大数据·人工智能·区块链研究（十三）

区块链：面向新一代互联网的基础设施

金澈清[1]　张　召[1]　潘　斌[2]

（1. 华东师范大学 数据科学与工程学院；2. 哲学系，上海 200062）

摘　要：互联网的普及与发展极大地促进了经济发展和社会进步。但互联网基础设施也暴露出数据质量低、可信度低、信息孤岛等问题。新时代的社会发展对互联网基础设施提出了新要求，即为面向海量客户群体构建可信交往关系。鉴于区块链技术能够在分布式不可信的环境中达成共识，构建不可篡改的账本数据，促进智能合约交易自动化，因此基于区块链技术打造面向新一代互联网的基础设施意义重大。文本阐述了新一代互联网基础设施建设的背景，介绍了区块链技术以及将其作为互联网基础设施的举措，并指出区块链技术所面临的挑战。

关键词：区块链；互联网基础设施；智能合约；共识机制

中图分类号：TP311.13；TP393.09　　**文献标识码：**A　　**文章编号：**1005-9245（2020）05-0103-11

DOI:10.14100/j.cnki.65-1039/g4.20200313.001

一、引　言

科技革命推动社会变革，进而带来文明进步。18世纪中叶，蒸汽机的发明使人类首次大规模使用机器代替人工，生产效率显著提升，促成了第一次工业革命。及至19世纪，电的发现以及电器的大规模使用进一步提升了生产力，点燃了第二次工业革命，人类进入电气化时代。进入20世纪以来，以信息技术（包括互联网）、新能源技术、生物技术等为代表的一大批新技术在多个领域爆发，使人类社会的发展进入到崭新的阶段。

互联网源于1969年美国的阿帕网（ARPA Net），之后在全世界迅速普及。智能手机的普及使普通民众能更便捷地使用互联网，进而从根本上改变了人类的交往方式，提升了协同式工作的效率。前互联网时代，人们进行远距离沟通的成本昂贵，生产生活被局限在相对狭小的范围之内，无法进行大范围合作。尽管电报、电话、传真等可以传输少量关键信息，但是信息传输带宽较小、处理数据比较单调、能支持的应用比较有限，因而进行远距离相互协作工作的难度很大。互联网的出现为人们增加了一种认识与进入世界的方式，使人能以不在场的方式完成以往只有面对面才能实现的目标任务。而“互联网+”时代的到来使社会生产与人际交往发生了颠覆性变革。例如跨国公司可以将业务进行切分之后在不同国家分别研发，联合完成一项大型任务；科学家可以通过互联网进行合作研究，促进科学技术的发展；人们可以更加便捷地通过互联网来表达自己的观点，等等。

（一）当前互联网基础设施的特点

互联网基础设施是指支撑互联网服务的所有软硬件的统称。在硬件方面，它包括数据存储设备、网络连通设备、微处理芯片等；在软件方面，它包括通信协议、操作系统、数据库等。经过多年发展，我国在互联网基础设施方面取得了突破性进步。据2019年7月由中国互联网络信息中心

收稿日期：2020-02-14

基金项目：本文系国家自然科学基金联合基金项目重点支持项目“教育大数据的获取、管理与知识构建方法研究”（U1811264）、国家自然科学基金联合基金项目重点支持项目“政府治理大数据共享与融合技术研究”（U1911203）、国家自然科学基金面上项目“面向高吞吐处理的区块链数据管理”（61972152）的阶段性成果。

作者简介：金澈清，华东师范大学数据科学与工程学院教授、博士生导师；张召，华东师范大学数据科学与工程学院副教授；潘斌，华东师范大学哲学系教授。

发布的《第44次中国互联网络发展状况统计报告》显示，截至2019年6月，我国的IPv6地址数达50286块/32，跃居全球第一；移动通信基站总数达732万个，其中4G基站占总数的60.8%。在用户规模方面，我国的网民规模已经达到8.54亿，互联网普及率达61.2%，其中手机网民规模达8.47亿；我国的域名总数达4800万个。另外，2019年6月，我国已经正式发布5G商业牌照，标志着互联网基础设施即将进一步升级换代。但成就与问题并存，主要显现在如下几个方面。

第一，海量信息杂糅且真假难辨。虽然互联网极大地促进了信息传播，但是互联网信息的真实性一直为人所诟病。互联网是焦点问题与重大事件的热点地带，但也可能发生信息发酵变形。当代社会每有重大事件发生，互联网世界随时可能出现各种信息汇聚，这其中既有科学客观的解析，也不乏各种假象谎言，更有恶意中伤与诋毁之辞。例如当前围绕新冠肺炎病毒疫情，各种言论铺天盖地，但不是所有信息都客观真实，为此诸多APP与新闻门户网站开辟了鉴真与辟谣专栏。甚至还有少数信息发酵变形而误导公众，实际上不利于疫情的控制。2016年的“魏则西事件”也曾引起互联网热议，诸如此类的案例反映出的共同问题是互联网作为信息传播的载体，难以有效负载查验信息真实性的责任。

第二，互联网信息的价值度有待完善。高质量的数据能够有助于科学决策，而低质量的数据可能会误导决策方向。大数据的可用性包括一致性、精确性、完整性、时效性、实体同一性[①]。一致性是指信息不包括语义错误或互相矛盾的数据；精确性是指信息能够准确表达显示世界中的实体；完整性是指信息足够回答各种问题和支持各种运算；时效性是指信息能够与时俱进；实体同一性是指信息在各种数据源中的描述统一。互联网数据的来源既多且杂，不同数据源通常自主发布数据，因而降低了数据的可用性。例如，当高校教师更换工作单位时，可能会存在多个个人主页，有些主页的信息较为陈旧。同时，基于管理主体归属问题而形成的数据孤岛现象，极大地妨碍了互联网的协作工作。以医疗信息化为例，当医疗信息未共享时，患者在一家医院就诊时的记录和检验结果可能不被其他医院认可，还需要重复检验。近年来，我国致力于推进数据开放共享，破除数据孤岛现象。2016年5月，李克强总理在全国推进简政放权放管结合优化服务改革电视电话会议上指出：“目前我国信息数据资源80%以上掌握在各级政府部门手里，‘深藏闺中’是极大浪费。”破除数据孤岛涉及办事流程调整、数据隐私等领域。

第三，数字资源的复制易和确权难等特性制约了数字资源传播的广度和积极性。一方面，数字资源易于复制。无论是歌曲、电影还是文献，经过复制之后就很容易进行传播；但另一方面造成的困境是知识产权无法得到合法合理的保障，创作者亦难以获得应有的知识回报，进而影响了数字资源创造的积极性与投入度。虽然目前知识产权法强调保护电子资源，但其被剽窃时仍存在取证难、执行难等问题。

（二）宏观视野中新一代互联网基础设施的机遇

党的十九届四中全会聚焦“推进国家治理体系和治理能力现代化”，从传统的“管理”到现在的“治理”，一字之差体现了党的治国理政总体方略的重大变化。政府治理是国家治理的重要内容，更加强调多部门基于协作和分工网络所进行的协同式行政[②]。在互联网时代，协同式行政需要多个部门之间的信息紧密沟通，实时反映治理现状。在治理过程中，各个治理实体之间可能并不存在上下级隶属关系，只是基于特定治理任务而发生有机联系。

现有的互联网基础设施能够让不同的治理实体保持信息沟通顺畅，但仍不足以支持顺畅快速的协同式行政。首先，互联网基础设施需要能够使治理主体确信所采集到的数据的真实性和准确性。治理主体依据从各治理实体处收集到的数据进行施政，若所得到的数据真实、准确，施政效果便能凸显出来；而一旦数据不真实、不准确，那么治理举措的科学性就会饱受质疑。换言之，治理主体迫切需要能够高效验证数据真伪的机制，而人工核验方法的成本过高，不适宜作为面向大规模数据的常规性核验机制。其次，互联网基础设施还需要能够保证治

①李建中、刘显敏：《大数据的一个重要方面：数据可用性》，《计算机研究与发展》，2013年第6期。

②唐坚：《从“政府管理”到“政府治理”——论新时代如何持续推进政府治理体系与治理能力的现代化建设》，《决策支持》，2019年第6期。

理实体之间传输数据时不会发生数据隐私泄露问题。治理实体之间可能不存在上下级关系，而仅仅是针对特定任务在一起的协同式工作，这就决定了各治理主体具有一定的自主性，有保护与本主体相关的用户隐私的内生动力。隐私泄露造成的恶性社会事件不胜枚举，确保隐私得到保护非常重要。最后，互联网基础设施需要能够支持对治理过程的溯源评价。在复杂的政府治理场景之中，需要客观评价各项治理举措。特别是当出现负面效应时问责机制要及时启用。客观评价治理举措需要能够确保治理举措不被其他用户篡改，即已经发生的举措无法被修改或者撤销，从而能够确认各施政者的职责。

互联网的发展促进了数字经济的发展。数字经济是指以使用数字化的知识和信息作为关键生产要素、以现代信息网络作为重要载体、以信息通信技术的有效使用作为效率提升和经济结构优化的重要推动力的一系列经济活动[①]。数字经济是时代转型的重要标志，其内涵包括生产方式变革、生产关系再造、经济结构重组、生活方式巨变等。分享经济旨在通过互联网盘活闲置资源，提倡“使用而非拥有”的消费理念，从而提升资源利用和配置效率[②]。分享经济也是供给侧改革的新经济方案，可以增加生产资料的利用率，激发经济活力。

互联网经济的目的是充分利用信息技术对分享经济进行革命，传统的分享经济局限于熟人社会之中的利益共享，但互联网时代的分享经济并非面向小规模的熟人社会，而是超越熟人圈子进入陌生人社会甚至全球用户中。在整个过程之中，需要界定平台企业、资源提供者、消费者的权利、责任和义务，使互联网用户能够安全地共同分享物品。然而，以信息传播为主要目标的互联网基础设施难以有效实现以上目标，新的互联网基础设施能够确保分享的信息真实准确、物品能够健康流转、分享者与被分享者的身份隐私信息能够得到保障。

（三）以信任构建为要旨的新一代互联网的基础设施

现有面向信息传播的互联网设施显然不能符合新时代的发展要求，新一代互联网基础设施将在场与不在场、线上与线下、熟人与陌生人、地方群体与全球用户的有机联系作为一个动态的有机场景，其中信任机制至关重要。信任机制的建构路径各不相同，最常用的模式有两种。一种是集中式架构，即在该架构中存在一个中心节点和多个普通节点，普通节点之间不互信，但是所有普通节点都信任中心节点，因而普通节点可以通过中心节点传递信任。中心节点可以是政府、大型企业，也可以是有名望的个人。把信任寄托在中心节点上，会存在一定的风险。在极端情况之下，部分中心节点会通过背弃契约来获利，从而令整个信任体系崩溃。2019年发生的多起网贷平台跑路事件，致使许多将资金放在网贷平台牟利的用户损失惨重，这是中心节点信用丧失的典型案例。

另外一种信用体系结构是投票策略，假定参与平台的大多数人是诚实的，能够客观地表达观点。则经由多数人共同认定的事件被认为是真实的事件。在一些更细化的场景之中，以加权方式表达不同用户的意见重要程度。这种朴素的观点在政府治理和分享经济中广泛存在。当诚实用户占多数时，整个生态是良性的，且能够逐步成长；而一个由不诚实用户占多数的生态是无法维系的。

基于技术构建信任关系是一个被研究多年的问题。密码学领域的研究成果，特别是公开秘钥密码体系，使人们得以通过互联网来传递信息，并确保信息的完整性和真实性。在公开秘钥密码体系中，每个参与者持有一对秘钥（公钥、私钥），公钥向所有人公开，私钥则仅由该用户自己保存。基于该用户的私钥加密信息只能由相应的公钥进行解密，从而验证了信息的真实性和完整性。但是，单独密码学的成果还不足以作为互联网的基础设施。近期，综合密码学、分布式系统、网络、数据库等多种技术的区块链已经在金融、物流等多个领域中得到应用。区块链是一份在不可信环境之中由多个用户共同维护的分布式账本，账本的真实性由共识机制保障，而且账本不可被篡改，因而可以作为基础设施为互联网构建可信的生态环境。

二、区块链技术

2008年，一位化名为“中本聪”（Satoshi

①《二十国集团数字经济发展与合作倡议》，http：//www.cac.gov.cn/2016-09/29/c_1119648520.htm，2016年9月29日。

②吴晓隽、沈嘉斌：《分享经济内涵及其引申》，《改革》，2015年第12期。

Nakamoto）的学者在密码学邮件组发表了一篇关于“比特币”的论文；隔年，基于该论文的比特币正式面世[①]。近十年来，由于比特币在发展过程中不断暴露出一些问题，人们开始将目光投入到比特币背后的核心技术——区块链上，并意识到这项技术所拥有的巨大潜力。简而言之，区块链是一份去中心、去信任、不可篡改的分布式共享账本（Distributed Shared Ledger）。

根据应用场景和准入规则的不同，区块链还可以划分为公有链和许可链两种。公有链面向全网公开，完全去中心化，所有节点无需授权即可加入或者离开，但是交易速度较慢，例如比特币和以太坊（Ethereum）[②]；许可链是部分去中心化，整个区块链受许可机制管理，仅允许授权节点加入或者离开，未获得授权的节点无法加入到区块链系统之中，典型系统包括超级账本（Hyperledge Fabric）[③]。许可链又可细分为联盟链和私有链，私有链针对单个个人或者实体，而联盟链针对多个个人或者实体。许可链的交易速度要快于公有链（见表1）。

（一）数据结构与智能合约

在区块链系统中，数据以链式结构存储，相邻区块使用密码学技术进行链接。每个区块包括区块头和区块体两个部分，区块头包括了前一区块的根散列、Merkle树根散列、时间戳以及其他信息。顾名思义，前一区块的根散列是前一个区块的树根散列，Merkle树根散列是本区块的Merkle树的根散列，时间戳记录本区块的发生时间。鉴于哈希函数的碰撞概率极低，前一区块稍微改动一点，产生的哈希值就会完全不同，因而能够防止数据被篡改。

人们把区块链的发展历程分为三个阶段。在区块链1.0模式下，主要用于构建基于区块链的数字货币（如比特币），用途较为单一。在区块链2.0模式下，主要涉及金融领域的革新，例如债券发行、证券发放等，功能进一步丰富。到了区块链3.0模式，区块链被用到更广的领域，被作为“万

表1　区块链分类

	公有链	许可链	
		联盟链	私有链
准入机制	无	有	有
共识机制	POW/POS	投票/多方共识	投票/多方共识
中心化程度	去中心化	部分去中心化	中心化
交易速度	慢	快	快
代表性系统	比特币、以太坊	超级账本	超级账本

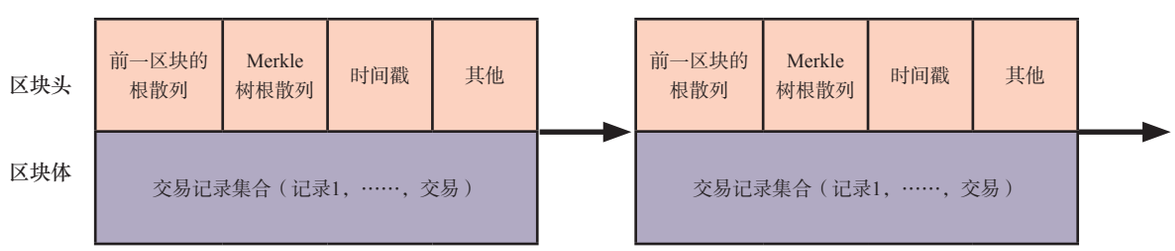


图1　区块链的链式结构

① Nakamoto S. Bitcoin：a peer-to-peer electronic cash system，https://bitcoin.org/bitcoin.pdf，2008.

② Ethereum，https：//ethereum.org/zh/.

③ Hyperledger Fabric，https：//www.hyperledger.org/projects/fabric.

物互联”的底层协议，例如医疗、政府治理、经济等领域[①]。智能合约是支持区块链从 1.0 模式跳跃到 2.0 和 3.0 模式的关键技术。智能合约是一段用程序语言编写的代码。当外部条件满足时，智能合约自动触发执行，无需人工干预。比如，可以编写一个支持物品分享的智能合约，当用户借用物品并且归还之后，能够根据用户的使用时间生成费用并且进行结算。智能合约无需经过第三方中介机构，可以在参与方之间直接执行。使用智能合约之后可以预先将规章制度、契约转化成代码，并向区块链的参与者公开，确保这些规章制度、契约能够透明地运行。

（二）共识机制

共识机制旨在使数据在分布式节点之间达成一致。区块链出现之前，一些学者研究如何在具有“故障—停止”模式的分布式网络系统之中达成共识。所谓的“故障—停止”模式，是指分布式系统之中不存在主动作恶的节点，在正常情况之下，节点按照预先约定的规则提供服务；当节点遇到软硬件故障时，停止提供服务，但不会提供虚假服务。典型的支持“故障—停止”模式的协议包括 Paxos 和 Raft 等。但是，这些协议并不适用于区块链场景，因为在不可信环境之中还存在主动作恶的节点，可能发送假消息、串谋、故意沉默等，这远比“故障—停止”模式复杂。如何在这种分布式场景之下达成全局共识，就是“拜占庭将军”（Byzantine Failures）问题。

为了应对“拜占庭将军”问题，目前已经研发出多种共识协议。针对公有链的共识协议包括工作量证明机制（POW）、权益证明机制（POS）等。POW 机制是比特币的共识机制，其本质是算力竞争，即各节点分别运用自己的算力解一道数学难题（该题目求解过程耗时较多，但是正确性验证的过程很快），最先解决该难题的节点可以赢得一次记账权利和一定数额的比特币奖励。为了确保区块链系统的安全，诚实节点的总算力需要占到 51% 以上。POW 机制的性能比较有限，平均每秒处理 7 个事务。由于 POW 机制不仅耗费大量能源，且效率低下，人们又提出 POS 机制，在该机制中，不同用户根据其拥有的权益换算权重，以此来达成共识。用户所拥有的权益越高，在共识阶段就享有更高权重。POS 机制降低了能耗，提升了共识效率。

但整体而言，POW 机制和 POS 机制能耗高、效率低，而在许可链场景下对效率和能耗要求更高，因而使用拜占庭协议（PBFT）和若干变种。PBFT 协议由 Castro 和 Liskov 提出，最多可以容忍 1/3 的不诚实节点。例如，假设区块链网络中共有 100 个节点，则可以容忍最多存在 33 个不诚实节点。PBFT 算法的另外一个优势在于不存在区块链分叉问题，区块一旦正式生成，会一直有效。由于 PBFT 协议需要节点之间频繁通讯，因此耗费较大的网络开销，其共识效率优于 POW 机制和 POS 机制，但仍有待进一步提升。

三、区块链作为新互联网基础设施发展的技术支撑

区块链是新互联网基础设施发展的技术支撑，构建超越传统模式的新互联网社会，前提是区块链技术的发展创新，具体而言，表现在以下四个方面的创新。

（一）区块链即服务（BaaS）

云计算可将计算、存储资源等整合起来向公众提供服务，极大节约了企业的开发与维护成本。当前典型的云计算基础架构理念是：基础设施即服务（IaaS）、平台即服务（PaaS）、软件即服务（SaaS）、数据即服务（DaaS），等等。在这些基础架构中，用户信任云服务提供商提供的设施、平台、软件和数据，主要适用于比较封闭的场景中。但由于云服务提供商的软硬件故障和管理等原因，数据篡改、泄露等事故屡见不鲜。在开放式场景之下，参与用户增多，人们不再依靠传统的单个大型厂商来提供服务，而是由多个企业自主联合起来进行服务。

区块链技术具有防篡改、去信任等优势，将区块链技术与云计算相融合，提供基于区块链的云服务（BaaS），能够支持开放式场景之下的可信服务，并减轻各个厂商的开发与部署成本[②]。BaaS 以区块链协议维护去中心化的平台，记录分布式账本，同时提供一些辅助功能，用以向用户提供云服

①邵奇峰、金澈清、张召等：《区块链技术：架构及进展》，《计算机学报》，2018 年第 5 期。

②朱昱锦、姚建国、管海兵：《区块链即服务：下一个云服务前沿》，《软件学报》，2020 年第 1 期。

务。BaaS 通常包括三个层次：基础设施层、中间层、服务层。基础设施层提供数据维护功能，包括数据的存储、查询和索引等。中间层是扩展协议层，包括网络协议、链上链下通讯协议等。服务层向用户提供服务，提供了相对统一的访问接口 API。

（二）驱动下一代人工智能

智能社会是人工智能发展的新阶段。与个体智能相比，群体智能更加侧重于社会性，要求各个主体之间实行有机协作，因此不同主体之间如何平等地共享信息、安全地传输信息以及可靠地保障隐私都成为人工智能发展要解决的重要问题。

图 2 显示了作为群体智能的基础设施，区块链技术总共分为 5 层，包括信息共享层、信息安全层、信任机制层、协作机制层和智能社会层（群体智能）。在信息共享层，首要解决的是个体之间信息共享和交换的问题，使用区块链技术能够使不同个体之间无障碍地、安全地共享与交换信息。在信息安全层，需要能够确保个体在信息交换和处理过程之中的数据安全及隐私保护。区块链使用密码学技术对数据进行加密和匿名化处理，有助于达成此目标。在信任机制层，需要使个体在缺乏第三方中介的前提之下仍然能够相互信任、互相配合。区块链中的智能合约技术能够让每个个体客观、公正地沿着预设的规章制度来实施。在协作机制层，需要每个个体之间相互协调、解决冲突、达成一致。区块链所提供的共识机制、投票机制有助于进行分布式协调。在顶端的智能社会层（群体智能）通过大规模合作与协同，将个体智能转化为群体智能，其中区块链的激励机制和监管机制能够成为有效的动力根源，区块链技术将为人工智能的未来发展提供全面支持。

图2　区块链作为群体智能的基础设施

（三）重构组织形式

传统的组织模式通常以中心化方式出现，即存在一个中心节点和多个普通节点（见图 3〈a〉）。普通节点完全信任中心节点，并在中心节点的指导之下开展工作。例如，一家大型企业包含总部和若干分公司，总部（中心节点）是上级单位，分公司（普通节点）是下级单位，总部指导分公司的工作，分公司完全信任总部。房产中介也是典型的中心化例子，房产中介公司就是中心节点，买方和卖方（普通节点）通过房产中介公司进行房产交易，将房产中介视为可信的一方。房产中介籍此收取费用。

第二种组织形式是去中心化，各个节点之间完全独立，相互协作共同完成任务（见图 3〈b〉）。在分享经济的时代，企业的组织结构逐渐扁平化、柔性化、开放化、互联网化、智能化、虚拟化[①]。扁平化和柔性化会减少管理层级，降低管理成本；开放化使得更多用户和组织加入到生态之中，使之更具活力；互联网化、智能化和虚拟化使得企业的组织形态变得不再重要。去中心化的组织结构并不需要中心化的中介机构，极大地降低了成本；相关的账本信息保存在区块链之中，使各参与方之间的交易通过智能合约执行。

第三种组织形式是弱中心化，在组织中存在许多普通节点和少量监管节点（见图 3〈c〉）。普通节点协同完成工作，监管节点不参与具体工作，但是扮演监管者的角色。这个监管者的权限与集中式组织架构的中心节点不同。中心节点起核心作用，会承受单点错误等，能够删除、修改数据。在弱中心化组织结构中，监管者只有有限的管理权限，无法在不经过其他节点同意的情况之下修改、删除数据，只是在必要时以第三者的身份进行仲裁。在社

①曲强、林益民：《区块链＋人工智能：下一个改变世界的经济新模式》，北京：人民邮电出版社，2019 年版。

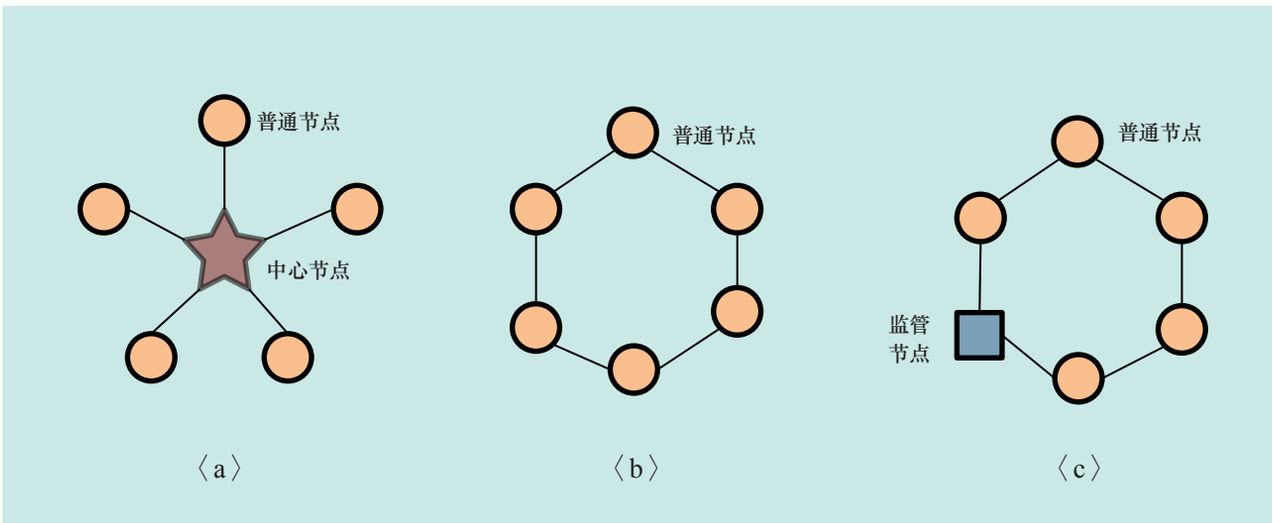


图3 区块链的三种组织结构

会治理过程中，社会组织作为主体管理具体事务，政府能够以监管者的角色促使社会治理顺利进行，在出现纠纷时政府能以仲裁者的身份调节纠纷，但这需要赋予政府监管权限（但非管理权限）。

传统的以信息传播为主要特征的互联网基础设施能够有效支持集中式组织架构，但是并不容易支持去中心化组织架构和弱中心化组织架构。在去中心化组织架构之中，不同节点之间协同式工作，同时各个节点也需要能够从机制上信任其他节点传过来的信息。在弱中心化组织架构中，监管节点能够从技术上针对其他节点的信息进行监管、溯源分析、确权，这也正是新互联网基础设施的特质所在。

（四）支持跨组织机构之间的协作

区块链解决单个组织内部的信任问题，但是有些任务需要在多个组织之间紧密协作，而不仅仅是由一个组织来承担。例如，社会上有很多公益基金会，接受公众捐赠，然后发放给受捐者。为了确保捐赠操作的透明性，各基金会可以基于区块链技术进行管理。在出现突发事件时（例如抗击新型冠状肺炎疫情），多家基金会同时介入。当各基金会自身以区块链管理捐赠事务时，受捐者可协调多个基金会的资源，以提升应对效率，这就涉及跨组织机构之间的高效协作。

跨链技术是实现区块链互联互通、提升可扩展性的重要手段。但缺乏跨链技术时，使用区块链技术将在单位之间、单位内部分别维护不同的区块链。特别是许可链对于节点的加入与否有严格的准入机制，导致各个区块链之间相互隔绝，无法达成数据流通[①]。

当前，主流的跨链技术可以划分为三类：公证人机制（Notary Schemes）、侧链 / 中继（Sidechains/Relays）、哈希锁定（Hash-locking）[②]。在公证人机制中，首先从当前组织中选择一个或多个作为公证人，公证人监听一条链中发生的事情，并在另一条链中复现；侧链 / 中继基于轻客户端验证技术验证在一条链上的交易或者事件是否发生；哈希锁定则在两条链上同时运行智能合约。

四、新一代互联网基础设施的多元应用

新一代互联网基础设施，可以在教育信息化、政治治理、互联网金融等诸多关乎国计民生的重要领域中进行创造性运用并推动行业变革。

（一）“互联网 +”教育

中国虽然是教育大国，但尚未步入教育强国行列。教育信息化是教育强国的必经之途，教育部于 2018 年 4 月发布《教育信息化 2.0 行动计划》，提出要积极推进“互联网 + 教育”，建设人人皆学、处处能学、时时可学的学习型社会。教育本质上是

①李芳、李卓然、赵赫：《区块链跨链技术进展研究》，《软件学报》，2019 年第 6 期。

② Buterin V. Chain interoperability，https：//www.r3.com/download/chain-interoperability.

一个复杂性的动态参与过程，良好的教育体制应该彰显教育的全员、全程、全方位特征。全员指教育的参与者面广，包括全日制学生、非全日制学生等，能基本覆盖全体公民；全程指教育的过程长，涵盖学前教育、基础教育、高等教育和终身教育；全方位指教育需要家庭、社会、学校三位一体全方位介入。

当前我国在互联网教育方面深耕厚植且成就突出。截至 2019 年 4 月，我国上线慕课数为 1.25 万门，学习人数多达 2 亿多人次，慕课的数量和应用规模位居世界第一[①]。但是，互联网教育还存在着诸多发展短板：教育信息技术的相对滞后、学习平台的有限性、一流师资与课程的稀缺性以及优质在线学习场景的匮乏。在线学习的动态性、开放性决定了学习者需要接入多重平台并灵活切换，智慧教育也有实时追踪学习者的大数据信息而及时提供个性化的学习方案与教育资源，借助于信息技术平台为全部学习者提供质量一流、技术便捷、课程丰富而应用广泛的学习路径，这显然是应该建立在新一代互联网基础设施充分发展的基础之上。

知识确权是推动智慧教育健康发展的制度保障。数字教育资源有着程度不一的交换价值，只有充分承认与有效保障该类资源的知识产权才能保护教育主体的合法权益、激励更多的资本投入以及完善尊师重教的社会习俗，进而推动知识经济的健康发展。区块链为数字教育的知识确权提供了一流且高效的技术支撑，它将所有的资源都记录在区块链上，并且按照时间顺序串接起来，可以根据在链上的时间顺序进行确权。同时，由于有智能合约的存在，交易就变得自动化和程序化。

（二）政府治理现代化

政府治理实质上是以政府作为治理主体与其他治理主体之间的协同式工作。治理过程是双向协作，需要从不同治理主体发回信息，然后作出科学决策，再反馈到治理对象，在此过程中如何从多元途径汇聚可信数据就显得尤为重要。但现代政治治理的复杂性使得要及时作出科学合理的决策就必须汇聚海量的、可信的优质数据，这显然不能为传统的人工方式所担负，而区块链则能为其提供一流的技术保障。

在区块链技术服务下，政治治理的所有施政行为都能完整地记录在区块链之中，而且基本不能修改，这既对施政行为提出了更高要求，也利于事中监管与事后追责，可以非常清楚地界定所有事情的责任和演化过程。例如，在民政扶贫领域，需要根据扶贫对象的经济情况确定款项如何发放，而扶贫对象的经济情况通常散布在不同的领域之中，例如税收、零售、房产等。如果利用区块链将不同的数据整合起来，就能够提高扶贫款发放的准确度和科学性。其次，区块链技术有助于廉政风险防控。为了逃避监管，违规违纪违法行为通常会对档案记录进行修改篡改，但在区块链技术支撑之下的记录是无法被修改的，这实际发挥了事前威慑、事中监管与事后问责的长效机制。

（三）分布式商业活动

在多方参与而又缺乏互信的商业环境之中开展分布式商业活动时，区块链技术有助于在原本信用匮乏或信用度较低的商业场景中建构可信的多边商业信任机制。以人们非常熟悉的电子商务交易系统为例，其将客户、商家、物流公司、银行四方处于一个分布式环境中。从客户方看，客户首先提交订单，然后通过网上银行向商家支付货款，最后从物流公司处收货；从商家方看，首先接受订单，然后通过物流公司发货，最后通过银行收款；从银行方看，首先执行从客户处收款，然后向商家付款；从物流公司方看，首先需要从商家处收货，然后收取物流款，最后向客户发货。在整个交易流程中每一个环节都需要记录相关的操作和信息。假设各方之间并不存在完全信任关系，最终以哪一方记录的数据为确认的信息是一个重要问题。目前的方法是这四方相信一个电子商务交易服务平台，将其作为公正的第三方进行统一记录各种账目信息，所有对于交易信息的查询操作全部在这个平台上进行处理，物流公司和银行的部分数据也以接入的方式添加至交易服务平台。传统的集中式记账方式与之相比存在显著差异，主要的交易信息存储在单一的记账方，这是一种集中式存储模式，即交易数据存储在唯一的某业务参与方并由其负责管理。如果两相比较

①《我国慕课数量世界第一 超 2 亿人次在线学习》，https：//js.qq.com/a/20190412/003205.htm。

则可发现，集中式记账方式已不足以应付高频且快速的现代商业活动，必然产生系列问题：记账方为了保证可靠性需要存储数据的多个副本，从而造成数据存储的性能瓶颈；交易数据可能被记账方篡改且无法验证，因此各参与方需要完全信任记账方；记账方受到攻击后数据难以恢复。因此，为了克服传统集中式记账方式的效率低、可信性差、易受攻击等弊端，引入区块链技术势在必行。

（四）应对突发公共卫生事件

突发公共卫生事件会对社会造成（或可能造成）巨大危害。据世界卫生组织（WHO）报告显示，截至2003年5月5日，全球SARS和疑似SARS患者累计为6583人，而中国占了一大半[①]。新型冠状病毒引发的肺炎影响的范围更大更广，截至2020年3月6日4时，31个省（自治区、直辖市）和新疆生产建设兵团累计报告确诊病例80651例，累计死亡病例3070例。在惨痛的损失面前，人们需要认真反思总结在应对此次突发公共卫生事件中的经验教训，从而对未来有所启示。作为一种新型技术，区块链在应对突发公共卫生事件方面也能发挥重要作用。

首先，区块链技术有助于慈善捐赠事项公开透明，增强公众信心。疫情暴发之后，全国各地踊跃捐款捐物，短时间内汇集到湖北省和武汉市。但是在物资接收和发放过程之中也出现了一些不透明现象，引发公众质疑。如果使用区块链技术记录所有接收和发放日志，就能够有效避免公众的质疑。

其次，区块链可用于物品追溯。在联防联治的紧急关头，医疗物资（特别是口罩）非常紧缺，有热心人士不远万里从国外抢购相关医疗物资支援国内抗疫一线。但是，仍有少数不法商家兜售假冒口罩等医疗物资，造成恶劣的社会影响。区块链技术可以用于物品追溯，记录每单商品的生产、流通过程，从而防止假冒商品的流通。

最后，区块链有助于治理过程透明化。疫情暴发后，公众的直接疑问在于是否存在瞒报缓报导致疫情大范围蔓延。不可否认，对于疫情的准确判断需要时间，但如果所有信息均建立在区块链上，有助于消解公众的疑问。

五、区块链技术的前景与挑战

区块链不是包治百病的灵丹妙药，它本质上是人类认识世界与改造世界的实践方式，是人类认识自我与社会的重要方式，在充分彰显理论价值与现实意义之际也面临诸多限制与挑战。

（一）有效保护用户隐私

所有账本数据均被记录在区块链上，如何保护用户隐私是一项重大挑战。当前存在两种保护隐私的解决路径：其一是从软件角度出发，结合区块链系统的不同层次开发隐私保护技术。例如，在网络层关注按需配置的安全防护机制，在交易层设计基于密码学算法的隐私保护机制，在应用层使用安全秘钥技术提高可靠性[②]；其二是采用硬件技术保护用户隐私。可信执行环境（Trusted Execution Environment，TEE）可用于构造一个安全的飞地环境，能确保在TEE中执行的程序安全可靠，不会被攻破。基于可信执行环境的区块链技术能够在保护隐私的前提下提高执行性能。

（二）为有效监管提供手段

区块链的最大价值是为分布式不可信环境构建可信的价值链。去中心化的组织结构并没有赋予任何节点特殊权限，所有节点一律平等，这种理想化的生态系统无法解决一些特殊需求，例如突发意外、录入错误等等。弱中心化的组织结构为监管节点提供一定的管理权限来维持整个生态的良性发展。监管节点并不是管理员的角色，无法在不受制约的情况之下删除、修改数据，但是可以在获得授权的前提下管理数据。例如，在大型突发公共卫生事件中，网络舆情密集，有真有假。监管节点的介入有助于提升区块链的质量。通常情况之下，区块链具有不可篡改的特性。为了提升监管能力，部分学者致力于研究可更改的区块链技术，使得在特定情况之下，区块链系统能够按照预先制定的规则修改相关历史数据。显然，可更改的区块链与“不可篡改性”并不一致，如果区块链上的数据可以被任意修改，使用区块链将会失去其意义。良好的生态不能滥用更改权限，必须在经过充分协商之后才可

①薛澜、张强：《SARS事件与中国危机管理体系建设》，《清华大学学报》，2003年第4期。

②祝烈煌、高峰、沈蒙等：《区块链隐私保护研究综述》，《计算机研究与发展》，2017年第10期。

以使用（例如，一半以上节点同意）。目前，一些学者使用变色龙哈希和可验证秘密分享构建可更改的区块链，但总体而言相关研究工作尚处于起步阶段[①]。

（三）进一步降低存储消耗

典型的区块链系统中包括全节点和轻节点两类节点，全节点维护一份完全的数据拷贝，而轻节点仅维护所有区块的区块头。由于区块头远远小于区块，因而整个区块链的存储开销取决于全节点的数量。假设单条区块链的规模是100GB，若一个区块链系统中有100个全节点和若干个轻节点，则整个区块链系统将至少耗费10TB的存储空间；而当全节点的数量增加到1000时，整个区块链系统将至少耗费100TB的存储空间。区块链系统的存储策略是一把“双刃剑”，一方面由于多个实体均保留数据备份，增加了数据可信性；另一方面，将产生严重的数据冗余现象，浪费存储资源。由于构建在不可信平台之上，区块链系统的数据冗余度远远高于传统的分布式数据库系统。在传统的分布式数据库系统中，通常只需要维护3个数据副本就足以应付硬件故障。如何同时达成数据可信和存储高效这两个目标就显得意义重大。纠删码是提升分布式存储效率的有效手段，可以在少部分节点失效的情况之下复原数据。基于纠删码的区块链存储策略既可确保数据可信，又能显著提升存储效率[②]。

（四）进一步提升执行性能

互联网应用通常是规模化应用，参与用户多，交易量大。特别是在特殊情况下，交易量呈现爆发式增长。例如，2019年双11期间，天猫平台的交易峰值是每秒钟54.4万笔[③]。然而，现有区块链系统的交易速度却远远低于这个量级。比特币每秒大约能处理7笔交易，以太坊大约每秒能处理10–20笔交易。联盟链的处理能力要大一些，但是当节点增加时，交易速度会明显降低。影响区块链执行性能的因素主要有两个：第一个制约因素是共识协议的效率。针对公链的POW和POS机制共识效率低下，典型的针对联盟链的共识机制（例如PBFT）受节点数量影响较大，当节点增加时，共识效率急剧下降。如何设计新型共识机制，提升共识效率，是一个公开的难题[④]；第二个制约因素是智能合约的执行效率。区块链系统不仅要求在所有节点上批量执行一组智能合约，而且要求在不同节点上的智能合约执行次序是等价的。一些学者使用并行化策略来提升执行性能，但由于不同合约之间还存在数据依赖等因素，如何进一步提升智能合约的执行效率还有待探索[⑤]。

（五）链上/链下数据的综合联动

区块链可用于构建“价值链”，但这并不意味着所有数据均需要上链。区块链的目的是使各参与方相信特定交易的真实性，这并不等同于各参与方的所有信息均需要在区块链上进行公示。陈纯院士认为链上/链下数据协同技术是联盟链的重要发展方向。2019年10月，陈纯院士在中国区块链技术大会上举例说，全国491个城市实行公积金联合管理之后，极大地便利了异地操作；但同时各个城市的公积金信息管理系统仍然按照传统方式运作，并未上链，因此链上数据和链下数据联动之后进行协同处理，能够分析、处理更多问题[⑥]。华东师范大学开发了师大链区块链数据库（SEBDB），将链上和链下数据采用同一个数据模型进行建模，并开发数据查询处理引擎，为链上/链下数据的联动处理提供了一种新的思路[⑦]。但总体而言这一问题尚处于起步阶段，有待理念的变革、技术的创新与市场的激励。

① 李佩丽、徐海霞、马添军等:《可更改区块链技术研究》,《密码学报》，2018年第5期。

② Xiaodong Qi、Zhao Zhang、Cheqing Jin et al. Automatic Calibration of Road Intersection Topology using Trajectories. ICDE 2020.

③《2684亿！2019年天猫双11收官 新消费引爆新增长》，http：//it.gmw.cn/2019-11/12/content_33311924.htm。

④袁勇、倪晓春、曾帅等:《区块链共识算法的发展现状与展望》,《自动化学报》，2018年第11期。

⑤贺海武、延安、陈泽华:《基于区块链的智能合约技术与应用综述》,《计算机研究与发展》，2018年第11期。

⑥ 陈纯:《链上、链下数据协同技术是联盟链发展重要方向》，https：//tech.ifeng.com/c/7rBjob6uhG4。

⑦ Yanchao Zhu、Zhao Zhang、Cheqing Jin et al. SEBDB：Semantics Empowered BlockChain DataBase. ICDE 2019：1820-1831.

Blockchain: Infrastructure of New-generation Internet

JIN Che-qing[1] ZHANG Zhao[1] PAN Bin[2]

（1.School of Data Science and Engineering；2. Department of Philosophy，East China Normal University，Shanghai 200062）

Abstract: The popularization and progress of the Internet have greatly promoted economic development and social progress. Although successful in promoting the dissemination of information，the current Internet infrastructure has some drawbacks，such as low data-quality，low credibility，and the creation of Information Island. The further development of economy and society demands new Internet infrastructure to build trusted relationships for large-scale users. It is critical to build an infrastructure for the next generation of the Internet based on blockchain technology which has the capability to reach consensus in a distributed，untrustworthy environment，to make ledger data non-tamperable，and to automate executing smart contracts. This article begins with the background of the next generation of Internet infrastructure，describes blockchain technology and its initiatives as Internet infrastructure，lists a number of application cases，and concludes with the challenges.

Key words: Blockchain；Internet Infrastructure；Smart Contract；Consensus Protocol

[责任编辑：马瑞雪]
[责任校对：王文秋]

本期特色栏目作者

AUTHORS FOR SPECIAL ISSUES

肖金明，山东大学二级教授、博士生导师；法治中国研究所所长，党内法治研究中心主任；享受国务院政府特殊津贴专家，国家社科基金重大项目首席专家。主要从事公法原理与制度、法治社会理论与实践、党内法治理论与实践等研究。出版著作 20 余部，发表论文 120 余篇。

（文章内容详见第 7–18 页）

胡鞍钢，清华大学国情研究院院长，清华大学公共管理学院教授、博士生导师；兼任国家“十三五”发展规划专家委员会委员等社会职务，是国内外享有盛誉的中国国情研究专家和学术带头人。出版《中国道路与中国梦想》《2020 中国：全面建成小康社会》等各类中英文专著、编著 60 余部，发表学术论文近 500 篇。

（文章内容详见第 54–63 页）

沈骑，同济大学外国语学院教授、博士生导师；同济大学语言规划与全球治理研究中心主任，上海市“曙光学者”，国际 SSCI 期刊 Language Policy 编委。主持并完成国家社科基金和省部级项目 8 项，出版个人专著 3 部，在国内外 SSCI、A&HCI 及 CSSCI 等高水平学术期刊发表论文百余篇，获得省部级科研与教学成果奖 2 项。

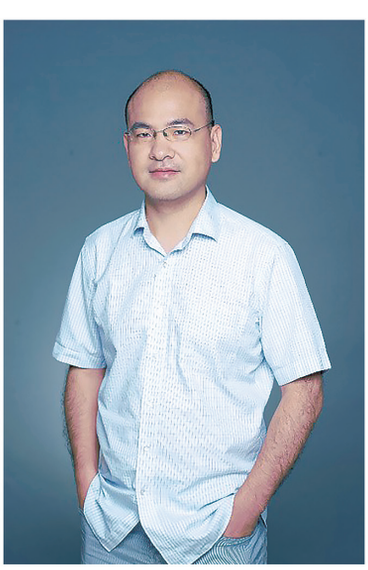
（文章内容详见第 64–74 页）

崔红志，中国社会科学院农村发展研究所农村组织与制度研究室主任，研究员、博士生导师；兼任中国社会科学院城乡发展一体化智库秘书长、全国社科农经协作网络大会理事会秘书长。出版学术著作 6 部，在核心期刊发表论文 70 余篇。

（文章内容详见第 75–85 页）

陆汉文，华中师范大学社会学院教授、博士生导师，中部地区减贫与发展研究院院长；曾担任联合国粮农组织、国际农业发展基金等国际机构减贫项目咨询专家。主持完成国家社会科学基金、国家自然科学基金、教育部、国务院扶贫办等研究咨询项目 80 余项，出版著作（含主编、合著）近 20 部，发表论文 70 余篇。

（文章内容详见第 86–94 页）



金澈清，华东师范大学数据科学与工程学院副院长，教授、博士生导师。主要从事区块链技术、数据科学与工程等研究。主持国家自然科学基金多项，出版英文专著 1 部、发表学术论文 100 余篇，曾获得省部级科研奖励和优秀论文奖励多次，并获霍英东教育基金会青年教师奖。

（文章内容详见第 103–113 页）