

大数据：未来的“新石油”

本报记者 刘松柏

热点追踪

就像空气和水一样，数据在我们周围自然而然地存在着。每一次点击鼠标，每一次刷卡消费，每一次拨打电话，每一次驾车出行……数据便已生成。我们每一个人既是数据的消费者，也是数据的生产者，每天都

生成海量数据。研究表明，过去5年里，人类行为产生的数据量增长了10倍，而在接下来10年中，这一增长将达到29倍。软件巨头EMC公司预计，到2020年，全球数据使用量将达到大约40ZB

(1ZB=10亿TB)。这些数据意味着什么，对我们有何价值？这就是当前最热门的“大数据”研究。中国工程院院士高文说：“不管你是否认同，大数据时代已经来临，并将深刻地改变着我们的工作和生活。”

AE模式的价值

□ 杨阳腾

近年来，“中国制造”冲出国门，我国制造业规模跃居全球首位。反观国内，“中国工程”却屡屡出现这样那样的质量问题。“中国工程”虽已初步实现了市场化发展模式，但仍存在着粗放管理、缺乏协同发展机制等诸多问题。如何堵住“中国工程”建设管理问题的缺口，核电AE公司(Architect Engineering Industries Inc.)的管理模式，值得我们认真借鉴。

核电AE公司管理模式的价值，逐渐为人所知。以中广核为例。中国广核集团通过设计建造一体化的核电AE公司模式承建的岭澳核电站二期工程，自2011年8月全面投产后，两台机组没有发生非计划自动停机停堆事件，在首个燃料循环周期内分别连续运行301天、286天，机组换料大修后持续保持满功率运行，2012年岭澳二期两台机组的世界核营运者协会(WANO)9项关键业绩指标，达到世界先进水平。此外，中广核还先后承担了辽宁红沿河、福建宁德、广东阳江、台山、广西防城港等多个核电工程建设任务，在建核电机组数量与装机容量位居世界第一。总之，10多年来，中广核在核电站建设方面保持领先，得益于其为客户提供专项技术服务的工程管理模式。

应该说，核电站建设不同于其他一般的工程，首先是安全要求与环境要求特别高，公众的敏感性特别强；需强化专门的核安全监管；需严密的质量保证与质量控制，独特的核安全文化等等。其次是高科学技术含量、庞大而复杂的系统和众多接口，建造实施的高难度等核电本身特性，对核电工程项目管理提出了不同于其他建设项目的更高要求。第三是资金需求大、建设周期长、风险因素多，使核电站的建设面临很多不确定因素。正因为如此，通过成立核电AE公司，加强核电项目从设计、采购、建设到运行期间技术支持的专项化管理，对促进核电工业健康快速发展十分重要。

核电工程公司即AE公司的项目管理模式，1981年起源于美国，其实是一种成熟而有效的国际工程管理模式。核电工程公司的服务范围包括核电建设的前期(初可、可研、概念设计、立项报告、招标评标等)、中期(初步设计、施工图设计、设计管理、设备采购及建造、施工管理、工程监理等)和后期运营、技术支持、技术改造等。从形式上看，AE模式是指设计建造一体化，更是设计、制造、建设一体化，有建造监督管理和设计构思的含义。从内容上看，AE模式就是总体工程管理的。国际上的AE模式有两个共同点：一是AE公司的集成作用，他们能集成全国乃至全世界的资源和技术能力，以建设好的核电站。二是AE公司的引领作用，他们在核电建设中起主导和“龙头”作用，能够引领行业内其他企业进行创新发展。

目前，我国有着大量的单一功能的设计院所，仅就核电站工程建设而言，大多数设计院所仅给出一份二维或三维的图纸，描绘着核电站建成后的形状、功能抑或是产品功能等。然而图纸却没有指出实现电站建设的路径或是今后做出有效的知识管理等现实和未来的重要课题。AE公司模式则是改变这种状况，要求设计院不仅描绘出二维或三维的图纸，还必须跟生产管理、施工运营结合起来，指出如何推进工程建设，同时还要统筹安排好工程安全、质量、进度、成本、技术、环境等六大方面内容。这成为AE公司模式的核心所在。除此之外，AE公司模式实施中还注重“大工程、大项目、大团队”的构建，强调质量管理应该覆盖工程建设全过程、贯穿核电工程全产业链。例如，为了从根源上提高工程质量，中广核致力于将碎片化的资源整合起来，并将自身掌握的技术与合作商进行分享，同时切实履行对上下游合作商的能力培育与监管责任，提升核电工程全产业链的质量。由此，清晰的质量保障体系也随着AE模式延伸至全产业链的每一家企业中去，并确保了核电安全质量的可控。

在实现市场化协同发展时，中广核AE模式布局强调全产业链上的各板块，均需在对合同负责的基础上，遵照“基于合同、超越合同”的理念协同创新，实现共赢。所谓“基于合同”即需各单位对合同负责、有着按合同办事的理念和心态；“超越合同”则是为了能更好地推动工程建设的一种权变安排。由于核电工程建设周期长的特点，最初签订的合同可能由于原材料成本上涨等因素导致最终无法实施，这时就需要以“超越合同”的理念在原合同无法执行的情况下，确保资金链的完整。

基于“基于合同、超越合同”的机制，中广核以创新管理，促进战略合作伙伴在各自项目上实现自我价值的同时，也实现了既定统一的目标。这正是在AE公司模式指导下，中广核能够使核电发展规划建立在有一个有效的组织体系之上，从而实现标准化生产并按照运行实践经验进行不断地改进创新，一方面减少核电项目的工程耗资，提高资源利用率；另一方面提高核电项目的工程效率，缩短建设周期，从而促进核电从设计到运营整体流程效率的提高。

AE公司模式实践，对我们解决“中国工程”建设管理中一些难题，具有一定的启发意义。特别是针对目前粗放的工程管理模式及缺乏协同、条块分割严重的行业格局，AE公司模式中的全产业链管理方式，不仅清晰地明确了各主体责任，而且有效地打破了各行业协会间的隔阂，实现了全产业链内外协同创新发展，有效推动了相关产业的技术进步，对我们广大工程企业进一步推进科学管理、实现协同创新，具有直接的参考价值。



被视为战略资源

一家电信运营商运用社交分析，筛选了3.65亿段电话记录，找出可能流失的客户并提供针对性的服务，大大提高了季度收益；一家金融服务公司从570亿笔ATM交易中检测出了欺诈模式……这是体现大数据价值的一些典型例子。

凯格总裁兼首席科学家杰米·霍华德说，与传统的、结构化的、关系型数据不同，“电话记录”等数据是非结构化的。据国际数据公司调查，非结构化数据在大数据中约占80%至90%，包括网络日志、语音、视频、图片、地理位置、传感信息等。作为新的尚未开发的信息源，非结构化数据分析可以揭露以前很难或无法确定的重要相互关系。

“这是最关键的。”大数据的出现让人们放弃了对因果关系的渴求，转而关注相关关系，人们只需知道“是什么”，而不用知道“为什么”，这颠覆了千百年来人类的思维模式，对人类的认知方式和交流方式提出了新挑战。

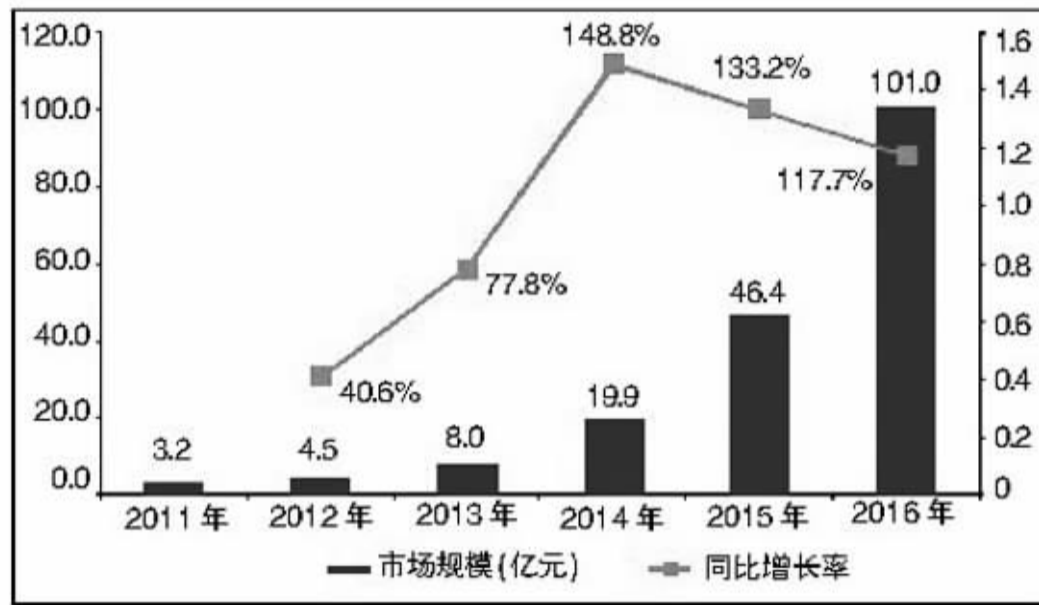
除了体量大、数据类型繁多这两个特征外，能否快速实时处理海量数据是衡量是否是大数据的关键维度。杰米·霍华德说：“大数据并非只有数百个TB才算得上，有时候数百个GB的数据也可称为大数据，这主要要看数据的处理速度。如果能在1秒之内分析处理通常情况下需要花费1个小时的数据，这种变化带来的价值是非常巨大的。”

目前，随着移动互联网、云计算、物联网等新一代信息技术的普及应用，智能手机、平板电脑、PC以及遍布各个角落的传感器，正在越来越多地接入到网络，成为数据来源和承载方式。各种交互数据、传感数据源源不断从各行各业迅速生成。这些数量庞大、种类广泛、迅速产生和更新的大数据，蕴含着前所未有的社会价值和商业价值。

杰米·霍华德说，像能源、原材料一样，大数据已成为提高未来竞争力的关键要素。比如，通过遍布各地的电子眼和监控摄像头，警务系统可以基于大数据识别一些危险的人或危险的行为，发现苗头进行预警。比如，通过车上的GPS和手机，当交通进入拥堵状态，交管部门可以通过手机信号的大数据运用，提示车主进入某个路段要收费，从而控制交通流量，等等。麦肯锡的最新调查研究显示，大数据可以在任何一个行业内创造更多价值。如，零售业可以利用大数据提高60%的运营利润，欧洲政府利用大数据每年可以减少1000亿欧元的开支，同时避免偷税漏税行为。

鉴于大数据潜在的巨大影响，很多国家或国际组织都将大数据视为战略资源，并将大数据提升为国家战略。美国政府把大数据看成是“未来的新石油”。2012年3月，美国奥巴马政府宣布了“大数据研发计划”，并设立了2亿美元的启动资金，希望增强海量数据采集、分析萃取能力，认为这事关美国的国家安全和未来竞争力。联合国推出的“全球脉动”项目，希望利用“大数据”预测某些地区的失业率或疾病爆发等现象，以提前指导援助项目。

2011年—2016年中国大数据应用市场规模与增长



(数据来源：赛迪顾问2013年2月)

主要是商业应用

尽管大数据应用在国防、反恐、安全以及社会管理等诸多领域，不过，企业应用还是大数据的主要推动者。IBM商业价值研究院的一项研究表明，在各个行业中，大数据业务主要集中在满足以客户为中心这一目标上。被调查企业认为，大数据有能力更好地了解 and 预测客户行为，并因此改善客户体验。

在互联网时代，在网上留下的每个“足迹”，聚集了大量的“前兆性”行为数据，如购买商品前先浏览、比价，观看电影前先搜索影评等，把这些数据搜集起来，通过进一步分析便可快速获取影响未来的信息。“这就意味着，借助大数据技术，公司可以比以往任何时候都更加了解消费者。”

阿里小贷就是运用大数据了解和预测客户行为的典范。

对于大多数小微企业或个体户而言，规模较小，风险大，若无抵押、无担保，从银行获得贷款的难度很大。不过，阿里小贷却与众不同，借款者凭借的是企业的“信

誉”这一资源，不用提交任何抵押与担保。

信誉从何而来？阿里金融高级专家年劲飞说，阿里巴巴、淘宝、天猫这些“阿里系”的电商平台，每天记录着各个商家、网店等小微企业的销售额、信用记录、顾客流量、评论、商品价格和存货等各类数据。显然，阿里对此了如指掌。“通过分析大量企业以及消费者行为数据，信用体系建立起来了，阿里小贷由此而生。”

像阿里小贷一样，在金融、电信、媒体、制造、医疗、地产等行业，大数据颠覆传统的故事轮番上演。沃尔玛基于数据分析的结果，将某些商品放置在一起；eBay通过数据分析技术，精确计算出广告中的每一个关键字为公司带来的回报以优化广告投放，2007年以来eBay产品销售的广告费降低了99%，而顶级卖家占总销售额的百分比却上升至32%。

除了作为大数据服务消费者的企业跃跃欲试，作为大数据服务提供者企业也

早已卷入其中。IT巨头如英特尔、IBM、甲骨文、微软等，早已对大数据的存储、处理和应用进行布局，在大数据的管理和分析上投入了巨额资金。2010年以来，欧美IT巨头掀起了一轮与大数据相关的并购潮。英特尔并购了安全软件制造商McAfee等大数据处理以及架构厂商来弥补自己在软件上的能力；IBM通过并购数据分析公司来增强自己在海量数据处理上的实力。这些IT巨头几乎抢占了搜索服务、数据库、服务器、存储设备、数据挖掘等核心价值环节。

国内的淘宝、百度、腾讯以及金蝶软件等也加入了大数据的生产、研究、应用行列。金蝶国际软件集团有限公司首席科学家、高级副总裁张良杰告诉记者：“我们参与搭建的全国中小企业信息平台，汇集了4000万家企业，通过对这些企业海量数据的挖掘和分析，能够对经济运行状况作出准确的预警，有助于国家相关部门作出应对决策。”

无论作为消费者还是供应商，企业开拓了大数据广阔的市场前景。据预计，从2013年到2018年，全球大数据市场将会出现年均26%的增长率，即从今年的148.7亿美元增长到2018年的463.4亿美元。

瓶颈有待突破

大数据在带来机遇的同时，也带来了不可忽视的一系列挑战。

“存不下”和“查不快”，是我们面临的两大难题。”中国工程院院士高文说。现在数据源源不断地产生，除非增加存储设备，但如此一来势必会增加存储成本。“目前，街头摄像头拍摄的视频数据，通常只会存一段时间，有的保存三个月，有的保存一个月，就被覆盖掉，为什么会这样？就是因为存不起，费用太高。”高文表示，提高视频的压缩效率是解决存储问题的途径之一，目前科研人员正在从事第三代编码技术的研究，以减少空间占用。

在数据处理上，由于大数据中大部分都是非结构化的数据，现有的软件和工具

主要适用于以结构化数据为主的传统数据，要想及时捕捉、存储、聚合和管理这些大数据，以及对数据进行深度分析和挖掘，需要新的技术和能力。而我国数据存储、处理技术基础薄弱，总体上以跟随为主，难以满足大数据大规模应用的需求。

此外，掌握非结构化数据或流数据的分析技能，对大多数企业来说是一个持续的挑战。在这些领域中，尽管硬件和软件已经成熟，但技能仍然缺乏。在美国，对高技能数据分析人才的需求，可能超出目前预测供应量50%至60%。到2018年，美国需要新增多达14万至19万名专家。“人才紧缺的状况如果不解决将严重影响大数据应用。”凯格总裁兼首席科学家杰米·霍华德说。

“大数据”对个人信息获取渠道拓宽的需求引发了另一个重要问题：隐私和便利性之间的冲突。一方面，大数据需要共享、数据开放、平台利用，而目前我国一些部门和机构拥有大量数据但受行政垄断和商业利益所限，数据开放程度也较低，有些部门把一些数据当成自己的，放那里既不使用，也不提供给研究者。专家指出，这是我国大数据研究的软肋和需要解决的大问题。

另一方面，消费者在享受大数据带来好处的同时，个人购买偏好、健康和财务情况的海量数据被收集，对隐私的担忧也在增大。高文表示，从纯技术角度来说，美国的“棱镜”项目是一个典型的通过分析海量通讯数据获取安全情报的大数据案例，但严重侵犯了个人隐私。

“这些瓶颈与挑战，是国内外大数据发展当中都面临的问题，只是我国在某些方面表现得更为突出。”高文说，“因此我们应该下定决心迎头赶上。”

发展大数据应用时不我待

□ 刘松柏



近年来，无论是国内还是国外，无论是学界还是商界，都在谈论大数据、畅想大数据。有人称，谁拥有了数据以及对数据的发掘能力，谁就将占领下一个10年全球经济发展的制高点。还有人称，在大数据领域的落后，意味着国家安全将在未来的国际竞争中落后于人。毋庸置疑的是，大数据正带来深刻的思维变革、商业变革和管理变革。

大数据固然重要，但我国大数据应用发展形势并不喜人，可谓内忧外患。一方面，目前我国大数据应用刚刚起步，基于大数据的商业模式还在萌芽阶段，从需求来看，很多产业对大数据的使用还没有意识，从供给来看，由于技术和人

才储备上的落后，也缺乏深厚的数据分析手段来支撑需求。另一方面国外IT巨头纷纷抢滩我国大数据市场。如，英特尔在上海成立团队，专门开发Hadoop大数据处理架构；EMC公司组建EMC中国研究院，在大数据、云基础构建、云平台与应用这3个方面推动大数据在中国的应用；IBM在北京成立了全球首个大数据智慧赋能中心，旨在构建一个完整的从商业评估、论证、方案部署到人员培训的大数据战略体系。

是观望，是退却，还是迎头赶上，事关国家的一些战略选择。国际上，不仅是美国，其他一些国家和地区都把大数据提升到国家战略层面予以支持。2013年法国发布了其《数字化路线图》，将大数据列为5

项将予大力支持的战略性高新技术之一；日本也发布了新ICT战略，提出“通过大数据和开放数据开创新市场”。这些国家认为未来国家层面的竞争力将部分体现为一国拥有数据的规模及运用数据的能力。

我国尽管已经意识到大数据的竞争力和战略性，却并未在真正意义上将其提升到应有的战略高度，数据财富尚未形成全民意识。因此，有必要从全局层面上制定大数据发展规划，明确大数据产业的发展重点、空间布局和保障措施，推动和改善与大数据相关的收集、储存和分析工具及技术，并在公共服务领域开展大数据应用示范，提高应急处置能力和安全防范能力。

事实上，我国大数据发展面临着难得

历史机遇。专家称，在这次云计算与大数据的新变革中，中国与世界的距离最小，在很多领域甚至还有着创新与领先的可能。一是大数据技术以开源为主，迄今为止，尚未形成绝对技术垄断。二是中国人口和经济规模，决定中国的数据资产规模将冠于全球，客观上为大数据技术的发展提供了广阔的发展空间。

抓住这一难得机遇，推动大数据发展，显得十分迫切。可喜的是，国内相对强势的互联网企业、电信运营商、电信设备供应商已经开始启动产业布局，以互联网应用服务为切入点抢占大数据制高点。我们期待在政府部门和企业共同努力下，我国大数据将迎来发展的春天。