



2.6 大数据与区块链

技术发展日新月异，行业创新层出不穷。继大数据、云计算、物联网、人工智能等新兴技术之后，区块链技术在全球范围内又掀起了新一轮的研究与应用热潮。区块链的出现，实现了从传递信息的“信息互联网”向传递价值的“价值互联网”的进化，提供了一种新的信用创造机制。区块链开创了一种在不可信的竞争环境中低成本建立信任的新型计算范式和协作模式，凭借其独有的信任建立机制，实现了穿透式监管和信任逐级传递。区块链源于加密数字货币，目前正在向垂直领域延伸，蕴含着巨大的变革潜力，有望成为数字经济信息基础设施的重要组成部分，正在改变诸多行业的发展图景。

本节内容从区块链的起源“比特币”说起，然后，介绍区块链的原理、定义和应用，最后，介绍大数据与区块链的关系。

2.6.1 从比特币说起

区块链的诞生与受到广泛关注，与比特币的风靡密切相关。事实上，区块链技术仅仅是比特币的底层技术，是在比特币运行很久之后，才把它从比特币中抽象地提炼出来。从某种角度来看，可以把比特币认为是区块链最早的应用。

关于区块链的一切故事，要追溯到比特币的诞生。2008年10月31日，至今匿名的神秘技术极客、比特币的创造者——中本聪，向一个密码学邮件列表的所有成员发送了一封电子邮件，标题为“比特币：点对点电子现金论文”。在邮件中，他附上了比特币白皮书的链接，论文题为“比特币：一个点对点电子现金系统”。中本聪在2008年发表的这篇论文可能是互联网发展史上最重要的论文之一，其他重要论文包括：利克里德写的开启互联网前身“阿帕网”的“计算机作为一种通信设备”（1968年）、蒂姆·伯纳斯·李写的万维网协议（WWW）建议书“信息管理：一个建议”（1989年）、谷歌联合创始人谢尔盖·布林与拉里·佩奇写的搜索引擎论文（1998年）等。2009年1月3日，在位于芬兰赫尔辛基的服务器上，中本聪生成了第一个比特币区块，即所谓的比特币“创世区块”，由此，比特币正式诞生。

比特币与传统的货币是完全不同的。传统的货币的发行权掌握在国家手中，存在着货币滥发的风险。比特币这个电子现金系统是同时去中介化和去中心化的。所谓的“去中介化”是指，个人与个人之间的电子现金无须可信第三方中介的介入。所谓的“去中心化”是指，这个电子现金的货币发行也不需要一个中心化机构，而是由代码与社区共识来完成。有了比特币以后，就无需中心化平台做信任的桥梁，区块链通过全网的参与者作为交易的监督者，交易双方可以在无需建立信任关系的前提下完成交易，实现价值的转移。如果说互联网TCP/IP协议是信息的高速公路，那么区块链的诞生则意味着货币的高速公路第一次建设已经初步形成。

2.6.2 区块链原理

1. 从记账开始讲起

区块链是比特币背后的技术，比特币和区块链是同时诞生的。比特币背后的技术被单独剥离出来，称为“区块链”。那么，区块链是如何运作的呢？这里需要从记账开始讲起。

货币最重要的行为就是交易，交易会产生记录，就需要记账。比如，如表2-3所示，这个账本就记录了很多条交易。比如，编号为501的记录，表示王小明给陈云转账了20元钱，

编号为 502 的记录，表示张一山给刘大虎转了 80 元钱，编号为 505 的记录，表示央行发行了 1000 元货币。

表 2-3 一份交易记录

编号	转账人	收款人	金额
.....
501	王小明	陈云	20
502	张一山	刘大虎	80
503	林彤文	司马鹰松	500
504	李文全	赵明亮	180
505		央行	1000
506	央行	某某	1000
.....

从数据的结构来说，每次转账其实就是一条数据记录，我们把这种方式叫做记账方式，就是记了一笔某人转给另一个人钱的账，这个记账叫“中心式记账”。法币是由我们信任的中心化机构（政府、银行）记账。几乎所有的银行都是用中心化记账的方式维护巨大的数据库，这个数据库保留我们所有钱的记录。

中心化记账有很多的好处，比如，数据是唯一的，不容易出错。你如果足够信任它的话，转账效率特别高，在同一个数据库里，瞬间就可以完成转账。但是，我们的信任往往会被辜负，可能会存在中介系统瘫痪、中介违约、中介欺瞒、甚至是中介耍赖等风险。

2. 比特币要解决的第一个问题：防篡改

为了避免以上问题，是否有一种货币可以不用中心化机构来记账呢？这也是比特币发行的初衷。不由传统的“可信”的中介机构记账，那么由谁来记账呢？怎样保证新的记账者不会篡改交易记录呢？黑客攻击篡改交易记录怎么办？这就是比特币要解决的第一个问题：防篡改。

为了实现“防篡改”，就需要引入哈希函数。哈希函数的作用是将任意长度的字符串，转变成固定长度的输出（比如 256 位），输出的值就被称为“哈希值”。哈希函数有很多，比特币使用的是 SHA256。哈希函数必须满足一个要求，就是计算过程不能太复杂，用现代计算机去计算，应该可以很快得到结果。

比如，图 2-27 中，输入字符串是“把厦门大学建设成高水平研究型大学”，经过哈希函数转换以后，输出是“EFC15F94D49618595530E6283465702D04E7F6718DA0B6843DC199209848FBF5”。当输入字符串是“把厦门大学建设成高水平研究型大学！”，经过哈希函数转换以后，输出是“17846B37261F9AF3B4C3F9B99DE378601E9D11FD8C2ACDF82DCB3E0ECD56DC3A”。可以看出，只要输入字符串发生微小变化，哈希函数的输出就会完全不同。

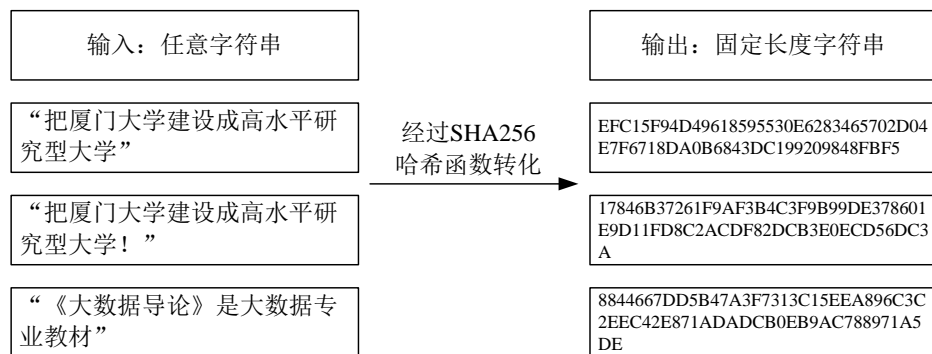


图 2-27 使用哈希函数转换的效果

哈希函数有两个非常重要的特性：

(1) 第一个特性：很难找到两个不同的 x 和 y ，使得 $h(x)=h(y)$ ，也就是说，两个不同的输入，会有不同的输出；

(2) 第二个特性：根据已知的输出，很难找到对应的输入。

这里重点看一下第一个特性。输入字符串是一个任意长度的字符串，是一个无限空间。而哈希函数的输出是固定长度的字符串，是一个有限的空间。从无限空间映射到有限空间，肯定存在多对一的情况，所以，肯定会存在两个不同输入对应于同一个输出值的情况。也就是说，肯定存在两个不同的 x 和 y ，使得 $h(x)=h(y)$ 。虽然这种情况在理论上是存在的，但是，实际上不知道用什么方法可以找到。因为，这里面没有任何规律可言，需要用计算机把所有可能的字符串都遍历一遍，但是，即使用目前最强大的超级计算机去尝试，也几乎要花费无穷无尽的时间，才能找到这样一个字符串。现在计算机找不到，那么，将来计算机发展了，是不是可以很容易找到呢？一样找不到，因为，如果计算机变得更强大了，那么只要把哈希函数输出值的长度变得更大即可。

了解了哈希函数以后，现在就来了解一下什么是区块链。

上面已经介绍过，所有的交易记录都被记录在一个账本中。这个账本非常大，于是，我们可以把这个大账本进行切分，切分成很多个区块进行存储，每个区块记录了一段时间内（比如 10 分钟）的交易，区块与区块之间就会形成继承关系。以图 2-28 为例，区块 1 是区块 2 的父区块，区块 2 是区块 3 的父区块，区块 3 是区块 4 的父区块，依此类推。每个区块内记录的交易条数可能是不同的，因为，每 10 分钟生成一个区块，如果 10 分钟内发生的交易次数较多，则这个区块记录的交易条数就较多。

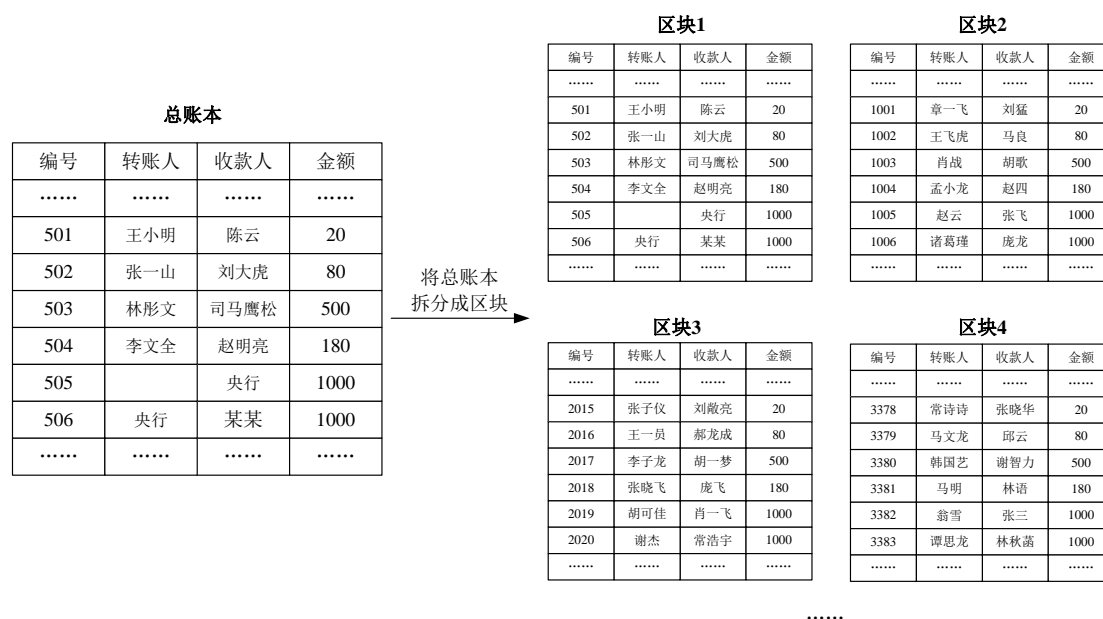


图 2-28 把一个账本拆分成很多个区块

然后，在每个区块上，增加区块头，其中记录了父区块的哈希值。通过在每个区块存储父区块的哈希值，就可以把所有区块按照顺序组织起来，形成区块链。如图 2-29 所示，区块 45 包含了一个区块头和一些交易记录，这些实际上都是一些文本，我们将这些文本内容打包，打包之后计算得到一个哈希值，这个哈希值就是区块 45 的哈希值，然后把这个哈希值记录在区块 46 的区块头里面。每个区块都如此操作。每个区块的区块头，都存储父区块的哈希值。这样就将所有区块按照顺序连接了起来，最终形成了一个链条，就叫“区块链”。

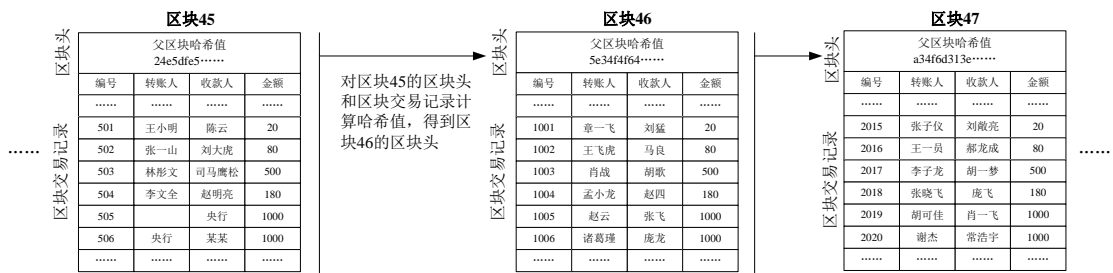


图 2-29 区块链示意图

那么，区块链是如何防止交易记录被篡改的呢？

假设我们修改了区块 45 的一点内容，这个时候，当有其他人来检查的时候，他很容易就可以发现，区块 46 中已经记录的关于区块 45 的哈希值，和最新计算得到的区块 45 的哈希值，二者不一样了。由此，他就可以知道，区块 45 已经被修改过。假设修改区块 45 的人的权限很大，他不仅把区块 45 的内容修改了，他还同时把区块 46 中的区块头的内容也修改了，那么实际上，我们也能够发现篡改信息的行为。因为区块 46 的头部被修改了以后，重新计算得到的区块 46 的哈希值，就和保存在区块 47 中的头部的哈希值不同了。假设这个篡改的人很厉害，他把信息一直篡改下去，不仅改 45 区块，也改了区块 46 的头部，也改了区块 47 的头部，一直改下去，一直改到最后一个区块，也就是最新的一个区块，那么也没有什么问题。因为，他只有获得最新区块的写入权，才可以做到。而要想获得最新区块的写入权（也就是记账权），他就必须控制网络中至少 51% 的算力。但是通过硬件和电力控制算力的成本十分高昂。比如，以 2017 年 11 月 16 日的价格计算，在比特币网络进行 51% 攻击，每天的成本包含大约 31.4 亿美元的硬件成本和 560 万美元的电力成本。

因此，我们可以说，区块链能够保证里面记录的信息和当时发生的时候一模一样，中间没有发生篡改，这就是区块链可以防篡改的原理。

3、在比特币的世界中如何进行交易

现在介绍一下在比特币的世界中如何进行交易？

进行交易，需要账号和密码，分别对应于公钥和私钥。在比特币中，私钥是一串 256 位二进制数字。获取私钥不需要申请，甚至不需要计算机，可以自己抛硬币 256 次来生成。地址由私钥转化而成，但是，根据地址不能反推出私钥。地址即身份，代表了比特币世界的 ID。一个地址（私钥）产生以后，只有进入区块链账本，才会被大家知道。

为了方便理解，可以做个比喻（如图 2-30 所示），比特币世界中的地址，就相当于现实世界中的银行卡号，比特币世界中的私钥就相当于现实世界中的银行卡密码。但是，它们之间还是有一些区别的，具体如下：

- （1）银行卡密码可以修改，但是，私钥一旦生成，就无法修改；
- （2）银行卡需要申请，而地址和私钥自己就可以生成；
- （3）银行卡实名制，地址和私钥是匿名的；
- （4）个人申请银行卡有限制，但是地址和私钥可以无限生成。

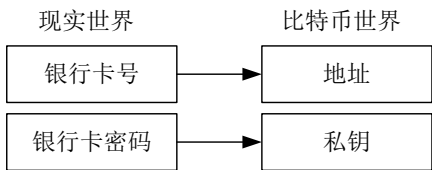


图 2-30 银行卡和比特币的对比

这里需要重点强调的是，在比特币世界中，私钥就是一切。首先你的地址是由私钥产生

的，然后你的地址上有多少钱，别人都是知道的，因为这些账本都是公开的。然后，只要他知道你的私钥，那么，他就可以发动一笔交易，把你的钱转到他自己的账户上面去。所以，一旦丢失了私钥，就丢失了一切。它不像银行卡一样，对于银行卡而言，如果别人只是知道了你银行卡的密码是没有用的，还要知道你的卡号。而且你还可以挂失银行卡。但是，在比特币的世界中，一旦你丢失了私钥，你就什么都没有了。所以，如何保管你的私钥，是一个很有讲究的问题，很多人都因为丢失了私钥，而造成了很大的损失。

现在假设张三已经有了地址和私钥，想要转账给李四 10 元钱，如何将这条交易记录添加到区块链中呢？

在把一条交易记录添加到区块链之前，首先就需要确认交易记录的真实性，这时就需要用到数字签名技术。如图 2-31 所示，张三调用签名函数 `Sign()` 对本次转账进行签名，然后，其他人通过验证函数 `Verify()` 来验证签名的真实与否。也就是说，张三通过签名函数 `Sign()`，使用自己的私钥对本次交易进行签名，任何人都可以通过验证函数 `Verify()`，来验证此次签名是否由持有张三私钥的张三本人发出的（而不是其他人冒用张三的名义），是就返回 `True`，反之返回 `False`。`Sign()` 和 `Verify()` 由密码学保证不被破解。

签名函数 `Sign(张三的私钥, 转账信息: 张三转10元给李四) = 本次转账签名`

验证函数 `Verify(张三的地址, 转账信息: 张三转10元给李四, 本次转账签名) = True`

图 2-31 签名函数和验证函数

签名函数的执行都是自动的，并不需要我们手工去处理这些事情。比如我们安装了比特币钱包 APP，它就会帮我们去做这样的事情，因为钱包 APP 是知道我们的私钥的，所以，我们只要告诉这个 APP，我想转账 10 元给李四，那么这个钱包 APP 会帮我们自动生成这次转账的信息以及签名，然后向全网发布，等待其他人使用 `Verify()` 函数来验证。

4. 比特币要解决的第二个问题：去中心化记账

一条交易记录的真实性得到确认以后，接下来的问题是，由谁负责记账呢？也就是说，由谁负责把这条交易添加到区块链中呢？

首先，我们就会想到，由银行、政府或支付宝这些权威机构负责记账，也就是采用“中心化方式”来记账。然而，历史上所有由中心化机构记账的加密数字货币尝试，都失败了。因为中心化记账的缺点很多，主要如下：

- （1）拒绝服务攻击。对于一些特定的地址，记账机构拒绝为之提供记录服务。
- （2）厌倦后停止服务。如果记账机构没有从记账中获得收益，时间长了，就会停止服务。

- （3）中心机构易被攻击。比如破坏服务器和网络、监守自盗、法律终止、政府干预等。

正是因为中心化记账会存在很多问题，因此，比特币需要解决第二个问题：去中心化。

在比特币区块链中，为了实现去中心化，采用的方式是：人人都可以记账，每个人都可以保留完整账本。任何人都可以下载开源程序，参与 P2P 网络，监听全世界发送的交易，成为记账节点参与记账。当 P2P 网络中的某个节点接收到一条交易记录时，它会传播给相邻的节点，然后相邻的节点再传播给其他相邻的节点，那么通过这样一个 P2P 的网络，这个数据会瞬间传遍全球。

采用去中心化记账以后，具体的分布式记账流程如下：

- （1）某人发起一笔交易以后，他向全网广播；
- （2）每个记账节点，持续监听、传播全网的交易。收到一笔新交易，验证准确性以后，将其放入交易池，并继续向其他节点传播；
- （3）因为网络传播，同一时间，不同记账节点的交易池不一定相同；
- （4）每隔 10 分钟，从所有记账节点当中，按照某种方式抽取 1 个，将其交易池作为下

一个区块，并向全网广播；

（5）其他节点根据最新的区块中的交易，删除自己交易池中已经记录的交易，继续记账，等待下一次被选中。

在上面的这个分布式记账的步骤中，还有一个很重要的问题就是，如何分配记账权。在比特币区块链中，采用的是 POW 机制，也就是“工作量证明机制”。记账节点通过计算数学题（如图 2-32 所示），来争夺记账权。

找到某随机数，使得以下不等式成立
 $\text{SHA256}(\text{哈希函数}(\text{随机数}, \text{父区块哈希值}, \text{交易池中的交易})) < \text{某一指定值}$

图 2-32 POW 机制的数学原理

上面这个数学公式的计算，除了从零开始遍历随机数碰运气以外，没有其他办法。解题的过程，又叫“挖矿”，记账节点被称为矿工。谁先解对，谁就获得记账权。某记账节点率先找到解，就向全网公布，其他节点验证无误之后，在新区块之后，重新开始下一轮计算，这种方式被称为“POW”。

总而言之，比特币的全貌就是，采用区块链（数据结构+哈希函数），保证账本不能被篡改；采用数字签名技术，保证只有自己才能够动自己的账户；采用 P2P 网络和 POW 共识，保证去中心化的运作方式。

2.6.3 区块链定义

前面以比特币为例对区块链原理做了基本介绍，现在就可以给区块链下一个定义。区块链是利用块链式数据结构来验证与存储数据、利用分布式节点共识算法来生成和更新数据、利用密码学的方式保证数据传输和访问安全的一种全新的分布式基础架构与计算范式。

区块链的三要素是交易、区块和链，具体如下：

- （1）交易：是一次操作，它会导致账本状态的一次改变，如添加一条记录。
- （2）区块：一个区块记录了一段时间内发生的交易和状态结果，是对当前账本状态的一次共识。
- （3）链：由一个个区块按照发生顺序串联而成，是整个状态变化的日志记录。

可以看出，区块链的本质就是分布式账本技术，是一种数据库。区块链用哈希算法实现信息的不可篡改，用公钥、私钥来标识身份，以去中心化和去信任化的方式，来集体维护一个可靠数据库。

2.6.4 区块链的应用

从科技层面来看，区块链涉及数学、密码学、互联网和计算机编程等很多科学技术问题。从应用视角来看，简单来说，区块链是一个分布式的共享账本和数据库，具有去中心化、不可篡改、全程留痕、可以追溯、集体维护、公开透明等特点。这些特点保证了区块链的“诚实”与“透明”，为区块链创造信任奠定了坚实的基础。而区块链丰富的应用场景，基本上都基于区块链能够解决信息不对称问题，实现多个主体之间的协作信任与一致行动。

总体而言，区块链在各个领域的主要应用如下：

- （1）金融领域。区块链在国际汇兑、信用证、股权登记和证券交易所等金融领域有着潜在的巨大应用价值。将区块链技术应用在金融行业中，能够省去第三方中介环节，实现点对点的直接对接，从而在大大降低成本的同时，快速完成交易支付。以跨境支付为例，跨境

支付涉及多种币种，存在汇率问题，流程繁琐，结算周期长；传统跨境支付基本都是非实时的，银行日终进行交易的批量处理，通常一笔交易需要 24 小时以上才能完成；某些银行的跨境支付看起来是实时的，但是实际上，是收款银行基于汇款银行的信用做了一定额度的垫付，在日终再进行资金清算和对账，业务处理速度慢。接入区块链技术后，通过公私钥技术，保证数据的可靠性，再通过加密技术和去中心，达到数据不可篡改的目的，最后，通过 P2P 技术，实现点对点的结算；去除了传统中心转发，提高了效率，降低了成本。

（2）物流领域。区块链可以和物流领域实现天然的结合。通过区块链可以降低物流成本，追溯物品的生产和运送过程，并且提高供应链管理的效率。区块链没有中心化节点，各节点是平等的，掌握单个节点无法实现修改数据；区块链天生的开放、透明，使得任何人都可以公开查询，伪造数据被发现的概率大增。区块链的数据不可篡改性，也保证了已销售出去的产品信息已永久记录，无法通过简单复制防伪信息蒙混过关，实现二次销售。物流链的所有节点上区块链后，商品从生产商到消费者手里都有迹可循，形成完整链条；商品缺失的环节越多，将暴露出其是伪劣产品概率更大。

（3）物联网领域。当区块链技术被应用于物联网时，智能设备将以开放的方式接入到物联网中，设备与设备之间以分布形式的网络相连接。在这个组织中，不再需要一个集中的服务器充当消息中介的角色。具体来说，以区块链技术为基础的物联网组织架构的意义在于，物联网中数以亿计的智能设备之间可以建立低成本、点对点的直接沟通桥梁，整个沟通过程无需建立在设备之间相互信任的基础上。

（4）版权保护。传统的版权保护方式存在两个缺点：第一，流程复杂，登记时间长，且费用高；第二，个人或中心化的机构存在篡改数据的可能，公信力难以得到保证。采用区块链技术以后，可以大大简化流程，无论是登记还是查询都非常方便，无需再奔走于各个部门之间，而且，区块链的去中心化存储，可以保证没有一家机构可以任意篡改数据。

（5）教育行业。在教育行业，学生身份认证、学历认证、个人档案、学术经历和教育资源等等都能够与区块链紧密契合。比如说，可以将学生的个人档案、成绩、学历等重要信息放在区块链上，防止信息丢失和恶意篡改，这样一来，招聘企业就能够真实可靠地得到学生的个人档案，有效避免了学历造假等问题。

（6）数字政务。区块链可以让数据跑起来，大大精简办事流程。区块链的分布式技术可以让政府部门集中到一个链上，所有办事流程交付智能合约，办事人只要在一个部门通过身份认证以及电子签章，智能合约就可以自动处理并流转，顺序完成后续所有审批和签章。

（7）公益和慈善。区块链上分布存储的数据的不可篡改性，天然适合用在社会公益场景。公益流程中的相关信息，如捐赠项目、募集明细、资金流向、受助人反馈等信息，均可以存放在一个特定的区块链上，透明、公开，并通过公示达成社会监督的目的。

（8）实体资产。实体资产往往难以分割，不便于流通，实体资产的流通难以监控，存在洗黑钱等风险。用区块链技术实现资产数字化后，所有资产交易记录公开、透明、永久存储、可追溯，完全符合监管需求。

（9）社交。区块链应用于社交领域的核心价值：让用户自己控制数据，杜绝隐私泄露。想想为什么我们刚刚浏览完某个购物网站，总会其他社交平台上收到类似的广告弹窗，因为数据隐私被垄断的大数据平台进行了贩卖。区块链技术在社交领域的应用目的，就是为了让社交网络的控制权，从中心化的公司转向个人，实现“中心化”向“去中心化”的改变，让数据的控制权牢牢掌握在用户自己手里。

2.6.5 大数据与区块链的关系

区块链和大数据都是新一代信息技术，二者既有区别，又存在着紧密的相互关系。

1. 大数据与区块链的区别

大数据与区块链的区别主要表现在以下几个方面：

（1）数据量。区块链技术是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式，区块链处理的数据量更小，是细致的处理方式。而大数据管理的是海量数据，要求广度和数量，处理方式上也会更粗糙。

（2）结构化和非结构化。区块链是结构定义严谨的块，通过指针组成的链，是典型的结构化数据，而大数据需要处理的更多的是非结构化数据。

（3）独立和整合。区块链系统为保证安全性，信息是相对独立的，而大数据的重点是信息的整合分析。

（4）直接和间接。区块链是一个分布式账本，本质上就是一个数据库，而大数据指的是对数据的深度分析和挖掘，是一种间接的数据。

（5）CAP 理论。C（Consistency）是一致性，它是指任何一个读操作总是能够读到之前完成的写操作的结果，也就是在分布式环境中，多点的数据是一致的。A（Availability）是可用性，它是指快速获取数据，可以在确定的时间内返回操作结果。P（Tolerance of Network Partition）是分区容忍性，它是指当出现网络分区的情况时（即系统中的一部分节点无法和其他节点进行通信），分离的系统也能够正常运行。CAP 理论告诉我们，一个分布式系统不可能同时满足一致性、可用性和分区容忍性这 3 个需求，最多只能同时满足其中 2 个，正所谓“鱼和熊掌不可兼得”。大数据通常选择实现 AP，而区块链则选择实现 CP。

（6）基础网络。大数据底层的基础设施通常是计算机集群，而区块链则是基于 P2P 网络。

（7）价值来源。对于大数据而言，数据是信息，需要从数据中提炼得到价值。而对于区块链而言，数据是资产，是价值的传承。

（8）计算模式。在大数据的场景中，是把一件事情分给多个人做，比如，在 MapReduce 计算框架中，一个大型任务会被分解成很多个子任务，分配给很多个节点同时去计算。而在区块链的场景中，是让多个人重复做一件事情，比如，P2P 网络中的很多个节点同时记录一笔交易。

2. 大数据与区块链的联系

区块链的可信任性、安全性和不可篡改性，正在让更多数据被释放出来，区块链会对大数据产生深远的影响：

（1）区块链使大数据极大降低信用成本。人类社会未来的信用资源从何而来？其实中国正迅速发展的互联网金融行业已经告诉了我们，信用资源会很大程度上来自大数据。通过大数据挖掘建立每个人的信用资源是很容易的事，但是现实并没有如此乐观。关键问题就在于，现在的大数据并没有基于区块链存在，这些大的互联网公司几乎都是各自垄断，导致了“数据孤岛”现象。在经济全球化、数据全球化的时代，如果大数据仅仅掌握在互联网公司的话，全球的市场信用体系建立是并不能去中心化的，如果使用区块链技术让数据文件加密，直接在区块链上做交易，那么我们的交易数据将来可以完全存储在区块链上，成为我们个人的“信用之云”，所有的大数据将成为每个人产权清晰的信用资源，这也是未来全球信用体系构建的基础。

（2）区块链是构建大数据时代的信任基石。区块链因其“去信任化、不可篡改”的特性，可以极大地降低信用成本，实现大数据的安全存储。将数据放在区块链上，可以解放出更多数据，使数据可以真正“流通”起来。基于区块链技术的数据库应用平台，不仅可以保障数据的真实、安全、可信，而且，如果数据遭到破坏，也可以通过区块链技术的数据库应用平台灾备中间件进行迅速恢复。

（3）区块链是促进大数据价值流通的管道。“流通”使得大数据发挥出更大的价值，类似资产交易管理系统的区块链应用，可以将大数据作为数字资产进行流通，实现大数据在更加广泛领域的应用及变现，充分发挥大数据的经济价值。我们看到，数据的“看过、复制即被拥有”等特征，曾经严重阻碍数据流通。但是，基于去中心化的区块链，却能够破除数据被任意复制的威胁，从而保障数据拥有者的合法权益。区块链还提供了可追溯路径，能有效破解数据确权难题。有了区块链提供安全保障，大数据将更加活跃。