

数据安全治理的几个基本问题

Several basic questions about data security governance



杜跃进 (1972-), 男, 博士, 阿里巴巴集团技术副总裁、首席安全专家、数据安全研究院负责人, 中国网络空间安全协会副理事长, 中国通信协会副主席, 中国保密协会个人隐私保护专家委员会副主任, 中国计算机学会计算机安全专业委员会常务委员, 中国互联网协会网络与信息安全工作委员会副主任委员。曾担任网络安全应急技术国家工程实验室主任、国家网络信息安全技术研究所所长、国家计算机网络应急技术处理协调中心副总工程师等职务。拥有近二十年互联网安全经验, 在我国最早开展国家级网络安全能力建设和大规模网络安全事件的应急响应, 推动了我国互联网应急体系建设与国际合作, 完成多项国家级科研项目。两次获得国家科技进步奖一等奖, 并获得新世纪百千万人才工程国家级人选、全国青年岗位能手、信息产业十大杰出青年等荣誉称号, 获得国务院特殊津贴。

中图分类号 :TP183

文献标识码 :A

doi: 10.11959/j.issn.2096-0271.2018064

* 本文为2018中国国际大数据大会演讲约稿

1 引言

数据安全如今在全球受到极大的关注，大家都在讨论这个话题。但对于什么是数据安全、该如何做数据安全等基本问题，大家还存在很多混乱甚至误解。本文围绕与当前数据安全相关的一些基本概念进行探讨。

2 为何强调“数据安全”，而非“大数据安全”

很多人在讨论大数据安全这个话题的时候，会纠结于“这是不是大数据”或者“这是不是大数据系统”。这会让人们忽略现在面临的真正问题：在数据无处不在、无处不用的情况下，如何保证数据安全。如果用大数据的定义限定大数据安全的范围，而忽略非大数据系统更加容易成为坏人得手的重灾区这一现实，那么就不会让数据安全的现状有任何改善。例如让全社会开始高度重视数据安全的“徐玉玉”案件，显然和大数据以及大数据系统都无关系。如果大家搞了半天所谓“大数据安全”，却不能降低下一个“徐玉玉”事件的发生概率，那么这样的“大数据安全”并没有什么意义。因此当前全社会需要关注的其实是“大数据时代下的数据安全”。当人们今天讨论数据安全问题的时候，不能用是否是“大数据”或“大数据系统”来限定范围。而且，大数据时代下数据的存在形式、使用方式、流转共享模式等都和原来极为不同，因此大数据时代下的数据安全是个新问题，而不是传统的数据安全概念。

3 “数据安全”不仅仅是防窃取

现在很多人说的数据安全问题仅仅是实际要解决的问题的一小部分。从用户的角度来看，当前数据安全至少包括以下三方面问题。

一是数据被窃取，即担心自己的数据被坏人偷走。实际上坏人可能从外面偷数据，也可能从内部偷数据。按照以往安全行业的基本共识，来自外界的安全威胁只占三分之一左右，三分之二的的安全威胁是从组织内部发起的。根据笔者团队的调查，在一些特定行业中，内鬼窃取数据的比例比外界窃取数据的比例更高。然而迄今依然有很多人把防范外部数据窃取作为数据防窃取的全部内容，而忽视对内部攻击的应对。仅仅强调用户侧应用软件的安全、数据通信过程加密、防火墙设个大门等，以为门外管好就行了，这是远远不够的。

二是数据被滥用，即用户对拥有数据的服务方不放心，担心他们会滥用自己的数据。用户的数据在服务提供方那里，他们的员工会不会滥用权限、随便访问用户数据？他们会不会因为好奇或者朋友要求查看某个用户的个人信息？他们会不会把用户的个人信息倒卖出去？与通过内部网络攻击窃取数据的行为不同，这里说的是服务方工作人员在授权范围内从事了不符合业务场景的数据访问。例如，用户要求客服帮忙，客服人员在这种场景下访问该用户的数据解决其问题，这是正常的业务场景。但是如果没有用户请求，客服人员擅自访问用户数据，就属于滥用行为。在大数据时代，数据和业务都在高频率地产生和变化，若对每一次数据访问行为都重新进行权限申请与审核，将直接扼杀业务，用户也无法接受这样的效

率。因此数据安全的工作需要包括对数据滥用行为的识别、报警甚至阻止,并且应建立制度,保障对实施滥用行为的员工进行严厉制裁。从目前曝光的数据“黑灰产”案件中可以看到,大量案件是通过买通内部人员滥用数据权限,从而进行数据倒卖的。而这些情况都是案发之后才被调查发现的,说明这方面能力比较欠缺。

三是数据被误用,即在大数据加工使用的过程中会不会侵犯用户利益,也就是很多人谈之色变的“用户画像”“精准营销”等。在各种描绘大数据给人类带来美好机会的故事中,实际上都离不开精准的个性化服务。而这些故事已经在金融、健康、医疗、教育等很多领域中开始切实发生,甚至在制造领域也在快马加鞭地发展,未来所有的领域都会走到这一天。这种精准的个性化服务的背后其实就是“用户画像”“精准营销”这些大数据加工技术。技术本身是中立的,关键在于技术被如何应用。“精准营销”是被用来“杀熟”还是被用来更好地满足用户的个性化需求?“用户画像”是用于支持个性化服务或者保护用户安全,还是用于满足其他不良动机?在大数据分析加工的过程中,有没有人能够从中窥探到某个特定人的个人信息或者隐私?这些是防误用的内容。当前的技术和管理总体上能够控制大数据加工过程中的误用,使人们在享受大数据带来的好处的同时,把危险“关在笼子里”。但是现实中,还有很多企业和组织在发展的过程中忽略了这个问题,让用户感到大数据是个“恐怖的恶魔”。数据防误用的问题,现在被关注得更少。

4 数据安全为何必须“以组织为单位”

目前有一个比较常见的误区是把手机

上的移动应用软件(App)安全等同于数据安全,例如通过评估手机上的移动应用软件安全来判断个人数据是否得到保护。手机App一直都是安全的重灾区之一,很多App确实存在不经用户许可秘密采集用户数据、植入广告、秘密产生费用或者自身安全做得不好导致其数据被窃取等情况。但是App就算不存在这些问题,也不等于用户的数据就是安全的。

如今,用户的数据并不都是沉淀在移动应用软件内部、存在用户终端上的。移动应用软件背后连接的是云端,而且可能连接不同业务。因此即便是在“知情同意”“最小够用”等原则下取得了用户的授权,用户数据还是会存在后端,并且在今天的社会化大协作的模式下共享使用。因此,大数据时代下数据的边界是在产品、设备、业务、人员还是系统?显然都不是。数据至少会在提供服务的组织内部的不同产品、业务、设备、系统、人员中流动,甚至为了完成用户的一个服务需求,数据还必须在产业链上的不同组织之间流动。例如,为了完成用户的一个购物需求,数据就必须在商家、平台、物流、独立软件供应商、金融等多个环节中流动。

因此,更合理的方式是以一个组织为单位来衡量数据安全的情况。这里的“组织”指的是拥有数据、提供服务的企业或者机构,其具有相对独立和完整的管理,也能够对业务和安全负责。数据在一个组织内的不同产品业务中形成流转闭环,组织是数据流动的最小边界。组织与组织之间通过可控的制度程序或者接口实现数据的跨组织流动、共享、交易等,这时候也可以以单个组织的数据安全能力为基础,进行责任的划分或者数据流动风险的控制。

如今,政府在大力推动多部门数据打通和共享利用,很多企业也在进行组织

变革，建立更强大的平台能力，其原因是数据只有打通和流动起来，才能更好地发挥价值。在这个过程中，单个产品或者技术平台的安全都不代表数据本身就是安全的。只有以组织为单位，才可以跳出频繁变化的产品、业务、人员等带来的困惑，寻找到支撑今天数据安全需求的方法。

5 为何传统安全方法不适用于大数据时代下的数据安全

大数据时代下的数据安全是一个全新的问题，无法简单地用原来的安全方法解决。这主要体现在以下两个层面。

一是不能用“以系统为中心的安全”思路解决问题。以系统为中心的安全是大家熟悉的安全方法，例如看某个软件、某个服务器或某个手机终端安全与否。这主要是看这些系统在各种人为干预下是否会出现与预期设计不符合的功能，从而导致运行状态失控。如今，数据要在不同的系统之间流动，若某个系统出了问题，可能影响到当时在这个系统中的数据（包括被窃取），但这些数据也可能在别的系统中出问题。数据本身并不存在运行状态，数据出问题的概念和系统出问题的概念也不同。这两者的关系，有点像医院里“心血管科”和“血液科”一样，前者解决的是血液循环系统本身的安全（运转正常），后者则是要保障血液自己的安全。两者显然有关系，但又有很大不同。单个系统的安全并不等价于数据的安全，系统被入侵也不等于数据一定会被偷走，每个系统都固若金汤也不等于数据就不会被滥用或误用。解决数据自身的安全问题，需要切换到“以数据为中心”的安全思路上来。

二是不能用传统的“数据安全”方法解决问题。数据安全是最古老的安全概念。从古代战场上就产生了数据安全的需求，并推动了相关技术的不断发展。对一个文件、一个数据库的记录的保护等都是数据安全的概念。但是，如今的数据安全的概念和方法已经和过去完全不一样了，数据的存在形式、使用方式和共享模式与过去有了极大的变化，数据的权属也不都是数据处理者的。数据可能以文件、记录、字段等方式在不同的环节中被快速打散、重组、流动，在这个过程中还会源源不断地产生新的数据。在一个业务里，数据可能涉及很多设备、服务器、产品、用户和不同部门的人的信息，然而真正需要回答的是数据在这么复杂的全过程中，从用户的角度来说安全不安全？显然，这和传统的“数据安全”概念有很大区别。

6 数据安全为什么需要强调治理而非管理

这里说的管理，指的是根据事务的规律制定一整套规则，然后自上而下地进行规则的执行落地，控制系统达到预期状态。而治理指的是找到若干关键抓手和基本逻辑，调动多方力量和资源形成某种协同或者生态，通过社会协同引导整个系统达到预期状态。大数据时代下的数据安全治理无法使用传统的管理模式达到目标，必须走协同治理的道路。

全社会所有的行业都在进行数字化转型，一切都将步入数据驱动，数据将成为所有领域的基本生产资料。因此数据安全问题也将涉及所有行业，并且涉及产品、业务、人员、共享机制等，并不是某个垂直领域的知识或者某个层面的单一方法就可以

解决的。如果按照过去管理某个垂直特定行业的方式，设立若干部门，自上而下进行管理，不但成本无法承担，效率也无法适应今天的实际情况。因此，政府、行业、企业、第三方机构、安全等需要发挥各自的优势形成有效的配合，才能建立适应当今数字时代的协同治理模式，共同提升全社会的数据安全水平。

7 数据安全治理的关键目标是让安全成为竞争力

人们不希望数据安全工作堵住所有行业的“血管”，阻碍大数据时代的创新和发展。如果简单化地看如何“消除数据安全问题”的话，那么不要有数据、不要有数据流动和应用是最安全的。数据安全治理的最终目标是实现数据安全和数字经济发展之间的平衡，甚至是二者的相互促进。人们通过法律法规、政策标准、技术产品、产业发展、测评培训、监督机制等进行综合数据安全治理模式设计的时候，如果能够瞄准以下两个关键目标，就可以推动数据安全治理最终目标的实现。

一是让数据安全成为组织的竞争力而不是成本，实现“能者多得”，即数据安全做得好，意味着有资格得到更多的业务机会。安全在过去一直是成本，因此几乎毫无例外，每一个创新业务、每一个创业公司、每一个创新产品一开始都不愿意做安全，因为大家首先考虑的是活下来的问题，没有精力和资源管活得好不好的问题，而且若对安全进行投入，可能在竞争中由于开发速度更慢、开发成本更高而失败。甚至很多大型项目系统的建设也不愿意做安全。于是等产品逐渐成熟、业务逐渐做大、工程项目投入运行之后，慢慢发现很多安全的坑已经难以填补了。

谁不做安全或者少做安全，谁就更可能赢得竞争，从安全的角度来说，这就是“劣币驱逐良币”。在数据安全领域通过建立科学的治理模式可以改变这个现象，要点是让一个组织能处理数据的类型和规模与其数据安全能力水平挂钩。例如健康医疗行业迫切需要利用大数据技术大幅提升技术和业务水平，造福百姓，但是这类数据又非常敏感，那么行业主管部门可以规定：处理哪些类型或多大数量的数据的组织必须证明其达到相应的数据安全能力级别要求。这样，当一个组织想要使用健康医疗数据开展研究或业务之前，就需要先具备足够的数据安全能力。于是数据安全能力越高的企业，意味着有权处理更多类型和数量的数据。这样企业才会积极而且认真地提升自己的数据安全能力，实现业务竞争力与安全的正向挂钩，这样才会带动整个数据安全产业发展。

二是让提升数据安全水平成为自发需求，而不是被动合规。2015年讲数据安全的人还不多，今天忽然全世界都在讲数据安全了。这主要是由很多相关法律出台带来的合规要求导致的。满足合规的要求是一种被动需求，即便做到合规要求，也不代表一个组织能够很好地应对不断变化的新风险。而真正从自身内在需求出发，不断提升数据安全能力的组织，才能够更好地应对不断变化的风险。如今，越来越多的企业开始认识到了数据安全的重要性，意识到这是数字经济时代保护企业自身利益以及建立用户对企业信任的重要工作，从而形成了内生需求。如果数据安全治理能让更多组织产生内生的、实实在在提升数据安全能力的动力，而不仅仅是被动满足安全检查时的合规要求，才会让大家真正直面数据安全威胁，降低数据安全风险。

8 数据安全治理的基本抓手是数据安全能力成熟度

用数据安全的“能力成熟度”而不是安全风险衡量一个组织的数据安全能力,能够更好地适应风险的变化情况。如果能力不够,即便今天做到了合规或解决了已知风险,若明天出现新规或者产品或威胁手段发生变化,还是会导致不合规或风险失控。因此,能力成熟度是更加内在的指标。通过科学的方法衡量一个组织的数据安全能力成熟度等级,用这个等级决定一个组织能够做什么、不能够做什么。当用户选择一个服务的时候,可以根据服务方数据安全能力的等级,判断把自己的数据给到对方的风险大小,在可以获得同样功能的情况下,用户会更愿意选择数据安全能力成熟度等级更高的服务方。在数据共享、交换、交易、流通的过程中,可以通过双方数据安全能力成熟度等级的情况分析数据风险的变化,发起方可以据此决定是否要继续与对方进行数据流动。政府建立多部门数据共享流通促进大数据利用的机制时,可以通过组织的数据安全能力成熟度级别决定允许数据流动的方向,从而实现总体数据安全风险可控。

能力成熟度的概念如今在很多领域中被使用,而数据安全能力成熟度模型(data security maturity model, DSMM)标准已经是国家标准化管理委员会的报批稿,并且也在ITU、ISO等国际标准化组织中通过或者立项。

9 结束语

大数据时代的数据安全是“旧瓶装

新酒”。看上去是一个老的概念,但实际上是完全不同的东西。因此不能用过去的经验解决今天的数据安全问题,需要创新。

今天数据安全的基本思想是,技术上以数据为中心,管理上以组织为单位,治理的基本抓手是能力成熟度。数据安全治理的关键是让数据安全能力和组织所能处理的数据类型和规模挂钩,让数据安全成为竞争力而不是成本,让数据安全成为内生需求而不是被动合规。

数据安全需要多方面工作的协同。仅仅靠法律法规解决不了数据安全问题,因为无论怎么样都会有数据进入服务方,法律能解决数据来源的合法性问题,但不能解决这些合法获得的数据是否会被安全保存或安全使用的问题,过去的信息安全行业管理经验不适用。今天的数据安全是全社会、全行业的事,过去把安全行业当作特种行业进行管理的思路是行不通的,很多数据安全前沿技术还有待突破。

数据安全需要政策、技术、学术的结合,需要多方参与,形成治理模式。另外,这个领域才刚起步,需要创新和持续改进。在今天的发展速度下,没有办法在给出一个完美的方案以后再执行,只能是小步快跑,不断迭代。DSMM中最高能力级别要求的也是持续改进的能力,即不断发现新的问题,不断调整自己的能力。

参考文献:

- [1] 杜跃进. 数据安全能力将成为大数据时代的重要竞争力[J]. 中国信息安全, 2017(5): 72-75. DU Y J. Data security capability will become an important competitiveness in the era of big data[J]. China Information

- Security, 2017(5): 72-75.
- [2] 杜跃进. 以数据为中心的安全[J]. 网络安全技术与应用, 2016(10): 10.
- DU Y J. Data centric security[J]. Network Security Technology & Application, 2016(10): 10.
- [3] 李克鹏, 梅婧婷, 郑斌, 等. 大数据安全能力成熟度模型标准研究[J]. 信息技术与标准化, 2016(7): 59-61.
- LI K P, MEI J T, ZHENG B, et al. Research on security capability maturity model standard of big data[J]. Information Technology & Standardization, 2016(7): 59-61.
- [4] 杜跃进, 郑斌. 大数据安全能力实践[J]. 大数据, 2017, 3(5): 30-37.
- DU Y J, ZHENG B. Security capability practice of big data[J]. Big Data Research, 2017, 3(5): 30-37. □

《大数据》2018年（第4卷）

总目次

◎ 聚焦

大数据驱动 5G 网络与服务优化 邬贺铨 6 [2018055]

◎ 专题：区块链与数据共享

区块链在数据流通中的应用 闫 树, 卿苏德, 魏 凯 1 [2018001]

基于区块链的供应链金融服务平台 马小峰, 杜明晓, 余文兵, 王 意 1 [2018002]

面向大数据的区块链在清算系统中的应用 蔡维德, 郁 莲, 袁 波, 邓佑权, 李 琪, 郭 斌 1 [2018003]

区块链与分享型数据库 钱卫宁, 金澈清, 邵奇峰, 周傲英 1 [2018004]

区块链交易数据隐私保护机制 祝烈煌, 董 慧, 沈 蒙 1 [2018005]

零知识证明应用到区块链中的技术挑战 李 康, 孙 毅, 张 珺, 李 军, 周继华, 李忠诚 1 [2018006]

◎ 专题：数据自治开放

数据自治开放模式 朱扬勇, 熊 贇, 廖志成, 叶雅珍 2 [2018013]

数据自治开放与治理模式创新 沈 逸, 姚 旭, 朱扬勇 2 [2018014]

面向数据自治开放的数据盒模型 熊 贇, 朱扬勇 2 [2018015]

数据自治开放的软件开发和运行环境 吴毅坚, 陈士壮, 葛佳丽, 赵文耘 2 [2018016]

数据自治开放模式下的隐私保护 王智慧, 周旭晨, 朱 云 2 [2018017]

数据自治开放的加密技术挑战 黄 霖, 黎 源, 汪星辰, 赵运磊 2 [2018018]

数据自治开放应用平台设计与实践 陈德华, 潘 乔, 王 梅, 乐嘉锦 2 [2018019]

◎专题: 生物医疗大数据

- 宏基因组大数据分析的质量控制流程规范·····郑广勇, 杨 桢, 曹瑞芳, 刘 婉, 李亦学, 张国庆 3 [2018025]
- 区域卫生数据用于临床疗效分析的可用性研究
·····叶 琪, 赵 亮, 阮 彤, 冯东雷, 高 炬, 刘 珉 3 [2018026]
- 生物效应大数据评估聚类算法的并行优化
·····彭绍亮, 杨顺云, 孙 哲, 程敏霞, 崔英博, 王晓伟, 李 非, 伯晓晨, 廖湘科 3 [2018027]
- 基于 ICD-10 诊断编码的慢性病并发症聚类算法·····王晓霞, 蒋伏松, 王 宇, 熊 贇 3 [2018028]
- 多中心临床大数据平台建设及深度应用·····朱立峰, 刘淑君, 陈德华, 乐嘉锦 3 [2018029]
- 基于环境气象因素影响的异常就诊量预测·····于广军, 熊 贇, 彭思佳, 阮 璐 3 [2018030]

◎专题: 大数据驱动的智能计算体系架构

- 基于 RDMA 高速网络的高性能分布式系统·····魏星达, 陈 榕, 陈海波 4 [2018036]
- 面向大数据的异构内存系统·····王孝远, 廖小飞, 刘海坤, 金 海 4 [2018037]
- 树状结构大数据类型的高效支持·····陈世敏 4 [2018038]
- 图数据流的模型、算法和系统·····李友焕, 邹 磊 4 [2018039]
- 深度学习编程框架·····王秉睿, 兰慧盈, 陈云霁 4 [2018040]

◎专题: 大数据创新实践

- 结合深度学习的工业大数据应用研究·····李 广, 杨 欣 5 [2018046]
- 基于公开数据的特殊人群在线活动特征挖掘·····刘楚楚, 吕 欣 5 [2018047]
- 基于社会化大数据的音乐文化研究进展·····耿 屿, 韩筱璞, 吕琳媛 5 [2018048]
- 利用网络数据预测企业失信行为

.....周 涛, 李艳丽, 李 倩, 陈端兵, 谢文波, 吴 桐, 曾 途 5 [2018049]

在线社交媒体信息冗余现象建模与实证研究..... 秦 烁, 吕 欣, 孟凡辉, 胡延庆 5 [2018050]

◎专题: 新工科背景下的大数据人才培养及课程体系设计

新工科背景下的大数据教学体系建设探析.....王元卓, 于建业 6 [2018056]

“数据科学”课程群与“数据科学导论”课程建设初探

.....覃雄派, 陈跃国, 李翠平, 柴云鹏, 徐 君, 文继荣, 杜小勇 6 [2018057]

大数据技术原理与应用课程建设经验分享.....林子雨 6 [2018058]

新工科背景下大数据专业导论课程的改革与探索.....张祖平 6 [2018059]

新工科背景下的计算机类专业人才培养探讨.....马 礼, 张永梅 6 [2018060]

◎专栏: 大数据领域国家工程实验室

大数据分析系统创新平台与生态建设.....王元卓, 靳小龙, 程学旗 1 [2018010]

大数据系统计算技术展望.....王海涛, 毛 睿, 明 仲 2 [2018022]

大数据驱动的安全协同生态建设.....鲍旭华, 曲晓东, 郑新华 3 [2018033]

大数据分析与应用技术创新平台.....张平文, 鄂维南, 袁晓如, 傅毅明 4 [2018043]

大数据系统软件创新平台与生态建设.....王建民, 王 晨, 刘英博, 刘 璘 5 [2018054]

综合交通大数据应用技术创新平台.....刘晓波, 蒋阳升, 唐优华 6 [2018063]

◎专栏: 2017年度大数据应用实践 Top10

基于关联分析与机器学习的配网台区重过载预测方法.....张国宾, 王晓蓉, 邓春宇 1 [2018011]

金融科技在互联网金融行业性风险防范领域的应用.....胡鹏飞 1 [2018012]

大数据平台在金融行业的典型应用.....罗圣美, 戚 晨, 王 敏, 叶郁文 2 [2018023]

垃圾短信大数据自动识别技术·····	林华生, 钱 岭, 周 莹, 杨 希, 李 婧	2	[2018024]
混合异构营销数据平台实践·····	卢亿雷	3	[2018034]
一种节能大数据平台方案·····	夏 刚, 王立忠, 刘亚恒	4	[2018044]

◎ 研究

面向大数据应用的混合内存架构特征分析·····	李 鑫, 陈 璇, 黄志球	3	[2018031]
工业大数据分析综述: 模型与算法·····	王宏志, 梁志宇, 李建中, 高 宏	5	[2018051]
基于密度的停留点识别方法·····	李毓瑞, 陈红梅, 王丽珍, 肖 清	5	[2018052]
基于主动学习和克里金插值的空气质量推测·····	常慧娟, 於志文, 於志勇, 安 琦, 郭 斌	6	[2018061]

◎ 应用

基于关联规则的气象服务智能推荐·····	赵文芳, 刘亚楠, 余东昌	2	[2018020]
基于 Gephi 的航运招聘信息可视化分析·····	王 扬, 田 野, 李铁山, 陈俊龙, 彭东成, 周义华	3	[2018032]
混凝土泵送机械大数据挖掘与应用·····	赵 鑫, 吴德志, 周志忠	4	[2018041]
冒烟指数: 大数据监测互联网金融风险·····	李崇纲, 许会泉	4	[2018042]
DeepEye: 一个基于深度学习的程序化交易识别与分类方法·····	徐广斌, 张 伟	5	[2018053]

◎ 论坛

新一代互联网安全管理体系研究框架: 阳光互联网倡议 ·····	唐 鑫, 吴晓松, 黄 伟, LEE Jae Kyu, 赵 玺, 王嘉寅	1	[2018007]
互联网环境下分布式事务处理系统现状与趋势·····	马鹏玮, 魏 凯, 姜春宇	2	[2018021]
从数据的属性看数据资产·····	朱扬勇, 叶雅珍	6	[2018062]

◎ 前沿

CCF大专委 2018年大数据发展趋势预测 周 涛, 卞超轶, 潘柱廷, 查 礼, 程学旗 1 [2018008]

浅论区块链的可运维性 白 硕 1 [2018009]

◎ 动态

人工智能赋能金融科技 王健宗, 黄章成, 肖 京 3 [2018035]

数据竞赛助力大数据产业生态系统健康发展 陈 娟 4 [2018045]

数据安全治理的几个基本问题 杜跃进 6 [2018064]

BIG DATA RESEARCH

Content

2018 (Vol.4)

Application of blockchain in data circulation	<i>YAN Shu, QING Sude, WEI Kai</i>	1	[2018001]
Blockchain based supply-chain finance service platform	<i>MA Xiaofeng, DU Mingxiao, YU Wenbing, WANG Yi</i>	1	[2018002]
Big data-oriented blockchain for clearing system	<i>TSAI Wei-Tek, YU Lian, YUAN Bo, DENG Youquan, LI Qi, GUO Bin</i>	1	[2018003]
Blockchain and sharing database	<i>QIAN Weining, JIN Cheqing, SHAO Qifeng, ZHOU Aoying</i>	1	[2018004]
Privacy protection mechanism for blockchain transaction data	<i>ZHU Liehuang, DONG Hui, SHEN Meng</i>	1	[2018005]
Technical challenges in applying zero-knowledge proof to blockchain	<i>LI Kang, SUN Yi, ZHANG Jun, LI Jun, ZHOU Jihua, LI Zhongcheng</i>	1	[2018006]
Framework of next-generation internet security management: bright internet initiative	<i>TANG Xin, WU Xiaosong, HUANG Wei, LEE Jae Kyu, ZHAO Xi, WANG Jiayin</i>	1	[2018007]
Developing tendency prediction of big data in 2018 from CCF TFBD	<i>ZHOU Tao, BIAN Chaoyi, PAN Zhuting, ZHA Li, CHENG Xueqi</i>	1	[2018008]
Brief comments on the operability of blockchain	<i>BAI Shuo</i>	1	[2018009]
Innovation platform and ecology construction of big data analysis system	<i>WANG Yuanzhuo, JIN Xiaolong, CHENG Xueqi</i>	1	[2018010]
Prediction method in distribution transformer heavy and overload supply areas with relevance analyze and machine learning.....	<i>ZHANG Guobin, WANG Xiaorong, DENG Chunyu</i>	1	[2018011]
Application of FinTech in internet financial industry risk prevention	<i>HU Pengfei</i>	1	[2018012]
Self-governing openness of data.....	<i>ZHU Yangyong, XIONG Yun, LIAO Zhicheng, YE Yazhen</i>	2	[2018013]
General design of self-governing openness of data and the exploring of the new mode of governance.....	<i>SHEN Yi, YAO Xu, ZHU Yangyong</i>	2	[2018014]
Data box: a novel data model for self-governing openness of data	<i>XIONG Yun, ZHU Yangyong</i>	2	[2018015]
Software development and runtime environment for self-governing openness of data	<i>WU Yijian, CHEN Shizhuang, GE Jiali, ZHAO Wenyun</i>	2	[2018016]

- Privacy preservation in self-governing openness of data
 *WANG Zhihui, ZHOU Xuchen, ZHU Yun* 2 [2018017]
- Challenge of encryption technology for self-governing openness of data
 *HUANG Lin, LI Yuan, WANG Xingchen, ZHAO Yunlei* 2 [2018018]
- Platform design and practice for self-governing openness of data application
 *CHEN Dehua, PAN Qiao, WANG Mei, LE Jiabin* 2 [2018019]
- Intelligent recommendation of meteorological service based on association rules
 *ZHAO Wenfang, LIU Yanan, YU Dongchang* 2 [2018020]
- State of the art and future research on distributed processing system in Internet environment
 *MA Pengwei, WEI Kai, JIANG Chunyu* 2 [2018021]
- Perspective on big data system computing technology
 *WANG Haitao, MAO Rui, MING Zhong* 2 [2018022]
- Typical application in financial industry of big data platform
 *LUO Shengmei, QI Chen, WANG Min, YE Yuwen* 2 [2018023]
- Automatic identification technology for spam message
 *LIN Huasheng, QIAN Ling, ZHOU Ying, YANG Xi, LI Jing* 2 [2018024]
- Quality control of big data analysis for metagenomics
 *ZHENG Guangyong, YANG Zhen, CAO Ruifang, LIU Wan, LI Yixue, ZHANG Guoqing* 3 [2018025]
- Usability research of regional health data for clinical efficacy analysis
 *YE Qi, ZHAO Liang, RUAN Tong, FENG Donglei, GAO Ju, LIU Min* 3 [2018026]
- Parallel optimization for clustering algorithm of large-scale biological effect evaluation
 *PENG Shaoliang, YANG Shunyun, SUN Zhe, CHENG Minxia,
 CUI Yingbo, WANG Xiaowei, LI Fei, BO Xiaochen, LIAO Xiangke* 3 [2018027]
- Chronic disease complications clustering based on ICD-10 diagnoses code
 *WANG Xiaoxia, JIANG Fusong, WANG Yu, XIONG Yun* 3 [2018028]
- Construction and deep application of multi-center clinical big data platform
 *ZHU Lifeng, LIU Shujun, CHEN Dehua, LE Jiabin* 3 [2018029]
- Abnormal detection of hospital admissions based on meteorological factors
 *YU Guangjun, XIONG Yun, PENG Sijia, RUAN Lu* 3 [2018030]
- Analysis on hybrid memory architecture for big data application
 *LI Xin, CHEN Xuan, HUANG Zhiqiu* 3 [2018031]
- Visualization analysis of shipping recruitment information based on Gephi
 *WANG Yang, TIAN Ye, LI Tieshan, CHEN C.L.Phipil, PENG Dongcheng, ZHOU Yihua* 3 [2018032]
- Big data driven security collaborative ecological construction
 *BAO Xuhua, QU Xiaodong, ZHENG Xinhua* 3 [2018033]
- Practices of hybrid heterogeneous marketing data platform
 *LU Yilei* 3 [2018034]
- Artificial intelligence energize Fintech *WANG Jianzong, HUANG Zhangcheng, XIAO Jing* 3 [2018035]
- Optimizing distributed systems with remote direct memory access
 *WEI Xingda, CHEN Rong, CHEN Haibo* 4 [2018036]

Big data oriented hybrid memory systems	<i>WANG Xiaoyuan, LIAO Xiaofei, LIU Haikun, JIN Hai</i>	4	[2018037]
Efficient support of tree-structured data types	<i>CHEN Shimin</i>	4	[2018038]
Graph stream: model, algorithm and system	<i>LI Youhuan, ZOU Lei</i>	4	[2018039]
Programming frameworks for deep learning algorithms	<i>WANG Bingrui, LAN Huiying, CHEN Yunji</i>	4	[2018040]
Big data mining and application of concrete pumping machinery	<i>ZHAO Xin, WU Dezhi, ZHOU Zhizhong</i>	4	[2018041]
Smoke index: big data technologies monitor Internet financial risks	<i>LI Chonggang, XU Huiquan</i>	4	[2018042]
Big data analysis and application technology innovation platform	<i>ZHANG Pingwen, E Weinan, YUAN Xiaoru, FU Yiming</i>	4	[2018043]
An energy saving big data platform scheme	<i>XIA Gang, WANG Lizhong, LIU Yaheng</i>	4	[2018044]
Data competition helps healthy development of big data industry ecosystem	<i>CHEN Juan</i>	4	[2018045]
An industrial big data application research using deep learning	<i>LI Guang, YANG Xin</i>	5	[2018046]
Analysis of online activity characteristics of hidden populations based on public data	<i>LIU Chuchu, LU Xin</i>	5	[2018047]
Progress of social-big-data-based studies of music culture	<i>GENG Yu, HAN Xiaopu, LÜ Linyuan</i>	5	[2018048]
Predicting the discredited behavior of enterprises via large-scale investment network	<i>ZHOU Tao, LI Yanli, LI Qian, CHEN Duanbing, XIE Wenbo, WU Tong, ZENG Tu</i>	5	[2018049]
Modeling and empirical research of information redundancy on online social media	<i>QIN Shuo, LU Xin, MENG Fanhui, HU Yanqing</i>	5	[2018050]
Survey on industrial big data analysis: models and algorithms	<i>WANG Hongzhi, LIANG Zhiyu, LI Jianzhong, GAO Hong</i>	5	[2018051]
Stay point identification based on density	<i>LI Yurui, CHEN Hongmei, WANG Lizhen, XIAO Qing</i>	5	[2018052]
DeepEye: a deep learning-based method of recognition and classification of program trading	<i>XU Guangbin, ZHANG Wei</i>	5	[2018053]
Big data system software eco-system and platform construction	<i>WANG Jianmin, WANG Chen, LIU Yingbo, LIU Lin</i>	5	[2018054]
Big data driven 5G network and service optimization	<i>WU Hequan</i>	6	[2018055]
Construction of big data teaching system under the background of emerging engineering education	<i>WANG Yuanzhuo, YU Jianye</i>	6	[2018056]
Toward construction of “data science” course group and “introduction to data science” course	<i>QIN Xiongpai, CHEN Yueguo, LI Cuiping, CHAI Yunpeng, XU Jun, WEN Jirong, DU Xiaoyong</i>	6	[2018057]

Course construction experience sharing of the principles and applications of big data technology	<i>LIN Ziyu</i>	6	[2018058]
Reform and exploration for introduction course of big data professional under the background of emerging engineering education	<i>ZHANG Zuping</i>	6	[2018059]
Discussion on training program of computer majors under the background of emerging engineering education	<i>MA Li, ZHANG Yongmei</i>	6	[2018060]
Air quality estimation based on active learning and Kriging interpolation	<i>CHANG Huijuan, YU Zhiwen, YU Zhiyong, AN Qi, GUO Bin</i>	6	[2018061]
Defining data assets based on the attributes of data	<i>ZHU Yangyong, YE Yazhen</i>	6	[2018062]
Innovation platform of integrated transportation big data application technology	<i>LIU Xiaobo, JIANG Yangsheng, TANG Youhua</i>	6	[2018063]
Several basic questions about data security governance	<i>DU Yuejin</i>	6	[2018064]