑、修改和删除操作，支持自动调度任务和手动执行任务；
- d) 数据加工：应能对数据源表进行多维度工作，包括但不限于数据源重命名、纵连接、横连接、数据过滤和字段映射；
- e) 数据属性：数据的属性应至少包含主键标识和增量标识，增量标识用于数据汇聚时增量同步，宜采用时间戳作为增量标识。

10.2.2.2 数据清洗

数据清洗应符合以下要求：

- a) 数据清洗任务管理：支持对数据清洗任务的创建、编辑、删除和查询功能；
- b) 待清洗数据多源头接入：支持多种数据源接入方式，包括但不限于加载系统数据模型中已有数据、加载质量评估报告中最新一次评估的脏数据以及通过 EXCEL 方式导入的外部数据；
- c) 数据清洗应用功能：在数据清洗工程任务中，包括但不限于编辑、标记列、数据动态加载、排序、分组、模糊查询、列移动、列隐藏、大小写转换、日期格式转化和暂存操作；
- d) 数据清洗结果及历史查询：对数据清洗结果提供展示界面，并可查看多次清洗的历史过程或结果；
- e) 数据质量检查规则复用：支持经数据质量检查模块已完成的数据质量检查规则和指标的复用。

10.2.2.3 数据质量

数据质量应符合以下要求：

- a) 质量检查规则：可自定义规则及待检字段范围，支持新增检查、导入检查和接口检查多种模式。并且检查规则配置可复用；

- b) 质量检查指标：包括非空规则指标、一致规则指标、唯一规则指标、组合唯一规则指标、正则规则指标、阈值规则指标和组合规则指标；
- c) 数据质量报告：能够对数据质量评估时间、评估数据量、合格数据量、不合格数据量以及合格率进行说明，并且提供脏数据明细的 EXCEL 导出功能；
- d) 数据质量跟踪：能够对数据质量做跟踪追溯，从时间维度统计各模型数据质量提升情况，直观展示各模型数据质量变化趋势，并能对每次执行的数据质量评估任务生成相应版本的数据质量报告，保留数据质量任务日志，用于数据治理质量的跟踪对比；
- e) 数据质量检查任务管理：能够创建、修改和删除，并按照配置的质量指标对指定数据模型的生效数据定时自动进行规范性检查，生成数据质量报告，同时，系统也应具备手动执行数据质量检查任务的功能。

10.2.3 数据资源管理

10.2.3.1 元数据管理

元数据管理应符合SF/T 0050—2019中5.4.2.1有关元数据管理的规定。

10.2.3.2 模型管理

模型管理应符合以下要求：

- a) 资源分类管理：提供自定义数据资源类型，形成管理数据对象的资源体系分类，支持建立具有层次结构的资源目录结构。数据资源分类应符合 SF/T 0050—2019 中 6.1 有关司法信息资源分类的规定；
- b) 数据字典管理：对各类基础数据进行自定义，供数据建模引用；
- c) 数据建模：提供数据模型自定义，包括但不限于属性校验规则的建立、数据层级关系展现、模型属性元素定义和约束条件定义；
- d) 属性维护：提供数据对象属性定义功能，可配置数据对象的属性编码、属性名称和属性类型等信息；
- e) 数据关联关系：可通过配置的方式实现数据对象之间的关联关系管理，并能支持多个关联对象的设置，同时支持新数据对象与历史数据对象之间的关系配置；
- f) 模型导入导出：将模型导出为 XLS 和 XML 等文件，提供模型导入模板；
- g) 模型启用及停用：包括模型的启用和停用。

10.2.3.3 主题库管理

主题库管理应按照数据资源的治理和分析应用的实际情况，建立包括但不限于基础汇聚库、标准业务库和主题分析库。要求如下。

- a) 基础汇聚库：汇聚司法行政各业务条线数据，形成基础汇聚库，数据表经过处理后具有的字段应包括但不限于主键、业务字段、是否必填、时间戳和数据生效标记；
- b) 标准业务库：应按照 SF/T 0014、SF/T 0015、SF/T 0016、SF/T 0017、SF/T 0018、SF/T 0019、SF/T 0020、SF/T 0021、SF/T 0022、SF/T 0023、SF/T 0025、SF/T 0026 和 SF/T 0052 各业务应用数据要求，结合各地区数据质量要求，将基础汇聚库的数据进行质量标准定义和数据清洗，形成结果数据并归至标准业务库中；
- c) 主题分析库：应以标准业务库为源头建立主题分析库，结合各业务应用需求和内外部数据共享需求，进行统计和融合分析，形成分析模型和共享交换模型，并归至主题分析库，主题分析库的分类应符合 SF/T 0050—2019 中 6.1.1 的规定，下级分类宜结合实际应用需求进行细分和扩展。

10.2.4 数据资源目录

数据资源目录应符合以下要求：

- a) 目录式管理、查阅和下载数据资源：能够对司法行政信息资源中心所拥有的数据统一梳理，形成目录清单，并在数据资源目录界面将所有数据资源进行展示，方便用户浏览、检索和下载；
- b) 数据资源目录范围内容如下：
 - 1) 部级数据资源目录包括但不限于部级资源目录和省级数据资源目录两级数据资源；
 - 2) 省级数据资源目录包括但不限于本级数据资源目录。

10.3 数据共享

10.3.1 基本要求

数据共享应包括但不限于数据接口管理、系统管理、统一用户管理、数据共享子平台和数据交换，为各级司法行政信息资源中心、司法行政系统和外部单位提供共享服务支撑。

10.3.2 数据接口管理

应提供主动与被动接口服务能力，其中主动接口服务可通过集成和服务配置，主动分发数据；被动接口服务提供本地各类查询接口，供外部系统调用。具体功能包括：

- a) 本地接口服务：提供的供外部系统调用的接口包括但不限于查询数据接口、接收数据接口和申请编码接口；
- b) 集成系统管理：对有数据共享需求的系统进行注册；
- c) 服务接口管理：提供第三方系统的接口注册功能，定义接口的系统名称、接口用途和接口参数；
- d) 数据集成管理：对目标系统和使用的接口进行管理，并提供数据交换规则的配置功能；
- e) 数据分发管理：定义自动分发条件，能查看的内容包括但不限于每类数据的分发状态、分发时间和数据详情；
- f) 接口用户访问控制：支持对访问的数据范围、字段范围、可访问的数据量进行控制和安全性验证；
- g) 接口调用日志：内容包括但不限于满足审计业务需要的操作类型、操作数据、操作时间、操作人和操作 IP 地址。

10.3.3 系统管理

系统管理应符合以下要求：

- a) 权限管理：实现系统中权限资源信息的“新增、删除、修改和排序”操作，且具备菜单权限配置、系统角色权限配置、权限资源维护和权限分配的统一维护；
- b) 安全管理（数据容灾备份等管理）：能定期对数据进行备份，支持导入导出功能，在发生应急情况时，可通过快速备份文件来恢复资源中心的数据应用。

10.3.4 统一用户管理

统一用户管理要求如下：

- a) 应支持司法行政用户的分部门和分权限管理；
- b) 应支持上级司法行政信息资源中心的用户经授权后跨级访问下级司法行政信息资源中心；
- c) 包括但不限于统一组织用户、统一身份认证和单点登录功能；
- d) 各级司法机关宜利用已具备的统一组织用户、统一身份认证和单点登录功能，实现统一用户管理。

10.3.5 数据共享子平台

数据共享子平台建设应符合SF/T 0050—2019中第7章的规定。

10.3.6 数据交换

数据交换应保证数据交换接口统一，接口应符合SF/T 0011的规定。数据交换的内容应符合以下要求：

- a) 司法行政内部数据交换包括司法行政各业务应用之间的数据交换、部省两级司法行政信息资源中心的数据交换以及部级统建系统（省级统建系统）对下级司法行政机构的数据回流；
- b) 司法行政数据与外部数据交换包括司法行政各业务应用与司法行政以外的单位之间的数据交换，如与政法委、公安机关、检察院和法院等单位的数据交换。

10.4 运行监控

运行监控支持对司法行政信息资源中心运行过程进行实时状态监控，包括系统运行监控和数据运行监控，应符合以下要求：

- a) 系统运行监控：对系统运行状态进行实时监控，包括但不限于登录人员信息、浏览信息和操作信息的全程留痕；
 - b) 数据运行监控：对数据治理、数据交换的数据信息进行监控，包括但不限于数据来源、数据量和数据接口信息。
-