

大数据协作征信的组织机制与发展对策研究

李迟芳 马毅

(佛山职业技术学院, 广东省、佛山市, 528225)

摘要: 征信是增强企业诚信、促进社会互信、减少经济矛盾并提高金融、商业、行政管理效率的有效手段。文章梳理了大数据给征信带来挑战, 以及相互融合的原则与组织协作过程。重点是立足组织协作理论, 对自主征信协作、链主约束型协作征信、多中心协作征信三种模式展开分析, 发现了三种模式在组织对象、组织层级、资源约束与交流等组织机制特征, 并就大数据征信协作的技术与内容提出了相应的对策。

关键词: 征信; 大数据; 协作; 组织; 层级约束

中图分类号: F832.5 **文献标识码:** C

一、引言

伴随着物联网、云计算、移动互联网技术的发展, 金融征信、商业征信、行政管理征信的需求越来越频繁。与此同时, 征信数据来源已不再局限于传统的业务数据和国家及相关部门的统计数据, 数据来源渠道从企业内网、微博、B2B 和 B2C 网站逐渐扩展到社交网络、政务网络、传感网络、智能终端网络等。它们将产生海量的多源、异构的数据元, 可以辅助人们从不同方面洞察、验证、评价征信对象的资质、动机、意图和优劣势等, 甚至是勾勒出完整的个性画像。但是, 我国正处于经济发展转型期, 相关利益主体的多元化使得“数据割据、数据孤岛、数据质量”问题严重, 这制约了数据的开发与利用。另据 Gartner 的一项调查指出: 现代企业 80% 的数据都是传感器、音视频、图片、GPS 等非结构化数据, 且这些数据每年都按照 60% 的速度增长[1]。在海量数据不断涌现的同时, 如何打通各数据渠道, 实现不同利益主体的大数据征信协作成为当前社会各界关注的焦点。本文拟围绕征信的大数据协作问题, 研究大数据发展下征信协作的机制与特征, 为相关利益主体规划与布局提供一定借鉴。

二、大数据开启了征信协作的新篇章

随着计算机计算速度的快速提升, 为解决“大规模数据”的应用问题, 数据库、数据仓库、数据集市等技术依次出现。2011 年麦肯锡环球研究院提出大数这个概念, 它在研究报告中指出: 数据已渗透到了每一个行业, 并将成为重要的生产因素[2]。这得到了全国互联网巨头认可, 同时激发了大数据技术开发与应用浪潮。与前几代“大规模数据”应用技术不同, 大数据不只在数据数量的巨大, 它还体现了从大容量、多结构和不同类型数据中, 高速挖掘、分析和提炼出新知识、新信息的架构与技术。国际数据公司 (IDC) 归纳了大数据的 4V 特征, 即数据规模巨大 (Volume)、数据类型多样 (Variety)、价值密度低 (Value)、高速性 (Velocity)。IDC 同时指出: 到 2024 年, 全球电子形式存储的数据至少将达到 40ZB,

且每两年翻一倍；这意味着传统需要大量数据积累的应用有可能实现，但同时数据量和非结构化数据应用也迎来了新挑战[3]。

任何一个社会组织出于保持或提升竞争力，都有可能从某些特定维度抽取或提供全面征信数据的一个子集，而如何从海量的、跨组织数据中获取有价值的信用信息是征信服务组织所面临的最迫切的挑战。另外，从企业、国家和政府的角度，大数据也开启了智能决策的新篇章。大数据能够强化贸易、经济、国家安全、政府公共决策的预见性和响应性。2016年3月，美国发布了《联邦大数据研发战略计划》“大数据研究和发展计划”，形成涵盖技术研发、数据可信度、基础设施、数据开放与共享、隐私安全与伦理、人才培养以及多主体协同等七个维度的系统的顶层设计，打造面向未来的大数据创新生态，其中核心工作就是基于大数据技术整合多种数据源，并提高信息处理的能力[4]。可以看出，基于大数据技术的数据共享与协作是大数据应用的关键问题。

三、大数据与征信协作体系

（一）大数据在征信协作应用中的挑战

虽然国外 Equifax（艾可飞）、Experian（益百利）、D&B（邓白氏）和国内阿里等为代表的征信服务机构都有着众多大数据征信成功的优秀案例，但是大数据技术在征信领域的全面应用还存在如下难题：

1、大数据存在不均衡和不公平问题。目前几乎所有的大数据网络平台或多或少的受制度漏洞、地方保护主义等干预，被人为的数据分散，造成了“数据群雄割据”；与此同时，由于技术差距、历史遗留，物理上的各自存储、维护，以及各利益部门对数据的自我理解和定义，使得数据的收集、存储及应用存在“物理和逻辑上的孤岛”；再加上匿名的网络环境及商业利益关系让大量数据的真实性、完整性、一致性的“数据质量”问题，都是大数据不均衡、不公平的表现。

2、系统建设与运营成本极高。Equifax、Experian、阿里等之所有能够运用大数据开展征信活动，除了其特有的数据入口基因优势外，还离不开强大的数据收集、计算、存储和数据分析技能。而开展大数据存储和分析的代价也是极高的，这不仅需要巨额建设投资，运营成本也十分高昂，仅用电消耗就很是惊人。以阿里的数据中心为例，不算空调、安防、防火墙等，该中心一天服务器的耗电量就超过 150 万千瓦时，占了整个杭州用电量的 1/80，相当于一个小型城市的用电需求¹。少超过半数的受访机构将内部数据视为“大数据”的主要渠道，其包括了企业及相关机构一定年限内的生产、交易、业务、事件和电邮等内部数据，大部分企业主要通过这些数据的大数据分析为企业决策提供洞察依据。包括阿里巴巴开展的大数据征信活动，其使用的数据也是以其生态体系内部的闭环数据为主，外部数据的利用也

¹数据来源：杭州统计调查信息网，作者整理。

是寥寥可数。

3、数据用户端的成本被忽视。由于“数据割据、数据孤岛、数据质量”，数据使用者的成本不断叠加，用户可能需要在—项服务中重复支付成本，这直接影响到了用户的体验，严重影响了大数据的市场服务活力。

4、大数据人才严重缺乏。据预测，美国到2024年具有高度分析技能的专业人才需求缺口为14-19万人，而懂得应用大数据工具进行管理和分析人员的缺口达150万人[5]。

在现实情况下，要实现强大的大数据征信服务，在数据基因、资产投资、效能管理、专业人才等各个方面都面临众多困难。同时，我国大多数产业、企业组织集中度比较低，行业、产业组织也相对比较弱小，也很难由一个组织或平台独立且全面的实施大数据方方面面的服务。故金融、商务、行政管理征信服务的提供者需要转变思路，合理规划和利用分散的大数据资源，基于风险管理、商务服务、政务协同等需要进行大数据的征信资源的协作与共享。

（二）大数据征信协作体系的组织原则

从全局来看，任何一个组织都是全面数据中某一个维度的生产者、采集者，同时也可能是全面数据的研究者或受益者。例如商务活动过程中，双方的交易历史、交易特征、交易场景等都是数据的综合体现，而交易双方如能实时、准确且低成本的把控对手的信用状况，可以有效的提高商务活动效率，而这些信息被金融机构、行政单位或第三方捕捉到后，可有效的促进商务、政务规模，并形成优胜劣汰机制，将奠定一个高效地市场化环境。

考虑到技术或信用数据资产保护的需要，形成这种高效的商业化大数据环境不可能是对用户共享开放数据本身，也不可能将所有组织的大数据资源实现某种形式的聚合，而是提供一种组织协作的大数据协作系统，按照用户需求挖掘分析数据并返回相关结果。同时，由于大数据具有价值多样性、不确定性等特点，不同层级的数据得到的分析结论不存在必然相关，所有不同组织所采用的分析模型是有效结论获得的关键要素。在此情形下，即使是大数据体系被竞争对手相互使用，其分析结论也具有自我独特性，可以有效保障其商业存在。

（三）大数据征信协作的具体过程

在共享协作原则下，每一个致力于大数据征信服务的组织，都需建立大数据技术系统，以收集内、外部数据，并运用云计算架构对各自数据开展存储、分析和共享服务。当用户提出分析任务后，服务组织直接在自身内部数据中执行挖掘分析；当需要跨组织或需要更大规模大数据集时，则按照下述步骤执行：首先是用户任务的系统分解；其次是系统分发子任务至其他组织共享服务；再次是其他组织将接受到的任务映射为本地任务，并利用本级大数据挖掘和分析；最后由主服务组织对本地和分发的任务结果进行汇总，并最终向用户提交分析结果。整个过程用户不需了解不同服务组织的技术，各服务组织也只需要执行映射的本地任务即可，这种分布式架构可以很好地降低组织协作之间的技术耦合性，并保证数据的独立和可控。

四、组织协作相关的理论与借鉴

组织协作的经济学理论内涵主要分为三大理论学派：第一个是威廉姆森的契约协作理论，他认为协作就是为了事后租金分配过程中履行契约的各种复杂因素的综合，他提出协作由一系列契约设计、制度安排来推动，最终实现降成本、减冲突及共同获利。但是他忽略了非正式制度（如学习、模仿、信任等关系）的作用，以及组织机制的发展。第二个是麦克内尔的关系协作理论，他认为企业内部资源和能力有限，必须动态的依赖新的外部关系协作。他比较强调政治学、社会学领域的协作，尤其注重相互之间信用及信任等非正式制度的作用，有些过犹不及。第三个是网络协作理论，他重新纳入了交易成本、社会关系等资源与活动，认为协作是利益相关者的资源关系的均衡交易。可以看出，协作的三大理论都是围绕着对象、组织层级、规则与约束三个方面的内容展开的分析。

所以，大数据背景下征信协作体系应该聚焦到不同征信服务的对象、组织构建的层级、不同构架下的资源交流机制和约束等方面，从而挖掘出大数据征信协作的本质。

五、大数据背景下征信协作的模式研究

借鉴国内外大数据金融产业链的相关案例，依据数据来源及使用过程，大数据背景下融资征信协作的模式可以归纳为三种：自主征信协作模式、链主约束型协作征信模式、多中心协作征信模式。其主要特征如表 1。

表 1 大数据背景下不同融资征信协作模式的主要特征

	自主征信协作模式	链主约束协作征信模式	多中心协作征信模式
协作发起主体	某维度的数据平台 或利益第三方	大型行业龙头或数据平台	跨行业协调组织或政 府部门
组织关系	大数据的短供应链 及产业互补和上下游组 织	链主企业主导 生态+闭环	稳定且有弹性的战略 关系
数据资源交换	临时性、频次较少 且单向居多	频次较多、双向	较多、双向
协作约束要素	价格 地方和企业的数据规范	链主的核心产品或规则	价格与数据权威
协作机制	市场>人际网络>层级	层级>人际网络>市场	层级>市场>人际网络
典型案例	阿里、科技小贷平台等	京东、海尔等	全球网、信联网等

1、自主征信协作模式

自主征信协作是一种应急性的征信中介行为，其发起征信调查的一般是征信对象所依附的某个数据平台或利益第三方，由数据平台或利益第三方基于某个维度的大数据，对该征信

对象开展大数据征信挖掘与分析,并提供资金、咨询、商务等服务的一种中介活动。这种模式的层级关系简单,只受数据平台的简单层级约束,行业、产业组织参与较少。

很多科技小贷平台以及早期的阿里小贷都是该模式的典型代表。大数据发展下自主征信协作的特征包括:(1)征信对象主要是数量众多的中小微企业、甚至个人,他们没有足够的资产或资源,只能寄希望于第三方大数据方式来自证其信用。(2)为提升金融、商务、政务活动效率,征信对象间可能会出现大量互助性、非战略性的双边信用组合,如友情、亲情;它们主要来源于短供应链的上下游或互补单位,同时为平衡征信成本及利益,这种互助组合更容易发生在声誉和信用规模相当的企业间。(3)在市场竞争或者突发事件下,由于缺少层级约束,征信发起机构的征信程序及决策过程容易出现非集约化问题,出现机会主义替代的风险。如部分 P2P 平台甚至认为简单的舆论和社会身份识别就可以有效的监督贷款人,最终滋生了大量恶意欺诈贷款人,最终导致 P2P 平台跑路的事件。

结合上述特征,自主征信协作的大数据活动应建立如下对策:(1)在数据渠道协作上,多数征信活动都是应急性的商务、金融、政务需求,使得大数据征信资源协作具有临时性、频次较少且单向交流居多的特征。(2)非战略性的互助关系实际上是一种互相学习、互相模仿的行为扩散,故对其大数据的呈现某个维度或者某类数据的整体趋向性挖掘和分析,是征信实施机构实施大数据挖掘和分析的主要思路。(3)考虑到行业、产业层级资源参与较少,自主征信协作的发起机构应聚焦于自身优势大数据挖掘与分析,并适当承担跨组织的大数据映射任务以降低成本、提高效率。

2、链主约束征信协作模式

随着规模经济、产业经济的发展,资源和数据将逐渐地集中并汇集到少数几家平台或产业龙头,形成大数据链主或平台。链主可能是缔造某个产业产品的大型龙头企业,也可以是集聚了流通或者交易入口的网络平台或政务平台,它们控制着这个产业产品或数据的核心环节,其他组织或服务机构接受产品或服务的辅助外包。这种模式下包含了不同规模的产业组织和机构,其层级约束更多,征信协作的对象和活动也更复杂。

本模式的典型代表是京东金融、海尔供应链金融等。这种模式一般是链主居于产业链的焦点位置,并与跨行业、跨层级的企业建立了稳定的契约或信任关系,各层级商流、信息流、资金流的规则 and 标准都由链主把握,在大数据分析、挖掘上呈现以下几个特征:(1)因为上游企业比下游企业的融资、商务议价能力更强,很容易出现“上游单位借下游单位的‘壳融资’”现象。(2)外部资源的交换基本上都由链主决定或者承担,同时由于对地方经济可能产生影响,链主的外部交易容易得到地方政府,公共部门的支持。

这种征信协作体系下,我们的大数据征信协作活动应该:(1)大数据的协作征信实施要侧重对议价能力更强层级的上游单位的数据分析与挖掘。(2)这种模式具有明显的自上而下的层级活动特点,在大数据征信监控上,应找准并加强对核心企业数据递进式的关联与监控分析。(3)在征信协作组织构架上,获得链主一定维度的数据支持是实施大数据征信

协作的关键。

3、多中心协作征信模式

多中心协作征信接近于奥斯特罗姆提出的“多中心协作理论”的思想[7]。数据割据、数据孤岛和数据质量问题让海量数据分散集合于企业、政府部门、科研机构、中介机构手中。

国内的“信联网”和“全球网”是该模式探索的国内先驱者。多中心协作征信模式的层级结构和活动主体更具有弹性。在实施大数据征信协作上，这种模式具有以下特征：（1）除了大数据征信资源提供的企业和服务组织外，还存多个第三方大数据服务机构，各征信组织主要进行用户任务的分解和分发工作。（2）由于存在多个数据平台博弈，不同规模的征信活动对单一维度数据的依赖性将下降。例如水电、经营产所甚至员工风貌的大数据挖掘分析也可能替代传统的财务、商务流分析。（3）由于多中心数据来源与分析，使得信息对称度大幅提升，平行企业间的金融担保、商务互助等行为特性会更容易挖掘与分析，使得大数据关联与聚类的有效性大幅提升。

多中心协作的组织层级更复杂，在征信协作上要注重以下策略：（1）数据平台的经济性和政府的权威性是决定征信协作开展的最关键要素。（2）政府、征信组织、数据核心机构之间由于存在利益差异，政府的牵线可能最终演变成政府主导，最终丧失市场化动力，故如何保证政府的第三方位是长期协作的关键策略，例如通过建立大数据征信产业发展基金或相关组织的方式可以有效规避政府的行政权力扩散。

六、不同模式的演进与发展建议

动态来看，大数据发展下的征信协作不一定会固定地遵循某种模式，而是人为或自发的成长、迁移。当现有的协作在对象、层级、规则与约束机制出现效率低下时，企业间协作的制度框架必然会被打破而向新的协作模式演变。

大数据作为重要的资产已经在潜移默化的影响着征信的方方面面。不同规模、性质的企业和组织应利用自身行业、区域的资源优势，重视和参与数据的收集、分析、共享等工作，还可以通过发挥数据协调、管理等职能提升企业影响，占领下一个风口。同时，政府要加强民营征信组织“国民化”待遇的步伐，形成良好的大数据市场化环境，真正倒逼金融、商业、政务环境的改革。另外，各行各业要加强大数据行业协会的建设，通过行业组织制定相关标准，一方面维护数据安全，另一方面促进大数据跨行业、区域、平台的共享。

参考文献

- [1]揭秘隐藏在非结构化数据背后的真相[EB/OL].
<http://tech.it168.com/a2012/0305/1320/000001320266.shtml>,2012-03-06.
- [2] Big data: The next frontier for innovation,competition,and,productivity[R/OL].Mckinsey Global Institute.2011.
- [3]Philip Russom.Big data analytics[EB/OL].<http://www.Docin.com/p-240502098.htm>,2014-07-08.
- [4]于磊.基于应急管理视角的政府应急专业处置能力建设的思考[J].中国管理信息化, 2012, (22): 92.
- [5] 李雪梅. 我国政府数据开放的困境及对策——基于大数据征信视角[J]. 征信, 2020(9).
- [6] 李仪, 陶宇. 大数据技术下消费者个人信用信息共享的治理机制[J]. 中国信用, 2020(2).

Research on organization mechanism and Development Countermeasures of big data cooperative credit investigation

Li ChiFang Ma Yi

Foshan technical college, Foshan City, Guangdong,528225

Abstract: Credit investigation is an effective means to enhance enterprise integrity, promote social mutual trust, reduce economic contradictions and improve the efficiency of finance, commerce and administration. The article combs the challenges brought by big data to credit investigation, as well as the principle of mutual integration and the process of organizational cooperation. Based on the organization cooperation theory, the focus is to analyze the three modes of independent credit investigation cooperation, chain master constraint cooperation credit investigation and multi center cooperation credit investigation, find the organizational mechanism characteristics of the three modes in organizational objects, organizational levels, resource constraints and exchanges, and put forward corresponding countermeasures for the technology and content of big data credit investigation cooperation.

Keywords:Credit investigation big data; Collaboration; Organization; Hierarchical constraints