



中国大数据 技术与产业发展报告 (2015)

CCF大数据专家委员会
2015年12月10日





《发展报告》形成

1. 自2012年成立以来，中国计算机学会大数据专家委员会已经连续两年发布了《中国大数据技术与产业发展报告》。《发展报告》汇聚大数据专家委员会一百多位专家的知识与智慧，为业界梳理大数据应用现状及发展趋势，为政府制定产业发展政策提供建议；同时，探讨大数据研究面临的科学问题和技术挑战，为科研机构 and 科研人员提供参考指南。
2. 今年大数据专家委员会在国务院出台的《促进大数据发展行动纲要》的背景下，对《纲要》进行专家解读，并据此编写了《2015中国大数据技术与产业发展报告》，深入阐述国家大数据战略的重要性以及相关的历史背景和国内外现状，并讨论技术难点和应对策略，一定程度上反映了我国大数据学术界和产业界的共识。



《发展报告》编委会

主 编：李国杰

副主编：程学旗 王晓阳 王建民

编 委：何震瀛 靳小龙 李德芳 潘柱廷
孙彦广 谭建荣 王 晨 王元卓
杨东日 周 涛 周晓方 赵劲松
(按姓名拼音排序)





《发展报告》大纲

第一篇 中国大数据开放共享发展报告

- 第1章 引言
- 第2章 数据开放共享是社会发展的驱动力
- 第3章 数据开放共享国内外现状
- 第4章 数据开放的风险与对策
- 第5章 促进中国大数据开放共享的探索

第二篇 中国工业大数据发展报告

- 第7章 工业大数据之新一轮工业变革
- 第8章 《中国制造2025》与工业大数据
- 第9章 工业大数据发展历程
- 第10章 离散工业大数据
- 第11章 流程工业大数据

第三篇 大数据发展趋势预测

2015中国大数据技术大会正式发布后收录

附录



意见征集

今天首发布《发展报告》样版，藉大会影响力广泛征集各方建议！

希望拿到《发展报告》样版的与会者，积极给CCF大专委反馈意见，修订版将提交出版社正式出版，感谢！

CCF大数据专家委员会邮箱：bigdata@ccf.org.cn





民生、多样、政策、生态

2016年大数据发展趋势预测

CCF大数据专家委员会官方发布
2015年12月10日



调研的目标、方法和过程



- 委员大数据专家委委员
 - 学术界、产业界、政府、海外、跨界特邀委员
- 中关村大数据产业联盟会员



2016年大数据发展趋势十大预测



发展趋势十大预测

民生

多样

政策

生态

1. 可视化推动大数据平民化
2. 多学科融合与数据科学的兴起
3. 大数据安全与隐私令人忧虑
4. 新热点融入大数据多样化处理模式
5. 大数据提升社会治理和民生领域应用
6. 《促进大数据发展行动纲要》驱动产业生态
7. 深度分析推动大数据智能应用
8. 数据权属与数据主权备受关注
9. 互联网、金融、健康保持热度，智慧城市、企业数据化、工业大数据是新增长点
10. 开源、测评、大赛催生良性人才与技术生态



可视化推动大数据平民化

1

可视化让大数据靠近国计民生

可视化让技术高深的大数据在政策决策者和普通百姓中得到更好的理解，进而在国计民生中发挥价值。

各种可视化技术和工具提升大数据分析

除了传统的可视化技术（各种图形）之外，如何表示Graph，特别是Graph相当大的情况下，并且方便用户进行交互，以及如何有效地对其它类型的数据进行可视化(社区的演变，时空轨迹等)，值得关注。

从电子表格到时间线动画再到3D可视化，大数据创业公司们提供了各种各样的分析工具和界面，有的面向数据科学家，有的选择绕过数据科学家直接面向业务部门。



多学科融合 数据科学雏形初显

2

数据科学带动多学科融合

在大数据时代，许多学科表面上看来研究的方向大不相同，但是从数据的视角来看，其实是相通的。随着社会的数字化程度逐步加深，越来越来多的学科在数据层面趋于一致。可以采用相似的思想来进行的统一的研究。

数据科学雏形初显、学科有待突破

数据科学作为一个与大数据相关的新兴学科出现。

各种数据实验室、数据研究院出现，新学科趋于成型。

针对大数据处理的理论研究上，新型的概率和统计模型将是主要的研究工具，学科基础理论的突破还难于在2016年出现。



大数据的安全与隐私保护

3

大数据的安全威胁、副作用、发展阻碍等问题越来越凸显

大数据的安全持续令人担忧

相对于传统的数据模式，大数据更加容易成为网络攻击的目标，大数据的不断增加，对数据存储的物理安全性要求会越来越高，数据的多副本与容灾机制也面临更高的要求。同时，大数据分析技术也更容易被黑客利用到攻击当中去。如今网络和数字化生活使得犯罪分子、网络黑客以及并无恶意的大数据服务提供商都更容易获得关于人的信息，也有了更多不易被追踪和防范的犯罪手段，可能会出现更高明的骗局。

推动数据立法，重视个人数据隐私



新热点持续融入大数据 多样化处理模式

4

大数据的处理模式依然多样化

大数据处理模式持续丰富，新旧手段不断融合

流数据、内存计算成为新热点

内存计算将继续成为提高大数据处理性能的主要手段。以Spark为代表的内存计算逐步走向商用，并与Hadoop融合共存。

Spark大数据平台显得非常有吸引力。其优势首先在于先进的内存计算技术，对比传统的硬盘处理方式，在性能上有了数量级的提升；其次，Spark生态系统可以同时支持批处理计算、流计算、交互查询计算、图计算等多种计算框架，数据使用效率大大提高

专为大数据处理优化的系统和硬件出现

CCF大数据专家委员会



大数据提升社会治理和 民生领域应用

5

基于大数据的社会治理，成为业界关注热点，涉及智慧城市、应急、税收、反恐、农业农村等多个民生领域

大数据从来都是应用驱动，技术发力。

在最易获得大数据应用成果的互联网环境之后，大数据走进国计民生成为必然。

而在2016年，与民生有关的应用将成为热点。

当然，国计与民生并不互斥。涉及民生的国计，将是快速发展的热点中的热点。



《促进大数据发展行动纲要》 驱动产业生态

6

《促进大数据发展行动纲要》成为产业生态快速发展完善的催化剂和标杆

在国务院2015年出台《促进大数据发展行动纲要》的背景下，各地政府将纷纷出台大数据产业和科研的利好政策。

政府牵引产业生态、带动数据共享交换

政府牵引的产业生态发展成为大数据发展历程在2016的特点。

政府带动的数据共享，将成为数据流转的源动力，让数据开放共享、交换交易成为产业生态的新态势。

大数据相关基础设施建设、大数据创业公司成为产业发展两大引擎

大数据就业岗位涌现

CCF大数据专家委员会



深度分析推动大数据智能应用

7

深度分析推动大数据智能应用

预测决策、精准推荐、语义化，这些涉及人的思维、影响、理解的延展，都成为大数据深度分析的关键应用方向。

深度学习成为大数据智能分析的核心技术

相比于传统机器学习算法，深度学习借助深层次神经网络模型，能够更加智能的提取数据不同层次的特征，对数据进行更加准确有效的表达。而且数据量越大，深度学习算法越有优势，可以得到更好的结果。目前，深度学习已经在图像分类检索、语音识别等领域产生重大突破。深度学习的兴起凸显出复杂机器学习模型在利用大数据方面的突出优势，进一步考虑大数据动态性、分布性、关联性的新型机器学习技术将很快涌现。



数据权属与数据主权备受关注

8

大数据价值化凸显数据权属的难题

大数据凸显了数据的巨大价值。而数据的权属问题并不是传统的财产权、知识产权等可以涵盖的权属问题。

数据成为国家间争夺的资源，数据主权成为网络空间主权的重要形态

数据成为重要的战略资源。人口红利、地大物博、经济实力、文化优势等等都纷纷体现为数据资源储备和数据服务影响力。

数据资源化价值化是根源



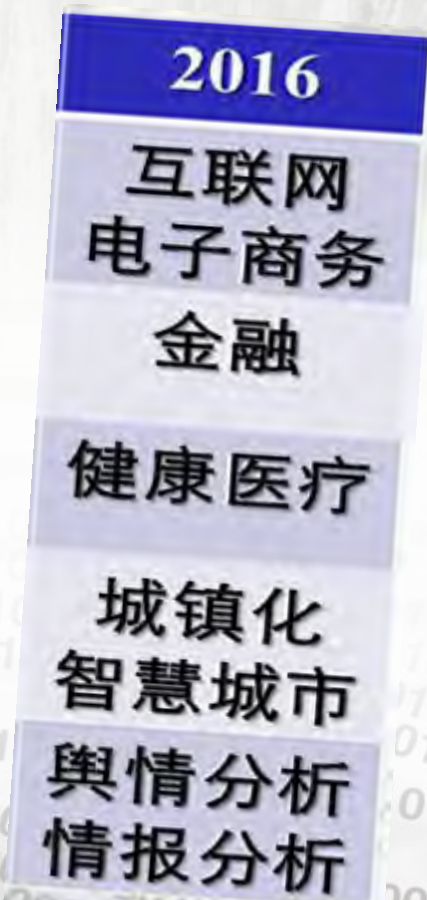
老三样持续热度 新三样是新增长点

9

互联网、金融、健康保持热度

智慧城市、企业数据化、工业大数据是新增长点

应用是大数据真实的推动力。





开源、测评、大赛 催生良性人才与技术生态

10

开源系统将成为大数据领域的主流技术和系统选择

以Hadoop为代表的开源技术拉开了大数据技术的序幕，大数据应用的发展又促进了开源技术的进一步发展。开源技术的发展降低了数据处理的成本，引领了大数据生态系统的蓬勃发展，同时也给传统数据库厂商带来了挑战。

测试、评估、标杆比对等第三方形态出现

对数据处理的能力、性能等进行测试、评估、标杆比对的第三方形态出现并逐步成为热点。相对公正的技术评价有利于优秀技术占领市场，驱动优秀技术的研发生态。

各类创业创新大赛完善人才生态

各类创业创新大赛纷纷举办，大赛为人才的培养和选拔提供了。



2013 - 2014 趋势对比

2013

1. 数据的资源化
2. 大数据的隐私问题突出
3. 大数据与云计算等深度融合
4. 基于大数据的智能化的出现
5. 大数据分析的革命性方法
6. 大数据安全
7. 数据科学兴起
8. 数据共享联盟
9. 大数据新职业
10. 更大的数据

2014

- 1、大数据从“概念”走向“价值”
- 2、大数据架构的多样化模式并存
- 3、大数据安全与隐私
- 4、大数据分析可视化
- 5、大数据产业成为战略性产业
- 6、数据商品化与数据共享联盟化
- 7、基于大数据的推荐与预测流行
- 8、深度学习与大数据智能成为支撑
- 9、数据科学的兴起
- 10、大数据生态环境逐步完善

大数据从“概念”走向“价值”



2014 - 2015 趋势对比

2014

- 1、大数据从“概念”走向“价值”
- 2、大数据架构的多样化模式并存
- 3、大数据安全与隐私
- 4、大数据分析可视化
- 5、大数据产业成为战略性产业
- 6、数据商品化与数据共享联盟化
- 7、基于大数据的推荐与预测流行
- 8、深度学习与大数据智能成为支撑
- 9、数据科学的兴起
- 10、大数据生态环境逐步完善

2015

1. 大数据分析成为数据价值化的热点
2. 数据科学带动学科融合，但自身尚未成体系
3. 与各行业结合，跨领域应用
4. “物云移社”融合，产生综合价值
5. 平台架构与基础设施
6. 大数据的安全与隐私保护
7. 计算模式：深度学习、众包计算
8. 可视化分析与可视化呈现
9. 大数据人才与教育
10. 开源系统将成为主流选择

跨界融合 基础突破



2015 - 2016 趋势对比

2015

1. 大数据分析成为数据价值化的热点
2. 数据科学带动学科融合，但自身尚未成体系
3. 与各行业结合，跨领域应用
4. “物云移社”融合，产生综合价值
5. 平台架构与基础设施
6. 大数据的安全与隐私保护
7. 计算模式：深度学习、众包计算
8. 可视化分析与可视化呈现
9. 大数据人才与教育
10. 开源系统将成为主流选择

2016

1. 可视化推动大数据平民化
2. 多学科融合与数据科学的兴起
3. 大数据安全与隐私令人忧虑
4. 新热点融入大数据多样化处理模式
5. 大数据提升社会治理和民生领域应用
6. 《促进大数据发展行动纲要》驱动产业生态
7. 深度分析推动大数据智能应用
8. 数据权属与数据主权备受关注
9. 互联网、金融、健康保持热度，智慧城市、企业数据化、工业大数据是新增增长点
10. 开源、测评、大赛催生良性人才与技术生态

民生
多样
政策
生态



2016年大数据发展趋势单项调研



最令人瞩目的应用领域

2013 - 14 - 15 - 16

2013	2014	2015	2016
医疗	互联网 电子商务	互联网 电子商务	互联网 电子商务
金融	金融	金融	金融
电子商务	健康医疗	健康医疗	健康医疗
城市管理	舆情分析 情报分析	城镇化 智慧城市	城镇化 智慧城市
		社会安全 犯罪侦查	舆情分析 情报分析



取得应用和技术突破的数据类型 2015 - 16

2015	2016
社会化媒体数据	城市数据
视频数据	互联网交易相关数据
互联网日志与电商交易数据	企业数据
语音数据、图形图像 设备测量和控制数据	视频数据 图形图像数据
人体数据、宏观经济	人体数据



在数据资源流转上会有什么举措

2016

2016

会自己收集大量数据

会利用数据提供服务

会买数据集

会只下载和获得免费数据集

没有数据流转考虑

会免费提供数据集

会提供数据流转和交易的中介服务

会卖数据

会进行国际数据交换



大数据的最佳拍档概念是 2016

2016

互联网+

云计算

智慧城市

物联网

移动互联网

大众创业万众创新
工业互联网（工业4.0）

智能生活设备

一带一路

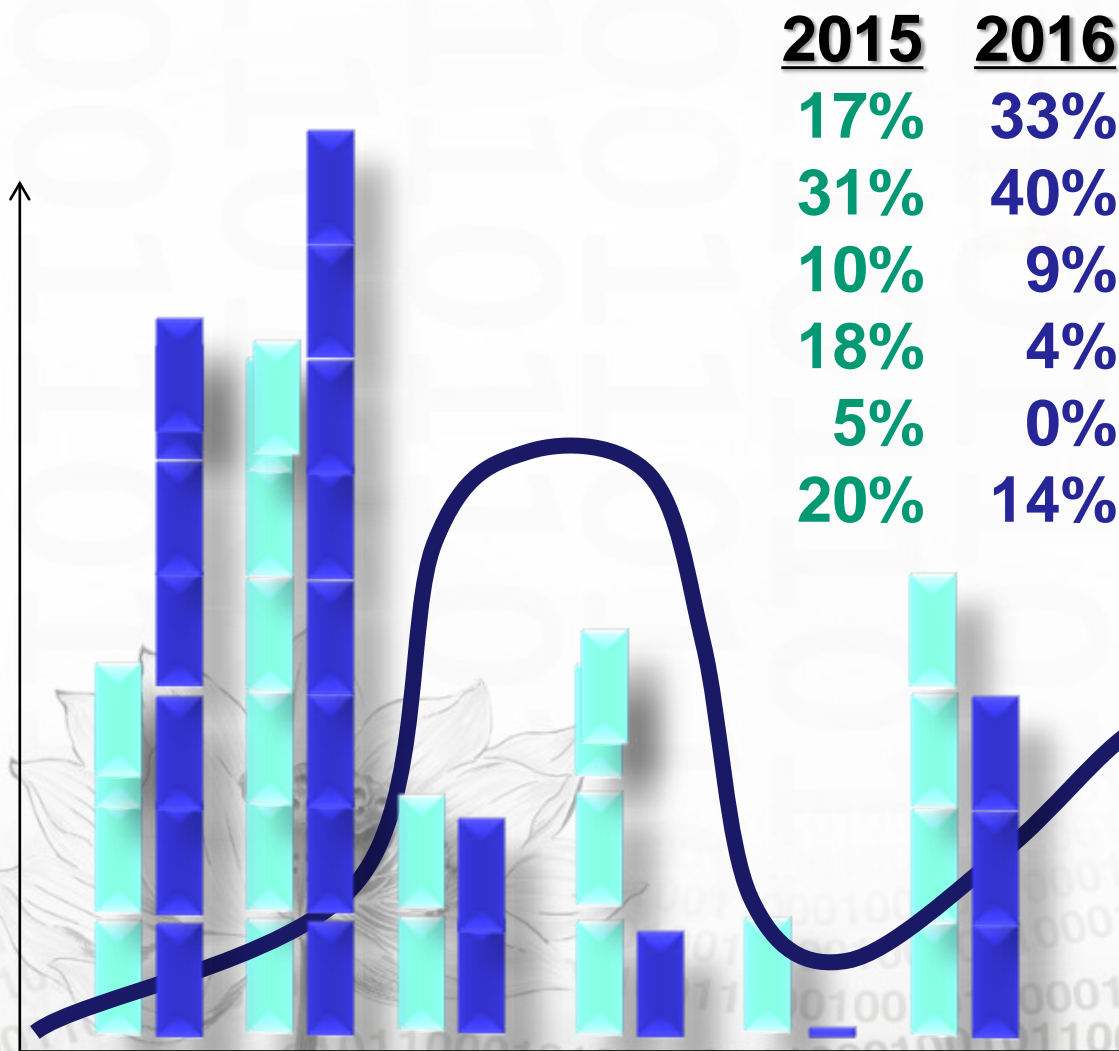


我国大数据发展的最主要推动者来自于 2015 - 16

2015	2016
大型互联网公司	大型互联网公司
政府机构	政府机构
国内大学和科研院所	创业企业
公共服务机构	
创业企业	



2015-16年大数据发展阶段判断



极为初级

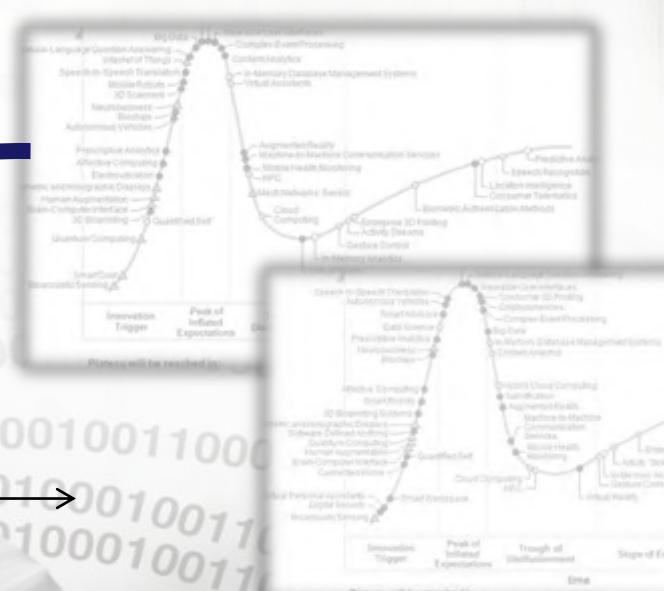
即将快速扩张

爆发增长中

达到一个顶峰，上升乏力

达到一个顶峰，将下降和幻灭

稳步成长中





多谢关注！

详细解读敬请关注高峰论坛
以及下一期《中国计算机学会通讯》《大数据》

CCF大数据专家委员会
2015年12月10日

