区块链与金融大数据整合实战

王静逸◎编著

资深专家呕心沥血之作，分享10年开发经验

金融与科技领域的9位重量级大咖点评并推荐

机械工业出版社
China Machine Press

区块链与金融大数据整合实战

王静逸◎编著

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

区块链与金融大数据整合实战/王静逸编著. —北京：机械工业出版社，2019.3

ISBN 978-7-111-62318-2

Ⅰ. 区… Ⅱ. 王… Ⅲ. ①电子商务－支付方式－研究 ②金融－数据处理 Ⅳ. ①F713.361.3 ②F830.41

中国版本图书馆CIP数据核字（2019）第053577号

本书从区块链的概念、原理、核心技术和应用等几个方面系统地介绍了区块链的相关知识，重点介绍了公链、联盟链和DApp开发等内容，并对区块链的海量存储和下一代公链扩容技术做了必要讲解。另外，本书对金融大数据的相关知识也做了详细介绍，并结合公链DApp和联盟链Fabric金融系统的开发，从实战角度介绍了区块链与金融大数据的整合应用，便于读者了解未来区块链技术在金融领域中的应用前景。

本书共10章，涵盖的主要内容有区块链的概念、原理与底层技术；大数据的概念与应用技术；密码学原理与加密算法；区块链的核心技术；搭建以太坊系统；公链DApp系统开发；Fabric超级账本与金融数据系统；多链和海量存储研究——金融大数据区块链架构；金融大数据的现状；区块链赋能金融大数据。

本书内容丰富，讲解通俗易懂，案例典型，实用性强，特别适合区块链技术爱好者和金融科技的相关从业人员阅读，也适合DApp开发者和区块链底层研究人员阅读。另外，本书还适合作为区块链培训机构的教材。

区块链与金融大数据整合实战

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号　邮政编码：100037）

责任编辑：欧振旭　李华君　　责任校对：姚志娟

印　　刷：中国电影出版社印刷厂　　版　　次：2019 年 4 月第 1 版第 1 次印刷

开　　本：186mm × 240mm　1/16　　印　　张：24.25

书　　号：ISBN 978-7-111-62318-2　　定　　价：99.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：（010）88379426　88361066　　投稿热线：（010）88379604

购书热线：（010）68326294　　读者信箱：hzit@hzbook.com

前言

大数据和人工智能的结合对金融的赋能让金融业可以尝试预判风险，根据用户的行为习惯建立数据模型，提前预测风险。区块链的公开可追溯特性，则让一切有迹可循，理论上可以让“坏人”“坏事”从根本上杜绝，任何“坏”的行为都无法发生，或者让潜在的金融犯罪需要付出高于收益的代价。区块链大数据时代，消除金融风险，从被动防护变为主动出击，全程可追溯，让个人和机构对自己的所有经济行为更加慎重，更加负责，从而让整个金融市场稳定、安全地发展。

目前图书市场上关于区块链的图书已经出版了多部，但却鲜见一本介绍区块链技术与金融大数据整合应用的图书。作为一个对区块链技术和金融大数据技术都有较为深入研究的研究人员，笔者觉得很有必要，也有义务编写一本介绍区块链与金融大数据整合应用的图书，这就是笔者编写本书的缘由。

本书系统地介绍了区块链的相关知识，包括公链、联盟链和 DApp 开发等，并对金融大数据相关知识也做了详细介绍，而且重点介绍了区块链与金融大数据的整合应用。相信通过阅读本书，读者可以大幅提升对区块链的认知能力，并系统掌握区块链与金融大数据整合的相关知识。

本书特色

1. 内容新颖、全面，知识体系完整

本书系统介绍了当前非常热门的区块链与金融大数据整合的相关知识，内容非常新颖，涵盖公链、联盟链、应用系统开发和金融大数据系统架构四大部分内容，可以帮助读者系统地掌握金融大数据系统开发所需要的知识。

2. 讲解由浅入深，循序渐进

本书按照“基础知识→底层技术原理→实战开发”的模式讲解，带领读者先掌握基础知识，再深入理解技术原理，最后进行区块链金融大数据系统的开发，学习梯度非常平滑。

3. 给出大量的原理图

俗话说，一图胜千言。本书在讲解区块链和大数据的底层原理时，绘制了多幅原理图

进行讲解，便于读者更加深入、清晰地理解区块链和大数据的底层架构，从而避免被晦涩的语言文字绕晕。

4．案例精讲，深入剖析

本书结合当前热门的区块链 DApp 应用与金融交易系统，通过案例带领读者实战区块链与金融系统的开发，帮助读者迅速提高开发水平。

5．提供完整的工程级性能源代码

本书源代码是基于作者开发的工程级系统简化而来，完成了从基础知识到工程应用的研发落地，可以帮助读者实际体验工程开发所需要的知识，最终实现从菜鸟到架构师的转变。

本书内容

第 1 章初识区块链，主要对区块链的历史、概念、分类和商业应用做了必要介绍，为后续学习打好基础。

第 2 章初识大数据，主要介绍了大数据的概念、发展和关键技术，并介绍了 Hadoop、分布式数据库和分布式计算等内容，最后对大数据的商业价值和发展趋势做了必要阐述。

第 3 章加密算法与区块链，主要介绍了区块链中的常用加密算法，涉及密码学、哈希算法、哈希链表、Merkle 树、公钥、私钥和椭圆加密算法等相关内容。

第 4 章区块链技术与特点，主要介绍了区块链的结构体系、去中心化、共识机制、POW 算法机制、POS 算法机制、DPOS 算法机制、拜占庭容错算法机制、数字货币的底层技术、智能合约、以太坊公链技术、超级账本联盟链、Token 经济与金融应用、区块链技术的缺陷与发展等内容。

第 5 章搭建本地以太坊环境，主要介绍了以太坊环境的搭建过程，帮助用户创建自己的私链。

第 6 章开发宠物 DApp 应用，从系统开发和金融大数据业务方面带领读者开发一个宠物应用系统，它是区块链的热门概念之一——DApp，通过宠物 Token 化，将宠物的交易和 ECR-20 代币结合起来。

第 7 章 Fabric 超级账本与金融数据系统，主要介绍了联盟链的代表 Fabric 超级账本的开发。银行等金融机构的业务都在大规模研究和使用超级账本，它是未来金融区块链大数据的重要利器。

第 8 章多链与海量存储——金融大数据区块链架构，主要介绍了下一代多链扩容方案及海量存储扩展方案。当前，区块链技术在与金融系统和大数据的整合中还存在许多的问题，比如它无法支撑高 TPS（Transaction Per Second）、高吞吐量和海量数据文件存储等，本章将对这些问题进行探讨。

第 9 章金融大数据的现状，介绍了金融业的产生和发展现状，以及现代金融业面临的挑战，并对金融大数据应用技术、核心需求、技术架构、行业应用及面临的挑战做了系统阐述。

第 10 章区块链赋能金融大数据，介绍了区块链应对金融商业环境的挑战、区块链应对金融大数据实施的挑战、区块链应对金融大数据应用的挑战及区块链应对金融大数据安全的挑战等内容。

本书读者对象

- 区块链技术爱好者；
- 金融大数据研究人员；
- 区块链底层开发人员；
- 区块链 DApp 应用开发人员；
- 区块链金融系统的开发人员；
- 区块链和金融大数据整合应用研究人员；
- 对数字货币感兴趣的人员；
- 计算机和金融专业的在校生和实习生。

配套资源及获取方式

本书涉及的源代码等配套资源需要读者自行下载。请登录华章公司的网站 www.hzbook.com，在该网站上搜索到本书，然后单击“资料下载”按钮即可在页面上找到“配书资源”下载链接。

售后服务

因作者水平和成书时间所限，本书可能还有疏漏和不当之处，敬请指正。读者可以通过 langkexiaoyi@gmail.com 和 hzbook2017@163.com 两个电子邮箱与作者或编辑取得联系。

目录

第 1 章　初识区块链

本书是以区块链为核心的技术书籍，其中最重要的就是区块链技术。在学习系统开发和业务应用之前，读者应该先了解区块链的概念和分类，本节就来介绍这部分知识。

1.1　什么是区块链

在工厂的概念中，区块链是由存储块单元通过指针连接，组成一个链状的账本结构。在传统分布式的基础上，区块链加入了共识机制，形成了可信的 P2P 网络。

1.2　区块链的“前世今生”

区块链发展到今天，已经进入了 2.0 时代。了解区块链的“前世今生”有利于熟悉它的演进规则，以帮助大家梳理学习流程，了解产业进程。

区块链的发展历程，可以总结为以下几个阶段。

- 密码学的网络支付：1982 年由 Chaum 提出理论，并且扩展成最初匿名现金的密码学系统。
- 比特币的奠基：2008 年，中本聪发表了一篇论文 *Bitcoin: A Peer-to-Peer Electronic Cash System*，奠定了比特币的理论基础。
- 比特币客户端：2009 年 1 月 3 日，中本聪发布了第一版开源比特币客户端，并且挖出了 50 个比特币，宣告区块链技术上线。
- 以太坊：2013 年 11 月，Vitalik Butern 发起了开源的 Etherum 项目，开启了区块链 2.0 项目，它的目标是一个全球的分布式计算机。
- 区块链的大航海时代：从 2011 年至今，以太坊、莱特币、Corda、超级账本和 EOS 等群雄并起，区块链的共识算法技术、存储技术及业务技术飞速发展，数字货币在社会和多个国家的政府层面也开始得到认可，区块链进入大发展时期。

区块链大发展时代如图 1-1 所示。

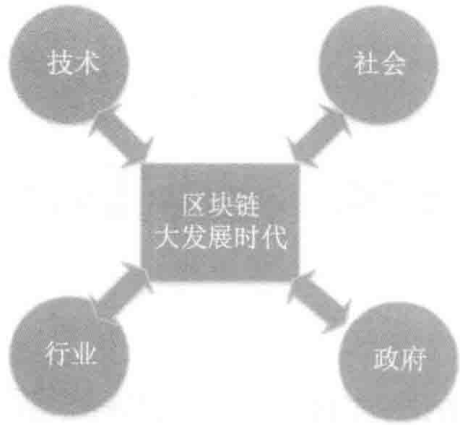

图 1-1　区块链大发展时代

区块链的大发展时代可以归纳为技术、行业、社会和政府 4 大方面的发展。

- 在技术发展层面：出现了 SHA-256 哈希算法，成为了区块链“挖矿”的主要算法；在共识算法方面出现了如 POW、POW、DPOS 和拜占庭容错算法等；在分布式和存储方面也有了大量的理论和实际研究成果。
- 在行业发展方面：出现了大量的公链、私链和联盟链。超过 500 家公司获得了超过 20 亿美元的投资，全球计算机行业已经兴起了区块链技术研究和应用的热潮。区块链在票务、保险、供应链、游戏、知识产权、溯源追踪、证券和货币等领域都有了成功应用，金融机构、银行和互联网企业也纷纷建立了自己的区块链项目。
- 政府方面：一些国家开始承认数字货币，并且允许数字货币进行交易和流通。我国政府也发起了区块链技术研究计划，并且投放了相应项目。工信部还指导发布了分布式账本技术的参考架构。
- 在社会方面：区块链引领了数字经济发展的热潮。有统计显示，数字货币的总市值已经超过 300 亿美元。一些国家的支付机构已经开始支持数字货币的支付，大大促进了数字经济的发展和应用。

区块链的发展已经成为了一股洪流，我们有理由相信，在不久的将来，区块链将会进入世界的各个角落，改变我们的社会和经济形态，重构互联网生态环境。

1.3　区块链的分类

区块链进入了大发展时代，大量的概念被提出，使其逐渐进入了细分领域。针对其分类，需要确定分类原则，具体说明如下：

- 以开放和权限划分；
- 以链的应用领域划分；
- 以程序独立划分；

- 以链的层级关系划分。

根据上面 5 类分类原则，下面进行详细的解说。

1. 以开放和权限划分

主要有三种类型：公链、联盟链和私链，具体说明如下。

- 公链：人人都可以加入，系统全面开放，所有节点的权限和等级平等，任何人都可以参与区块链，完全去中心化，不受机构控制，主要代表有 BTC 和 ETH。
- 联盟链：系统属于半开放，加入需要权限许可，使用仅限于联盟成员，一般应用在机构企业之间，主要代表有超级账本（Hyperledger）和 R3 联盟。
- 私链：仅限个人或企业内部使用，系统封闭，外部无法加入，不完全解决去中心化信任问题，但是可以改善企业内部的流程制度，主要代表有 Multichain。

2. 以链的应用领域划分

主要有两种类型：基链和行业链，具体说明如下。

- 基链：区块链使用底层通用的协议和 API，方便用户在链的基础上开发 DApp，一般来说，公链都是基链，主要代表有 ETH 和 EOS。
- 行业链：在底层的通用性上不如基链，一般是为某些行业定制协议和工具，可以解释为行业专用公链，主要代表有 BTM 和 SEER。

3. 以程序独立划分

- 主链：一般指的是正式上线以后的区块链，独立自主并且承担主要的业务，在自有的网络之中自成生态环境，主要代表有 BTC 和 ETH。
- 侧链：不特指某个区块链，一般是作为主链的补充，侧链是为了实现加密货币在链之间的互相转移，与主链使用同样的协议传输，主要代表有 Mixin Network。

4. 以链的层级关系划分

- 母链：能够不断生出新链，是某些区块链底层中的底层，主要代表有 NULS。
- 子链：基于母链的基础，再次构建的区块链，主要代表有 Press one。

随着技术的发展，区块链的种类越来越多，以后也可能会出现新的概念。读者熟悉链的划分，会有利于在未来的学习和工作中，更好地理解不同链存在的意义。

1.4　区块链的其他成员

区块链技术是由许多传统技术发展而来，传统技术和业务方案的发展，对于区块链的影响也是举足轻重的。这一节就来了解区块链的“兄弟”成员。

针对区块链的特性，分为以下几个重点：

- P2P 网络；
- CDN 分布式存储，共享带宽；
- 分布式云计算和边缘计算；
- 数字货币；
- ICO（首次币发行）；
- DAO（分布式自治组织）。

对于区块链的概念和发展，以上列出的几部分是极其重要的，下面对这些要点进行详细介绍。

1. P2P网络

P2P 网络称为对等网络。对等网络模型与服务型网络不一样，它的每个节点是逻辑平等的，没有特定的客户端与服务器。每个节点既对外提供服务，也在使用外部服务，而且每个节点的权限相同，任何单一节点丢失，都不会对整个网络稳定造成致命破坏。区块链是一个去中心化的分布式网络应用系统，它们通过 P2P 网络进行通信。

如图 1-2 所示是一个 P2P 网络。

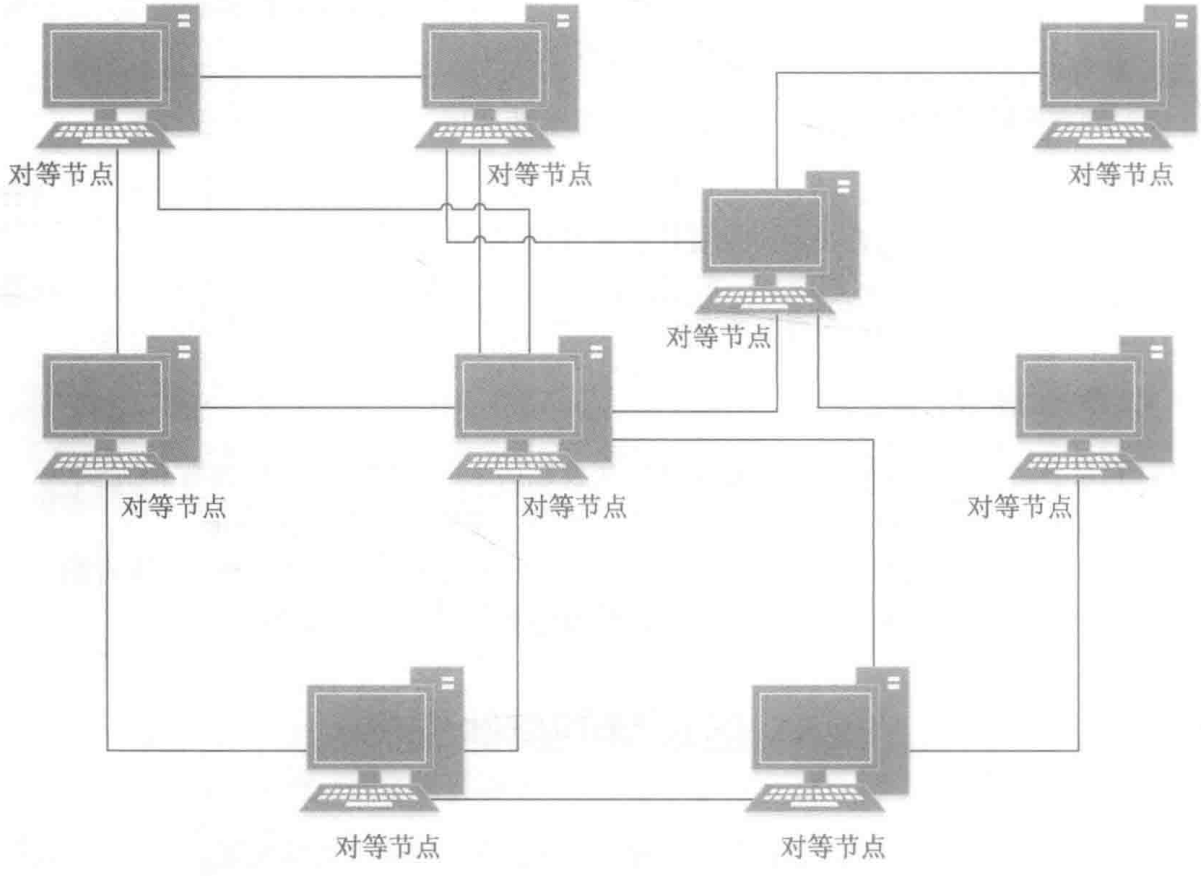

图 1-2　P2P 对等网络

2. CDN分布式存储，共享带宽

CDN 的技术系统主要由 3 部分组成，下面详细说明。

- 分发服务系统：由最基本的 Cache 缓存设备构成工作单元，由 Cache 响应用户的请求，快速提供缓存内容。Cache 也负责同步源节点，更新内容并且缓存到本地。Cache 的数量、规模和能力是 CDN 最基本的指标。
- 负载均衡系统：负责对用户的请求进行调度，提供用户最终的访问节点地址。它分为全局负载和本地负载。全局负载根据就近原则，提供最优服务节点。本地负载负责节点内部负载分配。
- 运营管理系统：为用户提供业务层面的管理分析、计费管理和数据统计分析等。

在区块链层面，根据负载和分配原则，也可以提供类似的下载和同步功能。

3. 分布式云计算和边缘计算

云计算是一种互联网集群的网络计算方法。通过这种计算方法，可以共享网络生态中的计算机终端资源。它是未来大型计算的重要技术。

边缘计算是一种云计算的优化方式，它通过网络边缘的设备和计算单元进行协作，允许附近的传感器和中央数据之间进行通信，比如笔记本电脑和智能手机等。

区块链作为一种开放性的网络系统，可以加入无数的设备，它是一种未来激励和促进边缘计算的有力手段。

4. 数字货币

区块链基于密码学技术，创造出了密码货币，通过区块链的激励和发行机制，让货币得以流通，是目前应用最广泛的领域。比如比特币、莱特币、元宝币和以太币等，都有大量的流通量和应用业务。

5. ICO（首次币发行）

如果读者关注数字货币市场，就会经常听到 ICO 这个词语。ICO 是区块链兴起的术语，英文为 Initial Coin Offering，即首次币发行，是一种加密数字货币和区块链项目的资金募集方式，对投资者使用数字货币进行抵押和回报。当它的流通量足够高时，区块链项目受到认可，数字货币就具有了市场价值，可以兑换法币。

同样，在资本市场也有 IPO（Initial Public Offering，首次公开募股）。下面来看一下 ICO 和 IPO 的共同点和不同点。

共同点：

- 二者都可以通过股份来募集资金。
- 投资者有可能通过潜在收益而参与其中。

不同点：

- ICO 的投资者大部分是不专业的投资者，是以炒币为主的投资者。
- 现阶段 ICO 监管牌照还未正式下发。
- ICO 是第三方平台，投资者自己承担风险，IPO 则是国资背景交易所来承接。

6. DAO（分布式自治组织）

DAO 是区块链改变世界的核心技术之一。DAO 的英文是 Distributed Autonomous Organization，代表了区块链时代的一种组织结构形式，通过区块链的共识和激励形式，可以在无人干预的情况下运转社区组织。它是以开源形式出现，每个人都可以贡献开发、购买权益股份、获得激励和共同推广。这是一种改变未来互联网的组织形式。

DAO 的表现形式如图 1-3 所示。

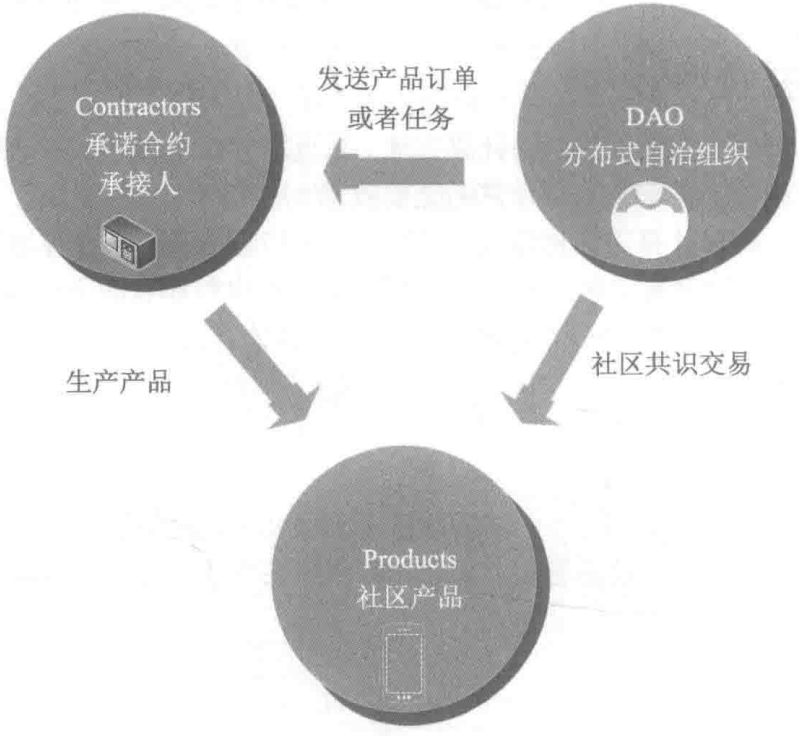

图 1-3　DAO 社区组织形式

1.5　区块链的商业价值

区块链是未来数字经济变革和重构互联网的重要技术，本节就来了解它的商业应用场景，以及未来的发展趋势。

区块链的商业应用场景主要集中在以下几个方面：

（1）通过区块链的不可篡改和溯源特点，可以应用在以下几部分。

- 信息防伪：链上记录产品的地址和重要信息，可以精准追溯，并且不可篡改。
- 食品安全：根据区块链，记录和追溯食品供应链的每个环节。
- 信息安全：防止信息篡改和 DDOS 攻击等。
- 身份验证：通过区块链来保存、鉴别和验证身份权限。

（2）通过区块链的去中心化和 Token，变革金融数字经济。

- 数字货币：帮助发行映射现实资产的货币，不可增加，也不可减少。
- 跨境支付：减少中间交易环节，减少手续费用，实现点对点直接交易。
- 通证和供应链金融：提供资产凭证 Token 和交易合约，减少供应链金融融资和供应风险。
- 股票发行和交易：通过区块链技术，减少股票中间渠道和人为错误，提供更公平、更可信、更透明的交易平台。
- 众筹：应用最为广泛的就是 ICO 筹集项目资金。

（3）在组织治理和管理方面，区块链也给出了新的解决方案，它通过数字资产、数字交易、智能合约、共识激励和仲裁服务等，一体化给出了未来智能自治组织方案，变革结构，极大节省了成本，提高了工作效率，自动产生社会经济价值。

区块链的应用生态，如图 1-4 所示。

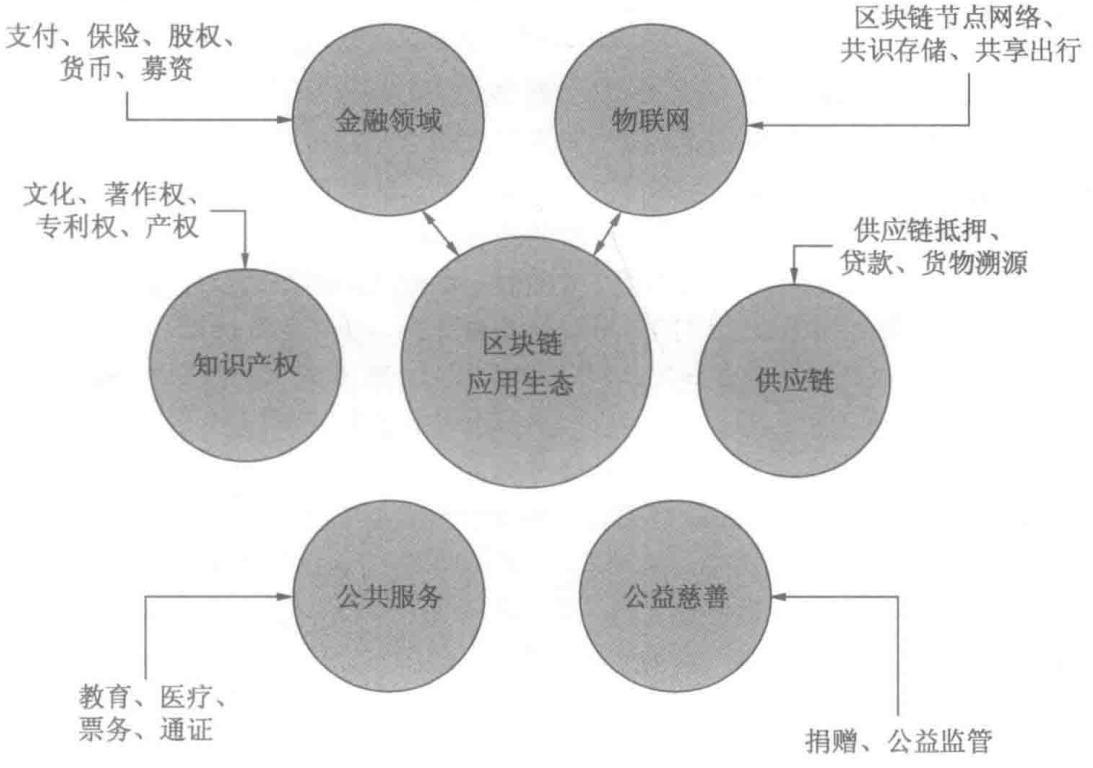

图 1-4　区块链应用生态

1.6 本章总结与思考

本章介绍了区块链的重要概念，帮助大家初步了解了区块链知识。结合本章所介绍的内容，思考以下问题：

1．区块链是什么？它的命名和重要组成是什么？

2．区块链的类别可分成哪几类？分别应用在哪些方向？

3．区块链在产业和技术上有哪些相近的地方？区块链重要的活动内容有哪些？

4．区块链在商业上有哪些应用？哪些是区块链未来的重要发展趋势？

区块链是未来商业和互联网的革命性技术，了解区块链的整体概念，有利于把握清晰的脉络方向，定位关键目标。