

大数据在民营企业信用体系建设中的应用研究

■ 占小凤

摘要：互联网的快速发展为民营企业信用体系建设带来了新的机遇和挑战，作为互联网的重要产物——大数据，参与了民营企业信用体系的全过程。本文重点研究大数据在民营企业信用体系建设的应用现状、存在的问题和对策建议。

关键词：大数据 民营企业 信用体系 应用

一、大数据在民营企业信用体系建设中的应用现状

（一）大数据及大数据产业内涵

“大数据”是指数量巨大、类型众多、结构复杂、有一定联系的各种数据所构成的数据集，其特点可以总结为5V（IBM提出）：Volume（大量）、Velocity（高速）、Variety（多样）、Value（低价值密度）、Veracity（真实性）。我国大数据主要运用于商业智能、政府决策、公共服务三大领域，其应用离不开大数据产业的发展。所谓大数据产业，是指以数据生产、采集、存储、加工、分析、服务为主的相关经济活动，包括数据资源建设、大数据软硬件产品的开发、销售和租赁活动，以及相关信息技术服务。目前，我国大数据产业链的上游是电信运营商和BAT（指百度、阿里巴巴、腾讯三大互联网公司巨头）等数据存储层企业，中游是拓尔思、同有科技、浪潮、华为等数据分析处理层企业，下游是贵阳大数据交易所、数据堂交

易平台等数据交易层企业和百分点集团、明略数据等数据应用层企业。

（二）民营企业信用体系建设

民营企业信用体系建设与企业信用体系不可分割，多项工作重合，如企业信用标识制度建设、企业信用信息基础数据库建设等等。具体来说，企业信用标识制度确立上，早在2017年国务院就在全中国部署推进多证合一工作，建立全国统一社会信用代码制度；企业信用信息采集上，目前中国人民银行征信中心以基本信息、借贷信息为主，第三方征信公司在公共信息、金融信息的基础上引入了媒体评价信息、行业评价信息、市场反馈信息、经营信息等多类信息；企业信用信息系统建设上，目前较为常用的有国家市场监督管理总局开发的国家企业信用信息公示系统、第三方开发的企查查和绿盾企业征信系统；企业信用信息应用上，目前企业信用信息主要被行政机关、利益相关者、协会商会、金融机构和信息主体自身这五类群体使用。尽管民营企业信用体系

建设成效显著但也有诸多问题：首先，民营企业数量多，信息采集难且数据失真；其次，信息共享难，各征信机构采集标准的差异和数据的资源性阻碍了数据共享；最后，信用市场发育不完善，信用机构多且前期投入较大，不少信用机构无法盈利，且信息产品单一，信用产品还有待开发。

（三）大数据在民营企业信用体系建设中应用及其作用

在大数据方兴未艾的背景下，社会信用信息收集、整理、归纳、分析、评判离不开大数据技术。信用体系建设的各个环节都用到了各种大数据技术，如数据采集和预处理技术、数据存储技术、数据分析技术、数据挖掘技术、数据可视化技术等等。大数据技术推动了民营企业信用体系的发展，其作用主要是：大数据征信能纳入更为多样性的行为数据，拓宽了服务的信用主体覆盖度；还可以带来更为时效性的评判标准；大数据征信模型可以使信用评价更精准。

二、大数据在民营企业信用体系建设中的应用存在的问题

大数据推动了民营企业信用体系建设，同时也带来了巨大的隐患，它放大了传统征信手段下的信息泄露、信息非标准化问题，也同样不可避免地面临着信用产品单一的问题，具体来说，大数据在民营企业信用体系建设中存在以下问题：

（一）数据采集难且存在非法采集现象

如前所述，大数据时代，民营企业信用信息除了传统的借贷信息、公共信息外，还包括五花八门的媒体评价信息、行业评价信息、市场反馈信息、经营信息等等，加上我国民营企业数量多、分布广、家族化经营、财务管理不规范等等，这些都大大增加了民营企业信用信息采集难度，且造成了信用信息的可靠性和有效性不足。此外，数据源的丰富带来了数据获取方式的多样化。数据采集除了交换、购买外，还可以通过应用程序接口即API和网络爬虫法进行网络数据采集，或通过调研等线下方式获取。数据获取的多样化不可避免地带来了数据非法采集的多样化，在大数据时代，征信公司非法采集民营企业信用信息的方式有：以欺骗、胁迫、诱导的方式采集，未征得企业同意采集非公开的企业信用信息，向被采集的

企业收费的方式，“一次授权、无穷采集、无限使用”，过度采集等等。

（二）数据采集标准各异

2008年6月，中国国家标准化管理委员会出台《企业信用信息采集、处理和提供规范》(GB/T22118-2008)，作为企业信用信息采集方面的权威文件，它对基本信息、财务信息、经营管理信息、银行往来信息、提示信息和其他信息的采集原则和要求都做了详细的规定，但对媒体评价信息、市场反馈信息等网络数据的采集尚未提及。我国民营企业存在数量众多的信用白户，它们的信用报告往往只有基本信息，市场评价信息、媒体评价信息等非金融信息对其信用信息的补充具有一定意义，因国家明确其采集标准，各企业目前对其进行了不同探索，如“企查查”对新闻评价信息从舆情情感类型占比、新闻舆情趋势、近期媒体映像、新闻类型分布TOP10四个维度进行描述，而“绿盾征信”从媒体热评信息和舆情监测信息两个维度进行描述，“启信宝”则仅有新闻舆情一项与新闻评价信息有关。可见，各征信公司的数据采集标准差异较大，更不利于数据共享。

（三）数据共享不足

民营企业信用信息的来源主要为金融机构、政府和互联网公司，就目前而言，信用信息共享不足，这种不足既存在于三者之

间，也存在于它们各自的内部。共享不足的原因主要有：第一，数据采集标准各异。数据采集的统一性既决定于统一口径是否一致，即是否有统一的采集标准，又决定于采集人员的业务水平。如采集人员对数据的理解，如前所述，部分数据的采集标准尚未统一，采集数据涉及到水平参差不齐的众多企业和个人，难以保证数据的统一性。第二，未建立相应的数据使用管理办法，数据版权、产权很难得到保护，阻碍了数据的共享。第三，数据共享机制有待探索和建设。目前，作为企业信用信息的“主力”——公共信息与金融信用信息尚未融合，更谈不上与互联网征信数据的融合。他们之间如何融合，谁来牵头、谁来配合、如何考核等核心问题尚待研究，信息保护、系统开发、风控模型改进等关键问题有待解决，数据共享机制尚需探索。

（四）数据分析能力有待加强

大数据背景下，更多的数据涌入企业信用信息数据库，而且据邓白氏首席数据官统计，全世界的非结构化数据占到了85%，如何对数据进行归类、分析、挖掘是个很难的命题。以行业巨头邓白氏为例，邓白氏数据云包括超过4亿家企业记录、3,500个数据字段以及1,200个数据交换。在强大的数据处理能力基础上，邓白氏推出了商业资信报告、付款信息交流计划、合



规风险评估报告、企业受益所有权报告和全球市场风险解决方案五类产品，相比之下，国内很多第三方信用服务机构的数据获取能力很强但数据分析能力相对薄弱，很多机构提供的信用产品还停留在企业信用报告层面。我国数据的特点有别于国外，我国民营企业的特点也有别于其他企业，如何构建适合我国民营企业的数据分析模型是需要重点解决的问题，这需要第三方信用服务机构、科研机构、政府、金融、企业等各方的共同努力。

三、大数据在民营企业信用体系建设中的应用对策及建议

（一）规范数据采集

首先，统一数据采集标准。按照政策定向、市场定需、专家定位的工作思路，开展信用标准化研究，尽快对原版《企业信用信息采集、处理和提供规范》进行更新、补充，同时要注意在经济全球化的背景下数据采集标准尽量与国际接轨，为信用服务机构走出国门做准备。其次，对数据采集人员进行专业培训，提高数据采集质量，统一数据采集口径。大数据背景下，数据更为复杂、繁琐，通过多频次的专业培训尽可能地提高数据的准确性。最后，立法保护信息主体的知情权、同意权、异议权，规范异议处理流程，明确信息提供机构在

异议处理中应履行的职责和具体要求、解决时间等，保护信息主体权利的同时提高数据的准确性。

（二）建立信用信息共享机制

民营企业尤其是小微企业融资难、融资贵、融资慢，特别是信用贷款难的问题依然不容忽视。而增加民营企业信用贷款关键是增加银行的风险管理能力，让银行更加准确地判断企业的还款能力和还款意愿，减少抵押担保的依赖。因此，建立信用信息共享机制，有助于解决银行信息不足的难题，对银行更全面了解企业生产经营状况，更准确地判断企业风险具有关键意义，是破解民营企业融资难题的一项重要基础设施工作。2021年政府工作

报告首次提出加快信用信息共享步伐，2021年6月银保监会会同相关部门研究制定《普惠金融高质量发展的实施意见》，把加强信用信息共享作为“十四五”期间健全普惠金融基础设施的重点工作。建立信用信息共享机制可以从四个方面着手：首先，统一数据采集标准，提高数据采集人员业务水平；其次，加强《数据安全法》的落地和宣传，保护数据版权，促进数据共享；再次，鼓励征信公司探索建立企业互联网信用信息数据库，弥补传统征信信息；最后，银行配合各级政府建立各级信用信息共享及金融服务平台，同时发挥监管考核指挥棒的作用，将银行参与信用信息共享、创新产品服务模式的情况纳入综合评价指标，指导银行运用金融科技手段将公共信用信息与内部金融数据有机结合，改进业务审批技术和风控管理模型，提升对小微企业等长尾客户能贷、惠贷的能力。

（三）加强民营企业信用评价模型的研究

民营企业信用评价模型是信用报告的思路和灵魂。数据是错综复杂的，数据该如何取舍，信息该如何组织，企业该如何画像，这些问题都需要民营企业信用评价模型来解决。从全球三大巨头标普、穆迪、邓白氏（其中标普和穆迪主要针对资本市场规模较大的企业，邓白氏偏重于中小企业）的实践看，三者均认为

信用评价是对企业还款能力和意愿的考察，他们所构建的指标体系均含有经营因素和财务因素，邓白氏还加入了对企业还款意愿的考察。比较三大机构企业信用评价指标体系可知：评价模型与企业的发展阶段、经营战略、评价目的、评价对象有关，具体指标的选择和侧重会有所不同；现金流指标不易被操作，更能提示企业的违约风险。民营企业信用评价模型研发既可以来自征信公司也可以来自行业协会、哲学社会科学工作者，或多方的共同努力。

（四）提高数据分析能力

一份好的信用报告离不开好的数据和好的算法。数据是基础，没有足够量、足够真的数据，再好的算法也是白搭；算法是核心，没有从数据中提取信息的方法，数据终究是一堆数字、字符，不能为人们创造财富。好数据通过数据共享实现，好算法通过技术人员的努力实现。互联网背景下，数据更为复杂，如何去伪存真，去粗取精需要技术的支持。专业人士指出数据分析的步骤可以分为五步：明确问题、理解数据、数据清洗、数据分析和数据可视化。可见数据分析人员不仅需要较强的业务知识，还需要对财务、运营等企业管理知识有一定的了解，业务知识有助于对数据开展分析，企业管理知识有助于理解数据之间的逻辑关系，提高数据的挖掘能力。那么

数据分析能力如何提高呢？一方面，加大对专业人员的培养，为数据分析行业提供人才支撑，另一方面，业务培训中加强对新数据分析技术的培训和基本企业管理知识的培训。■

参考文献：

- [1] 吴晶妹. 从信用的内涵与构成看大数据征信[J]. 首都师范大学学报(社会科学版), 2015, (6).
- [2] 熊治东. 改革开放以来中国社会信用体系建设: 成就、经验、问题与展望[J]. 征信, 2020, 38(10).
- [3] 何玲, 孟佳惠. 大数据征信浪潮——大数据时代下征信业现状与发展前景扫描[J]. 中国信用, 2021, (3).
- [4] 徐广军, 倪晓华, 肖运香. 标普、穆迪、邓白氏企业信用评价指标体系比较研究[J]. 浙江金融, 2007, (3).
- [5] 李晓. 非银行信用信息数据采集存在的问题及对策研究[J]. 纳税, 2019, 13(32).
- [6] 吴杰. 英国《2015年中小企业信用信息条例》解析及对我国的启示[J]. 武汉金融, 2018, (5).
- [7] 宋书彬, 朱红. 互联网金融大数据平台信用评价与信息共享机制[J]. 河北金融, 2018, (11).

（作者单位：安徽省社会科学院）

责任编辑：康伟