



CDAC

中国数据分析行业
年度发展报告

二零一四年度

前 言

中国商业联合会数据分析专业委员会，简称中商联数据分析委（英文译名：China Data Analysis Committee, China General Chamber of Commerce，缩写 CDAC，以下简称“协会”），是以项目数据分析师事务所等为主体，以从事项目数据分析相关的企业决策分析、企业经营数据分析、项目数据分析和大数据分析挖掘等业务方向的科研院所、大专院校、经营性企业、服务性企业及相关团体与个人自愿组成的全国性行业组织。2008 年 4 月 8 日经国务院国有资产监督管理委员会审核同意、中华人民共和国民政部批准和登记，正式成立的中国数据分析行业唯一的行业协会。

2014 年 11-12 月期间，协会对全国数据分析行业进行了一次全国范围的调研。根据调研情况，协会编写了《2014 年度中国数据分析行业年度发展报告》（以下简称“报告”）。这是协会连续五年对国内数据分析行业进行调研并发布的行业白皮书。目的是希望通过分析行业发展的现状和趋势，为数据分析行业制定发展战略提供咨询意见，为事务所开展工作和拓展业务提供思路，为国家政府主管部门的监管提供参考意见。

本报告主要分成以下四个部分：

第一部分：2014 年中国数据分析行业发展现状

第二部分：2014 年中国数据分析行业发展的特点与问题

第三部分：中国数据分析行业未来发展进阶

第四部分：数据分析专业委员会 2015 年工作规划

在编写本报告时，我们主要依据了协会在全国范围内通过问卷调查、访谈、事务所实际运营数据等调研方法所取得的数据，结合公开发布的数据和文献研究得出结论。本报告仅作为行业内使用，不代表任何国家管理部门观点和意见。由于编写时间紧迫，本报告难免有疏漏、错误之处，敬请业内同仁、广大读者提出宝贵意见和建议。

序 言

继“大数据元年”之后，大数据这把“金钥匙”，在 2014 年正在逐渐打开其商业价值的“潘多拉魔盒”。一股变革与颠覆的力量正在逐渐释放出来，将对整个社会带来深远的影响。

大数据浪潮的到来，大大推动了数据分析理念的传播，体现在社会各界对数据分析的关注度越来越高，在各种载体的媒体上曝光度也越来越高，促进了数据分析行业的发展。

同时我们也注意到一些大数据认识上的误区。这种误区集中体现在一谈起大数据就把大数据当作是 IT。多名知名学者和院士都指出，大数据的核心价值在于分析，大数据时代的核心在于数据分析能力的提高。之所以产生这种误区，是因为大数据的发展阶段还处于初期阶段，类似于互联网刚兴起的阶段，也就是类比于“web 1.0”的“大数据 1.0”时代。随着大数据的发展，会逐渐回归到大数据的价值本位，也就是大数据的分析上来。正如互联网的发展是逐渐回归到商业上来，互联网最核心的价值是商业的组织形式，以互联网的方式组织商业形式，如电子商务、互联网金融等，这才是其价值本位。

随着大数据技术的普及，大数据必将逐渐回归到其本位价值——数据分析的价值。而基于大数据分析，将能改变现有的社会和商业的组织形式，这是继互联网之后又一股变革的力量。

目 录

第一部分 2014 年中国数据分析行业发展现状	1
一、数据分析行业从业机构情况	1
1、项目数据分析师事务所发展情况	1
a、事务所地域分布情况	1
b、事务所团队建设情况	2
c、事务所宣传推广情况	6
d、事务所业务情况	6
e、事务所自身认识情况	12
f、典型事务所情况	14
2、技术类混业公司情况	17
a、百度大数据	18
b、阿里大数据	19
c、腾讯大数据	19
3、调研类混业公司情况	20
a、互联网发展促使数据采集方式发生转变	20
b、大数据的发展使得市场调研业务逐渐萎缩	21
二、数据分析行业建设情况	21
1、提出“技术+咨询”的发展战略	21
2、大数据智能分析平台“Datahoop”的研发	22
3、项目数据分析师培养课程改革	23
第二部分 2014 年中国数据分析行业发展的特点与问题	26
一、2014 年中国数据分析行业发展的特点	26
1、数据价值逐渐被认可但仍需进一步推广普及	26
2、数据技术在新兴技术的推动下逐渐走向成熟	26
3、数据生态系统逐渐丰富并影响企业商业模式	27
4、数据分析加速向传统领域拓展但仍处于初期	28
5、我国大数据发展战略顶层设计开始受到重视	28
6、数据共享时代各种区域大数据联盟不断涌现	29
二、2014 年中国数据分析行业发展的的问题	30

1、专业从业机构存在的问题.....	30
2、整体面临的问题.....	31
第三部分 中国数据分析行业未来发展进阶.....	34
一、中国数据分析行业发展的全球背景.....	34
1、数据逐渐成为具备重要价值的资产.....	34
2、大数据在更多的传统行业得以应用.....	34
3、行业定制化大数据解决方案将涌现.....	34
4、数据开放的步伐将进一步加快推进.....	34
5、大数据安全问题将会越来越受重视.....	35
6、大数据将成为智慧城市的发展引擎.....	35
7、大数据将会催生一批新的工作岗位.....	35
8、大数据将在多方面改善我们的生活.....	36
二、中国数据分析行业发展的国内环境.....	36
1、北京迎来大数据发展机遇.....	36
2、上海聚力发展大数据产业.....	37
3、广东打造大数据产业基地.....	37
4、贵州大数据产业亮点频闪.....	38
5、福建多措施促大数据发展.....	38
6、重庆挖掘大数据产业商机.....	38
三、2015 年中国数据分析行业发展趋势.....	39
1、结合智能计算的大数据分析成为热点.....	39
2、数据科学逐渐兴起并带动多学科融合.....	39
3、跨学科交叉的数据融合分析成为趋势.....	39
4、与热点技术的融合将产生综合性应用.....	40
5、一体化融合的数据处理平台渐成趋势.....	40
6、大数据安全和隐私继续成为关注重点.....	40
7、可视化技术和工具提升数据分析水平.....	40
8、数据技术标准和人才培养需高度关注.....	40
第四部分 数据分析专业委员会 2015 年工作规划.....	42
一、加大服务力度.....	42
1、进一步加大事务所业务转型扶持力度.....	42
2、有针对性的举办数据分析专业论坛，有针对性的解决地区问题.....	42
3、加强对会员单位和个人的执业教育培训，进一步提升研究水平.....	42
4、完善 CPDA 信息系统的信息服务，加大对会员单位的数据服务工作.....	42

5、大力扶持重点项目数据分析师事务所，起到以点带面的作用.....	42
6、面向企业，提供专业的数据分析咨询和培训服务.....	42
7、加大全国各地项目数据分析师事务所的走访.....	43
二、严格行业监管.....	43
1、加强 2015 年的各项年检工作.....	43
2、举办优秀项目数据分析师事务所评选.....	43
3、加强对会员单位出具的数据分析报告质量的检查.....	43
4、加强会员入会审批流程的质量监督.....	43
5、加强对个人和团体从业会员行为的监督.....	43
6、严格监督和查处个人及团体会员的违规现象.....	43
三、进一步做好行业推广工作.....	44
1、编写并发布行业发展白皮书.....	44
2、与媒体合作建立行业专栏，不定期发布行业新闻和资讯.....	44
3、有针对性地对相关地区进行推广.....	44
4、继续加强项目数据分析师人才培养的推广.....	44

第一部分 2014 年中国数据分析行业发展现状

在中国,项目数据分析师事务所是中国数据分析行业唯一被认可从事数据分析服务的专业机构。除了此类型专业机构,还有混业经营的机构,如市场调研公司、IT 公司、网络公司和系统集成公司等商业机构,这些商业机构的特点是数据分析是其重要的业务组成部分,与其他业务有机组合,混业经营,按照他们的业务特点分为调研类和技术类公司。

本部分研究以数据分析行业的专业机构为主,附带讨论混业经营的咨询类和技术类公司。

一、数据分析行业从业机构情况

1、项目数据分析师事务所发展情况

项目数据分析师事务所是我国从事数据分析服务的专业机构,接受中国商业联合会数据分析专业委员会的监管。项目数据分析师事务所的设立需经中商联数据分析委的严格审批,在运营接受中商联数据分析委的监督和检查。

截止 2014 年底,中商联数据分析委共收到设立项目数据分析师事务所的有效申请 558 份,经考察和资质审核共有 54 家项目数据分析师事务所通过审批,审批通过率不足 1/10。审批通过率不高的部分原因是因为中商联数据分析委“保质控量,宁缺毋滥”的审批原则,另外一方面因为市场被“大数据浪潮”催热,有很多盲目追热,导致申请数量的激增。

a、事务所地域分布情况

经审批通过的事务所遍布全国超过 20 个省市自治区,数量较多的区域主要有北京、陕西、山东、安徽和重庆,其中北京地区的事务所数量将近占全国事务所总数的三分之一,其余地区事务所数量相对较少。

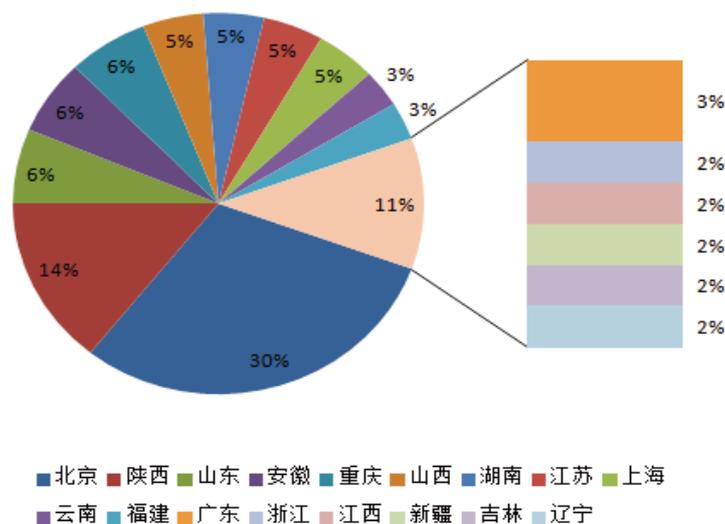


图 1-1 事务所以北京最为集中

b、事务所团队建设情况

(1) 经营规模

从调研数据来看，目前的项目数据分析师事务所规模普遍偏小，人数最多的一家事务所为 45 人（但其兼职人员占到一半），人数低于 12 人的事务所数量超过总调研事务所数量的一半，占比达 52%，且多为 9-12 人。

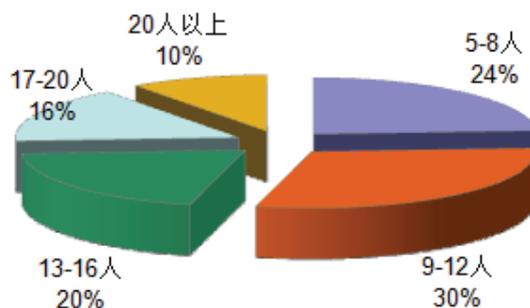


图 1-2 事务所普遍规模偏小

(2) 人员学历构成

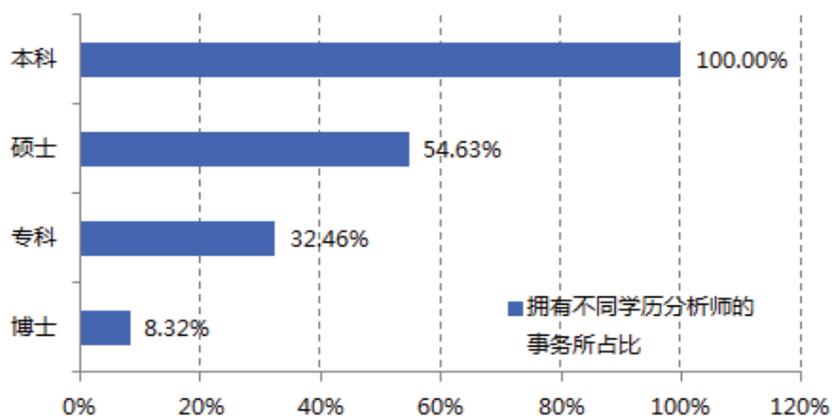


图 1-3 所有的事务所都有本科学历的分析师

各事务所的项目数据分析师的学历主要由专科、本科、硕士和博士构成，由调研数据可看出，每个事务所都有本科学历的分析师，有 32.46% 的事务所有专科学历的分析师，有 54.63% 的事务所有硕士学历的分析师，有 8.32% 的事务所有博士学历的分析师。博士硕士学历的分析师比去年增加 11.78%，事务所整体人员水平较以前提高很多。

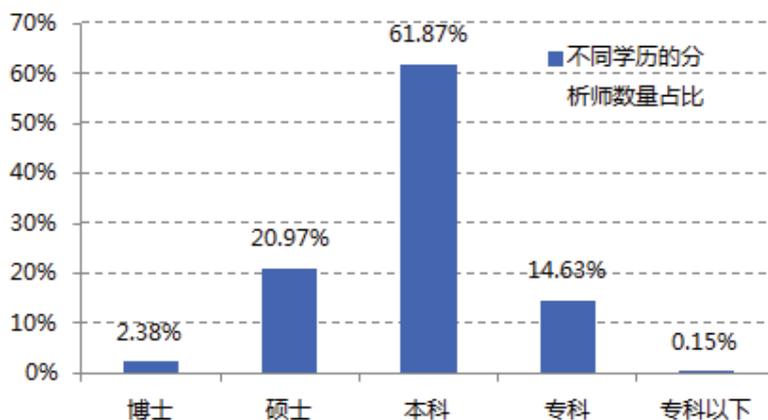


图 1-4 从业项目数据分析的人员学历以本科为主

(3) 人员从业年限

大部分事务所都有从业 1-5 年的分析师，各事务所中有从业 3-5 年的分析师的事务所占比达 79.96%，各事务所中有从业 1-3 年的分析师的事务所占比达 76.42%，同时也有 17.58% 的事务所拥有还未从事项目数据分析工作的刚拿到证书的分析师。

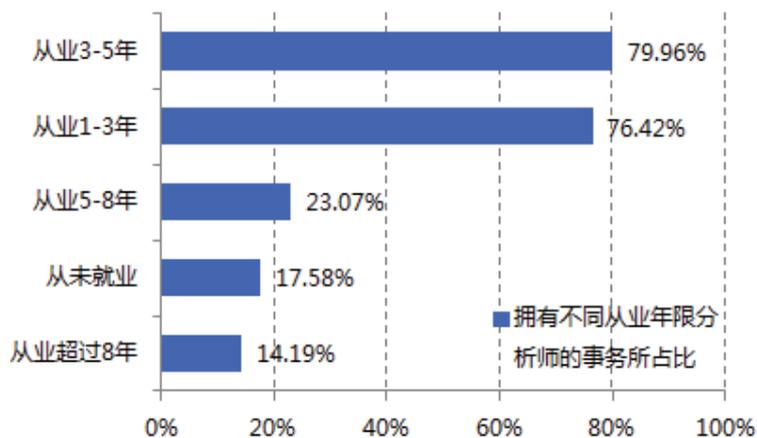


图 1-5 大部分事务所都有从业 1-5 年的分析师

从就业人数的分布来看，目前的项目数据分析师的从业年限主要集中在 1-5 年（占比 87%），其他就业年限的分析师数量较少，目前的项目数据分析师从业队伍相对年轻。

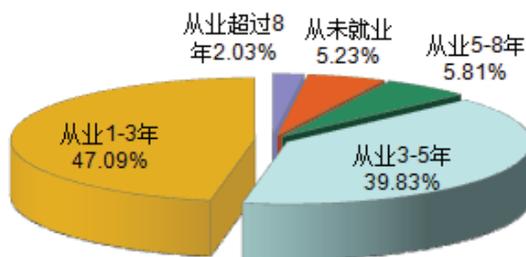


图 1-6 从业人数中以从业 1-5 年的分析师为最多

(4) 人员专业背景

从事务所分析师的专业情况看，拥有财务、税务相关专业分析师的事务所占比最大，高达 85.22%；而有统计、数学相关专业分析师的事务所占比 69.83%；有工商管理、企业管理相关专业分析师的事务所占比为 56.05%；拥有计算机相关专业分析师的事务所占比为 38.40%；拥有经济贸易相关专业分析师的事务所占比为 31.44%。

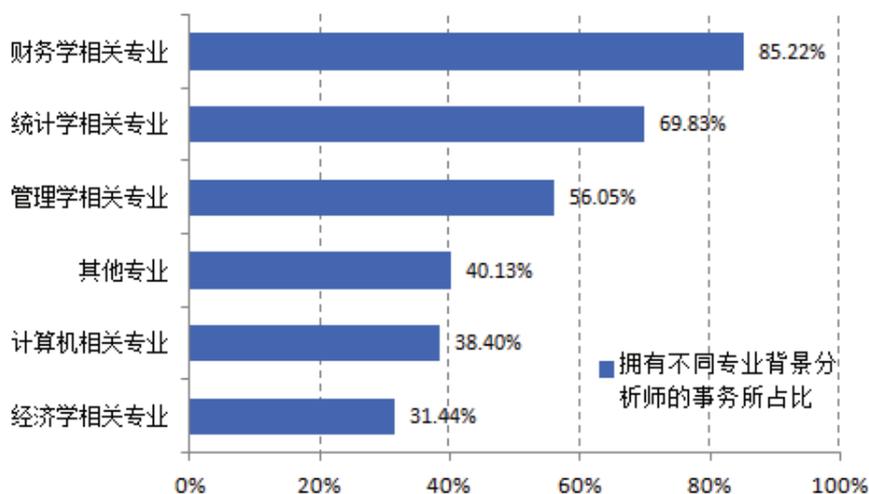


图 1-7 超过 85% 的事务所拥有财务、税务相关专业的分析师

从从业人数来看，有财务税务相关专业的人数占比为 30.79%，有统计、数学相关专业的人数占比为 19.49%，而计算机相关专业的人数从之前几乎没有到去年增加了 14.69%，其余类相关专业的人数占比相对较低。

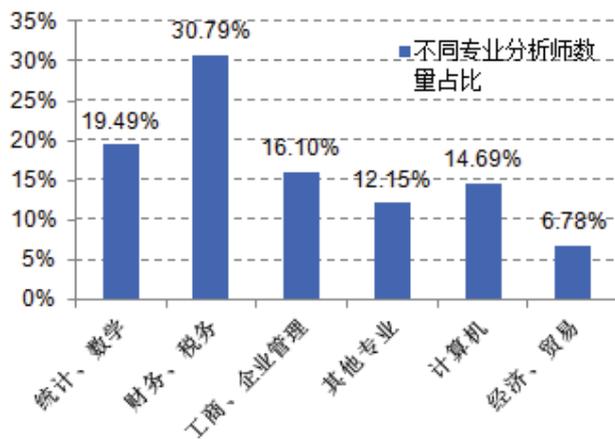


图 1-8 从业分析师中财务、税务相关专业背景人数最多

(5) 人员行业背景

从事务所分析师的从业背景来看，大部分事务所都有财务注会、咨询行业以及企业高管的背景，拥有这些从业背景分析师的事务所比例分别为 66.67%、56.86%和 49.02%，同时也有一部分事务所拥有税务审计、风险评估背景、计算机背景的分析师，还有少部分事务所拥有一些其他行业背景的

分析师。

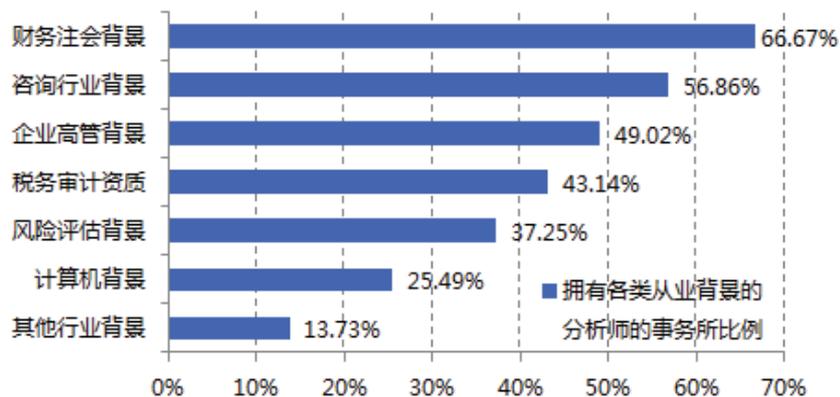


图 1-9 大部分事务所都拥有财务、咨询行业背景的分析师

从从业人员构成背景看，事务所以有咨询行业背景的人员为最多，占比为 27.47%，财务、注会资质的人数次之，占比为 21.60%，拥有企业高管背景的人员占比为 14.81%。

其他从业背景中有电商行业背景、通信行业数据分析背景、工程造价分析背景、计算机行业背景等，占比不足 10%。



图 1-10 拥有咨询行业背景的分析师数量最多

(6) 人员综合素质

从项目数据分析师所拥有的技能证书看，大部分的分析师只有 3 种以下技能证书。

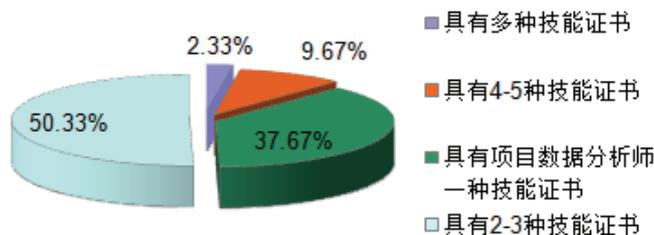


图 1-11 拥有 2-3 种技能证书的分析师占比最大

c、事务所宣传推广情况

(1) 与协会网站的链接情况

有 42.22% 的事务所网站、微博、微信均与协会网站、微博、微信做了链接；有 29.25% 的事务所网站与协会网站做了链接，但微博、微信没有做链接；有 19.06% 的事务所网站、微博、微信均没有与协会连接；且有 6.98% 的事务所仍然没有自己的网站、微博和微信。

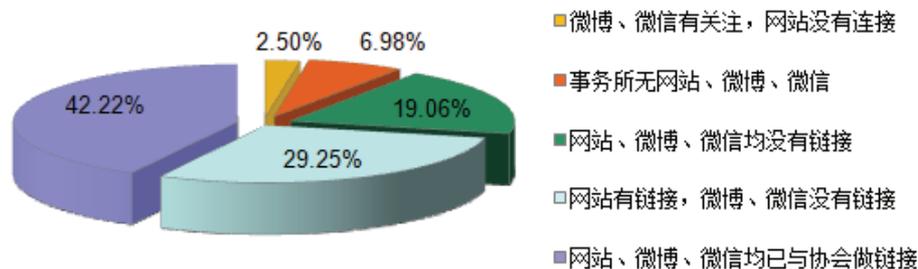


图 1-12 只有不到 70% 的事务所与协会网站有链接

(2) 事务所自身宣传推广

目前几乎所有的事务所都在通过一些相关渠道和方法进行宣传和推广，完全没做宣传的企业仅占 2.33%，在所有的宣传渠道中，通过一些有针对性的宣传方式进行宣传是事务所普遍的做法，如直接选定一些特定的企业做宣传，有 58.14% 的事务所通过这种方式做过宣传，另外有 53.49% 的事务所通过与银行、投融资机构合作进行宣传，有 51.16% 的事务所与政府公关进行过合作，这种针对性的宣传方式效率高、宣传费用相对较低，是目前事务所主要的宣传方式。

另外一部分事务所采取了广泛撒网的推广方式如将宣传业务外包给专业的网络公司，参加地区会议，做宣传册等方式。

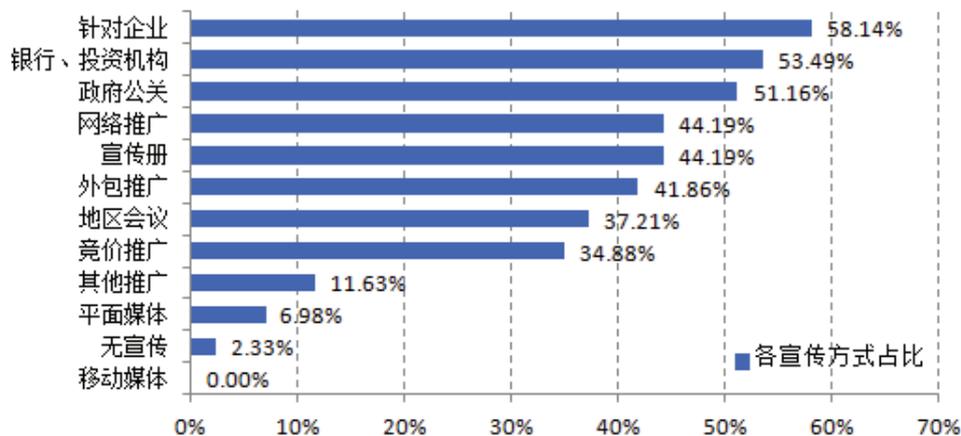


图 1-13 比较有针对性的宣传方式是事务所的首选

d、事务所业务情况

(1) 业务类型

各事务所业务定位以企业经营决策为主，兼顾其他业务的事务所占比为 36.51%；以投融资业务为主，兼顾可研、评估类业务的事务所占比为 14.29%；以投融资业务为主，兼顾企业经营决策

业务的事务所占比为 12.70%；以可行性研究报告业务为主的事务所占比为 11.11%；只做过银行信贷评估、商业计划书的事务所占比为 4.76%；只做投融资业务的事务所占比 3.17%，只做企业经营决策类业务的事务所占比 1.59%。

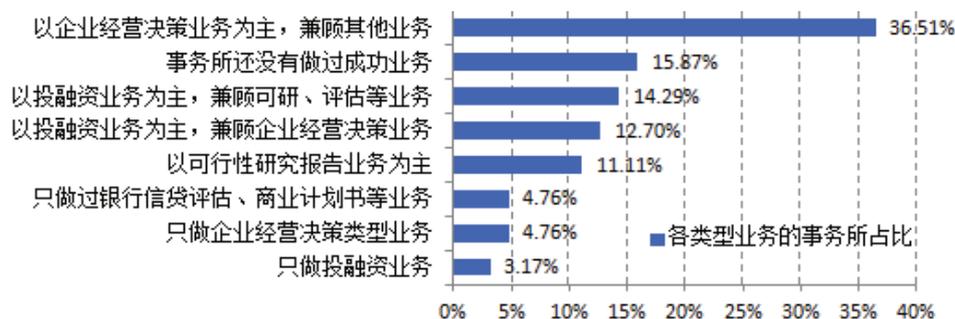


图 1-14 以企业经营决策业务为主、兼顾其他业务的事务所占比最高

(2) 业务资质情况

事务所中拥有其他业务资质的占比都不高，其中 26.98% 的事务所有大使馆认证，22.22% 的事务所有注会类业务资质，19.05% 的事务所有审计类业务资质，17.46% 的事务所有评估类业务资质，14.29% 的事务所有税务类业务资质，9.52% 的事物所有工程咨询类业务资质，还有 11.11% 的事务所有其他类业务的资质。

涉外资质认证的国家中几乎都有美国，其次几个认证较多的国家包括英国、法国、德国，还有一些事务所拥有瑞士、荷兰、澳洲、俄罗斯、香港等国家和地区的认证。

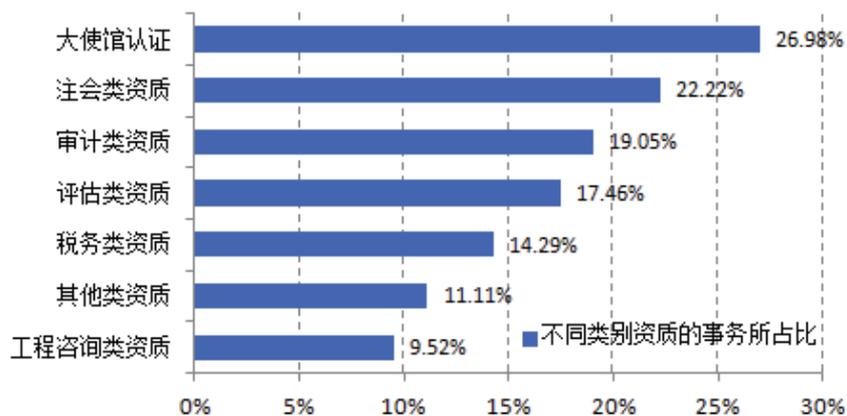


图 1-15 拥有各类其他资质的事务所均占比不高

只有项目数据分析类业务资质而没有其他业务资质的事务所占比为 28.57%，只有一种其他资质的事务所占比 36.51%，有两种其他资质的事务所占比为 15.87%，超过两种其他业务资质的事务所占比为 19.05%。总的来说，事务所拥有的业务资质相对单一。

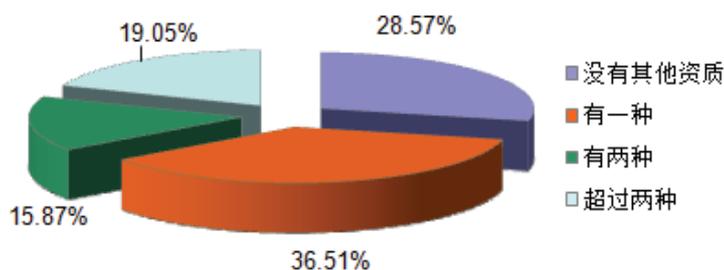


图 1-16 事务所拥有的业务资质相对单一

(3) 新增行业情况

2014 年各事务所新增的业务中，涉及的行业主要有房地产行业、制造业和农林牧渔业。建筑业和电子商务业也有一定业务，其他行业的业务相对较少。

新增行业	新增数量
房地产业	18
制造业	14
农、林、牧、渔业	12
建筑业	12
电子商务	11
信息传输、计算机和软件业	10
文化、体育和娱乐业	9
金融业	8
采矿业	6
教育	6
批发和零售业	5
科学研究、技术服务和地质勘查业	4
医疗卫生、社会福利业	4
水利、环境和公共设施管理业	4
交通运输、仓储和邮政业	3
公共管理与社会组织	3
住宿和餐饮业	3
居民服务和其他服务业	3
电力、燃气及水的生产和供应业	2
租赁和商务服务业	2
其他行业	2
国际组织	0

表 1-1 事务所涉及的房地产业、制造业的业务相对较多

(4) 业务获得方式

从调研结果看，事务所获得业务的来源主要通过网络宣传推广和与政府合作，通过这两种方式获得业务的事务所数量占比分别为 72.09%和 43.22%。

通过金融、投融资公司推荐，获得业务来源的事务所数量占比为 34.88%。通过其他方式获得业务的事务所数量占比相对较小。

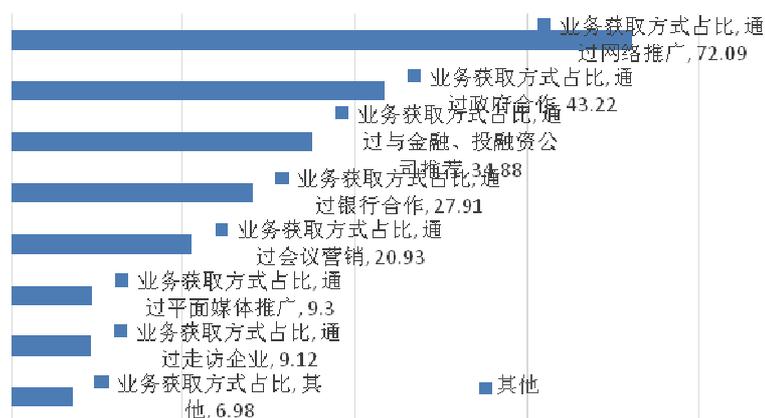


图 1-17 网络宣传是事务所获取业务的主要来源

(5) 数据来源渠道

从数据来源渠道来看，网络渠道和企业提供两种渠道的占比最高，需要自己从网络上搜索数据的事务所占比高达 61.90%，直接从企业获取数据的事务所占比为 55.56%，其他的主要来源渠道有亲自做市场调查（占比 33.33%），从统计局、税务等官方购买（占比 30.16%），从行业协会获得（占比 23.81%），从图书馆、文献杂志等渠道获得（占比 22.22%）。

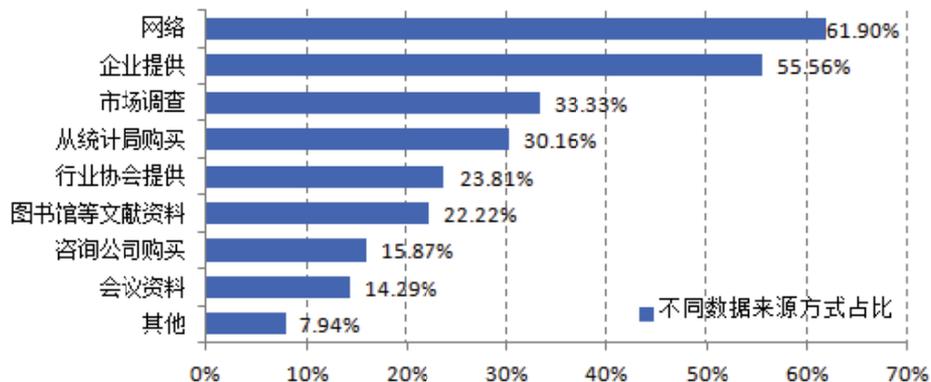


图 1-18 企业提供和网络查找是事务所获取数据的主要方式

(6) 合作企业情况

目前各事务所达成经营决策类的合作企业仍然不多，暂时没有经营类项目合作企业的事务所占比为 19.05%；有 1-2 家合作企业的事务所占比为 15.87%；有 3-6 家合作企业的事务所占比为 47.62%；有 6 家以上合作企业的事务所数量占比 17.46%。

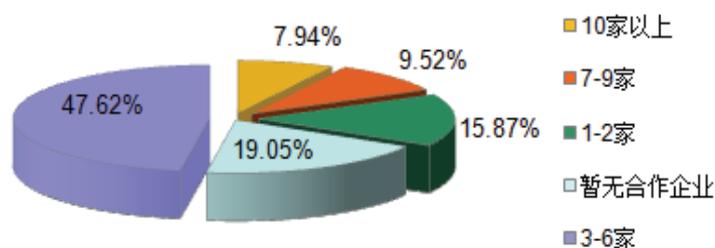


图 1-19 目前有超过一半的事务所没有经营业务类的合作企业

(7) 项目合作时间

从已签订的项目来看，目前大部分项目的合作时间是 1-6 个月，占比为 49.02%，6-12 个月的项目占比为 23.53%，1-3 年的项目占比为 17.65%，签订 3 年以上合作的项目数量占比为 9.80%。同去年相比，长期项目数量增长百分之三左右。

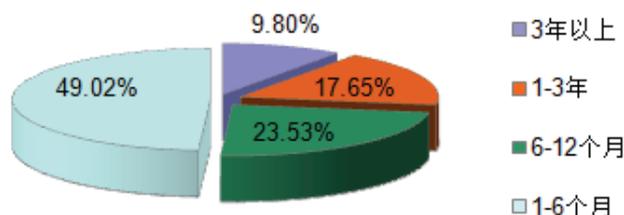


图 1-20 目前与事务所合作项目的的时间普遍较短

(8) 合作企业类型

从事务所合作过的企业类型来看，目前事务所主要合作的企业类型仍然是国内中小型私企，有 55.56% 的事务所与中小型私企进行过合作；有 44.44% 的事务所与股份制企业进行过合作；与政府、部委等机构有合作的事务所占比为 31.75%；与国有企业和国内大型私企有过合作的事务所占比例约为 28.57%；与外资企业有过合作的企业占比为 14.29%。

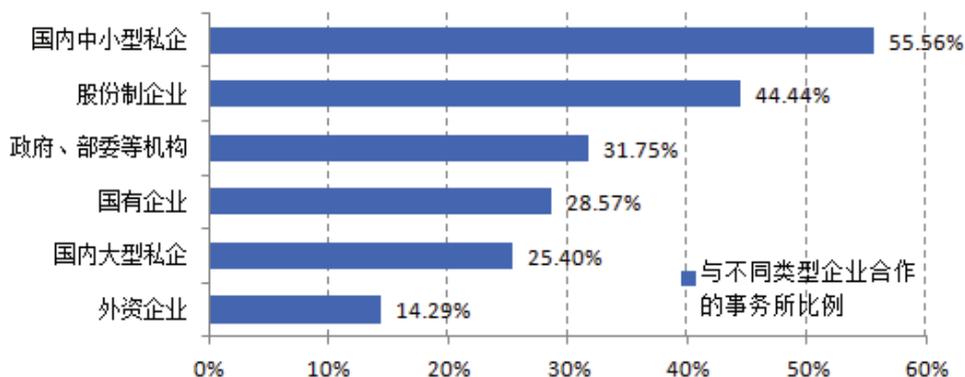


图 1-21 目前与事务所合作的企业主要是中小型私企

(9) 服务项目类型

从事务所已经提供过服务的项目类型来看，市场进入类分析和竞争分析类的项目数量占比最大，

精准营销类、产品定价类、客户满意度以及用户使用习惯及态度分析类项目相对较少，品牌分析类和广告效果评估类项目更少，目前还没有事务所从事过产品测试类和渠道暗访类的项目。

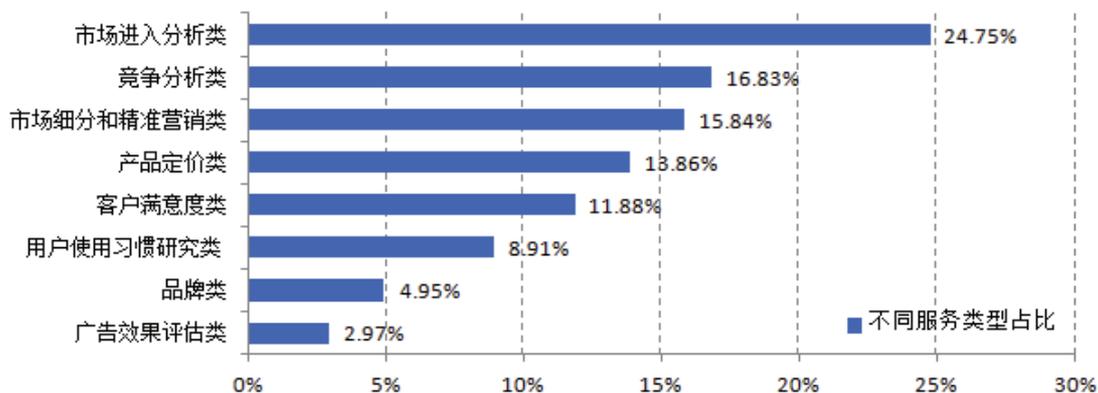


图 1-22 不同服务类型的项目数量差距不大

(10) 经营分析数量

目前有 34.85% 的事务所尚未出具经营决策类的报告，完成过 1-2 份该类型报告的事务所占比 15.15%，完成过 3-6 份报告的事务所占比 30.30%，完成超过 6 份报告的事务所占比 30.23%。

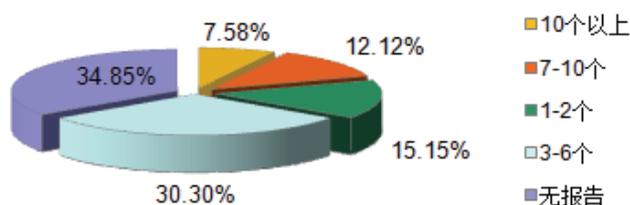


图 1-23 有超过 35% 的事务所没有出具经营决策类的服务报告

(11) 经营分析内容

目前事务所涉及的经营决策服务类的内容主要集中在基础的数据预测分析、数据录入整理和数据的分类与收集上，提供过该种内容服务的事务所占比均为 31.75%；其他类型的一些内容服务相对较少。

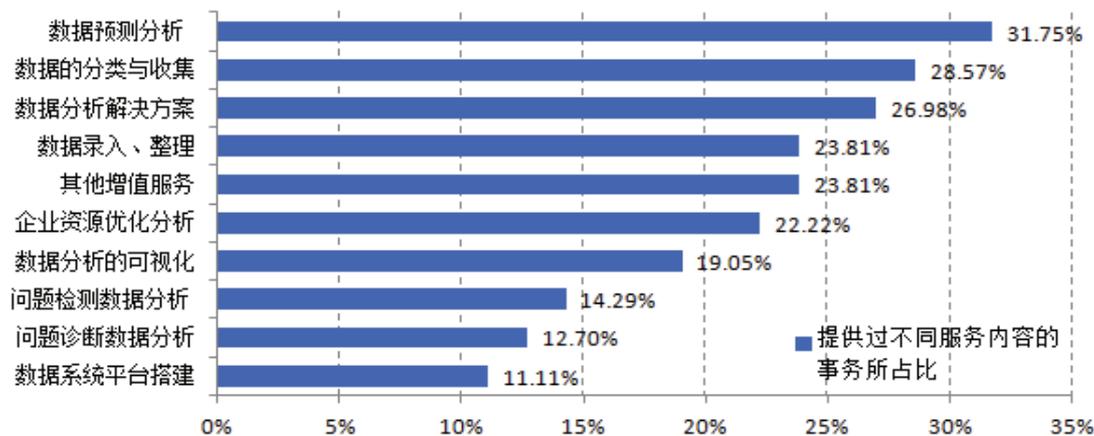


图 1-24 事务所的服务内容仍然比较基础

e、事务所自身认识情况

(1) 事务所自身对优势的定位

从事务所的优势列表可以看出，事务所的优势主要集中在事务所专业性、事务所资源以及事务所的内部管理上。

事务所优势	得分
事务所的专业性（专业团队、数据分析研究）	18
事务所人才涉及范围较广、结构合理	15
业务资质较多，业务范围较广	10
不同的区域性优势	9
数据分析行业的专注性（只做数据分析）	8
规模不大、业务转型灵活	5
已确立企业经营的数据分析方向	4
政府资源较多	3
有比较专注的行业	3
经营管理制度健全	3
市场优势明显	3
注重国内市场宣传	3
长期从事投融资业务，经验较充足	3
有较充足的企业管理经验	2
人员对新事物接受程度较高	2
与当地金融机构有良好的合作关系	2
事务所宣传推广做的比较好	2
自身品牌的可塑性强	1
有专业的网络推广公司做宣传	1
各项证书、资质齐全	1
在建设自己的数据平台	1
组织优势，有工会组织、党支部	1
已形成专业的报告产品	1
业务量、财务较稳定	1
找到对口行业快速打开市场	1

表 1-2 专业性优势是事务所最主要的优势

(2) 事务所自身面临的问题和困难

从事务所的调研情况看，行业认可度低、市场难打开、业务难开展是目前事务所普遍存在的一个问题。

事务所资深的问题和困难	得分
地区企业对数据分析行业认可度不高，业务难开展	14
事务所知名度较小，客户资源非常有限	9
市场开拓比较困难	8
分析师专业水平有待提高	8
目前以投融资业务为主，转型到企业经营较困难	7
人才储备不足、招聘困难	5
数据积累不足、数据收集难度大	5
经营性核心技术缺乏，缺乏系统的培训，业务难开展	4
事务所宣传力度较小、缺乏有效宣传	3
目前事务所还没建立自己的核心竞争力	3
市场竞争激烈、业务开拓难度较大	3
工作形式和内容偏于模板化，人才积累受限	1
事务所刚起步，行业经验较少	1
对于市场扩展人员没有系统的培训方案	1
对经营决策数据分析认知不清	1
事务所决策定位不准	1
事务所与市场接合度不高	1
品牌还有待进一步树立稳固	1

表 1-3 行业认可度不高业务难开展是事务所面临的最主要困难

(3) 事务所对量化经营与投资的认知

从调研问卷的情况看，事务所对经营决策类业务和投融资业务的认识主要体现在以下几点：

	企业经营类业务	投融资类业务
业务价值	经营决策类业务可以更好的提升事务所从业人员的业务能力和事务所的专业水平	量化研究类业务从广义上说应该归属在经营决策类以内，一份高水平的量化研究类的报告，可以为项目方的投资决策和风险规避提供重要的参考
业务特点	经营决策类业务能跟企业建立长期的业务联系，服务时间较长，有利于数据分析服务的价值体现，也有助于事务所业务的稳定	量化（投融资）研究类业务所涉及的领域虽然也很广阔，但量化（投融资）研究类业务比较单一、是短期的、一次性的服务
业务开展	经营决策类业务将会是各事务所未来发展的主要方向，但目前开展经营决策类业务也存在比较客观的难题，市场认知度不高，需要协会及各事务所同心协力共同宣传影响市场的认知度	在经营决策类业务得到较好地发展之前，量化研究类业务不能彻底抛弃
业务趋势	随着数据分析行业的发展，会有更多企业认识到精准决策的重要性，必然会产生海量的企业经营决策业务需求，因此经营决策类业务未来必将达到一个高峰，行业协会目前倡导全国事务所向经营决策类	量化（投融资）研究类业务目前不是很好开展，因为投融资行业，在市场中信誉度不高，很多投融资客户需要数据分析报告时，都附带要求提供其他公司的资质，并

	业务转型具有一定的前瞻性，是正确的引导。	不是为了项目而要数据分析，损坏了数据分析行业的科学性和价值
--	----------------------	-------------------------------

表 1-4 企业经营类业务与投融资业务的比较

(4) 事务所对行业协会建议

从提建议的数量来看，加大行业宣传力度（占比 25%）、建立全国事务所沟通平台（占比 14%）、建立行业规范和标准（占比 9%）和加大行业监管力度（占比 8%）是事务所关注度最高的几项建议。

建议协会加大对会员服务力度的条数也较多，如直接提出希望协会加大对会员的服务力度、完善数据库系统、对分析师进行持续的后续教育、多组织相关专业的培训、进行经营决策类业务培训等。

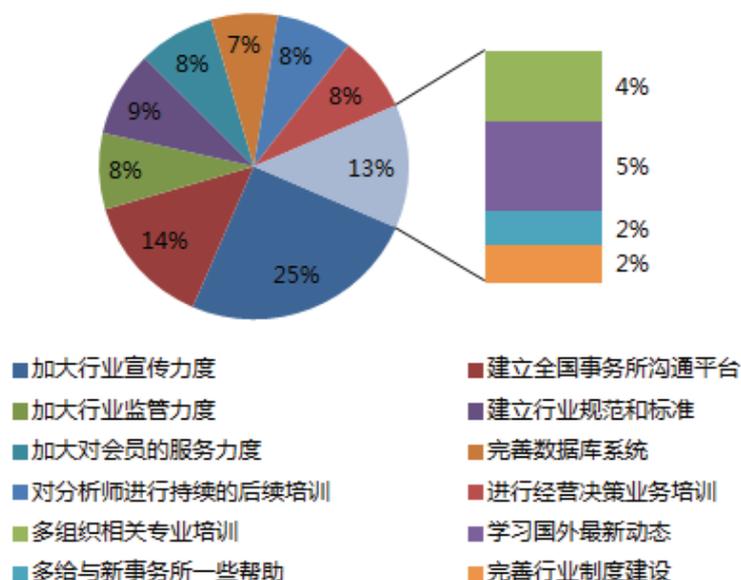


图 1-25 加大行业宣传力度是事务所共性最大的建议

还有一些个性化的建议，如按季度在协会内部开展数据分析论坛，解读行业政策与咨询，分析有代表性的业务案例，协会提供一些资源共享平台，适当提高基层的准入门槛等。

f、典型事务所情况

(1) 安徽量衡项目数据分析师事务所

安徽量衡项目数据分析师事务所有限责任公司是经国家行政管理部门核准注册，经中国商业联合会数据分析专业委员会备案（中数委团证第 067 号）的独立法人机构。事务所现拥有项目数据分析师、注册会计师、注册税务师、物流管理师、工程师等专业技术人员。业务涉及农业、文化、房产、物流、餐饮、科技、烟草、能源、通讯、金融等国民经济领域，以专业的数据分析服务为核心，致力于为政府机构、企业、国内外银行、风投公司等提供项目投资类数据分析、经营类数据分析、项目数据软件、投资项目评价、市场调研、项目经济效益评价、项目策划、投资咨询、经济信息咨询等专业的、系统的服务，并可为委托方提供可行性报告、商业计划书、项目数据分析报告等多种

具有客观性、公正性、实用性的报告。

安徽量衡项目数据分析事务所有限责任公司本着“专业、专注、专心”的职业态度，根据市场需求，从数据分析的角度为客户做好数据的调研、统计、挖掘、分析等服务。

专家点评：

安徽量衡项目数据分析师事务所作为安徽省第一家事务所，自 2011 年底成立起，便积极开拓市场，确立中小企业为服务目标，致力于在安徽省内普及数据分析理念、为企业做好经营决策数据分析服务。

安徽量衡项目数据分析师事务所充分洞察到数据分析行业的新变化，敏捷应变，在大数据分析、新媒体（微信）数据分析、经营分析领域积极探索，成果喜人。

安徽量衡提交的数据分析报告业务逻辑清晰，数据结构合理，数据资料较为详实，有足够的支撑，研究论证相对严谨。报告研究论证相对专业度较高，可以为企业提供客观、系统地了解项目和企业自身的实际情况，帮助企业做出更符合企业长期发展的决策意见。在数据分析深度挖掘上有提升空间。

(2) 中颢润（北京）项目数据分析师事务所

中颢润（北京）项目数据分析师事务所有限公司是经过北京市工商局登记注册的具有独立法人资格的专业项目数据分析机构，且经中国商业联合会数据分析专业委员会备案核准的常务理事单位（中数委团证号 001）。事务所拥有数据采集、数据分析、金融研究、工商管理等方面的复合型人才及专家，聘请大量数据分析业专家作为资深顾问。事务所运用先进的理念和独到的见解，突破行业的数据分析局限，从决策性数理研究入手，为企业提供专业化的分析和科学的决策依据的同时，更精准、全面客观的营销战略成为指导企业规避项目风险、保持强大竞争优势的科学依据。

中颢润拥有包括专业项目数据分析师、资产评估师、注册会计师和专业市场调查人员构成的强大团队，致力于专业数据分析深度服务，坚持投资数据分析报告、经营数据分析报告、决策数据分析报告的权威性、客观性、中立性和实用性，为众多的投融资和企业决策者提供了准确的数据分析成果和决策建议。

专家点评：

中颢润（北京）项目数据分析师事务所作为数据分析行业的常委单位，自成立以来一直坚持以经营决策数据分析业务为主要发展方向，致力于专业数据分析深度服务，坚持投资数据分析报告、经营数据分析报告、决策数据分析报告的权威性、客观性、中立性和实用性。

作为北京地区的项目数据分析师事务所，在面对激烈的市场竞争时，中颢润（北京）项目数据分析师事务所领导人提出要以互联网思维经营事务所，采取微信、微博、线上线下结合等多种新的营销方式，获得了优异的成果，为全国事务所的经营提供了宝贵的经验，且中颢润事务所在大数据分析领域已有一定的储备，专业实力较强。中颢润的数据分析报告思路清晰，结构严谨，分析细致，数据资料详实，方法得当，论证深入，整体质量较高。为企业做出经营决策提供科学依据。虽然在大数据分析领域有了一定储备，但有待进一步拓展大数据分析业务。

(3) 重庆传晟项目数据分析师事务所

重庆传晟项目数据分析师事务所是经过国家行政管理部门核准注册,具有独立法人资格的专业项目数据分析机构,且经中国商业联合会数据分析专业委员会备案核准的事务所会员单位(中数委团证号 084)。重庆传晟项目数据分析师事务所主要致力于房地产政府管理部门全局解决方案、软件开发、咨询服务和档案整理加工,以及房产数据分析统计处理。公司技术实力雄厚,拥有高素质的项目数据分析师团队、软件开发队伍和强大的技术支持力量。主要核心人员从事房地产政府管理部门行业软件开发和咨询工作多年,成功推出了拥有自主知识产权的档案管理系统、数据清理系统、搜索引擎、档案生命周期管理平台、档案整理加工系统、房地产管理部门便捷化服务平台、权证防伪、短信平台等系列产品。公司包括成都和重庆两个研发中心,公司已扎根西南,并不断扩展中。

专家点评:

重庆传晟项目数据分析师事务所一直充分发挥着带动当地项目数据分析行业快速发展的榜样作用,积极拓展业务范围,将企业经营决策数据分析作为事务所发展的长远目标,不断提高自身专业水平和业务能力,获得了当地政府和房地产行业的一致认可。

重庆传晟走的是一条“技术+咨询”的道路,重庆传晟具备较强的信息化技术能力,结合数据分析师事务所得天独厚的数据分析专业实力,这是其独特的核心竞争力。所内从业的数据分析师均具备多年的项目数据分析、软件技术工作经验,秉承以诚为本、客观公正的职业操守,为重庆地区的房地产政府管理部门提供提供专业、全面、详实、精准的项目数据分析报告。重庆传晟撰写的经营类分析报告宏观分析到位,推论严谨,思路较为清晰。公正客观,可以为需求单位提供最佳决策建议依据。事务所通过自身的实力和专业度获得了良好的行业口碑。

(4) 上海天元项目数据分析师事务所

上海天元项目数据分析师事务所是经国家行政管理部门核准注册,经中国商业联合会数据分析专业委员会备案(中数委团证第 065 号)。事务所是一家专业从事项目数据分析的服务性机构,公司由多位高级项目数据分析师发起成立,并拥有一支集经济研究、金融投资、财务分析、工商管理等多领域的复合型团队。

目前,上海天元项目数据分析师事务所业务网络遍布海内外与全国 30 多个省市自治区,已与多家国内外金融机构、大型财团、银行、上市公司、商会等建立了业务合作意向及业务往来,业务涉及农业、林业、建材、石油、石化、化工、通信、电子、电力、电网、水利、铁路、民航、煤炭、冶金、轻工、纺织、医药、机械、市政、交通、房地产、卫生、教育、文化娱乐、环保、旅游等诸多行业领域,在为客户提供服务的过程中,秉承诚信、客观、科学、实效、公正的经营理念,坚持高水准的专业品质,为客户提供最有价值的服务,成为了客户发展历程中值得信赖的合作伙伴!

专家点评:

上海天元项目数据分析师事务所自成立以来,一直致力为中小型企业、国内外银行、投融资公司、政府组织等机构提供投融资方面专业的、系统的服务,并为项目投资方以及融资方提供具有经济性、权威性、客观性、公正性、实用性的项目数据分析报告。上海天元撰写的报告内容结构严谨,分析到位。其专业的方法、谨慎的作风、客观的态度、公正的原则以及热情的服务,帮助企业做出切实有效的决策。为推动中国项目数据分析事业的发展贡献着自己的力量。

(5) 湖南翰林项目数据分析师事务所

湖南翰林项目数据分析师事务所是经国家行政管理部门核准注册,经中国商业联合会数据分析专业委员会备案(中数委团证第 080 号)的独立法人机构。湖南翰林项目数据分析师事务所是湖南省成立较早规模较大附和资质较齐的项目数据分析师事务所,是一家专业从事项目投资融资咨询、各类数据分析及管理咨询的专业机构。

事务所以专业的项目数据分析服务为核心,同时与湖南南方会计师事务所有限公司、湖南新融达咨询(集团)、邵阳南方资产评估有限公司、邵阳南方工程造价咨询有限公司、邵阳科信税务师事务所有限公司、邵阳南方司法鉴定所、邵阳华信房地产评估有限公司等执业联盟单位紧密联合,在会计审计、各类价值评估等业务领域弥补项目数据分析师事务所在传统法定业务上的短板,同时也改善了上述各类中介机构在新兴数据分析业务上的不足,构筑数据分析大平台,编织一张数据分析的大网,拓宽数据分析宽度,增强数据分析的深度,提升企业自主了解自身数据的意识。

专家点评:

湖南翰林项目数据分析师事务所自成立后,一直注重自身行业品牌及事务所形象的推广宣传,事务所拥有多位企业管理方面的复合型项目数据分析师,在提供科学、专业的数据分析服务时有了最基础的保障。事务所的数据分析报告思路清晰,内容详实,专业度较高,数据资料丰富,市场需求部分研究方法得当,为企业科学决策提供了依据。

2、 技术类混业公司情况

技术类混业公司指数据分析业务是其重要组成部分,与技术业务有机组合,混业经营。此类公司最具有代表性的有阿里巴巴、百度、中国电信等。见下图,有产业分布及代表公司。

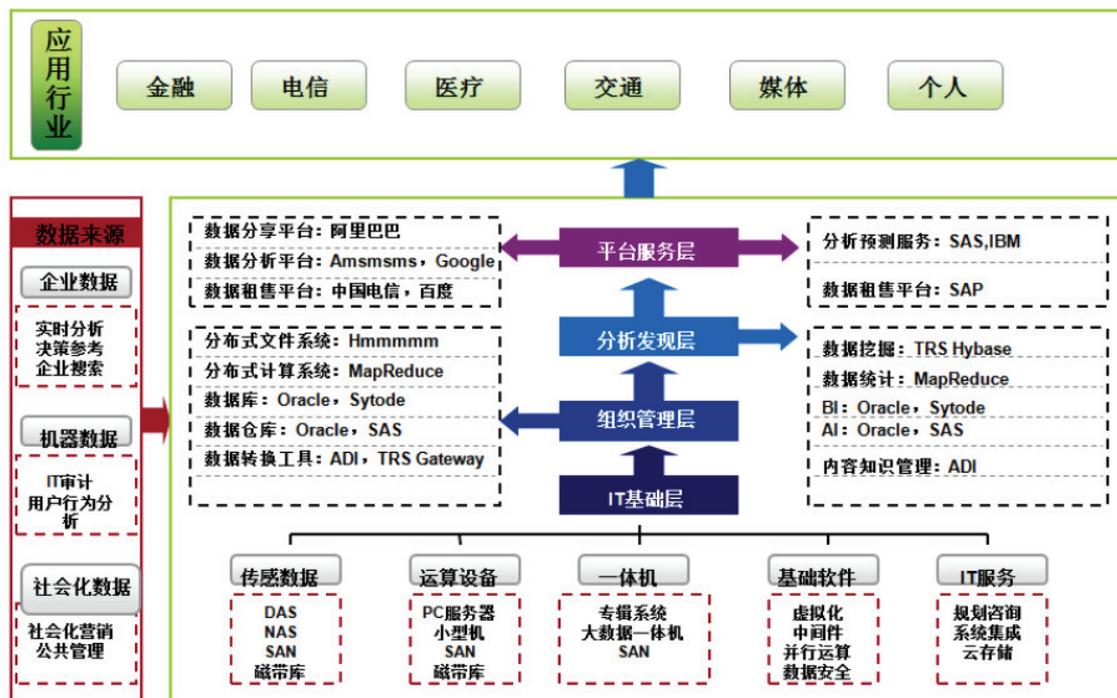


图 1-26 大数据产业链及代表企业

典型公司概况

a、百度大数据

百度大数据最重要的来源是通过爬虫搜集的 100 多个国家近万亿网页数据，数据量是在 EB 级的规模。百度的数据非常多样化，既有非结构化的或者半结构化的数据，包括网页数据、视频和图片等数据，也有结构化的数据，如用户的点击行为数据，广告客户的付费行为数据等。

百度大数据主要服务三类人群：一类是互联网网民，通过大数据和自然语言处理技术让网民的搜索更加准确；第二类是广告主，通过大数据让广告主的广告和搜索关键词的匹配度更高，或者和网民正在看的网页内容匹配度更高；第三类是在重点推进的百度大数据引擎，重点是服务传统行业中拥有一定规模数据的企业。百度大数据引擎代表了互联网企业数据服务能力开放和合作的趋势，由开放云、数据工厂和百度大脑三方面构成。

开放云是百度的大规模分布式计算和超大规模存储云，开放云大数据开放的是基础设施和硬件能力。过去的百度云主要面向开发者，大数据引擎的开放云则是面向有大数据存储和处理需求的“大开发者”。据百度相关人员称，百度开放云还拥有 CPU 利用率高、弹性高、成本低等特点。百度是全球首家大规模商用 ARM 服务器的公司，而 ARM 架构的特征是能耗小和存储密度大，同时百度还是首家将 GPU（图形处理器）应用在机器学习领域的公司，实现了能耗节省的目的。

数据工厂为百度提供将海量数据组织起来的软件能力，与数据库软件的作用类似，不同的是数据工厂是被用作处理 TB 级甚至更大的数据。百度数据工厂支持超大规模异构数据查询，支持 SQL-like 以及更复杂的查询语句，支持各种查询业务场景。同时百度数据工厂还将承载对于 TB 级别大表的并发查询和扫描，大查询、低并发时每秒可达百 GB。

百度大脑将百度此前在人工智能方面的能力开放出来，主要是大规模机器学习能力和深度学习能力。此前它们被应用在语音、图像、文本识别，以及自然语言和语义理解方面，并通过百度 Inside 等平台开放给了智能硬件。现在这些能力将被用来对大数据进行智能化的分析、学习、处理、利用，并对外开放。

百度将基础设施能力、软件系统能力以及智能算法技术打包在一起，通过大数据引擎开放出来之后，拥有大数据的行业可以将自己的数据接入到这个引擎进行处理。从架构来看，企业或组织也可以只选择三件套中的一种来使用，例如数据存放在自己的云，但要运用百度大脑的一些智能算法或者数据存放在百度云，自己写算法。

百度大数据引擎的作用可以从对于政府和对于企业两方面来分析。

(1) 对于政府机构：如交通部门有车联网、物联网、路网监控、船联网、码头车站监控等地方的大数据，如果这些数据与百度的搜索记录、全网数据、LBS 数据结合，在利用百度大数据引擎的大数据能力，则可以实现智能路径规划和运力管理；卫生部门拥有流感法定报告数据、全国流感样病例哨点监测和病原学监测数据，如果和百度的搜索记录及全网数据结合，便可进行流感预测、疫苗接种指导。

(2) 对于企业：很多企业也拥有海量大数据，不过很多企业的大数据处理和挖掘能力比较弱，如果应用百度大数据引擎，则可以对海量数据进行低成本的可靠存储，进行智能化的由浅入深的价值挖掘。如在 2014 年 4 月的百度技术开放日上，中国平安便介绍了如何利用百度的大数据能力加强消费者理解和预测，细分客户群制定个性化产品和营销方案。

b、阿里大数据

阿里最为突出的是电商数据，尤其是用户在淘宝和天猫上的商品浏览、搜索、点击、收藏和购买等数据，其数据最大特点是从浏览到支付形成的用户漏斗式转化数据。一方面，阿里巴巴的消费数据覆盖广、累积深；另一方面，阿里巴巴的云计算技术和数据挖掘能力很强。论商业价值，淘宝、天猫等购物网站产生的消费数据，是领域数据里边比较值钱的。

阿里巴巴大数据整体发展方向是以激活生产力为目的的 DT(data technology, 数据技术驱动) 数据时代发展。阿里巴巴大数据未来将由“基于云计算的数据开放+大数据工具化应用”组成：

(1) 基于云计算的数据开放。云计算使中小企业可以在阿里云上获得数据存储、数据处理服务，也可以构建自己的数据应用。云计算是数据开放的基础，云计算可以为全球的数据开发者提供数据工作平台，阿里分布式的存储平台和在这个平台上的算法工具，可以更好的为数据开发者所用，同时，阿里巴巴还需要做好数据的脱敏，把数据的商业定义，每个标签打得足够清晰，能够让全球的数据开发者在阿里巴巴平台展开数据思维，让数据为政府所用、消费者所用以及行业所用。阿里的大数据开放之后，线上线下的数据能够串联起来，所有人都是数据提供方，也是数据的使用者。

(2) 数据应用两方针

第一个方针：从 IT 到 DT（数据技术），DT 就是点燃整个数据和激发整个数据的力量，被管理所用，被社会所用，被销售所用，为制造业所用，为消费者信用所用。阿里巴巴的数据资产是以电商为主，其中，淘宝和天猫每天会产生丰富多样的数据，阿里巴巴已经沉淀了包括交易、金融、生活服务等多种类型的数据。这些数据能够帮助阿里巴巴进行数据化运营

另外一个其最为重要的应用是金融领域——小微金融。在小微金融企业融资领域。由于银行无法掌握小微企业真实的经营数据，不仅导致很多企业无法拿到贷款，还因为数据类型的不足导致整个判断流程过长，阿里已经通过其电商数据中的交易、信用、SNS 等多种数据来决定是否可以发放贷款以及放贷的额度。

第二个方针：让阿里巴巴的数据、让阿里巴巴的工具能够成为中国商业的基础设施。阿里巴巴已经开始在转型，阿里将由自己直接面对消费者变成支持网商面对消费者，阿里会根据其已有的运营和数据经验，开发更多的工具，帮助网商成长，让网商们更懂得用最好的工具、服务去服务好消费者。正如马云所言“我相信没有一个网商不希望拥有自己的客户，没有一个网商不希望知道客户对自己的体验到底好还是坏，如何持久的拥有这些客户，我们觉得一个国家的经济，应该让给企业家群体去做，我们觉得淘宝网商未来的经济，是应该留给网商们去决定，而不是我们去做决定”。

c、腾讯大数据

从数据类型看，腾讯数据最为全面，这与其互联网业务全面相关，其最为突出的是社交数据和游戏数据，其中：社交数据最为核心的是关系链数据、用户间的互动数据、用户产生的文字、图片和视频内容；游戏数据主要包括大型网游数据、网页游戏数据和手机游戏数据，游戏数据中最为核心的是游戏的活跃行为数据和付费行为数据，腾讯的数据最大的特点是基于社交的各种用户行为和娱乐数据。其中的社交数据，是腾讯自身最擅长的东西。它可以通过大数据分析得知，你的社会关系、性格禀赋、兴趣爱好、隐私绯闻甚至生理周期和心理缺陷都尽在其中。

腾讯的大数据目前更多的是为腾讯企业内部运营服务，相对于阿里和百度，数据开放程度并不

高。腾讯 90% 以上的数据已经实现集中化管理，数据集中在数据平台部，有超过 100 多个产品的数据已经集中管理起来，而且是集中存储在腾讯自研数据仓库（TDW）。腾讯大数据从数据应用的不同环节可以分为数据分析、数据挖掘、数据管理和数据可视化四个层面。

（1）数据分析层有四个产品：自助分析、用户画像、实时多维度分析和异动智能定位工具。自助分析可以帮助非技术人员通过简单的条件配置实现数据的统计和展示功能；用户画像则是对某一类用户或者某一业务的用户实现自动化的人群画像；实时多维度分析工具则是可以对某一指标可以实现实时的多个维度的切分，方便分析人员从不同角度对某一指标进行多维度分析；异动智能定位工具则实现数据异动问题的智能化定位。

（2）数据挖掘层面的产品应用有：精准广告系统、用户个性化推荐引擎和客户生命周期管理。精准广告系统如广点通，是基于腾讯大社交平台的海量数据为基础，通过精准推荐算法，以智能定向推广位导向实现广告精准投放；用户个性化推荐引擎根据每位用户的兴趣和喜好，通过个性化推荐算法（协同过滤、基于内容推荐、图算法、贝叶斯等），实现产品的个性化推荐需求；客户生命周期管理系统，则是基于大数据，根据用户/客户所处的不同生命周期进行数据挖掘，建立预测、预警和用户特征模型，以根据用户/客户所处的不同生命周期特点进行精细化运营和营销。

（3）在数据管理层面则有：TDW（腾讯数据仓库）、TDBank（数据银行）、元数据管理平台和任务调度系统和数据监控。这一层面主要是实现数据的高效集中存储、数据的业务指标定义管理、数据质量管理、计算任务的及时调度和计算以及数据问题的监控和告警。

（4）在数据可视化层面有：自助报表工具、腾讯罗盘、腾讯分析和腾讯云分析等工具。自助报表工具可以自助化的实现结构相对简单和逻辑相对简单的报表。腾讯罗盘分为内部版和外部版，内部版则是服务于腾讯内部用户（产品经理、运营人员和技术人员等）的高效报表工具，外部版则是服务于腾讯合作伙伴如开发商的报表工具。腾讯分析是网站分析工具，帮助网站主进行网站的全方位分析。腾讯云分析则是帮助应用开发商决策和运营优化的分析工具。

总的来看，百度、阿里巴巴和腾讯三大互联网企业都拥有大数据，三大互联网巨头的的数据都用来优化自己业务的运营效果，从这个层面看，其数据价值应用场景比较类似。但由于其业务和商业模式的不同决定了三者数据资产的不同，也决定了三者未来大数据策略的不同，尤其是基于大数据的开放和合作角度看，百度和阿里巴巴相对更加开放。对于重视大数据开放和合作的互联网企业，他们最为期待的是借着大数据开放的策略，与更多的传统行业交换更多的数据，从而更好的丰富其在线下数据，形成线上和线下数据的协同，从中拓展新的商业模式，如智能硬件和大数据健康。

3、 调研类混业公司情况

调研类混业公司指数据分析业务是其重要组成部分，与调研业务有机组合，混业经营。此类公司有代表性的有跨国调研公司 AC 尼尔森、特恩斯（TNS）、益普索（Ipsos）、盖特纳（Gartner）等，国内的有央视市场（CTR）、零点、赛诺等。这些市场调研公司都设置了数据分析部门对其市场调研的数据进行处理和分析。

近年来有两个新的变化：

a、 互联网发展促使数据采集方式发生转变

随着市场调研行业的发展，传统的数据采集方式已经越来越无法满足现在企业的需要，越来越

多的网上调查、App 调查和大数据采集等的数据采集方式取代了传统调研获得的数据。

随着互联网对各行各业的影响扩大，越来越多的企业依赖于互联网组织和开展业务，最大的影响领域在于零售行业，电商对零售及相关行业造成颠覆性影响。这些行业的调查和研究数据直接来源于互联网，对市场调研行业的数据采集方式提出了新的要求。

b、大数据的发展使得市场调研业务逐渐萎缩

2013 年被称为“大数据元年”，2014 年大数据从概念慢慢落地为很多企业的战略，特别是市场导向的企业。

一些依赖于传统“小数据”的市场调研公司在大数据浪潮下业务逐渐萎缩。根据我们的调研结果，2014 年北京地区调研类公司业务萎缩程度在 10% 左右，有的调研公司业务萎缩量高达 30% 左右。

这两个变换都反应了技术对商业的巨大影响力。从工业革命的机械化到如今正经历的计算机革命，技术创新始终支撑着经济变革，颠覆人类的工作和生活方式，改变着商业的生态系统。

二、 数据分析行业建设情况

中国商业联合会数据分析专业委员会是中国数据分析行业唯一的行业协会。其宗旨是：搭建专业平台，促进国内外同行及业界内外的交流；制定行业标准，规范行业市场；加强内外协调，监督行业自律；服务于会员和行业的市场需求，维护会员的合法权益；加强行业会员及各企业之间、行业与行业之间、行业与政府之间的协作，维护行业的正当有序竞争；积极推动数据分析技术的普及和应用，培养专业人才，促进中国数据分析行业健康发展。

在 2014 年，中商联数据分析委主要行业建设工作：

1、 提出“技术+咨询”的发展战略

2014 年 11 月 29 日至 30 日，由中国商业联合会数据分析专业委员会针对全国事务所会员单位举办的“2014 年经营决策数据分析实务专题研讨会”在北京举办。在此行业研讨会上，协会提出了“技术+咨询”的发展战略。

会议指出：

在数据大爆炸的今天，收集数据并进行有效分析，从而指导企业决策，帮助企业成功发展，这是大数据的核心价值所在。然而，自大数据概念提出以来，一直活跃在大数据领域的并非数据分析师，而是一些 IT 人士，甚至很多人认为大数据行业就是 IT 行业。之所以会有这样的认知偏差，是由国内企业数据化程度普遍较低和对大数据核心价值认知模糊决定的。基于这样的现状，协会在此次研讨会上明确提出了“大数据 1.0 时代”事务所应该坚持“技术+咨询”的商业模式。

数据分析咨询能力和技术手段是从事数据分析行业的两个必备武器。首先要具有技术能力，这是实现数据收集、存储、结果展示的工具和手段；二是要具备数据分析咨询能力，因为只有数据分析才是真正能够为客户创造价值的核心。二者相辅相成，缺一不可。在数据分析行业的不同发展阶段，这两者的侧重点会有所不同。在欧美等发达国家，数据分析行业已经发展得较为成熟，咨询能力在两者中占有主导作用。

然而在我国，数据分析行业发展不过十余年的时间，严格说还处于数据分析行业的初级阶段。电商、互联网、零售、金融这些天然具有数据资源优势的行业，已经有计划地完成了从数据收集到

数据分析的体系建立，但是对于国内多数企业，尤其是中小企业来说，虽然已经意识到数据的重要性，但由于数据缺乏以及从业人员自身原因，这些企业还远远没有完成数据化建立。因此在这个时期，首先帮助企业构建数据化平台变得成为重要，这期间必然会涉及 IT 技术的应用。但我们应该清楚地认识到，IT 技术只是实现数据收集、数据挖掘的工具或者手段，它并不能直接指导企业决策，真正发挥作用的是数据分析能力，它可以帮助企业发现数据之间的关系，从而指导企业做出正确地决策行为，实现企业盈利目标。

“技术+咨询”既是国际上流行的商业模式，也是国内事务所将数据分析业务向精深化发展过程中必然采用的商业模式。在本次研讨会上，我们看到一些优秀事务所已经在按此种模式运作，并且已经取得了令人欣喜的成果。协会作为行业的领导者，在推进整个数据分析行业发展以及指导事务所开拓业务方面具有不可替代的作用。目前协会正在研发的 Datahoop 大数据智能分析平台，正是基于这样的初衷，希望通过先进技术的使用，帮助事务所在推动新商业模式过程中遇到技术壁垒而停滞不前时，可以顺利跨过技术难关，将业务重点转移至数据分析咨询服务。相信随着数据分析行业的不断发展，“以咨询业务为主导，以技术为展现工具”的商业模式会被越来越多的事务所采用。

2、大数据智能分析平台“Datahoop”的研发

为了帮助全国事务所突破大数据技术门槛，进入到大数据分析与挖掘领域，中国商业联合会数据分析专业委员会自主研发了一款大数据智能分析平台，切实帮助事务所从业。



这个平台有以下 6 个特点：

十多年实践经验积累，集成行业顶尖算法

大数据的核心是分析，只有分析才能让数据发挥价值。而现在大部分大数据平台都没有很好的算法，甚至全盘照抄书上的算法，严重脱离实际。

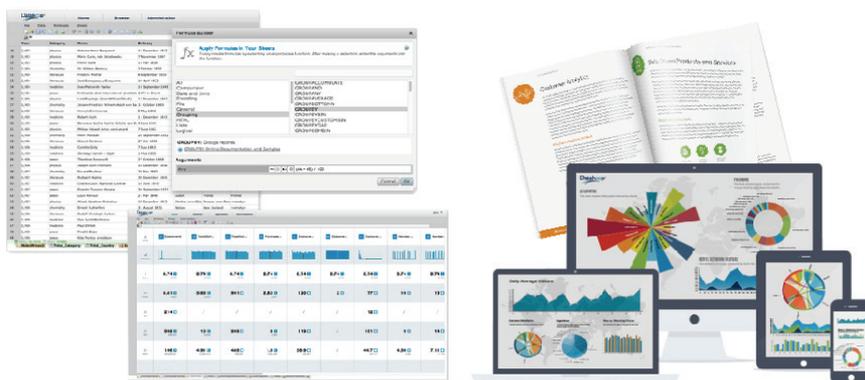
Datahoop 是由中商联数据分析委员会结合 11 年国内外数据分析行业实战经验自主研发的一款大数据智能分析平台，它结合了行业顶尖专家的经验智慧。内置丰富的数据分析和数据挖掘算法，实现算法参数的自动智能调优和升级，同时包含最完善的行业应用模型，使之可以应用于各行各业。这是目前市面上任何一款软件或平台都无法比拟的优势。

支持多种数据接口，真正实现无缝对接

Datahoop 数据接口丰富，集成数据转化及预处理功能，提供实时/非实时统一接口，能与企业现有的 ERP、CRM、OA、财务软件（金蝶、用友、SAP）以及公司网站等资源实现无缝对接，简单设置就可以对企业现有数据进行数据分析和挖掘，节省了大量的重新开发成本，节省了时间。

平台功能支持无限扩展及优化

Datahoop 不断集成新的算法和新的功能模块，终身维护算法库，不断调优。这使得它的应用可以随使用者的需求不断扩展。



独创的安全系统，提供五重防护

Datahoop 私有云尤其重视数据安全，自主研发的安全管理体系可以提供多达五重的防护，通过认证、加密、监控和追踪等手段在传统 PC 终端和移动终端提供数据保护解决方案。

操作简单，无技术要求

Datahoop 让数据挖掘和数据分析操作简单，它独创的一键报表生成功能，使得非专业人士也可轻松发现数据价值，模型建立和使用无需编写任何代码。

跨平台移动终端支持，可以随时操纵数据

Datahoop 提供多终端支持，手机也能访问，老板打开手机就能知道企业状况和解决方案，数据分析人员通过手机就能实现数据分析工作。

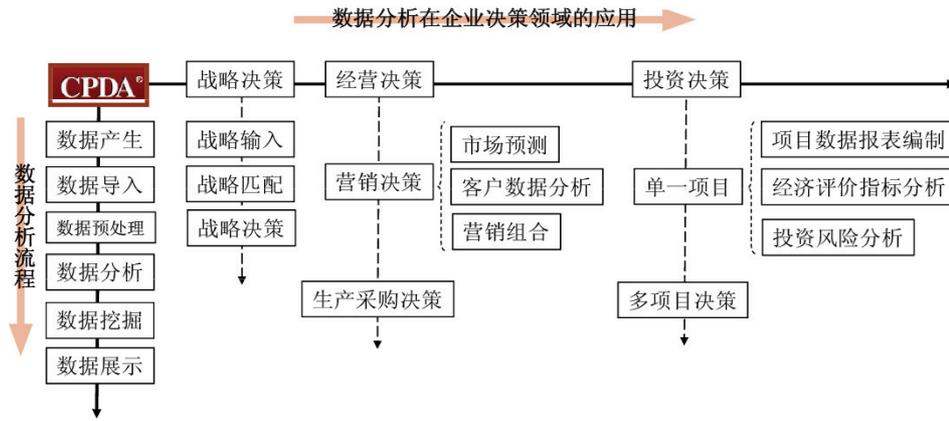
3、项目数据分析师培养课程改革

“项目数据分析师”（CPDA）是协会自主研发的数据分析行业的专业培训课程，旨在为社会培养并输送具有一定专业技能的数据分析人才。该课程自 2003 年在中国正式开办以来，已经在全国二十多个省市培养了近万名项目数据分析师。目前已具有一套成熟的课程体系及与之匹配的专业教材、远程学习平台及完善的学员服务体系。同时，协会本着专业负责的态度，对 CPDA 课程不断进行优化，使之更加切合时代发展的需要。

本次课改是自 2003 年以来的第三次大规模课改，主要在实效性、前沿工具、知识体系完善以及体系内部的关联紧密性方面做了重点调整，充分扩充“客户数据分析”、“市场分析与预测”两大模块，结合最新的技术，使课程内容更加适应社会发展需求，更加贴近生活，易于应用。

本次课改的成果已在 2014 年 12 月 27 日北京地区第 56 期面授课中进行展现，并取得了良好的效果，得到了师生们的一致好评。

课程架构



课改内容:

新版课程内容 NEW		旧版课程内容	
目录	课程概要	目录	课程概要
1. 数据分析概述	<ul style="list-style-type: none"> 数据与数据分析的发展脉络：从小数据到大数据 数据分析的应用领域-数据分析产业链 数据分析人才培养体系 CPDA认证培训系统介绍 	1. 数据分析概述	<ul style="list-style-type: none"> 什么是数据分析？ 数据分析以及在国内发展现状 数据分析对企业经营和项目投资的重要作用 数据的品质 项目数据分析师未来职业优势和机会 项目数据分析师认证培训课程体系 and 考试方式
2. 战略管理 (企业战略管理的模型与方法)	<ul style="list-style-type: none"> 战略输入 战略匹配 战略决策 	2. 战略管理	<ul style="list-style-type: none"> 企业战略管理是什么？ 战略管理对企业重要的意义 企业战略框架 企业战略管理的模型与方法
3. 数据产生	<ul style="list-style-type: none"> 什么是数据 数据的类型 数据分析流程与建模思想 大数据生成 传统数据收集 	3. 数据采集	<ul style="list-style-type: none"> 什么是数据？数据有哪些类型？ 什么是数据采集？数据采集有哪些作用？ 如何进行数据采集？
4. 数据导入与预处理	<ul style="list-style-type: none"> 大数据导入 传统数据导入 数据清洗 数据加工 	4. 数据处理	<ul style="list-style-type: none"> 什么是数据处理？ 怎样进行数据录入？ 如何进行数据清洗？ 数据的分析前加工
5. 数据分析	<ul style="list-style-type: none"> 远程预习：R语言操作基础 数据分析方法 描述型数据分析 	5. 基础数据分析	<ul style="list-style-type: none"> 基础数据分析方法 对比分析 集中度分析 离散性分析 分布分析 交叉分析
6. 数据挖掘	<ul style="list-style-type: none"> 挖掘概述 Kmeans 购物篮 决策树 朴素贝叶斯 	6. 预测数据分析	<ul style="list-style-type: none"> 什么是市场预测？ 为什么进行市场预测？ 如何进行市场预测？ 定性预测方法 定量预测方法
7. 数据展示	<ul style="list-style-type: none"> 定性数据图标展示 定量数据图标展示 数据展示技巧 	7. 营销决策	<ul style="list-style-type: none"> 什么是营销决策？ 客户数据分析 企业品牌研究 营销组合决策——新产品开发决策 营销组合决策——价格决策 营销组合决策——促销决策 营销组合决策——渠道决策

新版课程内容 NEW		旧版课程内容	
目录	课程概要	目录	课程概要
8.营销决策	<ul style="list-style-type: none"> 营销决策的理论背景 营销决策的实务思路 	8. 生产运营决策	<ul style="list-style-type: none"> 生产管理的内容和数据分析需求 生产计划和产品混合决策 分厂选择模型 经济生产量模型EPQ 设备最优化使用寿命决策 经验曲线模型
9.市场分析与预测	<ul style="list-style-type: none"> 市场分析的研究领域 什么是市场预测 为什么进行市场预测 如何进行市场预测 预测方法概述 定量预测方法详解—回归分析与预测 定量预测方法详解—时间序列分析与预测 	9. 采购优化决策	<ul style="list-style-type: none"> 采购决策的5R体系 供应链管理决策 最有采购量EOQ模型 物流需求计划MRO模型 采购优化模型ABC分析法 供应商选择和评估（层次分析法）
10.市场细分与客户数据分析	<ul style="list-style-type: none"> 客户数据分析概述 客户细分与市场定位 客户细分方法与模型选择 客户细分技术：聚类分析详解 客户细分技术：主成分分析与因子分析详解 客户数据分析综合案例：从问卷到结论 	10. 数据报表编制	<ul style="list-style-type: none"> 报表编制的基础数据 总投资构成及估算 总成本费用构成及估算 收益构成及估算 现金流量表的编制
11.营销组合数据分析	<ul style="list-style-type: none"> 营销组合决策的意义与基本流程 产品策略 价格策略 促销策略 渠道策略 	11.经济评价指标分析	<ul style="list-style-type: none"> 为什么要对项目进行经济指标评价？ 静态评价指标 动态评价指标
12.生产采购决策与库存优化	<ul style="list-style-type: none"> 生产与采购决策概述 分厂生产策略 生产计划与产品混合决策 供应链管理决策 库存管理决策 供应商决策 生产质量控制 交货期管理 	12.投资风险分析	<ul style="list-style-type: none"> 不确定性与风险分析 项目盈亏平衡分析 投资收益风险敏感性分析 风险概率分析
13. 投资决策概述	<ul style="list-style-type: none"> 量化投资相关概念 量化投资框架 量化投资原则 现金流量为基准的原则 	13.投资选择与评估	<ul style="list-style-type: none"> 资本限量决策 公司价值分析
14.项目投资经济收益分析	<ul style="list-style-type: none"> 现金流量表的解读 经济评价 投资估算 资金构成 成本估算 收入预测 	14.量化经营实战案例	
15.投资风险分析	<ul style="list-style-type: none"> 项目不确定性与风险分析 项目盈亏平衡分析 项目敏感性分析 项目风险概率分析 	15.数据分析项目流程及工作方法	
16.投资选择与评估	<ul style="list-style-type: none"> 独立项目决策 互斥项目决策 混合项目决策 	16.数据分析技术指导	
17.数据分析项目流程及工作方法	<ul style="list-style-type: none"> 项目建议书撰写 项目投资计划书 		
18.复习课	<ul style="list-style-type: none"> 知识串讲 实操模拟题 		

第二部分 2014 年中国数据分析行业发展的特点与问题

从 2003 年底信息产业部电子行业职业技能鉴定指导中心（现为“工信部教育与考试中心”）正式设立“项目数据分析师”培训项目，并制定出项目数据分析师培训、考试及管理办法到 2014 年，中国的数据分析行业已经走过了 11 个年头。这期间中国的项目数据分析师、项目数据分析师事务所、行业协会从无到有，发展越来越快，业务领域也从最初的投资数据分析逐步转向经营数据分析。2013 年 12 月 10 日，由中国商业联合会数据分析专业委员会主办的中国数据分析行业 10 周年庆典活动在北京广播大厦隆重举办，活动围绕“回顾过去，展望未来”这一主题，总结我国数据分析行业发展十年来所取得的重要成绩，并与与会人员一起谋划数据分析行业的下一个十年大计。

2014 年是中国数据分析行业第二个十年发展的开局之年，随着整个社会对数据价值的认识越来越深刻，随着数据产业链上不同领域的技术进步和资本投入，中国的数据分析行业有着它此阶段的发展特点和阶段性的发展困难。

一、2014 年中国数据分析行业发展的特点

1、数据价值逐渐被认可但仍需进一步推广普及

数据的资源化是指数据在企业、社会和国家层面成为重要的战略资源。2014 年数据成为新的战略制高点，是大家抢夺的新焦点；数据不断成为机构的资产，成为提升机构和公司竞争力的有力武器。从全国各地数据分析师事务所的项目经验和与企业沟通了解可以看出，数据既可以作为企业的核心资产驱动决策和流程，也可以作为业务板块驱动产品。

在数据驱动决策方面，过去往往只是简单汇总结果数据、二手信息和高层经营决策，凭主观与经验对市场进行评估和制定策略，而在以数据分析为主的量化决策时代，企业往往是通过收集、分析大量内外部的数据，进行智能化决策，制定更加行之有效的战略；在数据驱动流程方面，过去企业主要是一种被动思维模式，往往是出现问题、逻辑分析、找出因果关系、提出解决方案，而量化经营时代，企业需要具备主动思维模式，即收集数据、量化分析、找出影响关系、提出优化方案；在数据驱动产品方面，过去往往是把企业收集的数据打包，授权给别的服务公司使数据增值，现在，企业则开始利用数据把现有产品变得更加高效、智能或更具有洞察力，从而直接或间接产生额外的收入。

从数据分析协会对行业的研判以及各事务所的实际情况来看，数据在企业决策、流程管理和产品改进方面的重要价值已经是逐步得到认可，但不同行业的认知水平仍有较大差异，要想使全社会充分认识到数据的价值，仍需要进一步的推广普及。数据分析行业协会和从业单位将在这方面继续加大宣传力度。

2、数据技术在新兴技术的推动下逐渐走向成熟

数据技术离不开云计算技术，云计算为数据分析提供弹性可扩展的基础设施支撑环境以及数据服务的高效模式，数据分析则为云计算提供了新的商业价值，因此从 2014 年开始数据技术与云计

算技术进入更完美的结合期。总体而言，云计算、物联网、移动互联网和社交网络等新兴计算形态，既是产生大数据的地方，也是需要大数据分析方法的领域。

从数据分析的技术角度，主要有三个层面：第一个是数据预处理，包括数据的抽取、转换和加载；第二个是数据经过预处理之后进行存储、计算、分析的过程，这是体现数据价值的关键环节；第三个是数据的可视化，即如何把数据转换成大众易于理解、易于形成决策的方式，类似我们在 2014 年春节的时候新闻中不断在播放的百度对春运的大数据分析形式。

我们认为，不论是大数据处理还是传统的数据处理，同样会经历这三个环节，但大数据处理由于数据量的爆发和非结构化数据的引入，使得不同环节在技术上面对不同的挑战。在数据预处理阶段面对的主要挑战是多源、多模态数据融合及数据质量和可用性的处理；在数据存储阶段面对的主要挑战是高效数据查询访问和低成本超大容量的存储；在数据计算阶段面对的主要挑战是数据密集计算的效率和不同计算场景的处理；在数据分析阶段面对的主要挑战是非结构化数据的分析、分析深度及分析的自动化实现；在数据可视化阶段面对的主要挑战是海量数据的直观理解。

大数据的整体技术发展层面，美国等传统 IT 强国的大数据战略也都处于试错阶段；硬件方面，相对于其他一些高精尖技术，大数据对于硬件基础设施的要求相对较低，不会受困于基础设备核心元件的相对落后。软件方面，与在传统数据库操作层面的技术差距相比，大数据分析应用的中外技术差距要小得多。

3、数据生态系统逐渐丰富并影响企业商业模式

数据分析行业的整个产业链可以从基础数据的采集开始定义。包括数据采集、数据存储、数据处理（含数据清洁）、数据分析，直到数据分析结果呈现和商业应用。通过对市场的了解和判断，我们认为目前国内整个数据分析产业链的布局相对完整，但局部环节的竞争程度差异化明显。

在数据采集领域，综合型大数据源市场处于结构化整合阶段，数据源平台企业进行以数据资产为核心的整合，最终会形成几大平台型的数据源企业，或是纯粹的用户场景数据平台。而垂直型大数据源市场处于布局阶段，全新行业细分市场的用户场景数据企业诞生，目前还处于积累用户量的阶段，不断有新的企业成长起来，抢夺已有的入口或者开辟全新的入口；数据存储及挖掘市场结构较为稳定，国际巨头垄断，寡头格局已经形成，国内企业短期内很难超越。据易观智库的数据显示，在硬盘存储器领域，EMC、IBM、NetApp 等前五家厂商在全球的市场占有率合计超过 70%，在服务器领域，HP、IBM、Dell 等前五厂商市场占有率超过 85%，在数据库软件领域集中度更好，Oracle、IBM、Microsoft 等前五厂商超过 90%；在数据应用市场，国内企业机会较多，但技术仍不成熟。其中行业应用主要包括商业智能，企业大数据应用系统和在线数据分析平台等，辅助应用包括数据可视化、视频、语音识别等通用性的大数据应用。

与此同时，数据分析产业链的不同环节正在影响企业商业模式的转变。出现了四种主要的商业模式：一是数据存储租用，主要是利用存储能力进行运营，满足企业和个人面临海量信息存储的需求，典型厂商如阿里巴巴、金山云、腾讯云、百度云、IDC 等；二是数据增值服务，主要是对数据进行挖掘分析后预测相关主体的行为，以开展业务，典型厂商如电信业 IZP、金融业 Opera、阿里金融、GFK 等；三是租赁信息服务，主要是直接进行信息租赁或提供信息租售平台，典型厂商如 Twitter、VISA、MasterCard、思维图新、广联达、彭博等；四是数据技术服务，IT 服务提供商提供大数据出租模式，通过出租一个虚拟空间，从简单的文件存储，逐步扩展到数据聚合平台，针对

企业需求，为运营某一环节或某一业务问题提供解决方案，实施单点技术，例如向零售商提供大数据分析技术，获得营销电子，还可以针对企业系统需求，提供整体解决方案，例如 IBM 提供软硬一体的大数据解决方案，华为基于 IT 基础设施领域在存储和计算的优势，提供整体大数据解决方案，还可以通过云服务提供在线大数据技术或解决方案，称为数据应用即服务的模式，简称 BDaaS (Big Data as a Service)。

4、数据分析加速向传统领域拓展但仍处于初期

在 2014 年，大数据应用的典型行业仍然是以互联网为基础的电子商务，主要借助用户行为分析，产品分析、推荐系统和广告追踪优化，用于市场精准营销和个性化导购，未来该行业数据分析的应用方向主要是基于数据资产化的数据化运营和行业垂直整合。除了电子商务行业相对成熟的应用之外，数据分析已经逐渐向金融、电信、物流、医疗以及交通、教育等领域拓展并取得了初步成效。

金融行业的数据分析主要应用在客户行为分析、金融风险分析和运营分析。目前银行、保险、证券等拥有大数据的金融机构正在推进相关工作；电信行业的数据分析主要应用在提升网络服务质量、精准判断客户需求、升级行业信息化解决方案以及提供数据安全服务等方面，目前在支持营销运营管理、客户体验管理以及促进智能管理运营等方面已经取得了初步成效，未来将在社交网络分析、客户体验分析以及客户价值分析方面进一步的应用；在医疗行业，数据分析主要应用在各地医院资源的统一协调，保健咨询服务系统，新药研发以及便于医患沟通，提高诊疗效率等方面，目前已在药品研发、疾病治疗、公共卫生管理和居民健康管理方面取得了初步效果，未来将进一步应用在智能医疗平台化和电子健康服务生态系统；在政府公共服务方面，数据分析可以帮助相关部门加强统筹规划，提高智慧城市感知水平，提高经济社会智慧化水平，目前在智慧交通和智慧安防方面的应用已经初见成效，未来在政府的智能化，促进政府的数据治理以及数据驱动政府公共服务方面必将发挥更大作用。

总的来看，大数据的理念、技术和应用模式等正在向更广阔的领域扩散，但整体还处于初级阶段，谈概念的多，真正用的少，而且案例散发，尚未形成燎原之势。传统企业在理解什么是大数据，制定大数据战略以及整合多种数据源方面还存在诸多问题，大数据技术与模式的改造和适配还要经历长期过程。

5、我国大数据发展战略顶层设计开始受到重视

大数据热潮席卷全球，争夺新一轮技术革命制高点的战役已经打响，大数据正在引发传统行业的变革。当前，全球多个国家都相继出台了大数据战略规划，我国对大数据的重视也日渐提升，在产业链各环节的积极推动下，我国大数据正在加速“落地”。

中国政府在美国提出《大数据研究和发展计划》的 2012 年也批复了“十二五国家政务信息化建设工程规划”，总投资额估计在几百亿，专门有人口、法人、空间、宏观经济和文化等五大资源库的五大建设工程。我国的开放、共享和智能的大数据的时代已经来临！

2012 年 8 月份国务院制定了促进信息消费扩大内需的文件，推动商业企业加快信息基础设施演进升级，增强信息产品供给能力，形成行业联盟，制定行业标准，构建大数据产业链，促进创新链与产业链有效嫁接。

2014 年 9 月，经全国信标委 2014 年第二次主任委员办公会审议，全国信息技术标准化技术委

员会大数据标准工作组成立，负责制定和完善我国大数据领域标准体系。根据目前发布的体系框架来看，大数据标准将主要分为基础标准、数据处理、数据安全、数据质量、产品和平台、应用和服务 6 大领域，每一个领域又都被分出多个小项目。

工业和信息化部电子工业标准化研究院副院长高林表示，发布的一些数据中是否会涉及到人的隐私、用户应该通过何种渠道获取、是否该进行“脱敏”处理，这些数据的安全问题是最基本的问题之一。一系列标准的制定也是为了更好的规范大数据的使用。“目前已经进入标准草案的制定阶段，根据预期，大约在 2016 年陆续发布。”高林说。

2014 年 12 月份，发改委高技术司与工信部软件司共同组织召开促进大数据发展和应用工作研讨会，会议深入讨论了大数据发展的形势机遇和存在的问题，提出了加强我国大数据发展的顶层设计，完善相关法律法规和标准规范，强化政府数据共享开放和社会化利用机制，加强人才建设等方面的措施建议。

6、数据共享时代各种区域大数据联盟不断涌现

数据共享联盟在 2014 年逐渐壮大成为产业的核心一环。数据是基础，之前在科技部的支持下，已建立了多个领域的数据共享平台，包括气象、地震、林业、农业、海洋、人口与健康、地球系统科学数据共享平台等。之后，数据共享扩展到企业和区域层面。

(1) 中关村大数据产业联盟

2012 年 12 月 13 日，北京中关村举办“中关村大数据日”活动，会上宣布成立了中关村大数据产业联盟，发布了云天使基金、中云融汇基金、大数据实验室孵化基金三支产业投资基金，并为“中关村云广场”正式命名。活动同时发布信息：将把今后每年的 12 月 13 日确定为中关村的大数据日，集中开展以大数据为主题的各项交流活动。中关村大数据产业联盟成为推动中关村大数据产业飞速发展的重要力量。联盟组织成立中关村大数据工程中心，发布大数据年度发展报告，并策划一系列业界交流活动，进一步加快中关村大数据产业生态环境的构建，构建广阔的产业交流平台，推动产业更快发展。

(2) 中国企业大数据联盟

2014 年 11 月 4 日，由中国电信、中国东方航空、中航信、工信部电信研究院、互联网协会互联网金融工作委员会、易车、亚信科技、晶赞科技等 8 单位发起，45 家企业联合组成的中国企业大数据联盟在北京举行成立大会。该联盟致力于推动跨界大数据标准制定与合作，积极促进大数据技术成果在各行业应用推广，全面提升成员单位大数据应用的整体水平；汇聚各方力量，借鉴世界先进经验，使联盟成为推动大数据技术进步、应用创新的中坚力量，为我国大数据产业健康发展做出贡献。

(3) 长三角大数据联盟

2014 年 11 月 1 日，在上海大数据产业技术创新战略联盟主办的 2014 数据中国产业创新峰会上，我国首个最具投资价值的大数据企业排行榜，以及中国大数据产业图谱 1.0 新鲜出炉。11 月 1 日也因此成为首个“1101 中国大数据开放日”。为整合区域资源，长三角大数据联盟也在此次峰会上成立。长三角大数据联盟是由上海大数据产业技术创新战略联盟和上海现代服务业联合会共同发起成立的产学研线上创新组织，通过定期活动、交流、展览、讲座、众筹等多种方式，打造长三角大数据企业创业投资的融合创新生态。

(4) 深圳市大数据产业发展促进会

2014年9月，由腾讯、华为、联想、宝德科技、中兴等深圳高新技术企业共同发起的深圳市大数据产业发展促进会正式成立，该协会会长是深圳赛迪工业和信息化研究院，深圳南油有限公司任常务副会长，深圳新视界传媒与华为、联想和国家超算深圳中心同为副会长单位。在全球信息化高速发展的今天，作为国家级创新城市，深圳市科技创新委将大数据作为重点扶持的领域之一，在政策和具体行动上给予大力支持。同时，一大批从事大数据产业实践的知名企业在深圳扎根，也为大数据产业发展创造了得天独厚的条件。深圳市大数据产业发展促进会将这些优秀大数据研究机构和企业聚集起来，充分发挥产业联动作用，打造完整大数据产业链，未来将对深圳的发展具有重要战略意义。该协会的成立，标志着深圳在促进大数据科技创新、引领中国大数据产业变革方面迈出了坚实的一步。

二、2014 年中国数据分析行业发展的的问题

1、专业从业机构存在的问题

项目数据分析师事务所是我国从事数据分析服务的专业机构，2014年总体发展情况良好，但在调研过程中我们也发现其发展存在以下几个问题：

(1) 网络推广是事务所主要的业务来源

从调研统计结果看，事务所获得业务的来源主要有通过网络宣传推广和政府合作，通过这两种方式获得过业务的事务所数量占比分别为 72.09%和 43.22%。

通过与金融、投融资公司推荐获得业务来源的事务所数量占比为 34.88%。通过其他方式获得业务的事务所数量占比相对较小。

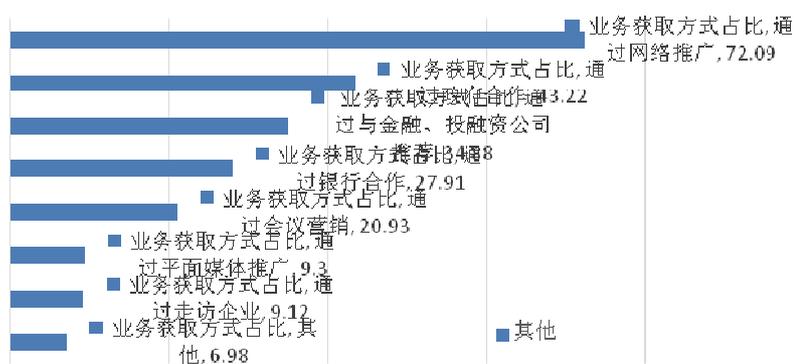


图 2-1 通过网络宣传是事务所获取业务的主要来源

随着网络的普及，互联网成为人们主要的信息来源，网络推广成为事务所获取业务的主要途径。根据我们的调研，还有很多事务所在做纸媒和硬广的推广，存在着投入产出比、转换率偏低的现象。在这种情形下，建议事务所在市场上应重点做好品牌推广，特别是网上的品牌推广。适当的减少企业拜访的业务获得方式，减少无效的市场活动。

（2）有部分事务所未开展经营类业务

目前有 34.85% 的事务所尚未出具经营决策类的报告，完成过 1-2 份该类型报告的事务所占比 15.15%，完成过 3-6 份该类型报告的事务所占比 30.30%，完成超过 6 份该类型报告的事务所占比 12.12%。目前企业经营类的项目较去年比有所增长，但仍然偏少，市场未得到全面打开。跟踪调研发现并不是事务所没有业务来源，一方面是有些事务所在经营类业务领域未有研究积累，导致不敢接项目，另一方面也有部分事务所只想赚投融资类项目的“快钱”，缺乏发展的眼光。

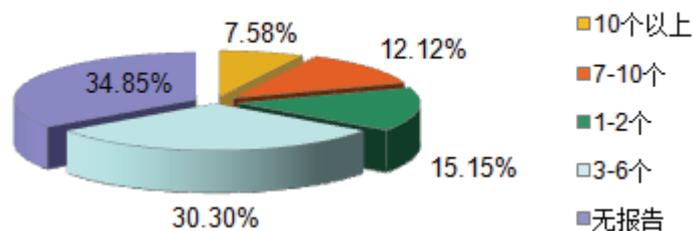


图 2-2 有部分事务所未开展经营类业务

这个问题的出现，一方面启示协会在审批新的事务所的时候，应坚持“保质控量，宁缺毋滥”的严格审批原则，另外一方面也提示行业协会需要进一步引导，提醒事务所应大力加强自身的能力建设，强化研究能力，提高业务水平，以发展的眼光看待市场的发展，积极转型，抢占市场先机。

（3）其他问题

事务所现在面临的问题是分析师的专业水平仍有待进一步提高，行业分析师人才仍有待培养。各事务所都有自己专业的分析师团队，但普遍存在分析师的从业经验有限，专业能力不系统、不全面的问题，所以分析师的综合能力仍有待提高，仍然需要多参与项目，多进行行业交流，以期分析师整体水平的提升；事务所遇到的另外的困难是业务转型困难，转型困难主要有几个原因，一是以前的分析师主要从事投资类、评估类项目，事务所和分析师对经营业务的认可度和了解度都不高；二是经营决策类业务目前缺乏核心技术，缺乏统一培训；三是企业经营类的数据分析市场认可度太低，目前大部分企业没有认识到这种需求。

一些事务所还存在一些其他的问题，如数据采集难度大、事务所缺乏有效的宣传手段，事务所的行业经验较少等。

2、整体面临的问题

在中国项目数据分析师事务所是中国数据分析行业唯一被认可从事数据分析服务的专业机构。除了此类型专业机构，还有混业经营的机构，如调研公司、IT 公司、网络公司和系统集成公司等商业机构，这些商业机构的特点是数据分析是其重要的业务组成部分，与其他业务有机组合，混业经营。这三种主体构成了中国数据分析行业的主要从业机构。2014 年，这些从业机构总体上发展态势良好，但在发展中也遇到了以下几个困难：

（1）数据需求模糊导致数据资产流失

数据分析协会与事务所在与企业的实际接触过程中发现，很多企业的业务部门不了解数据分析，

也不了解数据分析的应用场景和价值，因此难以提出数据分析的准确需求。由于业务部门需求不清晰，数据分析部门又是非盈利部门，企业决策层担心成本问题，导致很多企业在搭建数据部门时犹豫不决，或者很多企业都处于观望尝试的态度，从根本上影响了企业在数据分析方向的发展，也阻碍了企业积累和挖掘自身的数据资产，甚至由于数据没有应用场景，删除很多有价值历史数据，导致企业数据资产流失。因此，这方面需要数据分析从业者和行业协会一起，推动和分享数据应用场景，让更多的业务人员了解数据分析的价值。

(2) 数据孤岛严重导致数据价值受限

在实际接触过程中我们发现，很多企业启动数据战略最大的挑战是数据的碎片化。在很多企业中尤其是大型企业，数据常常散落在不同部门，而且这些数据存在不同的数据仓库中，不同部门的数据技术也有可能不一样，这导致企业内部自己的数据都没法打通。如果不打通这些数据，数据的价值则难以充分发挥。数据分析需要不同数据的关联和整合才能更好的发挥理解客户和理解业务的优势。将不同部门的数据打通，并且实现技术和工具共享，才能更好的发挥企业数据的价值。

(3) 数据质量差导致数据处理耗时长

很多中型以及大型企业，每时每刻都在产生大量的数据，但很多企业对数据的预处理阶段很不重视，导致数据处理很不规范。数据预处理阶段包括抽取数据并把数据转化为方便处理的数据类型，对数据进行清洗和去噪，以提取有效的数据等操作。有些企业甚至在数据的上报过程就出现很多不规范不合理的情况。以上种种原因，导致企业的数据库可用性差，数据质量差，数据不准确。而大数据的意义不仅仅是要收集规模庞大的数据信息，还有对收集到的数据进行很好的预处理，才有可能让数据分析和数据挖掘人员从可用性高的数据中提取有价值的信息。Sybase 的数据表明，高质量的数据应用可以显著提升企业的商业表现，数据库可用性提高 10%，企业的业绩至少提升在 10% 以上。

(4) 传统数据库技术不适应海量数据

大数据分析在技术架构的挑战包含以下几方面：(a) 传统的数据库部署不能处理 TB 级别的数据，快速增长的数据量超越了传统数据库的管理能力。如何构建分布式的数据仓库，并可以方便扩展大量的服务器成为很多传统企业的挑战；(b) 很多企业采用传统的数据库技术，在设计开始就没有考虑数据类别的多样性，尤其是对结构化数据、半结构化和非结构化数据的兼容；(c) 传统企业的数据库，对数据处理时间要求不高，这些数据的统计结果往往滞后一天或两天才能统计出来。但大数据需要实时处理数据，进行分钟级甚至是秒级计算。传统的数据库架构师缺乏实时数据处理的能力；(d) 海量的数据需要很好的网络架构，需要强大的数据中心来支撑，数据中心的运维工作也将成为挑战。如何在保证数据稳定、支持高并发的同时，减少服务器的低负载情况，成为海量数据中心运维的一个重点工作。

(5) 数据安全问题令人担忧需要重视

网络化生活使得犯罪分子更容易获得关于人的信息，也有了更多不易被追踪和防范的犯罪手段，可能会出现更高明的骗局。如何保证用户的信息安全成为大数据时代非常重要的课题。在线数据越来越多，黑客犯罪的动机比以往都来的强烈，一些知名网站密码泄露、系统漏洞导致用户资料被盗等个人敏感信息泄露事件屡屡发生，这提醒我们，要加强数据网络安全的建设。另外，数据的不断增加，对数据存储的物理安全性要求会越来越高，从而对数据的多副本与容灾机制也提出更高的要求。目前国内很多传统企业的数据安全令人担忧。

（6）数据分析人才匮乏制约行业发展

数据产业的每个环节都需要依靠专业人员完成，因此，必须培养和造就一支掌握数据技术、懂管理、有数据应用经验的数据建设专业队伍。目前数据相关人才的欠缺严重阻碍数据市场发展。据 Gartner 预测，到 2015 年，全球将新增 440 万个与数据相关的工作岗位，且会有 25% 的组织设立首席数据官职位。数据分析的相关职位需要的是复合型人才，能够对数学、统计学、数据分析、机器学习和自然语言处理等多方面知识综合掌控。未来，数据分析将会出现约 100 万的人才缺口，在各行各业，数据分析中高端人才都会成为最炙手可热的人才，涵盖了大数据的数据开发工程师、数据分析师、数据架构师、数据后台开发工程师、算法工程师等多个方向。因此需要高校和企业共同努力去培养和挖掘。目前最大的问题是很多高校缺乏大数据，所以拥有大数据的企业应该与学校联合培养人才数据分析人才。

（7）数据开放与隐私的权衡需要解决

在数据应用日益重要的今天，数据资源的开放共享已经成为在数据大战中保持优势的关键。商业数据和个人数据的共享应用，不仅能促进相关产业的发展，也能给我们的生活带来巨大的便利。由于政府、企业和行业信息化系统建设往往缺少统一规划，系统之间缺乏统一的标准，形成了众多“信息孤岛”，而且受行政垄断和商业利益所限，数据开放程度较低，这给数据利用造成极大障碍。另外一个制约我国数据资源开放和共享的一个重要因素是政策法规不完善，大数据挖掘缺乏相应的立法，既保证共享又防止滥用。因此，建立一个良性发展的数据共享生态系统，是我国数据发展需要迈过去的一道坎。同时，开放与隐私如何平衡，也是数据开放过程中面临的重大难题。如何在推动数据全面开放、应用和共享的同时有效地保护公民、企业隐私，逐步加强隐私立法，将是大数据时代的一个重大挑战。

第三部分 中国数据分析行业未来发展进阶

一、中国数据分析行业发展的全球背景

全球范围来看，大数据的发展呈现以下趋势。

1、数据逐渐成为具备重要价值的资产

随着大数据应用的发展，数据的价值得以充分体现，数据在企业和社会层面逐渐成为重要的战略资源，成为新的战略制高点，是大家抢夺的新焦点。《华尔街日报》在一份题为《大数据，大影响》的报告中指出，数据已经成为一种新的资产类别，就像货币或黄金一样。Google、Facebook、亚马逊、腾讯、百度、阿里巴巴和 360 等企业正在运用大数据的力量获得商业上更大的成功，此外，金融和电信企业也在运用大数据来提升自己的竞争力。我们相信大数据将不断成为更多机构和企业的重要资产，成为他们提升竞争力的有力武器。

2、大数据在更多的传统行业得以应用

一种新技术的出现，往往首先在少数行业取得好的应用效果，继而对其他行业产生强烈的示范效应。目前大数据在大型互联网企业已经得到较好的应用，其他行业的大数据尤其是电信和金融也逐渐在多种应用场景取得效果。因此，我们相信，大数据作为一种从数据中创造新价值的工具，将会在更多行业的企业得到应用，带来广泛的社会价值。大数据将帮助企业更好地理解 and 满足客户需求，更好地应用在业务运营智能监控、企业精细化运营、客户生命周期管理、精准营销等方面。

3、行业定制化大数据解决方案将涌现

传统商业智能一般只能处理少量数据时的一种方法。大数据用户更希望能获得一种整体的解决方案，即不仅要能收集、处理和分析企业内部的业务数据，还希望能引入互联网上的网络浏览、微博、微信等非结构化数据。除此之外，还希望能结合移动设备的位置信息，这样企业就可以形成一个全面、完整的数据价值发展平台。无论是大数据还是商业智能，目的都是为分析服务的，数据全面整合起来，更有利于发现新的商业机会，这就是大数据商业智能。同时，由于行业的差异性，很难研发出一套适用于各行业的大数据商业智能分析系统，因此，未来在一些规模较大的行业市场，大数据服务提供商将会以更加定制化的商业智能解决方案提供大数据服务。我们相信更多的大数据商业智能定制化解决方案将在电信、金融、零售等行业出现。

4、数据开放的步伐将进一步加快推进

大数据越关联，越开放才会越有价值。如果每一个医院想获得更多病情特征库以及药效信息，那么就需要整个国家，甚至全世界的医疗信息共享，从而可以通过平台进行分析，获取更大的价值。未来，公共事业和互联网企业的开放数据将越来越多，美国、英国、澳大利亚等国家都在政府和公共事业上的数据做出努力。而国内的一些城市和部门也在逐渐开展数据开放的工作。由北京市经济和信息化委员会牵头建设，北京市各政务部门共同参与的北京市政务数据资源网，于 2012 年 10 月开始试运行。该网站致力于提供北京市政务部门可公开各类数据的下载与服务，为企业和个人开

展政务信息资源的社会化开发利用提供数据支撑,推动信息资源增值服务业的发展以及相关数据分析与研究工作的开展;上海在 2012 年启动了政府数据资源开放试点工作,数据涉及地理位置、交通、经济统计和资格资质等数据;2014 年,贵州省也加入数据开放之列,10 月份云上贵州正式上线。对于不同的行业,数据越共享才越有价值。我们相信未来数据会呈现一种共享趋势,不同领域的的数据联盟将越来越多的出现。

5、大数据安全问题将会越来越受重视

像苹果、索尼和 Snapchat 等用户在近年来所遭遇的漏洞一样,重大安全漏洞一直以来并没有影响到大众在社交媒体和网络中分享隐私生活细节的行为。实际上,从未有过如此多的人认为,向公司提供个人信息只是享受新技术的便利所付出的小代价。我们能不能承受“完全风暴”。如今,黑客已经能够威胁到最安全的系统,而政府和执行部门防止数据泄漏,将不法之徒绳之以法的进程却非常缓慢。灾难性的黑客攻击或信息泄漏可能将会足以改变人们的态度,让人们恢复保护个人数据的意识。

随着数据价值的越来越重要,大数据的安全问题也将会逐渐被重视。网络和数字化生活使得犯罪分子更容易获取关于他人的信息,也有更多的骗术和犯罪手段出现,所以,在大数据时代,无论对于数据本身的保护,还是对于由数据而演变的一些信息的安全,对大数据分析有较高要求的企业将至关重要。大数据安全是跟大数据业务相对应的,与传统安全相比,大数据安全的最大区别是安全厂商在思考安全问题的时候首先要进行业务分析,并且找出针对大数据业务的威胁,然后提出有针对性的解决方案。比如,对于数据存储这个场景,目前很多企业采用开源软件如 Hadoop 技术来解决大数据问题,由于其开源性,其安全问题也较突出。因此,市场需要更多专业的安全厂商针对不同的大数据安全问题来提供专业服务。

6、大数据将成为智慧城市的发展引擎

智慧城市是通过物与物、物与人、人与人的互联互通能力、全面感知能力和信息利用能力,通过物联网、移动互联网、云计算等新一代信息技术,实现城市高效的政府管理、便捷的民生服务、可持续发展的产业发展。由于人口聚集给城市带来了交通、医疗、建筑等各方面的压力,需要城市能够更合理地进行资源布局和调配,而智慧城市正是城市治理转型的最优解决方案。智慧城市相对于之前数字城市概念,最大的区别在于对感知层获取的信息进行了智慧的处理。由城市数字化到城市智慧化,关键是要实现对数字信息的智慧处理,其核心是引入了大数据处理技术。随着技术的发展,大数据在智慧城市的建设过程中将发挥着越来越重要的作用。大数据是智慧城市的核心智慧引擎。智慧安防、智慧交通、智慧医疗、智慧城管等,都是以大数据为基础的智慧城市应用领域。

7、大数据将会催生一批新的工作岗位

一个新行业的出现,必将在工作职位方面有新的需求,大数据的出现也将推出一批新的就业岗位,例如,大数据分析师、数据管理专家、大数据算法工程师、数据产品经理等等。具有丰富经验的数据分析人才将成为稀缺的资源,数据驱动型工作将呈现爆炸式的增长。而由于有强烈的市场需求,高校也将逐步开设大数据相关的专业,以培养相应的专业人才。企业也将和高校紧密合作,协助高校联合培养大数据人才。如 2014 年,IBM 全面推进与高校在大数据领域的合作,引入强大的研发团队和业务伙伴,推动“大数据平台”和“大数据分析”的面向行业产学研创新合作以及系统化知识体系建设和高价值人才培养,建设符合中国教学特色及人才需求的大数据相关学分课程,为

未来建设特色专业方向做准备。

8、大数据将在多方面改善我们的生活

大数据不仅用于企业和政府，也应用于我们的日常生活。在健康方面：我们可以利用智能手环监测对我们的睡眠进行追踪，了解睡眠质量，我们可以利用智能血压计、智能心率仪远程监控身在异地的家里老人的健康情况，让远在异地的外出工作者更加放心；在出行方面：我们可以利用智能导航出行 GPS 数据了解交通状况，并根据拥堵情况进行路线实时调优；在居家生活方面：大数据将成为智能家居的核心，智能家电实现了拟人智能，产品通过传感器和控制芯片来捕捉和处理信息，可以根据住宅空间环境和用户需求自动设置控制，甚至提出优化生活质量的建议，如我们的冰箱可能会在每天一大早建议我们当天的菜谱。

二、中国数据分析行业发展的国内环境

IDC 发布的报告预测，中国大数据技术与服务市场规模将从 2011 年的 7760 万美元增长到 2016 年的 6.17 亿美元，5 年的复合增长率达 51.4%。中国市场的规模之大为大数据发展提供了大空间、大平台。我国在省（市）一级对数据分析的支持，特别是大数据分析的支撑走到了世界的前列。

1、北京迎来大数据发展机遇

大数据能够实现巨大的商业价值，可以推动产业结构升级，提高发展质量，还可以加快改造传统产业，推动产业体系整体升级，培育发展大数据产业符合北京产业结构调整升级的要求。与此同时，北京科技资源集中，研发实力丰厚，政府机构及大公司云集，有很强的数据生产能力，在培育大数据产业方面具有原发优势。通过“科技北京”、中关村国家自主创新示范区建设等一系列规划、战略的实施，北京积聚了全国最大规模和最有价值的数据资产，为发展大数据奠定了良好的数据基础。

北京在海量数据存储、挖掘、云计算等技术处于国内领先地位，与国外大数据技术热点保持一致，发展大数据遇到的关键技术相对容易突破。12 月 10 日，以“点击数据 启动未来”为主题的大数据交易服务平台发布会在京召开，北京大数据交易服务平台正式发布上线，同时北京大数据交易服务平台与大数据标准规范合作启动。

总的来看，北京市在大数据发展方面已经提前布局，形成了卓有成效的成果，如推动政府信息公开、鼓励大数据产业集群、建立中关村的大数据产业联盟及交易平台、搭建首都科技大数据平台等。下一步，北京可结合自身巨大优势，布局大数据。尤其要完善大数据相关的法律法规和技术规范，继续鼓励大数据通用关键技术研发，进一步开放政府等公共部门数据，统筹谋划大数据应用，培育相关产业，开展“数据驱动型创新”，把北京打造成智慧的创新中心。

在政策层面，密切关注国际立法理念的演变趋势，结合大数据技术发展趋势和具体的国情、市情对相关制度进行前瞻性研究；在行业技术标准层面，加快制定大数据分类、文件格式、传输协议、访问应用程序接口等；在大数据通用技术研发方面，北京可在前沿及共性基础技术上增加研发投入，激励企业和科研机构参与技术开发，聚集产学研用力量形成合力，力争在大数据平台级软件上实现突破；加强大数据技术研发方向的前瞻性和系统性，通过技术引进、自主研发、创新实践等培育自主的大数据技术和产品，在世界大数据行业培育核心的技术水平；结合北京市创新中心的建设，可采取措施，支持数据开放共享，鼓励公共部门和私营部门开展数据驱动型创新，开创北京智慧型创

新新局面；此外，北京还应顺应趋势，大力推动数据开放，按敏感性对政府的公共数据进行分类，确定开放优先级，制定分步骤的数据开放路线图；积极推动政府和公共部门应用大数据技术，通过大数据增强社会管理水平，建立与市民沟通的智能行政服务，推动北京成为智慧的创新之城。

2、上海聚力发展大数据产业

作为中国的经济中心，上海各界正在大力为建设全球科技创新中心而努力。当前，在政府部门数据对外开放，由企业系统分析大数据进行投资经营方面，上海无疑是先行一步。

2013年7月，上海市科委编制并颁布了《上海推进大数据研究与发展三年行动计划(2013-2015年)》，又称“汇计划”，同时也成立了推进办公室和“上海大数据产业技术创新战略联盟”作为具体落地推进。现在，“汇计划”实施过半，科委基金已经分两批支持了20多个项目，实施效果良好。

上海推进大数据产业主要突出企业创新主体地位，建设6个以上行业大数据公共服务平台，支持6类以上大数据商业应用系统的研制，培育一批带动本地数据产业发展的行业龙头企业，探索“数据、平台、应用、终端”四位一体的新型“大数据创新”商业模式。上海希望在大数据的全新技术、产业和思维浪潮下，能够汇聚全球、全国的创新资源，充分发挥上海的国际金融、航运中心的业务、数据、人才优势，以大数据开放的心态来海纳百川和百企。上海还努力推动各级政府部门将数据对外开放，并鼓励社会对其进行加工和运用。

根据上海市经信委印发的《2014年度上海市政府数据资源向社会开放工作计划》，目前已确定190项数据内容作为2014年重点开放领域，涵盖28个市级部门，涉及公共安全、公共服务、交通服务、教育科技、产业发展、金融服务、能源环境、健康卫生、文化娱乐等11个领域。其中市场监管类数据和交通数据资源的开放将成为重点，这些与市民息息相关的信息查询届时将完全开放。这意味着企业运用大数据在上海“掘金”的时代来临，企业投资和上海民生相关的产业如交通运输、餐饮等，可以不再“盲人摸象”。

3、广东打造大数据产业基地

广东省大数据产业发展处于我国前列，并成立了我国首个地方大数据管理局——广东省大数据管理局，挂靠在广东省经信委。目前，全国首个政府数据开放规划项目落地广州，浪潮已成功竞标成为服务提供商。广东省经信委2014年12月9日公布《广东省大数据发展规划(2015-2020年)》(征求意见稿)。提出到2020年，大数据在国民经济和社会各领域广泛应用，推动全省信息化总体水平再上新台阶，迈入世界先进水平。推动制定一批国家、行业和地方大数据标准，引进和培育50家以上大数据龙头企业，形成2-3个大数据产业基地，以大数据为主要内容的现代信息服务业产值超过1万亿元，成为新兴支柱产业。

规划提出建设宽带网络、云计算基础设施，建立大数据采集机制。广东将开辟数据采集渠道，综合行政收集、网络搜取、自愿提供、有偿购买、传感收集等方式建立自动、精准、实时的大数据采集体系。加强对企业信息系统规划与大数据采集需求的融合指导，鼓励制造业企业和商业机构加强对生产经营活动中数据的采集，推动无线射频识别(RFID)、近距离无线通信(NFC)、二维码、环境监测、人体传感设备的普及应用。

规划还提出，大力优化产业组织结构和区域结构，建立大数据产业发展集聚区，促进大数据产业向规模化、创新化和高端化发展。重点支持与省内有条件的高新技术园区、电子信息产业园、软件和信息服务业产业园、数据中心基地等共建大数据产业基地，发展各具特色的大数据产品和装备制

造业、大数据软件和服务业。大力培育具有国际竞争力的骨干企业，积极发展创新型中小企业。加快大数据与云计算、物联网、移动互联网等新一代信息技术的集成应用，鼓励打破行业壁垒，推动产业链协同发展。

4、贵州大数据产业亮点频闪

2014 年，贵州省自主搭建全国首个省级政府数据统一管理、交换、共享的云服务平台——“云上贵州”统一系统平台，面向省内外政务部门和企事业单位提供云计算、云存储和宽带资源等服务。借助“云上贵州”统一系统平台，贵州省经信委协同有关部门有序推进电子政务云、交通云、旅游云等“7+N”云工程建设。贵州省还成立了由省长任组长的省大数据产业发展领导小组，成立了平台性公司——“云上贵州”大数据产业发展有限公司；率先在全国出台省级层面“规划纲要”、“实施意见”等。并通过举办“云上贵州”大数据国际年会，聚集了一批国内顶尖专家来黔论道，举办“云上贵州”大数据商业模式大赛，搭建开放创新平台。国家工信部还批复同意贵州省创建全国第一个国家级贵阳·贵安大数据产业发展集聚区。这是贵州省扎实推进大数据产业创新发展取得的又一成果。

当前，贵州大数据产业发展稳步推进、硕果累累，数据显示：2014 年贵州省组织开展了 4 次规模较大的大数据产业招商活动，签约项目 160 多个，签约投资总额超过 1700 亿元，引进了一批项目。其中三大运营商数据中心、富士康第四代产业园、中关村贵阳科技园等一批重大项目落地建设；催生了一批企业，带动智能终端、可穿戴设备等高端产品制造业，推动形成配套产业集群；集聚了一批人才，成立由中国工程院院士领衔的省大数据专家咨询委员会，集聚一批行业顶尖专家，有力推动全省电子信息产业实力实现新提升。

5、福建多措施促大数据发展

2014 年 11 月份，福建省政府出台十条措施支持大数据产业重点园区加快发展。该十条措施包括完善园区发展规划、引进培育产业龙头、推动资源汇聚开发、建设大数据创新平台、加强人才引进培养、做好园区用地保障、确保园区用电需求、强化园区网络支撑、实施财税优惠政策、提高安全保障能力方面。

早在 2014 年 6 月底，国家统计局中国统计信息服务中心与厦门市信息化局共建的“大数据研究服务基地”就在厦举行了揭牌仪式，这是国家统计局的首个大数据基地。国内首个“大数据研究实验室”也同期落户赛凡信息科技（厦门）有限公司。厦门市以大数据为重要方向的软件产业发展由此将进一步推进。此外，厦门市信息化局目前正加紧研究大数据产业发展行动规划，以期推进大数据应用、发展大数据产业、保障大数据安全。厦门市还将建设涵盖市民、法人、地理空间的全市大数据资源库，同时强化政务信息资源整合与共享，并在此基础上发展智慧交通、智慧医疗，让市民共享大数据成果。

6、重庆挖掘大数据产业商机

重庆市作为我国西部唯一直辖市，目前正积极融入国家发展格局，全力践行“两带一路”战略，特别是依托“渝新欧”大通道和长江黄金水道，进一步扩大向西向东开放，加快“两带”战略支点建设。2010 年，重庆市政府按照中央转变发展方式、调整产业结构的要求，以信息化带动工业化、城市化、农业现代化的战略思想为指导，推出了“云端计划”，经过四年努力，端产业方面已经形成“5+6+800”的世界级产业集群和生产基地。

2013 年，重庆市在全国首先出台《大数据行动计划》，计划在 2017 年左右，使大数据技术在民生服务、城市管理及全市支柱产业发展等领域广泛应用，大数据产业成为重庆市经济发展的重要增长点，成为具有国际影响力的大数据产业基地

2014 年 11 月，包括甲骨文、太平洋电信、英业达旗下的无敌科技、贝格数据等十余家大数据云计算企业“云集”重庆，斥资 16 亿元挖掘大数据产业商机。

三、2015 年中国数据分析行业发展趋势

2014 年 12 月 12-14 日，由中国计算机学会 (CCF) 主办，CCF 大数据专家委员会承办，中科院计算所与 CSDN 共同协办，以推进大数据科研、应用与产业发展为主旨的 2014 中国大数据技术大会 (Big Data Technology Conference 2014, BDTC 2014) 暨第二届 CCF 大数据学术会议在北京新云南皇冠假日酒店盛大开幕。

中科院计算所研究员，CCF 大数据专家委员会秘书长程学旗发表了大数据白皮书与发展趋势报告。《中国大数据技术与产业发展白皮书 (2014 年)》由来自 20 多家单位 60 多名专家直接撰稿，集中了 130 多名专家委员和产业联盟会员的知识和智慧，在一定程度上反映了我国大数据学术界和产业界的共识。白皮书主要介绍了大数据的背景与动态，大数据典型应用，大数据技术进展，大数据 IT 产业链与生态环境，以及大数据发展趋势与建议等内容。

同时对于 2015 年数据行业发展趋势进行了预测：

1、结合智能计算的大数据分析成为热点

包括大数据与神经计算、深度学习、语义计算以及人工智能其他相关技术结合，将成为大数据分析领域的热点。大数据分析的核心是从数据中获取价值，价值体现在从大数据中获取更准确、更深层次的知识，而非对数据的简单统计分析。要达到这一目标，需要提升对数据的认知计算能力，让计算系统具备对数据的理解、推理、发现和决策能力，其背后的核心技术就是人工智能。近些年，人工智能的研究和应用又掀起新高潮，这一方面得益于计算机硬件性能的突破，另一方面则依靠以云计算、大数据为代表的计算技术的快速发展，使得信息处理速度和质量大为提高，能够快速、并行处理海量数据。

2、数据科学逐渐兴起并带动多学科融合

数据科学作为新兴的学科，其学科基础问题体系尚不明朗，数据科学自身的发展尚未成体系。在大数据时代，许多学科表面上看来研究的方向大不相同，但是从数据的视角来看，其实是相通的。随着社会的数字化程度逐步加深，越来越来多的学科在数据层面趋于一致。可以采用相似的思想来进行的统一的研究。数据科学作为一个与大数据相关的新兴学科出现，真正支撑大数据发展的学科跨越还没有出现。针对大数据处理的理论研究上，新型的概率和统计模型将是主要的研究工具，学科基础理论的突破还难于在 2015 年出现。

3、跨学科交叉的数据融合分析成为趋势

大数据技术发展的目标是应用落地，因此大数据研究不能仅仅局限于计算技术本身。由于现有的大数据平台易用性差，而垂直应用行业的数据分析又涉及到领域专家知识和领域建模，目前在大数据行业分析应用与通用的大数据技术之间存在很大的鸿沟，缺少相互的交叉融合。因此，迫切需

要进行跨学科和跨领域的大数据技术和应用研究,促进和推动大数据在典型和重大行业中的应用和落地。

4、与热点技术的融合将产生综合性应用

大数据将与物联网、移动互联、云计算、社会计算、等热点技术领域相互交叉融合,产生很多综合性应用。近年来计算机和信息技术发展的趋势是,前端更前伸,后端更强大。物联网与移动计算加强了与物理世界和人的融合,大数据和云计算加强了后端的数据存储管理和计算能力。今后,这几个热点技术领域将相互交叉融合,产生很多综合性应用。

5、一体化融合的数据处理平台渐成趋势

大数据多样化处理模式与软硬件基础设施逐步夯实。内存计算将继续成为提高大数据处理性能的主要手段。以 Spark 为代表的内存计算逐步走向商用,并与 Hadoop 融合共存,专为大数据处理优化的系统和硬件出现,大数据处理多样化模式并存融合,一体化融合的大数据处理平台逐渐成为趋势。其中有一个观点这种多元化一定程度上成为一体化,未来大数据多样化处理模式并存并且有可能成为一体化的平台。

6、大数据安全和隐私继续成为关注重点

在连续三年关于大数据热点问题趋势的预测中,大数据安全和隐私问题每一年都是非常靠前,这个反映相关专家和用户的一种期盼,一种理解和一种关注度,但是目前在大数据的安全和隐私保护方面,以及大数据涉及到资源国家主权这层面,实际上技术层面没有比较多的长足进步,这方面有一定的问题的,所以说大数据的安全持续令人担忧。

Gartner 预测,到 2016 年,25%的全球大型企业将部署大数据分析系统,专门应对信息安全或网络欺诈,而今天这一比例仅为 8%。企业在部署大数据分析系统的前 6 个月就可获得有效的投资回报。

7、可视化技术和工具提升数据分析水平

进行分析之前,需要对数据进行探索式地考察。在此过程中,可视化将发挥很大的作用。对大数据进行分析以后,为了方便用户理解结果,也需要把结果展示出来。市场已经出现了让数据实现可视化的专业软件,它们可以让我们更容易地发现其中的规律,找到因果联系。

这些软件将变得越来越复杂并被广泛使用。这类软件市场的增长速度将是其它商务智能软件产品市场增长速度的 2.5 倍。

8、数据技术标准和人才培养需高度关注

没有标准就没有系统。应建立面向不同主题、覆盖各个领域、不断动态更新的数据建设标准,为实现各级各类信息系统的网络互连、信息互通、资源共享奠定基础。此外,数据建设的每个环节都需要依靠专业人员完成,因此,必须培养和造就一支懂指挥、懂技术、懂管理的数据建设专业队伍。

2014 年 7 月份,IBM 宣布,与中国教育系统携手开展重要合作,以实际行动应对国内蓬勃涌现的大数据与分析人才需求。在这项名为“IBM U-100”的合作计划中,IBM 将向 100 所中国高校捐赠一系列价值 1 亿美元的大数据及分析软件,并提供专业知识,支持这 100 所中国高校培养下一代数据科学家。合作包括以下三个方面:在这 100 所高校设立大数据与分析技术中心;在其中 30 所高校开设本科和硕士课程;以及在 5 所大学设立“卓越中心”。针对当前中国市场急剧攀升的大数

据与分析技能需求和人才缺口，U-100 项目每年将会涵盖 4 万名学生，帮助他们获得大数据与分析的专业知识。同时，IBM 与慧科教育旗下的知名品牌开课吧签署协议，双方共同成立“IBM 大数据与分析专区”，提供大型开放式网络课程，培养具备大数据技能和专业知识的新一代人才。

第四部分 数据分析专业委员会 2015 年工作规划

一、加大服务力度

1、进一步加大事务所业务转型扶持力度

去年全国事务所业务正式向企业经营决策领域转型，2014 年是转型的第二年，全国事务所积极主动开拓业务，通过自身的努力，部分事务所取得了可喜的成绩，为下一步加大业务转型奠定了良好的基础。

2015 年，协会将加大鼓励全国事务所加大转型力度，同时协会也会将加大扶持力度，扶持力度将从数据分析技术在企业的运用价值、业务来源、商务谈判、撰写研究报告、后期关系维护等多个方面进行扶持。

2015 年，协会将向所有年检合格的事务所提供大数据分析平台的技术支持，帮助事务所跨越技术的壁垒，进入到大数据分析领域，推进大数据分析与挖掘领域的业务。

2、有针对性的举办数据分析专业论坛，有针对性的解决地区问题

2015 年将在之前相关活动的基础上继续挑选数据分析行业发展的重点地区，结合当地行业发展、事务所经营和人才培养等实际情况，举办有针对性的专业论坛活动。活动将邀请数据分析行业专家、当地的相关政府、企业、事务所、数据分析师及准备进入数据分析行业从业的个人，就当地的实际问题进行座谈；此外协会还将邀请业务比较成功的事务所现场讲解如何开拓业务、如何开展业务谈判、如何为企业做出更科学的数据支持等内容。

3、加强对会员单位和个人的执业教育培训，进一步提升研究水平

2015 年针对会员单位和个人的执业教育，将继续增加数据分析在企业经营决策方面的技术和实战案例，帮助会员单位和个人提高专业技术、开拓业务思路，从而开展全面的数据分析业务。

4、完善 CPDA 信息系统的数据库，加大对会员单位的数据服务工作

2015 年协会将进一步与数据公司进行合作，完善 CPDA 信息系统的数据库，为会员单位提供更及时、更全面的数据信息及基础性分析报告。

5、大力扶持重点项目数据分析师事务所，起到以点带面的作用

2014 年年底协会根据对全国各地区的行业发展情况，经多方研讨在重点地区选拔出重点事务所，在 2015 年将给予点对点的大力扶持。旨在通过对重点事务所的扶持，帮助事务所在经营理念、发展方向、业务拓展、专业技术、宣传推广等方面都有更好的提升，从而达到树立榜样、以点带面促进更多事务所更健康的发展。

6、面向企业，提供专业的数据分析咨询和培训服务

为了更好地满足企业在数据分析方面的需求，解决有数据分析意识，但是人才相对短缺企业的现实问题，协会将借助协会的资源平台，联合协会专家、业内专家以及科研、教育等系统内的专家，为企业提供具有专业性高、技术要求高的咨询和培训相关的服务，以让企业享受大数据时代背景下，

数据分析所能彰显的数据价值，助力企业量化经营。

7、加大全国各地项目数据分析师事务所的走访

2015 年协会将继续加大对各地区事务所的走访，真正深入到事务所的日常经营和发展中，了解事务所的业务和经营，根据事务所的具体情况帮助事务所解决问题，从而促使事务所更好地开展业务、建立自己的品牌。

二、严格行业监管

1、加强 2015 年的各项年检工作

在 2014 年的年检中，协会加大了对会员单位整体情况的审查力度，对于不规范从业或问题较多的会员单位撤掉了其从业资质，对树立行业风气、带动行业健康发展起到的重要作用。为此，在 2015 年年检中，协会还将继续保持严格的年检制度，对会员单位进行全面审查，进一步规范行业发展。

2、举办优秀项目数据分析师事务所评选

2015 年优秀事务所评选，除了事务所在制度建设、业务经营、专业技能、宣传推广、人才培养等多方面综合评比外，将增加经营类数据分析业务占比、客户满意度和规范性从业方面内容，从而大力提高协会行业服务水平和事务所自身的行业竞争力。然后，通过协会平台，对优秀事务所进行宣传，让优秀的事务所能够发展壮大，催生在规模、实力、能力和素养上更高的事务所，形成行业标杆企业。

3、加强对会员单位出具的数据分析报告质量的检查

2015 年协会将继续不定期抽查会员单位撰写的项目数据分析报告，对报告内容的专业性、系统性、客观性、合理性、科学性等进行综合审查。对于有错误的部分将给予会员单位具体的指导。

4、加强会员入会审批流程的质量监督

随着数据分析行业在社会中获得越来越广泛的认可，各地区向协会申办项目数据分析师事务所的社会单位越来越多。为了保证数据分析行业健康、平稳地发展，协会在 2014 年的工作中，将更严格地规范事务所申办制度与流程。如对申请报告和资料进行严格地审阅，加大对申请单位的实地考察力度和全面性，加强对申请单位的背景调查，全面评估申请单位对其所在地区的数据分析行业发展的重要性，严格把控完成入会手续的时间要求等。

5、加强对个人和团体从业会员行为的监督

已经成为正式会员单位的事务所，协会也将就其成立初期的业务方向、经营理念、管理模式、宣传策略等重要方面给予指导和关注。对于新成立的事务所，协会在 2015 年也将落实一些行之有效的扶持政策，帮助事务所顺利打开经营局面，进入平稳、健康的发展阶段。

6、严格监督和查处个人及团体会员的违规现象

近年来随着行业发展，少数违法违规的现象开始显露，如恶意压价侵害事务所利益、某些社会单位通过伪造行业从业资质骗取业务等恶劣现象进行严格的监督和管理。对于违反行业自律公约、损害行业信誉的个别事务所加强管理和处罚力度，一经查处直接取缔从业资质；对于对行业发展有负面影响的社会单位，协会将采取法律手段予以追究。协会要求各事务所单位严格遵守从业规范和执业操守，保证数据分析的中立性和独立性，确保行业健康发展。

三、进一步做好行业推广工作

1、编写并发布行业发展白皮书

行业白皮书是数据分析行业每年工作的总结和对下一年度工作的规划的重要表现形式。2015 年的行业白皮书将增加更多的国内行业发展情况、会员单位经营情况、业务情况等，将成功的经验提炼，对存在的问题进行分析，使会员单位、政府机关、企业单位等更直观地了解数据分析行业，同时，对行业内从业机构的业务经营起到纲领性指导的作用。

2、与媒体合作建立行业专栏，不定期发布行业新闻和资讯

加强与媒体的合作，不定期将协会的动态、行业发展的情况等以新闻、资讯或文章等形式发布，让更多的人了解数据分析行业。也将尝试与部分媒体建立专栏，系统、全面地刊登数据分析行业有关的内容。

3、有针对性地对相关地区进行推广

在 2015 年，协会将继续挑选一些经济比较发达、具有地区特色等地区作为数据分析行业发展的重要地区，在当地开展一系列的宣传推广，提高行业的知名度，扩大行业在当地的影响力，为会员单位和需求单位搭建桥梁。

4、继续加强项目数据分析师人才培养的推广

2015 年，协会将继续以开放式的合作方式带动全国项目数据分析师人才培养的推广，并将辅助各地区授权管理中心积极开展学员沙龙、论坛、技术研讨、案例研究等多类型活动，吸引更多人才成为项目数据分析师，促进地方数据分析行业的发展。



中国商业联合会数据分析专业委员会
China Data Analysis Committee