

2017年中国大数据发展趋势预测及解读

关键词：大数据发展趋势

CCF大数据专家委员会

CCF 大数据专家委员会日前发布了 2017 年中国大数据发展趋势十大预测。第一，机器学习继续成为智能分析的核心技术。第二，人工智能和脑科学相结合，成为大数据分析领域的热点。第三，大数据的安全和隐私持续令人担忧。第四，多学科融合与数据科学兴起。第五，大数据处理多样化模式并存融合，流计算成为主流模式之一。第六，数据的语义化和知识化是数据价值的基础问题。第七，开源成为大数据技术生态主流。第八，政府大数据发展迅速。第九，推动数据立法、重视个人数据隐私。第十，可视化技术和工具提升大数据分析工具的易用性。

2017年中国大数据发展趋势特点

十大预测分为三类：

第一，纯技术相关的预测占七条。与以往预测有很多关于产业、资本、政策、生态、人才教育、职业等方面的条目不同，这次预测大家更加关注技术本身，包括机器学习、人工智能、学科融合、处理多样性、流计算、语义化、开源、可视化等。

第二，安全和隐私问题。安全和隐私在每年的预测中都会排在第三或第四的位置上。目前，安全和隐私几乎成为现在阻碍大数据发展的唯一制约因素。对于安全和隐私的担忧使得数据的开放步伐放慢，使得大家认为对大数据的深度应用可能会带

来意想不到的负面价值。针对这个问题，除了要通过安全保障措施来解决安全和隐私的侵害问题外，还要对我们的担忧加以慰藉，从而松绑大数据的发展。但是，现在还没有一个特别好的办法来解除这个问题，尚处在大数据侵害爆发前的窗口期，是一个非常令人担忧的时期。

第三，政策法规的动力和规范作用。从第八条和第九条预测可以看出，专家们除了关注技术突破，还期望在政策和法规上能给予大数据发展有效的帮助。换句话说，我们在技术之外的发展诉求中，在市场力量和政府力量的天平中，大家更偏重于谋求政府的力量给予支持。这种支持可能直接通过第八条政府的应用需求来产生动力，也可能通过推动立法和规章制度形成合规性力量来构建良好的环境。

五年趋势预测的回顾

CCF 大数据专家委员会是从 2012 年底开始进行预测中国大数据发展趋势的。现在简单地回顾一下这五次预测之间的变化。

当时对 2013 年中国大数据发展趋势的预测，都是概念化、框架化的。如：数据的资源化；隐私问题突出；大数据与云计算等深度融合等。

2014 年开始关注价值、数据背后的价值。用一句话概括，就是“大数据从概念走向价值”。如：大数据架构的多样化模式并存；大数据分析与可视化等。

2015年的预测，出现了一些更加具体的趋向性的引导，强调跨界，强调融合，强调在基础科学和基础技术上寻求突破。如：大数据分析成为数据价值化的热点；数据科学带动学科融合，但自身尚未成体系等。

2016年的预测体现了一些政策、战略方面的认识，比如关注民生。关注多样化的技术，同时对政策和生态给予比较多的关注。如：可视化推动大数据平民化；大数据提升民生治理等。

而对于2017年的预测，可以用下面这些关键词来表述：

机器学习、人工智能与脑科学、安全隐私、学科融合与数据科学、多样化处理、流计算、语义化和知识化、开源、政府大数据、立法、可视化。

上述预测并没有超出大家的总预期，没有出现所谓的“黑天鹅”或者是突破性的变化。总之，2017年预测可以总结为：回归技术突破，安全担忧犹在，期待政府力量。

大数据专项调研

CCF大数据专家委除了给出大数据发展趋势的十大预测之外，还做了一些专项分析调研。大数据发展的根本在于应用。从“最令人瞩目的应用领域”的单项调研中，互联网和金融一直是排在前三的领域。在这次的调研中有了变化，排在其后的城镇化和制造业等领域与前三相去甚远，而健康超越金融排在第二。这也许和互联网金融及其所面临的风险

最令人瞩目的应用领域

	2013年	2014年	2015年	2016年	2017年
医疗		互联网 电子商务	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	健康医疗
电子商务		健康医疗	健康医疗	健康医疗	金融
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	城镇化 智慧城市
			社会安全 犯罪侦查	舆情分析 情报分析	制造业 工业BD

有某种关联，专家们对应用领域的关注度有了变化。

大数据最重要的是数据本身。“哪个数据类型会得到重要的突破”？城市数据、图形图像数据、语音数据和视频数据这四方面，专家们认为会有突破的点。这里强调的是突破，而不是取得巨大的成绩。这也可以解释为什么2017年互联网交易数据从突破数据中消失。

取得应用和技术突破的数据类型

2015年	2016年	2017年
社会化媒体数据	城市数据	城市数据
视频数据	互联网交易相关数据	图形图像数据
互联网日志与电商交易数据	企业数据	语音数据 视频数据
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	

在数据资源流转上的举措

2016年	2017年
自己收集大量数据	自己收集大量数据
利用数据提供服务	利用数据提供服务
买数据集	只下载和获得免费数据集
只下载和获得免费数据集	买数据集
没有数据流转考虑 免费提供数据集 提供数据流转和交易的中介服务 卖数据 进行国际数据交换	免费提供数据集

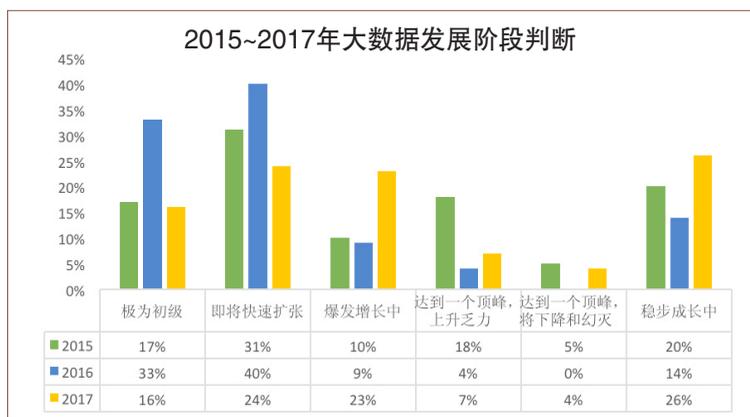
做研究和应用都有一个切实难题就是“没有数据”、“数据从哪里来”。这次关于“数据资源流转”的调研有一个有意思的小变化，“买数据集”退到第四位，而且得票数和前面三项的差距很大。这可以解释为大家现在对数据流转和数据交易交换并不抱太大的期望；数据的流转并没有真正形成一个良好的环境和市场。因此大家转而倾向于自己搜集数据，下载并获得免费数据，而不寄希望于通过交易和价值交换带来有用的数据集，尤其对量大的数据集的期望值在降低。当然，还有人会免费提供数据集，能够排到第五位，也让大家嗅到了一点共享的

气息。这条预测所反应的问题其实还是比较严重的，数据只有流转起来才能产生巨大价值，解决这一窘境，对整个大数据领域来说极为重要。

大数据的最佳拍档概念

2016年	2017年
互联网+	智能计算或认知计算
云计算	云计算
智慧城市	机器人和人工智能
物联网 移动互联网	物联网 智慧城市
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	

到底是什么力量在推动大数据的技术、产业、应用的发展？2016年的调研中，科研院所从“大数据发展的最主要推动者”名单中消失了。大家认为大学和科研院所并没有直接带来资本力量和各种资源。也许随着国家提出的“双创”概念，创业公司成为大家希望的推动者；但今年创业公司也从这个名单中消失了。此单项调研结果与“大数据发展趋势十大预测”中的第八条“政府大数据发展迅速”相符。因为互联网应用是大数据的推动力量，所以真正新的推动力量就来自政府。期望政府在治理理念上的变化，在治理方法上的更新需求，带来对大数据应用的真实需求，大家对这方面寄予比较高的期望。



“大数据发展阶段判断”单项调研用参加投票专家的百分比分布来看大家对大数据的发展阶段的看法。从分布来看，大数据专家委的专家们对大数据的发展前景还是持乐观态度，第二、第三和第六阶段占大部分，超过了70%。而且有相当多的专家(26%)认为大数据已经过了幻灭期，进入稳定增长，即低风险高创造价值的时期。虽然对于大数据的泡沫破裂期是不是真的已经度过，大数据专家委的专家

们还有一定分歧，但总体来说看好大数据的发展是毋庸置疑的。■
(整理：潘柱廷 程学旗 周涛 靳小龙 杨婧)

一个热门话题，总是伴随着一些概念。从“大数据的最佳拍档概念”调研中可以看出，2017年延续了2016年的热门话题“人工智能”。

(整理：潘柱廷 程学旗 周涛 靳小龙 杨婧)