

AliResearch
阿里研究院

智谱·AI

2023

全球数字技术 发展研究报告

数字技术前瞻 (第29期)



报告作者



叶静芸

jingyun.ye@aminer.cn

智谱 AI 研究咨询部科技产业研究员，生态学博士。曾经多次主持或参与“北京市十大高精尖产业研究”、“各地企业技术尽调”、“人工智能系列报告”等多个项目，积累了科技政策、人才政策、科技评估等多方面研究经验。重点关注人工智能、先进制造等领域。



鄂克

ke.wu@aminer.cn

智谱 AI 研究咨询部科技产业研究员、负责人，外交学院法学硕士。在核心期刊、知名媒体发表论文和评论性文章 20 余篇；10 余年科技产业研究咨询经验。



苏中

suzhong.sz@alibaba-inc.com

阿里研究院资深技术专家，计算机博士。阿里研究院资深技术专家，阿里研究院未来技术中心负责人，中国中文信息学会理事，中国计算机学会人工智能专委会常务委员。关注下一代互联网技术与产业发展。

前言

近年来，数字经济已经成为推动全球经济发展的新动力，各国纷纷出台战略规划，依托数字化技术实现新一轮产业变革，进而抢占全球价值链的制高点。中国政府也做出了重大战略部署，相继出台了多项数字科技发展引导政策，全面推进数字科技发展，并明确提出数字中国战略，构筑国家竞争新优势，加快建成社会主义现代化强国。

2022年11月，习近平总书记在世界互联网大会乌镇峰会的贺信中指出：“当今时代，数字科技作为世界科技革命和产业变革的先导力量，日益融入经济社会发展各领域全过程，深刻改变着生产方式、生活方式和社会治理方式。”国家十四五规划纲要专门设置“第五篇 加快数字化发展 建设数字中国”章节，对加快建设数字经济、数字社会、数字政府，营造良好数字生态做出明确部署。2021年12月，中央网络安全和信息化委员会日前印发《“十四五”国家信息化规划》，提出要建立健全规范有序的数字化发展治理体系，为推动开放、健康、安全的数字生态，加快数字中国建设进程提供方向。

学术论文和专利是科技成果的重要表现形式，也是研发实力的重要衡量指标。为了更好地展示全球数字科技的发展和知识产权保护情况，为推动数字科技产业发展提供建议以及引导制造业升级提供理论支持。我们基于AMiner¹科技情报平台的数据，利用文献计量方法，通过持续跟踪领域重要学术论文，研究分析论文的聚类和被引用模式，发现数字科技研究前沿；通过分析数字科技领域专利的申请情况和趋势，揭示相关技术创新活跃程度。在综合应用多种情报分析方法的基础上，系统、客观地揭示数字科技前沿热点。

¹J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. Arnetminer: Extraction and mining of academic social networks. In Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD'08), pages 990-998, 2008.

目录

一、报告要点	05
1. 科技实力对比	06
2. 科技人才储备实力研究	12
3. 数字科技十大趋势	20
二、全球科研实力对比	23
1. 数字科技论文	24
2. 数字技术专利	30
三、数字科技人才分布格局	38
1. 全球数字科技人才总体情况	39
2. 中国数字科技人才分布	42
3. 数字科技人才流动与合作	45
4. 全球数字科技顶尖科研团队	49
四、数字科技发展趋势分析	55
1. 研究方法	56
2. 数字科技 10 大技术趋势	57
3. 2022 数字科技热点技术	72
五、我国数字科技发展问题总结与建议	81
1. 中国数字科技论文两项关键指标将登顶全球第一	82
2. 中国是数字技术专利大国，而非强国	82
3. 中国数字科技发展前景在于向价值链高端跃升	82
4. 高层次人才不足需加大人才培养力度	82
5. 人才流失严重需采取措施引留人才	83
6. 企业人才储备薄弱需鼓励企业引育人才	83
六、附录	84
附录 I 技术趋势研判方法说明	85
附录 II 数字科技主题词	86

Chapter 1

报告要点

1. 科技实力对比
2. 科技人才储备实力研究
3. 数字科技十大趋势

1

科技实力对比

（一）全球数字技术论文的“量与质”比较

1. 中美数字科技论文整体影响力大致相当。尽管中国数字科技领域论文数量（506,775 篇）与美国（525,794 篇）还存在 4% 的差距，但是平均被引量中国比美国多 4 次 / 篇，这说明中国数字科技领域论文整体影响力与美国大致相当。英国和加拿大论文数量大幅落后于中美两强，但其论文平均被引量明显高于中美，这说明英加两国数字科技论文影响力在全球仍占有重要地位。




















2. 中国在卓越研究成果方面明显落后于美国。某领域全球被引量 Top 1% 论文（以下简称“Top 1% 论文”）被称为“顶尖论文”，其代表该领域卓越的学术研究成果。在数字科技领域，中国 Top 1% “顶尖论文”数量（7,096 篇）明显少于美国（9,634 篇），且平均被引量也明显落后于美国。

3. 在全球数字化浪潮中，欧美国家是数字科技基础研究创新的引领者。除中国外的亚洲国家——印度、日本和韩国也跻身全球论文数量前 10 强国家行列，但它们的身影没有出现在 Top 1% 论文数量前 10 强行列。在 Top 1% 论文数量前 10 强国家中，除中国和澳大利亚外，其他 8 国（美、英、德、加、法、意、荷、西）均为欧美国家，且平均被引量均在 300 次 / 篇以上，而中国仅为 202 次 / 篇。

4. 中国科学院²的数字科技论文发表量高居全球第一。在全球数字科技论文量前 10 强机构中，排名前三的机构依次是中国科学院（59,487 篇）、美国加州大学（49,111 篇）和法国研究型大学联盟（UDICE）（48,217 篇）；除中国科学院外，中国还有一家机构即中国科学院大学（18,381 篇）也跻身前 10 强；平均被引量最高的是哈佛大学和美国能源部，前者为 50 次 / 篇，后者为 45 次 / 篇，显示两者在数字科技领域显示出较强的研究和创新能力。

²注：中国科学院是一个庞大的科研学术机构，下设 11 个分院、100 多家科研院所、130 多个国家级重点实验室和工程中心，目前有正式职工 6.9 万余人，在学研究生 7.9 万余人。

表 1 全球数字科技论文产出国家排名

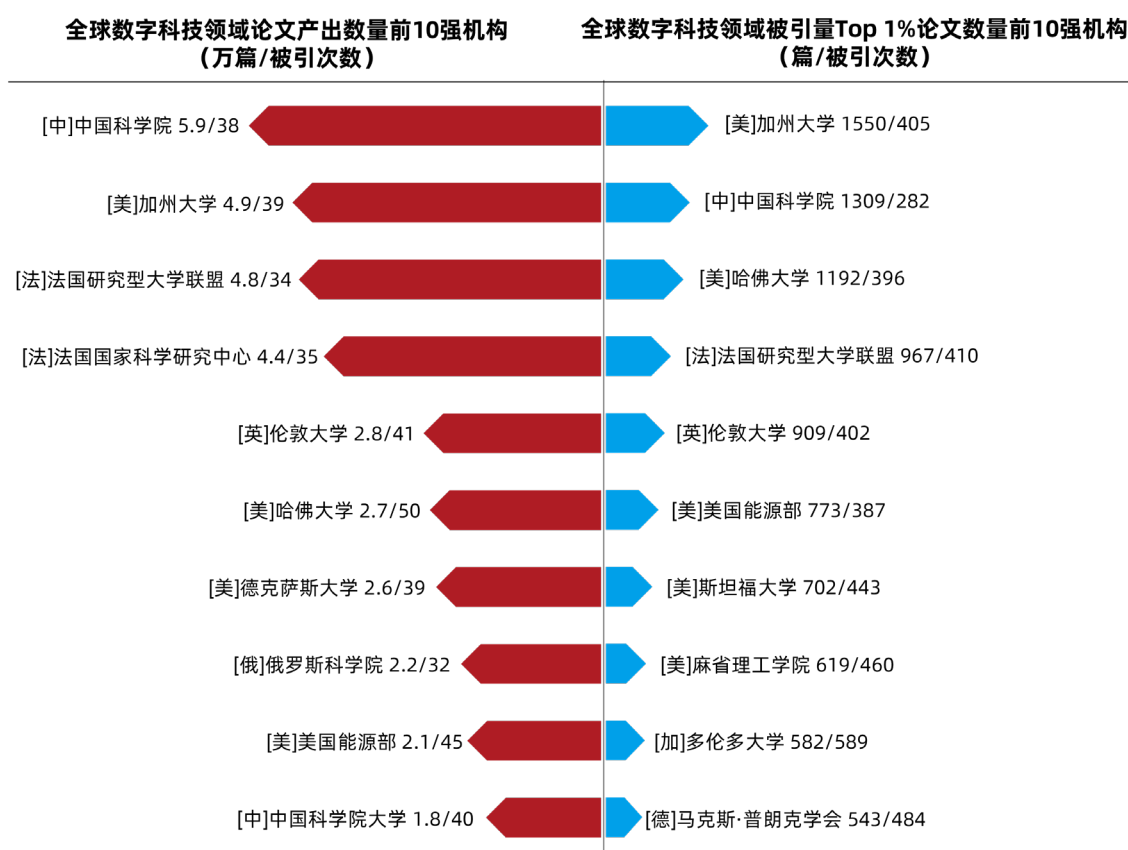
数字科技论文总产出					引用量 Top1%的数字科技“顶尖论文”				
排 名	国家	论文量 (万篇)	被引量 (次/篇)		排 名	国家	论文量 (篇)	被引量 (次/篇)	
1	 美国	52.6	14		1	 美国	9,634	312	
2	 中国	50.7	18		2	 中国	7,096	202	
3	 德国	15.1	12		3	 英国	3,611	308	
4	 英国	13.8	25		4	 德国	2,760	304	
5	 印度	12.3	15		5	 加拿大	2,065	357	
6	 日本	9.9	9		6	 澳大利亚	2,033	282	
7	 意大利	9.8	17		7	 法国	1,642	387	
8	 法国	9.7	15		8	 意大利	1,551	331	
9	 加拿大	9.6	22		9	 荷兰	1,382	386	
10	 韩国	8.0	11		10	 西班牙	1,325	377	

说明：论文检索时间范围为 2012 年 1 月至 2021 年 12 月。
数据来源：AMiner 科技情报平台。

5. 美国加州大学 Top 1% 论文发表量位居全球第一。从数字科技 Top1% 论文数量来看，美国加州大学最多，为 1,550 篇；其次是中国科学院，为 1,309 篇；第三是美国哈佛大学，为 1,192 篇。多伦多大学和马克斯·普朗克学会的 Top 1% 论文数量分别排名第 9 名和第 10 名，但平均被引量分别名列第 1 和第 2 名，这说明这两所学术机构在全球数字科技领域生产出最具影响力或者卓越成果。

6. 令人感到鼓舞的是，中国在总体论文和 Top 1% 论文数量增长上与美国的“黄金交叉”均已出现。2019 年，中国在总体论文发表数量上与美国实现“黄金交叉”，而 2020 年中国在 Top 1% 论文数量上与美国再次实现“黄金交叉”，这说明中国不仅在论文总量超过美国，而且在“顶尖论文”数量上超越美国，并逐年扩大与美国的优势。

图 1 全球数字科技领域论文产出数量前 10 强机构



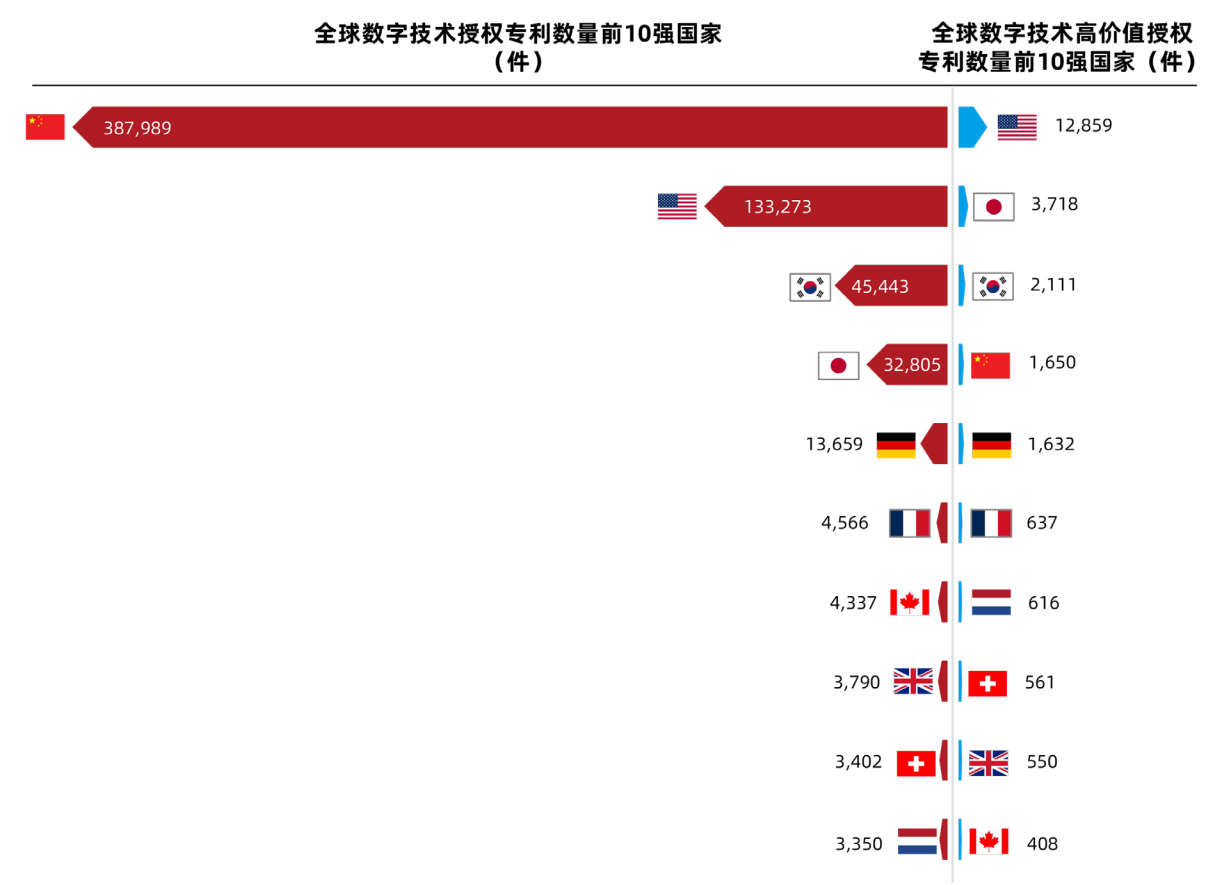
说明：1. 上图条形长度表示论文产出数量大小；2. 论文检索时间范围为 2012 年 1 月至 2021 年 12 月。
数据来源：AMiner 科技情报平台。

（二）全球数字技术专利比较，中国是数字技术专利大国，而非强国

1. 中国数字技术专利数量全球遥遥领先。中国共 387,989 件数字技术专利，是排名第 2 美国的 2.9 倍，是排名后 9 位国家总和的 1.6 倍，中国数字应用技术整体研发能力在全球首屈一指。

在数字科技领域专利数量排名前 10 机构中，中国机构占一半，其分别是百度、国家电网、腾讯、平安科技和华为。另外，老牌信息技术公司——国际商业机器公司（IBM）专利储备仍高居全球机构之首，显示深厚的数字技术积淀。分析发现，中国数字技术实力最强机构来自于能源、互联网、通信、金融等行业的科技巨头，以及 211、985、双一流和以理工为特色的全国重点大学，如西安电子科技大学、浙江大学、东南大学、华南理工大学等。

图 2 全球数字技术授权专利数量前 10 强国家排名



说明：专利检索时间范围为 2012 年 1 月至 2021 年 12 月；
数据来源：AMiner 科技情报平台。

2. 中国数字技术高价值专利（市场价值 100 万美元以上）数量在全球相对落后。从数量上看，中国以 1,650 件高价值专利数量排名全球第 4 位，落后于美国的 12,859 件、日本的 3,718 件以及韩国的 2,111 件。

3. 在全球高价值专利前 10 强机构中，中国的华为和阿里巴巴入榜。韩国三星、微软公司和谷歌公司是全球高价值专利前三大巨头，数量依次为 1,061 件、630 件和 592 件。其他前 10 强机构依次是、高通、华为、IBM、苹果、飞利浦、阿里巴巴以及索尼。中国的华为和阿里巴巴入榜。

4. 中国数字科技巨头高价值专利储备实力与国际巨头差距仍然较大。中国高价值专利数量最多的是华为，阿里巴巴和腾讯分别位列第 2 和第 3 位，专利数量分别为 261 和 127 件，与三星、微软和谷歌等国际巨头的专利储备实力相差较大。

表 2 中国数字科技高价值专利数量前 5 强机构

排名	公司名称	高价值专利数 (件)	简介
第 1 名	 HUAWEI 华为技术有限公司	373	1987 年创立，是全球领先的 ICT（信息与通信）基础设施和智能终端提供商
第 2 名	 Alibaba 阿里巴巴集团控股有限公司	261	1999 年创立，是一家以技术驱动，开拓数字经济时代商业基础设施的数字智能公司
第 3 名	 Tencent 腾讯 腾讯科技(深圳)有限公司	127	1998 年成立，中国最大互联网综合服务提供商之一，重要业务是致力于企业数字化升级
第 4 名	 ZTE 中兴 中兴通讯股份有限公司	80	1985 年成立，是全球领先的综合通信解决方案提供商，中国最大的通信设备上市公司
第 5 名	 tsmc 台积电(中国)有限公司	78	2003 年成立，台积电独资设立的子公司，生产线宽 0.35 微米及以下大规模集成电路

说明：1. 本报告将市场价值为 100 万美元及以上的专利视为高价值专利；
2. 专利检索时间范围为 2012 年 1 月至 2021 年 12 月。
数据来源：AMiner 科技情报平台。

5. 中美数字技术专利实力差距在价值链高端部分。虽然中国专利总量是美国的 2.9 倍，但美国高价值专利是中国的 8 倍，特别是越往价值链高端，美国的优势越明显。可见美国占据了全球数字科技产业链的高端部分，而中国专利集中在价值链低端。从近 10 年中美高价值专利数量走势看，中美之间的差距并非出现明显的缩窄态势。

6. 全球数字技术高价值专利储备实力在较长时间内仍然维持“一超多强”格局。从近 10 年高价值专利数量变化态势看，全球主要数字技术强国均处于下降态势（原因之一是专利价值的显现需要一定时长），但是美国与日本、韩国、中国、德国等数字技术强国仍保持较大优势，很难预期后者在较短时期内能赶超前者。

（三）中国数字科技发展前景在于向价值链高端跃升

专利是倾向于应用技术的科研成果，其市场价值直接反映该技术应用产品处于产业价值链的位置。从专利市场价值分布看，中国数字技术专利价值在 30 万美元以下的占 98%，因此中国数字技术产业仍处于全球价值链低端。论文是倾向于基础研究的科研成果，往往是科技创新突破的先导。未来中国数字科技的发展前景在于从日益强大的基础研究成果中实现实际应用转化，推动数字技术产业向价值链高端跃升。

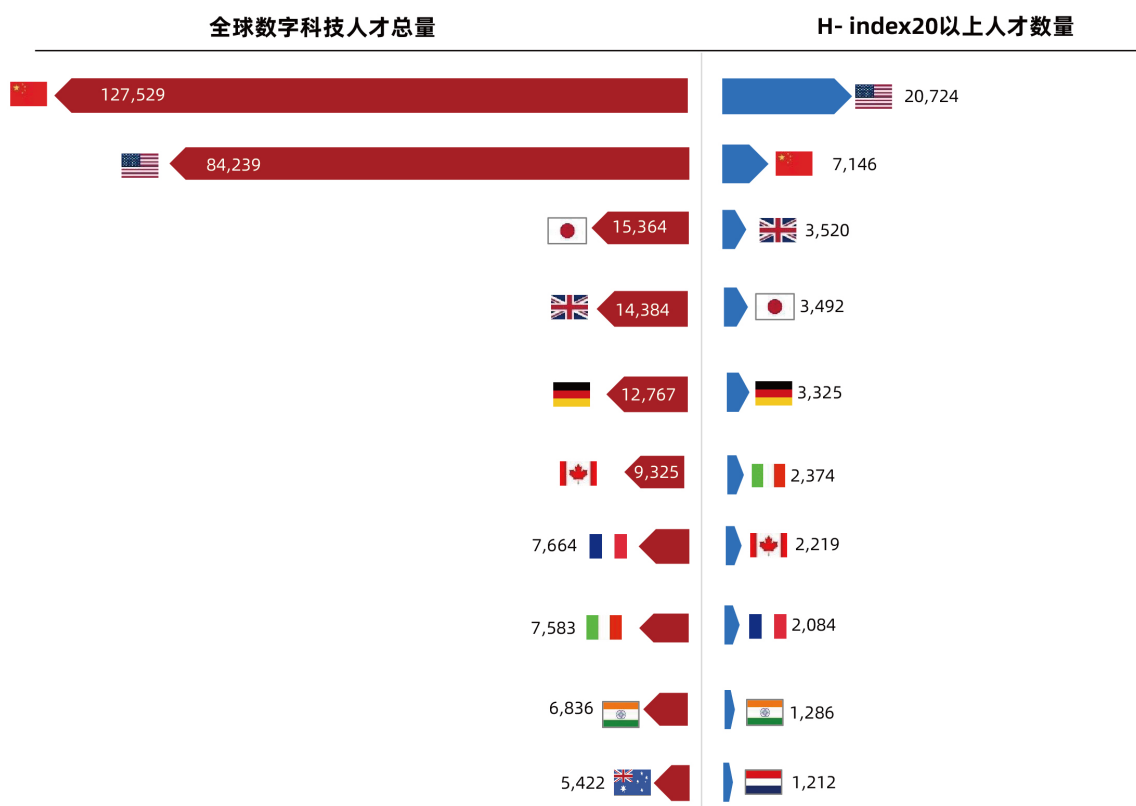
2

科技人才储备
实力研究

(一) 中国人才基数大，但高层次人不足，且严重落后于美国

全球数字科技人才总量为 77.5 万人，其中中国有 12.8 万人，位居第一，占全球总量的 17%，是排名第二的美国的 1.5 倍，是排名第三的日本的 8.3 倍。但中国数字科技高层次人才（即 H-index \geq 20，下同）只有 0.7 万人，仅占全球总量的 9%，仅为位居第一的美国（2.1 万，占全球 25%）的 35%。这说明我国数字科技人才基数很大，但高层次人才储备不足，且远落后于美国。

图 3 全球数字科技人才数量前 10 强国家



数据来源：AMiner 科技情报平台

（二）中国高层次人才集中在高校，而美国高科技公司人才储备不逊色于顶尖高校

从全球数字科技人才机构分布看，中国科学院以 4,722 人的数量名列第一名，以较大优势领先于排名第二的美国加州大学（2,623 人）；进入全球前 10 强的中国机构还有两家，即中国科学院大学（1,462 人，第六名）、清华大学（1,305 人，第八名）。

表 3 全球数字科技人才数量前 10 强机构

全球数字科技人才前 10 强机构				全球 H-index≥20 数字科技人才前 10 强机构			
排名	机构		人才储备 (人)	排名	机构		人才储备 (人)
1		[中]中国科学院	4,722	1		[美]加州大学	492
2		[美]加州大学	2,623	2		[美]谷歌	196
3		[法]UDICE 法国研究型大学联盟	2,534	3		[美]斯坦福大学	178
4		[法]法国国家科学研究中心	2,298	4		[美]微软	163
5		[英]伦敦大学	1,508	5		[中]中国科学院	144
6		[中]中国科学院大学	1,462	6		[中]北京大学	135
7		[美]哈佛大学	1,454	7		[美]卡内基梅隆大学	133
8		[中]清华大学	1,305	8		[美]华盛顿大学	131

9		[美]德克萨斯大学	1,209	9		[中]清华大学	129
10		[俄]俄罗斯科学院	1,123	10		[中]上海交通大学	119

数据来源：AMiner 科技情报平台

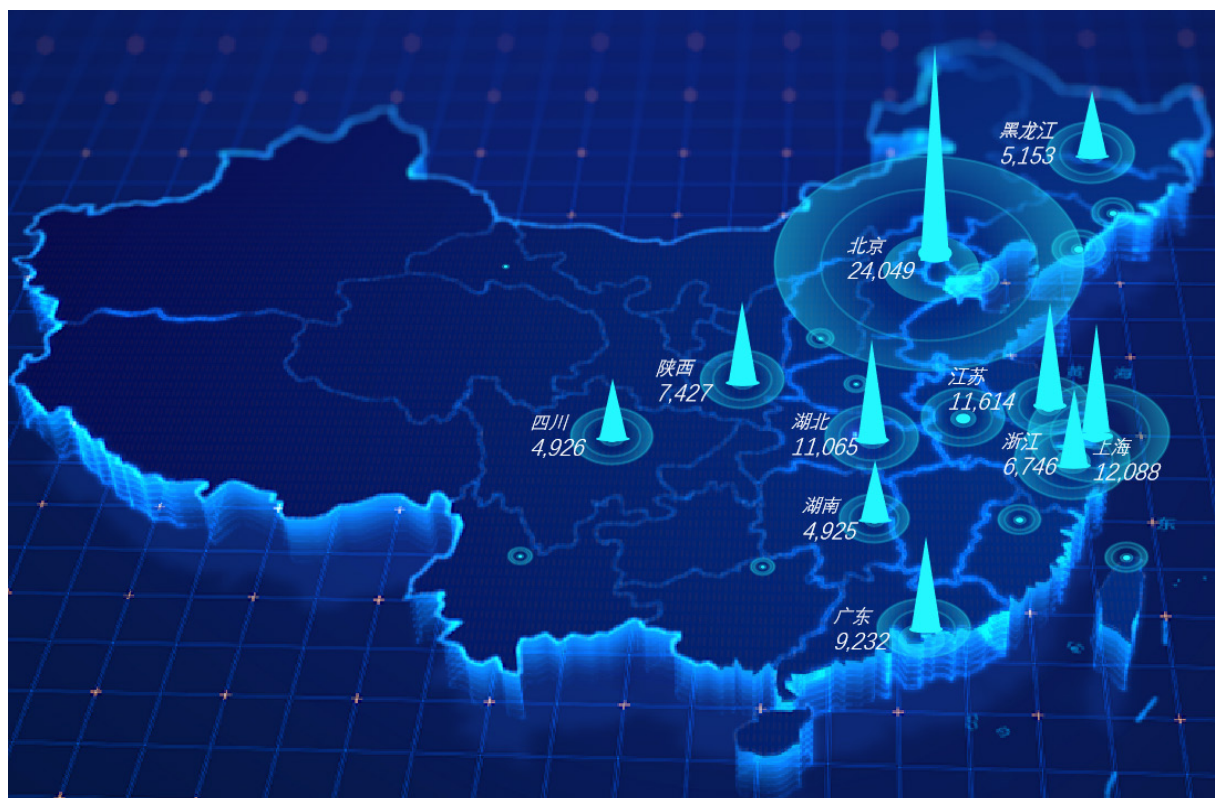
在全球高层次人才前 10 强机构中，美中各占 6 家和 4 家。第一名被美国加州大学摘取，其以 492 人的高层次人才储备大幅度领先于全球其他机构。4 家中国机构依次为中国科学院（144 人，第 5 名）、北京大学（135 人，第 6 名）、清华大学（129 人，第 9 名）、上海交通大学（119 人，第 10 名）。分析发现，中国科学院是中国数字科技人才储备结构的缩影：虽然人才基数大幅领先，但高层次人才明显落后（占其人才总数的 3%，落居全球第五）。

美国相对于中国的另一个优势是，身处科研一线的高科技企业的高层次人才储备实力强于中国，该国的谷歌、微软公司的高层次人才数量排名分别位居全球第二名和第四名，该优势将有力地推动美国数字技术持续创新发展，并引领全球数字技术应用新趋势。

（三）北京和上海是中国两大人才高地，前者人才储备为后者 2 倍

从国内数字科技人才分布看，北京人才总数优势明显，共 2.4 万人，占全国近 2 成，是排名第二的上海的 2 倍；北京的高层次人才数量为 1,076 人，占全国 1.5 成，仍然是排名第二的上海的 2 倍。另外，东部地区的广东、江苏、浙江、香港、黑龙江，以及高校资源较丰富的湖北、陕西、四川、湖南等内陆省份也是我国数字科技人才分布较多的地区。

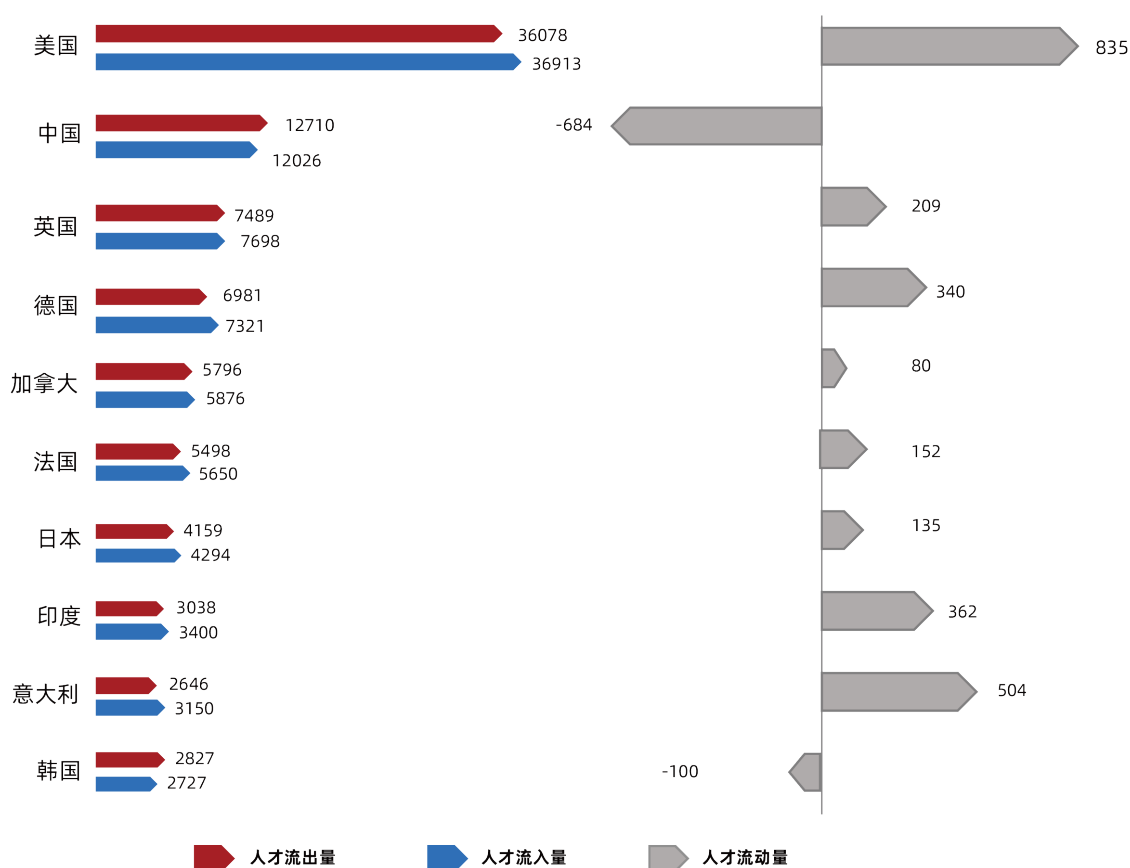
图 2 中国数字科技人才高地 (Top 10)



(四) 中美分别为全球最大的人才净流失国和净流入国

美国是数字科技人才流动最活跃的国家，也是人才净流入最大的国家。过去 10 年，其人才净流入 835 人，位居全球第一；其次是意大利，净流入 504 人。中国则为全球人才净流失第一大国，流失人数达 684 人；其次是韩国，净流出 100 人。由此可见，我国数字科技人才流失严重，在人才引进以及用好、留住人才等方面的工作显得尤为迫切。

图 4 2012-2021 年全球主要国家数字科技领域人才流动对比

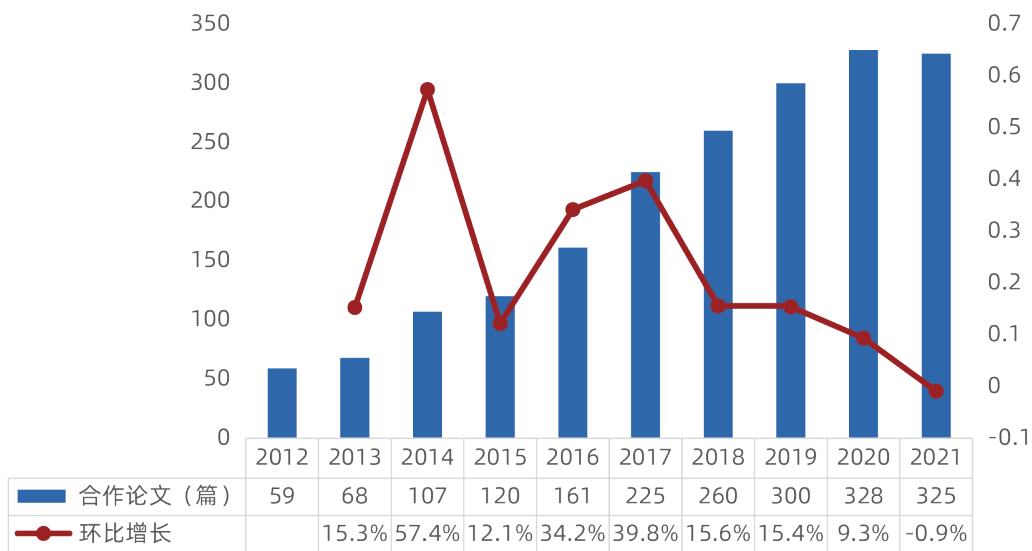


数据来源：AMiner 科技情报平台

（五）中美学者合作为全球最热络的双边合作，但从2018年开始出现降温

2012-2021年，在数字科技领域，中美学者共合作产出被引量 Top 1% 的论文 1,954 篇，为全球科研成果产出最丰硕的双边合作（第二为美英学者合作，共产出 1,528 篇）。但从2018年开始，中美两国学者合作论文增长率一直处于下降态势，在2021年更是出现负增长（-0.9%）。可见，自从2018年美国推出所谓“中国行动计划”后，中美在数字科技最卓越研究层次的合作受到严重冲击。

图 5 2012-2021 年数字科技领域中美学者合作关系走势



说明：图中合作论文为中美两国学者合作产出被引量 Top1% 的“顶尖论文”数量

（六）全球数字科技顶尖科研团队基本被美国机构包揽

全球数字科技顶尖科研团队（Top 10）基本来自美国顶尖高校（如斯坦福大学、卡内基梅隆大学、加利福尼亚大学、麻省理工学院、哥伦比亚大学、康奈尔大学等）、科技巨头（如谷歌、微软、IBM 等）下设的科研机构（如下表所示）。

表 4 全球数字科技顶尖科研团队 (Top 10)

排名	顶尖团队		重点研究方向
1		[美]谷歌 X 实验室	人工智能、人机交互
2		[美]斯坦福大学自然语言处理组	自然语言处理
3		[美]卡内基梅隆大学机器学习系	机器学习
4		[美]加利福尼亚大学伯克利分校 人工智能研究实验室(BAIR)	人工智能
5		[中]微软亚洲研究院	人工智能、图形学、分布式系统
6		[美]MIT 计算科学与人工智能实验室 (CSAIL)	人工智能
7		[美]IBM Thomas J. Watson 研究中心	基础科学、计算机科学
8		[美]哥伦比亚大学数据科学研究所	数据科学
9		[美]康奈尔大学计算科学系	计算机科学
10		[美]美国贝尔实验室	网络、通信

（七）我国数字科技人才工作任重道远

首先，高层次人才不足，需加大培养力度。要注重复合培养模式，重点培养有潜力的青年人才；鼓励对外学术交流。其次，人才流失严重，需采取措施引留人才。吸引海外高层次人才特别顶尖华人青年人才是回国发展，用好并留住高层次人才。再次，企业人才储备薄弱，需鼓励企业引育人才。身处科研一线的企业是科技创新的重要主体，有必要鼓励企业加强内部创新环境建设，并在重大专项、重点研发计划论证和实施过程中，支持企业承担政府科研项目，为企业引育人才和提高基础科研能力注入动力。

3

数字科技 十大趋势

趋势一：生物大数据

随着对生命系统的不断深入探究和各种其他高通量组学技术的产生和发展，生物信息学的研究范畴不断扩大，各种组学数据（转录组、蛋白质组、非编码 RNA 组、表观遗传组、代谢组、宏基因组等）以及生物系统层面的解读不断扩展，生命科学从定性描述开始实现动态、精准和定量解读。

趋势二：生成式对抗式网络算法

作为人工智能学界的热门方向，生成式对抗式网络算法（Generative adversarial networks）已法被广泛应用于图像和视觉、语音和语言、信息安全等领域。未来，随着 GANs 为代表的深度学习不断迭代，AIGC 百花齐放，产出效果或将逐渐逼真接近至人类作品。

趋势三：沉浸式扩展现实娱乐平台

扩展现实是元宇宙连接虚拟与现实的关键设备，随着扩展现实产业链和技术不断发展、内容应用逐渐繁荣，整个元宇宙娱乐生态正在持续丰满，沉浸式扩展现实娱乐平台有望迎来爆发。

趋势四：量子原型样机和专用处理器的研制

当前，量子计算机的研制已从以院校、研究所为主的基础性研究阶段，过渡到以 Google、IBM 等创新企业为主体的“量子霸权”研究阶段。随着量子计算机研制成果的不断涌现，通用量子计算机研制问题将在未来五年内被攻破，可实际使用的量子处理器将会落地。

趋势五：AI 解码蛋白质结构

2021 年，Deepmind 公司的蛋白质解码预测系统——AlphaFold 横空出世。与此同时，公司还公布了约 35 万种蛋白质的结构，该项成果因此入选《科学》2021 年度十大科学突破。随着人工智能技术的不断发展，对具有内在无序特性的蛋白质以及通过翻译后修饰或环境条件改变结构的蛋白质建模问题在将来有望解决。

趋势六：移动边缘计算网络

随着物联网、5G、工业自动化、智能制造的兴起，处于物理实体和工业连接之间的移动边缘计算越来越发挥出重要的作用，也逐步实现了集中化和智能化，其低时延、高带宽、个性化、高安全性、高隐私性等特性满足了分布式服务、自动工业控制的需求。

趋势七：可解释的 AI

由于人类社会的价值观念和价值体系存在多元化的特点，未来发展具有可解释性的人工智能，加强人机之间的理解，让人工智能技术跳出“黑箱”，建立可解释、可理解、可信任的人工智能体系成为趋势。

趋势八：基于算法模型和安全隐私的联邦学习技术

近年来，联邦学习作为解决数据孤岛问题的重要技术引起业界广泛关注，并被广泛应用于金融、医疗健康以及智慧城市等领域。未来，联邦学习技术的发展将与边缘计算、区块链和网络安全等多个领域关联，更好的发挥其隐私性、高效性和便捷性。

趋势九：混合计算

随着万物互联市场的发展，产业界逐渐认识到单一的计算方式不能解决所有问题，“混合计算”借鉴了异构计算的思想——用不同的计算资源处理适合该结构的任务，构建出某领域专用的高效应用组件，从而更好地满足无线互联、视频处理、图像识别、智能制造等多领域的高效处理需求。

趋势十：能源区块链

针对能源互联网应用系统中各个模块存在的问题，区块链技术能够依靠自身的链式特征和独特的区块结构融合到能源互联网中的各个层面，解决能源互联网系统中的相关问题。未来，创建区块链一体化能源系统可有效促进清洁能源的产量和能源的及时高效消费利用，提高地区性能源的综合利用率，在保证经济性和稳定性的基础上，实现节能环保目标。

Chapter 2

全球科研 实力对比

1. 数字科技论文
2. 数字技术专利

1

数字科技论文

本篇以数字科技领域的 ESI³ 核心论文（以下简称“论文”）、被引量 Top 1% 论文（以下简称“Top 1% 论文”）数据为研究基础，分别从论文的国家 and 机构分布两个方面进行分析。基于 AMiner 科技情报系统，共检索到数字科技领域内论文 2,204,950 篇，其中 Top 1% 论文 23,138 篇。

（一）国家分布

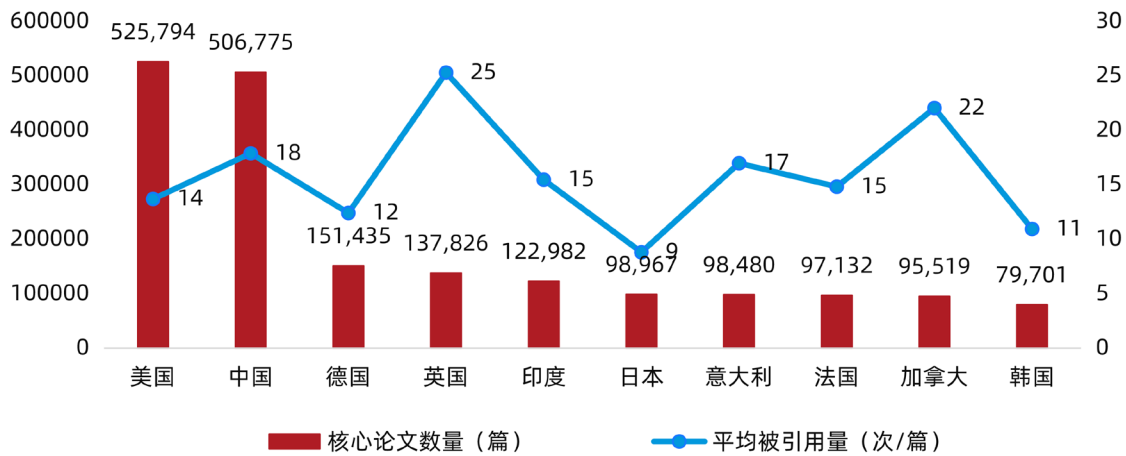
论文总量前 10 强国家

从论文数量上看，美国的论文数量为 525,794 篇，位居全球第一，论文平均被引量为 14 次；中国的论文数量为 506,775 篇，论文平均被引量为 18 次。虽然中国的相关论文数量低于美国，但是被引用数量更具优势。第三至五位的德国、英国和印度的论文数量分别是 151,435 篇、137,826 篇和 122,982 篇。

从论文的平均被引量上看，英国的平均被引量最高，为 25 次 / 篇。虽然加拿大的论文数量相对较少，但平均被引量高达 22 次 / 篇，位列第二。中国论文平均被引量位列全球第三。意大利（17 次 / 篇）和印度（15 次 / 篇）研究成果的认可度也较高，分列第四和第五位。

³ESI 被称为基本科学指标数据库 (Essential Science Indicators Database)，是国际公认的评价大学、学术机构和国家 / 地区国际学术水平和影响力的重要评价指标工具之一。

图 6 数字科技领域论文数量前 10 强国家



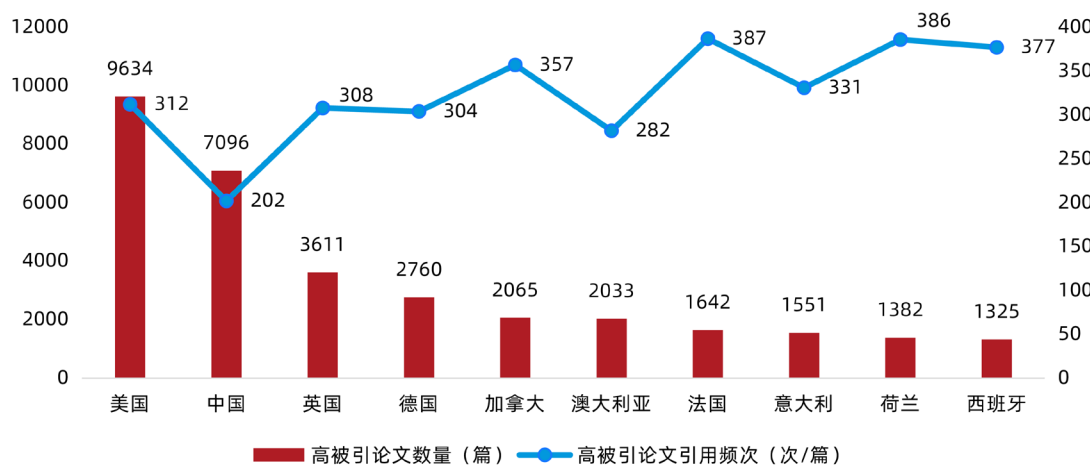
说明：论文检索时间范围为 2012 年 1 月至 2021 年 12 月。
数据来源：AMiner 科技情报平台。

Top1% 论文前 10 强国家

被引量进入前 1% 的“顶尖论文”被称为引领相关领域最卓越的研究成果。下图显示了数字科技领域 Top 1% 论文数量前 10 强国家状况。

从论文数量和平均被引量综合分析看，拥有最多 Top 1% 论文的国家是美国为 9,634 篇，但被引状况表现一般；第二名为中国，数量为 7,096 篇，约为第三名英国（3,611 篇）的两倍，但相比于英国 202 次 / 篇的被引量，中国研究论文的国际影响力还待提升。其余七个国家 Top 1% 论文数量明显少于前三名，但其中六个国家的平均被引量都在高于 300 次 / 篇，其中法国（387 次 / 篇）、荷兰（386 次 / 篇）、西班牙（377 次 / 篇）和加拿大（357 次 / 篇）的被引量 Top 1% 论文最受关注。整体而言，欧美国家的数字科技理论研究成果更受国际认可。

图 7 数字科技研究领域被引量 Top 1% 论文数量前 10 强的国家

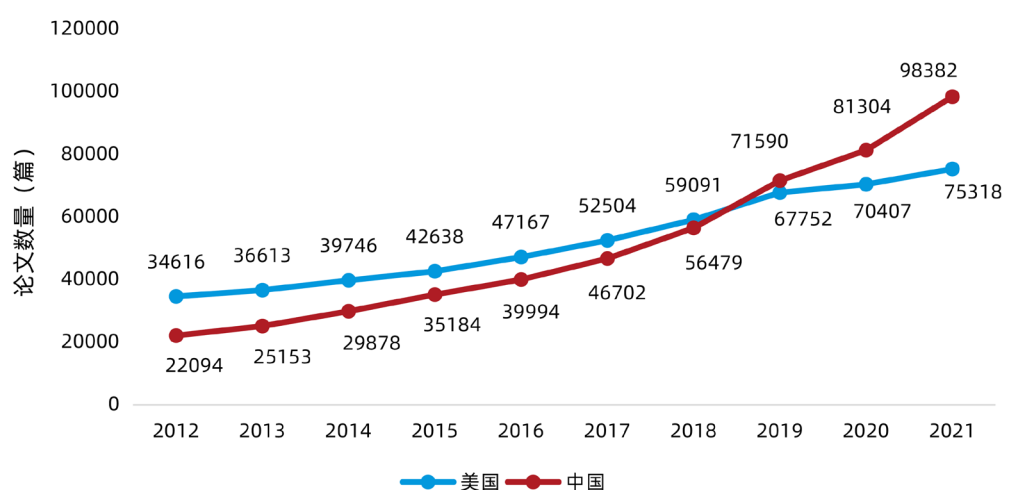


说明：论文检索时间范围为 2012 年 1 月至 2021 年 12 月。
数据来源：AMiner 科技情报平台。

中美论文发表数量变化态势

在 2012-2018 年间，美国的年产论文数量多于中国，但随着中国科研实力的逐步增强，中国逐年缩小与美国的差距。到 2019 年，中国的论文数量超过美国，并逐渐扩大对美优势。2021 年中国数字科技领域的论文为 98,382 篇，而美国为 75,318 篇。

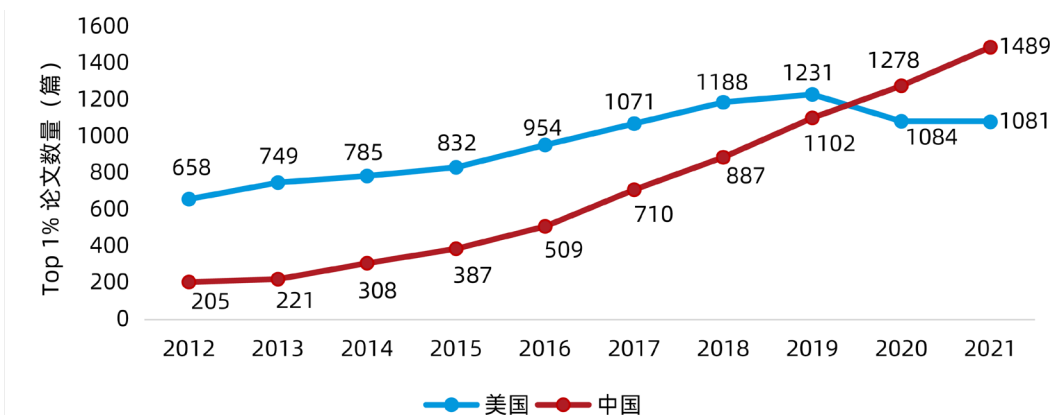
图 8 2012-2021 年中美两国论文发表数量变化态势



数据来源：AMiner 科技情报平台。

2020 年之前，美国 Top 1% 论文年产量多于中国；2020 年之后，中国的被引量 Top 1% 论文数量超过美国，并逐渐扩大优势，2021 年中国的论文数量为 1,489 篇，美国的论文数量为 1,081 篇。

图 9 2012-2021 年中美两国被引量 Top1% 论文数量变化



说明：论文检索时间范围为 2012 年 1 月至 2021 年 12 月。
数据来源：AMiner 科技情报平台。

从以上中美两国在数字科技领域论文发表态势看，2019 年中国在数量上超过美国，2020 年中国在质量上超过美国。

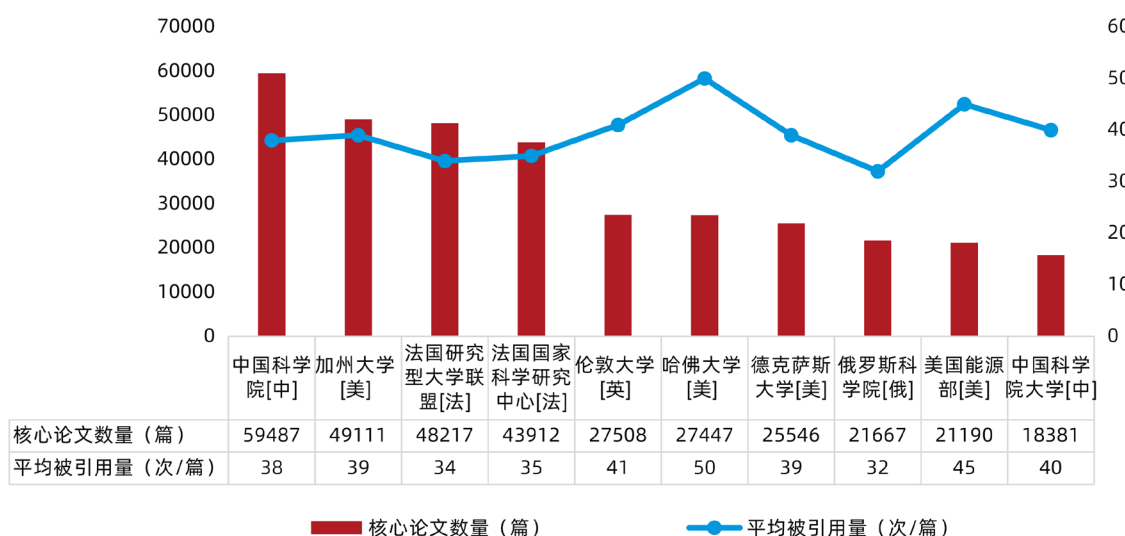
(二) 机构分布

论文发布量前 10 强机构

从机构发表的论文数量和引用率看，中国科学院拥有的论文数量最多，为 59,487 篇，引用率处于中等水平（38 次 / 篇）；论文发表量排名第二的是美国加州大学，共有论文 49,111 篇，被引量 38 次 / 篇；第三是法国研究型大学联盟（UDICE），共有论文 48,217 篇，被引量为 34 次 / 篇，该联盟于 2020 年由法国 10 所顶尖大学组成，显示在数字科技领域较强的研究能力。平均被引量最高的是哈佛大学，为 50 次 / 篇，论文量为 27,447 篇；平均被引量排名第二的机构是美国能源部，为 45 次 / 篇，论文量为 21,190 篇，该部是美国联邦政府在基础科学研究方面最主要的管理和资助机构，下设 24 个国家实验室和技术中心，在数字科技领域显示较强的研究和创新能力。

除中国科学院外，中国还有一家机构即中国科学院大学进入前 10 强机构行列，其论文量为 18,381 篇，被引量为 40 次 / 篇。

图 10 全球数字科技领域论文数量前 10 强机构



说明：论文检索时间范围为 2012 年 1 月至 2021 年 12 月。

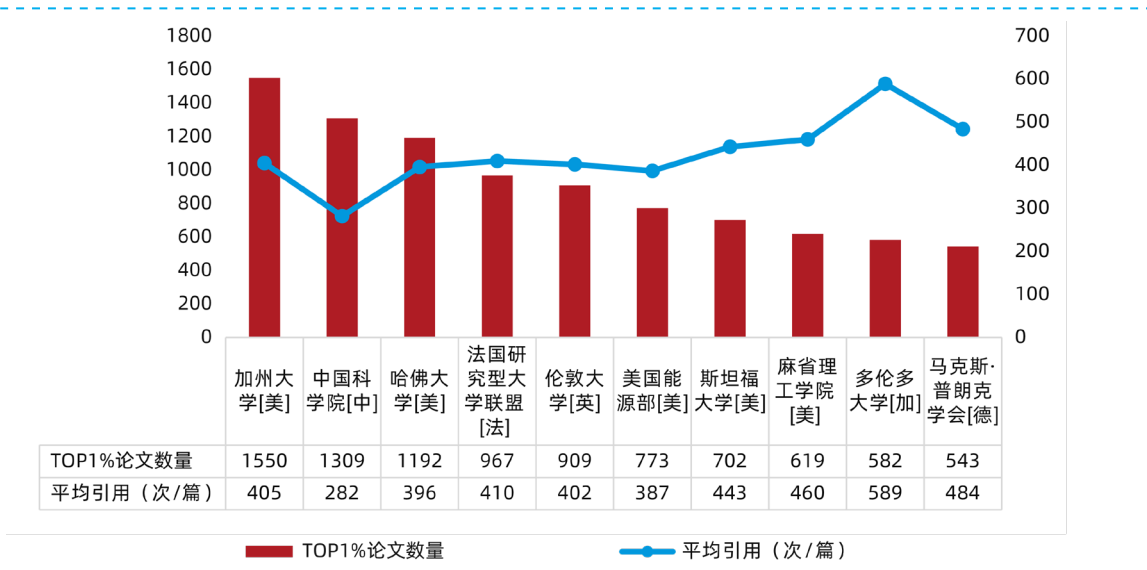
数据来源：AMiner 科技情报平台。

Top1% 论文数量前 10 强机构

从数字科技被引量 Top1% 论文数量来看，美国加州大学最多，为 1,550 篇；其次是中国科学院，为 1,309 篇；第三是美国哈佛大学，为 1,192 篇。进入被引量 Top 1% 论文数量前 10 强机构还有法国研究型大学联盟、英国伦敦大学、美国能源部、美国斯坦福大学、美国麻省理工学院、加拿大多伦多大学以及德国马克斯·普朗克学会。

其中，多伦多大学和马克斯·普朗克学会的被引量 Top 1% 论文数量分别排名第 9 名和第 10 名，但平均被引量分别名列第 1 和第 2 名，说明这两所学术机构在全球数字科技生产出了最具影响力的学术研究成果。

图 11 全球数字科技领域 Top1% 论文数量前 10 强机构



说明：论文检索时间范围为 2012 年 1 月至 2021 年 12 月。
数据来源：AMiner 科技情报平台。

2

数字技术专利

本篇以数字技术相关的专利数据为研究基础,从专利的国别分布、价值分布两个维度对进行详细分析。基于 AMiner 大数据平台,共检索到数字科技领域同族授权专利(以下简称“专利”) 669,942 件。

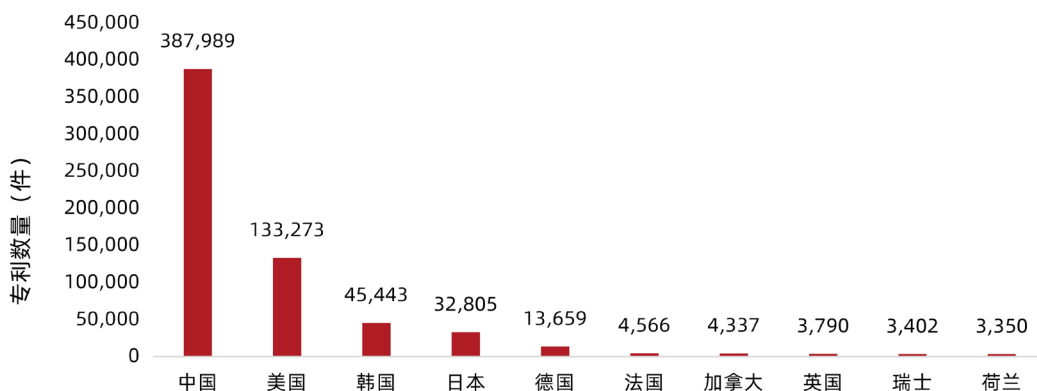
(一) 专利分布

全球专利数量前 10 强国家

在数字科技领域,全球专利数量前 10 强国家或地区如下图所示。从专利数量来看,中国共 387,989 件,遥遥领先于其他国家,是排名第 2 的美国的 2.9 倍,是排名后 9 位国家总和的 1.6 倍。可见,中国数字应用技术研发整体水平在全球首屈一指。

亚洲的韩国和日本分别以 45,443 件专利和 32,805 件专利位居第 3 名和第 4 名,其他跻身专利数量前 10 国家依次是德国、法国、加拿大、英国、瑞士和荷兰等欧美国家。

图 12 数字科技专利数量前 10 强国家

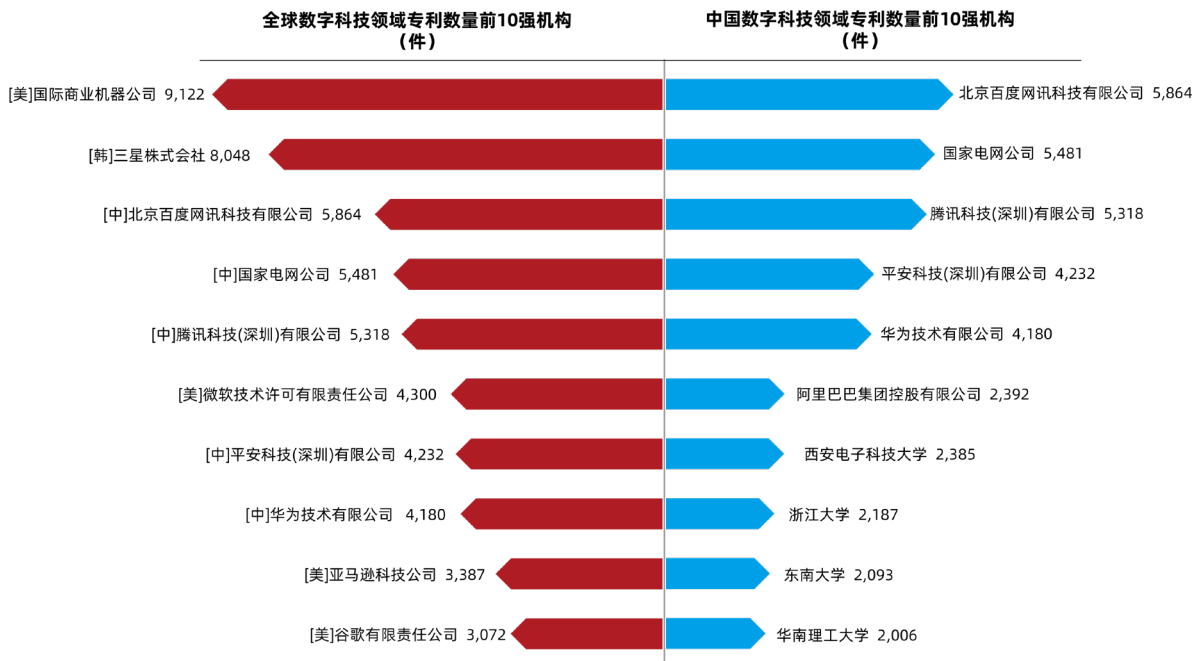


说明: 专利检索时间范围为 2012 年 1 月至 2021 年 12 月;
数据来源: AMiner 科技情报平台。

专利数量前 10 强机构

数字科技领域专利数量排名前 10 机构如下表所示。全球专利前十强机构中中国机构占据半数席位，其余为 4 家美国机构和 1 家韩国机构。其中，美国 IBM 公司拥有的专利数量最多，共 9,122 件，可见 IBM 在数字科技领域具非常深厚的技术积淀；韩国三星以 8,048 件专利优势居于第二，与 IBM 差距较小；第三名是百度公司，共有 5,864 件。进入前 10 强的中国机构还有国家电网、腾讯、平安科技和华为，它们分别位居全球第 4、第 5、第 7 和第 8 名。中国 TOP10 机构包括 6 家企业和 4 所高校，前者有百度、国家电网、腾讯、平安科技、华为和阿里巴巴，后者包括西安电子科技大学、浙江大学、东南大学和华南理工大学。由此可见，中国数字技术实力最强机构来自于能源、互联网、通信、金融等行业的科技巨头，以及 211、985、双一流和以理工为特色的全国重点大学。

图 13 全球数字科技领域论文产出数量前 10 强机构

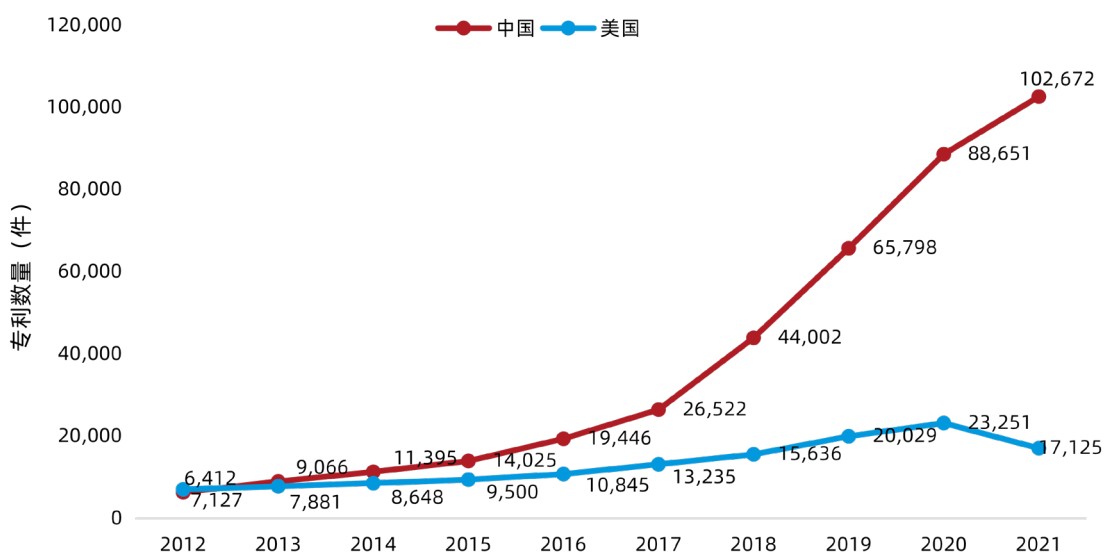


说明：1. 上图条形长度表示论文产出数量大小；2. 论文检索时间范围为 2012 年 1 月至 2021 年 12 月。
数据来源：AMiner 科技情报平台。

中美专利数量变化态势

中国一直是数字科技领域的专利研发大国。2012 年中美数字科技专利数量差距较小，2013 年之后中国的专利研发速度不断加快，而美国的增长幅度较小，中美之间的差距持续扩大。截至 2021 年，中国的数字科技专利数量达到了 102,672 件，是同年美国（17,125 件）专利的 6 倍之多。

图 14 2012-2021 年中美两国专利数量变化



说明：专利检索时间范围为 2012 年 1 月至 2021 年 12 月；
数据来源：AMiner 科技情报平台。

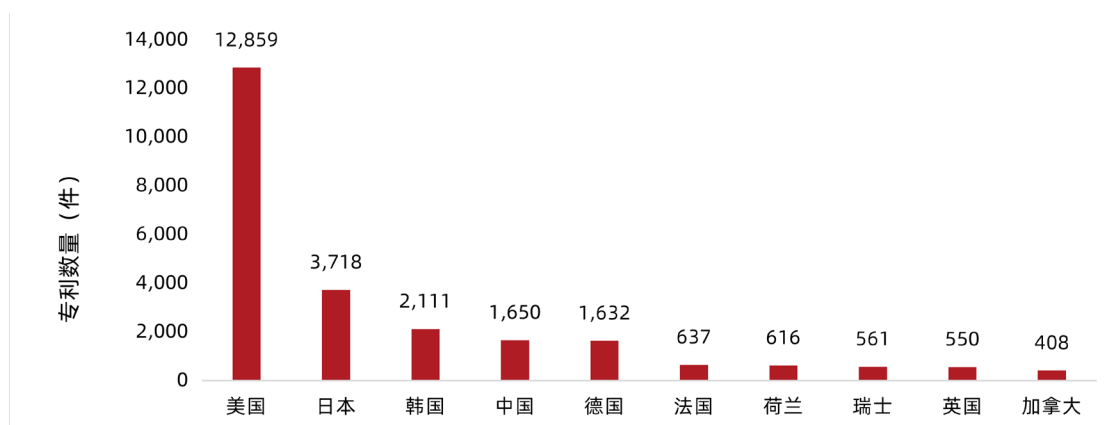
(二) 高价值专利分析

全球高价值专利前 10 强国家

本报告将市场价值为 100 万美元及以上的专利视为高价值专利。

全球高价值专利数量排名前 10 的国家如下图所示。其中美国高居第 1 名，共 12,859 件，是排名第 2 的日本（3,718 件）的 3.4 倍；第三是韩国，共 2,111 件；中国排名第 4，共 1,650 件，仅为美国的 12.8%；德国（1,632 件）以微小差距紧随中国而名列第五。第 6-10 名依次为法国、荷兰、瑞士、英国以及加拿大。由此可见，虽然中国数字技术专利数量遥遥领先其他国家，但是高价值专利数量大幅落后于美国，也落后于日韩两国。

图 15 全球数字科技高价值专利前 10 强国家



注：1. 本报告将市场价值为 100 万美元及以上的专利视为高价值专利；
2. 专利检索时间范围为 2012 年 1 月至 2021 年 12 月
数据来源：AMiner 科技情报平台

全球高价值专利前 10 强机构

全球高价值专利数量排名前 10 的机构如下图所示。其中美国企业数量共 5 家，中国 2 家，韩国、荷兰、日本各 1 家。韩国三星公司的高价值专利数量最多，共 1,061 件；第二是美国微软公司，共 630 件；第三是美国谷歌公司，共有 592 件。中国的华为公司和阿里巴巴分别以 373 件和 261 件专利的数量位居榜单的第五和第九位。

专利的市场价值随时间递进而具波动性，高价值专利能够反映当下技术创新的新风向。因此，三星、微软、谷歌等上述高价值专利前 10 强公司是当下全球数字技术创新的主要引领者。

表 5 全球数字技术高价值专利数量前 10 强机构

排名	公司名称	高价值专利数 (件)	简介
第 1 名	 [韩]三星电子株式会社	1,061	1938 年创立，是韩国最大的跨国企业集团，业务涉及电子、金融、机械、化学等众多领域
第 2 名	 [美]微软公司	630	1975 年创立，以研发、制造和提供电脑软件服务为主，为全球最大电脑软件供应商
第 3 名	 [美]谷歌有限责任公司	592	1998 年成立，被公认为全球最大的搜索引擎公司，提供大量基于互联网的产品与服务
第 4 名	 [美]高通公司	545	1985 年创立，为全球领先的无线科技创新者，主营无线电通信技术研发，芯片研发
第 5 名	 [中]华为技术有限公司	373	1987 年成立，是全球领先的信息与通信技术（ICT）解决方案供应商
第 6 名	 [美]国际商业机器公司	352	1911 年创立，计算机产业长期的领导者，全球最大的信息技术和业务解决方案公司
第 7 名	 [美]苹果公司	332	1976 年成立，目前为全球市值最大的高科技公司，主营业务有智能手机、计算机软件、在线服务等
第 8 名	 [荷]飞利浦公司	295	1891 年成立，世界上最大的电子品牌之一，在欧洲名列榜首。业务覆盖照明、消费电子、家用电器和医疗系统等领域。
第 9 名	 [中]阿里巴巴集团控股有限公司	261	1999 年创立，是一家以技术驱动，开拓数字经济时代商业基础设施的数字智能公司
第 10 名	 [日]索尼公司	260	1946 年成立，是全球知名的大型综合性跨国企业集团，以研制电子产品为主要事业



注：1. 本报告将市场价值为 100 万美元及以上的专利视为高价值专利；
 2. 专利检索时间范围为 2012 年 1 月至 2021 年 12 月。
 数据来源：AMiner 科技情报平台。

中国高价值专利前 10 强机构

中国高价值专利前 10 强企业如下图所示。其中，清华大学是唯一的高校机构，其余均为企业。排名第一的是华为技术有限公司，专利数量共 373 件；阿里巴巴和腾讯科技分别以 261、127 的专利数量位居第二、第三。第四名是中兴通讯，共有 80 件专利；第五名是台积电，共有 78 件专利。后四家公司分别是小米、京东方、OPPO 以及宏达电子。

表 6 中国数字科技高价值专利数量前 10 强机构

排名	公司名称	高价值专利数 (件)	简介
第 1 名	 HUAWEI 华为技术有限公司	373	1987 年创立，是全球领先的 ICT (信息与通信)基础设施和智能终端提供商
第 2 名	 Alibaba 阿里巴巴集团控股有限公司	261	1999 年创立，是一家以技术驱动，开拓数字经济时代商业基础设施的数字智能公司
第 3 名	 Tencent 腾讯 腾讯科技(深圳)有限公司	127	1998 年成立，中国最大互联网综合服务提供商之一，重要业务是致力于企业数字化升级
第 4 名	 ZTE 中兴 中兴通讯股份有限公司	80	1985 年成立，是全球领先的综合通信解决方案提供商，中国最大的通信设备上市公司
第 5 名	 tsmc 台积电(中国)有限公司	78	2003 年成立，台积电独资设立的子公司，生产线宽 0.35 微米及以下大规模集成电路
第 6 名	 mi xiaomi 小米科技有限责任公司	49	2010 年创立，是以智能手机、智能硬件和 IoT 平台为核心的消费电子及智能制造公司
第 7 名	 BOE 京东方科技集团股份有限公司	37	1993 年成立，是一家为信息交互和人类健康提供智慧端口产品和服务的物联网公司
第 8 名	 oppo OPPO 移动通信有限公司	31	2004 年成立，智能终端制造商和移动互联网服务提供商

第 9 名		宏达电子股份有限公司	29	1997 年成立，是一家位于中国台湾的手机与平板电脑制造商
第 10 名		清华大学	21	1911 年创办，211、985 和双一流重点大学，中国高层次人才培养和科学技术研究的基地

注：1. 本报告将市场价值为 100 万美元及以上的专利视为高价值专利；

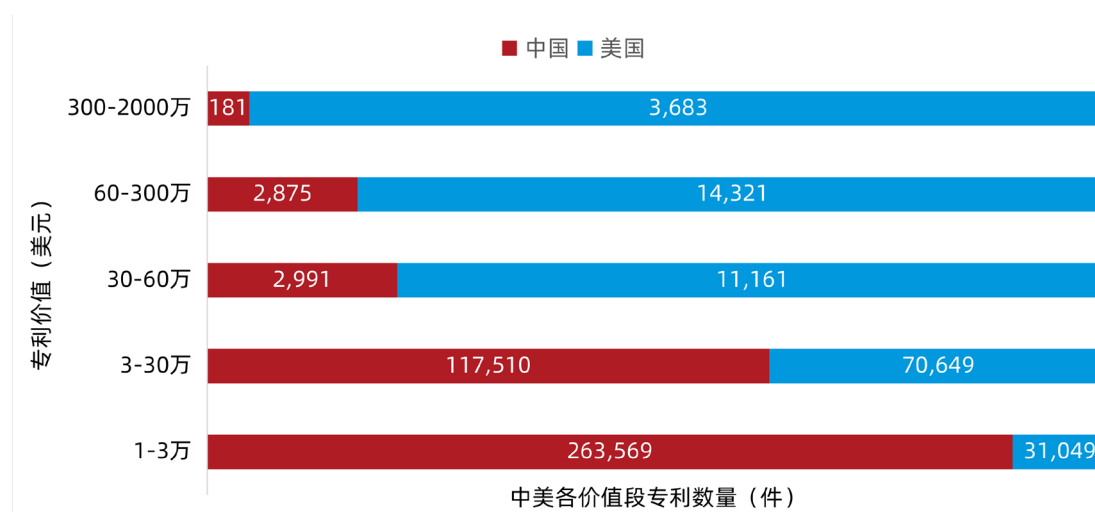
2. 专利检索时间范围为 2012 年 1 月至 2021 年 12 月。

数据来源：AMiner 科技情报平台。

中美专利价值比较分析

从中美两国专利价值分布看，中国拥有的 1-3 万美元的专利数量占绝对优势，共 263,569 件，是美国的 8 倍；其次是 3-30 万美元的专利，是美国的 1.3 倍。但是 30 万美元以上的专利，美国数量开始翻转。在 30-60 万美元的专利中，美国共 11,161 件，是中国的 3.7 倍；60-300 万美元的专利是中国的 5 倍；300-2000 万美元的专利是中国的 20 倍。可见，以 30 万美元为节点，该节点往下，中国优势越明显；该节点以上，美国的优势越明显。

图 16 2012-2021 年中美两国专利价值分布示意图



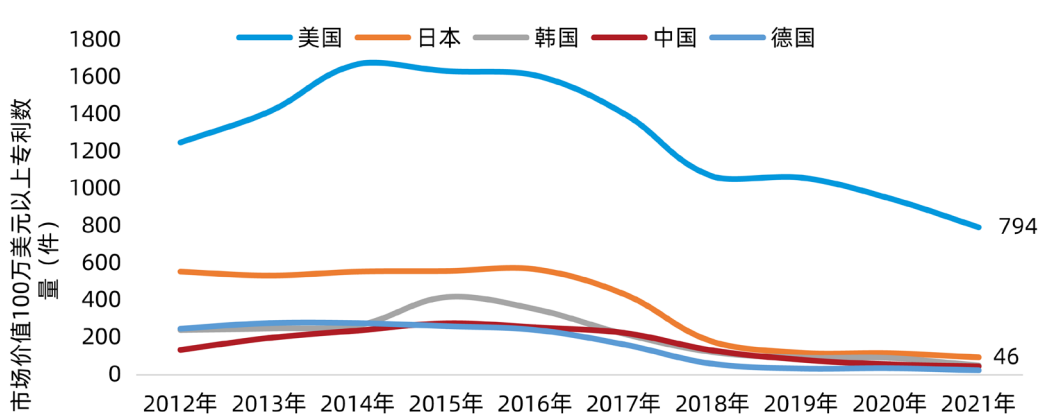
说明：专利检索时间范围为 2012 年 1 月至 2021 年 12 月；

数据来源：AMiner 科技情报平台。

中美等国高价值专利变化态势

从近 10 年高价值专利数量变化态势看，全球主要数字技术强国均处于下降态势（原因之一是专利价值的显现需要一定时长），美国与日本、韩国、中国、德国等数字技术强国仍保持较大优势。因此，全球数字技术领域的高价值专利储备实力格局在较长时间内仍然维持“一超多强”。

图 17 2012-2021 年数字技术高价值专利前 5 强国家专利数量走势



说明：专利检索时间范围为 2012 年 1 月至 2021 年 12 月；
数据来源：AMiner 科技情报平台。

Chapter 3

数字科技 人才分布格局

1. 全球数字科技人才总体情况
2. 中国数字科技人才分布
3. 数字科技人才流动与合作
4. 全球数字科技顶尖科研团队

1

全球数字科技人才总体情况

基于 AMiner 科技情报系统，共检索到全球数字科技领域人才 774,710 人，其中 H-index⁴ 在 20 以上的人才数量为 83,012 人。本章节从全球数字领域科技人才的国家 and 机构分布两个方面进行分析，并总结其特征。

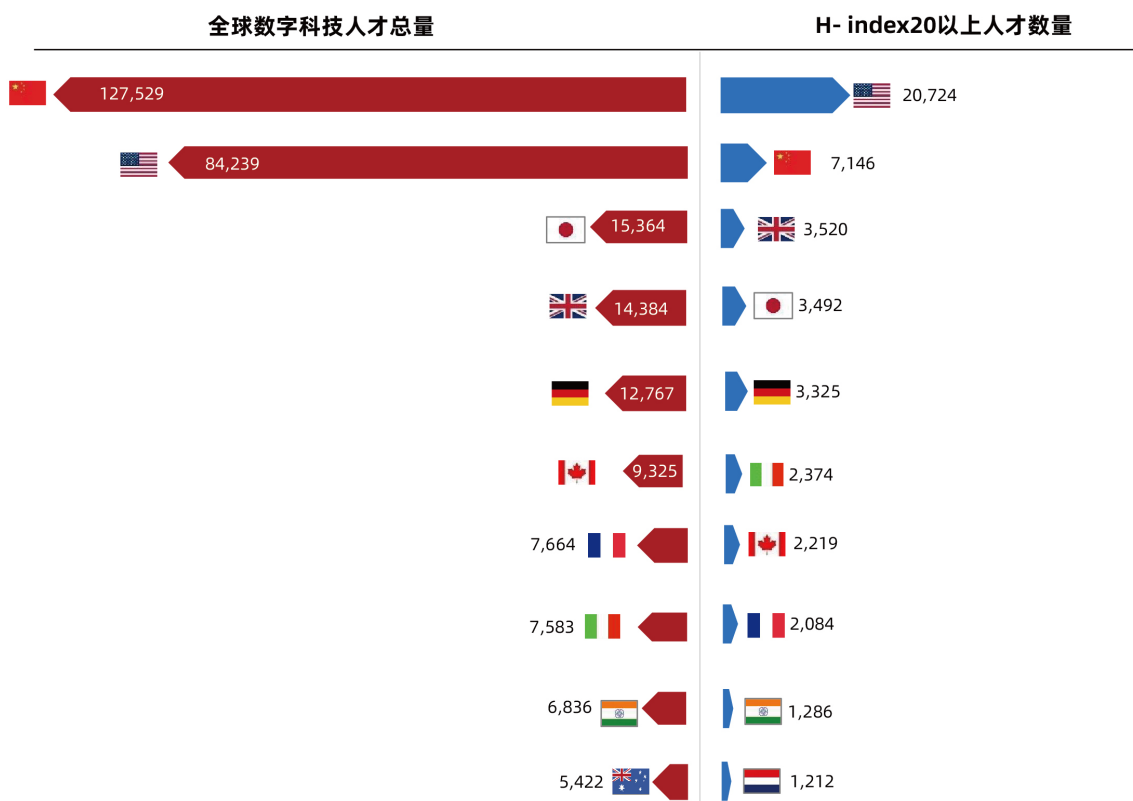
(一) 全球数字科技人才前 10 强国家

全球数字科技人才数量排名前 10 国家如下图所示。其中人数最多的国家是中国，共有 127,529 位，占全球人才总量的 17%；第二名是美国，共有 84,239 位，占全球 11%；第三名是日本，人才数量为 15,364 人，占全球 2%。可见，中国数字科技人才储备优势明显，是排名第二的美国的 1.5 倍，排名第三的日本的 8.3 倍。其他排名前 10 的国家依次为英国、德国、加拿大、法国、意大利、印度和澳大利亚。

从 H-index 20 以上的人才国别分布看，美国排名第一，共 20,724 人，占全球总量的 25%；中国名列第二，共有 7,146 人，占全球 9%；第三名是英国，共有 3,520，占全球 4%。由此可见，虽然中国数字科技领域人才总量是美国的 1.5 倍，但美国高层次以上人才数量是中国的 2.9 倍。

⁴H-index 是一个混合量化指标，目的是量化科研人员作为独立个体的研究成果。一名科研人员 H-index 是指其发表的 Np 论文中有 H 篇每篇至少被引 H 次、而其余 Np-H 篇论文每篇被引均小于或等于 H 次。H-index 能够比较准确地反映一个人的学术成就。其数值越高，则表明学者论文的影响力越大，美国学者赫希提出副高级职称的 H-index 一般在 20 以上。

图 18 全球数字科技人才数量前 10 强国家



数据来源：AMiner 科技情报平台

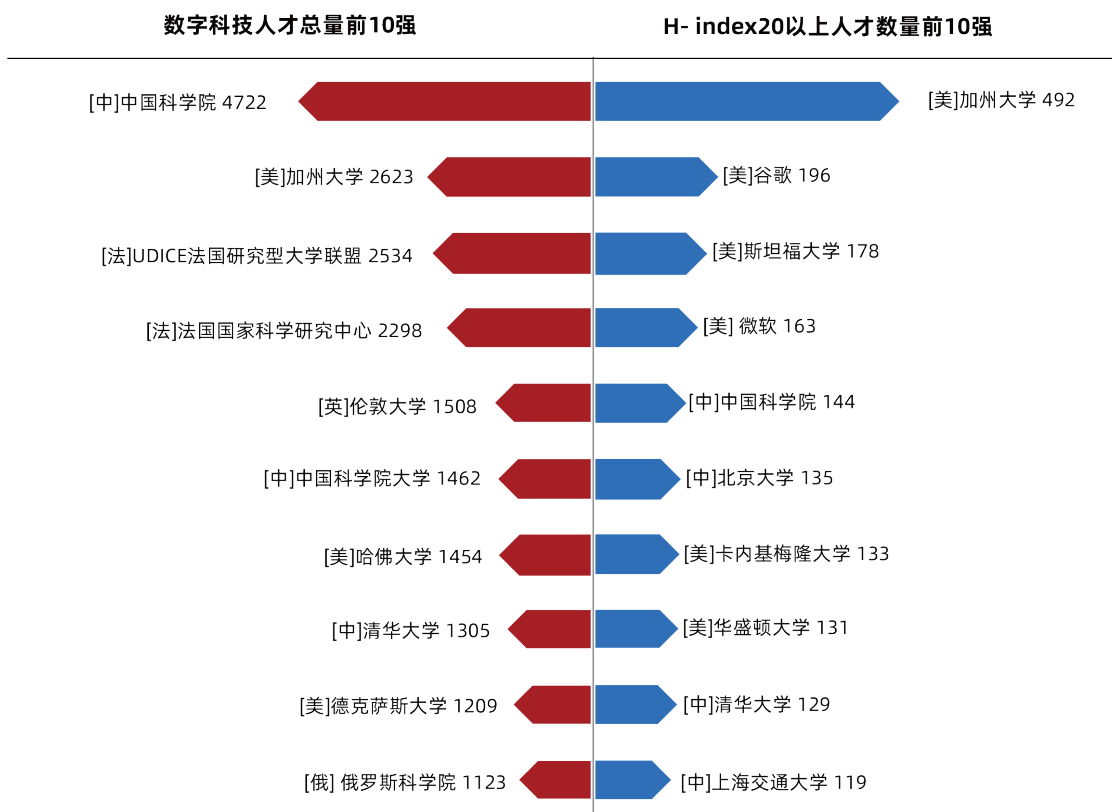
（二）全球数字科技人才前 10 强机构

在全球数字科技人才数量前 10 强机构中，中国科学院以 4,722 人位列第一，加州大学以 2,623 人位列第二，UDICE 法国研究型大学联盟协会以 2,534 人位列第三。其他前 10 强机构依次为法国国家科学研究中心、伦敦大学、中国科学院大学、哈佛大学、清华大学、德克萨斯大学和俄罗斯科学院。

在 H-index 20 以上的人才数量前 10 强机构中，美国和中国各占 6 家和 4 家。中国入榜机构依次是中国科学院（144 人，第 5 名）、北京大学（135 人，第 6 名）、清华大学（129 人，第 9 名）、上海交通大学（119 人，第 10 名）。值得注意的是，1) 虽然中国科学院人才总数高居全球第一，但 H-index 20 以上人才（144 人）数量排名下降为全球第五，H-index 20 以上的人

才数量仅占该机构人才总数的 3%；2) 虽然加州大学人才总数全球排名第二，但 H-index 20 以上人才（429 人）数量以较大优势高居第一，其 H-index 20 以上人才占比为 19%；3) 中国进入前 10 强的机构均为高校和科研院所，而美国有谷歌、微软两家高科技公司上榜。

图 19 全球数字科技人才数量前 10 强机构



数据来源：AMiner 科技情报平台

2

中国数字科技人才分布

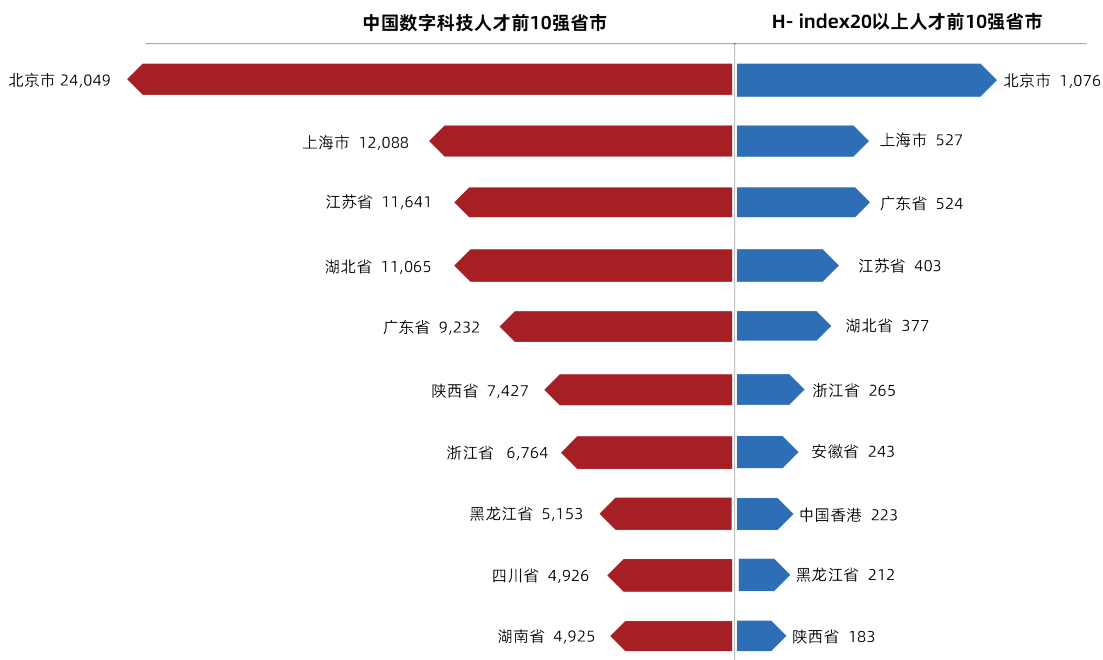
本篇节以中国数字科技领域科技人才为研究对象，分别从人才的省市和机构分布两个方面进行分析。

（一）中国数字科技人才前 10 强省市

根据 AMiner 科技情报平台显示，北京是全国聚集数字科技人才最多的省市，共有 24,049 人，占全国总量的 19%，是位列第二的上海（12,088 人）的 2 倍。江苏和湖北的人才资源也较多，分别为 11,641 人和 11,065 人。

在 H-index 20 以上的人才储备方面，北京依然排名第一，共有 1,076 人，占全国总量的 15%，仍然是排名第二的上海（527 人）的 2 倍；广东排名第三，共有 524 人。总体看，中国数字科技人才主要分布在经济发达的东部沿海省份，以及重点高校资源较丰富的湖北、安徽和陕西等内地省份。

图 20 中国数字科技人才前 10 强省市



数据来源：AMiner 科技情报平台

(二) 中国数字科技人才前 10 强机构

从中国数字科技人才机构分布看，中国科学院拥绝对领先优势，共有 4,722 人；中国科学院大学排在第二位，共计 1,462 人；清华大学则以 1,305 名的人才总数紧随其后排在第三位。

在 H-index 20 以上的人才储备方面，中科院、北京大学和清华大学依次以 144 人、135 人和 129 人名列前三甲；其他跻身国内前 10 强的机构依次为上海交通大学、浙江大学、南京大学、哈尔滨工业大学、复旦大学、香港科技大学和中国人民大学。可见除中国科学院和香港科技大学外，上榜的内地高校均为 211 和 985 院校。

表 7 中国数字科技人才数量前 10 强机构

中国数字科技人才前 10 强机构				中国 H-index≥20 数字科技人才前 10 强机构			
排名	机构	人才储备 (人)		排名	机构	人才储备 (人)	
1	 中国科学院 CHINESE ACADEMY OF SCIENCES	中国科学院	4,722	1	 中国科学院 CHINESE ACADEMY OF SCIENCES	中国科学院	144
2	 中国科学院大学 University of Chinese Academy of Sciences	中国科学院大学	1,462	2	 北京大学 PEKING UNIVERSITY	北京大学	135
3	 清华大学 Tsinghua University	清华大学	1,305	3	 清华大学 Tsinghua University	清华大学	129
4	 上海交通大学 SHANGHAI JIAO TONG UNIVERSITY	上海交通大学	1,116	4	 上海交通大学 SHANGHAI JIAO TONG UNIVERSITY	上海交通大学	119
5	 浙江大学 ZHEJIANG UNIVERSITY	浙江大学	1,103	5	 浙江大学 ZHEJIANG UNIVERSITY	浙江大学	103
6	 北京大学 PEKING UNIVERSITY	北京大学	1,028	6	 南京大学 NANJING UNIVERSITY	南京大学	95
7	 中山大学 SUN YAT-SEN UNIVERSITY	中山大学	845	7	 哈尔滨工业大学 HARBIN INSTITUTE OF TECHNOLOGY	哈尔滨工业大学	94
8	 中国科学技术大学 University of Science and Technology of China	中国科学技术大学	842	8	 复旦大学 FUDAN UNIVERSITY	复旦大学	66
9	 华中科技大学 HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY	华中科技大学	828	9	 香港科技大学 HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY	香港科技大学	49
10	 复旦大学 FUDAN UNIVERSITY	复旦大学	763	10	 中国人民大学 RENMIN UNIVERSITY OF CHINA	中国人民大学	45

3

数字科技人才 流动与合作

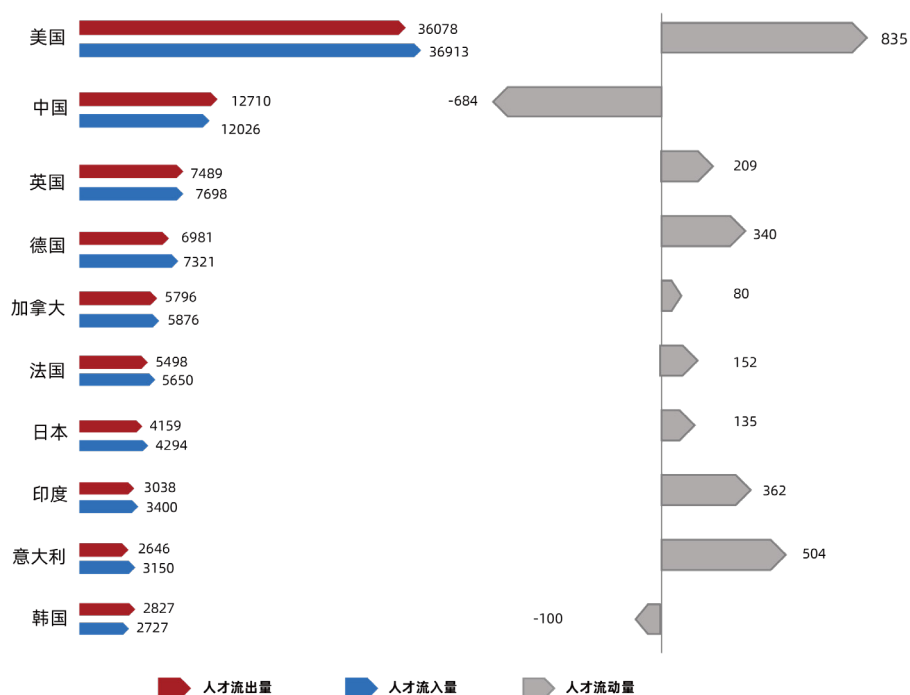
在全球化的背景下，人才合理流动在一定条件下可以优化资源配置，提升技术创新体系整体效能。根据 AMiner 科技情报平台统计，2012-2021 年间，全球数字科技领域人才流动量为 24.4 万人次。

（一）全球数字科技人才流动情况

从各国数字科技人才流动情况看，美国是人才流动最活跃的国家，流出和流入总量达到 72,991 人次；中国人才流动总量为 24,736 人次，位居第二；英国、德国、加拿大、法国的数字科技人才流动也较为活跃，流动总量达 1 万人次以上。

从各国数字科技人才净流入情况看，美国是全球最大的净流入国家，共净流入 835 人；其次是意大利，共 504 人；第三是印度，362 人；另外，德国、英国、法国、日本、加拿大等发达国家均为正向净流入国家。从净流出情况看，中国为全球最大的净流出国家，达 684 人次；其次是韩国，净流出 100 人。

图 21 2012-2021 年全球主要国家数字科技领域人才流动对比



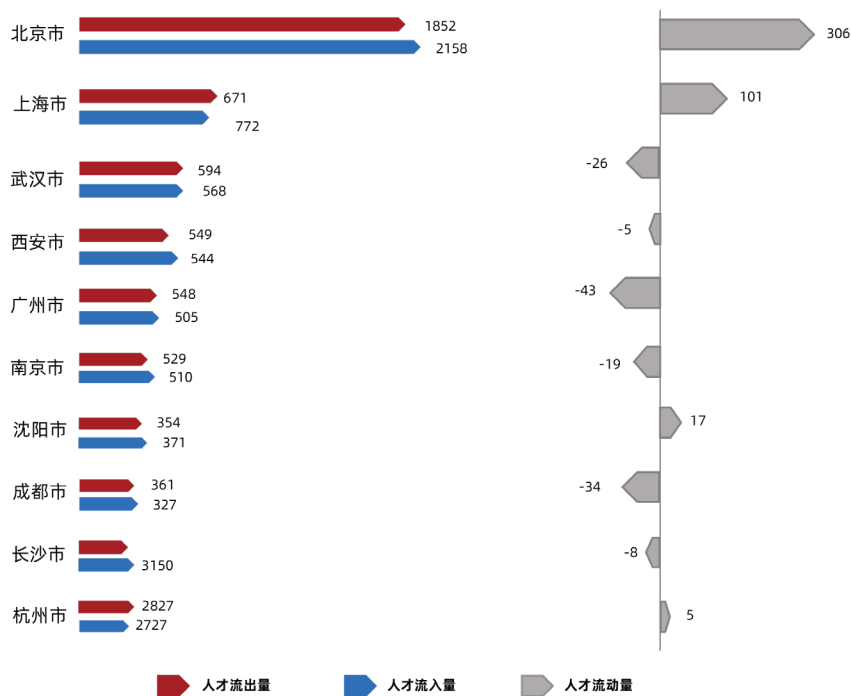
数据来源: AMiner 科技情报平台

(二) 中国数字科技人才流动情况

从国内数字科技人才流动情况看,北京数字科技人才流动最活跃,流出和流入总量达到 4,010 人次;上海第二,人才流动总量为 1,443 人次;武汉第三,流动总量为 1,162 人次。

从国内各主要城市数字科技人才净流入看,北京最多,共 306 人;其次是上海,共 101 人;另外,沈阳(17 人)和杭州(5 人)也是正向净流入城市。从人才净流出情况看,广州净流出最多,共 43 人;其次是成都,共 34 人;第三是武汉,共 26 人;另外,南京(19 人)、长沙(8 人)、西安(5 人)也出现一定程度的流失现象。可见,我国数字科技人才呈现由内陆城市向京沪一线城市流动的态势。

图 22 2012-2021 年中国主要城市数字科技领域人才流动对比



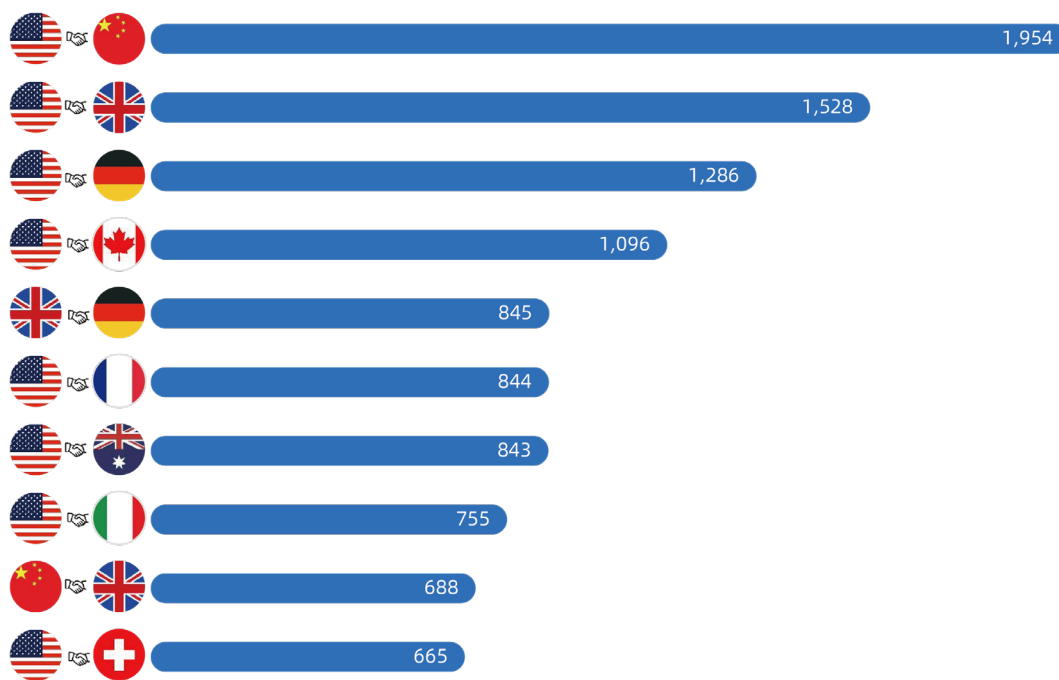
数据来源：AMiner 科技情报平台

(三) 全球数字科技人才合作情况

论文合作是科研合作最主要的合作形式之一。从各国被引量 Top1% 合作论文数量来看，中美两国的合作的论文最多，共为 1,954 篇；第二名为英美两国，合作数量为 1,528 篇；第三名为美德两国，合作论文数量为 1,286 篇。其他排名靠前的合作国家还有美国和加拿大（1,096 篇）、英国和德国（845 篇）、美国和法国（844 篇）、美国和澳大利亚（843 篇）、美国和意大利（755 篇）、中国和英国（688 篇）、美国和瑞士（665 篇）。整体而言，美国与各国都保持着紧密的合作，在全球合作最热络的前十强双边国家中，有 8 组双边关系由美国与其他 8 个国家组成。

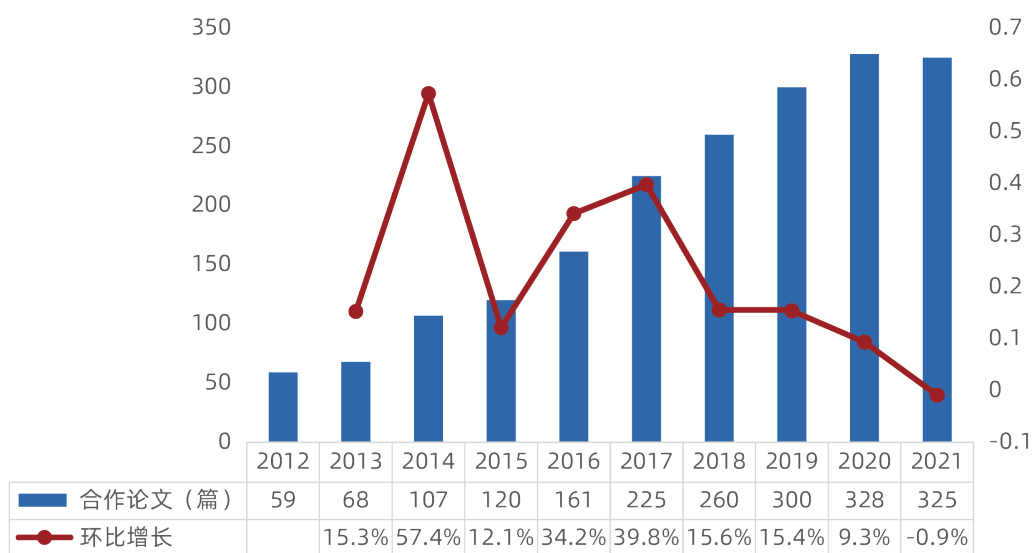
从中美学者合作的 Top 1% 论文数量的逐年分布来看，从 2012 年开始，中美两国学者合作论文数量一直处于增长态势，到 2020 年达到顶峰（328 篇），2021 年出现小幅下降现象。实际上，两国学者合作论文数量增长率从 2018 年开始就一直处于下降态势，在 2021 年更是出现负增长（-0.9%）。可见，中美两国学者在数字科技最卓越研究层次的合作热络度有所降温，究其原因，这与美国国内对中美在前沿领域科研合作和交流奉行限制性政策，以及全球新冠疫情大流行有关。

图 23 2012-2021 年全球数字科技人才国际合作关系热度 Top 10



说明：图中数据为两国学者合作产出被引量 Top1% 的“顶尖论文”数量
数据来源：AMiner 科技情报平台

图 24 2012-2021 年数字科技领域中美学者合作关系走势



说明：图中合作论文为中美两国学者合作产出被引量 Top1% 的“顶尖论文”数量
数据来源：AMiner 科技情报平台

4

全球数字科技 顶尖科研团队

坚持全球视野和开放创新，掌握国内顶尖团队及其学者最新概况是攻克行业关键技术或前瞻性技术的重要路径。本篇通过大数据分析和挖掘技术，对数字科技领域机构进行挖掘分析，筛选其中“h5 指数”⁵ 排名靠前的十大机构的创新研发团队进行介绍。

(一) 谷歌 X 实验室

谷歌公司 (Google Inc.) 成立于 1998 年，由 Larry Page 和 Sergey Mikhaylovich Brin 共同创建，被公认为全球最大的搜索引擎公司。谷歌旗下有数个研究实验室，其中包括研究方向有些神秘的 X 实验室、研究深度学习的 Google Brain、研究人工智能的 Google AI、研究机器人的 Robotics at Google 以及研究增强现实和智能织物的先进技术和项目集团 (Advanced Technologies and Projects)，自动驾驶汽车 Waymo、互联网气球 Loon、送货无人机 Wing 等研究成果在不同方向上推动了数字科技领域的进步。

重要成员：Astro Teller (人工智能，信号增强算法)、Sebastian Thrun (机器人，人工智能)、Johnny Chung Lee (人机交互)。

机构网址：<https://x.company>

⁵h5 指数 (h5-index) 指是收录在 AMiner 学术系统中的期刊和会议在最近 5 年的论文数量及各论文被引用的次数，可以体现一个机构的综合实力，是评价一个机构影响力的重要参考。例如，如果某机构在过去 5 年所发表的论文中，至少有 h 篇论文分别被引用了至少 h 次，那么这个机构的 h5 指数就是 h。机构排名详情见：<https://www.aminer.cn/ranks/org>。

（二）斯坦福大学自然语言处理组

斯坦福大学自然语言处理组是全球最重要和最突出的实验室之一，由斯坦福语言学系和计算机系共同组成，同时也是斯坦福大学人工智能实验室的重要组成部分。改组主要研究计算机处理和理解人类语言的算法，工作范围从计算语言学的基本研究到语言处理的关键应用技术均有涉猎，涵盖句子理解、自动问答、机器翻译、语法解析和标签、情绪分析和模型的文本和视觉场景等。该小组的一个显著特征是将复杂和深入的语言建模和数据分析与 NLP 的创新概率、机器学习和深度学习方法有效地结合在一起。

重要成员: Chris Manning (自然语言处理、人工智能、机器学习、信息检索); Dan Jurafsky(自然语言处理)、Percy Liang (计算机科学、人工智能、自然语言处理、机器学习)、Christopher Potts (人工智能、计算机科学、自然语言处理)、Tatsunori Hashimoto (自然语言处理、机器学习)、Monica S. Lam (自然语言处理、深度学习)、Diyi Yang (自然语言处理、人工智能)。

机构网址: <https://nlp.stanford.edu>

（三）卡内基梅隆大学机器学习系

卡内基梅隆大学机器学习系成立于 2006 年春季，是世界上第一个机器学习系。它是从 1997 年创建的自动学习和发现中心 (CALD) 组织演变而来的。2017 年 6 月，CMU 启动人工智能计划 (CMU AI)，整合了全校人工智能的研究资源，开展跨学科的人工智能协作，组建世界上规模最大、经验最丰富的 AI 研究团队之一，同时促进本校人工智能方向的人才培养，除覆盖计算机学科以外，还涉及生物、环境工程、哲学、艺术、公共政策等多个领域。

重要成员: Ruslan Salakhutdinov (机器学习)、Maria-Florina Balcan (人工智能、机器学习)、Yuanzhi Li (深度学习)、Henry Chai (机器学习、计算机科学)、Christos Faloutsos (机器学习)、Ziv Bar-Joseph (机器学习)。

机构网址: <https://www.ml.cmu.edu>

（四）加利福尼亚大学伯克利分校人工智能研究实验室 (BAIR)

加州大学伯克利分校是全球最负盛名的公立学校，人工智能研究 (Berkeley Artificial Intelligence Research, BAIR) 实验主要研究领域涵盖计算机视觉、机器学习、自然语言处理、规划和机器人等。机器人和智能机器实验室致力于用机器人复制动物的行为，自动化科学和工程实验室从事更广泛的机器人功能的研究，如机器人辅助外科手术和自动化制造。计算机可视化小组研究如何让机器人能“看得见”。

重要成员：Trevor Darrell（人工智能、机器人、计算机视觉）、Pieter Abbeel（人工智能、机器人、机器学习）、Alexei Efros（人工智能、计算机视觉）、Peter L. Bartlett（计算机科学、深度学习、神经网络学习）、John Canny（人工智能、机器人）、Ken Goldberg（机器人、计算机科学）。

机构网址：<https://bair.berkeley.edu>

（五）微软亚洲研究院

微软亚洲研究院是微软公司在亚太地区设立的基础及应用研究机构，也是微软在美国本土以外规模最大的研究机构。微软亚洲研究院从事自然用户界面、智能多媒体、人工智能、云和边缘计算、大数据与知识挖掘、计算机科学基础等领域的研究，致力于推动整个计算机科学领域的前沿发展，将最新研究成果快速转化到微软的关键产品中，着眼下一代革命性技术的研究和孵化。

重要成员：周礼栋（可靠、可信及可扩展的分布式计算系统）、洪小文（自然交互、语音识别）、郭百宁（计算机图形学、计算机可视化、自然用户界面以及统计学习）、刘铁岩（机器学习、信息检索）、潘天佑（智能芯片、信息系统安全）、田江森（智能芯片）、张冬梅（大数据分析、知识计算、数据可视化以及软件工程）、周明（自然语言处理）。

机构网址：<https://www.msra.cn>

（六）MIT 计算科学与人工智能实验室（CSAIL）

CSAIL 是麻省理工学院最大的实验室，专注于开发基础新技术，开展基础研究以推动计算领域的发展研究领域涉及电气工程，计算机科学，数学，航空航天，脑和认知科学，机械工程，媒体艺术与科学，以及地球，大气和行星科学，卫生科学与技术。实验室开创了计算领域的新研究，改善了人们工作、娱乐和学习的方式。

重要成员：Daniela Rus（人工智能、计算机科学、机器人）、Arvind Mithal（人工智能、计算机科学）、Joel Voldman（计算机科学）、Antonio Torralba（人工智能、计算机科学、计算机视觉、机器学习、计算机视觉）、Hal Abelson（计算机科学、人工智能）、Ted Adelson（人工智能、计算机科学）、Anant Agarwal（人工智能、计算机科学）、Pulkit Agrawal（人工智能、计算机科学）、Mohammad Alizadeh（人工智能、计算机科学）、Saman Amarasinghe（人工智能、计算机科学、机器学习）。

机构网址：<https://www.csail.mit.edu>

（七）IBM Thomas J. Watson 研究中心

IBM Thomas J. Watson 研究中心位于美国纽约，专门从事基础科学研究，并探索与产品有关的技术，科学家在这里工作，一方面推进基础科学，一方面提出对实际应用有益的科学新思想。研究范围包括计算机科学、输入 / 输出技术、生产性研究数学、物理学、激光物理、天文学和基本粒子。

重要成员：Dario Gil（人工智能）、Charles Bennett（量子计算、量子密码学、算法信息论）、Rogerio Feris（人工智能、计算机视觉、机器学习）、Mo Yu(于墨)（自然语言处理、机器学习）、Abhishek Kumar（机器学习、人工智能）。

机构网址：<https://research.ibm.com/labs/watson>

(八) 哥伦比亚大学数据科学研究所

美国哥伦比亚大学数据科学研究所致力于培养下一代的数据科学家，并开发创新性技术以造福于社会。该研究所设立了7个研究中心，这些研究中心是数据科学转化研究与教育的引擎，也是具备高度商业化潜能的技术的来源。1) 网络安全中心专注于开发能在数据整个生命周期保护数据安全与隐私的能力。2) 数据、媒体与社会中心（前新媒体中心）关注数据中的人，重点研究如何使用数据来理解人类行为、数据与数据处理如何塑造人们的工作和生活方式、数据对网络世界和数字化世界中的人的意义。

重要成员：Jeannette Wing（人工智能）、Chris Wiggins（应用数学、系统生物学）、Tian Zheng（机器学习）、Matthias Preindl（电力电子系统）、Alan West（电化学）。

机构网址：<https://datascience.columbia.edu>

(九) 康奈尔大学计算科学系

康奈尔大学计算机科学系成立于1965年，是最早成立的此类系之一。计算机科学系在理论计算机科学、编程语言、分布式系统、人工智能、信息检索、科学计算和计算机图形等方面处于全球领先地位。

重要成员：Kilian Weinberger（机器学习、数据挖掘）、Kenneth Borup（机器学习、深度学习、计算机视觉）、Todd Schmid（计算机科学）、Weijia Song（云计算）、Christoph Kreitz（计算机科学）。

机构网址：<https://www.cs.cornell.edu>

(十) 美国贝尔实验室

贝尔实验室原名贝尔电话实验室，始建于1925年，总部在美国纽约（后迁至新泽西州的墨里黑尔）。它是一个在全球享有极高声誉的研究开发机构，主要宗旨是进行通讯科学的研究。贝尔实验室自成立以来，共获重大科研成果50多项，如有声电影、晶体管、信息论、激光理论、3K宇宙背景辐射、可视电话、磁泡器件、光通信、数字计算机等。

重要成员：Peter Vetter（可编程网络系统和安全、移动无线电系统、光学系统）、Thierry Klein（先进传感技术、人工智能、机器人、图像和数据分析、自动化）、Gee Rittenhouse（移动通信、无线网络）、Alfred Aho（量子计算、程式语言、编译器、演算法）。

机构网址：<http://www.bell-labs.com/>

Chapter 4

数字科技 发展趋势分析

1. 研究方法
2. 数字科技 10 大技术趋势
3. 2022 数字科技热点技术

1

研究方法

前沿技术（或研究前沿）通常被视为“最具发展潜力的新兴研究领域或研究主题”，美国知名科研情报机构科睿唯安（Clarivate）将其定义为“研究前沿由一组共同引用的核心论文以及引用一篇或多篇这些核心论文的当前来源论文簇组成”⁶。近年来，前沿技术与挖掘受到学界广泛关注，论文数量（贡献度）和论文被引频次（影响度）成为衡量技术前沿度的重要评估指标⁷。

一般来说，技术趋势是指某种技术在一定时间跨度内的水平变化方向。为了有效表现先技术在某段时间的发展趋势，本报告引入论文数量、论文被引频次和论文发表时间作为重要评估指标，将数字科技领域被引量 Top 1%ESI⁸ 核心论文作为贡献度测算指数，文的被引用次数作为影响力测算指数，论文首次被引至最近一次被引时间间隔作为识别“学术睡美人”测算指数，论文被引用变化趋势作为判断技术未来发展潜力的测算指数。旨在通过论文、专利数据的被引用量、被引时间跨度、被引量变化趋势来发现技术影响力，且近年来学术影响力处于增长状态的技术。计算方式如下式所示：

$$\text{前沿度} = \frac{C_{total}}{T \cdot S_{trend}}$$

其中， C_{total} 代表科技成果的被引用量，反映科技成果产生的学术影响力； T 表示科技成果第一次被引用至最近一次被引用的时间间隔，对于被引用量相同的文献，该值越小，表示文献所包含的技术点在短期内的影响力越大，越具有前沿特性； S_{trend} 表示研究目标文献年度被引用量的变化趋势，与被引量时序数据的斜率有关，其他条件一致的情况下，被引用量呈现上升趋势的文献能够通过计算得到更大的前沿度数值。

⁶ 中国科学院科技战略咨询研究院等.《2022 研究前沿热度指数报告》[P]. 北京. 2022.12.27.

⁷ Fajardo-Ortiz D, Lopez-Cervantes M, Duran L, et al. The emergence and evolution of the research fronts in HIV/AIDS research[J]. PLoS ONE, 2017, 12(5):e0178293.

⁸ ESI 被称为基本科学指标数据库 (Essential Science Indicators Database)，是国际公认的评价大学、学术机构和国家 / 地区国际学术水平和影响力的重要评价指标工具之一。

2

数字科技 10 大技术趋势

通过对 2012 年—2021 年数字科技领域 Top 1% 核心论文的持续跟踪，聚类相关论文，挖掘技术前沿度测算指数，筛选出数字科技领域排名靠前的十大技术，以期为科技评价和管理提供决策支撑。

(一) 数字科技领域 10 大热点前沿技术

全球数字科技领域排名靠前的 10 大热点前沿技术分别是生物大数据、生成式对抗网络算法 (Generative adversarial networks, GANs)、沉浸式扩展现实 (Extended Reality, XR)、量子计算机、AI 破解蛋白质和基因结构、移动边缘网络、可解释的 AI、联邦学习、混合计算和能源区块链技术。其中，生成对抗网络算法、可解释的 AI、混合计算和能源区块链技术的相关科研成果的平均出版年均均在 2019 年以后，均为最新的新兴前沿技术。

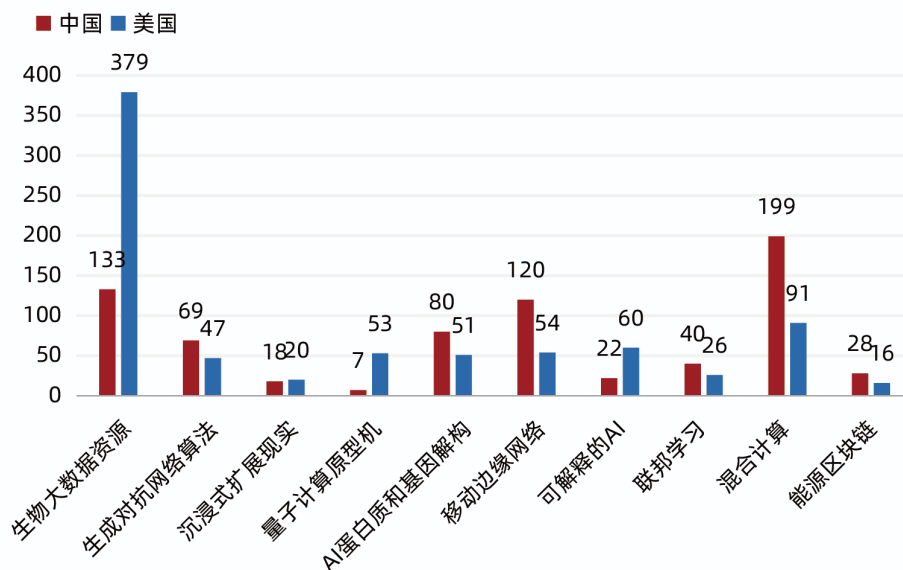
表 8 数字科技领域 10 大热点前沿技术

排名	热点前沿	TOP1% 论文数量 (篇)	高价值专利数量 (件)	前沿指数	平均出版年
1	生物大数据	694	13	0.99	2017
2	生成式对抗网络算法	141	150	0.98	2019
3	沉浸式扩展现实娱乐平台	54	1153	0.95	2018
4	量子计算机	81	1143	0.94	2017
5	AI 解码蛋白质结构	160	185	0.93	2018
6	移动边缘计算网络	175	1686	0.92	2018
7	可解释的 AI	126	41	0.92	2019
8	联邦学习	66	13	0.91	2020
9	混合计算	292	293	0.90	2018
10	能源区块链	67	33	0.86	2019

说明：数据检索时间范围为 2012 年 1 月至 2021 年 12 月；
数据来源：AMiner 科技情报平台。

中国和美国是发表数字科技领域前沿技术论文最多的两个国家，高被引论文数量占全球总量的 81.52%。盘点中美创新成果，从前沿技术论文数量来看，中国更具优势，尤其是在生成对抗网络算法、AI 解码蛋白质结构、移动边缘网络、联邦学习、混合计算和能源区块链六个方面。但是从论文的平均被引用量来看，中国仅在联邦学习领域超过美国。总体而言，相对美国，近年来中国在数字科技领域的创新产出方面获得了显著和广泛的进步，但是在创新产出质量方面还有待进一步提高。

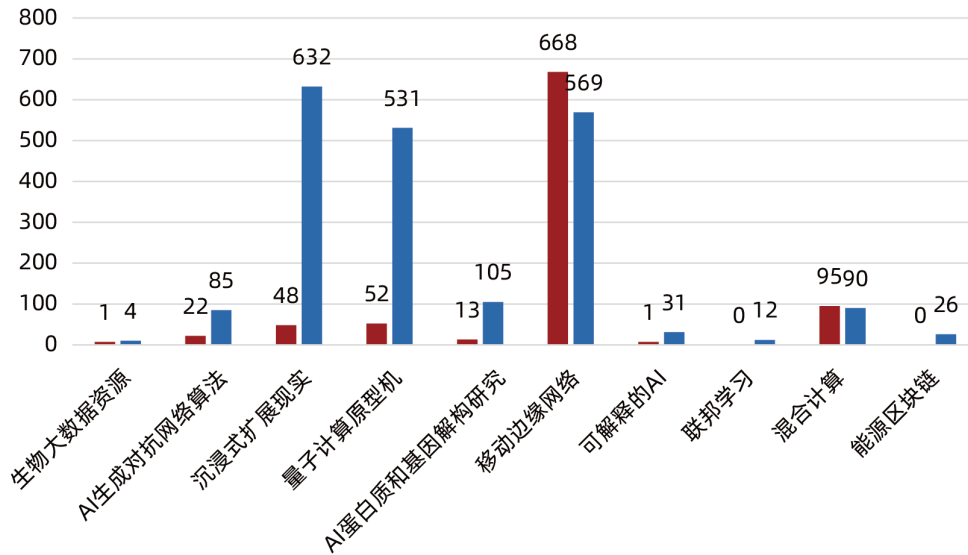
图 25 中美数字科技领域前沿技术论文数量对比



说明：数据检索时间范围为 2012 年 1 月至 2021 年 12 月；
数据来源：AMiner 科技情报平台。

中国和美国是申请数字科技领域前沿技术专利最多的两个国家。从核心专利数量来看，中国更具优势，尤其是在混合计算、AI 生成对抗网络算法、移动边缘网络、量子计算原型机、联邦学习五个方面。但是从高价值（专利价值五十万美元以上）专利数量来看，中国仅在移动边缘网络、混合计算两个领域超过美国。

图 26 中美数字科技领域高价值专利数量对比



说明：数据检索时间范围为 2012 年 1 月至 2021 年 12 月；
数据来源：AMiner 科技情报平台。

(二) 前沿技术解读

趋势一：生物大数据

随着对生命系统的不断深入探究和各种其他高通量组学技术的产生和发展，生物信息学的研究范畴不断扩大，各种组学数据（转录组、蛋白质组、非编码 RNA 组、表观遗传组、代谢组、宏基因组等）以及生物系统层面的解读不断扩展，生命科学从定性描述开始实现动态、精准和定量解读。

大数据是互联网技术快速发展的重要产物，包括数据量大、速度快、种类多、有价值等特点。随着大数据时代的来临，大数据技术在各行各业都发挥了极为重要的作用，具有十分广阔的发展前景。生物大数据按来源属性主要分为科学数据和真实世界数据两类。科学数据产生于科研活动中，主要是生物医学研究中的基因组数据、蛋白质组数据、代谢组数据等各类组学研究数据，质量较高，并包含对数据的机理性解读。

生物大数据领域 TOP1% 论文共计 694 篇，其中综述性论文 123 篇，主要综述了生物数据库建设及数据挖掘技术、生物数据库应用和生物大数据伦理问题。其余 563 篇研究性论文主要是



在探索和发现生物大数据方面做出研究，如对 OMICs 数据集、致病菌毒力因子 (VFDB) 数据库、TBtools 分析软件的研究。从中美论文数量看，中国论文 133 篇，美国论文 379 篇。

生物大数据领域专利共计 195 件，中国专利 65 件，美国专利 20 件。专利研究方向以生物数据的测定、数据库的设计和构建方法相关。全球高价值专利 13 件，中美高价值专利共计 5 件，中国高价值专利 1 件，美国高价值专利 4 件。

总体来说，生物大数据领域论文的数量远大于专利数量，该领域最重要的科研成果输出途径是论文。假设将论文作为衡量基础性研究生产力的指标，专利作为衡量应用性研发生产力，生物大数据领域的基础研究生产力远高于应用性研发生产力。



趋势二：生成式对抗式网络算法

作为人工智能学界的热门方向,生成式对抗式网络算法(Generative adversarial networks)已法被广泛应用于图像和视觉、语音和语言、信息安全等领域,并被图灵奖得主 LeCun 评为为“过去十年间机器学习领域最让人激动的点子”。未来,随着 GANs 为代表的深度学习不断迭代,AIGC 百花齐放,产出效果或将逐渐逼真接近至人类作品。

生成式对抗网络 GANs (Generative adversarial networks) 模型是 Goodfellow 等最早在 2014 年提出的一种生成式模型⁹。GANs 模型具有“无限”的生成能力,在图像和视觉计算、语音和语言处理、信息安全等领域具有重大的应用价值。GANs 模型不仅可以生成与真实数据分布一致的数据样本,例如图像、视频等。同时,GANs 可以解决学习过程中标注数据不足的问题,例如无监督学习(Unsupervised Learning)、半监督学习等(Semi-Supervised Learning)。此外,GANs 还可以用于语音和语言处理任务,例如对话生成(dialogue generation)、文本生成图像(text to image)等。

生成式对抗网络 GANs 领域 TOP1% 论文共计 141 篇,其中综述性论文 16 篇,主要综述了生成式对抗网络模型的原理;生成式对抗网络模型在视觉领域的应用及展望等。其余 124 篇研究性论文主要在图像、视频、生成对话和文本生成图像方面做出了研究。从中美论文数量看,中国论文 69 篇,美国论文 47 篇。

⁹Goodfellow, Ian, et al. "Generative adversarial networks." Communications of the ACM 63.11 (2020): 139-144.

AI 生成对抗网络算法领域专利共 8209 件，中国专利 5341 件，美国专利 1283 件。专利研究方向大多与图像分析、模型计算方法相关。全球高价值专利共计 150 件，中美高价值专利 107 件，中国高价值专利 22 件，美国高价值专利 85 件。

总体来说，生成式对抗式网络算法领域论文和专利的数量相近，假设将论文作为衡量基础性研究生产力的指标，专利作为衡量应用性研发生产力，生成式对抗式网络算法领域的应用性研发生产力与基础研究生产力发展处于平衡状态，随着相关创新研究工作不断普及，未来在该领域可能会取得明显推动科学发展的关键突破。

趋势三：沉浸式扩展现实娱乐平台

扩展现实是元宇宙连接虚拟与现实的关键设备，随着扩展现实产业链和技术不断发展、内容应用逐渐繁荣，整个元宇宙娱乐生态正在持续丰满，沉浸式扩展现实娱乐平台有望迎来爆发。

扩展现实（Extended Reality, XR）技术是虚拟现实（Virtual Reality, VR）、增强现实（Augmented Reality, AR）、混合现实（Mixed reality, MR）的合称，该技术通过计算机将真实与虚拟相结合，打造一个可人机交互的虚拟环境。沉浸式扩展现实利用图形系统和各种控制接口设备，在计算机上生成可交互和具有沉浸感觉的现实世界的模拟仿真，也就是利用电脑模拟产生一个三维空间的虚拟世界，给使用者在虚拟环境中提供关于视觉、触觉、听觉等感官模拟，让使用者身临其境。



沉浸式扩展现实娱乐平台领域 TOP1% 论文共计 54 篇，其中综述性论文 14 篇，主要综述了可穿戴设备系统设计和可穿戴设备材料研发。其余 25 篇论文主要在系统优化、娱乐体验和心理感受等方面做出了研究。从中美论文数量看，中国论文 18 篇，美国论文 20 篇。

沉浸式扩展现实游戏平台领域专利共 6264 件，中国专利 1883 件，美国专利 2156 件。专利研究方向以数据处理、图像识别为主。全球高价值专利共计 1153 件，中美高价值专利 680 件，中国高价值专利 48 件，美国高价值专利 632 件。

总体来说，沉浸式扩展现实游戏平台领域专利的数量远大于论文数量，假设将论文作为衡量基础性研究生产力的指标，专利作为衡量应用性研发生产力，沉浸式扩展现实游戏平台领域的应用性研发生产力要远高于基础研究生产力，该项技术有望在未来几年成功落地。

趋势四：量子原型样机和专用处理器的研制

当前，量子计算机的研制已从以院校、研究所为主的基础性研究阶段，过渡到以 Google、IBM 等创新企业为主体的“量子霸权”研究阶段。随着量子计算机研制成果的不断涌现，通用量子计算机研制问题将在未来五年内被攻破，可实际使用的量子处理器将会落地。

量子计算机 (Quantum computer) 是一类遵循量子力学规律进行高速数学和逻辑运算、存储及处理量子信息的物理装置。量子计算机的特点主要有运行速度较快、处置信息能力较强、应用范围较广等。与一般计算机相比，信息处理量愈多，量子计算机实施运算就越有利，更能确保运算具备精准性。量子计算机的优势使得其在加密和破译等领域有着巨大的应用。



量子计算机领域 TOP1% 论文共计 81 篇，其中综述性论文 11 篇，主要综述了量子计算原型机的发展。其余 70 篇论文主要在量子计算机架构、量子计算机的硅 CMOS 架构等方面做出了研究。从中美论文数量看，中国论文 7 篇，美国论文 53 篇。

量子计算原型机领域专利共 8148 件，中国专利 2958 件，美国专利 2824 件。专利研究方向与量子计算、网络安全相关。全球高价值专利共计 1143 件，中美高价值专利 583 件，中国高价值专利 52 件，美国高价值专利 531 件。

总体来说，量子计算机领域专利的数量远大于论文数量，假设将论文作为衡量基础性研究生产力的指标，专利作为衡量应用性研发生产力，量子计算机领域的应用性研发生产力要远高于基础研究生产力，该项技术有望在未来几年成功落地。

趋势五：AI 解码蛋白质结构

2021 年，Deepmind 公司的蛋白质解码预测系统——AlphaFold 横空出世。与此同时，公司还公布了约 35 万种蛋白质的结构，该项成果因此入选《科学》2021 年度十大科学突破。随着人工智能技术的不断发展，对具有内在无序特性的蛋白质以及通过翻译后修饰或环境条件改变结构的蛋白质建模问题在将来有望解决。

自上世纪五六十年代起，蛋白质序列、结构与功能间的关系就一直是生命科学的核心问题。作为这一信息链条的中心点，蛋白质结构既可以帮助人们理解生命活动的分子机理，也能有效地辅助蛋白质设计和基于结构的药物设计，因而结构解析已经成为生物物理领域最重要的研究方向之一。

自 20 世纪八九十年代起，人们就开始发展计算机算法，通过研究序列和结构间的关系，根据氨基酸序列预测蛋白质的三维结构。1994 年，约翰·莫尔特 (John Moult) 等人组织了第一届国际蛋白质结构预测评估竞赛 (Critical Assessment of protein Structure Prediction, CASP)，用于系统评测各种计算方法的预测准确性。2015 年，克里斯·桑德斯 (Chris Sanders) 等人提出可以从多重序列比对中获得氨基酸残基间的共进化关系，从而为结构预测提供额外信息。

2021 年，DeepMind 提出 AlphaFold2 算法，该算法采用了一种全新的端对端模型直接根据序列预测结构，可以提高绝大多数目标蛋白预测模型的精准。对有些蛋白质而言，AlphaFold2 预测的结果与实验解析的模型高度相似，甚至仅根据实验数据都很难区分孰优孰劣。



AI 解码蛋白质结构领域 TOP1% 论文共计 160 篇，其中综述性论文 34 篇，主要综述了各类蛋白质的结构技术方法。其余 126 篇研究性论文主要在蛋白质解构技术和应用两个方面做出了研究。从中美论文数量看，中国论文 80 篇，美国论文 51 篇。

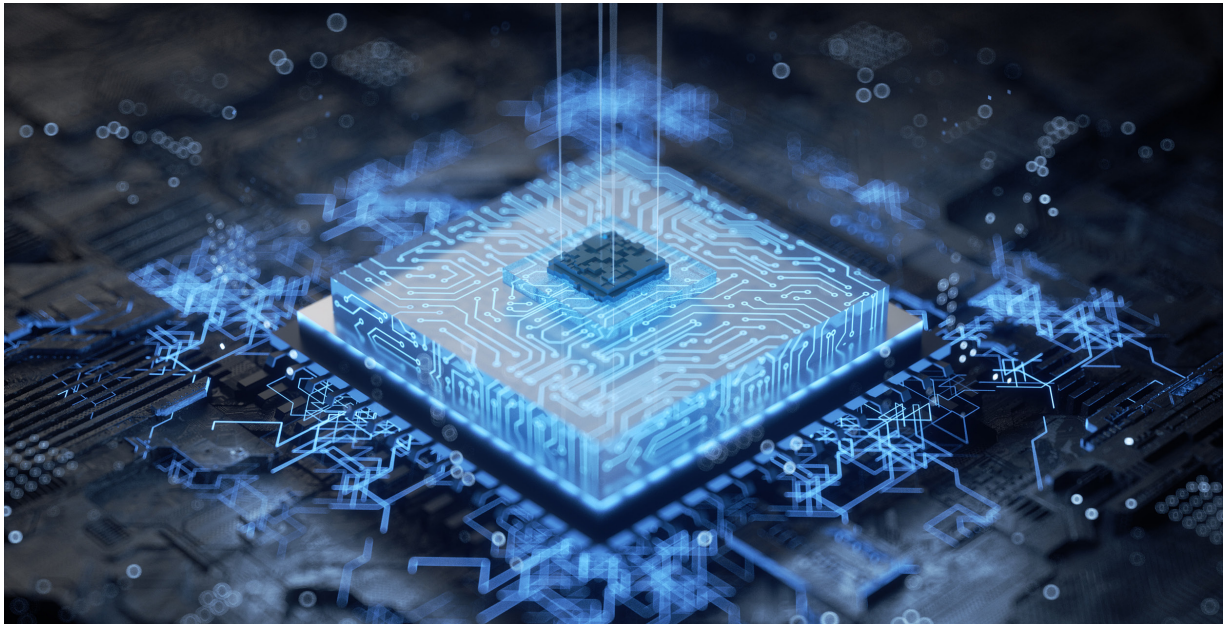
AI 解码蛋白质结构领域专利共 2007 件，中国专利 453 件，美国专利 866 件。专利研究方向以生物数据处理、测序方法为主。全球高价值专利共计 185 件，中美高价值专利 118 件，中国高价值专利 13 件，美国高价值专利 105 件。

总体来说 AI 解码蛋白质结构领域论文的数量远大于专利数量，该领域最重要的科研成果输出途径是论文。假设将论文作为衡量基础性研究生产力的指标，专利作为衡量应用性研发生产力，AI 解码蛋白质结构领域的基础研究生产力远高于应用性研发生产力。

趋势六：移动边缘计算网络

随着物联网、5G、工业自动化、智能制造的兴起，处于物理实体和工业连接之间的移动边缘计算越来越发挥出重要的作用，也逐步实现了集中化和智能化，其低时延、高带宽、个性化、高安全性、高隐私性等特性满足了分布式服务、自动工业控制的需求。

移动边缘计算 (Mobile Edge Computing, MEC) 技术作为第五代移动通信系统的核心技术之一，将云计算能力与业务平台下沉到网络边缘 (基站 / 终端)。靠近用户的无线接入网能就近向



终端用户针对计算密集型、时延敏感型业务提供云计算与互联网技术服务，比如车载导航、虚拟现实技术等。目前，云计算数据中心的集中式处理时代已逐渐转入以万物互联为核心的边缘计算时代，其核心思想是利用网络功能虚拟化和软件定义网络技术，使网络中的存储和计算资源更接近终端用户。随着物联网、5G、工业自动化、智能制造的兴起，移动边缘计算越来越发挥出重要的作用。

移动边缘网络领域 TOP1% 论文共计 175 篇，其中综述性论文 4 篇，主要综述了 5G 和物联网在移动边缘网络的应用。其余 170 篇论文主要在软件开发、网络配置和该技术在汽车等领域的应用为主。从中美论文数量看，中国论文 120 篇，美国论文 54 篇。

移动边缘网络领域专利共 7848 件，中国专利 3723 件，美国专利 2081 件。专利研究方向与数字信息传输、数据交换网络相关。全球高价值专利共计 1686 件，中美两国专利 1237 件，中国高价值专利 668 件，美国高价值专利 569 件。

总体来说，移动边缘网络领域专利的数量远大于论文数量，假设将论文作为衡量基础性研究生产力的指标，专利作为衡量应用性研发生产力，沉浸式扩展现实游戏平台领域的应用性研发生产力要远高于基础研究生产力，该项技术的落地应用场景应用还将不断向多源化、稳定化发展。

趋势七：可解释的 AI

由于人类社会的价值观念和价值体系存在多元化的特点，未来发展具有可解释性的人工智能，加强人机之间的理解，让人工智能技术跳出“黑箱”，建立可解释、可理解、可信任的人工智能体系成为趋势。

可解释人工智能 (explainable artificial intelligence, XAI) 是指人类可以理解人工智能做出的决定或预测的人工智能，是一套流程和方法，可以帮助人类理解模型决策，并发现黑盒模型的问题，如偏差或对抗攻击易感性。可解释性可以帮助 AI 模型的使用者理解机器学习模型是如何得出预测，这就像理解哪些特征驱动模型决策那样简单，但在试图解释复杂模型时会变得困难。



可解释的 AI 与机器学习中的黑匣子概念形成对比。在机器学习中，即使其设计者也无法解释为什么人工智能会得出一个特定的决定。通过完善人工智能系统用户的心智模型，拆除他们的错误观念，可解释的 AI 有望帮助用户更有效地执行。可解释的 AI 可能是对社会解释权的实施。即使没有法律权利或监管要求，可解释的 AI 也是相关的。目前，可解释的 AI 已经在许多领域得到了应用，包括神经网络坦克成像、天线设计、算法交易、医学诊断、自动驾驶汽车等。

可解释的 AI 领域 TOP1% 论文共计 126 篇，其中综述性论文 38 篇，主要综述了可解释 AI 在机器学习和生物医药方面的应用进展。其余 88 篇研究性论文主要分为两个方面。一方面是，可解

释 AI 原理和信任管理方面的理论研究；另一方面是，可解释 AI 在生物医药、自动驾驶方面的应用性研究。从中美论文数量看，中国论文 22 篇，美国论文 60 篇。

可解释的 AI 领域专利共 979 件，中国专利 440 件，美国专利 266 件。专利研究方向以机器学习、图像识别为主。全球高价值专利共计 41 件，中美高价值专利 32 件，中国高价值专利 1 件，美国高价值专利 31 件。

总体来说，可解释的 AI 领域论文的数量远大于专利数量，该领域最重要的科研成果输出途径是论文。假设将论文作为衡量基础性研究生产力的指标，专利作为衡量应用性研究生产力，可解释的 AI 的基础研究生产力远高于应用性研究生产力。

趋势八：基于算法模型和安全隐私的联邦学习技术

近年来，联邦学习作为解决数据孤岛问题的重要技术引起业界广泛关注，并被广泛应用于金融、医疗健康以及智慧城市等领域。未来，联邦学习技术的发展将与边缘计算、区块链和网络安全等多个领域关联，更好的发挥其隐私性、高效性和便捷性。

联邦学习是一种分布式机器学习技术，其本质上是一种分布式机器学习框架，在保障数据隐私安全及合法合规的基础上，实现数据共享，共同建模。它的核心思想是通过在多个拥有本地数据的数据源之间进行分布式模型训练，在不需要交换本地个体或样本数据的前提下，仅通过交换模型参数或中间结果的方式，构建基于虚拟融合数据下的全局模型，从而实现数据隐私保护和数据共享计算的平衡，即“数据可用不可见”、“数据不动模型动”的应用新范式。



随着社会对隐私安全的日益重视，联邦学习作为下一代人工智能大规模协作的基础理论，为目前发展人工智能面临的小数据和隐私等关键问题提供了有效的解决思路。同时，作为网络安全领域新的研究热点，联邦学习在边缘计算、物联网、智慧医疗、金融风控、智慧城市以及涉密数据的安全共享等领域均有着广阔的应用前景。

联邦学习领域 TOP1% 论文共计 66 篇，其中综述性论文 3 篇，主要综述了联邦学习在安全隐私和智慧医疗方面的应用进展。其余 63 篇研究性论文主要分为两个方面。一方面是，联邦学习的算法、安全等理论研究；另一方面是，联邦学习在生物医药、城市信息学方面的应用性研究。从中美论文数量看，中国论文 40 篇，美国论文 26 篇。

可解释的 AI 领域专利共 979 件，中国专利 440 件，美国专利 266 件。专利研究方向以机器学习、图像识别为主。全球高价值专利共计 41 件，中美高价值专利 32 件，中国高价值专利 1 件，美国高价值专利 31 件。

总体来说，联邦学习领域论文的数量远大于专利数量，该领域最重要的科研成果输出途径是论文。假设将论文作为衡量基础性研究生产力的指标，专利作为衡量应用性研发生产力，生物大数据领域的基础研究生产力远高于应用性研发生产力。

趋势九：混合计算

随着万物互联市场的发展，产业界逐渐认识到单一的计算方式不能解决所有问题，“混合计算”借鉴了异构计算的思想——用不同的计算资源处理适合该结构的任务，构建出某领域专用的高效应用组件，从而更好地满足无线互联、视频处理、图像识别、智能制造等多领域的高效处理需求。



混合计算（Hybrid Computing）指利用 5G 的万物互联能力，综合利用云计算、雾计算、边缘计算等计算方式，实现高效协同计算，主要解决云计算、边缘计算、家庭计算、个人计算的协议适配、海量连接、数据存储、设备管理、规则引擎、事件告警等应用场景的共性问题。混合计算三大核心能力包括算力的自协同、算法的自成长、赋能各场景的商业应用自构建。

混合计算领域 TOP1% 论文共计 292 篇，其中综述性论文 16 篇，主要综述了混合计算在物联网、深度学习和算法框架。其余 276 篇研究性论文主要在物联网、隐私和网络安全研究。从中美论文数量看，中国论文 199 篇，美国论文 91 篇。

混合计算领域专利共 11220 件，中国专利 8538 件，美国专利 1291 件。专利研究方向与通信处理、程序控制相关。全球高价值专利共计 293 件，中美高价值专利 185 件，中国高价值专利 95 件，美国高价值专利 90 件。

总体来说，混合计算领域论文和专利的数量相近，假设将论文作为衡量基础性研究生产力的指标，专利作为衡量应用性研发生产力，混合领域的应用性研发生产力与基础研究生产力发展处于平衡状态，随着相关创新研究工作不断普及，未来在该领域可能会取得明显推动科学发展的关键突破。

趋势十：能源区块链

针对能源互联网应用系统中各个模块存在的问题，区块链技术能够依靠自身的链式特征和独特的区块结构融合到能源互联网中的各个层面，解决能源互联网系统中的相关问题。未来，创建区块链一体化能源系统可有效促进清洁能源的产量和能源的及时高效消费利用，提高地区性能源的综合利用率，在保证经济性和稳定性的基础上，实现节能环保目标。

区块链（Blockchain）技术起源于“比特币”，是分布式数据存储、P2P 传输、共识机制、加密算法等计算机技术的新型应用模式。与传统数据库技术相比，区块链技术具有去中心化、数据不可篡改、集体维护、信息公开透明、智能合约、可追溯性、开放性、无需信任系统等特点。虽然，区块链在能源领域的应用案例较少且规模较小，并且实际运营经验不足。但是，随着新能源技术与信息技术的发展和成熟，能源区块链将成为双碳背景下能源结构转型的重要解决方案。

能源区块链领域 TOP1% 论文共计 67 篇，其中综述性论文 15 篇，主要综述了能源区块链管理、能源区块链技术和区块链技术在电力行业的应用。其余 52 篇研究性论文主要在物联网、安全交易方面的研究。从中美论文数量看，中国论文 28 篇，美国论文 16 篇。



能源区块链领域专利共 1761 件，中国专利 1200 件，美国专利 228 件。专利研究方向与支付协议、监督预测系统相关。全球专利共计 33 件，中国无高价值专利，美国高价值专利 26 件。

总体来说，能源区块链领域论文的数量远略高于专利数量，该领域最重要的科研成果输出途径是论文。假设将论文作为衡量基础性研究生产力的指标，专利作为衡量应用性研发生产力，能源区块链领域的基础研究生产力略高于应用性研发生产力。

3

2022 数字科技 热点技术

作为 21 世纪的重要变革力量，数字技术正在加速转型，深刻影响着世界各国的政治、经济、和文化等各个方面。面对数字技术高速发展的总态势，本章以 2022 年新发表的 Top 1%ESI¹⁰ 核心论文、高价值专利¹¹ 和趋势报告^{12 13 14 15 16} 中提到的热点新兴技术进行盘点，并从中选取热度靠前的 AIGC、云计算和量子计算机技术的最新进展、研究力量布局进行介绍。

（一）AIGC

当前，社会各界对 AIGC（AI-Generated Content，人工智能生成内容）这一概念尚无规范统一的定义。从技术角度看，国内对于 AIGC 的理解是利用人工智能技术自动生成内容的新型生产方式。国际上的理解是通过人工智能算法对数据或媒体进行生产、操控和修改的统称¹⁷。

¹⁰ESI 被称为基本科学指标数据库 (Essential Science Indicators Database)，是国际公认的评价大学、学术机构和国家 / 地区国际学术水平和影响力的重要评价指标工具之一。

¹¹ 高价值专利：市场价值为 100 万美元及以上的专利视为高价值专利。

¹² 腾讯研究院。《2022 年十大数字科技前沿应用趋势》[R]。2022-8-24。

¹³ 德勤。《2022 技术趋势》[R]。2022-4-16。

¹⁴ 中国信通院发布。《全球产业创新生态发展报告》[R]。2023-1-24。

¹⁵ 中国电子信息产业发展研究院。2022 中国数字经济发展研究报告 [R]。2021-11-17。

¹⁶ 阿里巴巴达摩院。达摩院 2022 十大科技趋势 [R]。2023-1-28。

¹⁷ 维基百科：“人工智能合成媒体 (AI-generated Media 或 Synthetic Media) ”https://en.wikipedia.org/wiki/Synthetic_media。

最新进展

2022年，在大模型和开源模式的推动下，AIGC(AI-Generated Content, 人工智能生产内容)迅速发展，生成内容百花齐放，模拟结果接近真人传达信息。例如：2月 Deepmind 推出 AI 编码引擎 AlphaCode；4月，Open AI 宣布推出其文本到图像模型 DALL-E 的优化版本 DALL-E2；5月，Google 在 2022 年 I/O 大会上公布了其对话式人工智能模型 LaMDA2；6月，Google 推出 Parti (Pathways Autoregressive Text-to-Image)，该模型最高可扩展至 200 亿参数，并且随着可使用参数数量的增长，其输出的图像也能够更加逼真；7月，Bloom 公司正式开源具备 1,760 亿个参数的大型语言模型 BLOOM；8月，清华大学知识工程实验室 (KEG) 与智谱 AI 共同研发的大规模中英文预训练语言模型 GLM-130B 正式发布；12月，Open AI 的大语言模型聊天机器人 ChatGPT 上线，建立在 GPT-3.5 模型之上，该模型在上线后的短短 5 天里，注册人数突破百万。详见下图：

图 27 2022 年 AIGC 技术大事年表



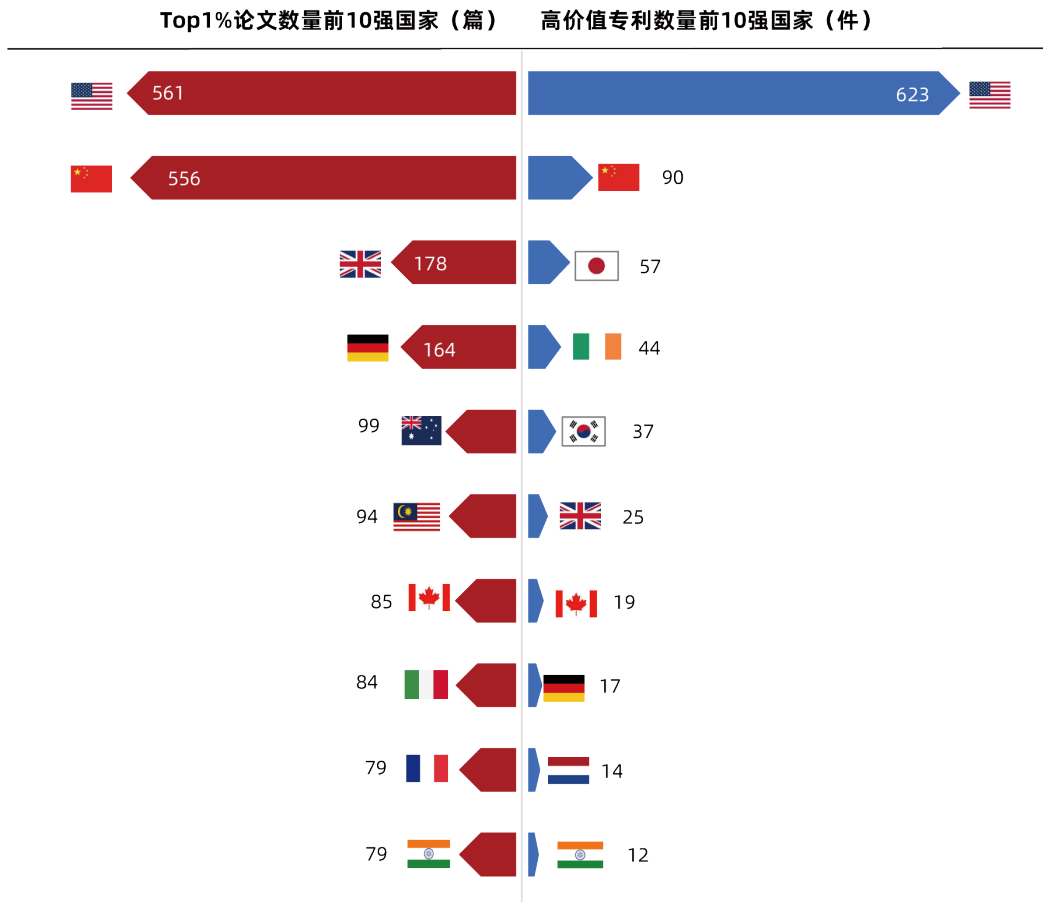
- 7月 ● Bloom 公司正式开源具备 1,760 亿个参数的大型语言模型 BLOOM
Meta 展示其新的图像生成器 “Make-A-Scene”
- 8月 ● Stability AI 发布文本到图像模型 Stable Diffusion
清华大学知识工程实验室 (KEG) 与智谱 AI 共同研发的大规模中英文预训练语言模型 GLM-130B 正式发布
- 10月 ● 微软开始将由 DALL-E2 提供支持的生成人工智能技术集成到其 Bing 搜索引擎、Edge 浏览器和新的 Microsoft Designer for Office 中
DeepMind 公布了 AlphaZero 的拓展版本 AphaTensor，用于发现高效矩阵乘法算法
- 11月 ● Open AI 以 API 的形式发布了 DALL-E2，允许开发人员将模型集成到应用中
- 12月 ● Open AI 的大语言模型聊天机器人 ChatGPT 上线，建立在 GPT-3.5 模型之上
由独立开发者开发的文本到视频应用 QuickVid 上线并开始流行，整合了多个生成式 AI 系统

说明：时间范围为 2022 年 1 月至 2022 年 12 月；
数据来源：AMiner 科技情报平台。

研究实力对比

下图显示了 AIGC 领域被引量 Top1% 论文和高价值专利前 10 强国家。其中 Top1% 论文数量和高价值专利数量最多的国家是美国，共有 561 篇 Top1% 论文，623 件高价值专利；中国的 Top1% 论文数量和高价值专利数量排名第二，共有 556 篇 Top1% 论文，90 件高价值专利；全球 Top1% 论文数量排名第三的是英国，为 178 篇；全球高价值专利数量排名第三的是日本，为 57 件。总体来看，虽然中国 AIGC 高价值专利数量排在全球第二，但是专利数量仅为排名第一美国的 14%，AIGC 应用落地还有待进一步发展。

图 28 AIGC 领域被引量 Top1% 论文和高价值专利前 10 强的国家



说明：检索时间范围为 2012 年 1 月至 2021 年 12 月。
数据来源：AMiner 科技情报平台。

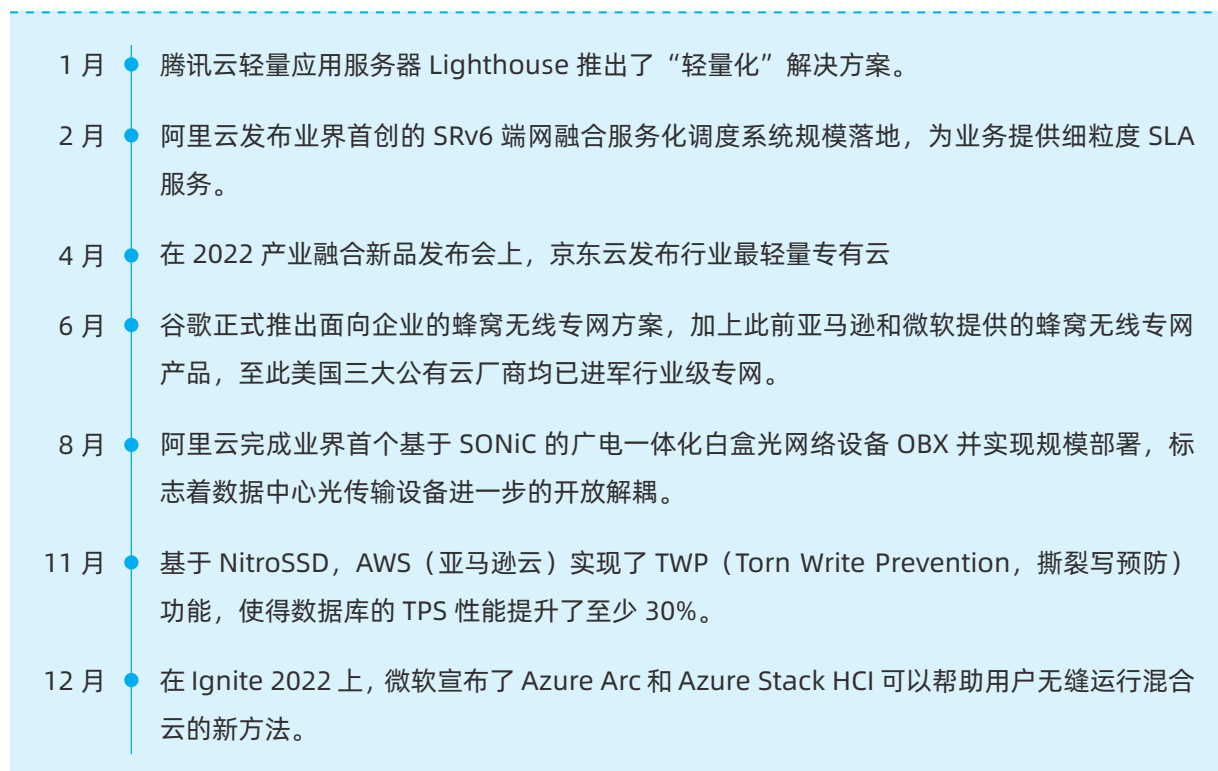
(二) 云计算

从虚拟机技术开始，云计算技术以实现 IT 资源池化为目标，即将所有计算、存储资源整合成统一的计算、存储资源池，从而实现 IT 资源的按需自助服务，实行统一的资源管理。今天，云计算技术已经发展到 2.0 阶段，以容器、微服务、Serverless 为代表的新一代云原生技术对外提供更加极致的弹性公共服务能力、服务自治、高可用性、跨数据中心调度与跨多种硬件的兼容适配能力。

最新进展

随着信息技术与互联网技术的不断发展,云计算技术随之快速发展,为实现人工自动化技术在计算机网络中的有效应用,云计算的多源信息服务系统的发展还需要各种成熟技术来支持。2022年,1月腾讯云轻量应用服务器 Lighthouse 推出了“轻量化”解决方案。2月,阿里云发布业界首创的 SRv6 端网融合服务化调度系统规模落地,为业务提供细粒度 SLA 服务。8月,阿里云完成业界首个基于 SONiC 的广电一体化白盒光网络设备 OBX 并实现规模部署,标志着数据中心光传输设备进一步的开放解耦。11月,AWS(亚马逊云)实现了 TWP(Torn Write Prevention, 撕裂写预防)功能,使得数据库的 TPS 性能提升了至少 30%。12月,微软宣布了 Azure Arc 和 Azure Stack HCI 可以帮助用户无缝运行混合云的新方法。

图 29 2022 年云计算技术大事年表

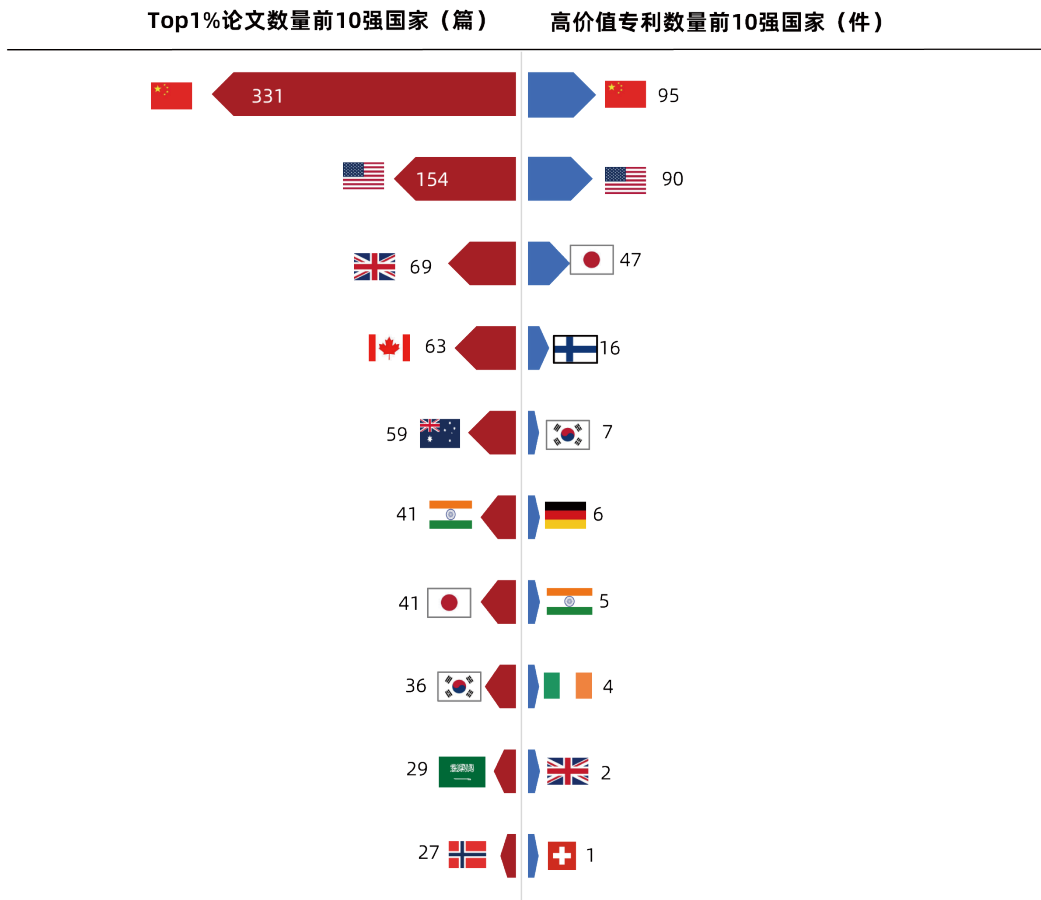


说明: 时间范围为 2022 年 1 月至 2022 年 12 月;
数据来源: AMiner 科技情报平台。

研究实力对比

下图显示了云计算领域被引量 Top1% 论文和高价值专利前 10 强国家。其中 Top1% 论文数量和高价值专利数量最多的国家是中国，共有 331 篇 Top1% 论文，95 件高价值专利；美国的 Top1% 论文数量和高价值专利数量排名第二，共有 154 篇 Top1% 论文，90 件高价值专利；全球 Top1% 论文数量排名第三的是英国，为 69 篇；全球高价值专利数量排名第三的是日本，为 47 件。在 Top1% 论文数量方面，中国遥遥领先，是美国的 2 倍；在高价值专利方面，中国同样占据优势。

图 30 云计算领域被引量 Top1% 论文数量和高价值专利前 10 强的国家



说明：检索时间范围为 2012 年 1 月至 2021 年 12 月。

数据来源：AMiner 科技情报平台。

（三）量子计算机

量子计算机（quantum computer）是一类遵循量子力学规律进行高速数学和逻辑运算、存储及处理量子信息的物理装置。量子计算机的特点主要有运行速度较快、处置信息能力较强、应用范围较广等。其最新进展和各国研究实力如下：

2022 最新进展

当前，量子技术研究已成为世界科技研究的一大热点，世界各主要国家高度关注量子信息技术发展，纷纷加大政策和资金支持，力争抢占新兴信息技术制高点。主要大事有：2022年1月，荷兰量子计算公司 QuTech 在量子纠错方面取得了重大进展：他们将编码量子信息的高保真操作与可扩展量子比特的解决方案集成在一起，从而实现重复信息稳定传输。3月，印度和芬兰制定了建立印度 - 芬兰量子计算虚拟网络中心的详细计划，并为该项目确定了 IIT Madras、IISER Pune 和 C-DAC Pune 研究所。7月，NVIDIA 宣布发布其统一计算平台 QODA(量子优化设备架构)，以加速人工智能、高性能计算、金融和健康等领域的量子计算研究。11月，芬兰 IQM 量子计算机公司、阿尔托大学和芬兰 VTT 技术研究中心的一组科学家发现了一种新的超导量子比特——unimon，可提高量子计算的准确性。12月，中国首个量子计算机和超级计算机协同计算系统解决方案（简称“量超协同”系统解决方案）发布。详见下图：

图 31 2022 年量子计算机技术大事年表

- | | |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1月 | <ul style="list-style-type: none"> 麻省理工学院 (MIT) 计算机科学与人工智能 (CSAIL) 的科学家们开发了一种名为 Twist 的量子计算编程语言。该语言依赖于纯粹的概念，通过强制没有纠缠来构建直观的程序，减少计算错误。 荷兰量子计算公司 QuTech 在量子纠错方面取得了重大进展：他们将编码量子信息的高保真操作与可扩展量子比特的解决方案集成在一起，从而实现重复信息稳定传输。 中国科学技术大学郭光灿院士团队将量子门测试理论与其近年来开发的多参数量子精密测量平台相结合，并通过改进量子门检验的数据处理算法，实现了在量子门缺陷检验上保留高效率的同时提升误差检测的可靠性。 |
| 2月 | <ul style="list-style-type: none"> 美国能源部 (DOE) 阿贡国家实验室和芝加哥大学的研究人员发现：量子态持续时间可超 5 秒，这成为量子科学领域的新纪录。该项突破将有望实现更复杂的量子计算，甚至实现“量子互联网”。 |
| 3月 | <ul style="list-style-type: none"> 日本 NTT 的子公司 NTT Research 和康奈尔大学联合研发出了一种算法，该算法将深度神经网络训练应用于可控物理系统，并在三种非常规硬件上进行了展示，并发表于 Nature。 |

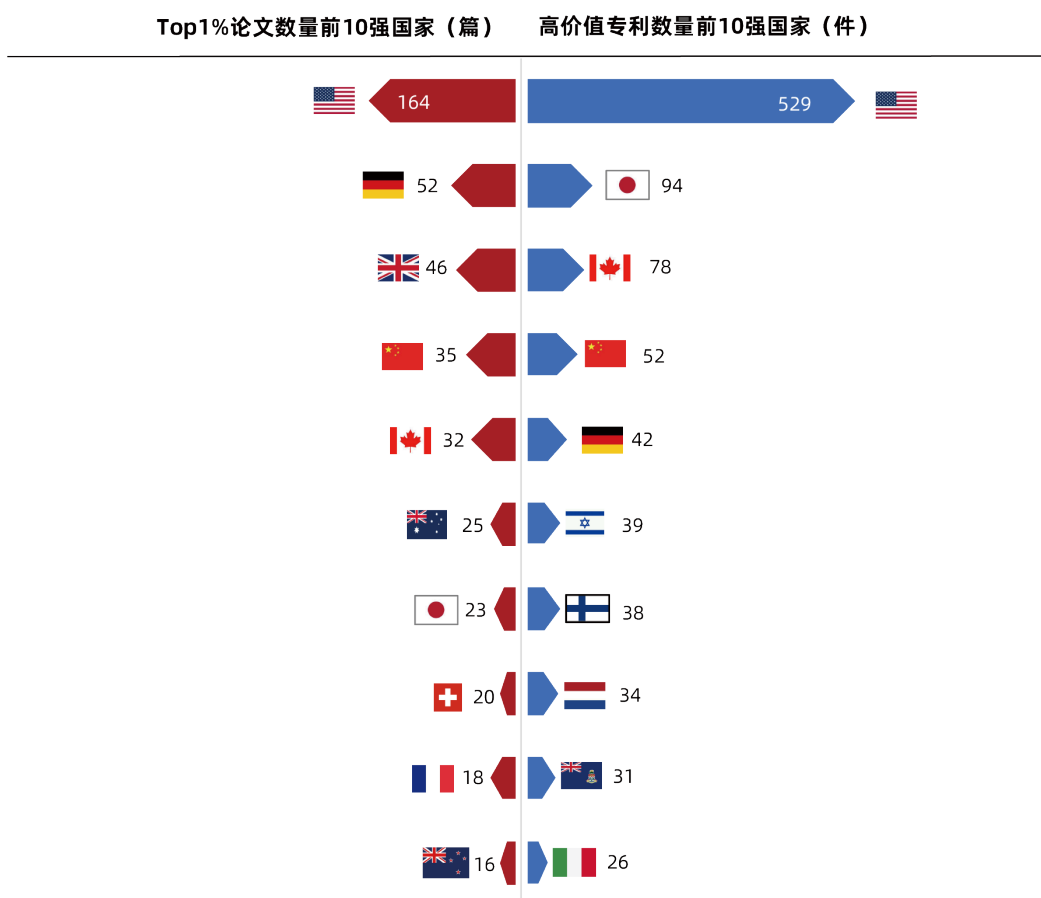
- 印度和芬兰制定了建立印度 - 芬兰量子计算虚拟网络中心的详细计划，并为该项目确定了 IIT Madras、IISER Pune 和 C-DAC Pune 研究所。
- 6 月 ● 谷歌量子人工智能项目团队使用 Sycamore（悬铃木）量子处理器开展了超大规模的化学模拟实验，并开发出一种新技术，这种技术有助于抵抗量子电路中常见的噪声，研究成果论文已发表于 Nature。
- 北京量子信息科学研究院、清华大学龙桂鲁教授团队和陆建华教授团队合作设计和实现了一种相位量子态与时间戳量子态混合编码的量子直接通信新系统，通信距离达到 100 公里，是当前世界最长的量子直接通信距离。
- 7 月 ● 哈佛大学、QuEra Computing、麻省理工学院和因斯布鲁克大学等的研究团队联合开发出一种处理量子信息的新方法：通过平移的方式动态改变系统中原子的布局，并在计算过程中将它们相互连接，成果发表在 Nature 上。
- NVIDIA 宣布发布其统一计算平台 QODA(量子优化设备架构)，以加速人工智能、高性能计算、金融和健康等领域的量子计算研究。
- 谷歌公开量子虚拟机 (QVM)，是一种用于原型、测试和优化量子电路的工具，具有类似处理器的输出，可以从 Colab 笔记本电脑部署。
- 8 月 ● 美国洛斯阿拉莫斯国家实验室在“量子 +AI”领域取得了重大突破：他们的最新研究表明，训练一种量子神经网络只需要少量数据集。相比于经典计算中 AI 对数据的巨大需求，这一研究成果将大大推进量子 AI 优越性的加速实现。
- 10 月 ● 谷歌人工智能量子团队的研究人员使用量子计算机计算了迄今为止最大规模的化学模型。该量子算法的进一步发展将使我们能高精度地预测化学反应的过程，帮助化学家在反复试验上节省大量时间。
- 新南威尔士大学研究团队延长量子相干时间实现基准增长 100 倍。
- 11 月 ● 芬兰 IQM 量子计算机公司、阿尔托大学和芬兰 VTT 技术研究中心的一组科学家发现了一种新的超导量子比特——unimon，可提高量子计算的准确性。
- TCS 在 AWS 上为企业客户提供了量子计算和应用程序的量子计算实验室。研究开发将由 AWS 为量子计算提供的全面管理服务“braket”提供支持。
- 中国量旋科技公司发布了三款便携式量子计算机 Gemini、Gemini Mini 和 Triangulum，可用于教育目的。这些计算机使用核磁共振 (NMR) 进行量子计算，利用原子的自旋运动。
- 12 月 ● 中国首个量子计算机和超级计算机协同计算系统解决方案（简称“量超协同”系统解决方案）发布，该方案可以将计算任务在量子计算机和超级计算机之间进行分解、调度和分配，实现量子计算和超级计算机的高效协同。
- Quantinuum 公司发布 InQuanto 2.0，新版 InQuanto 引入了提高效率的新工具，用于加速矢量计算和积分运算符类的高级算法。

说明：时间范围为 2022 年 1 月至 2022 年 12 月；
数据来源：AMiner 科技情报平台。

研究实力对比

下图展示了量子计算机领域被引量 Top1% 论文和高价值专利前 10 强国家。其中 Top1% 论文数量和高价值专利数量最多的国家是美国，共有 164 篇 Top1% 论文，529 件高价值专利；Top1% 论文数量排名第二和第三的依次是德国和英国，分别有 52 篇和 46 篇。高价值专利数量排名第二和第三的依次是日本和加拿大，分别有 94 件和 78 件。中国的 Top1% 论文数量为 35 篇，全球排名第四；高价值专利数量为 52 件，全球排名第三。

图 32 AIGC 研究领域被引量 Top1% 论文数量和高价值专利前 10 强的国家



说明：检索时间范围为 2012 年 1 月至 2021 年 12 月。
数据来源：AMiner 科技情报平台。

Chapter 5

我国数字科技发展 问题总结与建议

1. 中国数字科技论文两项关键指标将登顶全球第一

目前中国数字科技论文两项关键指标，即论文总量和 Top 1% 论文量均处于全球“老二”的位置，但中美该两项指标的“黄金交叉”已经出现，而且中国的增长势头强劲，因此中国数字科技论文总量和最卓越论文数量登顶全球第一的位置可期。

2. 中国是数字技术专利大国，而非强国

中国数字技术专利数量全球遥遥领先，专利数量前 10 强机构一半为中国公司和高校，但是高价值专利数量排名跌至全球第 4，仅为美国的 12%。中国数字技术专利胜在“量”上，输在“质”上，因此中国是数字技术专利大国，而非强国。

3. 中国数字科技发展前景在于向价值链高端跃升

专利是倾向于应用技术的科研成果，其市场价值直接反映该技术应用产品处于产业价值链的位置。从专利市场价值分布看，中国数字技术专利价值在 30 万美元以下的占 98%，因此中国数字技术产业仍处于全球价值链低端。论文是倾向于基础研究的科研成果，往往是科技创新突破的先导。未来中国数字科技的发展前景在于从日益强大的基础研究成果中实现实际应用转化，推动数字技术产业向价值链高端跃升。

4. 高层次人才不足需加大人才培养力度

我国数字科技人才基数全球第一，占全球 17%，但高层次数字科技人才仅占全球 9%，而美国占 25%，我国仅为美国的 35%。数字科技为世界关键前沿科技，高层次人才落后局面不利于我国数字科技总体能力达到世界领先水平的战略目标。扭转该不利局面，关键是要创新培养机制，提升人才质量，加大力度持续培养本土高层次人才。1) 注重复合培养模式。鼓励高校在原有学科基础上拓宽教育内容，注重数字科技各方向交叉融合的复合专业培养新模式。2) 重点培养有潜力的青年人才。把培育国家战略人才力量的政策重心放在青年科技人才上，制定实施基础研究人才专项，长期稳定支持一批在数字科技领域取得突出成绩且具有明显创新潜力的青年人才。3) 鼓励对外学术交流。鼓励和引导国内科研人才及其团队与全球顶尖数字科技研究机构开展交流与合作。

5. 人才流失严重需采取措施引留人才

我国是全球数字科技人才流失第一大国，而美国是最大的流入国。目前，全球频频上演“高科技战”、“人才争夺战”，在此的背景下，数字科技人才流失问题势必影响我国未来产业发展安全，甚至国家安全，必须引起高度重视。首先，吸引海外高层次人才回国发展。统筹现有人才计划，加强高层次人才特别是海外顶尖华人青年人才的引进工作，对于一时难以回国或来华工作的高层次人才实行柔性引进，采用学术交流、项目合作、技术咨询等方式，柔性引进高层次人才。其次，用好并留住高层次人才。优化科研环境，以信任为基础使用人才，为各类人才搭建干事创业平台，用好用活高层次人才，赋予人才更大的科研自主权，并完善人才评价体系。

6. 企业人才储备薄弱需鼓励企业引育人才

我国数字科技人才集中分布在高校和科研院所。全球数字科技人才前 10 强机构榜单，乃至国内前 10 强机构榜单均不见中国企业身影，而美国高科技巨头谷歌和微软的高层次人才储备实力高居全球第二和第四名。身处科研一线的企业是科技创新的重要主体。因此，1) 有必要鼓励企业加强内部创新环境建设，协助企业引进高层次人才，支持企业与高等院校和科研院所共同培养基础研究人才。2) 发挥政府科技计划的导向作用，在重大专项、重点研发计划论证和实施过程中，支持企业承担政府科研项目，为企业引育人才和提高基础科研能力注入动力。

Chapter 6

附录

附录 I：技术趋势研判方法说明

数字科技是基于物理世界和数字世界映射互动的体系提炼的一个新概念，是当今世界创新速度最快、通用性最广、渗透性和引领性最强的领域之一。本报告通过人工智能技术，从海量大规模非结构化的数据中抽取数字科技相关知识，生成结构化的知识数据，并根据关系和属性链接等构建数字科技知识图谱，确保了知识体系的完整性，本次研究范围覆盖与数字科技高度相关的八大方向，依次是人工智能、区块链、大数据、云计算、物联网、量子科技、新一代无线通信、XR 技术。

本报告围绕数字科技领域核心论文和专利，利用大数据分析挖掘、知识图谱、自然语言处理等技术，对领域论文和专利的分布格局、人才资源和技术趋势进行分析。

（一）科研实力与人才布局分析。时间范围：2012 年 1 月—2021 年 12 月。关键词：采用知识图谱技术抽取，经顾问组专家指导形成领域关键词表¹⁸。选取领域强相关关键词进行检索。检索范围：基于自由大数据服务平台，覆盖全球近 1.4 亿人次学者，3.2 亿篇学术论文，6.8 亿件专利。

（二）技术趋势分析。前沿技术（或研究前沿）通常被视为“最具发展潜力的新兴研究领域或研究主题”，美国知名科研情报机构科睿唯安（Clarivate）将其定义为“研究前沿由一组共同引用的核心论文以及引用一篇或多篇这些核心论文的当前来源论文簇组成”¹⁹。近年来，前沿技术与挖掘受到学界广泛关注，论文数量（贡献度）和论文被引频次（影响度）成为衡量技术前沿度的重要评估指标²⁰。

一般来说，技术趋势是指某种技术在一定时间跨度内的水平变化方向。为了有效表现先技术在某段时间的发展趋势，本报告引入论文数量、论文被引频次和论文发表时间作为重要评估指标，将数字科技领域被引量 Top 1%ESI²¹ 核心论文作为贡献度测算指数，文的被引用次数作为影响力测算指数，论文首次被引至最近一次被引时间间隔作为识别“学术睡美人”测算指数，论文被引用变化趋势作为判断技术未来发展潜力的测算指数。旨在通过论文、专利数据的被引用量、被引时间跨度、被引量变化趋势来发现技术影响力，且近年来学术影响力处于增长状态的技术。计算方式如下式所示：

¹⁸ 数字技术关键词见附录 II

¹⁹ 中国科学院科技战略咨询研究院等.《2022 研究前沿热度指数报告》[P]. 北京. 2022.12.27.

²⁰ Fajardo-Ortiz D, Lopez-Cervantes M, Duran L, et al. The emergence and evolution of the research fronts in HIV/AIDS research[J]. PLoS ONE, 2017, 12(5):e0178293.

²¹ ESI 被称为基本科学指标数据库 (Essential Science Indicators Database), 是国际公认的评价大学、学术机构和国家 / 地区国际学术水平和影响力的重要评价指标工具之一。

$$\text{前沿度} = \frac{C_{total}}{T \cdot S_{trend}}$$

其中, C_{total} 代表科技成果的被引用量, 反映科技成果产生的学术影响力; T 表示科技成果第一次被引用至最近一次被引用的时间间隔, 对于被引用量相同的文献, 该值越小, 表示文献所包含的技术点在短期内的影响力越大, 越具有前沿特性; S_{trend} 表示研究目标文献年度被引用量的变化趋势, 与被引量时序数据的斜率有关, 其他条件一致的情况下, 被引用量呈现上升趋势的文献能够通过计算得到更大的前沿度数值。

附录 II: 数字科技主题词

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词	
人工智能	模式识别	文字识别	OCR				
			汉字识别				
		图像识别	人脸识别				
			虹膜识别				
			指纹识别				
		生物特征识别	人脸识别				
			虹膜识别				
			指纹识别				
		语音识别					
		自然语言处理	自然语言处理基础研究	词法分析			
	句法分析						
	语义分析						
	认知语言学						
	语言表示						
	知识图谱						
	自然语言处理应用研究		文本分析	文本分析			
				文本聚类			
			信息抽取	命名实体识别			
				实体消歧			
				关系抽取			
				事件抽取			
			自动文摘				
	信息推荐						
自动问答							

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词	
			机器翻译				
			社交媒体处理				
	计算机视觉	图像处理					
		特征提取					
		图像分割					
		三维重建					
		图像拼接					
		图像识别					
		图像检索					
		目标检测					
		目标跟踪					
		运动行为分析					
		图像标注					
		视频标注					
		虚拟现实					
		增强现实					
		语音技术	语音合成				
			语音识别				
	声纹识别						
	语音增强						
	语音翻译						
	语音系统						
	机器学习	监督学习					
		半监督学习					
		强化学习					
		集成学习	Boosting				
			Bagging				
			随机森林				
			旋转森林				
		数据降维	主成分分析				
			主成分回归				
			因子分析				
			流形学习				
			线性判别分析				
			局部线性嵌入				
			拉普拉斯特征映射				
	度量学习						
	特征选择	稀疏表示					
		字典学习					

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词	
			压缩感知				
		计算学习理论	VC 维				
		概率图模型	隐马尔可夫模式				
			马尔可夫随机场				
			条件随机场				
			变分推断				
			马尔可夫链				
			马尔可夫蒙特卡洛采样				
			吉布斯采样				
			神经网络	神经元模型			
				反向传播			
				深度学习	深度信念网络	自由编码器	欠完备自编码
		稀疏自编码					
		去噪自编码					
		变分自编码					
		对抗自编码					
		卷积神经网络					
		循环神经网络					
		生成对抗网络					
		深度学习框架			Theano		
					PyTorch		
					MXNet		
					Torch		
					TensorFlow		
					Caffe		
		无监督学习	关联规则学习				
			聚类分析	模糊聚类			
				K-means 算法			
				层次聚类			
			k 近邻法				
	数据挖掘	数据预处理					
		关联规则学习	频繁项集挖掘				
			序列模式挖掘				
			高级模式挖掘				
		数据分析	线性模型	线性回归			
				对数几率回归			
				多项逻辑回归			

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词
				Softmax 回归		
		聚类分析	排序			
			模糊聚类			
			K-means 算法			
			层次聚类			
		数据异常检测				
		数据降维	主成分分析			
			生成回归			
			因子分析			
			流形学习			
			线性判别分析			
			局部线性嵌入			
			拉帕拉斯特征映射			
		数据可视化				
		数据仓库				
		信息推荐				
		信息检索				
		Kaggle				
区块链	密码学	Hash 算法	MD5、SHA-1 和 SHA-2			
			数字摘要			
		加密算法	对称加密算法	MD5、SHA-1 和 SHA-2		
			非对称加密算法	RSA		
				ElGamal		
			混合加密机制	椭圆曲线		
		数字签名和证书	数字签名			
			数字证书			
		PKI 体系	OpenSSL			
			EJBCA			
			OpenCA			
		Merkle 树	完整性校验			
			SPV 简化支付验证			
			快速比较大量数据			
			零知识认证			
		同态加密	函数加密			
	以太坊	智能合约	对比传统合约			

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词	
			交易模型				
		协议设计	以太坊虚拟机				
			账户				
			交易				
			一致性				
		开发实战	以太坊钱包	Geth			
			Solidity 语言				
	Hyperledger	组织结构	BlockChain Explorer				
				Fabric			
				Sawtooth Lake			
				Iroha			
				Cello			
			Fabric	机构设计	会员		
					区块链		
					链码		
				安装部署	集群部署		
					部署智能合约		
		开发实战	Dapp 实战开发				
	比特币	中本聪白皮书	账户/地址				
				交易方式			
				脚本系统			
				区块信息			
				双花问题			
				避免攻击			
			挖矿	矿机、矿工			
				算力			
				共识机制	POW		
			共识机制		POS		
				POW			
			闪电网络	POS			
				RSMC			
				HTLC			
			交叉	侧链技术			
		区块扩容					
		软分叉和硬分叉					
	分布式系统	一致性问题	顺序一致性				
				线性一致性			
				CAP 原理			
				FLP 不可能原理			

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词
大数据	基础技术	分布式系统共识	Paxos 算法			
			Raft 算法			
		拜占庭问题与算法	拜占庭问题			
			BFT 算法			
	数据采集	数据分片式路由				
		数据复制&一致性				
		大数据常用算法与数据结构				
	数据传输	系统日志采集				
		外部数据采集				
		IOT 设备数据采集				
	数据存储	消息队列				
		数据同步				
		数据订阅				
		序列化				
		物理存储				
		分布式文件/对象存储系统				
		分布式关系型数据库				
		分析型数据库				
		搜索引擎				
键值存储数据库						
图数据库						
列存储数据库						
文档数据库						
时序数据库						
RDF 数据库						
多媒体数据库						
事件存储数据库						
数据处理	数据计算					
	数据仓库					
	数据湖					
	数据分析					
	数据挖掘					
数据应用	平台工具					
	数据服务					
	数据可视化					
	数据共享					
		数据预警				

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词	
	数据治理	数据规范标准					
		元数据管理					
		数据质量					
		主数据管理					
		大数据架构管理					
		数据安全					
		业界框架产品					
云计算	SaaS	Web Server	Nginx	构架原理			
				负载均衡			
				反向代理			
				正向代理			
			Tomcat	架构原理			
				调优方案			
		缓存	本地缓存				
			客户端缓存				
			服务端缓存	Memcached			
				Redis			
		消息队列	RabbitMQ				
			RocketMQ				
			ActiveMQ				
			kafka				
	Redis 消息推送						
	搜索引擎	ES 的分布式架构原理	原理				
			倒排索引				
			集群化部署				
	LaaS	虚拟化技术	软件辅助虚拟化	CPU 虚拟化			
				内存虚拟化			
				I/O 虚拟化			
			安全虚拟化	硬件辅助虚拟化	CPU 虚拟化的硬件支持 (VT-x)		
					内存虚拟化的硬件支持 (EPT)		
IO 虚拟化的硬件支持 (VT-d)				IO 虚拟化的硬件支持 (VT-d)			
		时间虚拟化					
类虚拟化		CPU 虚拟化					
		内存虚拟化					
		I/O 虚拟化					
	时间与时钟管理						

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词	
		OpenStack	OpenStack 体系架构	OpenStack 发展历程			
				OpenStack 组件关系			
			计算组件	Nova 体系结构			
				Nova API			
				Rolling Upgrade			
				Shueduler 典型工作流程			
			存储组件	Swift			
				Cinder			
				Glance			
				Ceph			
			网络组件	Neutron 体系结构			
				Neutron API			
				网络插件	ML2 Plugin		
					Port Binding		
					Open vSwitch Agent		
					Service Plugin		
				其他	DVR		
					SDN		
			NFV/SRIOV				
			OVS & DPDK				
			安全组件	Keystone 体系结构			
				Keystone 启动流程			
				令牌及证书相关			
				可信计算池			
			计量与监控	Ceilometer			
				Aodh			
				Gnocchi			
				Panko			
			OpenStack 部署与运维	OpenStack 部署	Devstack 一键部署		
					docs.OpenStack.org 分布式部署		
		扩容&升级					
		Openstack 运维		Horzion 控制面			

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词	
					板		
					维护&故障&调试		
					网络排障&日志监控		
					备份和恢复		
	PaaS	Docker	Docker 基础	Docker 安装			
				Docker 基本操作			
			Docker 核心原理	内部资源隔离	namespace		
					cgroups		
				Docker 架构	daemon		
					client		
				libcontainer			
				存储管理			
				数据卷			
				网络管理	Docker 网络基础		
					Docker daemon 原理		
					libcontainer 原理		
			link 原理				
			pipwork 原理				
			容器安全				
			Docker 编排三剑客	Compose			
				Machine			
				Swarm			
		Kubernetes	Kubernetes 体系架构	Kubernetes 设计解读			
				Kubernetes 基本组件			
			Kubernetes 组件原理				
			Kubernetes 部署与运维	Kubernetes 部署	Kubernetes 快速部署		
					二进制文件部署		
					重置&升级		
				Kubernetes 运维	Kubernetes 集群管理	Dashboard	
						Kuboard	
						Prometheus & Grafana	

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词	
						Helm 包管理工具	
					Trouble Shooting		
			Kubernetes 开发指南	Kubernetes API 详解			
				Kubernetes AIP 扩展			
				新功能	调度器		
			Pod 的垂直扩容				
				GPU 支持			
		Kubeedge	边缘计算场景				
			Kubeedge 安装部署	X86			
				arm32 & arm64			
			Kubeedge 架构及其组件	Kubeedge 架构原理			
				Edge 端组件	EdgeD		
					EventBus		
					MetaManager		
					Edgehub		
				Cloud 端组件	Device Twin		
			Edge Contoller				
		CloudHub					
			Device Controller				
物联网	感知识别层	芯片	系统级芯片				
			传感器芯片				
			基带芯片				
			射频芯片				
		传感器	物理传感器				
			化学传感器				
			生物传感器				
		感知设备	RFID 读写器				
			条码扫描器				
			卫星定位装置				
	语音识别设备						
				视觉成像设备			
	网络传输层	通信模组	蜂窝通信模组	5/6G			
				NB-IoT/eMTC			
非蜂窝类模组			WiFi 模组				

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词
		通信网络	蜂窝通信网络	蓝牙模组		
				5/6G		
				NB-IoT/eMTC		
			非蜂窝通信网络	SIM卡		
				WiFi		
				蓝牙		
		基础通信设施	基站	Zigbee		
				LoRa/Sigfox		
				天线系统		
			服务器			
	平台管理层	物联网平台	连接管理平台			
			设备管理平台			
			应用使能平台			
			业务分析平台			
		系统&软件	操作系统			
			软件开发			
	应用服务层	智能终端	可穿戴设备			
			智能家电			
			消费电子			
			智能仪表			
			无人机			
			系统集成应用服务	车联网		
		智能家居				
		智慧城市				
智慧农业						
远程医疗						
量子科技		量子计算	量子因数分解			
	量子搜索					
	量子通信	量子密钥分发				
		量子隐性传态				
		通信终端	量子安全	量子密钥管理模块		
				终端加密		
			终端解密			
		终端通信				
	通信模式选择模块					
	量子测量	量子目标识别	量子干涉雷达			
			量子照射雷达			

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词
			量子增强雷达			
		量子重力测量	喷泉式原子重力仪			
			自由下落式重力仪			
		量子磁场测量	光泵磁力仪			
			原子 SERF 磁力仪			
			相干不居囚禁磁力仪			
			金刚石色心磁力仪			
		量子定位导航	量子导航系统	六星定位系统		
				三星定位系统		
			量子加速度计	原子干涉加速度计		
				金刚石色心加速度计		
			量子陀螺仪	原子干涉陀螺仪		
				超流体干涉陀螺仪		
				核磁共振陀螺仪		
				SERF 陀螺仪		
				金刚石色心陀螺仪		
			量子时间基准	量子时钟源	冷原子喷泉钟	
		原子/离子光钟				
		相干布囚禁钟				
		量子同步协议		量子保密时间同步协议		
				纠缠消除色散时间同步协议		
				符合测量纠缠光子对时间同步协议		
				HOME 干涉时间同步协议		
新一代无线通信技术	超密集组网	关键技术	多链接技术			
			无线回传技术			
		部署架构	微基站+微基站			
	宏基站+微基站					
	动态自组织网络:					

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词	
	SON						
	网络功能虚拟化： NFV						
	5G/6G						
	终端设备						
	eMBB 大流量移动 宽带业务						
	mMTC 大规模物 联网业务						
	uRLLC 无人驾驶						
XR (虚拟现实、混 合现实、增强现 实)	计算机图形学	数学基础	向量				
			齐次坐标				
			几何变化				
		建模	几何造型				
			曲线/曲面				
			消隐				
		渲染	光栅化				
			光照模型				
			交互技术	二维图形交互技 术			
		计算机仿真	仿真建模技术	机理建模			
				辨识建模			
				物理建模			
	智能建模						
	多媒体建模						
	仿真支持系统与 平台技术		仿真语言与工具				
			网络化仿真平台				
			仿真资源库管理				
			仿真计算机系统				
			仿真可视化系统				
	仿真应用技术		仿真应用共性技 术	仿真 VV&A 技术			
				仿真实验技术			
				仿真实验结果分 析与评估技术			
			应用领域仿真技 术	自然科学与工程 信息技术			
				社会科学仿真技 术			

一级主题词	二级主题词	三级主题词	四级主题词	五级主题词	六级主题词	七级主题词
				管理科学仿真技术		
				生命科学仿真技术		
				军事科学仿真技术		
	人机接口	脑信号记录				
		信号处理				
		模式识别				
		机器学习				
		感知反馈				
		刺激信号处理				
		脑刺激				
		基础神经科学				
	传感技术	光声光谱传感技术				
		分布光纤传感技术				
		法拉第磁光传感技术				
		液态金属传感技术				
		磁阻电流传感技术				
	多媒体技术	视觉	位图图像			
			动态图像	动态影像		
				真实感三维动画		
			矢量图形			
			文字			
		听觉	语音（人类语言）			
			声响（自然界）			
			音乐			
		触觉	振动			
			运动			

数字技术前瞻

DT时代 · 共创新知

审 核：高红冰

编 辑：苏中、安筱鹏

联系邮箱：ALIRESEARCH@ALIBABA-INC.COM

阿里研究院

WWW.ALIRESEARCH.COM

