

ICS 35.240  
L 67



# 中华人民共和国国家标准

GB/T 37721—2019

---

## 信息技术 大数据分析系统功能要求

Information technology—Functional requirements for big data analytic systems

2019-08-30 发布

2020-03-01 实施

---

国家市场监督管理总局  
中国国家标准化管理委员会 发布

## 目 次

前言 .....	III
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 总体要求 .....	2
6 数据准备模块功能要求 .....	2
6.1 数据抽取功能要求 .....	2
6.2 数据清洗功能要求 .....	2
6.3 数据转换功能要求 .....	3
6.4 数据加载功能要求 .....	3
7 分析支撑模块功能要求 .....	3
7.1 查询功能要求 .....	3
7.2 机器学习功能要求 .....	4
7.3 统计分析功能要求 .....	4
7.4 可视化功能要求 .....	4
8 数据分析模块功能要求 .....	5
8.1 分析模式 .....	5
8.2 分析类型 .....	6
9 流程编排模块功能要求 .....	6
9.1  workflow管理 .....	6
9.2 告警和日志 .....	6
附录 A (资料性附录) SQL 关键字 .....	7

## 前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

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这些专利的责任。

本标准由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本标准起草单位:浪潮电子信息产业股份有限公司、中国电子技术标准化研究院、勤智数码科技股份有限公司、杭州海康威视数字技术股份有限公司、浪潮软件集团有限公司、上海计算机软件技术开发中心、新华三技术有限公司、天津南大通用数据技术股份有限公司、中兴通讯股份有限公司、华为技术有限公司、星环信息科技(上海)有限公司、北京百分点信息科技有限公司、北明软件有限公司、哈尔滨哈工大大数据通用技术有限公司、国网上海市电力公司、陕西省信息化工程研究院、南京南瑞信息通信科技有限公司、广州浪潮大数据研究有限公司、杭州中奥科技有限公司、科大讯飞股份有限公司。

本标准主要起草人:苏志远、张东、赵江、卫凤林、张群、刘宇峰、李正、赵世范、黄先芝、王建华、陈敏刚、刘振宇、蔡立志、潘子健、赵伟、孙卡、吴文峰、刘蔚、王东、赵华、符海芳、周洪明、孙伟、汪疆平、王进宏、赵志强、王刚、王宏志、郭乃网、苏运、张勇、孙立华、汤宁、刘广庆、沈贝伦、陆韵、武新、张绍勇、赵乾、李冰、尹卓、孙嘉阳。

# 信息技术 大数据分析系统功能要求

## 1 范围

本标准规定了大数据分析系统的数据准备模块、分析支撑模块、数据分析模块和流程编排模块的功能要求。

本标准适用于大数据分析系统的设计、开发和应用部署。

## 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 35295—2017 信息技术 大数据 术语

## 3 术语和定义

GB/T 35295—2017 界定的以及下列术语和定义适用于本文件。

### 3.1

#### 大数据分析系统 **big data analytic systems**

在大数据存储和处理系统提供的原始数据和计算框架的基础上,集成了一系列数据分析生存周期过程中所用工具的系统。

### 3.2

#### 结构化数据 **structured data**

存储在数据库里,可以用二维表结构表示的数据。

### 3.3

#### 非结构化数据 **unstructured data**

除了结构化数据之外的没有明确结构约束的数据。

### 3.4

#### 分布式执行计划 **distributed execution plan**

分布式场景下的 SQL 查询计划,需要根据数据分布特点将 SQL 拆分成多个切片及多个步骤,提供调度给多节点并行执行。

## 4 缩略语

下列缩略语适用于本文件。

API:应用程序编程接口(Application Programming Interface)

GPU:图形处理器(Graphics Processing Unit)

JSON:JS 对象标记(JavaScript Object Notation)

OLAP:联机分析处理(On-Line Analytical Processing)

REST:表述性状态转移(Representational State Transfer)