

北京市人工智能基础研究人才特征分析

冯琬婧^{1, 2}, 徐 峰¹

(1. 中国科学技术信息研究所, 北京 100038;
2. 科学技术部科技人才交流开发服务中心, 北京 100045)

摘 要: 北京是我国人工智能发展的重要地区, 汇集了全国最为丰富的人工智能人才资源。本文以北京市人工智能基础研究人才为研究对象, 通过对其结构特征进行分析, 在一定程度上了解我国人工智能基础研究人才队伍的现状。本文利用文献计量方法遴选出北京人工智能领域基础研究的核心人才, 对其发文情况、核心作者结构、研究主题等方面进行了分析研究。结果显示, 北京人工智能基础研究人才大多集聚在高校和科研院所, 企业基础人才较少; 大部分为 30~50 岁的中青年男性; 研究方向聚焦于深度学习、神经网络、自适应控制、优化等多个研究主题。

关键词: 人工智能; 人才结构; 核心作者; 机构分布; 共词分析

中图分类号: G311 **文献标识码:** A **DOI:** 10.3772/j.issn.1009-8623.2020.12.008

当前, 全球人工智能发展进入新一轮快速发展的阶段, 成为世界主要国家竞相部署的战略领域。中国高度重视人工智能发展, 2017 年发布实施《新一代人工智能发展规划》, 全面推动了我国人工智能的发展。尽管有很多研究表明, 中国人工智能发展已经接近甚至超过美国等主要发达国家, 但总体而言, 我国在人工智能基础理论和关键核心技术方面仍存在较大差距。特别是在人工智能高端基础研究人才和产业人才方面与美国等西方发达国家的差距更加明显。与此同时, 我国人工智能人才结构不合理, 人才类型偏重应用型, 基础类与技术类人才存量严重不足^[1]。

北京是我国人工智能发展的重要地区, 在推动全国科创中心建设过程中, 也对人工智能进行了全产业链布局^[2], 成为第一个国家新一代人工智能创新发展试验区, 并把推动人工智能基础理论原创研究作为重要的建设内容之一。北京聚集了全国最为丰富的人工智能人才资源, 拥有人工智能相关人才

近 4 万人, 是全国的近 6 成^[3]。本文聚焦北京人工智能基础研究人才, 分析研究其群体特征、学术产出、所属机构、研究领域等方面特征, 以期从某些程度上掌握当前我国人工智能基础研究人才的基本情况, 同时为我国进一步推动人工智能领域基础研究人才的培养提供参考借鉴。

1 相关研究概述

科技人才研究一直是学术界研究的热点, 其中以诺贝尔奖获得者为代表的“科学精英”人才更受关注。美国科学社会学奠基人罗伯特·金·默顿首次对具体时间和空间范围内的科学家群体的社会组织形式进行了研究^[4]。朱克曼·哈里特^[5]1977 年发表的《科学界的精英——美国的诺贝尔奖金获得者》, 可以说是对“科学精英”人才群体的研究中最具影响力的研究成果。之后, 学者们陆续围绕诺贝尔奖获得者, 从科学家出生顺序和其社会地位的关系^[6]、年龄与产出效率之间的关系^[7]等维度展开

第一作者简介: 冯琬婧 (1987—), 女, 在读硕士研究生, 助理研究员, 主要研究方向为科技战略与科技政策、科技情报研究。

项目来源: 科技部科技创新战略研究专项“我国新一代人工智能应用模式与实施路径研究”(ZLY201808)。

收稿日期: 2020-10-12

研究。

随着科学计量学的发展,基于论文作者计量分析的人才研究也日益增多,对基础研究人才的研究也大多围绕与科学论文相关的指标展开。科学计量学之父普赖斯在洛特卡定律的基础上研究了科学家人数与科学文献数量的关系,提出了著名的普赖斯定律和其他重要结论^[8]。加菲尔德则通过引文统计,系统检验了世界上被引次数最多的科学家^[9]。近些年随着信息科学的发展,越来越多的学者开始针对“高被引学者”^[10, 11]或“特定范围的核心学者”^[12]群体展开研究,如基础研究人才群体的名单^[13, 14]或领域专家推荐名单^[15];研究内容主要围绕基础研究人才队伍建设^[16, 17]、成长^[13, 14]、结构分布^[11, 18]等方面。

近年来,随着人工智能成为世界主要国家关注的重点领域,人工智能人才相关研究也成为各界的研究热点。如Yuan等^[19]从人工智能的发展趋势、人才流动和学术协作的演变方面研究了人工智能的演变,研究发现人工智能的研究热点已经从理论层面转移到应用层面;美国拥有数量最多的杰出人工智能科学家,并且吸引了大量其他国家的人工智能人才流入,而发展中国家的人工智能人才外流问题日益严重;人工智能精英之间的联系高度集中在协作网络中。王婷婷等^[20]结合国内外人工智能人才战略规划,从基础研究人才、技术研发人才和应用人才的培育需求方面,提出中国人工智能人才培养体系框架。王双等^[21]利用履历分析和社会网络分析方法,对图灵奖人工智能领域获奖者的履历和学术产出开展实证研究,探讨了科技人才成长的规律。高子平^[22]在产业层面,通过信息采集与实证调查,分析了上海浦东新区人工智能人才队伍的基本特征及面临的问题与诉求。范漪萍等^[23]围绕北京人工智能产业发展现状,分析了北京人工智能产业发展的人才需求,并提出精准引才的对策建议。

许多机构发布的人工智能研究报告也对人工智能的人才情况进行了分析。如中国的人工智能人才拥有量仅次于美国,高校和科研机构是人工智能人才的主要载体^[24];中国人工智能学术人才储备相对充足,越来越多的人工智能学者开始流入产业界开展技术或工程创新^[25];中国人工智能高水平人才不足美国的1/10,产业界重点领域人工智能人

才短缺,企业领军人才以海归华人为主,本土培育的高水平人才严重缺乏等^[25]。其中,北京在人工智能人才方面的优势地位也得到了广泛认可,北京人工智能基础研究人才不论从数量还是质量上看,都在国内处于领先的位置。如北京市人工智能人才投入量名列全国第一,江苏第二;各省市人工智能人才学术差距不显著,北京市居第一,上海市居第二^[24]。

总体而言,针对基础研究人才的已有研究多以“诺贝尔奖获得者”“高被引学者”“领域核心作者”“国家杰出青年科学基金获得者”等作为研究对象,主要关注这些高端人才的建设、成长、结构分布、评价等方面。对人工智能人才的研究主要聚焦在人工智能人才的培养、人工智能产业人才现状等方面,或是利用文献计量学和履历分析等方法,研究人工智能人才的协作网络、流动轨迹等。本文以北京人工智能基础研究人才群体作为研究对象,利用文献计量等方法,在分析北京人工智能基础研究人才总体情况的基础上,进一步基于高产和高被引论文遴选出以核心作者为代表的人工智能基础研究人才,对这些人才的群体结构、机构分布和研究主题进行分析研究,以期更加全面深入地了解北京人工智能基础研究人才的情况。

2 研究方法 with 数据

对基础研究人才的研究,在研究对象的遴选上大多基于与科学论文相关指标,如高被引学者^[10, 11]、基础研究人才群体名单^[13, 14]和领域专家推荐名单^[15]等。《中国统计年鉴(2019)》中也强调,基础研究“成果以科学论文和科学著作作为主要形式”。因此,借鉴已有的研究,本文从发表科学论文作者的角度去界定和遴选人工智能基础研究人才。在综合调研人工智能领域学术论文发表情况的基础上,本文选取2019年4月发布的《中国计算机学会推荐国际学术会议和期刊目录》(以下简称《目录》)中的人工智能领域的期刊和会议文献作为文献源。该《目录》由中国计算机学会(CCF)学术工委主持并组织中国计算机学会相关领域的专家完成,它既考虑了会议和刊物的影响力,也考虑了不同领域间的平衡,能在其中发表论文的作者在一定程度上体现了其学术影响力。数据检索与遴选过程如图1所示。

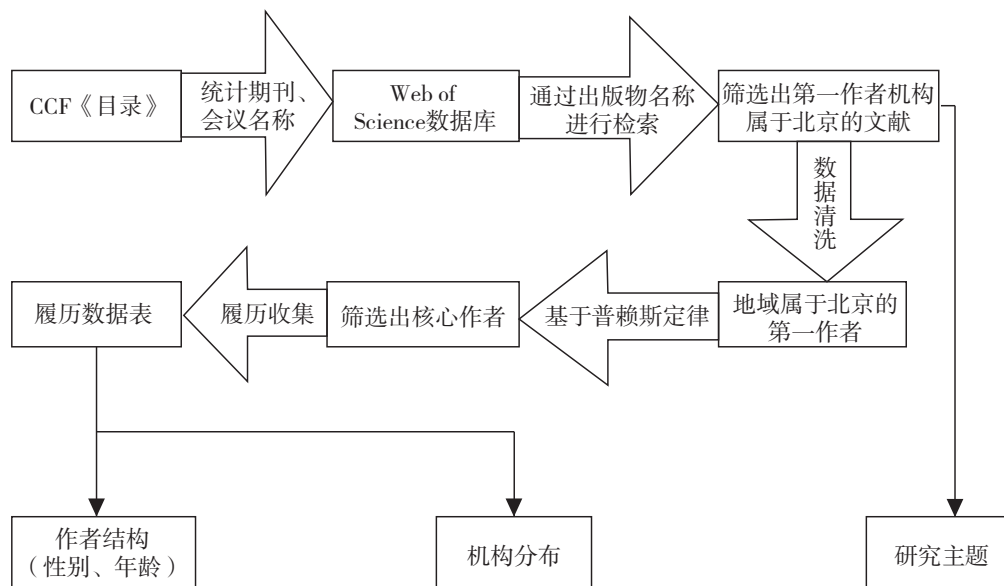


图1 数据获取及处理流程图

(1) 数据检索。基于《目录》人工智能领域中推荐的期刊和会议名称,从 Web of Science 数据库中检索获取文献数据。检索时间是 2020 年 1 月 27 日,限定的时间范围是 1990—2020 年,限定地区是中國大陸,共检索到文献 49 800 篇。

(2) 数据清洗。对检索到的 49 800 篇文献,去除综述、评论等其他类型文献,为了控制后续履历分析的数据量级,所以选取 2010—2019 年第一作者机构属于北京的文献,共 6 514 篇。之后,对 6 514 篇文献的第一作者姓名手工去重,得到 3 793 名作者。

(3) 核心作者遴选。基于普赖斯定律,对获得的 3 793 名作者从发文量和被引量两个维度^[26]计算核心作者范围阈值。其计算方式如公式 (1) (2):

$$K(p)=0.749\sqrt{\text{Max}(p)} \quad (1)$$

$$K(q)=0.749\sqrt{\text{Max}(q)} \quad (2)$$

$K(p)$ 表示核心作者的最低发文量,即发文量不少于 $K(p)$ 的作者才能够成为核心作者, $\text{Max}(p)$ 表示发文量最多的作者所发表的文献量; $K(q)$ 表示核心作者的最低被引量,即被引量不少于 $K(q)$ 的作者才能够成为核心作者, $\text{Max}(q)$ 表示被引量最高的作者的被引量。经计算,最后将发文量 3 篇

及以上、被引量 25 次及以上的作者作为核心作者,共有 329 人。

(4) 收集履历数据。通过文献、搜索引擎、科学家主页、机构官网等途径获取研究对象的履历信息。329 名核心作者履历数据最后确认了性别的有 285 人、博士毕业时间的有 275 人。将这些数据信息输入到统一格式的数据表之后,并逐条人工编码,形成如表 1 所示的北京 AI 基础研究核心作者履历数据表。

基于收集的履历数据,本文将主要从人才的群体结构、机构分布和研究领域三个大维度,综合运用文献计量分析、统计分析、社会网络分析等方法,对北京人工智能基础研究人才进行分析研究。

首先,基于机构发文量,从“量”的维度对北京人工智能领域基础研究的研究机构进行总体分析:这些机构在人工智能领域的学科优势是什么,以及这种优势是否出现梯度分层。

其次,聚焦核心作者,通过分析他们的性别结构和职业年龄结构,了解其结构特征;另一方面,通过他们的机构分布,从“质”的维度分析哪些机构有更强的研究力量和后发潜力,以及是否存在“二八定律”的分布现象。

表 1 北京 AI 基础研究核心作者履历数据表字段信息

基本信息				教育信息	
编号	英文姓名	中文姓名	性别	博士就读机构	博士毕业时间
1	Li Ming	李明	男	中科院自动化研究所	2011 年
工作经历					
首职时间	首职机构	现职时间	现职机构	其他就职时间和机构	
2011 年	美国罗德岛大学	2015 年	中科院自动化研究所	无	

再次,基于第一作者属于北京的论文,利用关键词共现、社会网络分析等方法,对这些论文的研究主题进行分析,以更深入地了解北京人工智能基础研究人才在人工智能领域的主要研究方向。

3 结果分析

基于上述方法遴选出发表机构属于北京的 6 514 篇文献,以及发文机构属于北京的第一作者 3 793 名。为了从中再挑选出较活跃、高贡献和高影响力的代表作者进行分析研究,基于普赖斯定律确认了 329 名核心作者。下文将结合数理统计及文献计量学相关方法,基于发文量、作者机构、机构发文量、职业年龄等数据,从北京人工智能领域基础研究核心作者结构特征、作者机构分布及研究主题三方面展开研究,以探析北京人工智能基础研究是“什么样的学者”“在哪”“研究什么”的问题。其中,“发文量”是指该作者于 2010—2019 年在《目录》人工智能领域期刊和会议论文中以第一作者身份发表文献的数量之和;“作者机构”是指作者发文时所在的机构;“机构发文量”是指这部分论文第一作者所属机构的计数之和;“职业年龄”^[27]是指从作者获得博士学位的时间算起的职业生涯年龄。

3.1 北京人工智能领域基础研究核心作者结构

为进一步了解北京人工智能基础研究人才的情况,下文基于所遴选出来的 329 名核心作者的相关履历信息,从性别结构、职业年龄结构和所属机构等方面进行了进一步分析。

3.1.1 性别结构

对 329 名核心作者进行履历调查,最终确定了 285 名作者的性别。如图 2 所示,其中男性 244 人,

占 85.61%,女性 41 人,占 14.39%。与加拿大 Element 人工智能公司 2019 年发布的《2019 年全球人工智能人才报告》(Global AI talent report 2019)^[28]中全球 21 个顶级人工智能学术会议女性作者大约占 18% 的比例相比,北京人工智能基础研究核心作者的女性比例还是偏低。

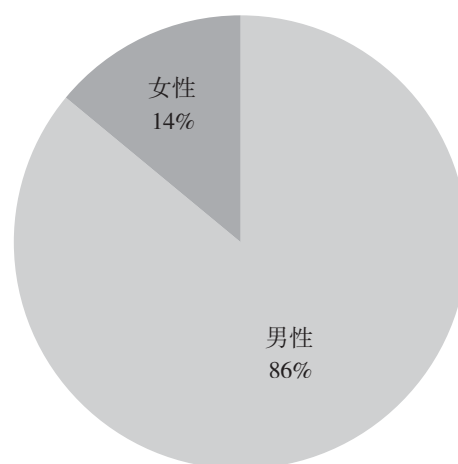


图 2 核心作者性别结构图

3.1.2 职业年龄结构

在 329 名核心作者中,最终确认了 275 名作者的博士毕业时间,并计算出相应的职业年龄。职业年龄为“1”表示该作者获得博士学位已有 1 年,以此类推。如图 3 所示,北京人工智能基础研究核心作者职业年龄 6~10 年的最多,有 98 人,其次为职业年龄 0~5 年的,有 83 人。故北京人工智能基础研究核心作者有超过 6 成职业年龄分布在 0~10 年,将近 3 成的职业年龄分布在 11~20 年。如果平均博士毕业年龄按 30 岁算,即北京人工智能基础研究核心作者超过 9 成是 30~50 岁的中青年学者,其中以青年为主力军。这与全球人工智

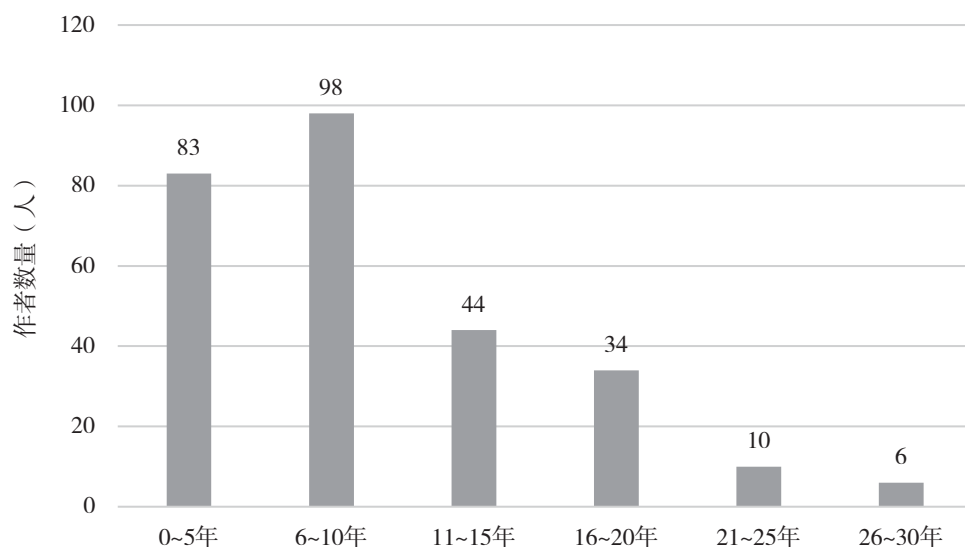


图3 核心作者职业年龄结构图

能领域活跃顶级学者超过6成为30~50岁的结果近似^[1]。

3.2 北京人工智能领域基础研究核心作者机构分布

从核心作者机构分布比例结构来看，329位核心作者分别就职于38家机构，如图4所示，23.4%属于中科院自动化研究所，13.98%属于清华大学，10.03%属于北京大学，8.81%属于北京航空航天大学

大学；排名前4位的机构核心作者拥有量之和占比已过半（56%），前10位的过8成，即20%的机构拥有80%的核心作者，体现了“二八定律”。

这些机构在AI领域发展和学科建设方面的规划各有侧重，但发展速度不同，有历史积累的机构相对发展更平稳，如中科院自动化研究所，该所是我国最早成立的自动化研究机构和最早开展类脑智能研究的研究机构，并设有包括模式识别

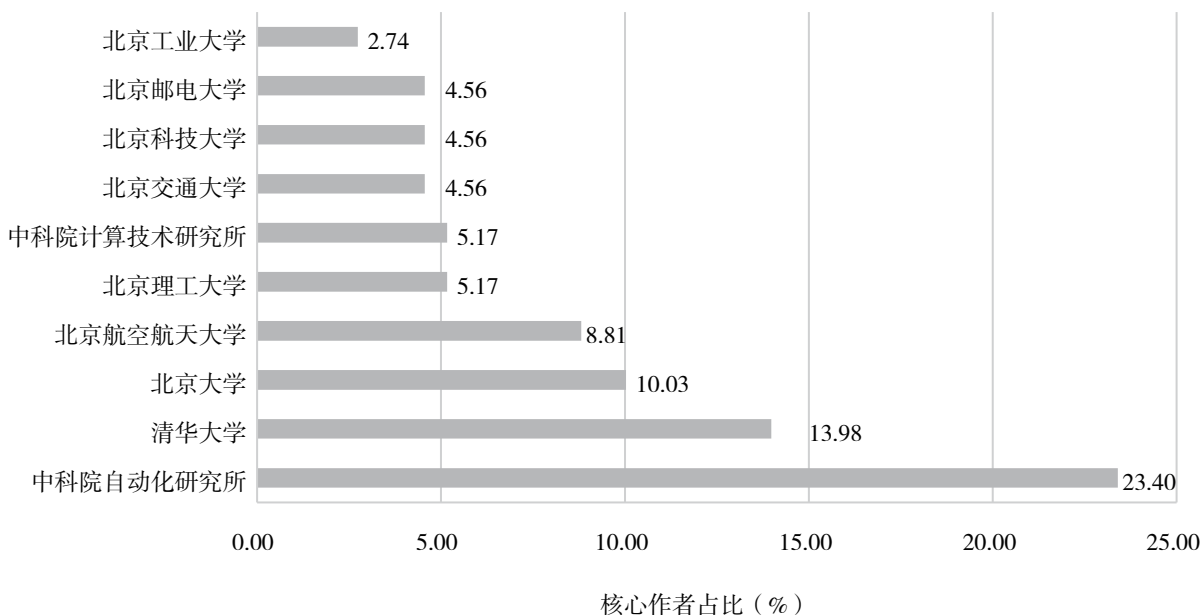


图4 核心作者占比前10机构

国家重点实验室、复杂系统管理与控制国家重点实验室等科研部门^[29]，上世纪90年代就率先布局了AI研究，2010年起又率先布局了类脑智能研究，在生物特征识别、机器学习、视觉计算、自然语言处理、智能机器人和智能芯片等领域取得丰富的研究成果。

高校是人工智能发展和人才培养的重要阵地，截至2019年1月底，我国已有59所高校成立了人工智能学院或人工智能研究院^[30]，所以高校在学术和人才资源等方面具有先天优势，高校的发文增量更加明显。例如，清华大学电子工程系早在1979年就开设了《人工智能导论》课，是国内高校开设最早的AI课程之一；上世纪80年代末，清华大学计算机系人工智能方向组建了“智能技术与系统”国家重点实验室，即著名的“人智所”；图灵奖得主姚期智院士是清华大学人工智能学堂班教授，同时还是清华大学学术委员会主任；清华大学人工智能研究院的院长则是计算机科学与技术专家张钹院士。北京大学在1988年就成立了视觉与听觉信息处理国家重点实验室，它是北大第一个国家重点实验室；2003年北大又率先成立了“智能科学与技术”本科专业，主要从事机器感知、智能机器人、智能信息处理和机器学习等交叉学科领域的学习。长期的积累加资源优势，使得学科发展速度立竿见影，特别是清华，2018—2019年，其AI课程注册人数比2012年增加了近16倍^[31]。

从机构性质角度来看（见图5），329位核心作者中，65%属于高校，32%属于科研院所，仅3%属于企业。这与全球人工智能人才所属机构类型的分布比例类似^[24]。值得注意的是，9位属于企业的

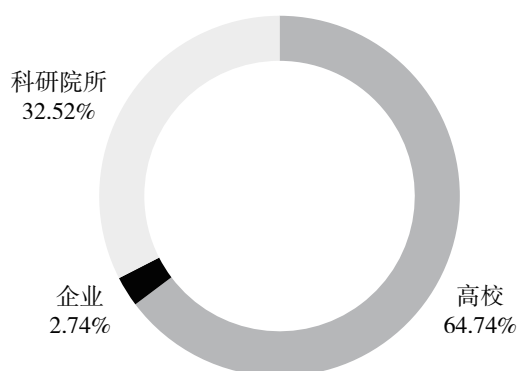


图5 核心作者所属机构性质分布图

作者中有8位都就职于外企（其中6位都属于微软亚洲研究院），仅1位来自国内互联网公司百度。虽然北京已成为全球人工智能企业数量最多的城市^[24]，然而从高产作者、高被引作者、核心作者分布数据来看，北京基础研究人才主要分布在高校和科研院所，企业基础研究人才投入少。

3.3 北京人工智能基础研究人才的研究主题及热点

为了进一步了解北京基础研究人才研究主题的变化情况，本文提取了6514篇北京基础研究人才发表论文的关键词，并将其按2010—2014年和2015—2019年两个阶段进行了划分，分别进行了词频分析和高频关键词的共现分析。

3.3.1 关键词词频分析

对6514篇文献中的关键词按频次排序进行统计，发现两个阶段高频关键词类似，但是第二阶段比第一阶段研究主题范围稍广。第一阶段排名靠前的关键词有：Face Recognition（面部识别）、Support Vector Machine（基本概念支持向量机）、Neural Networks（人工神经网络）、Classification（分类问题）、Adaptive Dynamic Programming（自适应动态规划）、Feature Selection（特征选择），这些关键词主要和计算机视觉与图像、机器学习、底层模型3个领域有关。第二阶段排名靠前的关键词有：Deep Learning（深度学习）、Face Recognition（面部识别）、Classification（分类问题）、Convolutional Neural Network（卷积神经网络）、Feature Selection（特征选择）、Support Vector Machine（基本概念支持向量机）、Neural Networks（人工神经网络）、Extreme Learning Machine（极限学习机器）等，这些关键词的涉及领域主要和机器学习、计算机视觉与图像、自然语言处理、底层模型4个领域相关。

3.3.2 高频关键词共现分析

关键词共现网络分析可以分析归纳出学科的研究热点、结构与范式^[32]。在分阶段词频统计的基础上，选取词频 ≥ 10 的高频关键词进行共现分析，利用Ucinet软件对共词矩阵进行可视化，如图6、图7所示。图中节点的大小代表不同的度（Degree），即与该节点相连线的条数；节点间线段的粗细则代表了节点间的关联强度。

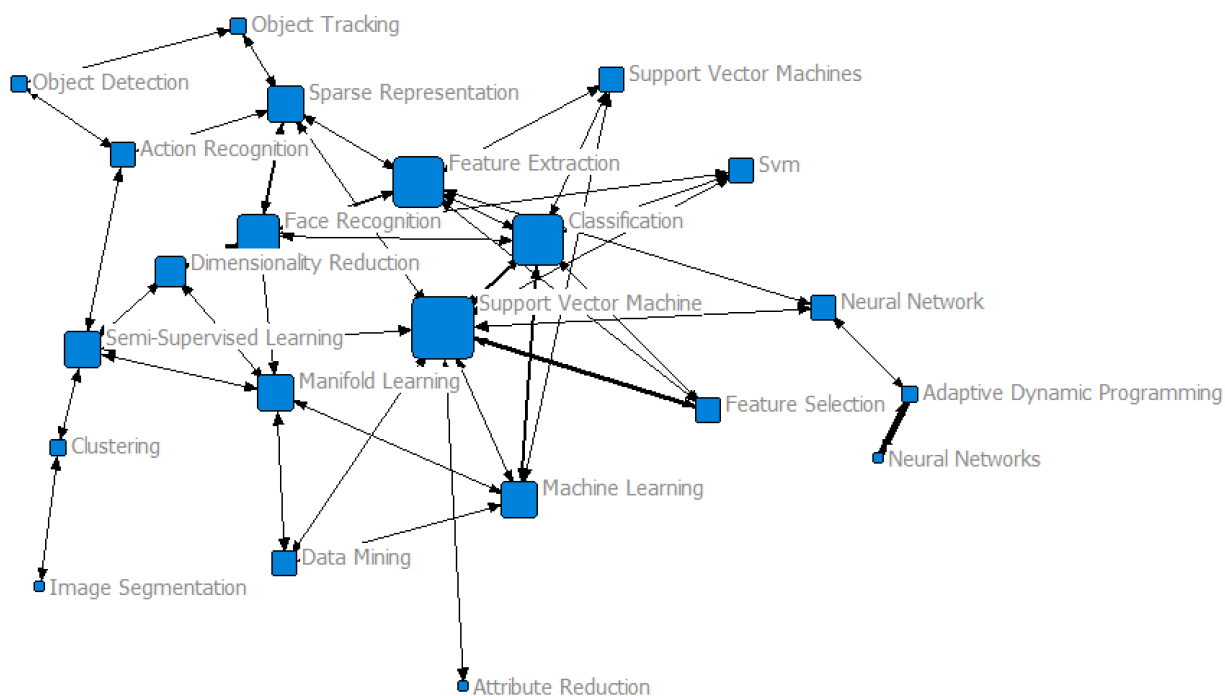


图 6 2010—2014 年高频关键词共现网络

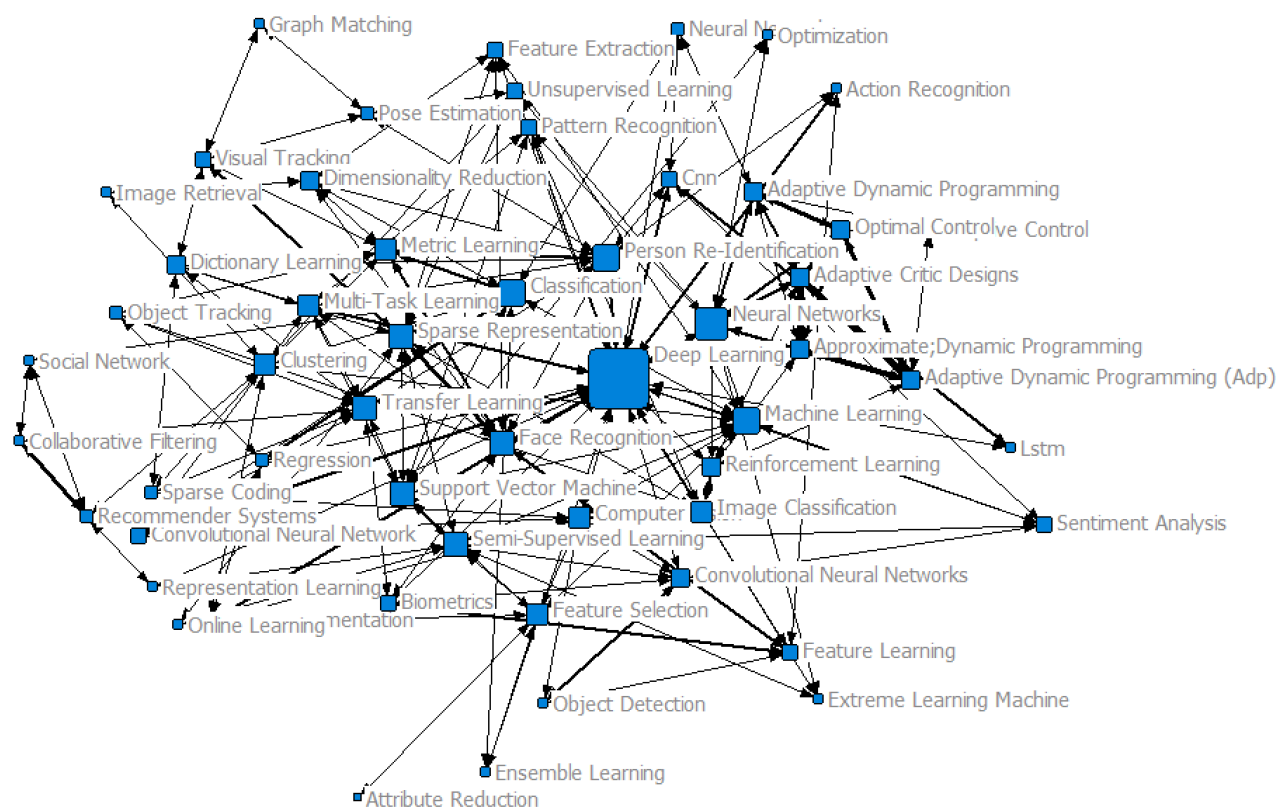


图 7 2015—2019 年高频关键词共现网络

度数中心度是研究网络图中节点影响力的重要指标,分为绝对度数中心度和相对度数中心度,“相对度数中心度”可以解决不同规模的网络图中点的局部中心度不可比较的问题^[33]。表2即是利用Ucinet进行网络中心度分析,生成点的度数中心度统计表。因为2010—2014年和2015—2019年的共现网络图规模不一致,用“相对度数中心度”更适合对比这两阶段的中心度。

保留两阶段相对度数中心度大于4的高频关键词,结果见表2。虽然2015—2019年的高频关键词共现网络规模更大(见图7),词频排序统计时高频词也明显多于2010—2014年的,但是经过相对度数中心度计算,发现2015—2019年有影响力的高频关键词反而不如2010—2014年的多。2010—2014年的相对度数中心度大于4的高频关键词有Support Vector Machine(基本概念支持向量

机)、Classification(分类问题)、Face Recognition(面部识别)、Adaptive Dynamic Programming(自适应动态规划)、Neural Networks(神经网络)、Dimensionality Reduction(降维)、Sparse Representation(稀疏表示)、Machine Learning(机器学习),这些关键词主要和机器学习、计算机视觉与图像、自然语言处理、神经网络等领域有关;2015—2019年的相对度数中心度大于4的高频关键词有,Deep Learning(深度学习)、Adaptive Critic Designs(自适应评价设计)、Dynamic Programming(动态规划)、Adaptive Dynamic Programming(自适应动态规划)、Optimal Control(最优控制),这些关键词则主要与机器学习相关。间接说明近些年北京人工智能学者对机器学习的研究更加聚焦和深入了。

机器学习是当今人工智能领域的热门话题,而

表2 高频关键词的度数中心度

2010—2014 年高频关键词度数中心度			2015—2019 年高频关键词度数中心度		
关键词	绝对度数中心度	相对度数中心度	关键词	绝对度数中心度	相对度数中心度
Support Vector Machine	17.000	8.586	Deep Learning	49.000	7.040
Classification	13.000	6.566	Adaptive Critic Designs	39.000	5.603
Face Recognition	11.000	5.556	Dynamic Programming	33.000	4.741
Adaptive Dynamic	11.000	5.556	Adaptive Dynamic Programming	33.000	4.741
Neural Networks	9.000	4.545	Optimal Control	28.000	4.023
Dimensionality Reduction	8.000	4.040			
Sparse Representation	8.000	4.040			
Machine Learning	8.000	4.040			

深度学习则是近10年机器学习领域发展最快的一个分支。《中国人工智能发展报告2018》^[24]也显示,近年全球高水平人工智能论文关注深度学习(Deep Learning)、神经网络(Neural Network)、自适应控制(Adaptive Control)、优化(Optimization)等多个方向,图灵奖得主约书亚·本吉奥(Yoshua Bengio)在神经信息处理系统大会(NeuIPS 2019)上的报告“From System 1 Deep Learning to System 2 Deep Learning”也讨论了深度学习发展的方向^[34]。大数据+深度学习正在成为当前人工智能各领域研

究的最重要范式,也是当前人工智能产业化的主要驱动力量^[35]。可见北京人工智能学者研究主题在一定程度上符合国际人工智能的研究方向。

4 总结与思考

本文基于2010—2019年在《目录》人工智能领域期刊会议上发表论文的北京作者,对这些作者的机构分布、研究主题,以及其中遴选出的核心作者的履历进行了计量分析,发现北京人工智能领域基础研究人才主要有以下几方面特征:

首先,在核心作者性别结构方面,北京 AI 基础研究核心作者的性别分布仍以男性为主。其中男性占比 88%,女性比例仅 14%。但相比科研领域女性学者的总体比例而言,女性学者在人工智能领域相对更加活跃。2020 年 3 月 8 日,清华大学-中国工程院知识智能联合研究中心、清华大学人工智能研究院与北京智源人工智能研究院发布的人工智能全球女性榜单(Women in AI)^[36],显示“AI 2000 最具影响力学者”^[37]中女性学者的整体比例仅为 9.7%。而具体到国家层面而言,中国入选榜单的学者中,女性仅占 6.9%,低于 9.7% 的国际基准线。这一现象也间接说明了,我国女性科技人才的“高位缺席”现象,即在高端科学家、科技领军队伍中,越往象牙塔的顶端,女性越少^[38]。这个现象受社会文化、用人机制、个人偏好等多种因素影响,需要相关的政策支持,以进一步促进女性科技人才的发展。

其次,在职业年龄结构方面,北京 AI 基础研究核心作者以 30~50 岁的中青年学者为主,发展潜力较大。如果平均博士毕业年龄按 30 岁算,即北京 AI 基础研究核心作者超过九成是 30~50 岁的中青年学者,其中以青年为主力军。不同于其它学科,AI 青年学者的学科影响力并不“轻”。以北京 AI 基础研究核心作者被引量最高的“80 后”学者何恺明为例,他同时位于“AI 2000 最具影响力学者”之“计算机视觉”领域和“机器学习”领域的榜单第一和榜单第二十。他目前就职于 Facebook AI 研究院,从“高考满分状元”到计算机视觉领域国际顶级会议 CVPR 最佳论文奖“首位华人得主”,他提出的 ResNet 和 Faster R-CNN 直接颠覆了整个计算机视觉和机器学习领域。

再次,在核心作者机构分布方面,中科院自动化所、清华大学、北京大学和北京航空航天大学的核心作者拥有量之和占比已过半(56%),而核心作者拥有量居前 10 位的机构,拥有了超过 80% 的核心作者,优势累积现象^[39]明显。其实在科学社会中,优势累积现象并不少见,如在 1901—1999 年获得诺贝尔奖最多的 20 家科研机构中,60% 位于美国^[40];我国国家自然科学基金中,“重点”“杰青”项目与“面上”“青年”项目相比,优势累积现象更明显^[41]。所以,预计未来北京人工智能基础研究领域的人才分层会更加突显。另一方面,北

京企业核心作者占比低,这也间接说明,北京企业的人工智能人才可能更多地投入于实验发展和应用研究,因为它能短期就为企业带来实际可见的利益。对比其他发达国家地区,除了伦敦 98% 的高质量论文作者分布在高校之外,硅谷、纽约和东京分布在企业的高质量论文作者比例相对较高^[23]。所以应尽快推进促进高校和科研机构人工智能基础研究人才与企业人才之间的流动“旋转门”机制,促进人才的有效流动。

最后,北京人工智能学者的研究主题在一定程度上符合国际人工智能研究方向。北京基础研究人才聚焦的深度学习、神经网络、自适应控制、优化等方向均是当前人工智能领域的研究热点。

本文的分析可以从一定程度上展示当前北京人工智能领域基础人才的基本情况,将有助于探索进一步推动北京人工智能基础人才发展的相关举措,如促进高校和科研机构与企业间人才的流动,鼓励更多企业开展人工智能基础研究,以更好地推动人工智能产业创新发展等。当然,本文的研究也存在一定的局限性:一方面,由于本文主要利用文献计量方法遴选基础人才,没有行业专家的参与,所以可能会导致研究对象缺少行业代表性,特别是产业界人工智能基础研究人才部分;另一方面,在研究主题方面仅以关键词为分析单元进行了研究主题的分析,数据的覆盖面尚不全面。后续,将进一步扩充数据来源,进一步分析挖掘人才的履历信息,以期更加系统地了解我国人工智能基础研究人才的发展状况。■

参考文献:

- [1] 腾讯研究院,BOSS 直聘. 2017 全球人工智能人才白皮书[R]. 北京:腾讯研究院,BOSS 直聘. 2017.
- [2] 任万霞. 按下科创中心建设加速键[J]. 北京观察, 2019(4): 22-23.
- [3] 北京日报. 北京拥有人工智能相关人才近 4 万人领跑全国[N/OL]. [2020-05-18]. <http://ai.people.com.cn/n1/2019/1118/c422228-31460589.html>.
- [4] 张煌. 中国现代军事技术创新高端人才研究[D]. 长沙:国防科技大学,2011.
- [5] 朱克曼·哈里特. 科学界的精英——美国的诺贝尔奖金获得者[M]. 北京:商务印书馆,1979: 1-336.

- [6] Clark R D, Rice G A. Family constellations and eminence: the birth orders of nobel prize winners[J]. The Journal of Psychology, 1982, 110(2): 281-287.
- [7] Stephan P E, Levin S G. Age and the Nobel prize revisited[J]. Scientometrics, 1993, 28(3): 387-399.
- [8] 邱均平. 文献计量学[M]. 北京: 科学出版社, 2019: 146-148.
- [9] 吴尔中. 利用“SCI”评选杰出科学家的情况分析[J]. 图书情报工作, 1984(3): 6-10.
- [10] Parker J N, Lortie C, Allesina S. Characterizing a scientific elite: the social characteristics of the most highly cited scientists in environmental science and ecology[J]. Scientometrics, 2010, 85(1): 129-143.
- [11] Parker J N, Allesina S, Lortie C J. Characterizing a scientific elite (B): publication and citation patterns of the most highly cited scientists in environmental science and ecology[J]. Scientometrics, 2013, 94(2): 469-480.
- [12] Laudel G. Migration currents among the scientific elite[J]. Minerva, 2005, 43(4): 377-395.
- [13] 陈晓剑, 李峰, 刘天卓. 基础研究拔尖人才的关键成长路径研究——基于 973 计划项目首席科学家的分析[J]. 科学学研究, 2011(1): 44-48.
- [14] 周建中, 赵璐. 我国基础领域科研人员职业发展状况研究[J]. 科学学研究, 2019(3): 476-483.
- [15] Peters H P F, van Raan A F J. A bibliometric profile of top-scientists a case study in chemical engineering[J]. Scientometrics, 1994, 29(1): 115-136.
- [16] 高润生, 曾国屏, 陈冬生, 等. 关于我国基础研究队伍建设的战略思考[J]. 中国软科学, 2001(4): 84-91.
- [17] 毛献峰, 宋福明, 王菲. 江苏省基础科学研究青年人才队伍建设现状分析——以 2012—2017 年江苏省获国家自然科学基金人才类项目资助情况为例[J]. 江苏科技信息, 2018(32): 4-9.
- [18] 杨岩, 姚长青, 刘志辉, 等. 我国主要城市群基础研究人才的时空分布研究[J]. 地理信息世界, 2019(3): 50-54.
- [19] Yuan S, Shao Z, Wei X, et al. Science behind AI: the evolution of trend, mobility, and collaboration[J]. Scientometrics, 2020, 124(2): 993-1 013.
- [20] 王婷婷, 任友群. 人工智能时代的人才战略——《高等学校人工智能创新行动计划》解读之三[J]. 远程教育杂志, 2018(5): 52-59.
- [21] 王双, 赵筱媛, 潘云涛, 等. 学术谱系视角下的科技人才成长研究——以图灵奖人工智能领域获奖者为例[J]. 情报学报, 2018(12): 1 232-1 240.
- [22] 高子平. 上海浦东新区人工智能人才队伍建设现状、问题及优化策略[J]. 北京教育学院学报, 2019(3): 66-72.
- [23] 范漪萍, 王晓迪, 张炜, 等. 北京人工智能产业发展人才需求分析及精准引才对策建议[C]. 北京科学技术情报学会. 科技情报创新缔造发展新动能论坛论文集. 北京: 北京科学技术情报学会, 2019.
- [24] 清华大学中国科技政策研究中心. 中国人工智能发展报告 2018[R]. 北京: 清华大学, 2018.
- [25] 中国科学技术发展战略研究院, 科技部新一代人工智能发展研究中心. 2019 中国新一代人工智能发展报告[R]. 北京: 科技部新一代人工智能发展研究中心, 中国科学技术发展战略研究院, 2019.
- [26] 钟文娟. 基于普赖斯定律与综合指数法的核心作者测评——以《图书馆建设》为例[J]. 科技管理研究, 2012(2): 57-60.
- [27] 刘俊婉. 科学合作对高被引科学家论文产出力的影响分析[J]. 情报科学, 2014(12): 77-81.
- [28] Kiser G, Mantha Y. Global AI talent report 2019[EB/OL]. (2020-01-30)[2020-06-09]. <https://jfgagne.ai/talent-2019/>.
- [29] 中国科学院自动化研究所. 研究所简介[EB/OL]. [2020-03-08]. <http://www.ia.cas.cn/gkjj/yjsjj/>.
- [30] 方兵, 胡仁东. 我国高校人工智能学院: 现状、问题及发展方向[J]. 现代远距离教育, 2019(3): 90-96.
- [31] 郑子亨, 王哲. 2019 年人工智能全景报告[N]. 中国计算机报, 2019-12-30.
- [32] 赵蓉英, 李飞. 基于社会网络分析方法的国内外信息计量比较研究[J]. 情报科学, 2013(2): 7-12.
- [33] 刘军. 整体网分析(第三版)[M]. 上海: 格致出版社, 上海人民出版社, 2019: 139.
- [34] 唐杰. 浅谈人工智能的下一个十年[J]. 智能系统学报, 2020(1): 187-192.
- [35] 李修全. 当前人工智能技术创新特征和演进趋势[J]. 智能系统学报, 2020(2): 409-412.
- [36] AMINER. AI 2000 | Women in AI[EB/OL]. [2020-03-08]. <https://www.aminer.cn/women-in-ai/country>.
- [37] AMINER. AI 2000 人工智能全球最具影响力学者[EB/OL]. [2020-03-08]. <https://www.aminer.cn/ai2000>.
- [38] 中国科协创新战略研究院中国科协调研(下转第 67 页)

Research on International S&T Collaboration Policy Elements in China's STI Policies

REN Xiao-ping¹, YANG Yun¹, LI Zi-yu¹, YANG Fan²

(1. Department of International Evaluation and Research, National Center for Science & Technology Evaluation, Beijing 100083;

2. School of Computer and Information Engineering, Central South University of Forestry and Technology, Changsha 410004)

Abstract: Based on the design of international science and technology cooperation (ISTC) policy analysis model and method, this paper cross-analyzes the ISTC elements of 96 policy documents issued in the 13th five-year period, summarizes the types, distribution, characteristics and main contents of China's ISTC policies, and then puts forward policy recommendations to the formulation of China's ISTC strategy in next period.

Key words: international science and technology cooperation; content analysis; national S&T strategy, plan and policy; major area

(上接第57页)

宣传部. 中国科技人力资源发展研究报告(2018)[M]. 北京: 清华大学出版社, 2020: 99-100.

[39] 赵万里, 付连峰. 科学中的优势积累: 经验检验与理论反思[J]. 科学与社会, 2014(2): 75-89.

[40] 刘俊婉. 从诺贝尔奖现象看科学创造的特征[J]. 科学

学研究, 2009(9): 1 289-1 297.

[41] 李丹, 马彪, 陈曦, 等. 国家自然科学基金分配中的“马太效应”及其变化分析[J]. 科技进步与对策, 2017(16): 20-25.

An Analysis of the Characteristics of Talents in Basic Research on Artificial Intelligence in Beijing

FENG Wan-jing^{1,2}, XU Feng¹

(1. Institute of Scientific and Technical Information of China, Beijing 100038;

2. Exchange & Development & Service Center of Science & Technology Talents of The Ministry of Science & Technology, Beijing 100045)

Abstract: As an important area for the development of artificial intelligence in China, Beijing has the most abundant AI talent resources in China. This paper, taking the AI talents in Beijing as the research object, aims to reveal the status quo of domestic talents teams for basic AI research through analyzing their structures. In this paper, the bibliometric method was used to select the core talents of basic research in the artificial intelligence field in Beijing, and their papers, core-author structures and research topics were analyzed and studied. The research results showed that most of the basic research talents of artificial intelligence in Beijing are concentrated in universities and scientific research institutes, and there are few basic talents in enterprises; they are mainly young and middle-aged males aged 30-50; and they mainly focus on research topics such as deep learning, neural network, adaptive control, optimization, etc.

Key words: artificial intelligence; talent structure; core authors; institution distribution; co-word analysis