

中国大数据发展调查报告 (2018年)

中国信息通信研究院

2018年4月



版权声明

本调查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调查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。

报告目录

前言	III
观点摘要	4
一、调查背景	6
(一) 调查方法及样本	6
(二) 报告术语界定	7
二、大数据产业规模	8
(一) 大数据产业总体规模	8
(二) 大数据核心产业规模	9
(三) 细分领域规模	9
三、大数据应用	10
(一) 大数据应用现状	10
(二) 大数据应用场景	13
(三) 大数据应用带来的效果	13
(四) 大数据应用的主要障碍	14
(五) 未来大数据投入趋势	15
四、大数据资源	16
(一) 数据规模	16
(二) 数据来源	16
(三) 数据类型	18
(四) 数据共享	19
(五) 数据管理	20
五、大数据平台	21
(一) 建设模式	21
(二) 自建大数据平台模式	22
(三) 采购公有云服务模式	24
(四) 大数据产品选型	25
六、政策需求和资源需求	26
(一) 政策需求	26
(二) 资源需求	27

前言

在国家政策的推动下，2017 年大数据市场发展迅速，与此同时国内重要软硬件企业陆续推出大数据相关产品。为进一步掌握中国大数据发展和应用情况，中国信息通信研究院开展了 2017—2018 年度中国大数据发展状况的调查。本次活动采用现场访问、电话访谈和在线调查相结合的方式，对中国大数据产业规模、应用现状、数据资源情况、大数据平台、政策和数据需求等问题进行了调研，共回收有效问卷 1572 份。本报告以调查数据为基础，结合行业专家的访谈，力争详实客观的反映中国大数据发展现状，为政府和企业了解中国大数据发展状况和制定相关决策提供参考，也为广大关注大数据的从业人员、专家学者和研究机构提供真实可信的大数据发展报告。

本报告的数据搜集工作得到了相关政府部门、大数据解决方案提供商、企业用户以及社会各界的大力支持。我们在此表示衷心的感谢！同时也对接受大数据调查访问的企业朋友表示最诚挚的谢意！

中国信息通信研究院

2018 年 4 月 16 日

观点摘要

大数据产业规模

- 2017 年中国大数据产业总体规模为 4700 亿元人民币，同比增长 30%；2017 年大数据核心产业规模为 236 亿元人民币，增速达到 40.5%，预计 2018-2020 年增速将保持在 30%以上。

大数据应用

- 接近 2/3 的企业已经成立了相关的数据分析部门，企业对数据分析的重视程度进一步提高。65.2% 的受访企业已经成立了数据分析部门；正在计划成立相关数据部门的企业占比为 24.4%。
- 近四成的企业已经应用了大数据。在接受调查的 1572 家企业中，已经应用大数据的企业有 623 家，占比为 39.6%，与 2016 年相比上升 4.5%，垂直行业中如金融等领域大数据应用增加趋势较为明显。
- 大数据应用帮助企业实现了智能决策、提高了运行效率和风险管理能力。应用大数据实现了智能决策的企业占比最高(55.2%)；其次，46.6%的企业表示应用大数据后提升了运营效果，35.9%的企业应用大数据后能够更好的管理风险。
- 接近六成的企业在大数据领域的投入将持续增加。调查发现，58.2%的受访企业预计未来将加大大数据投入力度，相比 2016 年上升了 6.7%；其中，大数据投入增加超过 50%的企业占比达到了 25.5%；另有 32.7%的企业在大数据方面的投入增加在 50%以内。

大数据资源

- 在受访企业中，内部生产数据和客户/用户数据是企业的两大数据来源。54.3%和 48.8%的企业数据来自其内部生产数据和客户/用户数据，与 2016 年相比分别上升 5.3%和 1%。
- 87.6%的企业非结构化数据已超过一半比例。其中，非结构化数据比例超过 90%的企业占比为 16.2%，与 2016 年相比上升 1.5%；非结构化数据占数据总量 70%-90%的企业占比 55.8%，比 2016 年上升了 4.4%；非结构化数据比例在 50%-70%之间的企业占比为 15.6%。
- 22.5%的企业已经聘用专职首席数据官 (CDO)。调查发现，12.3%的受访企业聘用专职首席数据官已经超过两年；聘用专职首席数据官 (在两年以内)的企业占比为 10.2%。此外，33.4%的受访企业表示将在一年内聘用或正在考虑聘用专职首席数据官。随着对数据管理和利用数据支撑战略

决策的企业越来越多，未来聘用专职首席数据官的企业比例有望进一步提升。

大数据平台

- **本地部署仍是企业进行大数据平台建设的首选方式，同时云服务商大数据服务的使用率有所提升。**选择自建大数据平台企业占比最大（49.1%）；35.6%的企业倾向选择云服务商提供大数据服务，较2016年上升了2.5%。
- **基于非结构化数据架构的大数据平台应用最为广泛。**38.6%和23.8%的受访企业分别选择非结构化的批处理架构（如Hadoop）以及非结构化数据的内存架构（如Spark）作为大数据平台的技术架构。
- **六成以上的受访企业自建大数据平台的服务器数量在50台以下。**调查显示，28%的企业部署的服务器数量在20台以下，与去年相比提高了0.5个百分点；服务器数量在20-50台的企业占比超过三成。
- **与其他产品相比，企业对在线可视化工具的选择增幅最大。**采用在线可视化工具的企业占比为23.3%，与2016年相比提高了6.5%；增幅次之的是在线大数据分析工具（2.3%）。

政策需求和资源需求

- **六成以上的企业认为完善行业标准、健全法律法规将有助于推动大数据的发展。**据调查，68.3%的受访企业认为完善行业标准及法律法规对于推动大数据市场的发展至关重要；其次，43.1%的企业表示应加强个人信息保护，比2016年上升了1.9%；此外，企业对开放更多政府信息公开资源（42.7%）和资助更多大数据领域的科研项目（36.3%）的呼吁较高。
- **企业对市政管理、教育科研以及交通服务类公共数据资源的需求增强。**在对政府公共数据资源需求情况调查中，2017年企业对市政管理（36.4%）、教育科研（36%）以及交通服务（35.2%）类数据资源需求均有一定程度的提升。

一、调查背景

(一) 调查方法及样本

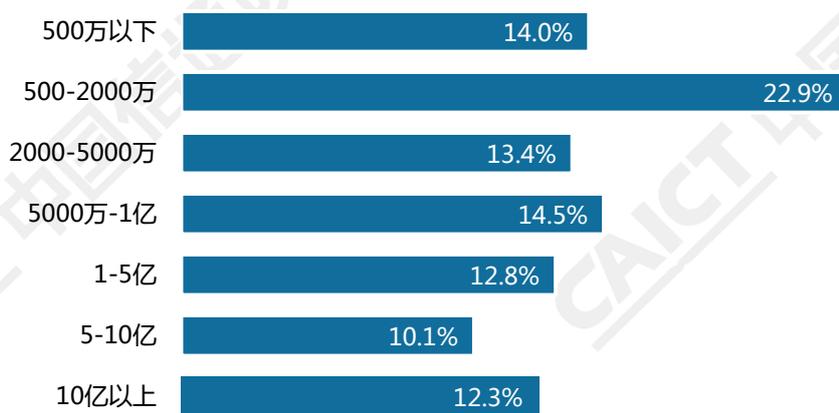
1、调查方法

本次报告信息收集采用现场访问、电话访谈和在线调查相结合的方式，收集到 1,572 份大数据用户的有效问卷。其中，通过现场访问收集问卷 712 份，通过电话访谈收集问卷 486 份，通过在线调查收集问卷 374 份。

2、样本描述

参与调查企业规模：以中小企业为主，年营业额在 1 亿元以下的企业占总体样本的 64.8%。

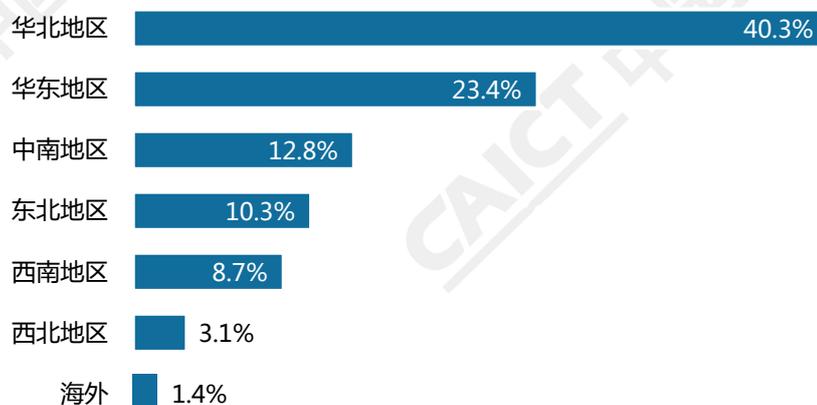
图 1 企业营业额分布 (N=1,572)



数据来源:中国信息通信研究院

参与调查企业所在区域：覆盖华北、东北、华东、中南、西南、西北以及海外地区。

图 2 企业区域分布 (N=1,572)



数据来源:中国信息通信研究院

参与调查企业所在行业：包括互联网、软件及信息技术服务、电信、金融、制造、交通及物流、能源、医疗、政府、科教文卫和农林牧渔等行业。

(二) 报告术语界定

大数据产业：本报告所称大数据产业，指以数据生产、采集、存储、加工、分析、服务为主的相关经济活动，包括数据资源建设、大数据软硬件产品的开发、销售和租赁活动，以及相关信息技术服务。

大数据核心产业：本报告所称大数据核心产业，指大数据直接相关的软件、硬件及专业服务。大数据市场规模的计算口径，相应的包含大数据软件、硬件及服务的直接产值，不含因大数据应用产生的附加值。

大数据软件：本报告所称大数据软件，主要指以 Hadoop、Spark 等技术架构为基础的新型分布式平台软件，以及基于这类平台软件的应用软件。

大数据服务：本报告所称大数据服务，主要包括大数据分析公有云服务及大数据系统相关的咨询、设计和系统集成服务等。

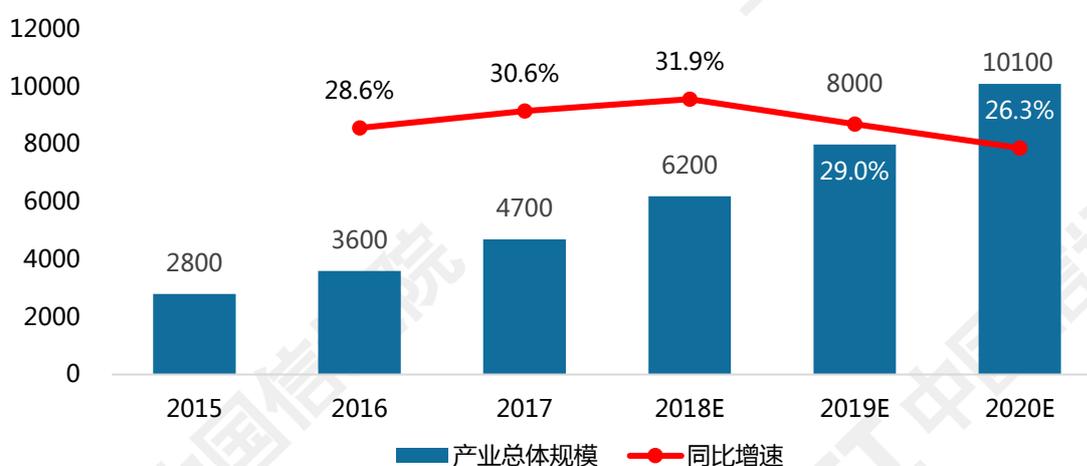
大数据硬件：本报告所称大数据硬件，是企业购买的用于支撑企业内部大数据相关业务的硬件基础设施，主要包括服务器、网络设备和存储设备等。特别说明，大数据系统一般使用通用硬件，技术上并无特殊性，但同时也包含部分经过专门优化的一体机设备。

二、大数据产业规模

(一) 大数据产业总体规模

对于我国大数据产业的规模,目前各个研究机构均采用间接方法估算。中国信息通信研究院结合对大数据相关企业的调研测算,2017年我国大数据产业规模为4700亿元人民币,同比增长30%。

图3 中国大数据产业总体规模及增速(单位:亿元)

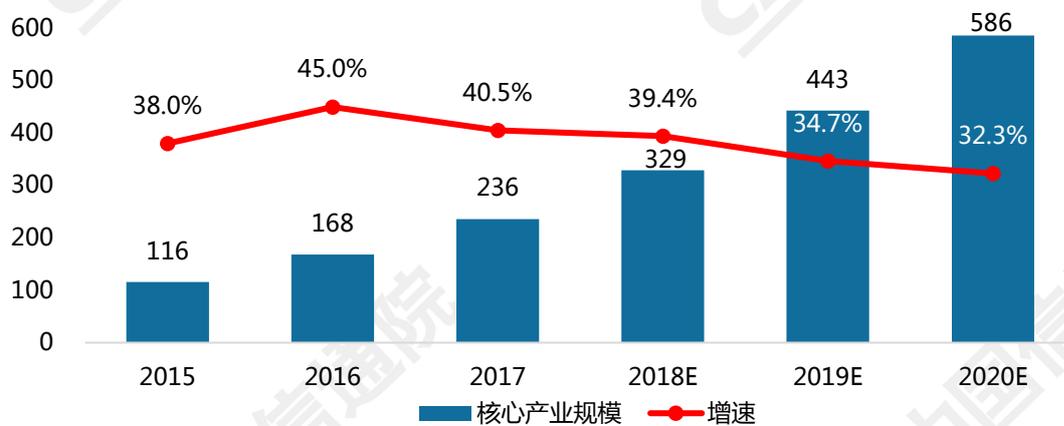


数据来源:中国信息通信研究院

(二) 大数据核心产业规模

中国信息通信研究院结合对大数据相关企业的调研得出,2017 年中国大数据核心产业规模为 236 亿元人民币,较 2016 年相比增速达 40.5%。预计未来几年,大数据市场仍将保持 30%以上的增速,到 2020 年中国大数据市场规模将达到 586 亿元人民币。

图 4 中国大数据核心产业规模及增速(单位:亿元)

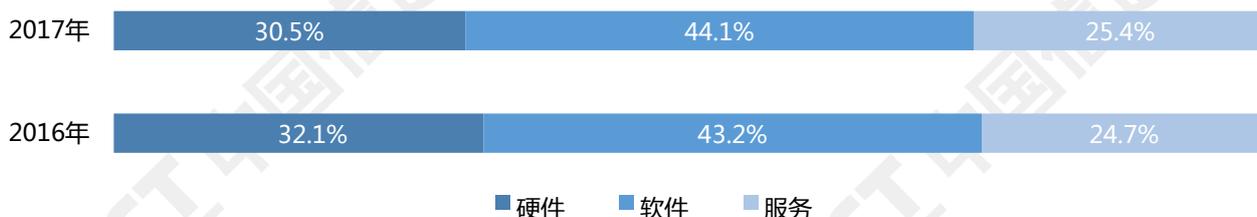


数据来源:中国信息通信研究院

(三) 细分领域规模

软件及服务依然占据大数据市场的主导地位,硬件比例略有下降。2017 年中国大数据硬件市场规模为 72.0 亿元,占比为 30.5%,与 2016 年相比下降了 1.6%;软件市场规模为 104.1 亿元,占比达到 44.1%,比 2016 年上升 0.9 个百分点;服务市场规模为 59.9 亿元,较去年上升了 0.7%。

图 5 中国大数据细分领域构成



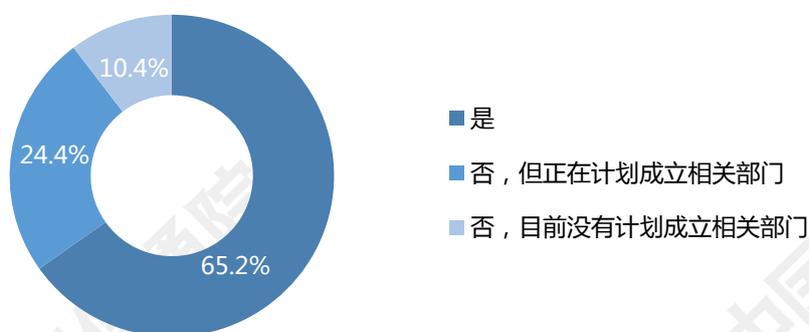
数据来源:中国信息通信研究院

三、大数据应用

(一) 大数据应用现状

接近 2/3 的企业已经成立了数据分析相关部门，企业对数据分析的重视程度进一步提高。调查结果显示，65.2%的企业已成立数据分析部门；另外，目前没有数据分析相关部门，但正在计划成立相关数据部门的企业占比为 24.4%。

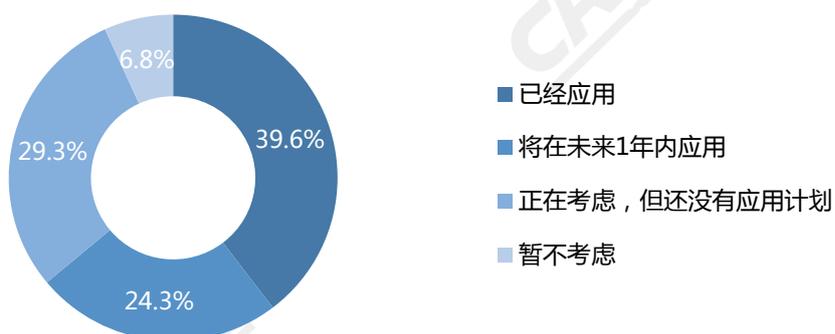
图 6 企业数据分析相关部门建设情况 (N=1,572)



数据来源:中国信息通信研究院

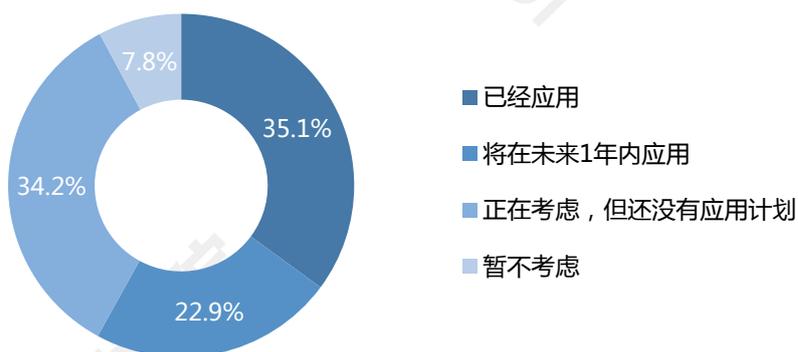
近四成的企业已经应用了大数据。在接受调查的 1572 家企业中，已经应用大数据的企业有 623 家，占比为 39.6%，与 2016 年相比上升 4.5%，垂直行业中如金融等领域大数据应用增加趋势较为明显。此外，24.3%的企业表示未来一年内将应用大数据。

图 7 2017 年企业对大数据的应用状况 (N=1,572)



数据来源:中国信息通信研究院

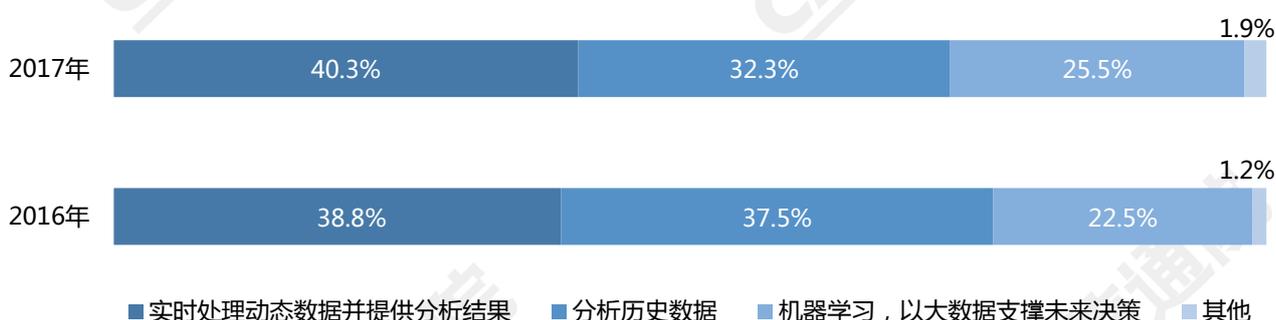
图 8 2016 年企业对大数据的应用状况 (N=1,465)



数据来源:中国信息通信研究院

40.3%的企业选择使用实时动态数据进行分析。对数据分析方式选择情况的调查显示，40.3%的企业采取实时处理动态数据并提供分析结果，占比最高；其次是分析历史数据和通过机器学习进行辅助决策，占比分别为 32.3%和 25.5%。未来，随着人工智能技术的发展和应用普及，选择机器学习进行辅助决策的企业占比有望进一步提升。

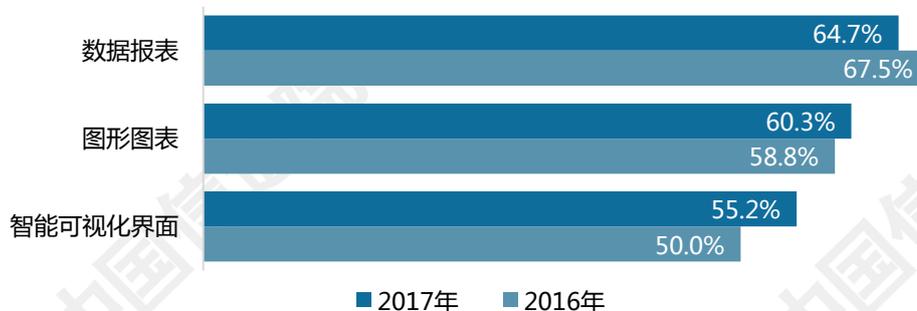
图 9 企业数据分析方式 (N=623)



数据来源:中国信息通信研究院

通过智能可视化界面展示数据分析结果的企业占比为 55.2%，与去年相比有较大幅度提升。数据报表依旧是企业选择展现数据的最主要方式，占比为 64.7%；其次为图形图表 (60.3%)。与 2016 年相比，利用智能可视化的方式展示数据分析结果的企业占比增幅达到 5.2%。

图 10 企业数据分析结果的展现方式 (N=623)

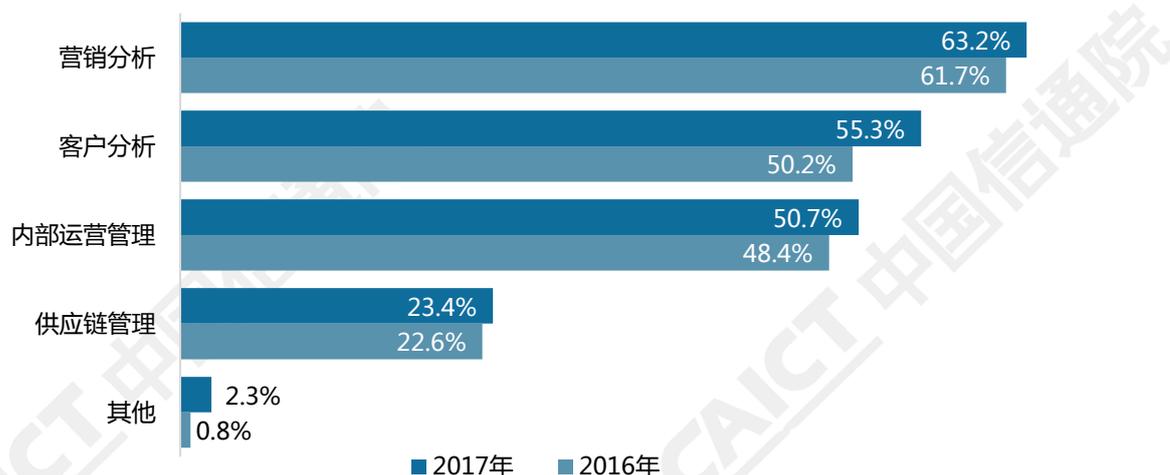


数据来源:中国信息通信研究院

(二) 大数据应用场景

营销分析、客户分析和内部运营管理仍是企业大数据应用场景最广泛的三个领域。已经应用大数据的企业中，将大数据用于营销分析的企业占比最高，达到 63.2%，与 2016 年相比上升了 1.5%；其次，55.3% 的企业将大数据应用于客户分析，比 2016 年提升了 5.1%，部分企业已将大数据应用于分析客户反馈并改进产品和服务等场景；另外，将大数据应用于内部运营管理的企业占比为 50.7%，与 2016 年相比上升了 2.3%。

图 11 企业大数据应用场景 (N=623)

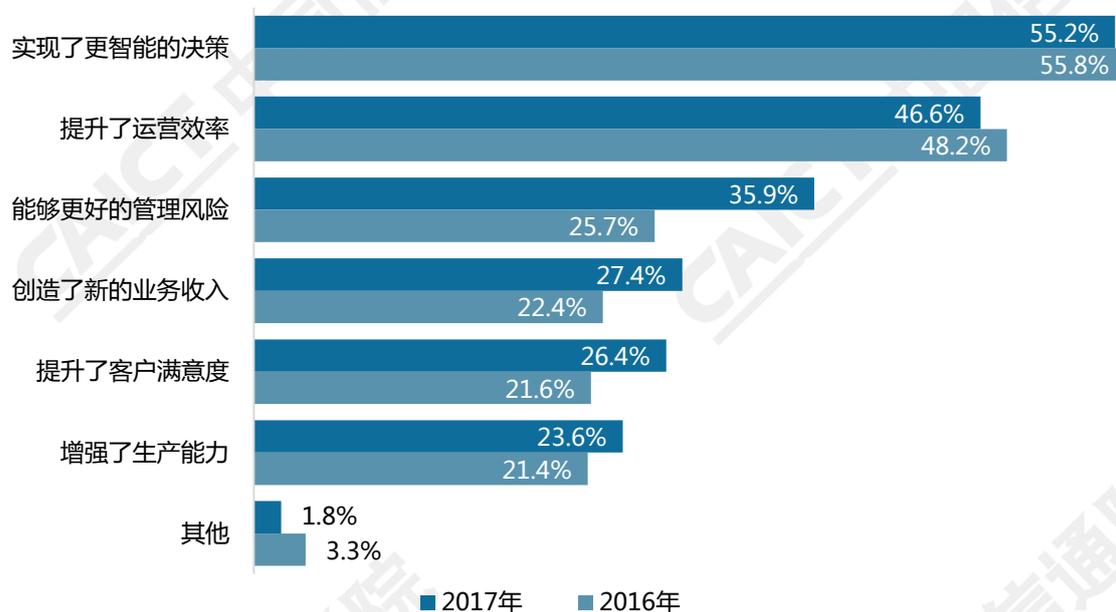


数据来源:中国信息通信研究院

(三) 大数据应用带来的效果

大数据应用在实现智能决策、提高运行效率和风险管理能力等方面具有较明显的效果。已经应用大数据的企业中，实现了智能决策的企业占比最高，达到 55.2%；其次，46.6% 的企业表示应用大数据后提升了运营效果；还有 35.9% 的企业认为应用大数据后能够更好的管理风险，与去年相比有较大幅度的提升 (10.2%)；此外，分别有 27.4%、26.4% 和 23.6% 的企业表示应用大数据后创造了新的业务收入、提升了客户满意度以及增强了生产能力。

图 12 企业应用大数据带来的效果 (N=623)

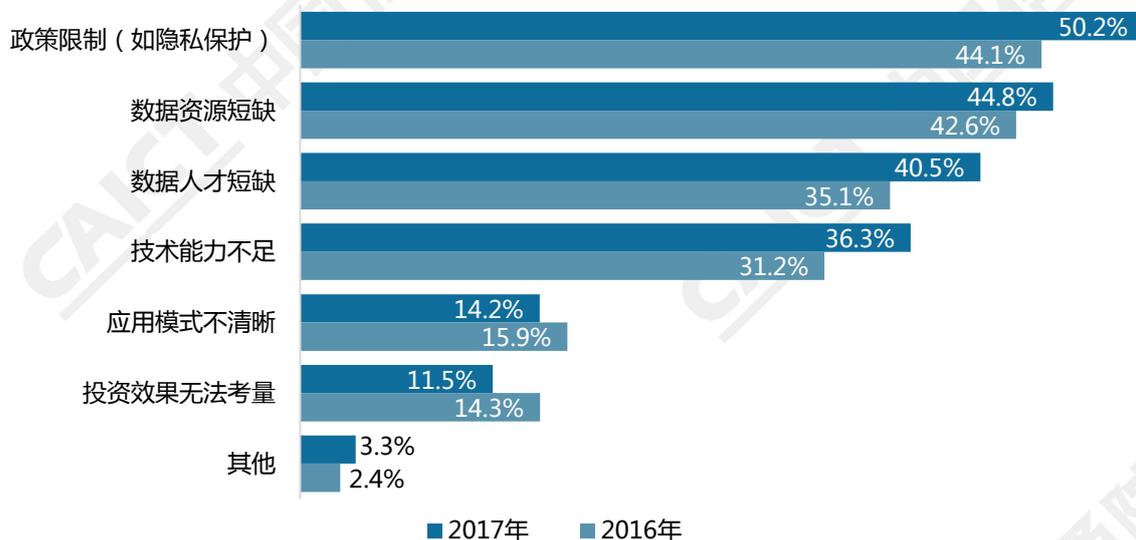


数据来源:中国信息通信研究院

(四) 大数据应用的主要障碍

政策限制、数据资源短缺和数据人才短缺是限制企业大数据发展最主要的三个因素。调查发现,隐私保护等政策不完善是制约企业大数据发展的首要因素,占比为 50.2%,比 2016 年上升了 6.1%。四成以上的受访企业认为数据资源短缺和数据人才不足是其发展大数据所面临的主要困难;其他制约因素还包括:技术能力不足(36.3%)、应用模式不清晰(14.2%)和投资效果无法考量(11.5%)。

图 13 制约企业大数据发展主要因素 (N=1,572)

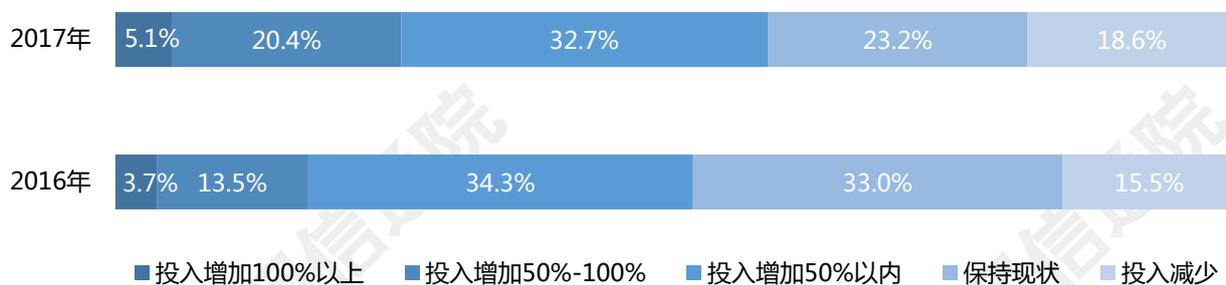


数据来源:中国信息通信研究院

(五) 未来大数据投入趋势

接近六成的企业在大数据领域的投入将持续增加。调查发现, 58.2%的受访企业预计未来将加大大数据投入力度, 相比 2016 年上升了 6.7%; 其中, 大数据投入增加超过 50%的企业占比达到了 25.5%; 另有 32.7%的企业在大数据方面的投入增加在 50%以内。

图 14 未来企业大数据投入情况 (N=1,572)



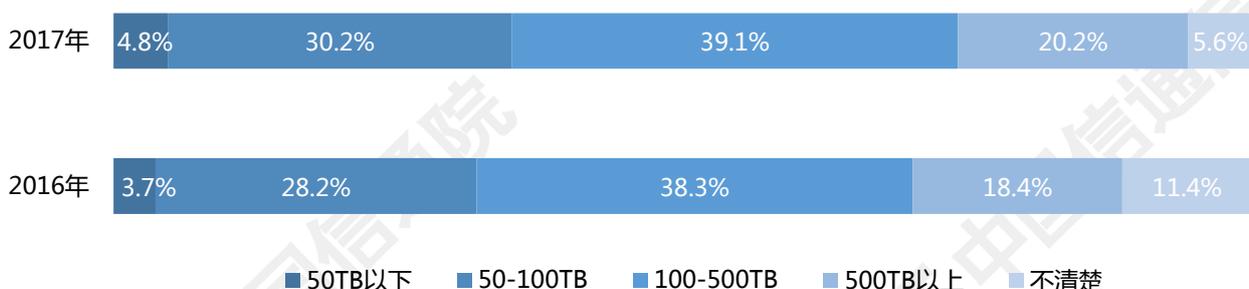
数据来源:中国信息通信研究院

四、大数据资源

(一) 数据规模

数据量在 50-500TB 的企业数量进一步增加。调查结果显示，30.2%的企业数据资源总量在 50-100TB 之间，与 2016 年相比提高了 2%；数据资源总量在 100-500TB 之间的企业占比为 39.1%，比 2016 年提高了 0.8%；与此同时，超过两成的企业数据资源总量达到 500TB 以上，该比例比 2016 年提高了 1.8%。

图 15 企业数据资源总量 (N=1,572)

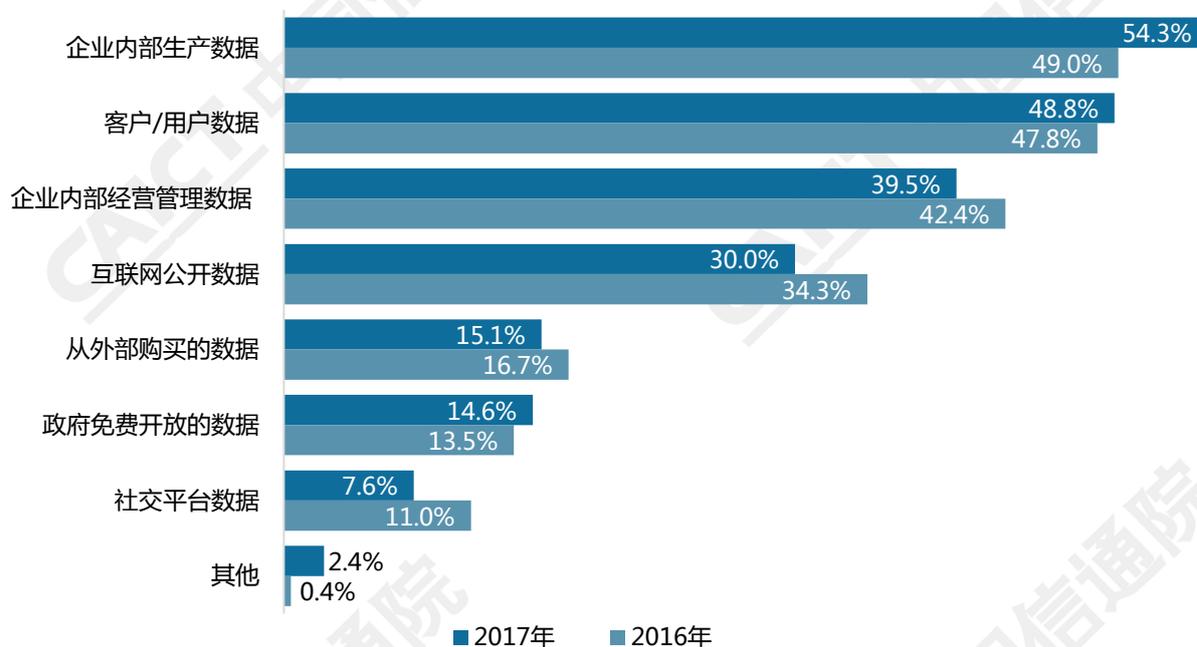


数据来源:中国信息通信研究院

(二) 数据来源

在受访企业中，内部生产数据和客户/用户数据是企业的两大数据来源。对企业数据来源的调查发现，54.3%和 48.8%的企业主要数据来源为内部生产数据和客户/用户数据，与 2016 年相比分别上升 5.3%和 1%；内部经营管理数据为主要来源的企业占比为 39.5%，与 2016 年相比降低了 2.9%；而将互联网公开数据（30%）、从外部购买数据（15.1%），以及政府免费开放的数据（14.6%）作为主要来源的企业占比仍处于较低水平。

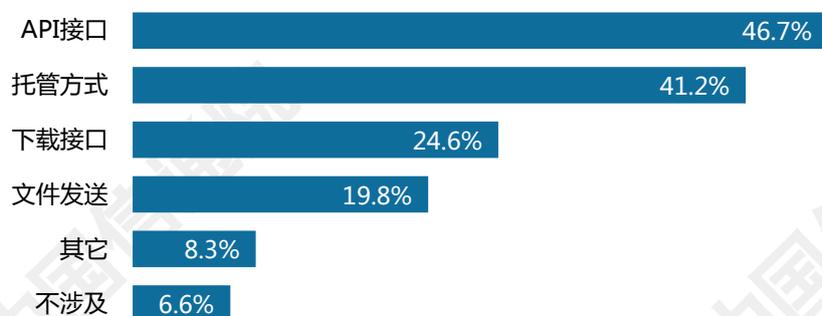
图 16 企业数据来源情况 (N=1,572)



数据来源:中国信息通信研究院

近半数企业选择采用 API 接口及托管方式吸纳外部数据。调查结果表明，API 接口和托管方式占比分别达到了 46.7%和 41.2%，远高于其他数据输入方式：下载接口（24.6%）、文件发送（19.8%）。

图 17 企业吸纳外部数据方式 (N=1,572)

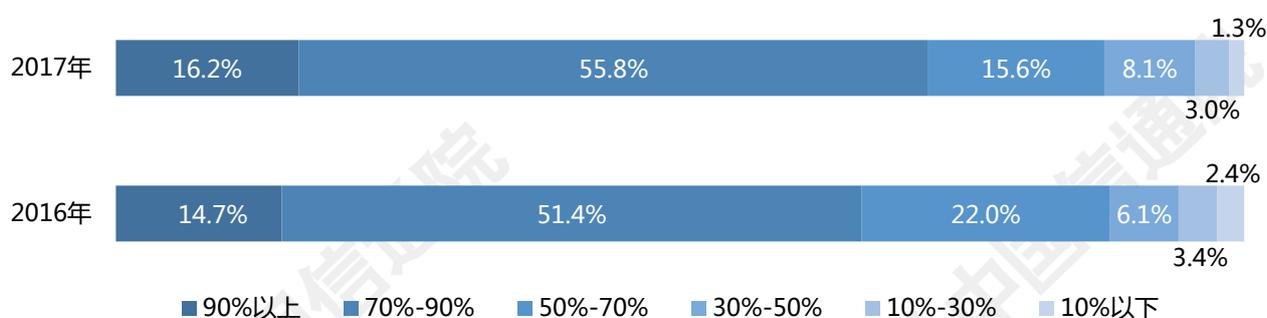


数据来源:中国信息通信研究院

(三) 数据类型

87.6%的企业非结构化数据已超过一半比例。其中,非结构化数据比例超过 90%的企业占比为 16.2%,与 2016 年相比上升 1.5%;非结构化数据占数据总量 70%-90%的企业占比 55.8%,比 2016 年上升了 4.4%;非结构化数据比例在 50%-70%之间的企业占比为 15.6%。此外,还有 4.3%的企业非结构化数据比例在 30%以下。

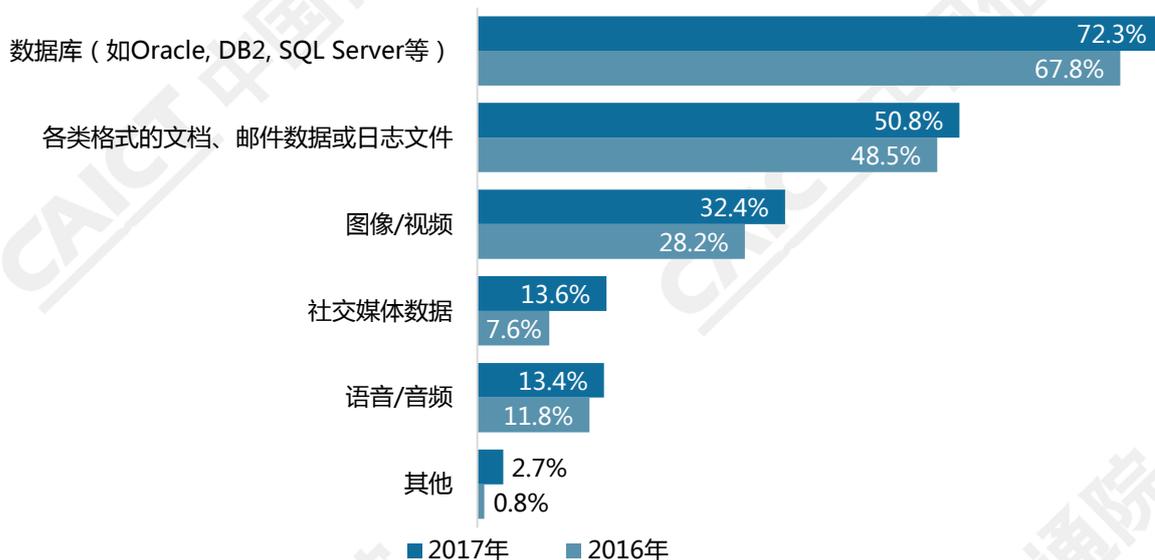
图 18 企业非结构化数据比例 (N=1,572)



数据来源:中国信息通信研究院

企业对数据库的分析利用率最高,与去年相比上升了 4.5%。在对企业各类数据分析利用情况的调查中,以结构化数据为基础的数据库进行大数据分析的企业占比为 72.3%,较去年上升了 4.5%,32.4%和 13.6%的企业对图像/视频和社交媒体数据等非结构化数据进行了分析,与 2016 年相比分别上升了 4.2%和 6%。近年来语音和图像识别技术发展迅速,在一定程度上提高了企业对于语音/音频和图像/视频等非结构化数据的利用率。

图 19 企业各类数据分析利用情况 (N=1,572)

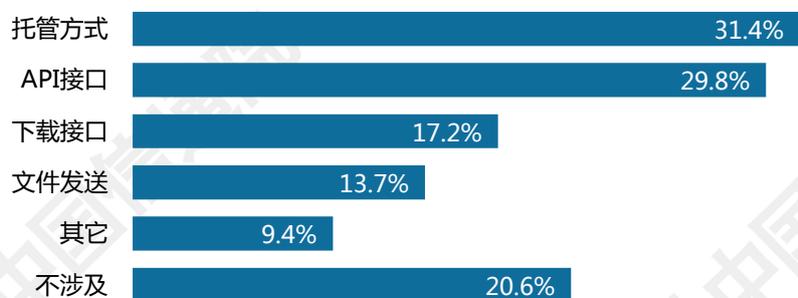


数据来源:中国信息通信研究院

(四) 数据共享

目前企业在数据共享过程中主要通过托管方式和 API 接口进行数据输出。企业输出数据方式的调查显示,受访企业中,分别有 31.4%和 29.8%的企业倾向选择托管方式和 API 接口进行数据输出;而采用提供下载接口进行数据输出的企业占比达到 17.2%;此外,使用文件发送方式输出数据的企业占比为 13.7%。

图 20 企业输出数据方式 (N=1,572)

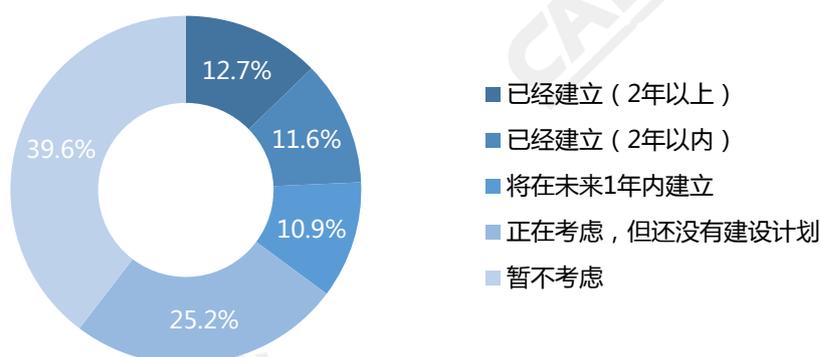


数据来源:中国信息通信研究院

(五) 数据管理

在受访的企业中，24.3%的企业已经设立了数据管理委员会。设立数据管理委员会在两年以上的企业占比达到了 12.7%；同时，10.9%的受访企业计划未来一年内成立数据管理委员会；另外正在考虑设立数据管理委员会的企业占比达到了 25.2%。

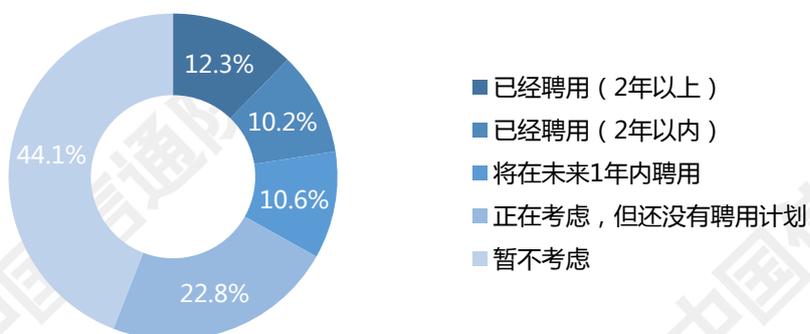
图 21 企业数据管理委员会建立情况 (N=1,572)



数据来源:中国信息通信研究院

22.5%的企业已经聘用专职首席数据官 (CDO)。调查发现，12.3%的受访企业聘用专职首席数据官已经超过两年；聘用专职首席数据官 (在两年以内) 的企业占比为 10.2%。此外，33.4%的受访企业表示将在一年内聘用或正在考虑聘用专职首席数据官。随着对数据管理和利用数据支撑战略决策的企业越来越多，未来聘用专职首席数据官的企业比例有望进一步提升。

图 22 企业专职首席数据官聘用情况 (N=1,572)



数据来源:中国信息通信研究院

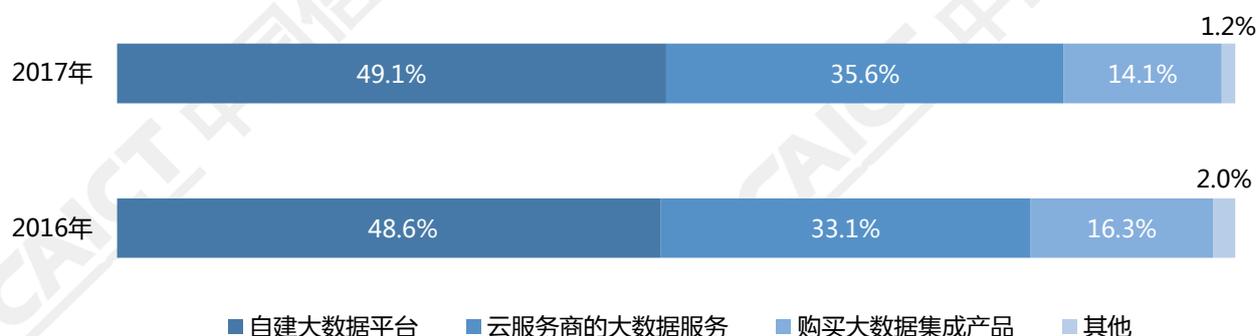
五、大数据平台

(一) 建设模式

本次调研根据大数据平台建设模式分为本地部署大数据平台和采购公有云服务模式。本地部署大数据平台包括：1、购买大数据相关产品或解决方案，在本地建设大数据相关基础设施、存储和计算平台、分析工具和应用展示等；2、采购大数据服务提供商的集成产品，如大数据一体机。采购公有云服务指的是购买某一家大数据云服务提供商所提供的在线大数据服务。

本地部署仍是企业进行大数据平台建设的首选方式，同时选择云服务商大数据服务的企业占比略有提升。在对企业大数据平台建设模式的调查中发现，选择自建大数据平台企业占比最多（49.1%）；而 35.6% 的企业倾向选择云服务商提供大数据服务，较 2016 年上升了 2.5%；另外购买大数据集成产品的企业占比为 14.1%，与去年相比降低了 2.2%。

图 23 企业大数据平台建设模式 (N=1,572)



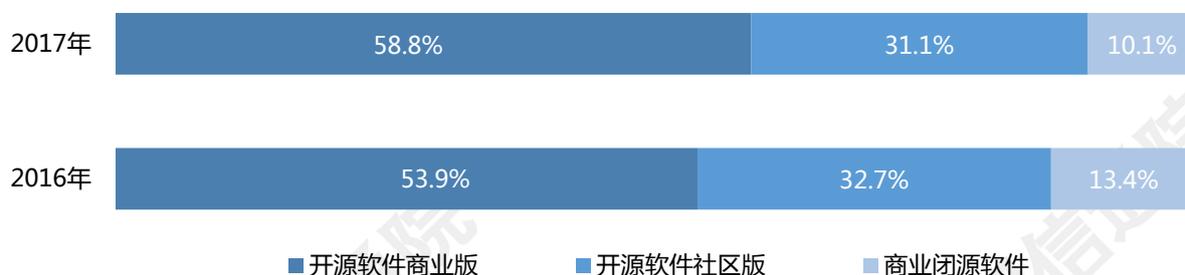
数据来源:中国信息通信研究院

(二) 自建大数据平台模式

1、软件类型选择

企业对开源大数据平台软件的接受程度大幅度提升。调查发现，开源软件依然是企业自建大数据平台时优先选择的软件，占比达到了 89.9%，与 2016 年相比增加了 3.3 个百分点。其中，倾向选择开源软件商业版的企业占比为 58.8%，比 2016 年提升了 4.9%。

图 24 自建大数据平台软件类型选择 (N=1,572)

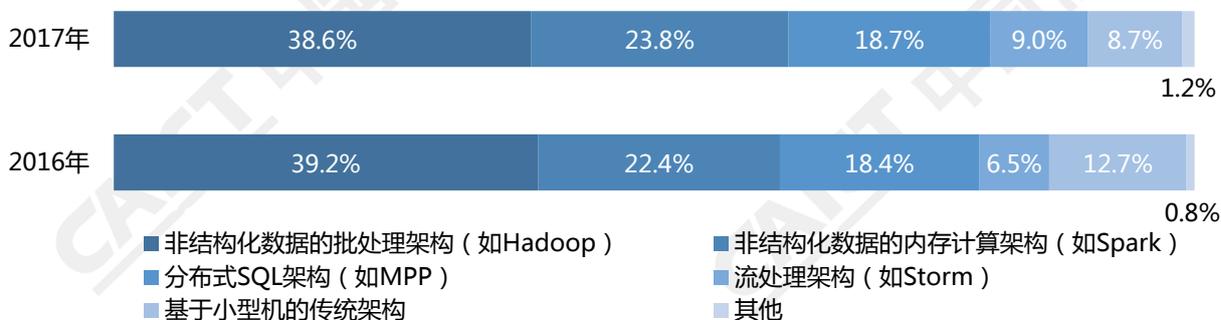


数据来源:中国信息通信研究院

2、平台架构选择

基于非结构化数据架构的大数据平台应用最为广泛。分别有 38.6%和 23.8%的受访企业选择非结构化的批处理架构(如 Hadoop)以及非结构化数据的内存架构(如 Spark)作为大数据平台的技术架构;其次, 18.7%的受访企业倾向使用分布式 SQL 架构(如 MPP);此外,采用流处理架构(如 Storm)的企业占比为 9%,与 2016 年相比上升了 2.5%。

图 25 自建大数据平台技术架构选择 (N=1,572)

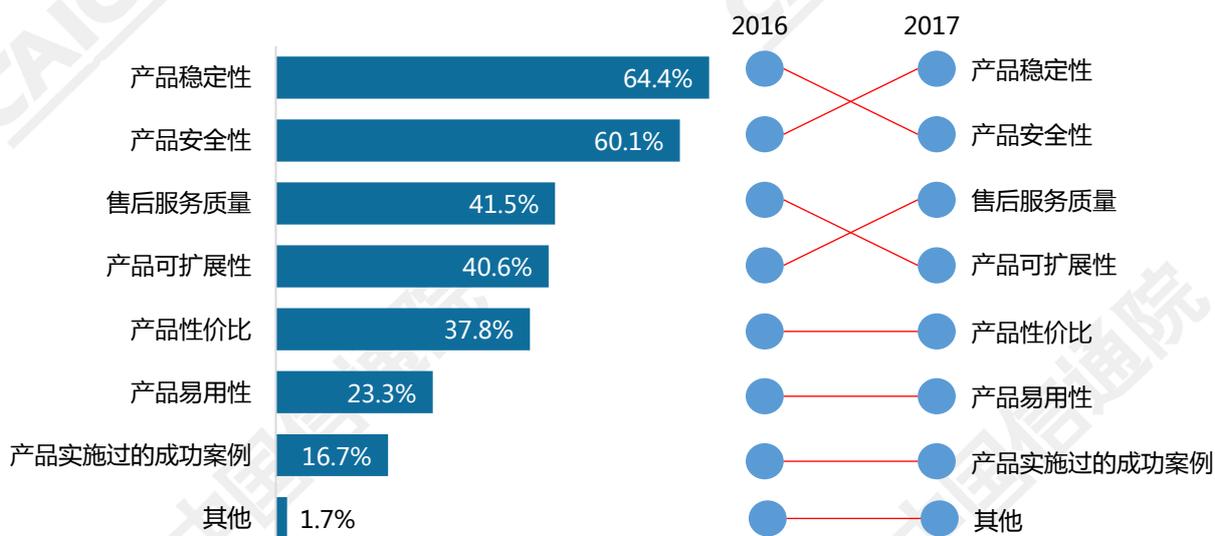


数据来源:中国信息通信研究院

3、选择大数据软件的考虑因素

产品稳定性和安全性依然是企业自建大数据平台软件的优先考虑因素。调查发现，64.4%的企业在选择大数据软件时首先考虑产品的稳定性；其次是产品安全性，该比例较2016年（71.4%）下降了11.3%。另外，超过40%的受访企业表示在选择大数据软件时会考虑售后服务质量和产品的可扩展性。

图 26 企业选择大数据平台软件的考虑因素 (N=1,572)

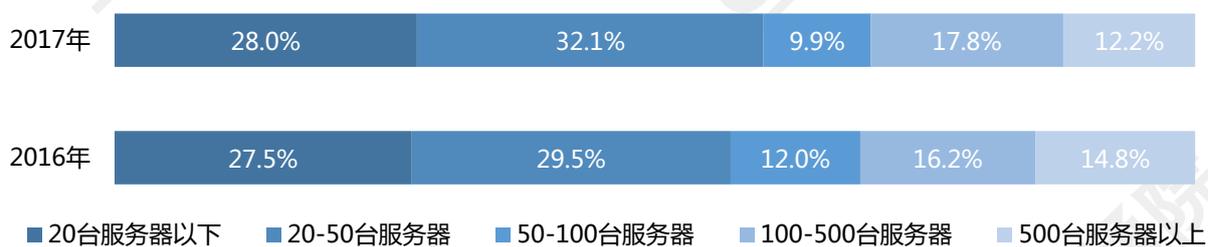


数据来源:中国信息通信研究院

4、服务器规模

六成以上的受访企业自建大数据平台的服务器数量在 50 台以下。调查显示,28%的企业部署的服务器数量在 20 台以下,与去年相比提高了 0.5 个百分点;服务器数量在 20-50 台的企业占比超过三成;部署服务器数量在 100-500 台的企业占比达到了 17.8%,比 2016 年上升了 1.6%。

图 27 自建大数据平台的服务器数量 (N=393)

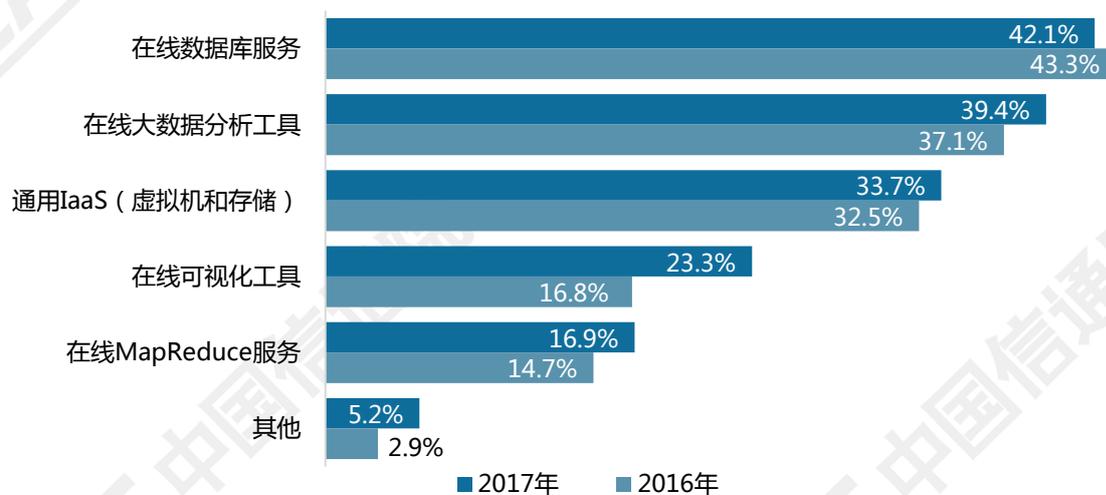


数据来源:中国信息通信研究院

(三) 采购公有云服务模式

与其他产品相比,企业对在线可视化工具的选择增幅最大。采用在线可视化工具的企业占比为 23.3%,与 2016 年相比提高了 6.5%;增幅次之的是在线大数据分析工具 (2.3%)。

图 28 企业对云服务产品类型选择 (N=1,572)

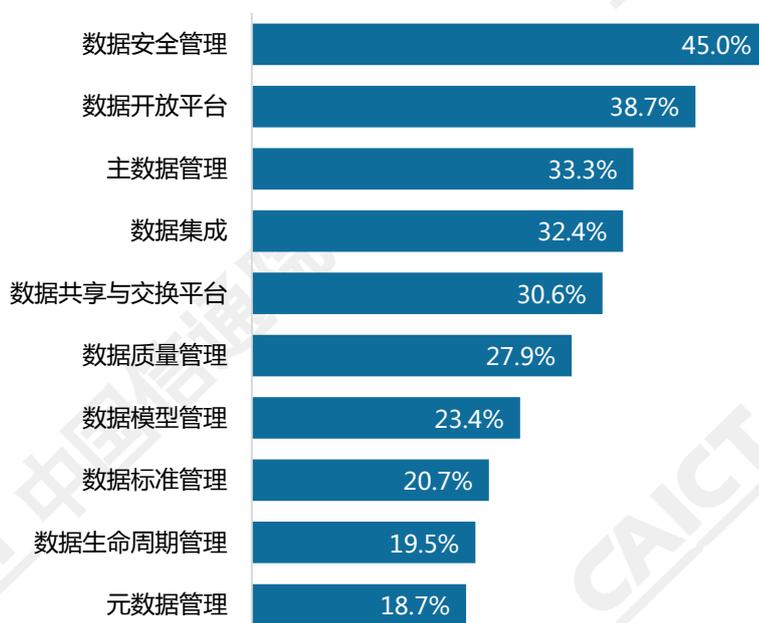


数据来源:中国信息通信研究院

(四) 大数据产品选型

在众多大数据产品及服务中，企业的对数据安全的需求最为强烈。调查结果显示，45%的企业倾向选择数据安全，占比最高。其次，企业对于数据开放平台和主数据管理的需求占比分别为 38.7%和 33.3%。其他还包括：数据集成（32.4%）、数据共享与交换平台（30.6%）、数据质量管理（27.9%）以及数据模型管理（23.4%）等。

图 29 企业大数据产品选型 (N=1,572)



数据来源:中国信息通信研究院

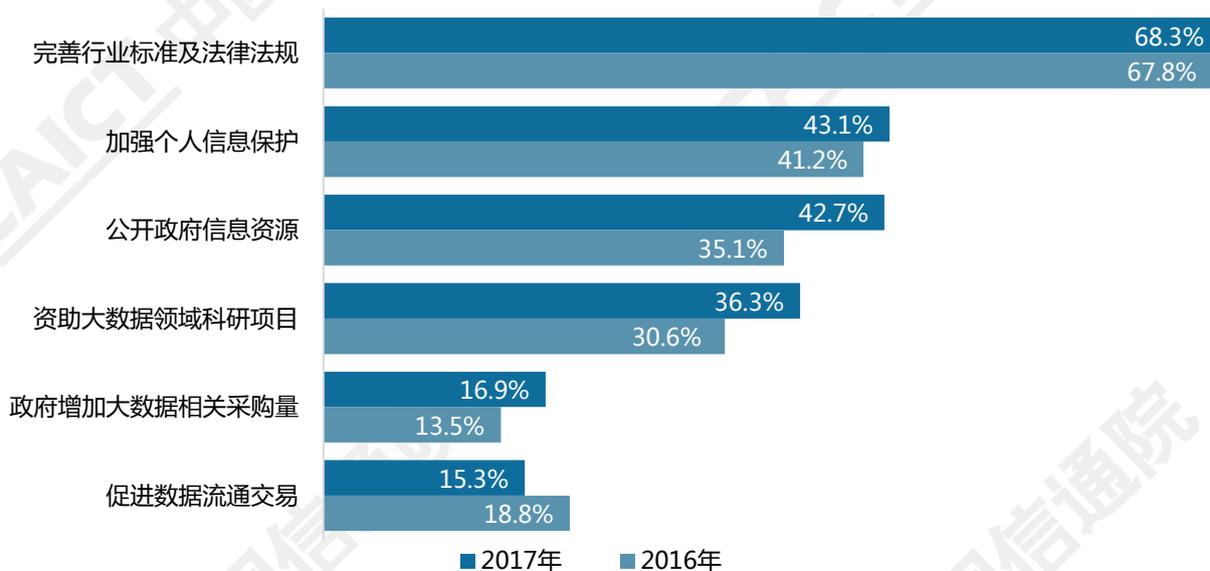
六、政策需求和资源需求

(一) 政策需求

为加快实施国家大数据战略，推动大数据产业健康快速发展，工业和信息化部编制并印发《大数据产业规划（2016-2020年）》（以下简称“《规划》”）。《规划》提出了大数据产业发展的目标：到2020年技术先进、应用繁荣、保障有力的大数据产业体系基本形成，大数据相关产品和服务业务收入突破1万亿元，年均复合增长率保持30%左右。

六成以上的企业认为完善行业标准、健全法律法规将有助于推动大数据的发展。据调查，68.3%的受访企业认为完善行业标准及法律法规对于推动大数据市场的发展至关重要；其次，43.1%的企业表示应加强个人信息保护，比2016年上升了1.9%；此外，企业对开放更多政府信息公开资源（42.7%）和资助更多大数据领域的科研项目（36.3%）的呼吁较高。

图 30 企业对政府和政策的需求 (N=1,572)

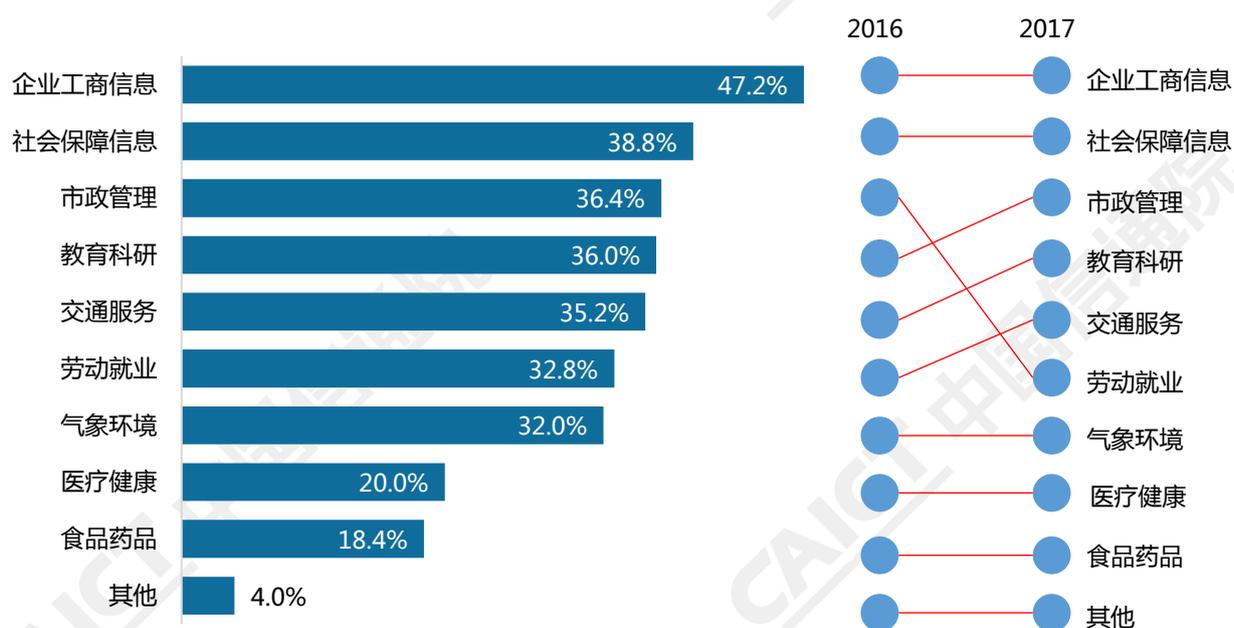


数据来源:中国信息通信研究院

(二) 资源需求

企业对市政管理、教育科研以及交通服务类公共数据资源的需求增强。在对政府公共数据资源需求情况调查中，2017 年企业对市政管理 (36.4%)、教育科研 (36%) 以及交通服务 (35.2%) 类数据资源需求均有一定程度的提升。

图 31 企业对数据资源的需求 (N=1,572)



数据来源:中国信息通信研究院

编后语

中国信息通信研究院已经连续三年开展中国大数据发展情况的市场调查工作，并发布《中国大数据发展调查报告》，在行业内也产生了重要影响。随着大数据技术的发展和大数据应用的深入，我院将持续跟踪大数据发展情况。欢迎大数据领域广大企业、研究机构、专家学者和从业人员与报告编写组联系（发邮件至：liuhan@caict.ac.cn），提出您感兴趣的调研问题，反馈您的宝贵意见和建议，帮助我们持续提升报告的针对性和实用价值。

感谢您对大数据产业的支持！

中国信息通信研究院

2018 年 4 月



关注我们



关注我们

中国信息通信研究院

地 址：北京市海淀区花园北路52号

邮政编码：100191

联系电话：010-62304839

传 真：010-62304980

