

编者按

2018中国国际大数据产业博览会于5月26日—29日在贵阳举行,此次大会以“数化万物,智在融合”为主题,围绕大数据技术创新与最新成果,探寻大数据发展的时代变革,打造高端专业的大数据交流交易服务平台。博览会同期举办多场论坛、高端对话等系列活动。本报梳理此次会议部分精彩观点,以饷读者。

中国科学院院士梅宏:

分层次推进大数据治理体系建设

- 从政府层面,我们需要明确数据的资产地位,在法律法规的层面确立它的资产地位。
- 从行业层面,需要在国家相关法律框架的约束之下,考虑本行业中企业的共同利益和长效发展,构建相应的行业大数据的治理规则。
- 从组织层面,每一个组织都需要考虑通过组织内部的规章,把数据确定为它的核心资产。

随着大数据的价值和潜力为全社会所认知、关注,大批的企业、开源基金以及风险投资纷纷进入大数据领域。到2017年,随着领域应用不断深入,数据作为基础性战略资源的地位也日益凸显。数据质量、数据安全、隐私保护、流通管控、共享开放等问题日益受到高度关注,大数据治理成为产业生态系统的一个新的热点。

大数据治理仍存缺陷

近年来,各界学者纷纷开展了大数据治理研究和实践,已经看到不少成功的经验。比如,在国家及政府层面,围绕促进数据共享开放、保障数据安全和保护公民隐私等方面陆续出台了相关的政策;有很多针对企业组织、机构如何改善数据管理能力成熟度的模型陆续出现;技术上,面向数据质量保证的方法与技术,促进数据互操作的技术和规范标准,有很多研究成果,特别是近几年,数据质量、方法和工具受到了广泛关注。

然而,分析当前现状,我们也看到大数据治理还存在若干问题和不足。

第一,大数据治理概念比较狭义。大都以企业或组织为对象,仅仅是从一个企业或组织的角度来考虑大数据治理相关问题。然而,目前的现状是,多元数据的聚集和跨组织、跨领域数据深度融合挖掘,才是展现大数据价值的前提。在价值的驱动下,各界普遍存在着对突破数据的组织边界进行流动的需求。我们可以看到,随着数据开通流通技术以及相关渠道逐步建立和完善,已经出现数据跨组织流动的现象,而且呈现日益普遍的趋势。这就需要大数据治理突破企业或组织编写,应从行业内到跨行业,从区域内到跨区域,从全国乃至到全球多个层次进行考虑。

第二,对大数据治理内涵的理解还

没有形成共识。不同的研究者、不同的组织机构,从组织业务和管理流程设计、信息治理规则、数据管理应用等不同视角都在尝试给出治理的不同定义。有的人认为大数据治理就是IT治理的延伸,是它的一部分;有的人认为大数据治理需要独立于IT治理,是数据管理的延伸;有的着重于大数据相关制度优化、隐私保护和数据变现政策。

第三,大数据治理相关的研究实践多条线索并行,关联性、完整性和一致性不足。比如,在国家层面,政策法规和法律规定等较少被学者纳入大数据治理视角。数据作为一种资产的地位还没有通过法律法规予以确立,这就给它的流通带来很大影响。大数据的管理已经有不少可用的技术与产品,但是缺少多层次管理的体制和高效管理机制,怎样有效结合相关技术和标准规范,建立大数据共享开放环境仍然在探索之中。除了不断完善相关技术应对各种安全问题、新兴攻击和挑战之外,企业安全保障制度、行业自律监管机制和国家通过法律确定的强制手段方面还需要完善。

全面系统考虑大数据治理

我认为需要构建一个大数据的治理体系。大数据治理体系必须要跳出单个组织的边界,从营造国家大数据产业发展环境的视角来进行全面和系统化的考虑。

我理解有四个方面主要内容,一是数据资产地位的确立,二是数据的管理体制和机制,三是促进数据共享和开放,四是要保障数据安全。这四大方面内容涉及每一个组织层面,是一个多层次的事情,具体的手段可以通过制度的建立,法律法规、标准规范的制定,大量的应用实践,以及相应技术支撑体系的建设。

从政府层面,我们需要明确数据的资产地位,在法律法规的层面确立它的

中国科学院院士潘建伟:

量子科学催生现代信息技术

- 量子保密通信可以提供一种在原理上无条件安全的通信方式,所以它能够让网络和信息传输更安全。
- 量子计算机有比较好的性能,它可以用在大数据和人工智能等方面。
- 某种意义上讲,量子科学可以算信息时代的“核武器”。

上世纪初,随着量子力学的建立,催生出以信息技术为代表的第三次产业变革,X射线、能源科学、信息科学、生命科学、材料科学都跟量子科学紧密相关。从某种意义上,我们可以说,正是量子科学催生了现代信息技术发展。

随着技术发展,目前遇到两个问题。第一,信息安全。为了实现信息安全,人们用尽了各种方法,但历史经验告诉我们,依赖于计算复杂度的经典加密算法,随着人们计算能力的提高,它都会被破解。

第二,计算能力问题。随着大数据和人工智能的发展,人们对计算能力的需求日益增大。大数据要把里面的有效信息提取出来需要做大量的计算。发展计算能力,传统的手段就是把计算机集成度提高,并发展更好的软件。但是,目前摩尔定律已经面临着终结。尽管摩尔定律告诉我们,单位面积集成电路上可容纳的半导体晶体管数目每18个月就会增加1倍,但是不到10年,这个规律就会停止,会达到原子的尺寸,到原子的尺寸后,数据就不能被很好地定义什么是0、什么是1。

量子科技的发展催生了信息技术,它经过近100年的发展,也为解决上述问题做好了初步的准备。

量子保密通信可以提供一种在原理上无条件安全的通信方式,所以它能够让网络和信息传输更安全;利用量子计算我们可以得到一种超快的计算能力,实现更快的计算速度。利用单位量子,我们可以在原理上实现一种无条件分发,利用没有窃听的密钥进行安全通信。量子计算机有比较好的性能,它可以用在大数据和人工智能方面,比如求解 10^4 变量构成方程组的话,用目前最快的超级计算机大概需要100年左右的时间,但是利用万亿次量子计算机只需要0.01秒就完成了。某种意义上讲,量子科学可以算信息时代的“核武器”,所以它的发展非常重要。

我们国家很早就布局了量子科学研究计划。所以,在国际上,我们在量子通信方面走得比较靠前,2012年量子通信系统已经在北京投入使用。但量子通信存在一个问题,虽然其信号是安全的,但是不能被放大,远距离通信是个瓶颈,所以实现

100公里或者几百公里通信非常难。为了解决这个问题,我们开展了基于自用空间的量子通信技术研究。从2003年开始,经过十多年的努力,我们于2016年发射了世界首颗量子科学实验卫星,并开展了相关工作。去年已经取得了比较好的成果,比如点对点的量子密钥分发,去年我们完成的结果是每秒发送1000个密钥,之后是10万个,最近已经达到了50万个,也就是说在下一代量子通信卫星中我们很快可以实现每秒1兆点对点密钥分发的计算速度。利用单位量子,我们可以在原理上实现一种无条件分发,利用没有窃听的密钥进行安全通信。量子计算机有比较好的性能,它可以用在大数据和人工智能方面,比如求解 10^4 变量构成方程组的话,用目前最快的超级计算机大概需要100年左右的时间,但是利用万亿次量子计算机只需要0.01秒就完成了。某种意义上讲,量子科学可以算信息时代的“核武器”,所以它的发展非常重要。

在量子计算方面,我们国家也取得了比较好的成果。2012年,我们让量子计算的容错率由原来的 10^{-3} 提高到 10^{-2} ,很好地推动了量子计算的发展。去年,我们在国际上首次实现了光量子计算的原型机,它超越了早期的经典计算机。估计到今年年底,我们利用光量子计算和超导计算的计算能力,在某些特定计算功能上将达到目前最快的商用CPU处理速度。从这个角度讲,它可能在不久的将来能够比较好地用于优化网络、优化治疗、理解图像。我们希望通过10到15年的努力,一方面,能够构建全球化量子网络;另一方面,在量子计算方面,希望能够实现数百个量子比特相关操纵,这样对某些问题的求解大概就能达到目前全球计算能力总和的100万倍。

中国工程院院士、中国互联网协会理事长邬贺铨:

大数据及AI构筑数字安全新秩序

中国特色互联网所建立的数字新秩序——中国电子商务、互联网金融发展速度全球领先,但随之而来的是个人信息泄露风险和金融监管难度的增加。

移动支付中国全球领先,发展规模是美国的70倍。移动支付使用非常便捷,但便捷伴随的就是风险。如果不重视信息安全的保护,黑客可以从网络、邮件、电子商务、家庭地址搜集用户数据。

中国科学院院士张钹:

人工智能技术变成产业取决于四个因素

人工智能就是研究和设计智能系统,而这个智能系统能够感受周围的环境,能够思考,能够做决策,最后还能对周围世界产生作用。如何把人工智能的技术变成产业,创造价值,取决于四个因素。

首先是数据。现在大家认为到处都是大数据,实际上不然。网络上的数据虽然多,但跟你这个任务有关系的、合法可以拿的、能够保证的数据,不见得是大数据。另外,数据质量也是一个问题。有人统计,网络上的数据只有34%有用,7%经过标记,

1%经过分析。换句话说,只有1%的数据有用和好用。要把人工智能技术转化为产业,创造价值,就必须解决数据的质量和数量问题。其次是人工智能的算法。现在社会上用得最多的人工智能算法是深度学习,但是深度学习的应用场景又有诸多限制。再次是计算资源。目前对大多数人来说,计算资源还是足够多的,但是某些条件下,这个资源还很紧张。最后,应用场景是最重要的问题。如果没有选好应用场景,就不可能发挥人工智能算法的作用。

中国工程院院士沈昌祥:

只有可信计算才能解决区块链安全

区块链是一种利用密码学技术,将系统内的有效交易进行编码的可附加账本。区块链的安全与其他重要信息系统是相同的,因为它是一个典型的信息应用系统。我们说要实行网络安全等级保护制度,那么它的需求要做到交易有效,达成共识。

这几年区块链发展过程中发生了很多重大的安全事件。2010年8月就曾发生利用整数溢出漏洞凭空造出1840亿个比特币的事件。

我们应从以下几方面考虑,一是计算资源可信。区块链是在互联网基础上的应用,下面的计算资源网络不可信是不行的。二是交易数据要可控。无中心化是不错的,是理想状态,我们必须说人人平等,不要一个人说话算数,相当于仲裁的影响。我们应该做到比特币等区块链数据能够安全可靠地存储与传输。三是交易过程可靠。交易过程真实可信,不可伪造,可信共管。只有可信计算才能解决区块链的安全。

腾讯公司创始人、首席执行官马化腾:

数字化广度深度精度密不可分

我想从数字化进程的三个角度谈谈对数字生态的理解,分别是数字化的广度、深度和精度。

第一,数字化的广度。数字化的广度能够给我们带来很多的创新机会,随着网络规模和用户增长,在政务民生方面也有很多机会,尤其是民生服务的应用,有赖于政府的数字化和数字平台的广度。

第二,数字化的深度。每个垂直领域都非常深、非常专,数字化转型需要数字技术和行业经验深度融合,数字化创新需要下沉,进入到各行各业的五脏六腑,在这个过程中我们希望能够提供更加丰富有效的数

字工具,帮助各行各业打通七经八脉,让整个链条数据流通起来。

第三,数字化的精度。我们有理由期望数字时代能够实现工业时代难以达到的工艺水平。数字化的精度一方面有赖于我们在底层和前沿技术上的突破,另一方面也有赖于我们发扬“数字工匠精神”,做出高水平的数字工具和产品。

数字化的深度、广度和精度密不可分,广度让我们每个人都受益,深度可以助力创新,精度则可以提高生活品质。一句话总结,数字化将创新地提高每个人的生活品质。

百度创始人、董事长兼首席执行官李彦宏:

AI将为人类带来更多自由和可能

AI时代,除了技术和产品的快速落地外,人们越来越会去思考AI能给人类带来什么结果?而我们也越来越觉得需要有新的规则、新的价值观、新的伦理。所有的AI新产品、新技术都要有一个我们大家共同遵循的理念和规则。

第一,我们认为AI的最高原则是安全可控。一辆无人驾驶车我们一定要让它它是安全的、可控的。

第二,我们应该让所有的企业和人都能够平等地获取AI的技术和能力,而不是让它成为少数大公司的专利。

海尔集团副总裁陈录城:

工业互联网的核心是大规模定制

用户需求越来越个性化、高端化,传统的大规模制造模式已不适应时代的需求。对于我们这样的企业来说,必须要由大规模制造模式向大规模定制模式转型,也就是由以产品为中心转换到以用户为中心。

海尔的智能制造解决方案,就是要打造一套自主创新、具有中国自主知识产权的工业互联网平台,核心是大规模定制。这种模式,颠覆了原有工业体系在研发、制造、营销等方面的传统做法。

过去,产品研发是先调研后开发,开发出来的产品能不能卖出去,无人负责;现在是先有用户,由用户需求驱动产品研发,这是一个新的发展理念。

原来的营销体系是用户选择产品、为用户推销产品,这个产品卖不出去就变成库存,库存就会打折,降低企业效应;现在我们让用户参与制造、让用户创造价值。

今天,定制模式在海尔已经得到了验证。