



CDAC

中国数据分析行业 年度发展报告

二零二一年度

目 录

前言	2
第一部分 2021 年全球数据分析行业发展概况	3
第一章 全球数据分析行业的发展现状	3
第二章 全球数据分析行业的发展趋势	14
第二部分 2021 年中国数据分析行业发展概况	22
第三章 2021 年中国数据行业发展现状及趋势	22
第四章 2021 年中国数据分析师事务所发展状况	26
第五章 2021 年数据分析人才培养概况	30
第三部分 中国数据分析应用领域应用发展及需求分析	35
第六章 中国数据分析应用领域应用发展及需求分析	35
第四部分 中国商联数据分析委 2022 年工作规划	41
第七章 中国商联数据分析委 2022 年工作规划	41

前 言

中国商业联合会数据分析专业委员会，简称中国商联数据分析委(英文译名:China Data Analysis Committee, China General Chamber of Commerce, 缩写 CDAC, 以下简称“我会”), 成立于 2008 年 4 月, 是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正式批准和登记的中国数据分析行业组织。是以数据分析师及数据分析师事务所等从事与数据分析行业相关的团体与个人自愿组成的全国性数据分析行业组织, 是中国较早坚持发展专业型、靠专业引领市场的行业组织。

随着数字经济在全球加速推进以及 5G、人工智能、物联网等相关技术的快速发展, 数据已成为影响全球竞争的关键战略性资源。伴随数字经济快速发展, 数据这座巨大“宝藏”正显示出前所未有的使用价值和发展潜力。人类社会发展的历史经验表明, 每一次经济形态的重大变革, 往往催生并依赖新的生产要素。正如劳动力和土地是农业经济时代主要的生产要素, 资本和技术是工业经济时代重要的生产要素, 进入数字经济时代, 数据正逐渐成为驱动经济社会发展的新的生产要素。

“十三五”期间, 我国大数据产业迅猛发展, 政策环境日益优化, 产业规模稳步提升, 产业价值充分释放, 管理机制不断健全。“十四五”规划将“加快数字化发展, 建设数字中国”作为独立篇章, 明确指出大数据是七大数字经济重点产业之一, 这是国家自《促进大数据发展行动纲要》颁布以来, 对大数据发展做出的又一重要战略部署, 为以大数据为重点的数字产业带来了新的发展契机。

2021 年 11 月期间, 我会对数据分析行业发展情况及对专业从业机构——数据分析师事务所进行了一次全国范围内的调研。这是我会针对行业的自身发展和市场环境变化出现的新情况、新要求、新变化所做的行业调研。根据调研情况, 我会编写了 2021 年度《中国数据分析行业年度发展报告》(以下简称“报告”), 主要目的是希望能够分析行业发展的规律和趋势, 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, 为行业发展提供方向性指引, 发挥数字经济的新引擎作用, 为数字中国建设赋能添力。

本报告主要分成以下四个部分:

第一部分: 2021 年全球数据分析行业发展概况

第二部分: 2021 年中国数据分析行业发展概况

第三部分: 中国数据分析应用领域应用发展及需求分析

第四部分: 中国商联数据分析专业委员会 2022 年工作规划

需要特别说明的是, 在编写本报告时, 我们主要依据了我会在全国范围内通过问卷调研、访谈、事务所实际运营数据收集等调研方法所取得的资料和数据, 以及相关行业新闻、行业报告、专家访谈等二手资料通过深度分析总结进而得出的结论。

本报告仅作为行业内部使用, 不代表任何国家管理部门观点和意见。

第一部分 2021 年全球数据分析行业发展概况

第一章 全球数据分析行业的发展现状

一、全球数据分析产业发展分析

数据是数字经济的核心内容和重要驱动力，数字经济是数据通过信息技术分析产生价值的全方位体现。关于数字经济的定义，许多国际机构和组织都做出了概括，以 2016 年 G20 杭州峰会发布的《二十国集团数字经济发展与合作倡议》中的定义最具代表性“数字经济是指以使用数字化的知识和信息作为关键生产要素、以现代信息网络作为重要载体、以信息通信技术的有效使用作为效率提升和经济结构优化的重要推动力的一系列经济活动。”

作为衡量数字经济发展水平的重要统计标准，2021 年 5 月中国国家统计局公布了《数字经济及其核心产业统计分类（2021）》（以下简称《数字经济分类》），首次确定了数字经济的基本范围。从“数字产业化”和“产业数字化”两个方面，确定了数字经济的基本范围，将其分为：数字产品制造业、数字产品服务业、数字技术应用业、数字要素驱动业、数字化效率提升业 5 大类。其中，前 4 大类为数字产业化部分，主要包括计算机通信和其他电子设备制造业、电信广播电视和卫星传输服务、互联网和相关服务、软件和信息技术服务业等，是数字经济发展的基础。第 5 大类为产业数字化部分，是指应用数字技术和数据资源为传统产业带来的产出增加和效率提升，是数字技术与实体经济的融合。

数据分析行业目前侧重于产业数字化部分，旨在新一代数字科技支撑和引领下，以数据为关键要素，以价值释放为核心，以数据赋能为主线，对产业链上下游的全要素数字化升级、转型和再造的过程。

1. 全球数据量产生规模分析

近年来，全球正大步迈向大数据新时代，数据的高效存储、处理和分析等需求也越来越旺盛。在此背景下，行业大数据得以高速发展，应用于各个领域。根据国际数据公司（IDC）的监测数据显示，2013 年全球大数据储量为 4.3ZB，2014 年和 2015 年全球大数据储量分别为 6.6ZB 和 8.6ZB，增速每年保持在 30% 以上，2016 年和 2017 年全球大数据储量分别为 16.1ZB 和 21.6ZB，2018 年和 2019 年全球大数据储量达到 33ZB 和 41ZB，国际数据全球机构 Statista 的统计，2020 年全球数据产生量达 59ZB，并将在 2024 年增至 175ZB，到 2035 年，这一数字或将达 2142ZB，较 2020 年增长约 35 倍。

随着数据量呈爆发式增长的便是由数据所创造的经济价值。目前，并没有权威机构对 2020 年全球各国数据经济发展进行详细统计及比较，但从国际数据公司 IDC2019 年的预测来看，到 2025 年全球数据市场会形成三大板块，一是中国数据市场，占有率约为 28%；二是美国市场，占有率约为 17.5%；三是 EMEA 地区（欧洲、中东与非洲地区）占有率约为 27.6%。

图表 1:2013-2024 年全球大数据储量及其增长预测情况 (单位: ZB)

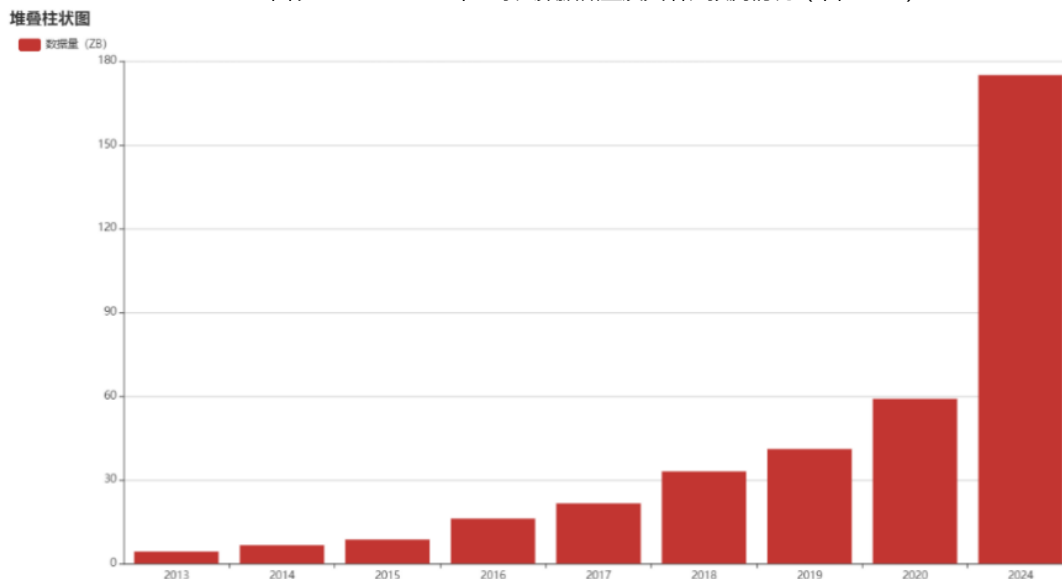


图: 通过 Datahoop 智能分析平台呈现

2.全球数字经济市场规模分析

•数字技术投资规模情况

疫情深刻改变了经济社会增长方式、国际分工合作态势以及全球竞争格局。为应对经济下行压力、应对国际格局重塑挑战,各国加快政策调整,全球数字经济加速发展。以美国、日本、中国等为代表的国家围绕人工智能、量子信息科学、先进通讯网络(5G)等关键领域,持续巩固创新并先行投资。

美国通过《2021 美国创新与竞争法案》,在《无尽前沿法案》部分,针对可能对国际科技竞争格局产生重大影响的十大关键技术领域,即人工智能与机器学习、高性能计算、量子计算和信息系统、机器人、灾害预防、先进通信、生物技术、先进能源技术、网络安全和材料科学,提出投入 1000 亿美元聚焦技术研发。

日本着手开发使用人工智能(AI)的多语言同时翻译程序,将投资约为 110 亿日元,完善研发基地,力争到 2025 年前完成开发,并在 2025 年就行的大阪世博会上投入使用。

英国先后发布《产业战略:人工智能领域行动》《国家计量战略实施计划》等一系列战略行动计划,截至 2020 年 12 月英国政府已向包括虚拟技术在内的沉浸式新技术研发投入 3300 万英镑、向数字安全软件开发和商业示范投入 7000 万万英镑、向下一代人工智能服务等投入 2000 万英镑的研发经费。

•大数据中心发展情况

根据 Statista 最新发布的统计数据,全球大数据中心主要集中在美国、中国及日本。截止 2020 年末,美国大数据中心数量占全球的比例达到 39%,中国占比达到 10%,日本为 6%。同时,由 2017-2020 年全球大数据中心数量的分布变化趋势来看,中国的份额越来越大,说明中国大数据产业的潜在空间巨大。美国和中国继续占主要云和互联网数据中心站点的一半以上。紧随其后的是日本、德国、英国、澳大利亚、加拿大、爱尔兰和印度,合计占总数的 25%。中国从 2G 跟随、3G 追赶、4G 并跑到 5G 领先,在一些领域成为

科技创新的前沿部队。中国 2020 年底已建成 71.8 万个 5G 基站，目前已经超过 115 万个，2025 年将实现中国境内全覆盖。未来的世界，5G+强大算力，或者 6G+量子计算，将构成巨大的不受边界限制的无垠网络空间。

•不同机构测量和预估的数字经济规模和增长情况

目前学术界与各国的研究机构开发了各种不同的指标来衡量数字经济的发展程度。对于数字经济规模的测算，通常有两个口径的指标，一是测算与数字经济技术直接相关的信息通信部门（ICT 部门）的增加值；二是测算包括传统产业数字化在内的广义数字经济体量。由于不同机构测算衡量体系存在差异，本白皮书引用了不同机构的数字经济发展测算指标来反映国内外数字经济发展程度。

中国信息通信研究院发布的《全球数字经济白皮书》显示，测算的全球 47 个国家数字经济增加值规模在 2020 年达到了 32.6 万亿美元，同比增长 3.0%，占 47 个国家 GDP 的比重上升到 43.7%。2020 年美国数字经济仍然蝉联全球第一，规模达到 13.6 万亿美元，全球比重高达 41.7%。中国居第二位，规模为 5.4 万亿美元。德国、日本和英国分列第三至第五。值得注意的是，德国、英国、美国数字经济在 GDP 中的占比均超过了 65%，中国比重不到 40%，但中国以 9.6%的增速位居全球第一。

根据 IDC 数据，2020-2024 年全球大数据市场规模五年内约实现 10.4%的复合增长率，预计 2024 年全球大数据市场规模约为 2983 亿美元，据此测算，2020 年全球大数据市场规模约为 1959 亿美元，到 2026 年，全球大数据市场规模将超过 3600 亿美元。

根据 IDC《2021 年 V1 全球大数据支出指南》数据，2021 年全球大数据与分析(BDA)解决方案支出规模预计将达到 2157 亿美元，比 2020 年增长 10.1%。

从地域来看，美国是最大的市场，2021 年 BDA 支出超过 1100 亿美元，其次是日本和中国，BDA 支出预计分别达到 124 亿美元和 119 亿美元，另一个预计 2021 年 BDA 支出超过 100 亿的国家是英国。

3.全球数字化人才发展分析

数字经济在全球经济增长中扮演着越来越重要的角色，以大数据、人工智能为代表的数字技术与传统产业的深度融合释放出巨大能量，成为引领经济发展的新动力。伴随着数字化转型的深入推进，各行各业对数字技能人才的需求急剧增长，人才短缺已经成为影响数字经济发展的关键因素。所谓数字化人才指拥有数据化思维，有能力对多样化的海量数据进行管理和使用，进而在特定领域转化成为有价值的信息和知识的跨领域专业型人才。

•数字人才向传统行业渗透，主要集中在制造、金融、消费品、公司服务四大行业

《全球数字人才发展年度报告（2020）》研究显示相对于 ICT 行业（包括软件与 IT 服务和计算机网络与硬件），数字人才在非 ICT 行业（制造、金融、消费品等 22 个传统行业）比例更高，其中，洛杉矶、纽约、香港、阿联酋、伦敦非 ICT 行业数字人才占比位居前五，且均超过 80%。班加罗尔、杭州、北京、旧金山湾区、南京、深圳、都柏林 ICT 行业数字人才占比位居前列，且均超过 30%。

在全球主要的创新型城市中，非 ICT 行业的数字人才在制造、金融、消费品、公司服务四大行业相对集中。制造业数字人才占比最高的五大城市包括：苏州、慕尼黑、上海、深圳、南京；金融业数字人才占比最高的五大城市包括：香港、伦敦、多伦多、纽约、新加坡；消费品行业数字人才占比最高的五大城市包括：深圳、广州、米兰、香港、巴塞罗那；公司服务业数字人才占比最高的五大城市包括：伦敦、华盛顿、阿姆斯特丹、米兰、香港、布鲁塞尔。制造和消费品行业数字人才占比亚太和欧洲相对领先，公司服务业数字人才占比欧洲相对领先，金融业数字人才占比全球各区域比较均衡。

•数字技能分析各地区差异较大

在全球主要的创新城市（地区）中，数字技能正表现出越来越高的代表性，但不同地区差异较大。中国大陆和印度创新城市代表技能以数字技能为主；欧洲和亚太地区其他创新城市代表技能既包括丰富的数字技能，也包括丰富的行业技能；北美地区代表技能以行业技能为主。

颠覆性数字技能是推动数字化深层次转型的核心力量，旧金山湾区、班加罗尔颠覆性技能的渗透率最高，在材料科学、人工智能等多个领域均处于全球引领地位。阿联酋、新加坡、上海、巴黎、慕尼黑、伦敦、纽约、波士顿、柏林、巴塞罗那等处于第二梯队，在单个（或少数几个）颠覆性技能领域具有突出的人才优势，例如新加坡的材料科学、机器人、基因工程，上海的材料科学、纳米技术和机器人，巴黎的材料科学和航空航天工程等技能具有较强的人才优势。

北美、欧洲和亚太地区颠覆性数字技能呈现差异化发展趋势。北美地区城市颠覆性技能的渗透率较高，且处于全球引领地位；欧洲地区整体排名都比较靠前，其中德国慕尼黑在航空航天、法国巴黎在航空航天和材料科学均具备领先优势；亚太地区印度班加罗尔、阿联酋、新加坡排名较高，而中国城市在颠覆性技能渗透率上的排名相对落后。

•欧洲城市和亚太城市对数字人才最具吸引力，中国城市数字人才吸引力有待提高，各地区间数字人才流动差异较大

从全球数字人才吸引力来看，欧洲城市和亚太城市居于前列。吸引力前五的城市依次是都柏林、柏林、新加坡、深圳、米兰，而北京、南京、广州、芝加哥、波士顿五个城市处于净流出状态。从数字人才流动量来看，班加罗尔、阿联酋、新加坡、旧金山湾区、纽约是全球数字人才流动中心枢纽。

北美和亚太地区的数字人才主要在本地区流动；北美地区和亚太地区数字人才双向流动频繁；欧洲的数字人才主要流向亚太地区。

4.全球数据分析行业的深度研究能力

为了从数据中生成可操作的实时见解，各行业组织已将人工智能和机器学习视为其增长战略的核心技术，希望通过人工智能和机器学习部署和深研等多种方法获得所需的结果。

机器学习是一种能够赋予机器学习的能力以此让它完成直接编程无法完成的功能的方法。最基本的做法，是使用算法来解析数据、从中学习，然后对真实世界中的事件做出决策和预测。与传统的为解决特定任务、硬编码的软件程序不同，机器学习是用大量的数据来“训练”，通过各种算法从数据中学习如何完成任务。

从提高运营效率到实现持续创新，机器学习已经成为企业发展的关键。然而，IDC 发布的《2020 年人工智能战略观察：执行摘要》显示，目前全球只有约四分之一的人工智能和机器学习技术方案被应用于生产中，这表明许多企业并未从这项技术中充分受益。企业机构唯有快速实施和扩展机器学习模型，使其能够支持整个企业内的各种应用，才能充分运用机器学习挖掘实际业务价值。

机器学习技术必将推动企业机构的变革，目前，许多机器学习应用已经为企业机构带来了实际的业务成果。机器学习可以实现流程自动化，发现新洞察，从而帮助企业创造新产品或增强现有产品及服务，提供更好的客户体验。然而，糟糕的数据质量和缺乏对相关数据的访问可能会阻碍企业对于机器学习的应用。因此，企业机构还需完成全方位运营转型，具备建立和开发机器学习模型以及部署和运营整个模型的能力。发掘机器学习的潜力。

5.全球数据分析交易现状

根据《贵阳大数据观山湖公约》，数据交易是将数据作为一种资产进行商品化后交易。首先，通过收集数据、加工处理数据、分析得出结论，使数据成为产品；然后，结合商业应用场景对产品经过再处理、再分析得出可视化结果；最后，通过合理的数据交易规则将数据产品和可视化结果出售给数据需求方。

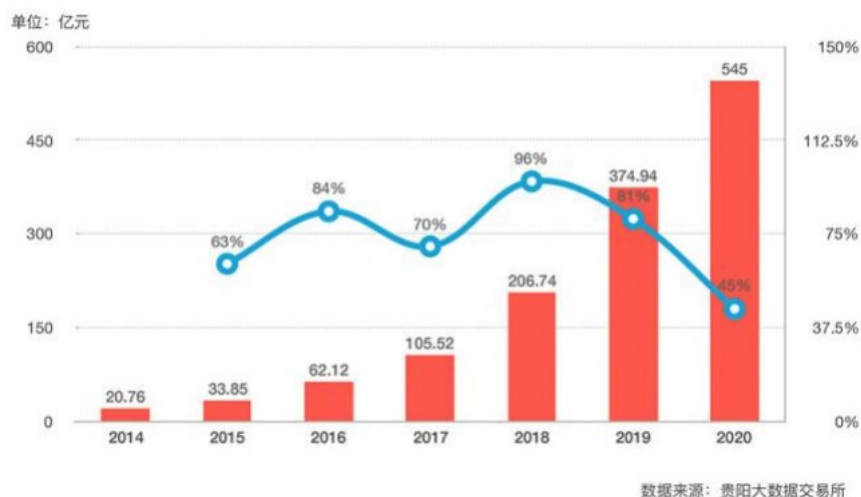
•各国数据交易模式各有特点

美国的数据交易市场主要有：（1）用户把自己的个人数据贡献给数据平台，数据平台给用户一定数额的商品、货币、服务等对价利益；（2）数据平台以中间代理人身份为数据提供方和数据购买方提供数据交易撮合服务；（3）数据平台以数据经纪商身份，收集用户个人数据并将其转让给他人。

欧盟通过“数据生产者权利”鼓励（特殊情况下强制）企业授权第三方访问其数据，促进数据流通和增值，同时基于 FRAND（公平、合理和不带歧视性）条款建立数据许可框架，为中小企业和初创公司提供更公平的数据利用机会。

日本通过开展大型科技企业数据垄断审查等一系列举措，推动形成政府、数据交流平台、数据银行、数据流通协会、数据流通运营商和国内外企业共同参与的数据流通市场。

中国数据交易一直处于探索阶段。2015 年，国务院在《促进大数据发展行动纲要》率先明确提出“引导培育大数据交易市场”；2016 年，工信部在《大数据产业发展规划 2016-2020》文件中进一步明确了大数据交易的发展目标、建设路径和保障措施；2019 年，十九届四中全会决议，首次将数据增列为生产要素，要求建立健全由市场评价贡献、按贡献决定报酬的机制；2020 年，国家《关于构建更加完善的要素市场化配置体制机制的意见》，明确提出要加快培育数据要素市场。2021 年，《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》，明确提出建立健全数据产权交易和行业自律机制，培育规范的数据交易平台和市场主体，发展数据资产评估、登记结算、交易撮合、争议仲裁等市场运营体系。上海、贵州等地方政府从 2015 年开始，一直在探索大数据交易机制，上海数据交易中心、贵阳大数据交易所等一批政府背景平台已经建成并投入使用，但也面临交易清淡等现实问题。随着数据要素明确成为生产要素，北京国际大数据交易所、北部湾大数据交易中心、粤港澳大湾区数据平台等交易所纷纷成立，数据交易所的建设热潮再次出现。阿里、腾讯、百度、京东、数据堂等企业也纷纷建立数据交易平台。



•数据交易存在的问题

数据产权确权问题：权属关系是生产要素至关重要的属性，也是数据交易的基础。现有的物权法、合同法、知识产权法等法律规范在信息革命和数字经济时代未能及时完善，加上数据具有数字化、易复制、易传输、易修改等特点，使数据确权更加复杂困难，原始数据和衍生数据的产权界定困难，而且对数据产权规制过严可能会限制企业创新。

数据定价机制问题：目前数据定价机制取得一定成果，但是由于数据自身存在的无形性、可复制性、可共享性等特点，使数据定价过程中存在以下三个难点：第一，数据买卖双方存在信息严重不对称，对数据质量、数据成本和收益难以计量，当数据交易主体较少时，自由市场难以定价，存在价值价格的偏离；第二，数据要素的生产过程涉及感知、采集、传输、存储、计算等多个环节，参与主体多元化使得数据要素的确权和定价尤为困难；第三，数据要素的异质性显著，包括数据结构异质性、搜集主体各不相同、价值高度依赖使用场景等，很难有一套统一的定价方法，容易形成“千人千价”的价格市场局面。

数据价值化问题：数据价值化按照资源化、资产化、资本化加速推进。目前数据价值化存在以下难点，第一，随着数据量增大、维度增多，数据复杂化、定义模糊化程度越来越高，对数据进行清洗、脱敏、分析、建模、可视化处理加工深度和难度增大，而中小企业不存在大量的、复杂的数据处理能力；第二，数据产品商业化程度低，数据与具体的业务融合程度低，未形成由市场决定的数据产品包装、数据产品售卖体系；第三，数据价值随时间变化，数据具有高时效性的特点，用户的需求和兴趣具有转瞬即逝的特点，数据一旦无法得到有效及时的处理便会丧失预期的价值。

数据安全保护问题：在数据交易过程中，第一，数据黑色产业链日益成熟，数据攻击、黑客侵入、数据破坏、数据泄露、数据篡改、数据被恶意利用和买卖等现象频发；第二，个人数据泄露事件不断发生，非法获取、泄露和非法出售个人信息行为不断发生；第三，数据垄断、数据和算法导致的人们对其过分“依赖”问题、社会“被割裂”等伦理问题以及“爬虫调查风暴”等数据安全事件频发。

6. 数据跨境流动和发展分析

2021 年联合国贸易和发展会议发布《2021 年数字经济报告(数据跨境流动和发展:数据为谁而流动?)》, 从全球视角分析了数字经济在跨境流动方面的现状和未来的发展。

首先, 数据流量全球增长迅速。2022 年的全球互联网协议流量 (包括国内和国际流量) 将超过截至 2016 年的互联网流量之和。2020 年全球互联网带宽提高了 35%, 是 2013 年以来增幅最大的一年。预计每月的全球数据流量预计将从 2020 年的 230 艾字节激增到 2026 年的 780 艾字节。

其次, 数据驱动的数字经济在全球国家和地区表现出极大的不平衡。最不发达国家使用互联网的人口占比仅 20%, 而且在国家内部, 农村与城市之间以及男女之间在网络使用上存在明显的鸿沟。从参与数据驱动的数字经济并从中受益的能力来看, 美国和中国脱颖而出。全世界的超大规模数据中心有一半在这两个国家, 5G 普及率最高, 占过去五年人工智能初创企业融资总额的 94%, 占世界顶尖人工智能研究人员的 70%, 占全球最大数字平台市值的近 90%。发达国家和发展中国家之间存在很深的传统数字鸿沟, 体现在互联网连接、接入和使用方面, 对发展构成经常性的挑战。随着数据作为一种经济资源以及跨境数据流动发挥越来越大的作用, 发展中国家可能处于从属地位, 因为数据及相关价值获取集中在几个全球性的数字企业和其他控制数据的跨国企业中。发展中国家可能会沦为全球数字平台的原始数据提供方, 要想获得数字智能则必须付费, 尽管这些智能来自它们自己提供的数据。

第三, 虽然数据在数字经济发展中十分重要, 但全球对数据的概念还没有达成一致的理解。然数据与国际贸易有密切的联系, 但是跨境数据流动本身既不是电子商务, 也不是贸易, 不应纯粹作为电子商务加以管制, 此外数字经济中的信息不对称、数据伦理、数据过滤和审查问题等全球都没有共识性文件。

第四, 世界范围内, 不同国家和地区政策差异较大。美国侧重于私营部门对数据的控制, 欧盟主张基于个人基本权利和价值对数据进行控制, 全球性的数字企业在积极扩大自身的数据生态系统。各方都在争夺技术发展的领先地位, 希望可以通过控制数据和相关技术, 特别是人工智能技术, 获得经济和战略优势。这种各自为政的数据驱动实际上违背了互联网的作为一个自由、去中心化和开放网络的初衷, 数字经济的治理, 还需要全世界的政策制定者从全局出发看待数据流动, 并考虑到对人权、国家安全、贸易、竞争、税收和互联网整体治理可能产生的影响, 共同去探索造福所有人的全球解决方案。

第五, 数据驱动的数字创造了全球性机遇, 也带来了全球性挑战, 需要建立全球数据治理框架, 以辅助利用数据驱动的数字经济促进发展的国家政策, 并与其保持一致。同时需要具有灵活性, 以便所有国家, 不论其数字化程度和从数据中获益的能力如何, 在数据驱动的数字经济下制定和实施其发展战略时, 都能获得必要的政策空间。数字经济下的国家政策或发展战略若没有全球视角, 就很可能失败。

二、各国数据战略和发展情况

1、美国

美国是全球最早布局数字化转型的国家, 多年持续关注新一代信息技术发展及影响, 奠定了其数据战略和发展的领先地位。近年来, 美国进一步聚焦大数据和人工智能等前沿技术领域。

2012 年 3 月 22 日，奥巴马政府宣布推出“大数据的研究和发展计划”，该计划承诺投资 2 亿美元推动和改善大数据相关的收集、组织和分析工具及技术。这是大数据技术从商业行为上升到国家科技战略的分水岭，表明大数据正式提升至战略层面。

2014 年 5 月美国发布《大数据：把握机遇，守护价值》白皮书，对美国大数据应用与管理的现状、政策框架和改进建议进行了集中阐述。从《白皮书》所代表的价值判断来看，美国政府更为看重大数据为经济社会发展所带来的创新动力，对于可能与隐私权产生的冲突，则以解决问题的态度来处理。

2019 年 12 月 23 日，美国白宫行政管理和预算办公室（OMB）发布《联邦数据战略与 2020 年行动计划》。

以 2020 年为起始，联邦数据战略描述了美国联邦政府未来十年的数据愿景，并初步确定了各政府机构在 2020 年需要采取的关键行动。

2021 年 1 月新美国安全中心（CNAS）发布《设计美国的数字发展战略》报告，指出技术政策逐渐成为外交中的争论焦点，美国及其盟友选择的数字发展道路将对未来美国国家安全发挥关键作用。面对发展中国家在 5G、云计算、电子商务等新兴技术领域的快速发展，报告提议美国政府应升级其数字发展战略，并提出了四项指导原则及五项策略。

美国信息技术和创新基金会（ITIF）发布《美国全球数字经济大战略》报告。提出为保持美国在全球科技领域的领导地位，美国政府必须制定一项全面的宏大战略来指导美国的 IT 和数字政策，战略首要目标应是限制中国在 IT 和数字领域的全球主导地位。美国面临的风险是，包括欧盟在内的世界许多地区可能会与美国的 IT 和数字利益相抵触，从而导致多对一的环境，带来不利后果。美国必须以各种方式争取志同道合的国家支持本国的利益，并在必要的时候向其施压。

美国是元宇宙领域的先行者，在成立了第一家元宇宙公司 Roblox 之后，2021 年 10 月底美国互联网巨头 Facebook 现任 CEO 马克扎克伯格宣布，Facebook 更名为 Meta，在全球范围内引爆了元宇宙。之后，微软、谷歌、英伟达等美国企业先后宣布布局元宇宙。

2、英国

大数据发展初期，英国在借鉴美国经验和做法的基础上，充分结合本国特点和需求，加大大数据研发投入、强化顶层设计，聚焦部分应用领域进行重点突破。2013 年英国政府加大了对大数据领域研究的资金支持，提出总额 1.89 亿英镑的资助计划，包括直接投资 1000 万英镑建立“开放数据研究所”。

2017 年英国政府发布《英国数字战略》，提出了多项数字化转型战略，包括连接战略、数字技能与包容性战略、数字经济战略、数字转型战略、网络空间战略、数字政府战略和数据经济战略，为数字化转型作出全面部署。同年，英国发布《产业战略：打造适合未来的英国》，旨在与产业界合作开展科技创新和应用研究，实现人工智能技术的创新应用，将英国建设为全球人工智能与数据驱动的创新中心。2018 年，英国政府出台《产业战略：人工智能领域行动》，再次强调支持人工智能创新以提升生产力，使英国成为全球创立数字化企业的最佳之地。

2020 年 9 月 9 日，英国数字、文化、媒体和体育部（DCMS）发布《国家数据战略》，支持英国对数据的使用，帮助该国经济从疫情中复苏，并在 2020 年 12 月之前面向社会进行公开咨询。《战略》将设定五项“优先任务”，研究英国如何利用现有资源来促进企业、政府和公民社会对数据的使用。政府必须充分利用这些任务来发挥数据带来的机会，创建一个蓬勃发展、快速增长的数字行业，以促进经济发展。

五项任务分别为：

- (1) 释放经济中的数据价值；
- (2) 确保促进增长和可信的数据体制；
- (3) 转变政府对数据的使用以提高效率并改善公共服务；
- (4) 确保数据所依赖的基础架构的安全性和韧性；
- (5) 倡导国际数据流动。

此外，《战略》还包括多项计划，如：

(1) 到 2021 年，对 500 名分析师进行公共部门数据科学方面的培训，并设立政府首席数据官进行领导，改变政府当前数据使用方式，从而提高效率并改善公共服务；

(2) 通过立法促进智慧数据计划的参与度，该计划可以使人们能够通过自身数据在电信和能源等领域获得更好的资费；

(3) 新建一个 260 万英镑的项目，解决当前数据共享中存在的障碍，并支持发展创新能力，以发现在线危害。

《战略》阐明了在英国如何释放数据的力量，为处理和投资数据以促进经济发展建立了框架。总体而言，《战略》中确定的步骤是基于英国发展现状来推动更好、更安全、更具创新性的数据使用。通过数据使用推动增长，改善社会公共服务，使英国成为下一轮数据驱动型创新的领导者。

3、欧盟

为应对未来发展，欧盟致力于平衡数据流动和广泛使用，坚持合作共赢原则，希望通过建立单一统一的数据市场，为成员国产业协同发展提供有力条件，确保欧洲在未来的数字经济中占据领先地位。

2015 年欧盟委员会启动《数字化单一市场战略》，通过一系列措施消除法律和监管障碍，着力将 28 个成员国打造成统一的数字市场。此后，欧盟委员会于 2016 年 4 月出台《产业数字化新规划》，计划在 5G、云计算、物联网和网络安全等重点领域加快建立共同标准，以统筹欧盟各成员国的产业数字化转型。基于传统行业数字化转型明显滞后的现状，欧盟在整合成员国已经出台的工业数字化战略基础上，逐步形成《欧洲工业数字化战略》，以打造开放协同的创新体系、培育创新型企业为路径，确保欧盟在产业数字化转型方面处于全球领先地位。

2020 年 2 月 19 日，欧盟委员会公布了《欧盟数据战略》，以数字经济发展为主要视角，概述了欧委会在数据方面的核心政策措施及未来五年的投资计划，以助力数字经济发展。

2021 年 3 月，欧盟委员会发布了《2030 数字罗盘：欧盟数字十年战略》（2030 Digital Compass: the European way for the Digital Decade）报告，为欧盟数字化发展提供战略建议。报告指出，欧盟的数字化发展落后于美国和中国，欧盟应该通过加强人员的数字化技能培训、加强数字化基础设施投资、鼓励企业进行数字化转型、实现公共服务的数字化等方法实现数字化转型。

4、亚洲

(1) 日本

日本的数字化进程可以追溯至 2001 年时任首相森喜朗(Yoshiro Mori)提出的“e-Japan 战略”，首次将数字化确立为日本国家级战略。为了在新一轮国际竞争中取得优势，日本制定和发布了一系列技术创新计划和数字化转型举措。

2016 年，日本发布《第五期（2016-2020 年度）科学技术基本计划》，提出利用新一代信息技术使网络空间和物理世界高度融合，通过数据跨领域应用催生新价值和新服务，并首次提出超智能社会“社会 5.0”这一愿景。互联工业作为“社会 5.0”的重要组成部分，得到政府的高度关注，日本产经省推动成立了工业价值链促进会并发布《日本互联工业价值链的战略实施框架》，提出的新一代工业价值链参考架构成为日本发展高端制造业的新抓手。2018 年 6 月发布的《日本制造业白皮书》强调“通过连接人、设备、系统、技术等创造新的附加值”，正式明确将互联工业作为制造业发展的战略目标，并通过推进“超智能社会”建设，抢抓产业创新和社会转型的先机。

2020 年 9 月，日本开展迅速的选举并交接权力之后，新任首相菅义伟(Yoshihide Suga)宣布，期待已久的数字改革将成为疫情后经济复苏的主要目标。计划日本 2021 财政年度的国家预算将超过 105 万亿日元(9940 亿美元)，重点支持数字化。

当前日本政府十分重视数字经济发展，全力推进“数字新政”。在“后 5G”信息通信基础设施、学校的 ICT 应用、中小企业信息化、ICT 领域研发等方面，加大资金投入力度，体现了其提升数字经济竞争力，全面推动社会数字化、智能化转型，激发中小企业数字时代新活力，促进经济新增长点的战略动机。日本“数字新政”具备硬件投资与软环境建设相结合、产业数字化与数字产业化相结合、数字技术各类研发投入并重等基本特征。

(2) 韩国

随着全球第四次工业革命浪潮的到来，韩国重新审视本国智能制造和信息技术的发展，大力推进智能工厂建设，为实现“制造业复兴愿景”全面谋划。2018 年 2 月，韩国政府发布《第四期科学技术基本计划(2018—2022)》，提出将无人驾驶、智能城市、VR/AR、定制化医疗保健、智能机器人、智能半导体等领域作为第四次工业革命的创新增长引擎。2018 年 12 月，韩国召开了“中小企业智能制造创新战略报告会”，提出到 2022 年建设 3 万个智能工厂及 10 个智能产业园，成立 3000 亿韩元规模的智能工厂设计建设企业基金。

2019 年 6 月，韩国公布了人工智能产业的发展目标和投资计划，提出到 2030 年建设 2000 家“人工智能工厂”，并建立数据中心促进关键软件、机器人、传感器等智能制造设施的发展。

2020 年 8 月 20 日，韩国财政部发布《基于数字的产业创新发展战略》。2018 年韩国制造业对大数据的利用率为 0.9%，与金融（20.5%）、通信（7.4%）、流通（2.2%）等产业相比利用率非常低。因此，韩国政府通过制定“数字+制造业”创新发展战略，将重点放在制造业这一韩国优势产业，提高制造业中产业数据（产品开发、生产、流通、消费等产业活动全过程中产生的数据）的利用率，以增强韩国主力产业的竞争力。

到目前为止，韩国的大数据政策主要集中在保护个人信息及建立大数据专门机构等大数据产业生态政策上，但通过此次政策，政府表明今后将考虑产业特殊性，制定针对性的大数据政策。目标实现通过数字创新跃升为世界四大产业强国之一，并计划开展：产业价值链智能化优先实现数据与人工智能应用的成功案例；民间主导推进体系；全球开放创新战略等 4 个实施战略和 9 个具体推进任务，此次战略暂不适用于所有产业数据，主要针对钢铁、半导体等企业需求，按产业类别逐步推进，并计划通过结合产业数据-个人信息-公共数据，提升数据利用率。

在全球范围内，韩国政府对元宇宙反应最快，已经成立了元宇宙协会。根据 2021 年 10 月份数据，包括三星在内的两百多家韩企已经加入了这个协会。此外，韩国首都首尔市政府宣布，将打造“元宇宙平台”，向市民提供公共服务。届时，首尔市政府将成为韩国首个提供元宇宙服务平台的地方政府。

(3) 印度

麦肯锡报告《数字化印度》显示印度是数字消费者规模最大、增长最快的市场之一，但印度企业的数字化并不均衡，存在很大发展空间。

政府部门是印度快速数字化的强大催化剂。2009 年推出的 Aadhaar 成为了全球最大的数字身份识别计划，并且加速了其他数字服务的普及。2013 年建立的商品和服务税网络（GSTN）促进了间接纳税企业的交易进入数字平台，为企业数字化运营创造了强大动力。

数字化背景下，印度互联网经济蓬勃发展，但是印度企业存在数字鸿沟，且行业内的差异高于行业间的差异。传统行业包括农业和建筑业中有数字化程度相当高的企业；而一般认为拥有更多数字化领先者的部门如信息通信技术，仍有公司处于数字化末端。

2017-2018 年，印度信息技术-商务绩效管理（IT-BPM），数字通信和电子制造等核心部门约为 1700 亿美元，占 GDP 的 7%。到 2025 年，IT-BPM 的增值贡献可能在 2050 亿美元到 2500 亿美元之间，电子产品制造业的贡献将为 1000 亿美元到 1300 亿美元之间，数字通信的贡献将为 500 亿至 550 亿美元之间。这三个部门的增值贡献总额在 3550 亿美元至 4350 亿美元之间，可占印度 2025 年 GDP 的 8%至 10%。

目前印度正在积极推动数字化转型，其中的 5 大方向为：打造完善的数字化公共数据基础设施、利用数字化技术促进良政、发展新创企业生态体系、推动洁净能源及生态保护，以及发展在地电信产业的 5G 与 6G 能力。

印度有 8 亿人口使用网络、7 亿 5000 万人拥有智能手机，印度政府正积极做用宽带网络连接 60 万座村庄。印度现阶段正聚焦于发展硬件，为发展半导体产业铺路。

(4) 新加坡

新加坡为加快数字化步伐，推出一系列“数字化蓝图”勾勒经济社会的整体转型发展计划，以服务业转型为重点寻求数字化新变革。2014 年，新加坡公布“智慧国 2025”发展蓝图，在该计划推动下，电子政务、智能交通等领域均达到全球领先。2016 年，新加坡公布产业转型蓝图，提出涵盖能源化工、航空业、医疗保健、金融、教育等 23 个具体行业的转型措施及发展目标，2018 年推出“服务与数字经济蓝图”，重点提升本国服务业领域的数字创新能力。为确保公民从数字化转型中获益，新加坡政府发布“数字化能力蓝图”，进一步提升本国人数字技能，并设立科技卓越中心培育数据分析、人工智能和网络安全等领域人才。

第二章 全球数据分析行业的发展趋势

一、全球数据分析产业发展趋势分析

1. 开源成为技术创新主要模式

开源已成为全球数字科技创新的大趋势。经过近 40 年的实践发展，开源逐渐成长成为一种强大的技术创新模式，并从最初的软件行业走向了硬件、芯片等多个领域。开源商业模式也在逐渐成熟。如今，新技术在开源、新架构在开源、新平台也在开源，就连顶尖的研究成果很多也都以开源形式发布。开源逐渐成为全球科技进步至关重要的创新渠道，从本质上讲，开源是数字经济的典型生产协作方式。

2021 年 1 月 5 日，美国国会研究服务局提出要重新审视对中国获得美国“开源技术”问题，并考虑对“开源技术”实施出口管制。红帽公司 2021 年发布的全球《企业开源现状》调查报告显示“最具创新能力的企业都在使用开源”。2021 年 3 月 13 日，我国发布的《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》指出，支持数字技术开源社区等创新联合体发展，完善开源知识产权和法律体系，鼓励企业开放软件源代码、硬件设计和应用服务。可以定，数字开源技术所构成的开源商业生态已经成为数字经济时代国家竞争力的重要关键创新应用平台，是一个国家或地区参与数字世界治理并发挥国际影响力的主要领域之一。其中“深度学习框架等开源算法平台构建”被列入新一代人工智能科技前沿领域攻关内容。

开源相当于工业时代的道路和桥梁，是数字经济的重要公共基础设施之一。开源的价值体现在：

一是直接经济效益，极大降低企业创新创业的成本；不管是工业互联网，还是智能制造，借助开源平台、开源项目，人们可以便捷地获得公共产品和服务，无需再做重复性劳动，这极大降低了企业及个人创新创业的门槛和成本，提高了数字商品资源配置的效率。

二是间接经济效益，可以快速形成标准及生态系统，极大降低社会互操作性成本；企业内部不同系统、企业与企业之间、企业与政府之间数据可以快速实现互通，无需第三方插件，极大降低不同系统（不同技术路线）之间的转换及数据链接成本。

三是极大激发个体参与创新创业的积极性，劳动者不再受物质及时空制约就可以参与数字商品生产协作。开源以开放、对等、分享以及全球运作方式，彻底颠覆了工业经济时代的商业逻辑，并深刻影响到人们的思维方式、企业商业模式和社会经济运行规则，开启了数字经济时代的新商业文明。

2020 年，欧盟委员会发布的研究报告指出，开放源码最显著的优势并不一定与节约成本或投资回报有关，而是开放标准和互操作性，以及独立于专有供应商的优势。安卓操作系统采用开源模式，获得了手机 87% 市场份额（IDC，2019 年）。丰田、特斯拉、大众、吉利等汽车制造商也在走开源之路，通过开放其标准及专利，建立自己的生态系统（产业链），以获取竞争优势。相对于美国，我国开源创新还有较大差距。我国是最大的开源技术“消费国”，但国际主要开源项目均由西方国家所主导，我国在“开源技术”领域缺乏话语权和核心掌控力，已经为我国软件供应链科技的安全敲响了警钟，也警示我国要高度重视开源数字经济的战略性作用。

2. 产业数字化转型在加速

全球产业信息基础和环境大幅加强，海量数据源源不断地持续产生，进一步推动劳动、技术、资本、市场等要素互联互通，带动数字化转型从被动变为主动，将数字化从用于提高生产效率的被动工具，转变为创新发展模式、强化发展质量地主动战略，将数字化从局部生产经营环节的参数获取和分析，转变为对全局流程及架构的诠释、垂构及优化，将数字化从聚焦于单一环节、行业和领域，转变为对产业生态体系的全面映射。

数字化转型加速推动产业链各环节及不同产业链的跨界融合，实现了组织架构和商业模式的变革重塑，构建起核心优势独具特色、运作体系不拘一格的各大平台，将企业间的竞争重点从产品和供应链层面推向生态层面，对数字化转型底层技术、标准和专利掌控权的争夺更为激烈。同时，数字化转型的快速推进为供需实时计算匹配提供了坚实基础，并通过高频泛在的在线社交，以及渐趋完善的信用评价体系，为部分产业提供了未有效配置资源的低成本共享渠道，弱化“所有权”而强调“使用权”，促使共享经济快速兴起。

产业数字化转型直接带动了技术开源化和组织方式去中心化，知识传播壁垒开始显著消除，创新研发成本持续大幅度降低，创造发明速度明显加快，群体性、链条化、跨领域创新成果屡见不鲜，颠覆性、革命性创新与迭代式、渐进式创新相并行。产业创新主体、机制、流程和模式发生重大变革，不再受到既定的组织边界束缚，资源运作方式和成果转化方式，更多地依托网络在线展开，跨地域、多元化、高效率地众筹、众包、众创、众智模式不断涌现，凸显出全球开放、高度协同地创新特质。

产业数字化转型地快速推进带来新兴地数字化产品、应用和服务大量涌现，对消费者的数字化资源获取、理解、处理和利用能力提取更高要素。符合用户根本需求，具备完整商业模式，持续迭代完善地各类数字化新兴产物，已开始有效引导消费者数字化技能和素养地提升及更新，更好地发掘数字化价值和享受数字化便利，逐步培育、形成及发展起新兴地数字消费群体和数字消费市场。世界各主要国家将日益高度重视对公民数字技能和素养地教育及培养，并逐步上升到维护国家在新时代打造新型核心竞争力的战略高度。

3. 前瞻技术大量涌现和现实需求尚待挖掘的市场培育问题越来越突出

以 5G、人工智能、大数据、区块链、量子计算等为代表的新一代信息技术蓬勃发展，竞相在关键环节和领域出现重大集中突破，从理论创新逐步延伸向实用创新，从科学研究加速迈进到产业应用，一批原本属于实验室范畴的前瞻性技术已经尝试性地走向市场和用户，并探索与部分应用场景进行结合的立足点和开拓方式，试图去培育、激发、引导新的市场需求。但在现实需求层面，大部分产业的工艺特性、业务流程和运营模式已经过长期沉淀相对稳定，并非简单引入和嫁接前瞻技术就能获得进一步的效益和质量提升，既需要仔细梳理和挖掘真实的需求痛点，更需要对前瞻技术的成熟度和适配性有着清晰的认知与判断。如何挖掘和培育兼具事件应用价值和示范推广意义的多类型市场需求，充分发挥前瞻技术持续迭代和演进的创新特性，这是越来越突出的现实问题。

4. 数据跨境流动是全球必然趋势

当前，世界正经历百年未有之大变局，新一代信息技术创新应用正引领新一轮科技革命和产业变革。在这场调整变革中，共享经济、平台经济、新个体经济等新业态加速兴起。尤其是在新冠肺炎疫情的突袭下，全球化进程遭遇逆流，加剧了各国之间的冲突和竞争。当前，数字空间成为大国发展与安全利益博弈的重要领域。中美欧三方鼎立的全球数字竞争版图已初现端倪，而如何在促进数据流动和共享中获得最大的发展红利，是各国实现数字发展的最终目标。在数字经济发展的今天，数据作为新的关键生产要素，只有实现在更大范围内的流动共享，才能更好地发挥对经济增长、社会发展、全球化进程的支撑推动作用。数据跨境流动已成大势所趋，各国纷纷针对本国国情和优先目标选择适当的数据跨境流动方案。受当前经济全球化和数字化的趋势影响，数据大规模的跨境传输不可避免，多国通过国内立法、签署国际协定的方式确立数据跨境流动规则。

5. 数据和新技术影响全球竞争格局

美国智库大西洋理事会 (Atlantic Council) 2021 年 5 月 26 日发布《新技术与数据的地缘政治影响委员会报告》。报告认为，新技术与数据能力将对地缘政治、全球竞争与合作产生巨大影响，未来十年全球将迎来“地缘技术十年” (GeoTech Decade)。报告认为，新技术和数据能力的发展速度、规模和复杂性在帮助或者破坏人类互联的世界方面达到了前所未有的程度。人类现在处在一个新技术和数据以新颖甚至出乎意料的方式重塑社会和地缘政治的时代。报告认为，2001-2011 年是“全球反恐十年”，2011-2021 年是对政府和大型科技公司“信任度下降十年”，2021-2031 年将是新技术和数据能力影响地缘政治、竞争和合作的“地缘技术十年” (GeoTech Decade)。未来十年将迎来“地缘技术十年”，世界将迎来新兴关键技术与数据能力决定世界地缘政治与竞争格局的十年，世界地缘政治、全球竞争和全球合作机会受到复杂影响、不断变化。

全球发展可能出现拐点，即当前的设想与方法不再适用、现有能力与优势基础减弱。但目前为止，没有一个国家或国际组织能够组建合适的治理结构，应对新兴技术的复杂性和不稳定性。因此，在未来的“地缘技术十年”，政府、行业和其他利益相关者必须共同努力，保持经济竞争力，维持社会福利和公共安全，保护人权和民主进程，维护全球和平与稳定，最终促进全球自由、安全与繁荣。因此，在未来十年，必须用新的方法开发和部署关键技术、刺激企业的创新；要培养人才，促进人才的多样性，要利用这些技术与数据能力，激发人类的独创力；要重建对国内外治理的信任，建立国际合作规范，扩大全球合作伙伴的范围；要解

决网络系统复杂而脆弱的问题，将弹性作为未来发展的必要基础支柱，维护国家和经济安全、民主价值观。而这将极大地依赖于志同道合的国家之间的合作。

6. 数据安全保障能力得到提升

世界主要国家均高度重视数据及其安全，近年来，针对数据的攻击、窃取、劫持、滥用等手段不断推陈出新，给经济、政治、社会等领域带来巨大风险。为切实保障数据安全，国际社会各方在实践中不断探索，在维护数据安全方面进行了诸多尝试。由于数据安全问题复杂，涉及主体多元，领域广泛，包括国家、企业乃至个人的各方均是数据安全维护的主要实践者，虽然各有专长或专注领域，但彼此之间又有着紧密关联，共同构成不断发展的数据安全治理生态。

• 国家层面

出于保护国家安全和公民隐私考虑，各个国家或地区不断推进数据本地化措施，对数据存储做出相应要求，尤其是发展中国家和新兴经济体。对本国数据境外传输采取相应规制，主要表现为需经数据主体同意或特定数据经审批后方可对外传输。

在数据领域的长臂管辖问题（主要源于网络犯罪的取证问题）上，网络犯罪因其数据转移更新快，通过正常审批流程周期长等问题，使得各国为了本国的司法便利而采取长臂管辖。

• 行业层面

不同行业对于数据安全类型的侧重和维护方式均有所不同，所有需要和个人交互的行业都需要保护的个人隐私数据之外，重点对各自“核心数据”采取了相应措施以提升数据安全能力。主要包括：通过加大数据安全架构优化力度，促进整个数据安全治理体系的持续优化；通过不断加强数据安全风险评估力度，将数据安全风险评估工作常态化、机制化；通过采取新技术手段，部署人工智能、机器学习、分析和其他形式的安全自动化技术，强化数据安全态势感知；同时通过加大人员数据安全能力培训，打造针对技术、自动化与安全专业人才团队建设，维护和减少系统内部漏洞以避免外部恶意攻击；通过规范数据处理流程，对于企业内部人员进行数据安全培训与道德培训以降低内部无意识泄露的风险。

• 国际层面

各国数据在跨境流动中，不断地通过机制对接，寻找流动与安全的平衡点。目前数据安全理念在国际范围内不断扩散，相应数据安全规则也在逐渐成形。虽然目前阶段各国囿于利益、价值观等方面的较大差异，短期内在联合国或 WTO 等多边机制下很难形成各方接受的单一数据流动规则，但各主要国家都在探索可能的路径。

一是对于个人数据的保护基本达成了共识。欧盟通过 GDPR 确立了被称为“世界上采用最广泛的个人数据保护模式”，大多数国家在国内立法中都在不同程度上适用其原则。美国推动 APEC 跨境隐私规则（Cross-Border Privacy Rules, CBPR）体系，增强个人信息保护机制兼容性。

二是通过“相互认定”设立数据安全“门槛”。欧盟 2018 年发布的《通用数据保护条例》(GDPR) 通过“充分性认定”规则确定数据跨境自由流动白名单国家,对具有充分保护水平的国家和地区进行充分性认定,从而确定能与欧盟进行数据自由传输的国家和地区。在一定程度上设立了数据安全保护的“门槛”,打造了一些“模板”。如韩国为与欧盟达成“充分性认定”,于 2020 年修改了其国内数据保护法。

三是力图形成兼顾发展与安全的治理框架。日本倡议的“可信数据自由流动”比较具有代表性,该框架通过 G20 机制得到美国和欧盟的认可。2019 年 6 月, G20 贸易与数字经济部长会议重申可信数据自由流动的概念,强调跨境数据流动对促进生产率、创新和可持续发展的重要性。

四是积极推出全球数据安全倡议。中国作为数据大国,积极为全球数据治理不断探索。2020 年 9 月 8 日,在北京举办的“抓住数字机遇,共谋合作发展”的国际研讨会上,中国提出《全球数据安全倡议》。在这份倡议的推动下,2021 年 3 月 29 日,中国外交部同阿拉伯国家联盟秘书处召开中阿数据安全视频会议,宣布共同发表《中阿数据安全合作倡议》,阿拉伯国家成为全球范围内首个与中国共同发表数据安全倡议的地区,体现了中阿在数字治理领域的高度共识。

二、全球数据分析技术发展趋势分析

1. 整体技术趋势

根据德勤发布的《2021 年技术趋势》,过去十年内,数字化体验、数据和分析技术以及云计算为各项技术赋能,成为全球众多企业有效推进战略和新商业模式的核心基础,这些技术已经推动了十年颠覆性变革,并且在变革中还将持续发展。

接下来十年中,数字现实、人工智能以及分布式平台正在逐步获得广泛应用,并迅速成为颠覆性驱动力。目前正在塑造下一个十年的业务创新。

而“环境体验”设想了无处不在的交互,与周围环境无缝集成,能够主动预见并满足人类需求,基于人工智能构建的指数智能,在识别和响应人类情绪、理解外部环境和执行任意任务方面,将拥有更强的能力。同时“量子技术”利用亚原子微粒的特性来解决当今超级计算机仍无法解决的复杂问题。宏观力量框架也将有助于直观显示交互、信息和计算等技术的演化发展路径。



(1) 战略支持技术

在日益复杂的竞争中，不仅战略变得日益数字化，战略制定和执行以及结果监控的流程也要数字化。当领导者被迫考虑更广泛的各种变量和未來情景时，技术驱动型战略平台有助于战略制定者更全面、更准确地思考未來的各种可能性。

(2) 核心系统激活技术

新兴技术正在帮助企业以更优的成本投入和更加方便运营管理的方式，将陈旧的核心系统向云平台进行迁移。以前大而全的 ERP 软件包，或一体化系统思路，逐步被以用户体验简化为核心，通过分布式，云化，API 接口化的一组服务能力和功能组合的架构思路所替代。

(3) 解放供应链技术

通过优化供应链的各个环节，为每位客户提供差异化服务，有助于公司找到成本和服务之间的最佳平衡点。经过汇总、挖掘洞见，并在生态系统中的伙伴之间实时交换，行成可交互数据，通过机器人流程自动化（RPA）、协作机器人和无人机让供应链现场的工作更安全高效，将传统的成本中心转变为价值驱动导向。

(4) MLOps: 人工智能产业化

精密的机器学习模型有助于企业更高效地发现规律、找出异常、进行预测以及做出决策，形成洞察，尽管采用机器学习技术的企业越来越多，但繁杂的开发和部署流程阻碍了生产团队、运营员工和数据科学家之间的合作，给企业带来巨大阻力。据 IDC 报告，28%的人工智能/机器学习项目会以失败告终，其主要原因是缺乏所需的专业技能、足以支持生产的数据以及集成开发环境。还有更多的失败项目（47%）甚至未能从实验阶段走向投产阶段。

MLOps 是一种方法论，它将模型开发和运营管理整合起来并实现自动化，以加速整个模型的生命周期进程。MLOps 通过加速模型实验过程和开发队列管理，提高了建模质量，简化了监控、维护生产模型及管理监管要求，从而提高模型的商业价值。预计到 2025 年，MLOps 的市场空间将上升至近 40 亿美元。

(5) 机器数据革命：颠覆机器学习时代的数据管理价值链

•**采集与存储**：在原有系统与传统数据库中寻找有价值的非结构化、非传统数据，将其聚合并存储至针对机器学习优化的云平台产品中。

•**发现与关联**：使用认知数据管理技术识别不同数据之间的潜在联系并获取洞见。

•**增强机器学习能力**：探索利用边缘计算和 5G 技术强化实时数据处理能力与降低延迟的机会。

(6) 零信任架构：可渗透边界时代的数据安全问题

• **加固网络安全基本面**：通过加深对网络卫生基本原则和实践的了解，帮助企业最大限度的利用零信任为自身创造价值。

• **重新思考网络组织**：零信任要求企业重新思考其组织的安全防护方式，和支持这种方式的相关技能、流程和技术。

• **安全技术自动化**：简化、整合安全技术堆栈并实现自动化，提升安保团队的效率，简化安保流程和运行措施。

(7) 启动数字化工作环境

数字化工作环境为管理提供了便利，数字化工作环境的技术和工具可洞见包括员工个人绩效、团队绩效，以及整个公司的士气，让他们能够识别新工作模式并进行预测，鼓励积极的员工行为，并微调个人、团队和组织绩效。当员工回到办公室时，这些数据还可以帮助远程团队和现场团队协同工作，并确保远程员工和现场员工获得平等体验。

(8) 数字化与实体相融

• **“个人专属”的设计**：设计师和开发人员根据个人行为、态度和偏好量身打造人感体验。

• **“实体+数字化”融合**：构建实体和数字化元素融合的客户旅程，打造独一无二的定制品牌体验。

• **技术扩展信任**：通过信任机制支撑的技术，来强化精心的设计，打造具备定制化人感体验的未来。

(9) DEI(解多样性、公平性和包容性)技术：公平工具

DEI 技术解决方案可以使用人工智能/机器学习和高级分析，深入了解人才引进、晋升、幸福感和保留等方面的情况。其他技术则可为领导者和决策者提供反馈和指导。所有技术都希望减少决策方面的人为差错和偏见，同时提高组织绩效，促进组织创新。以上只是几个示例，说明组织如何使用技术工具支持 DEI 成果。

2.人工智能技术总趋势

(1) 成熟技术的发展侧重于特殊场景下技术难题的解决

在人类语言技术方面，主要体现在语音技术下的语音识别以及自然语言处理方面的机器翻译、问答以及情感分析任务，这些技术任务领域都是在整体技术指标达到超越人类的一定阈值之后转入研究解决个别的具体实际问题。已经落地的技术任务还共同面临着场景切换的问题，主要是不同垂直领域的语言或图像识别任务需要采取不同数据训练出的模型进行应对，而未来基于无监督学习的研究会更加侧重这些技术的跨领域、跨场景融合，使得一个模型可以解决多个不同场景的问题。

(2) 无监督学习是未来技术突破的主要着力点

在语言技术和视觉技术这两个人工智能感知层应用落地最成熟的两个技术领域，无监督学习都是被提及频率最高的技术发展趋势。NLP 领域目前最成功的大规模预训练语言模型(BERT 等)证明了无监督学习在自然语言处理类任务里的可行性与发展潜力，transformer 体系结构已成为 NLP 任务的实际标准，开始扩展至最相关的语音技术领域。而计算机视觉领域相比 NLP 更早使用注意力机制，但对于 transformer 模型的应用仍然受到卷积网络结构本身的限制。

(3) 移动设备和物联网设备开始承担更多 AI/ML 计算

随着移动智能手机和物联网设备的计算能力越来越强，用户和监管机构对数据隐私的保护越来越重视，同时移动网络 5G 的逐步部署落地，移动设备和物联网设备会进行更多的 AI/ML 的计算和训练。这使得人工智能离数据更近，离用户更近，也能更迅速地进行迭代升级。一方面，云计算发展依旧迅猛，并逐渐向边缘计算发展，助力物联网智能化；另一方面，AI 芯片也逐渐走向云端、边缘端以及物联网设备终端。可以预见，随着 5G 技术的逐渐成熟，将会有越来越多的物联网设备和相关应用，进一步促使相关的云计算和芯片产业与之共同发展前进。

(4) 图神经网络、图深度学习、大规模图计算平台将有很大发展空间

近年来，受社交网络、知识图谱以及生物信息学和神经科学中图形的重要性所驱动，将深度学习模型应用在不规则图结构的数据上逐渐成为了一种新兴趋势。这些基于图的神经网络（图神经网络，GNN）通常将神经网络模型应用在图中与顶点和边相关的特征上，传播运算结果并进行聚合，从而生成下一级的特征，目前已经在分类、嵌入、问答等多种不同的目标应用中取得最佳表现效果。这种方法“将端到端学习与归纳推理相结合，业界普遍认为其有望解决深度学习无法处理的因果推理、可解释性等一系列瓶颈问题”。相比传统的文本、语音、图像数据格式，图结构是一种泛用性更强且更能代表现实世界真实情况的数据表示方式。

(5) 深度学习的可解释性与 AI 安全

黑盒问题是自深度学习诞生起便一直伴随其中的主要问题与瓶颈之一，尽管深度学习技术的发展日新月异，在准确率等技术指标上取得了惊人的成就，但是对于模型的可解释性方面提升仍然不大。这极大地影响了相关技术在一些对模型结果的可信度和可解释性有着高要求行业（如金融和医学）的应用。与此同时，伴随着 AI 技术的不断发展，用户对于技术的关注度越来越高，人工智能安全问题也便随之越来越受到大众关注。这与神经网络的可解释性一样是技术得到更广泛的应用之前所必须突破和解决的问题。

(6) 自动驾驶与机器人技术任重而道远

自动驾驶技术整体上处于上升趋势，但是总体热度相对其他 AI 技术而言却是逐年下降的。2015 年无人驾驶最火热的年代曾有多家公司和媒体预测 2020 年将会有高可用性的 L4-L5 级别的自动驾驶出现；然而随着高成本、复杂环境、以及一些无人驾驶事故的出现，人工智能带来的安全、隐私和道德问题也开始得到越来越多的关注。无论是技术上还是社会影响上，无人驾驶领域的泡沫开始逐渐散去，其研发开始转入了一个相对沉静的缓慢发展期。

第二部分 2021 年中国数据分析行业发展概况

第三章 2021 年中国数据行业发展现状及趋势

党中央、国务院高度重视大数据在推进经济社会发展中的地位和作用。2014 年，大数据首次写入政府工作报告，大数据逐渐成为各级政府关注的热点。

2015 年，国务院发布《促进大数据发展的行动纲要》，大数据正式上升至国家战略层面，对大数据整体发展进行了顶层设计和统筹布局，产业发展开始起步。

2016 年 3 月，《十三五规划纲要》正式提出“实施国家大数据战略”，国内大数据产业开始全面、快速发展。随着国内大数据相关产业体系日渐完善，各类行业融合应用逐步深入，国家大数据战略走向深化阶段。

2017 年 10 月，党的十九大报告中提出推动大数据与实体经济深度融合，为大数据产业的未来发展指明方向。12 月，中央政治局就实施国家大数据战略进行了集体学习。

2019 年 3 月，政府工作报告第六次提到“大数据”，并且有多项任务与大数据密切相关。

2020 年是历史发展进程中极不平凡的一年，世界正经历百年未有之大变局，特别是突如其来的新冠肺炎疫情为各行各业带来了前所未有的挑战，然而，危机之中，数字化技术驱动技术和产业变革仍加速发展，大数据技术、产业和应用逆势而生，数据的作用在助力疫情防控和复工复产中大放异彩，“数据驱动”的价值更加深入人心。

2021 年是“十四五”开局之年，也是数字化转型关键年。数字经济势不可挡，成为中国崛起、弯道超车的重大机遇，也是企业提升效率的必由之路。然而，企业在享受数字化带来红利的同时，挑战也随之而来，加强数字化治理迫在眉睫。2021 年 3 月，国务院《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》中“加快数字化发展，建设数字中国”独立成篇。迎接数字时代，激活数据要素潜能，推进网络强国建设，加快建设数字经济、数字社会、数字政府，以数字化转型整体驱动生产方式、生活方式和治理方式变革。2021 年 9 月 1 日，《中华人民共和国数据安全法》正式施行。《数据安全法》是以数据安全为核心，涵盖个人信息、政务数据等各类型数据，涉及数据利用与安全发展，规定了数据安全工作机制、职责与保护制度，兼顾政务数据安全与开放，是我国首部比较全面的、效力层级较高的、专门针对数据的法律。2021 年 11 月 1 日，《个人信息保护法》正式施行，这部法律不仅是进一步加强个人信息保护、维护网络空间良好生态的法治保障，也是促进信息数据依法合理有效利用、推动数字经济持续健康发展的法律推手。《个人信息保护法》标志着我国在个人信息保护领域法律框架的基本确立，向世界展示了数据跨境流动制度体系的中国方案。

一、2021 年中国数据分析产业发展概况

作为衡量数字经济发展水平的重要统计标准，2021 年 5 月中国国家统计局公布了《数字经济及其核心产业统计分类（2021）》，首次确定了数字经济的基本范围。从“数字产业化”和“产业数字化”两个方面，

确定了数字经济的基本范围，将其分为：数字产品制造业、数字产品服务业、数字技术应用业、数字要素驱动业、数字化效率提升业 5 大类。其中，前 4 大类为数字产业化部分，主要包括计算机通信和其他电子设备制造业、电信广播电视和卫星传输服务、互联网和相关服务、软件和信息技术服务业等，是数字经济发展的基础。第 5 大类为产业数字化部分，是指应用数字技术和数据资源为传统产业带来的产出增加和效率提升，是数字技术与实体经济的融合。

数据分析行业目前侧重于产业数字化部分，在新一代数字科技支撑和引领下，以数据为关键要素，以价值释放为核心，以数据赋能为主线，对产业链上下游的全要素数字化升级、转型和再造的过程。2021 年，数据分析产业也将迎来新趋势。目前来看，数据分析产业主要可面向数字政府、数字生态、数字科技等方面。

1、数字政府

在数字时代的背景下，在政府治理方面也迎来了前所未有的转变，“数字政府”的概念也被熟知。数字政府是指在现代计算机、网络通信等技术支撑下，政府机构日常办公、信息收集与发布、公共管理等事务在数字化、网络化的环境下进行的国家行政管理形式。是在现有政府组织架构下，遵循“体制创新+技术创新+管理创新”三维一体架构，所形成的一种新型政府运行模式。用数据对话、用数据决策、用数据服务、用数据创新。作为数字中国体系的有机组成部分，数字政府是推动数字生态建设、实现经济高质量发展、再创营商环境新优势的重要抓手。

2、数字生态

数字生态则是新的产业发展方向，是推动“互联网+”深入发展、促进数字经济加快增长的新动力。其具备价值循环体系、产业融合机制、社会协同平台这三大特征的新型经济单元。数字生态是数字经济的基本单元。大量异质性的企业，借助大数据等新兴技术，紧密的融合在一起，形成的共生、互生、乃至再生的价值循环体系，产生新的产业融合机制。与传统产业受时间、空间限制不同，数字经济体往往可以跨越地域、行业、系统、组织、层级形成广泛合作的社会协同平台。

3、数字技术

大数据、云计算、人工智能等技术的快速发展，加快了公众对大数据技术的认知度。当前，数字科技几乎已渗透到我们的生活的所有领域，包括政治、经济、文化、医疗预防、金融、农业、能源、城市管理等。全国各地利用新技术进行创新，促进产业数字化转型已成经济发展普遍共识。

二、2021 年中国数据分析行业面临市场机遇

数据分析行业以大数据行业发展为基础，2020 年，我国大数据产业迎来新的发展机遇期，产业规模稳步增长。根据中国信息通信研究院对大数据相关企业的调研数据，近年来我国大数据产业规模稳步增长。2016-2019 年，短短四年时间，我国大数据产业市场规模由 2841 亿元增长到 5386 亿元，增速连续四年保持在 20% 以上。根据近年来大数据行业市场规模增长态势，2020 年大数据行业规模约为 6670 亿元。2020 年，受新冠肺炎疫情倒逼，大数据技术、产品和解决方案被广泛应用于联防联控、产业监测、资源调配、行程跟踪等新兴领域。2021 年 1 月，工信部印发《工业互联网创新发展行动计划（2021-2023 年）》，到 2023 年，国家顶级节点具备标识、域名、区块链等综合服务能力，标识注册总量超过 150 亿，日均解析量达到千

万次量级。要求编制完善工业互联网平台监测评价指标体系，支持建设平台监测分析系统，提供平台产业运行数据分析服务。随着各行业领域大数据应用主体持续增加、与业务特点相匹配的个性化、定制化大数据解决方案日益受到青睐，2021 年数据分析的重要性进一步加强，数据分析师、数据分析师事务所持续增长，专业化认证进一步加强，在行业有利的大数据环境，国家政策支持，数据分析企业、数据分析人才的不断增加驱动数据分析行业蓬勃专业发展。

掌握数据资产的企业群是大数据的首批和直接受益者，可以方便地对大数据进行加工、消化、利用。而随着应用价值的逐步体现及大数据产业的发展，应用将必然扩张到传统产业的方方面面，不断创造新的应用场景，数据分析场景也不断延伸。

数据分析行业是一个跨学科的边缘科学行业。它以量化指标为利器，以数据为产品表现形式。现在大数据蓬勃发展的阶段，数据分析行业优势得以凸显：从横向看，只要有需求和满足需求的服务能力，数据分析行业的业务就可以涉足各行各业，不再受行政框框的约束；从纵向看，只要有需求和满足需求的服务能力，它的业务可以深入行业内部，上可以运营大数据，下也可以驾驭微管理，不再受资质等级的限制。

数据分析行业是一个理念先进，提前进入新形势的行业。企业的市场行为应该让市场去鉴定和评价。所以，对于有着明确的全市场特征的数据分析业而言，这种特殊的垄断是不可能形成的。有数据的、有能力的公司都可以进军这一行业。

大数据时代的到来，不仅为数据分析行业提供了大家均已经耳闻目睹的名义上的数据分析概念，也为数据分析行业提供了大家都能够身心体会的实质上的数据分析效益。数据分析行业市场机遇无限，占得时代先机，前景广阔。

三、中国数据分析行业未来发展趋势

大数据时代，数据成为关键生产要素，企业的数字化转型和数据应用也迈入新的阶段。数据量的爆发式增长和数据问题的暴露让越来越多的企业将目光转向数据分析，期待从数据金矿中采炼更多，向数据驱动决策转型。然而知易行难，数据分析带给企业的也不仅仅只有机遇和收益，还有各种问题和挑战。因此，如何深入挖掘数据应用的价值，让数据成为生产力成为了广大企业不断探索的关键所在。

按照数据应用深入程度的不同，可将众多的大数据应用分为三个层次：第一层，描述性分析应用，是指从大数据中总结、抽取相关的信息和知识，帮助人们分析发生了什么，并呈现事物的发展历程。第二层，预测性分析应用，是指从大数据中分析事物之间的关联关系、发展模式等，并据此对事物发展的趋势进行预测。第三层，指导性分析应用，是指在前两个层次的基础上，分析不同决策将导致的后果，并对决策进行指导和优化。

当前，在大数据应用的实践中，描述性、预测性分析应用多，决策指导性等更深层次分析应用偏少。虽然已有很多成功的大数据应用案例，但还远未达到我们的预期，大数据应用仍处于初级阶段。未来，随着应用领域的拓展、技术的提升、数据共享开放机制的完善，以及产业生态的成熟，具有更大潜在价值的预测性和指导性应用将是发展的重点。

1、融合发展，数字化转型，赋能企业数字化重塑

国家推进数字化转型的认识和决心不断强化，不断出台支持政策，数字化转型的发展环境不断优化。《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》将“加快数字化发展，建设数字中国”单独成篇，提出以数字化转型整体驱动生产方式、生活方式和治理方式变革，在顶层设计中明确数字化转型的战略地位。各省级“十四五”规划也都强调加快产业数字化转型，推动数字化赋能各行各业。

数字化转型是一个循序渐进的过程，是企业发展的必经之路，且“丰俭由人”。而对于数字化转型的认知，很多企业只停留在技术层面，认为其只是对数据和技术的简单利用。然而，数字化转型是由数字化和转型两个部分构成，其最终目的是要实现整个企业人员组织流程的转型。可以说，数字化转型的背后是企业战略、组织文化的变革，而不只是技术的应用。因此，忽视企业战略，盲目追求技术，将难以对企业的实际经营赋能。实现数字化转型需要与业务深度融合，因此数据化管理的思维是非常重要的。

2、数据人才培养

数据化转型是一个充分利用数据资源解决现有行业问题和社会问题的过程，这个过程不仅依赖于资金投入、资源整合，更高依赖于人才的技术、创造性以及管理能力，数据人才就是支撑企业数据化转型的重要基础和武器。如何培养数据人才、配置数据团队、制定相应的人才战略并吸引更多优秀的数据人才、激励数据人才发挥最大价值，是其成功数字化转型的关键，从而在新一轮的市场竞争中占据高地。

自 2016 年教育部第一次新增“数据科学与大数据技术”专业起，大数据相关专业新增数量在新增专业数量排行榜中均位居前列，认真分析这些专业更多体现了实用性与交叉性，反映了大数据相关专业朝着精细化、融合化的方向发展。然而，面对快速变化的商业和科技环境，学校和企业对于数据人才的培养始终面临更多形态的挑战：一方面数据人才需要大量实践经验积累和不同场景历练，才能将数据科学运用到真实的业务场景中，因此一个优秀的数据人才培养周期是非常漫长的；另一方面数据人才培养出现一定的教育偏差，数据思维能力的提升是重中之重，数据能力将是未来求职者的重要基本素质，同时企业也应鼓励人员提升数据思维能力，从提升数据质量出发从而培养企业数据文化底蕴；第三方面我国教育体制以注重培养专业化人才为主，导致现阶段既了解传统行业技术，业务流程与发展需求，又能够掌握和应用数字技术的复合型人才严重缺乏，有融合实践经验的高素质人才更是紧缺。数据化应用在每个企业中蔓延，每个企业都无法拒绝这种趋势的影响。对于企业而言，在这个不确定、模糊、复杂、易变的商业环境中能够帮助自身应对数据化“战争”的有力武器之一就是数据人才。在数字经济发展规划中，汇聚人才要素资源，培养高精尖数据人才，打造“数字工匠”，形成一支数量充足、素质优良、结构合理、富有活力的数字人才队伍。

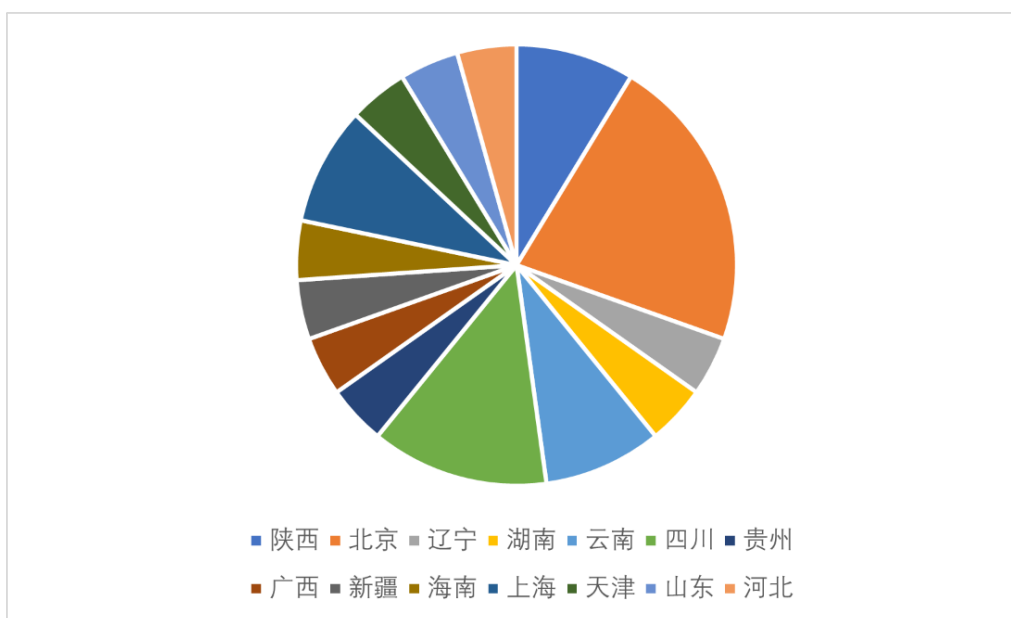
第四章 2021 年中国数据分析师事务所发展状况

数据分析师事务所经我会审批通过成为我会会员后，接受我会的监督和管理，数据分析师事务所会员具有统一的服务标准规范，是数据分析行业走向规范和自律的中流砥柱，是促进数据分析行业健康发展的中坚力量。目前，我国数据分析师事务所会员遍布在全国各省，虽然在规模、人员学历层次、服务的领域和深度以及业务模式方面还存在需要改进的地方，但随着社会数字化转型的需求，对数据价值的认可，数据分析师事务所必将有更广阔的发展空间。作为数据分析行业全国性行业组织，我会也将一如既往地加强行业服务和管理工作，引导中国数据分析行业的健康发展，今年我会组织制定《数据分析行业服务参考文件》，在业务范围和收费参考等服务方向上重新规范，为数据分析师事务所从业提供行之有效的建议。未来，数据分析行业依然面临着诸多的困难和挑战，我会将及时了解会员单位发展中的问题与难题，积极为会员单位提供政策咨询和专家指导，促进行业持续健康发展。与此同时，我会将继续积极协助政府有关部门逐步制定和完善各类行业标准，规范数据分析从业人员及企业，促进数据分析行业繁荣发展。

我会于 2021 年 11 月对全国数据分析师事务所进行了问卷调查，就事务所的团队建设情况、业务情况、宣传推广情况等进行调查，并收集了事务所关于我会工作的一些意见和建议，为下一步更好地指导事务所的健康发展提供了参考。

一、事务所地域分布情况

作为行业最先发展的试点地区，北京和陕西多年来一直保持着平稳的发展节奏。除此之外，四川、贵州、云南等经济增长迅速、各类行业较为集中的地区，受疫情影响，虽然 2021 年事务所和执业分析师的数量缓慢增长，但我会接到咨询事务所的情况有增无减，多数创业者持观望和准备状态。

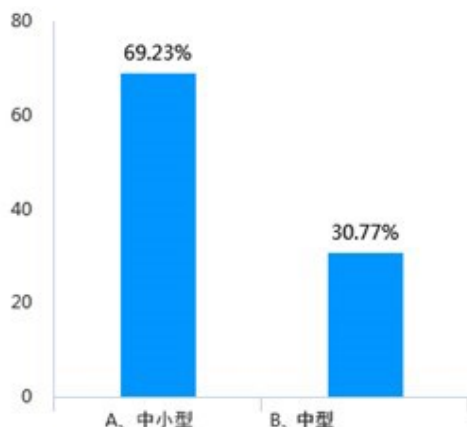


图：事务所会员分布图

二、事务所团队建设情况分析

1、事务所规模分析

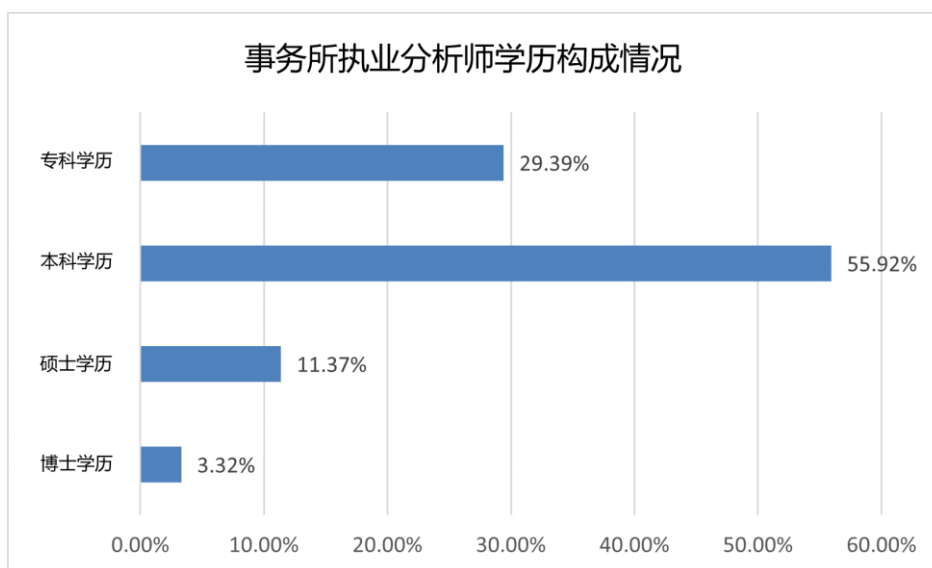
就调研数据来看，目前的数据分析师事务所规模普遍集中小型，占比 69.23%，中型规模，占比 30.77%。整体体现出事务所目前还属于创业期，规模整体偏小。



图：数据分析师事务所规模

2、事务所执业分析师学历构成分析

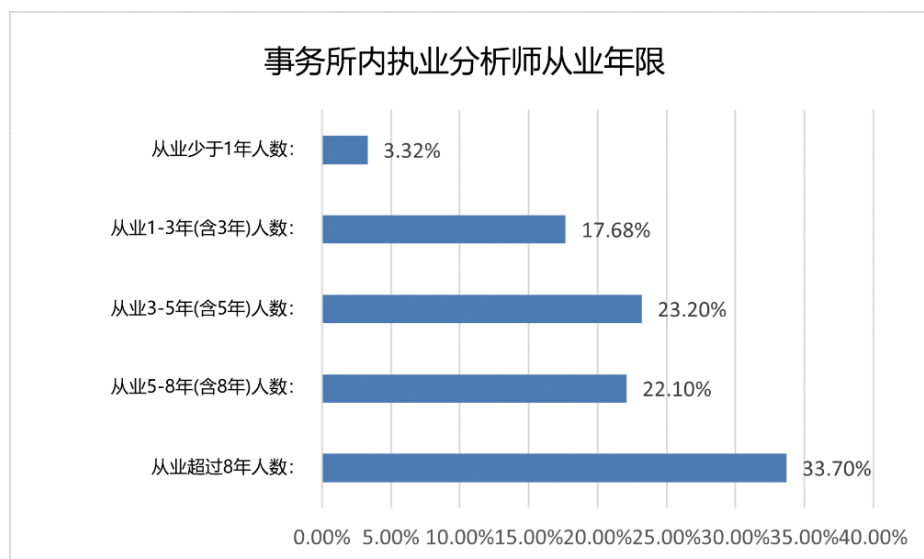
目前，事务所的数据分析师学历主要由专科、本科、硕士和博士构成，由调研数据可看出，有近 55.92% 的事务所有本科学历的分析师，占比最高，有 29.39% 的事务所有专科学历的分析师，有近 11.37% 的事务所有硕士学历的分析师，只有近 3.32% 的事务所有博士学历的分析师。而从全国执业分析师学历情况来看，本科学历依旧成为主要学历层次，硕士和博士学历有所增加。



图：数据分析师事务所执业分析师学历构成

3、事务所内执业分析师从业年限分析

随着行业快速发展，事务所内执业分析师的从业年限也进入稳定增长期。拥有 8 年以上从业经验执业分析师 33.7%，5-8 年从业经验的执业分析师近 22.1%，3-5 年从业经验的执业分析师接近 23.2%，而拥有 1-3 年从业经验的执业分析师也达到了 17.68%。

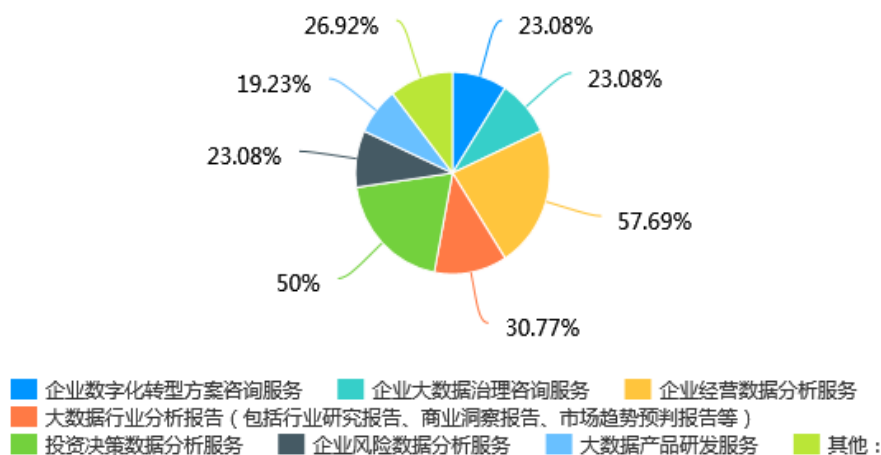


图：数据分析师事务所执业分析师从业年限构成

三、事务所业务情况分析

1、事务所主要业务类型占比

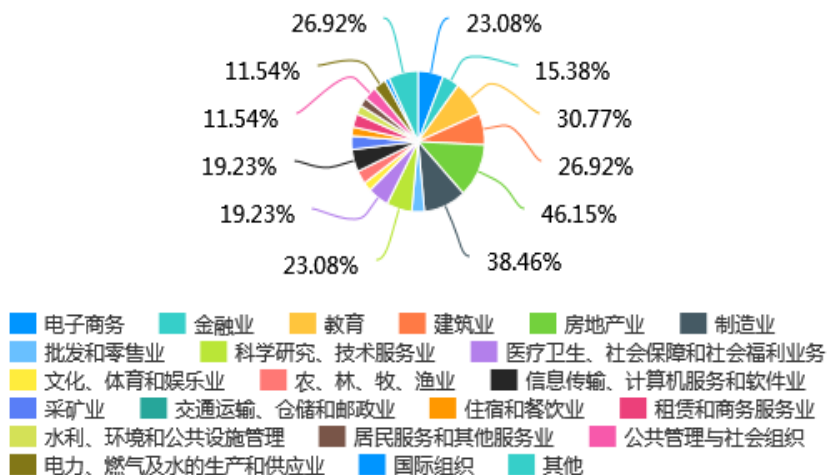
通过调查结果显示，目前事务所主要业务类型，涵盖了企业数字化转型方案咨询服务、企业大数据治理咨询服务、企业经营数据分析服务、大数据行业分析报告、投资决策数据分析服务、企业风险数据分析服务、大数据产品研发服务以及.....其中，企业经营数据分析占比 57.69%，投资决策数据分析服务占比 50%，企业数字化转型方案咨询服务占比 23.08%，大数据产品研发服务占比 19.23%。



图：数据分析师事务所主要业务类型

2、2021 年事务所从事数据分析业务的领域

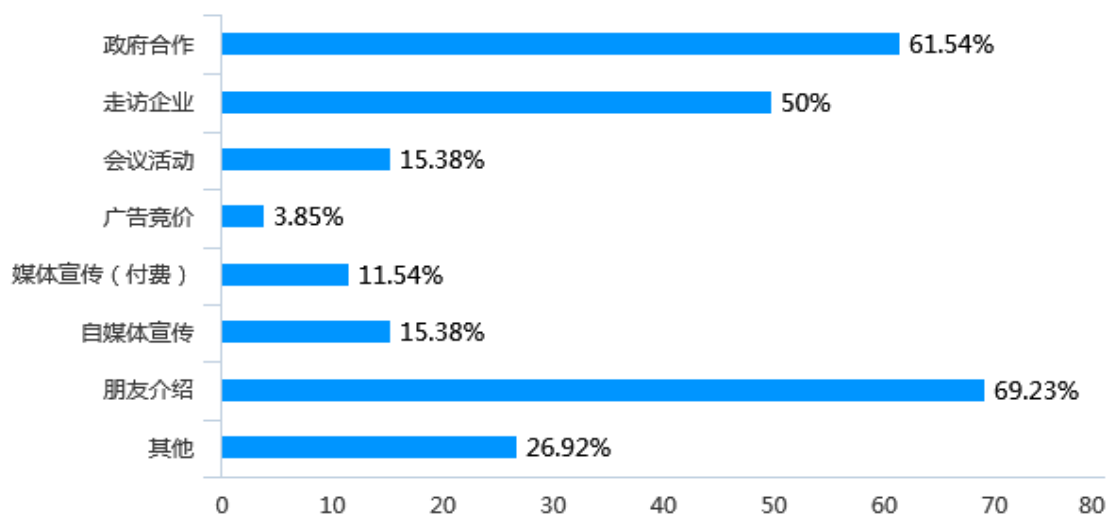
从调研结果看到，数据分析师事务所从事业务领域广泛，遍布各个行业。其中房地产占比最高，而制造业、金融业、电子商务、住宿与餐饮业、科学研究、技术服务业也明显较多。



图：2021 年数据分析师事务所业务领域

3、事务所获得业务来源

从调研统计结果看，事务所获得业务的来源主要有通过朋友或关系推荐和政府合作，其次主动走访客户。部分事务所开始注重自媒体运营、自主策划并组织活动。事务所在业务来源方面，整体比以往思路更开拓，逐渐具有互联网时代发展思路。

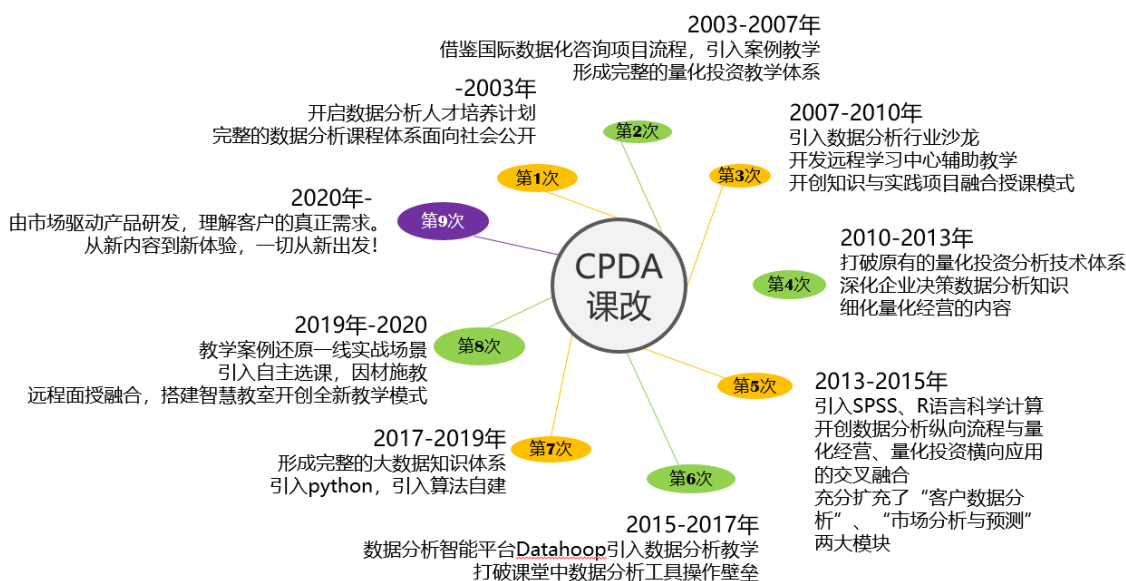


图：2021 年数据分析师事务所获得来源方式

第五章 2021 年数据分析人才培养概况

一、CPDA 数据分析师课程培训及课程改革

2003 年至今，CPDA 数据分析师认证培训用 18 年的时间为社会各行业输送了超过 5 万名数据分析师，助力上千家单位公司转型数字化建设，成为中国数据分析行业发展的中坚力量；CPDA 数据分析师认证培训历经 9 次升级，积淀了极具竞争力的数据分析专业知识传授体系。



经过十余年的推广并根据市场的反馈，我们建立了科学完整的课程体系，配套教学体系自主研发了数据分析专业教材。如《数据分析基础》、《客户与产品数据分析》、《供应链优化与投资分析》等。



二、Datahoop 智能大数据分析平台

2016 年我会为分析师免费开通 Datahoop 数据分析平台，运用先进的专业分析平台，在学习专业课的过程中，掌握智能平台的应用，提高业务水平，真正将学习与应用融为一体。Datahoop 数据分析平台是由我会授权专业的技术团队自主研发设计的数据分析平台，采用一键式算法自匹配功能，为用户节省了操作时间，也省去了前期大量先验知识的积累。全面支持 Python 语言，个人用户可以方便地提交个人的算法包，企业也可以根据个性化需求进行微场景搭建。大量的二次开发不但将极大地强化平台的深度分析研究能力，而且将有效地推动 Datahoop 平台的快速迭代与升级。

1. 无编码的拖拽式数据分析

Datahoop 平台提供了独立的分析工作台模式的功能服务，让用户用可视化组件的积木搭建方式实现数据分析项目，使整个分析工作直观、简单、便捷、逻辑感强、聚焦业务核心。

Datahoop 引入独立的试算工作台模式，数据分析项目就像搭积木一样简单便捷，同时通过可视化组件实现“数据导入→数据加工→模型构建→数据展示→数据导出”完整数据流过程。从数据管理到模型设计，全部为拖拽式搭建，无编程门槛，各种类型的数据分析人员都可轻松完成。特别适用于高校学生以及入门数据分析师进行实操训练。

2. 渐进式增长的算法资源

平台预置了多种分析算法，供数据分析师所在企业直接使用，也可基于 Python 脚本创建新的算法，更加适合分析师个性化需求。为企业提供准确有效的分析结果。并且通过平台可以得到各种分享的算法和知识，满足数据分析师的执业需求。

3. 丰富的数据分析能力体验

Datahoop 平台为了提供更加丰富的数据分析能力，通过算法分析数据结果并图表化展现。以 PaaS 的方式向用户输出大数据分析服务。平台把算法和场景作为应用管理，让用户聚焦业务核心，专注于数据、算法和场景，为用户的商业用途提供了更多的数据分析业务体验。

三、数据分析行业公益沙龙活动

- 4 月 11 日，我会在北京举办主题为“打破发展瓶颈，高效推动业务突破”的线下沙龙活动，由来自互联网、金融、房地产、医疗、联合办公等领域的大数据分析爱好者们共聚一堂，共同探讨这一主题。活动中京东何正松老师带来如何破题用户增长，科学评估用户生命周期的演讲；赵迅多老师为大家解读数据分析如何助力智慧生态规划高科技落地场景；我会科研中心赵玉莲老师深度剖析并分享学习思路和技巧，同时引入应用场景，学以致用。

- 5月15日，数据分析行业公益沙龙活动在昆明举办，本次活动特邀复旦 MBA、TOP 管理咨询公司分析顾问肖策予老师为大家带来主题为“500 强前沿案例交流分享”。肖老师与大家一起分享行业前沿的数据分析实战案例，总结经验。让数据分析师在工作中有例可鉴。结合曾经的工作经历，肖策予老师分享了各大企业基于数据分析价值地图所做的具体案例，案例企业涵盖了汽车制造业、电力行业、消费品行业等，信息量十足。此外，本次活动现场，还有来自民生银行、云南网、国开大学物业管理学院、中国人寿、中国平安、协合新能源集团、辛元集团、Max Tech 等行业大咖和数据爱好者，大家共同参与了本次交流。

- 7月4日，我会在北京与来自数据分析行业各领域的数据分析师嘉宾们一起现身「数据分析线师线下沙龙圆桌咖啡」，畅谈工作所需、圈层动态，建言行业未来发展。这次活动的分析师嘉宾分别来自银行、律所、综合业态集团、汽车、航空航天等不同行业，大多是拥有丰富工作经验，具备多年数据分析实战的职场精英，拥有较好的行业洞察力，短短几个小时，大家就自身所处行业对大数据、数据分析的应用现状和痛点进行了深刻、扼要的分享。大家共同认为，数据分析行业资源优势有必要通过一个“中间平台”实现更广的价值释放，并让更多数据分析师加以利用；成立一个由数据分析师组成的俱乐部，会对前述行业痛点提供一个良好的解决出口和交流环境，不断产生的“数据需求”和与之对应的“数据方案”会让这个“平台生态”得到持续的补充，成长。数据分析行业是一个朝阳行业，数据分析师更是一群赶在时代浪潮前的“先行者”，目前行业内存在的“信息壁垒”和“交流孤岛”需要我们去破除。经过本次活动的讨论，大家进行深入思考，并约定不定期举行“圆桌”讨论，为数据分析行业发展献计献策，添砖加瓦，为数据分析师的凝聚与未来提供更多的价值。

- 7月29日，我会组织数据分析行业线上公益沙龙活动，特邀海南数据分析师事务所创始人周召安为分析师们带来一场关于 PowerBI 解决方案数据平台部署实例的分享。次活动分享内容围绕服务器基本搭建、数据库部署、ETL 模型构建、可视化模型构建及发布和用户数据权限设置几个方面。通过本次分享交流的学习，使大家掌握到在技术支持下，一个人借助 PowerBI 框架，完成从搭建到数据同步，再到模型构建、用户数据权限设置，独立完成网页端和手机 APP 端数据平台的搭建。使我们的分析师们通过实操案例，将各个概念衔接在一起，加强了对 Power BI 的理解，更助力在实际业务中搭建企业数据平台。

- 8月20日，我会组织数据分析行业公益沙龙活动。特邀北京派可数据首席数据运营官、前天善智能联合创始人、36Kr 企业联盟|TO B 专家团成员吕品先生为数据分析行业分析师带来一场主题为商业智能 BI 的本质的线上公益沙龙活动。吕品先生从商业智能 BI 介绍，到业务管理决策思维与金字塔思维的剖析，最后通过客户流程分析的价值与意义的案例分析深入讲解商业智能 BI 的本质。分析师们通过直播，直接和业内专业人士进行在线对话，对数据分析、商业智能 BI 的发散性思考，讨论产品与数据之间的关系，启发对数据可视化的深度思考与研究。

- 9月9日，我会数据分析行业公益沙龙活动，特邀帆软软件九数云事业部运营总监 Jojo 先生为数据分析行业分析师带来一场主题为“让你的数据会说话——小白如何跨越数据分析学习鸿沟”的线上公益沙龙活动。随着企业越来越重视数据，从经验决策转变为以数据驱动业务发展，同时各岗位对数据分析能力有了更高要求，数据分析师是随着互联网时代发展起来的岗位，Jojo 先生从 Excel 开始的探索到瓶颈，到如何跨越数据分析的学习鸿沟，通过实际案例项目讲解如何短期提升效率和数据思维。

- 8月4日，我会特邀西安交通大学管理学博士李妹老师在线直播解读新版数据分析行业认证考试大纲。结合数据分析行业人才培养体系标准的要求，以及大数据应用场景的市场需求，我会对数据分析师是否具备

数据分析思维能力以及是否能利用数据分析技术解决实际业务问题等作为关键考核目标，通过对考试科目、形式以及考试范围进行详细解读，介绍由中国商业联合会数据分析专业委员会编著、中国商业出版社出版的教材内容，并通过对新旧大纲的对比更加充分的说明新版大纲的考核方向。从而帮助数据分析师对能力培养的准确理解和对知识的正确认知。

•9 月 23 日，我会特邀讲师王冲老师和赵丽老师为数据分析行业分析师带来一场主题为“5 分钟实现数据分析弯道超车”的线上公益沙龙活动。两位老师分别从数据分析发展应用趋势，学习路径现状到如何集中精力用思维指导业务，从 Datahoop 数据分析平台模块功能介绍到数据分析应用场景和完整案例演示。

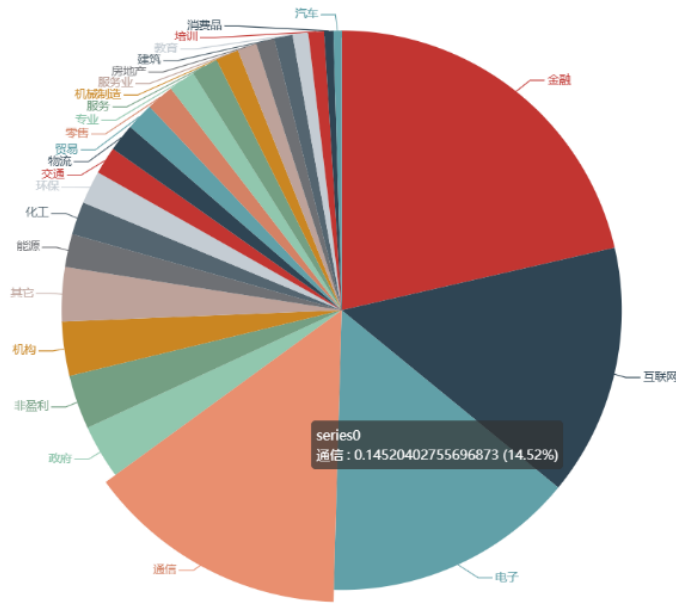
•11 月 10 日，我会特邀北京盈科律师事务所刘桂红律师为数据分析行业分析师带来一场主题为“个人信息与重要数据保护”的线上公益沙龙活动。在“十四五”规划和 2035 年远景目标纲要草案中，对营造良好数字生态进行阐述时也明确提出要“加快速度推进数据安全、个人信息保护等领域基础性立法，强化数据资源全生命周期安全保护”。大数据、人工智能、互联网新技术的广泛使用让隐私保护和个人信息安全问题日益突出，加强个人信息保护迫切性进一步加强。随着数据安全法的出台以及个人信息保护法的实施，无论是企业还是个人，我们应该做哪些准备。刘律师以实际案例方式，通过三个民、刑事案件的引入，进一步为数据分析师深度解读数据隐私新规下的数据使用。

•12 月 16 日，我会特邀数据分析师胡红艳女士为数据分析行业带来一场主题为“聊聊数据分析那些事儿”的线上公益沙龙活动。胡红艳女士曾任职于北京国际鲜花港数据中心主任、中旅北京生态旅游集团数据中心副主任，主要负责数据分析应用，一直致力于数据分析在司法实践中的应用研究。本次活动胡红艳女士从业务方向和技术方向如何选择，数据分析对企业的用途，怎么精确的做数据分析以及数据分析的边际效应等多个维度进行分享。

•12 月 29 日，我会特效永洪科技数据分析师陈辉杰先生为数据分析行业分析师带来一场主题为“企业数据治理实例分享”的线上公益沙龙活动。陈辉杰先生曾多年深耕客户交付一线，在大数据践行落地上积累较多经验，对零售、制造、金融行业的数据应用解决方案有深刻理解。本次活动中陈先生从对某全球头部 ICT 科技公司数据治理案例的项目背景，解决方案与价值体现多维度进行讲解，从公司实际部门业务场景，单部门试点推广到多部门应用，以及其完整项目流程。

四、2021 年数据分析师学员所属行业分布

随着数据分析在社会的各个领域的应用，社会对数据分析师的需求也越来越广泛，吸引了大量社会人员纷纷加入到数据分析行业中来，数据分析人才培养的规模也不断扩大。



图：2021 年数据分析师学员所属行业分布

我会一直加强对数据分析师学员及各会员单位分析师的培训，为他们营造更好的能力培养环境，使他们成为行业的专业人才。随着社会对数据分析需求剧增，企业对数据分析人员的能力要求也越来越明确。为此，我会在人才培养的过程中，通过数据分析相关知识来帮助分析师有效地了解数据分析工作的流程，提高在数据分析工作中真正用到的专业技能和职业素养，帮助其学有所用，进而整体推进全国数据分析人才队伍建设。

第三部分 中国数据分析应用领域应用发展及需求分析

第六章 中国数据分析应用领域应用发展及需求分析

一、金融行业数据分析应用发展及需求分析

风险控制是银行、证券等金融机构最核心的能力，大数据分析已成为金融机构提高风险控制能力的重要手段。金融机构汇集用户的贷款信息、交易记录、征信数据等多源异构数据资源，利用大数据画像、机器学习、知识图谱和复杂关系网络分析等先进技术，对多源数据进行实时处理，建立风险监测预测模型，大规模监测各关系数据中出现的非一致性，促使传统模式下难以浮现的风险显性化，从而及时有效地识别骗保、逾期、坏账等风险，实现潜在风险规避。

中国金融业向支持实体经济、创新驱动转换，数据分析在这个转换过程中的应用，成为中国金融业的新增长点和新亮点，数据分析尤其是大数据分析能够充分利用金融市场形成的海量数据来挖掘用户需求、评价用户信用、管理融资风险等，加上云计算等智能信息工具对序列之间的联系、动态信息数据序列进行分析，大幅度提高金融风险定价的效率，降低定价成本，有效减少信息不对称问题，进而对单个客户的信用信息、消费倾向、理财习惯分析成为可能。

我国的金融行业正处于应用数据的初级阶段，国内的金融机构经过多年发展与积累，拥有超过百 TB 的海量数据而且非结构化的数据量也在不断增长。金融机构在大数据应用方面具有天然优势：首先，金融企业在平时的业务开展中积累了大量高质量的数据，例如客户的身份、资金收付交易、资产负债情况等，这些数据经过专业技术挖掘和分析之后，将产生巨大的商业价值；其次金融机构相比之下有较为充足的预算，可以吸引到了解大数据技术的高端人才，也有能力采用大数据的最新技术。



为了促进金融业更好服务于国民经济各行业高质量发展，金融科技正在助力金融业数据化转型。在人工智能、大数据、区块链、物联网等技术驱动下，数字化金融服务在一定程度上已成为产业数字化解决方案的有机组成部分，如多个工业互联网平台嵌入了基于区块链的金融服务方案，可以对上游供应商、下游采购商信

用和信贷数据进行上链管理，将进一步提升工业互联网平台的服务能力，增加产业链黏性，尤其是对降低工业互联网平台上中小企业的融资难问题有明显作用。

二、医疗行业数据分析应用发展及需求分析

医疗数据是医生对患者诊疗和治疗过程总产生的数据，包括患者基本数据、电子病历、诊疗数据、医学影像数据、医学管理、经济数据、医疗设备和仪器数据等，以患者为中心，成为医疗信息的主要来源。目前我国医疗产业已经沉淀海量数据，并以 40% 以上的年均增速快速增长，这些数据早已超过了人力所能处理的极限，其中大部分医疗数据并未得到有效处理。同时，我国还面临着慢病发病率提升、临床决策失准及医疗资源配置不均衡、重复诊疗等问题。通过对医疗大数据的分析和加工，可以挖掘出和疾病诊断、治疗、公共卫生防治等方面的重要价值，为健康管理、辅助诊疗等场景提供解决方案，优化诊疗流程、提升医疗行为的效率，在“海量数据”与“医疗问题”之间架起一条通路。医疗大数据行业的发展关乎国计民生，具有重大的战略性意义，李克强总理表示发展医疗大数据产业已经“迫在眉睫”。随着市场的发展和政策的引导，未来 5 年将是医疗大数据行业发展的绝佳时机。

我国医疗体系的强监管性，大数据在行业内实现价值需要国家建立一套自上而下的战略方针，从而引导医院、药企、民办资本、保险等机构企业构建项目，相互合作，最终实现从“治疗”到“预防”的就医习惯的改变，降低从个人到国家的医疗费用。

我国已陆续出台关于扶持医疗大数据发展的相关政策，初步做好顶层设计并构建出医疗大数据发展的宏伟蓝图：2018 年，国务院 26 号文、《国家健康医疗大数据标准、安全和服务管理办法（试行）》先后颁布，2019 年 4 月，国家卫生健康委、国家中医药局联合发布《全国基层医疗卫生机构信息化建设标准与规范（试行）》，同年 9 月，国家发改委、国家卫生健康委等 21 部委联合印发《促进健康产业高质量发展行动纲要（2019-2022 年）》。

医院是医疗大数据最早产生和使用的主体，目前我国医疗大数据应用已经扩展到医疗相关的医药、流通企业，随着大数据技术的迅速发展，医疗与大数据结合更加紧密，在最近的几年，我国医疗大数据应用市场迎来快速增长期。2014 年我国医疗大数据市场规模仅 6.06 亿元，2016 年市场规模突破 10 亿元大关，2018 年市场规模达 56.3 亿元，年复合增长率达到 74.6%。其中，数据整合管理市场规模为 29.7 亿元，占比为 52.8%。预计 2025 年市场规模突破 200 亿元。

三、电信行业数据分析应用发展及需求分析

新一轮的产业转型将围绕从信息化向数字化的发展逐步展开，通信网络作为数据传播的底层通道重要性不言而喻。与其他行业一样，电信行业也面临着大数据的挑战。传统数据时代，电信企业已经占尽海量数据先机，较其他行业有更完善和发达的数据仓库系统，大数据时代，在原有传统数据采集和分析处理基础上，电信运行商有更广阔的数据处理渠道和数据源，不同于互联网的匿名性，电信企业获得的数据更为准确可靠。以移动用户为对象，通过 RFID、WIFI 和通话详单记录所记录的用户通信往来信息为代表的运营商数据更有分析的空间。通过这些数据的分析，可以发展和完善电信客户社交网络分析。通过 GPS 和移动终端商的各项应用信息，采集用户的位置及各项应用信息，为用户提供互联网、套餐等业务的推荐。

对于电信企业来说，主要掌握了用户的兴趣需求和喜好、交友等数据，根据这些数据种类进行社交网络分析、客户体验分析和客户价值分析等。

基于社会数字化转型的价值创造路径，电信企业数字化转型的关键在于两点。第一点是用户基数，也就是基础设施的完备度和先进度，例如便捷的通信网络和普及的智能手机是消费互联网爆发式增长的根基。第二个是新增市场的广度，关乎后续增长的潜力与空间，例如淘宝和滴滴这种全民级的应用场景就能带来更多的市场空间。上一轮 4G 技术红利为消费互联网的渗透打下了坚实的基础，而面向更高数量级的工业互联网、人工智能、车联网等赛道时，更强大的数据传输能力是其不可或缺的支撑。5G 的增强移动宽带、高可靠低时延和广覆盖大连接的属性可以带来高基数的终端客户，5G 的发展也将与云计算、大数据、人工智能、区块链、边缘计等其他信息通信技术一起，共同促进传统产业升级，孵化新应用，催生新业态，创造增量市场。2021 年国内新一轮 5G “新基建” 规模不减，远超全球其他国家建设进度。

四、交通行业数据分析应用发展及需求分析

人、车、路等是交通系统的三个基本要素，三者之间的相互关系演绎出复杂的交通系统，这三类数据是交通系统的基础数据，也是最核心的数据。当前重要的交通大数据有：

- **一卡通**，作为交通出行的重要结算方式，分析出行一卡通数据，能够全面了解城市交通出行规律。

- **移动手机**，一是移动手机信令数据，主要包括手机信令时间、定位信息等，来源于三大移动运营商；二是移动手机客户端定位信息，来源于移动客户终端互联网公司，如百度、腾讯、阿里巴巴等。利用手机移动通行网络中的定位信息来分析推算动态交通状况是一种新兴的广域动态交通探测技术。该技术利用现有的移动手机网络资源，利用已有手机通信中的少量信息，定位手机用户，进而获取实时交通数据和动态交通出行数据。

- **路网监控**，城市道路是城市的基本骨架，从某种意义上讲，城市道路监控网络就是城市治安监控系统的骨架。城市交通管理部门一般负责道路交通电视监控、机动车交通违法监测抓拍、智能卡口等系统建设，这三个系统的监控范围都是机动车道，监控对象都是机动车辆，是城市道路监控网络系统的重要组成部分，对构建城市治安防控网络系统都有着十分重要的作用。同时为改善交通拥堵状况，通过图像传输将路面交通状况上传至道路监控中心已经城市交管部门了解实施路口，及时指挥交通的重要工具。

- **车联网**，车联网是以车内网、车际网和车载移动互联网为基础，按照约定的通信协议和数据交互标准，在车与（车、路、行人及互联网等）之间，进行无线通讯和信息交换的大系统网络，是能够实现智能化交通管理、智能动态信息服务和车辆智能化控制的一体化网络。随着车联网的逐渐普及，以车为移动载体的用户出行数据规模将越来越大，成为交通大数据的重要组成部分。

目前在交通领域主要的数据应用方向包括：

- **提升交通服务**。大数据在交通行业的应用对于提升交通出行效率具有重要意义，也是解决城市交通拥堵的有利途径。（1）提高公众接受交通信息服务的便捷性、效率性、舒适度。（2）减少公众利用公共交通、私人交通出行的在途时间。（3）整体提高车辆平均行程速度，减少碳排放。

•**辅助政府决策。**通过大数据的融合分析，整合多方信息，实时掌握城市交通运行状况，准确预测交通发展态势，从而提高交通决策与管理水平。(1) 通过整合、共享道路交通、公共交通、对外交通等数据资源，提高交通行业各部门自身管理效率，避免由于部门间信息不畅通造成的重复建设，达到已建共享，新建统筹的目的；(2) 通过对多源异构数据的处理，形成能有效辅助政府决策的信息，辅助制定和优化公交-轨道交通换乘优惠、停车收费等政策，减少机动车出行量，提高公交出行量，缓解局部道路交通拥挤。(3) 全面、实时掌握城市的交通运行状况，有效应对突发状况。

•**改善交通环境。**大数据技术在减轻道路交通堵塞、降低汽车运输对环境的影响等方面有重要的作用。(1) 通过建立区域交通排放的监测及预测模型，共享交通运行与环境数据，建立交通运行与环境数据共享试验系统，大数据技术可有效分析交通对环境的影响。(2) 分析历史数据，大数据技术能提供降低交通延误和减少排放的交通信号智能化控制的决策依据，建立低排放交通信号控制原型系统与车辆排放环境影响仿真系统。

•**提高交通安全。**大数据技术的实时性和可预测性则有助于提高交通安全系统的数据处理能力，提高交通安全水平。(1) 通过边探测器检查车辆运行轨迹，大数据技术快速整合各个传感器数据，构建安全模型后综合分析车辆行驶安全性，从而可以有效降低交通事故的可能性。(2) 利用大数据，预测出群体出行的态势，对其可能出行的时间，出行路线，出行方式等等进行预测，从而为城市车辆调度、交通管理等提供决策支持，从而提前做好应急准备工作。

我国各城市交通建设、规划、管理、运营分散在各个行业主管部门，涉及部门较多，而每个部门根据自己业务需要建立各自信息系统。而这些信息系统仅应用于各个部门，缺乏与其他部门的信息共享，从而“信息孤岛”大量存在。通过大数据的体系，将不同范围、不同区域、不同领域的“数据仓库”标准化处理，从而构建交通行业基础数据库，构建交通行业，发挥整体性交通功能。

五、政府数据分析应用发展及需求分析

政府一直是新技术的最新应用者，因此“数字政府”逐渐成为全球各大经济体竞相追逐的目标。政府大数据的发展中，数据是核心资源和关键基础设施。通过构建大数据驱动的政务新机制、新平台、新渠道，可以全面提升政府在经济调节、市场监管、社会治理、公共服务、环境保护等领域的履职能力，形成“用数据对话、用数据决策、用数据服务、用数据创新”的现代化治理模式。

数字政府中的数据普遍包含政务数据、企业数据、社会数据三类数据，这其中政务数据处于核心地位。

公共部门是国家最大的数据源，政务大数据占据全部数据的 70%到 80%，并且政务数据具有权威性、可靠性，更具有开发价值。因此政务大数据被进一步开发利用，数据融合的趋势会越来越明显。

2017 年我国政府大数据市场规模为 220 亿元，占总应用层比例达到 27.2%，政府大数据已经广泛的应用到交通、公安、法检等各个专业领域。

预测在未来发展中，政务数据、企业数据、社会数据相互融合，通过三类不同数据的融合、开发、利用，使其派上更大的用场。当然，这中间最大的难题仍然是政务数据的开放问题，如何建立完善的政务数据开放制度框架，定期向社会开放政务数据，让企业开发利用，使数据产生新的价值，将是一个重要问题。

•**政府的政务大数据“拥”而难“用”**：当前发展存在“三难”，即互联互通难、数据资源共享难、业务协同难，导致海量数据价值难以挖掘，数据缺乏动态性和准确性；

•**政府的政务平台“统”而不“通”**：“三通”方面发力不足，即网络通、数据通、业务通尚有发展空间，许多地区大数据建设存在重复建设、低水平建设以及数据烟囱等问题，资源统建与共享是个难题；

•**数据资源“汇”而难“慧”**：政务服务存在明显短板，没有充分发挥互联网企业、基础电信运营商和咨询机构的技术优势；

•**政务治理体制机制“兴”而难“新”**：大数据整合融通交通、环保、城管、安防等各个方面，涉及层面广，统筹难度大，触及各部门深层次利益，而当前政务部门分工责任不明确，大数据统筹协调力度、督办力度不足，亟需建立健全机制体制保障。

六、能源行业数据分析应用发展及需求分析

当前，全球步入能源紧缺时代，提升能源生产和利用效率成为急需解决地突出问题，能源大数据是解决这一问题的有效手段。建设覆盖能源生产、传输、消费各个环节的数据资源池，搭建一体化的大数据可视化分析平台，直观展现贯穿能源生产运营全过程的数据变化趋势，同时基于对各环节实时数据的深度分析挖掘，在能源生产侧实现对能源生产设备的远程健康管理，在能源传输侧实现对能源传输网络的运行状态在线监控、自动识别和准确预测故障位置，在能源消费侧为用户使用行为特征挖掘、能源使用负荷异动识别、能源供需形势预判等复杂问题提供辅助决策。

七、物流能源行业数据分析应用发展及需求分析

连接物流上下游环节地仓储是物流全链路地中枢部分，在现代物流企业中占据举足轻重地地位。应用移动互联网、物联网、人工智能等先进技术对实体仓库进行智能化升级改造，已成为物流企业加快货物周转、减少物流成本以及增加服务附加值地关键战略。智能仓储使用大量地自动导引运输车、自主移动机器人、分拣机器人、码垛机器人等仓储物流机器人，极大地提升货物流通工作效率；使用三维扫描设备对所有入库物品进行扫描，并采集物品三维数据信息；使用智能物流管理系统实时建立入库物品三维数字模型，实现对库位、库存、物件状态地全程透明化跟踪和三维可视化调度。

八、零售业数据分析应用发展及需求分析

在传统零售与电商业务增速双双趋缓的背景下，如何打通线上线下渠道形成优势互补效应，帮助零售商准确触达用户消费需求成为破局关键。越来越多的零售商运用计算机视觉、深度学习等人工智能技术将实体门店升级为智慧门店，为消费者提供会员刷脸身份验证、小程序找货、扫码购、刷脸支付等多场景数字化服务，在打造便捷化无感化的用户购物体验的同时，收集用户店内行为数据，基于大数据、图像识别等技术勾勒用户精准画像，通过人工智能算法预测用户潜在消费需求，并推出及时、精准、个性化的线上营销，从而构建线上线下消费数据和营销数据的闭环。

九、畜牧业数据分析应用发展及需求分析

畜禽健康状况是畜牧业关注的焦点问题，以感知技术为切入点，对畜禽体征及行为进行监测、分析和预测成为农场实现精准养殖的可行选择。智能化的项圈、耳标、脚环等实行多样的动物可穿戴设备可实时采集畜禽体温、心率等体征数据和活动场所、运动量等行为数据，并将数据实时上传到畜禽大数据监管云平台，实现畜禽数据全天候、全流程的记录和跟踪。云端服务器上的配套数据分析软件将基于原始数据，挖掘深层次的健康信息和行为模式，将其转换为反映畜禽健康状态、繁殖预测、喂养需求相关的直观图表和信息，并发送至养殖人员手机端，从而实现对动物饲养、疫病防控、产品安全等全环节的精准质量管理。

十、矿产开采业数据分析应用发展及需求分析

矿产开采的地质环境条件复杂而危险，使用智能化的开采装备不仅可以更加精准地确定开采方位和开采工艺，提高开采效率，更有助于实现安全生产。智能开采设备基于机器视觉技术可智能感知和测量开采环境，并实时回传地址条件、设备方位等开采工艺数据具有较高自主控制能力地开采设备可根据开采条件地变化自动调节开采工艺，实现无人化地自主开采作业。

十一、建筑业数据分析应用发展及需求分析

建筑工程项目往往时间跨度大、参与机构多且质量要求高，项目管理的复杂程度极高。三维可视化建筑信息模型（BIM）的出现，为建筑工程项目管理提供了新途径。基于项目现场的各项信息数据，以 BIM 三维数字模型为载体，打造与项目现场精准映射的集成化项目管理平台，在时间维度上可实现建筑从图纸设计、施工建造、竣工交付到运行维护的全生命周期数据的集成与融合，同时在空间维度上将促进业主、设计、施工、供应商、运营方等各参与方的数据系统与业务协同，最终取得提升项目质量、缩短工程周期、降低建造成本、提高管理效率的多重效果。

十二、制造业数据分析应用发展及需求分析

制造企业对优化资源配置、创新生产模式、提升生产效率的需求日益迫切，作为新一代网络信息技术与制造业深度融合的产物，以及实现制造业数字化、网络化、智能化发展的重要基础设施，工业互联网通过数据、劳动、技术、资本、市场等全要素的全面互联，供应链、创新链、服务链、物流链、金融链等全产业链上下游的高度协同，以及从价值单点式迁移到价值网络化共生，从商品交换到信息交换再到知识和能力交换的全价值链革新重构，大幅提升工具的效用及效率，显著优化决策的流程与机制，持续创新以定制化、服务化、融合化为特征的生产方式、组织形式和商业范式，实现供给侧与需求侧的精准对接和实时匹配，构筑起数字化、网络化、智能化的新型生产制造体系和服务体系。

第四部分 中国商联数据分析委 2022 年工作规划

第七章 中国商联数据分析委 2022 年工作规划

2021 年，数字经济的发展，大数据的强大赋能，数字化转型的需要，数据分析的决策引导，正前所未有地促进各行各业价值创造方式。2022 年，作为行业组织，始终不忘初心，砥砺前行，以促进国家大数据战略目标实现为总原则，在规范从业行为、倡导行业自律、培养大数据人才，引领行业不断向前可持续性健康发展的道路上，不断审时度势，开拓创新、劈波斩浪、行稳致远。对 2022 年工作规划，我会将从以下几点说明：

一、推进会员发展、增强会员凝聚力，体现会员价值

作为数据分析行业引领者及会员权益的捍卫者，我会在大数据发展和数字化经济建设中应扮演重要角色，着力为行业发展保驾护航、建言献策，发挥积极的作用。2022 年，我会将着力于会员价值的挖掘，加强企业合作交流，为会员提供信息交流、资源获取的平台。

1、大力扶持重点数据分析师事务所，起到以点带面的作用。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。

2、加强我会对外合作交流，拓展信息渠道。与会员单位及相关企业交流，寻求合作机会。

3、组建数据分析行业案例库。搭建企业 and 专业机构等多方共同参与的平台和机制，通过数据分析应用案例的合作开发、共同交流、研究等方式，共建数据分析行业案例库，以帮助更多企业了解数据分析、应用数据分析，从而广泛促进企业数据化人才的建设、推动企业数字化转型的发展。

二、提升对行业内会员的监督与管理工作

1、引领行业，规范从业标准，组织行业的标准化工作的开展。

对《数据分析行业服务参考文件》广泛宣传、监督执行，为会员单位及数据分析行业内相关单位从业提供行之有效的建议，同时为维护市场秩序，推进行业健康发展而努力。

2、加强 2022 年的各项年检工作。

在 2021 年的年检中，我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 2022 年年检中，我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。

3、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,各地区向我会申请事务所会员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,我会在 2022 年的工作中,将更严格地规范事务所申办制度与流程,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。

4、加强对个人和团体从业会员行为的监督。

对于事务所发展,尤其是新成立的事务所,我也会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。

5、严格监督员及从业机构的违规现象。

近年来随着行业发展,少数违法违规的现象开始显露,部分事务所在巨大的发展机遇面前准备不足,虚假量化投资类业务屡禁不止,假协会、假培训机构炒作大数据概念、误导企业及分析师等等。对于违反行业自律公约、损害行业信誉的个别事务所会员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,我会将采取法律手段予以追究。我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、提升对行业内会员的服务工作

1、加强对会员单位和个人的执业教育培训及大数据相关培训,通过实战案例帮助会员单位和个人提高专业能力、研究水平、开拓业务思路,从而开展全面的数据分析业务。

2、搭建会员间交流、学习的桥梁,带领、指导会员技术能力提升。通过线上课程、社群服务等多种途径组织会员不断学习提升自身能力。

3、搭建业务对接平台,服务会员促进业务交洽与对接。

四、增强行业宣传力度,不断提升数据分析行业的影响力

1、举办第中国数据分析行业峰会。自 2010 年开始,中国数据分析行业峰会已成功举办七届,一路走过北京、上海、广州、贵州等地。每一届峰会都为当地带来了一场知识与实战的盛宴,受到当地政府、企事业单位的高度重视,更有力推动了当地数据分析人才培养进程。

2、通过网站、公众号、行业刊物、线上沙龙等各种资源平台及相关媒体平台宣传行业,让更多的人了解数据分析行业。优化升级的同时不断提升行业影响力。

3、事务所创业指导活动,将在全国范围内邀请业务比较成功的事务所为有意向创业成立事务所的人员讲解如何开拓业务、如何开展业务,以及案例分享、创业心得等内容。让事务所不断得到认知的同时,也能深度挖掘出适合自己的发展道路。

4、继续加强数据培训的推广。在国家大力主导大数据产业发展及大数据人才培养的背景下，为了更好的促进国家大数据人才培养战略的实施，近年来我会深入评估高校相关专业，对符合数据分析行业人才标准体系的高校专业开放考试条件，加快培养社会急需的具备大数据处理及分析能力的高级复合型人才。2022年，我会将继续评估其他院校相关专业的课程设置情况，对符合标准的院校和专业陆续公布，以引导更多人加入数据分析行业，促进行业快速发展。

附：本《中国数据分析行业年度发展报告》部分参考文献及资料，参考以下单位：

- 1、IDC（国际数据公司）国际数据全球机构 Statista
- 2、中国信息通信研究院《2021 年全球数字经济报告》
- 3、《全球数字人才发展年度报告》（2020 年）
- 4、软件和集成电路 2021 年第 6 期 如何通过机器学习挖掘实际业务价值？Cloudera（全球领先的企业级数据管理和分析平台提供商）
- 5、《全球数据交易实践、行业规范现状与政策法律问题研究》公安部第三研究所网络安全法律研究中心
- 6、中国电子学会《全球产业数字化转型趋势及方向研判》
- 7、人民日报/2021 年/8 月/31 日/第 009 版 理论以开源开放为抓手形成科技与产业新优势 江小涓
- 8、科技日报/2021 年/6 月/25 日/第 006 版 构建“开源”创新体系助力我国关键数字技术“弯道超车” 隆云滔 王晓明
- 9、第 20 卷第 3 期 2021 年 7 月北京交通大学学报（社会科学版）开源数字经济的创新逻辑：大数据合作资产视角 齐佳音 张国锋 王伟
- 10、中国电子学会《全球产业数字化转型趋势及方向研判》
- 11、德勤《2021 年技术趋势》
- 12、机器之心：2020-2021 全球 AI 技术发展趋势报告
- 13、《交通大数据需求分析与应用方向浅析》杨伟
- 14、中国电子学会《全球产业数字化转型趋势及方向研判》
- 15、<http://www.rmlt.com.cn/2021/0629/617612.shtml>
- 16、https://www.thepaper.cn/newsDetail_forward_15507104
- 17、<https://baijiahao.baidu.com/s?id=1714307003432866447&wfr=spider&for=pc>
- 18、https://www.sohu.com/a/448316216_99900352
- 19、<https://www.163.com/dy/article/G738KIUU053874C1.html>
- 20、<https://www.baogaoting.com/info/26775>
- 21、<https://baijiahao.baidu.com/s?id=1689589613297160806&wfr=spider&for=pc>
- 22、<https://max.book118.com/html/2021/0125/7124006101003046.shtml>
- 23、<https://baijiahao.baidu.com/s?id=1682326635054970592&wfr=spider&for=pc>
- 24、<https://www.doc88.com/p-91061766446170.html?r=1>
- 25、<https://baijiahao.baidu.com/s?id=1700793998968800869&wfr=spider&for=pc>
- 26、<https://zhuanlan.zhihu.com/p/341431556>
- 27、<https://ishare.ifeng.com/c/s/7zqHnZrK6vl>
- 28、<https://zhuanlan.zhihu.com/p/62482482>
- 29、<https://baijiahao.baidu.com/s?id=1714307003432866447&wfr=spider&for=pc>
- 30、https://www.sohu.com/a/348479375_472878
- 31、<https://weibo.com/ttarticle/p/show?id=2309404566851479797939&udaref=www.baidu.com>
- 32、<https://www.wxnmh.com/thread-9094870.htm>
- 33、<http://www.chinaaet.com/article/3000135950>
- 34、<https://wenku.baidu.com/view/954fdcf800768e9951e79b89680203d8ce2f6a9c.html>
- 35、<https://max.book118.com/html/2021/1101/8131034140004027.shtml>
- 36、<https://baijiahao.baidu.com/s?id=1702961389262572116&wfr=spider&for=pc>
- 37、<https://wenku.baidu.com/view/d1fddf0003020740be1e650e52ea551810a6c9b0.html>
- 38、<http://www.sykv.cn/m/view.php?aid=24060>



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