

# 大数据平台安全 研究报告

中国信息通信研究院安全研究所  
2021 年 1 月

---

## 版权声明

---

本报告版权属于中国信息通信研究院，并受法律保护。转载、摘编或利用其它方式使用本研究报告文字或者观点的，应注明“来源：中国信息通信研究院”。违反上述声明者，本院将追究其相关法律责任。

## 前 言

2020 年全球遭到了新冠肺炎疫情的打击，据经济合作组织 2020 年 9 月 16 日发布的《经济展望中期报告》，受疫情影响，2020 年全球 GDP 预计萎缩 4.5%，各国面临着经济逆行的巨大挑战。为应对经济压力，各国纷纷将目光聚焦于数字经济，将其作为经济逆境中的新动力。据中国信通院发布的《全球数字经济新图景（2020 年）》预计，2020 年全球数字化转型技术和服务支出将增长 10.4%，达到 1.3 万亿美元，是 2020 年中为数不多的经济增长亮点之一。

大数据是数字经济的重要基础，蕴含着巨大的潜在价值，为探索客观世界新规律、减少战略决策中的主观因素影响提供了可能。据 2020 年 9 月 10 日国家互联网信息办公室印发的《数字中国建设发展进程报告（2019 年）》，2019 年我国大数据产业规模超过 8100 亿元，同比增长 32%，依托大数据，人工智能、区块链、工业互联网等数字经济产业得到快速发展。

大数据平台为大数据提供了计算和存储的能力，这使得海量的静态数据“活动”起来，并释放出自身价值。然而，一旦缺少了平台安全这个前提，数据价值的释放将受到阻碍。如果将大数据平台比作大厦，其价值的释放能力和安全能力分别是大厦的地面建筑和地基，地基的深度决定了大厦地面建筑的高度，地基不稳的大数据平台，注定只能是“空中楼阁”。为了“测量”并“夯实”大数据平台的安全“地基”，我院在 2020 年发起了卓信大数据平台安全专项

行动。行动开展过程中，我们发现大数据平台在建设和运维方面确实存在一定的安全问题，本报告对此进行了深入分析。

报告全文以本次专项行动中积累的安全检测数据为基础，从平台配置安全隐患和安全漏洞的分布规律、产生原因、危害影响、修复难度等维度分析了大数据平台的安全现状。同时，详细分析了形成该安全现状的问题根源，并给出了相应的解决方案建议。最后，从监管、标准、技术研究等方面提出了大数据平台安全未来的工作方向。

“工欲善其事，必先利其器”，在数字经济蓬勃发展的大背景下，我们深刻认识大数据平台安全，不仅需要技术手段上的安全防护，还需要安全意识和运维水平的同步提升，亟需多方共同努力。希望本报告能够为企业大数据平台安全建设提供参考，为行业大数据的健康发展提供助力。

# 目 录

一、大数据平台概况.....	1
(一) 大数据产业蓬勃发展.....	1
(二) 大数据平台应用模式多样化演进.....	2
二、大数据平台安全现状.....	4
(一) 组件配置类安全隐患.....	6
(二) 组件安全漏洞.....	9
(三) 组件安全隐患统计分析.....	12
三、大数据平台安全问题分析.....	18
(一) 基于 Hadoop 的开源大数据平台安全配置复杂度较高.....	19
(二) 安全漏洞修复对平台运行影响较大.....	19
(三) 大数据平台建设过程中安全投入不足.....	21
(四) 大数据平台重视边界防护忽视内部安全.....	21
(五) 企业大数据平台安全管理制度滞后.....	22
(六) 企业技术人员安全能力不足.....	23
四、大数据平台安全解决方案建议.....	23
(一) 加强大数据平台安全基线管理.....	24
(二) 对大数据平台安全进行整体规划.....	24
(三) 大数据平台边界防护与内部安全建设并重.....	25
(四) 建立完善的大数据平台安全制度流程.....	25
(五) 增强企业技术人员安全能力.....	25
五、大数据平台安全未来发展建议.....	26
(一) 加强企业大数据平台安全防护工作的监管.....	26
(二) 强化大数据平台安全防护技术研究.....	27
(三) 推动大数据平台安全产品和服务市场发展.....	27
(四) 构建大数据平台安全生态.....	28

## 图 目 录

图 1 中国大数据市场规模预测 .....	2
图 2 大数据平台架构 .....	2
图 3 大数据平台类型分布占比 .....	5
图 4 大数据平台组件配置安全隐患统计 .....	6
图 5 大数据平台安全漏洞统计 .....	10
图 6 单个组件安全隐患占比 .....	13
图 7 单个组件配置安全隐患占比 .....	13
图 8 单个组件安全漏洞占比 .....	14
图 9 各类组件安全隐患占比 .....	15
图 10 各类组件配置安全隐患占比 .....	15
图 11 各类组件漏洞安全隐患占比 .....	16
图 12 传输类组件安全隐患等级分布 .....	17
图 13 存储类组件安全隐患等级分布 .....	17
图 14 计算类组件安全隐患等级分布 .....	18
图 15 平台管理类组件安全隐患等级分布 .....	18
图 16 漏洞修复方式选择流程图 .....	20

## 一、大数据平台概况

当今时代大数据产业持续发展，对于挖掘新的经济增长点大有益处，大数据成为了推动各行业发展的新动力。大数据平台作为大数据的载体起到了关键作用。

### (一) 大数据产业蓬勃发展

根据 IDC 最新预测，2020 年中国大数据相关市场的总体收益将达到 104.2 亿美元，较 2019 年同比增长 16.0%，如图 1 所示，增幅领跑全球大数据市场。

2020 年，大数据硬件在中国整体大数据相关收益中将继续占主导地位，占比高达 41.0%；大数据软件和大数据服务收入比例分别为 25.4% 和 33.6%。而到 2024 年，随着技术的成熟与融合、以及数据应用和更多场景的落地，软件收入占比将逐渐增加，服务相关收益占比将保持平稳，而硬件收入在整体的占比则将逐渐减少。硬件、服务、软件三者的比例将逐渐趋近于各占三分之一的比例。IDC 预计，在 2020-2024 年的预测期间内，中国大数据相关技术与服务市场将实现 19.0% 的年均复合增长率。

### 2019-2024年中国大数据市场规模预测



数据来源：IDC 中国，2020

图 1 中国大数据市场规模预测

## (二) 大数据平台应用模式多样化演进

大数据平台通常使用 Apache Hadoop 架构进行搭建，由存储、计算、平台资源管理、传输交换等类型组件组成整个平台，如图 2 所示。目前，大数据平台主要有基于开源技术的自建平台和商业化平台两种。自建平台灵活性强、自主性高，商业化平台安全性强、使用便捷。



预览已结束，完整报告链接和二维码如下：

[https://www.yunbaogao.cn/report/index/report?reportId=1\\_18207](https://www.yunbaogao.cn/report/index/report?reportId=1_18207)

