

用数据解读大数据政策

合肥兴泰金融控股集团

金融研究所

徐蕾 李庆

咨询电话：0551-63753813

服务邮箱：xtresearch@xtkg.com

公司网站：<http://www.xtkg.com/>

联系地址：安徽省合肥市政务区

祁门路 1688 号兴泰金融广场 2206 室

2017 年 11 月 17 日

更多精彩 敬请关注

兴泰季微博、微信公众号



内容摘要：

- ◆ 大数据作为信息资本和数据资源将对经济发展、国家治理、政府决策、公共服务、组织方式和业务流程等产生巨大影响。作为重要的国家公共信息资源，大数据应被提升到国家战略高度。当前多个国家已进行了政府大数据管理先驱性探索，积累了经验。为提升我国在数据资源应用、数据要素流通、大数据产业发展等方面的能力，中央政府先后批准了八家国家级大数据综合试验区和一家大数据基础设施统筹发展类综合试验区（以下简称“大数据试验区”）。试验区批复后，在大数据组织架构建设、政策创新等方面都展开了一系列积极探索。
- ◆ 本期将充分运用大数据思维，综合扎根理论、社会网络等分析方法，对九家大数据试验区出台的八十九条政策展开可视化研究，针对目前大数据政策的发展现状和存在的不足展开讨论，为后续的政策制定提供参考。

用数据解读大数据政策

大数据是通过量化角度认识世界的有效途径，是改变市场格局、组织结构，以及政府与公众关系的独特方法。在寻求量化与认识世界的过程中，资源不断被数据化，对大数据的透彻掌握能使其转化为信息资本和数据资源，进而成为经济价值的来源。大数据具有转化为国家重要公共信息资源的潜力，作为重要的信息资本和数据资源，俨然被各国政府提升到了战略高度。2016年2月，我国首家大数据综合试验区获批成立，标志着我国大数据产业进入全新发展阶段。大数据试验区旨在通过制度创新、产业培育等手段，探索大数据在政务、物流、医疗等众多领域的应用，形成可借鉴、可复制、可推广的实践经验，最终形成试验区的辐射带动作用 and 示范引领效应。

一、大数据试验区出台政策简况

截至目前，中央政府共批复了9家大数据试验区，除成立最早的贵州试验区外，其他各家由于批复成立时间较短，关于大数据出台的相关政策并不多，据统计，9家试验区共出台了89项省市级大数据政策。此外，本文根据政策的内容和性质对政策进行简要分类，将政策划分为宏观指导类、具体应用类等几个类别。具体结果如下表所示：

表 1：大数据试验区出台政策数量、类别情况

试验区	政策数量	政策类别
贵州	34	宏观指导、具体应用、人才培育、产业规划、优惠扶持、平台建设
北京	2	宏观指导
广东	17	宏观指导、产业规划、具体应用
河北	5	宏观指导、具体应用
河南	9	宏观指导、具体应用
辽宁	4	宏观指导、具体应用、平台建设
内蒙古	4	宏观指导、具体应用、优惠扶持
上海	5	宏观指导
天津	1	具体应用
重庆	6	宏观指导、具体应用、产业规划

表 1 中，从政策数量来看，首个批复的贵州试验区政策占有所有试验区出台数量的近三分之一，第二批批复的试验区中，除了广东的数量较多外，其他的试验区只有少量政策出台。在政策类别方面，贵州试验区已经脱离了基础的指导意见阶段，已经开始就大数据产业发展相关的人才建设、发展应用以及各种扶持优惠等基础支撑领域出台了一揽子政策；广东、辽宁、内蒙古和重庆也开始由宏观指导向具体操作转变；而北京、河北、河南以及上海在政策制定方面尚处于起步阶段。特别指出的是，天津作为大数据试验区，仅仅出台了 1 项关于推进大数据在医疗卫生领域应用的政策，还未出台大数据发展的宏观规划和实施纲要，发展较为滞后。

二、大数据政策量化分析

（一）政策量化过程

鉴于本文主要的研究重点在于考察各家试验区在大数据领域关注的重点与焦点，以及在大数据领域的发展特色。因此，将主要考察试验区内出台的政策，国家级的政策并不在研究范围内。以各试验区政府官方网站公开的大数据政策为基础，初步得到总计 101 篇政策文本。通过进一步的政策通读，删除其中 12 篇存在信息不完整以及与大数据无关的政策，最终得到 89 篇政策文本作为研究对象。添加若干重要属性，进一步将其形成试验区大数据政策数据库。

根据扎根理论，首先对所有 89 篇政策文本进行反复研读，寻找其间的意义流动，基于此提取出了一系列的初始编码。即每篇政策中提取 3—5 个能够概括政策核心内容的长短句，再从每个长短句中提取 1—2 个能够概括长短句的词组，共获取词组 623 个。其次，根据 TF-IDF 统计算法，测算出每一个词组在每一篇政策中的权重 Z_{ik} ，以 0.6 为权重值设置值域阀，选取权重值最高的前 92 个关键词作为代表试验区大数据政策的关注热点。再次，由于不同政策文本的初始编码可能存在着同意异形的情况，因此将其按意义进行归并，最终提取出 68 个核心关键词。最后，结合核心关键

词在每一篇政策中的权重，将关键词两两配对，以关键词为结点形成共词网络，借鉴 Ochiai 系数概念，共词网络中的对应系数 $W(j,k) = \frac{\sum_i \sqrt{Z_{ik}Z_{ij}}}{\sqrt{\sum_i Z_{ik}} \sqrt{\sum_i Z_{ij}}}$ 。

通过运用 Excel 和 MATLAB，即可得出试验区大数据政策核心关键词的共词矩阵。该矩阵以 68 个关键词为结点，以关键词间的加权值为权重，矩阵中的数字表示两词之间的相关程度。表 2 为共词网络矩阵的部分数据。

表 2： 大数据试验区出台政策核心关键词共现矩阵（部分）

	大数据	资源共享	开放利用	数据采集	标准化建设	政务应用	统筹治理
大数据	1	0.589	0.42788	0.24663	0.21766	0.30386	0.10835
资源共享	0.589	1	0.53135	0.30532	0.25031	0.66142	0.30747
开放利用	0.42788	0.53135	1	0.35967	0.191458	0.5460	0.34211
数据采集	0.24663	0.30532	0.35967	1	0.288754	0.36174	0.16122
标准化建设	0.21766	0.25031	0.19145	0.28875	1	0.15301	0.20428
政务应用	0.30386	0.66142	0.54600	0.36174	0.153015	1	0.28670
统筹治理	0.10835	0.30747	0.34211	0.16122	0.204284	0.28670	1

（二）数据可视化结果与分析

社会网络分析是对社会行动者之间的互动关系研究为基础的一种社会研究方法。通过使用社会网络分析软件 Ucinet 和其附属的制图软件 Netdraw，基于核心关键词共现矩阵绘制出国家级大数据综合试验区大数据政策的社会网络图谱。具体如下图所示：

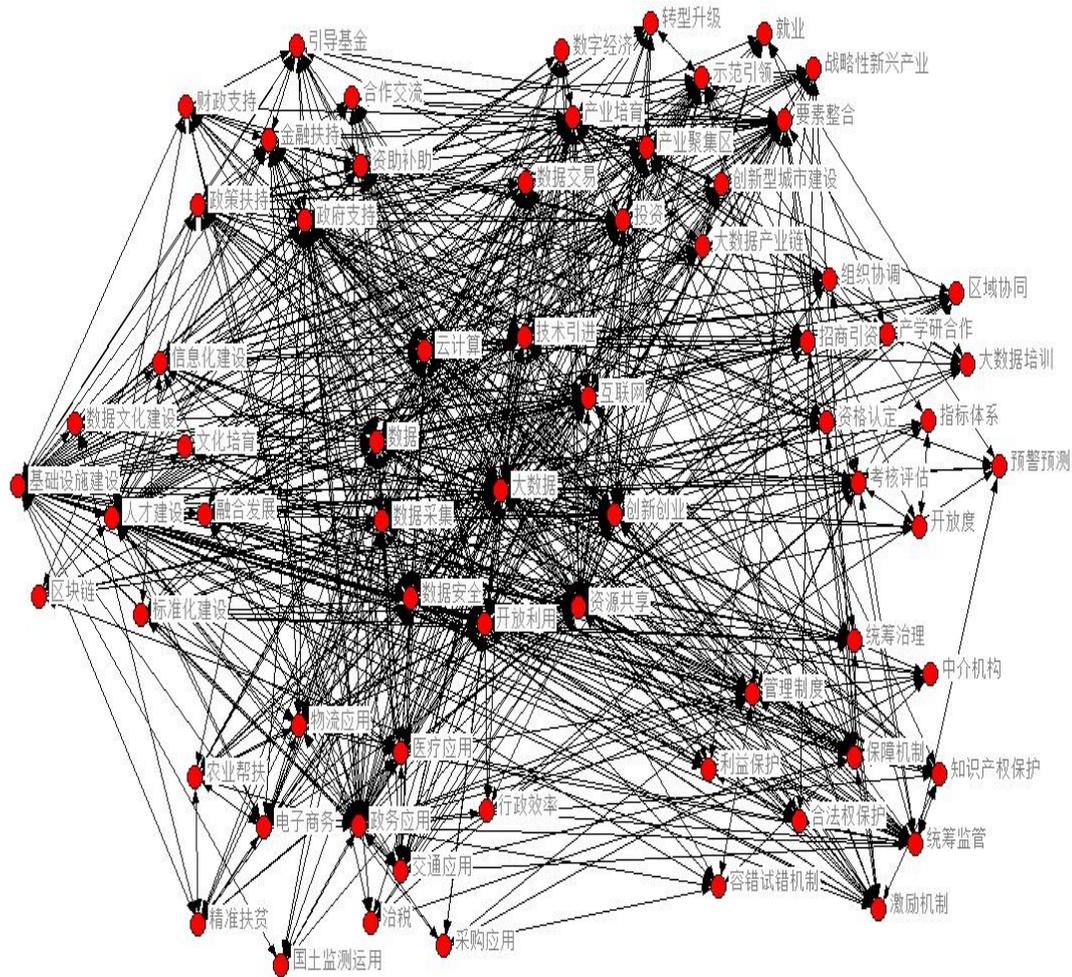


图 1：试验区大数据政策核心关键词共词网络图

通过对核心关键词进行主观的再次判别和划分主题类别，将图 1 中的大数据政策网络的概貌概括为“一个核心，六大主题”。一个核心是围绕大数据形成的以数据采集、数据安全、云计算、技术引进、互联网等 9 大关键词敏感地带，六大主题依据关键词的内涵和权重的高低，从核心敏感地带左上方开始逆时针方向依次为（1）外部支撑；（2）基本要素建设；（3）大数据的应用领域；（4）制度保障建设；（5）促进举措；（6）产业集聚。下文将结合试验区的大数据政策内容，对这六大热点主题分别做进一步讨论。

（1）外部支撑。试验区发展大数据产业离不开必要的外部支持，包括政府支持、金融扶持、财政补助、引导基金以及必要的外部交流与合作。这其中，当前试验区大数据政策主要关注的焦点在于寻求政府的支持、金融扶持、以及必要的财政资助补助，在泛金融领域包括引导基金的设立等

方面还存在一定的欠缺。因此，对于试验区而言，一方面大数据产业的发展离不开资金的支持，需要进一步扩大金融的扶持力度；另一方面，在推动融资发展的过程中应该要进一步强化引导基金、融资担保等在促进大数据产业发展过程中的重要作用。

(2) 基本要素建设。图 1 中，当前制定的政策中关于大数据的基本要素建设，除了必要的基础设施外，尤其强调对于大数据人才的建设。大数据行业的特殊性使得其对人才的要求，特别是高层次人才的需要往往高于其他行业。在试验区的发展中，多数试验区都提出了要强化人才的培育与引进，打造大数据人才集聚高地。同时，与已有的产业示范区不同，大数据试验区还比较重视对大数据文化的培育与建设工作，究其原因，笔者认为，大数据产业的发展打破了已有的社会认知和传统的治理模式，特别是在政务管理领域，对现有的政府管理模式和组织架构产生了巨大地冲击。因此，对于发展大数据产业来说，有必要在全社会形成数据化思维与行为范式，为大数据发展提供思想保障。

(3) 大数据的应用领域。众所周知，大数据以其客观性、大样本和预测性强等特点，能够应用于社会发展的各个领域。由于我国的大数据研究起步晚，大数据的应用尚处于起步阶段。目前试验区在税收治理、农业扶贫、医疗、政务活动、交通等大数据应用较为广泛的领域展开了先行先试的探索，就相关内容出台了一系列的政策法规保障数据应用的顺利推进，并取得了一定的成效。但是，在大数据应用的深层次领域，如跨行业的数据交叉等方面还有待提升。

(4) 制度保障建设。新兴产业的发展离不开必要的制度支撑，大数据行业的产权界定不清晰、隐私保护等方面的困境导致数据的使用、交易都存在制度障碍。因此，试验区制定了相关的利益保护机制和激励机制，并针对试验区的使命，建立了容错机制，建构一整套完善的制度保障体系，保障数据的应用和交易的顺利进行。

(5) 促进举措。该类主题中产学研合作、考核评估、招商引资等核心关键词尤其受试验区政策关注。对于多数试验区来说，大数据的发展基础较为薄弱，想要在快速提升产业的发展水平，一方面需要借助外在力量，通过招商引资的方式，引进行业内具有一定竞争力的机构，快速与先进水

平接轨；另一方面需要提升内在实力，积极拓展产学研合作，促进大数据科技成果转化和技术水平提升。此外，适当的考核评估也能够保证大数据产业发展方向和目标不发生偏转。

(6) 产业集聚。多数试验区在政策中明确提出了要建立大数据产业园、特色小镇、示范区等产业集聚区，从而实现产业升级转型和经济水平提升，并最终实现创新型城市建设的目的。同时，当前试验区的政策也希望通过要素整合的方式，建构完整的大数据产业链，实现产业从实体经济向数字经济转变。另外，在政策研究中发现，试验区的产业集聚政策侧重于产业的集聚效应和示范引领作用，在产业与现有发展基础的结合，核心企业的培育等方面亟待加强。

免责声明

兴泰智库成立于 2016 年 8 月，是由合肥兴泰金融控股（集团）有限公司（以下简称“公司”）发起，并联合有关政府部门、高等院校、研究机构、金融机构和专家学者组成的，以“汇聚高端智慧，服务地方金融”为宗旨的非营利性、非法人学术团体。

《兴泰智库研究报告》是兴泰智库自主研究成果的输出平台，内容以宏观报告、政策解读、行业观察、专题研究为主，对内交流学习，对外寻求发声，致力于为合肥区域金融中心建设提供最贴近市场前沿的前瞻性、储备性、战略性智力支持。

本报告基于兴泰智库研究人员认为可信的公开资料或实地调研资料，但对这些信息的准确性和完整性不作任何保证。报告中的信息或所表达意见仅代表研究人员观点，不构成对任何人的投资建议，公司不对本报告任何运用产生的结果负责。

本报告版权归公司所有，未经事先书面授权，任何人不得对本报告进行任何形式发布、复制。如引用、转载、刊发，需注明出处，且不得对本报告进行有悖原意的删节和修改。