

2019年第2期(总第2期)

国际战略与安全研究报告

INTERNATIONAL SECURITY AND STRATEGY STUDIES REPORT

预调研报告:

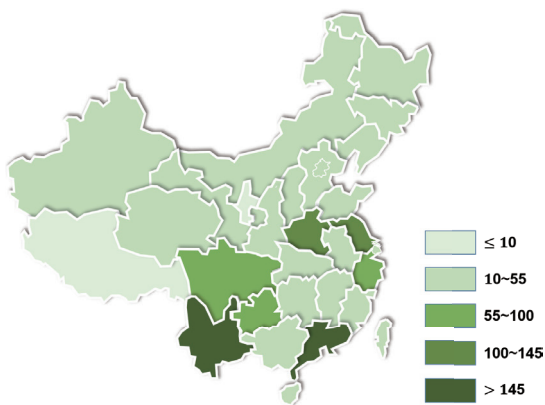
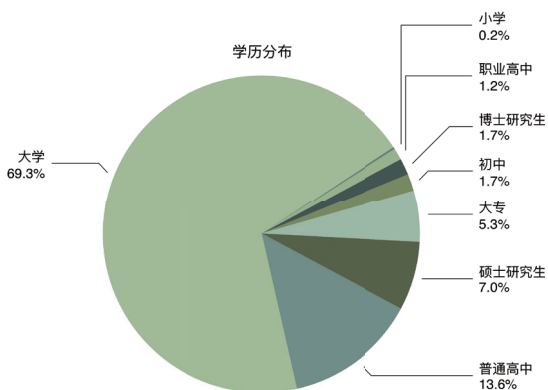
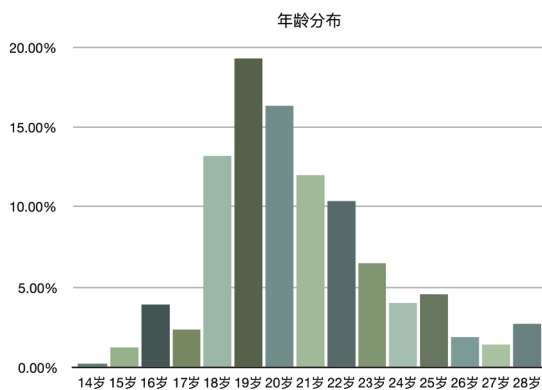
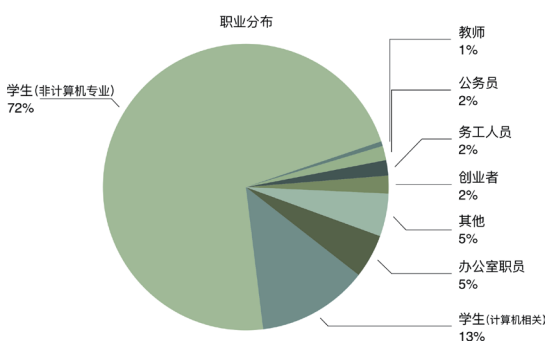
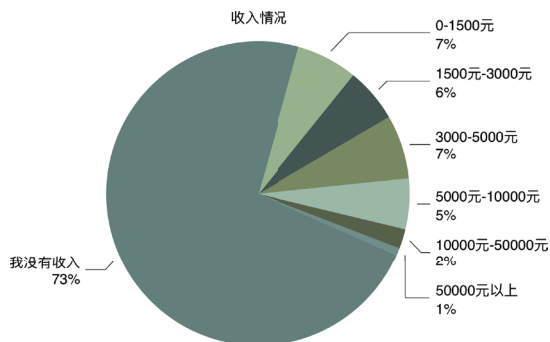
中国青年视角下的 AI 风险与治理



清华大学战略与安全研究中心

CENTER FOR
INTERNATIONAL SECURITY AND STRATEGY
TSINGHUA UNIVERSITY

本研究分为长问卷和短问卷，长问卷相比短问卷，多一道要求回答者对未来人工智能技术发展说一句话的题目。通过“番茄表单”发放，收回有效问卷总计 1491 份。其中，长问卷 738 份，短问卷 753 份。问卷的职业、学历、收入、年龄和地理分布的情况如下：



编者的话

在过去的一个月时间里,来自不同省份、不同民族,有着不同学科背景和基础的我们组合在一起,尽力完成了这一项预调研工作。在过程中,我们有两位同学经历了高考,也恰逢期末考试季。在一翻手忙脚乱后,我们从问卷设计、到第一波数据出路、再到数据分析不断深入。从无到有,看到一个个有趣的结论出现在了我们的面前,勾勒出了一副清晰度不高、分辨率不够、分析还很稚嫩“中国青年眼中的人工智能、风险和治理”的画像;虽然受限于学识、能力和时间,但这写初步的结果却激发了我们的兴趣。

在这个过程中,我们非常感谢清华大学战略与安全研究中心的支持和帮助。也感谢一路走来帮我们吆喝着发问卷的小伙伴们。我们也特别期待有老师和同学,能对我们幼稚而初步的结论进行批评和指导,希望有一天,我们能完整而系统的对全球青年眼中的人工智能、风险和治理进行调查。梦想很远,我们仍在路上。

主 编：

杨晶晶 北京大学外国语言与外国历史专业

主 笔：

周子超 吉林省通化一中

蒋建平 清华大学电子系

虞 宙 清华大学公共管理学院

调研与分析：

崔竞时 清华大学交叉信息院

敖腾隆 清华大学电子系

周子超 吉林省通化一中

编写团队：

袁令仪 中国人民大学附属中学

李欣宸 中国人民大学附属中学

智正元 内蒙古师范大学附属中学

刘 莹 北京市第十二中学

王南舜 北京大学附属中学

王元辰 清华大学环境学院

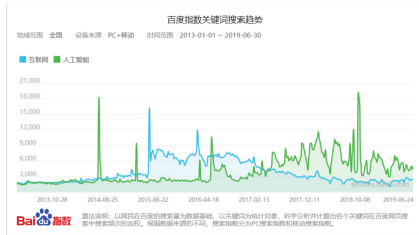
指导老师：

于 洋 清华大学交叉信息研究院

预调研报告：中国青年视角下的 AI 风险与治理

零、引言

从 2012 年神经网络在图像识别领域取得突破性的成功起^[1]，以深度学习为代表的的人工智能技术（AI）在互联网之后引领了新的一波技术浪潮。近年来，AI 在语音识别、机器翻译、人脸识别、自动驾驶等领域大放光彩，受到了人们广泛的关注。作为新兴经济体，也是人工智能技术产业化的热土，人工智能技术在中国引起了广泛的关注。对比的看：谷歌搜索趋势显示，人工智能技术在全球范围内的影响力仍不能匹敌互联网技术；然而，百度搜索趋势却显示出中国网民对人工智能的关注热度已经超过了互联网技术。



人工智能带来技术革命的同时，也带来了新的风险和治理困难。由于人工智能将人决策中的自由意志赋予机器，原有的制度、伦理与法律等在面对人工智能技术应用的管理、规制和治理上，遇到了前所未有的挑战。^{[2] [3] [4] [5]}

2016 年 1 月 20 日，一辆特斯拉轿车在京港澳高速邯郸段行驶时，与前方的道路清扫车发生追尾事故身亡，这是自动驾驶模式下致死的第一起案件，关于自动驾驶中的风险与追责问题一时成为焦点。^[6]2016

年剑桥分析公司 (Cambridge Analytica) 从 Facebook 上获取了数千万人的数据,并在当年美国总统大选中利用这些数据投放政治广告,这引发了一场“技术风暴”。^[7] AI 已不仅仅是学界的研究课题,它已经深入影响到政治、经济、工业、文化与个人生活领域。如何协调 AI 技术的发展与治理,保证 AI 技术安全可靠,促进社会、经济与生态可持续发展,这应该成为一个社会各界共同思考和努力的问题,而青年人作为国家未来的主人,也应当对 AI 所带来的机遇与风险保持关注。

人工智能的发展始终伴随着剧烈的社会变革,在为人类社会带来便利的同时,作为同一枚硬币的两面,人工智能带来的风险也逐渐引起世界各地不同群体的重视。在现有的针对人工智能的研究中,关注着人工智能在政府管理和社会治理中的应用与挑战,关注着在信息化时代中人工智能如何在医疗、教育和社交中发挥作用,关注着人工智能对女性和少数群体地位的改善,同时也注意到传统冲突论视角下既得利益者对规则的把控以及与公众的博弈等内容。

然而,作为人工智能产业化走在世界前列的国家,中国民众、特别是青年一代对人工智能及其风险治理的看法,尚未被分析和刻画。作为将长期身处人工智能技术变革带来的社会制度变迁中的一代,今天的中国青年如何看待人工智能的发展?他们担忧那些人工智能技术可能带来的风险?青年们又如何思考人工智能技术和产业的治理?这一系列攸关政策方向和设计的关键性问题亟待回答和分析。

然而,由于缺乏对中国青年对人工智能及其风险和治理看法的基础信息,因此本研究的第一步,是通过预调研为进一步大规模、系统性调研的问卷设计、采样设计和数据分析提供参考和支撑。

因此,本报告以中国社会中 14-28 周岁的青年人作为研究对象,采用调研问卷的方式搜集到中国不同地区的研究数据,以青年人的性别、年龄、受教育情况、职业、收入和地域等作为参考,就青年人对人工智能在日常生活中的应用、在不同层面可能存在的风险和如何对

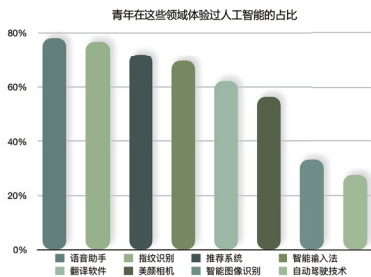
人工智能可能的风险进行防范和治理进行预调研,并基于预调研数据进行了初步研究。本报告试图呈现在日新月异的中国社会中,身处剧烈变动中的中国青年如何就人工智能可能带来的风险作出反应,而他们在人工智能的治理中又有哪些设想。

本报告主体分为五个部分,第一部分展示了中国青年对人工智能影响的评价和看法;第二部分则聚焦于中国青年关切的人工智能风险;第三部分讨论了中国青年对人工智能技术应该谁来治理、如何治理上的思考;第四部分对比了中国青年和国际学术界对人工智能治理问题的关切差异;第五部分总结了整个调研。

一、中国青年对人工智能的总体评价

1.1、人工智能已经深入和便利了中国青年的生活

从搜集到的有效问卷来看,人工智能的发展带给中国青年最直观的感受能够用一个词形容,那就是“方便”。这和人工智能已经深入青年人生活的方方面面是分不开的。从调查看在智能手机不断普及的背景下,他们对语音助手(78.05%)、指纹识别(76.82%)、推荐系统(72.13%)、智能输入法(69.54%)、翻译软件(62.15%)和美颜相机(56.29%)都有过或多或少的接触和体验。因此,当本研究请被调查者用三个字形容他们对人工智能的印象时,出现最多的词汇时“方便”(图一)。

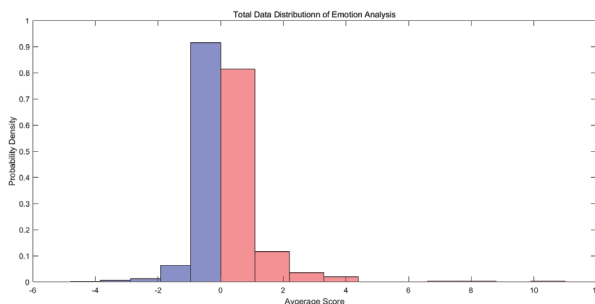


图一、青年日常生活接触的人工智能以及对人工智能的总体印象

但我们也注意到，青年们对人工智能的总体影响集中在技术、功能等客观描述性词汇上，较少使用具有情感色彩和价值判断的词汇。

1.2、中国青年对人工智能技术的关注高、情感正面（中美对比、情感分数）

为分析青年人对人工智能的情绪和观感，本研究请被调查的青年人寄语人工智能的发展。在针对青年人这段话的情感分析当中，我们发现积极情感分数高于消极情感分数，且分布更为均匀（图二）。这说明，对人工智能的发展，青少年的积极态度高于消极态度。人工智能本身蕴含的巨大潜力也使得这一阶段中国青年对它的发展持较为积极的态度。



图二、青年对人工智能的情感分数分布

通过交叉分析，我们发现：

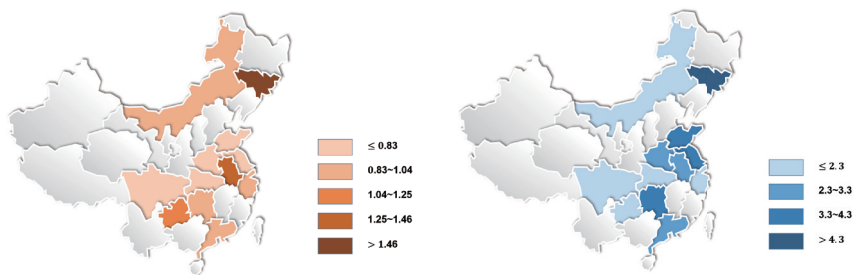
1、不同年龄段的青年人就人工智能有着各自的特点，进入大学前后的年轻人（17-19 周岁）对人工智能表现出更为积极的态度，而在接受大学教育之后的年轻人（20-22 周岁）则更多的看到人工智能的消极方面；

2、就中国不同的地域分布来看，青年人对人工智能在情感上的积极或消极并无明显规律，这可能与问卷的发布也存在一定的关联；而不同的性别在当今的中国社会中并未成为青年人对人工智能感受的

障碍，两性在这一方面并没有明显的区别；

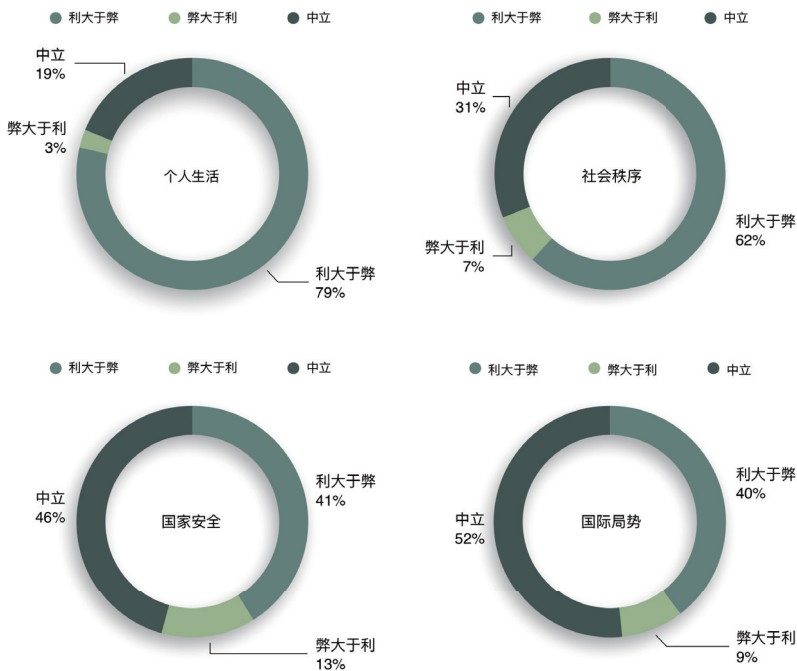
3、就青年人不同的职业选择来看，大学生作为教育程度最高的群体，始终是对人工智能的发展最具忧患意识的群体，他们在消极和积极态度中都占据主导地位。

4、同时。在调查的省份中，吉林省、贵州省、安徽省等的积极情感分（左图）和消极情感分（右图）都相对较高（图四）。地理分布的差异是否和该地区产业结构有关——例如贵州是中国云产业、数据产业发展快速的地区——还有待进一步的分析和调研。



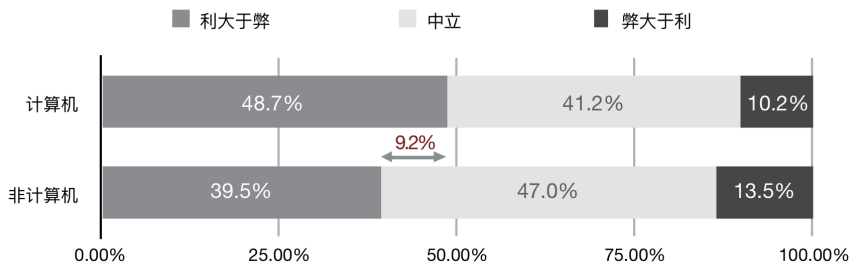
1.3、中国青年对人工智能的态度：微观积极 vs 宏观保守

问卷中调查了青年如何看待人工智能对“个人生活”、“社会秩序”、“国家安全”和“国际局势”四个不同尺度问题的影响。调查发现，80%的青年认为人工智能技术对个人生活的影响是利大于弊；然而，认为对国家安全、国际局势利大于弊的人群比例则只有40%（图三）。针对人工智能在不同维度的利弊影响，在个人生活（79%）和社会秩序（62%）两方面大多数人都持利大于弊的看法，但是当上升到“国家安全”和“国际局势”这样较为宏观的、个人难以把控的维度时，青年人的态度是逐渐趋于保守的，选择持中立看法的年轻人从19%（“个人生活”）和31%（“社会秩序”）上升到46%（“国家安全”）和52%（“国际局势”）。



图三、青年对人工智能对个人生活、社会秩序、国家安全、国际局势影响的看法

在问卷投放的过程中，尤其对计算机专业和非计算机专业的学生对这些维度对看法做了区分，研究中涉及到的计算机专业的青年在人工智能对国家安全看法上，表现出了比非计算机专业的青年更为乐观的态度。可能的解释包括：人工智能和相关的其他信息与计算科学领域的发展，使得计算科学相关专业的青年人更相信人工智能技术负面影响可控。



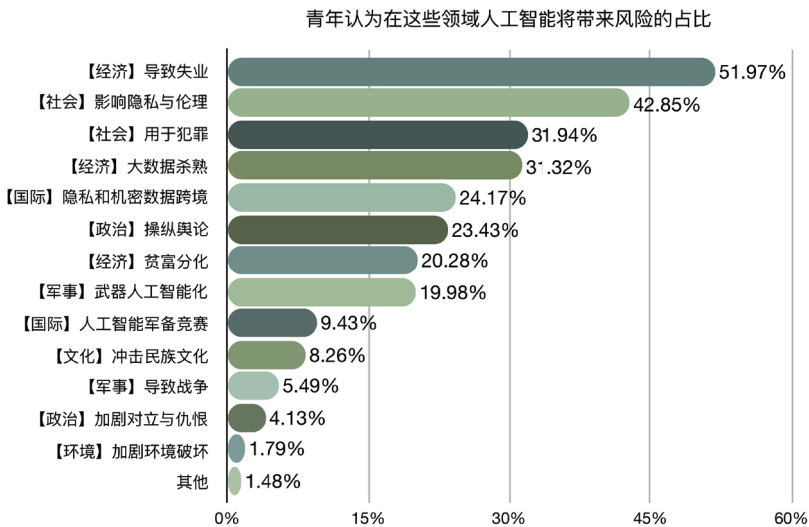
图四、计算机专业相关青年和非计算机专业相关青年对人工智能影响看法的差异

二、中国青年最关切的人工智能风险和治理问题

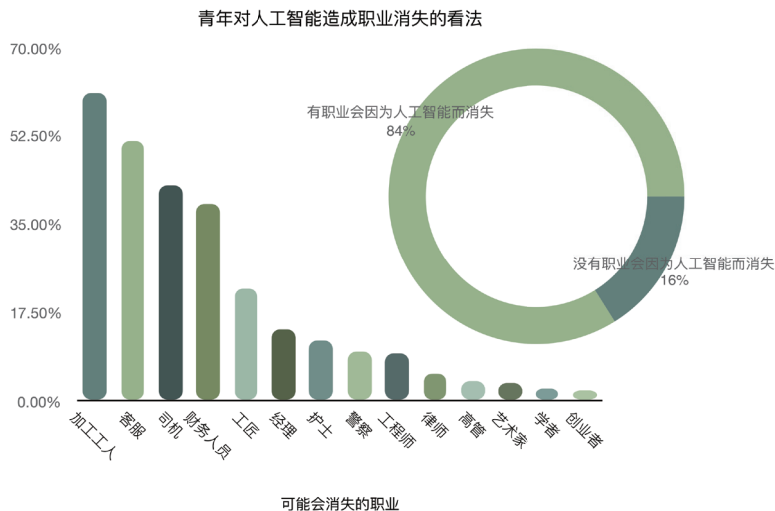
2.1、中国青年更关切人工智能的经济和社会风险

那中国的青年人是否看到人工智能在方便之外，对社会生活的各个领域可能产生的风险呢？他们最为关注的又是哪些领域呢？从搜集到的数据来看，**中国大多数青年人更关注的是人工智能在经济和社会等领域的影响**，尤其是人工智能的发展可能导致失业（51.97%）、影响隐私和伦理（42.85%）和被用于犯罪（31.94%）。

中国的青年人关注的是与自身发展息息相关的领域，尤其是未来社会中的就业问题。在关于人工智能的发展对未来职业的冲击中，工厂加工工人（61.10%）、客服/电话销售员（51.48%）和司机/邮递员/快递员（42.85%）等职业成为青年人眼中的“高危职业”，很可能会伴随着人工智能的发展而消失。中国青年的职业焦虑由于人工智能的发展而逐渐被转移，绝大多数人（84%）都认为人工智能的发展会引起社会上部分职业的消失。



图四、中国青年最关心的人工智能风险问题



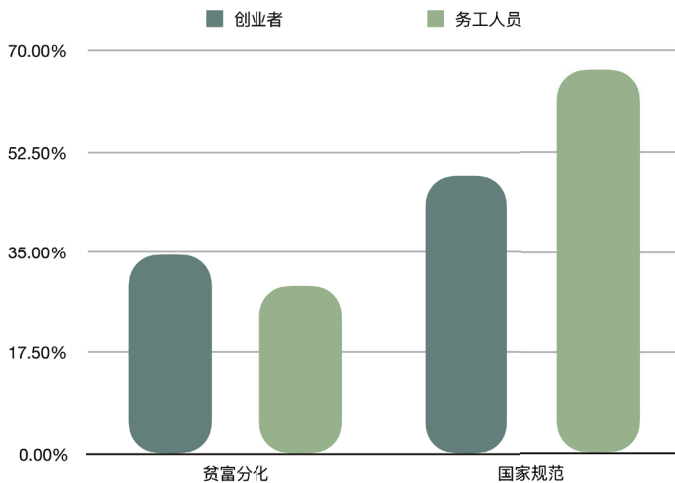
图五、中国青年认为会因人工智能技术而消失的职业

可以看到，样本中的青年，对人工智能的隐私和伦理风险也极为关注。同时，也有相当高比例的青年担忧人工智能用于犯罪的风险。不少青年对人工智能用于操纵市场等行为也保持着较高的警惕。

作为青年人中除学生外、和人工智能技术密切相关的、有代表性的两个群体：创业者（29人）和务工人员（24人）对人工智能的风险和治理又是什么样的反应呢？虽然我们的样本中，这两个群体的人数较少。但是对比这些仅有的样本，仍然一定程度上提醒我们，职业分工的不同，会造成人们对人工智能风险和治理的看法差异。针对人工智能可能带来的风险，超过半数的务工人员和创业者都对人工智能可能导致失业表示担忧；而创业者作为与现代信息技术密切接触的群体，则较务工人员有更高的比例担忧贫富分化。然而不论职业如何，被调查者对可能受人工智能发展而消失的职业有相同的看法，他们与总体的数据样本也保持着一致的观点：劳动力密集型和重复性工作更容易受到影响。在针对人工智能的治理方面，务工人员（66.67%）比创业者（48.24%）更寄希望于国家和政府通过法律进行严格的

规范，这在一定程度上说明传统行业在应对人工智能可能带来的风险是是不自信的，更希望能够在国家和政府的庇护下以实现平稳的过渡。

但是，需要指出的是：我们的样本中，创业者和务工人员在地域上分布不平衡，分散在中国不同的省份；与务工人员大多具有大学和大专的文凭相比，创业者在学历分布上更具包容性，几乎涵盖了从初中到博士研究生的所有群体；但女性在创业者中的占比是低于务工人员的，同时与具有稳定月收入的务工人员相比，在创业者中部分人是没有收入的。他们在人工智能在日常生活的体验中，智能手机的普及使得他们获得的服务不存在明显的差别。上述差别，使得我们关于创业者和务工人员对比分析的结果只适用于描述被调查到群体。



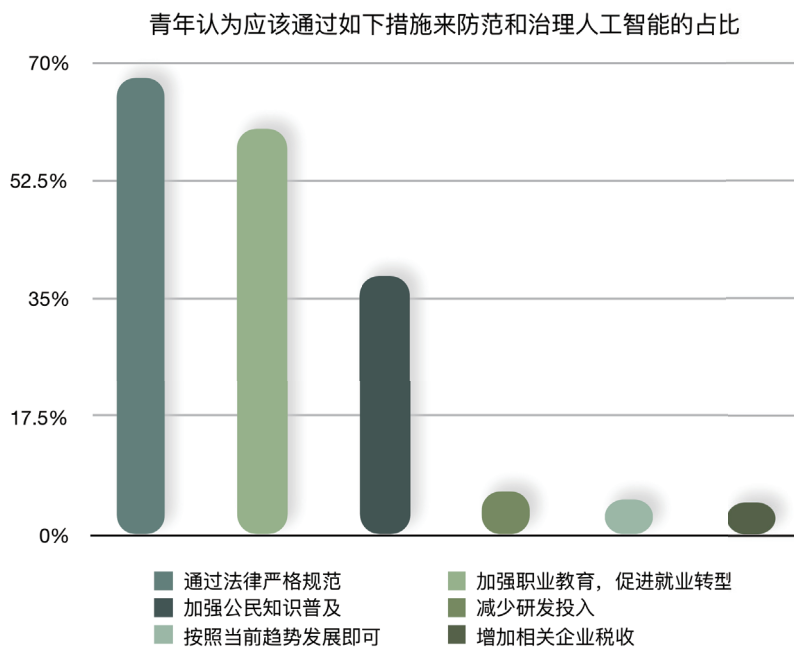
图六、创业者群体和务工人员群体在人工智能风险问题上的关切差异

2.2、中国青年忧虑的政治和国际性人工智能治理议题：数据跨界流动和舆论操纵

而在政治和国际领域，中国青年最担心的是数据跨界流动（24.17%）和舆论操纵（23.43%），这与谷歌学术指数（绝对值）有

一定的相似性,舆论操纵在后者的指数为19.92%,居于第二位。一方面,由于精准的算法使得对网络媒体资讯的筛选成本更低,无论在全球范围还是在青年群体中,对于舆论操纵的担忧都是一股不可忽视的社会情绪,在人工智能迅速发展的当下如何合理规范与界定其职权,如何提升公民对于信息源的安全感正是亟待解决的问题,另一方面,中国青年对于数据与舆论的关注也说明,除却实际的经济红利,各种隐性的精神与情感需求也是其考量人工智能这一新事物的重要指标,如何权衡与调和两者,缓解技术革命进程中的矛盾至关重要。

三、中国青年认可的人工智能治理： 法制化和政府的积极作为



图七、青年认为应该采取的人工智能治理措施

而如何来应对和防范人工智能可能带来的变化，中国的青年人显然也有自己的看法，多数人（67.82%）将希望寄托于本国的政府，渴望国家和政府能够通过法律的出台来严格规范人工智能的发展；青年人对人工智能发展可能导致的失业也在治理的措施中得以体现，60.11% 的接受问卷调查的青年人将加强职业教育，促进就业的转型作为降低人工智能带来的风险的重要手段；而通过较少研发经费和增加相关企业税收等经济学的调节方式却得不到青年人的青睐，市场“看不见的手”在人工智能风险的防范中在中国青年的认知中，远不如国家和政府这只“看得见的手”来得可靠。

四、国际学界 vs 中国青年：关切的共性和差异

通过对比调研结果和谷歌学术指数（绝对值）可以看到：此次调查的我国青年与学术界关切的人工智能风险虽然有一定共性，但也存在着许多差异。

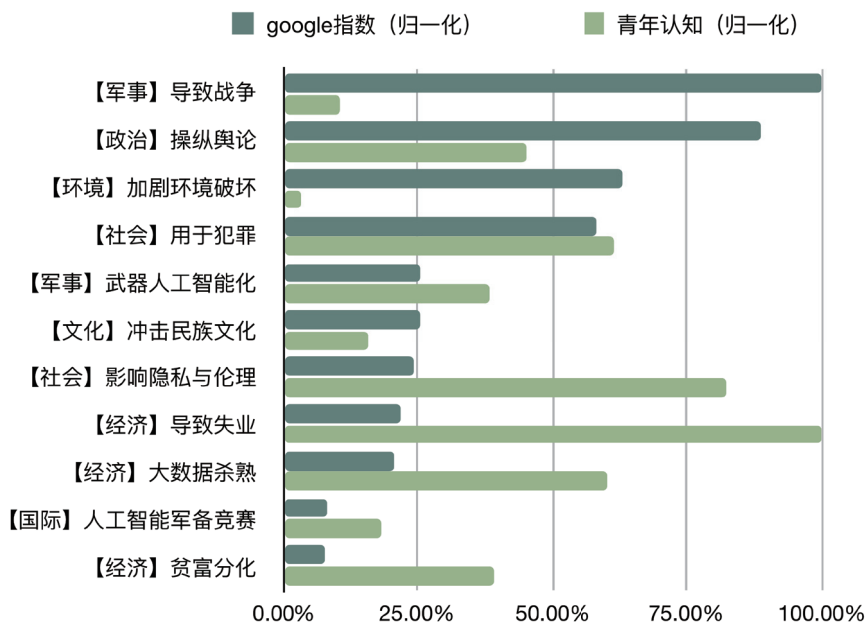
关切的共性方面：不可否认无论是国际学界的顶尖学者，还是中国各年龄阶段，不同教育背景的青年对于人工智能都有着—个明显的共识，那就是人工智能的出现与发展势必将对于我们所处的时代产生深远的影响。这由此也说明了研究 AI 治理以及可能带来的风险是有必要的，也是与未来青年人发展的目标是吻合的。

关切的差异方面：通过对于国际学界的文献总结以及中国青年的问卷调查，我们也不难发现，两类群体所聚焦的问题还是存在着—定的差异，主要表现在以下方面：

谈及受 AI 技术影响的群体，国际学界所聚焦的更为具体，而中国青年所聚焦的影响群体更为宏观。在国际学者眼中，性别，肤色都有可能是造成不同潜在影响的因素，因此对于某些群体的关切程度也会有所差异。^{[8] [9]}而在中国青年的眼中，更倾向于以宏观整体的思

维去看待问题，研究分析 AI 技术对整个人类可能造成的影响。

谈及受 AI 技术影响的领域，国际学者往往更关注国际政治、环境等领域，而中国青年则聚焦于社会、生活等方面。从文献中我们不难发现，国际学界对于 AI 科技前沿的研究，往往将目光聚焦于国际政治、环境等领域，无论是总统竞选、还是无人驾驶技术，在学者的眼中，AI 可能带来的风险都是不可忽视的一大因素。从数据上来看，学者们关注人工智能的发展可能导致战争（22.48%）、社会舆论被操纵（19.92%）和加剧环境的破坏（14.16%）等国际政治和环境问题。而中国青年则更关注 AI 技术对于生活和社会造成的影响，包括失业、伦理、犯罪等方面。



图八、学术研究焦点和青年关注焦点的差异

谈及对 AI 技术的态度，国际学者更为消极，倾向于扩大 AI 的负面影响，而中国青年则相对更为乐观（虽然随着维度的扩大，态度逐渐趋向保守）。在国际的视野下，一个由 AI 引发的事故都会被

无限放大，由此带来社会对于 AI 技术本身的质疑。然而在中国的情境下，在生活中和人工智能技术市场接触的中国青年则表现的更为乐观；他们能清楚地看到 AI 技术给生活所带来的各种便捷，同时面对可能带来的风险，也表现出了一种积极的态度。而当评价的维度逐渐扩大时，中国青年的态度则趋向于保守。

五、总结和展望

习近平总书记 2017 年 5 月 3 日在中国政法大学考察时的讲话中强调，“中国的未来属于青年，中华民族的未来也属于青年。”2018 年 5 月 2 日在北京大学师生座谈会上的讲话里也提到，“每一代青年都有自己的际遇和机缘。”新时代的青年在人工智能不断发展和普及的背景下，更应该把握好自身的机遇和机缘，在人工智能提供的便利之外，更应该注意到其中可能伴随着的风险，特别是对个人生活和社会秩序可能产生的消极影响。因此，青年人对与自身息息相关的经济和社会生活的关注也就是合理的，同时青年人承载着中国的未来和中华民族的未来这两重的担当，更应当在培养自身过硬本领的同时，与国家和政府一道，共同筑起抵御人工智能可能的威胁的防洪堤，为人工智能的治理贡献力量！

2015 年 9 月 25 日，联合国可持续峰会在纽约总部召开，联合国 193 个成员国在峰会上正式通过了 17 个可持续发展目标，旨在解决从 2015 年到 2030 年间以综合方式彻底解决社会、经济和环境三个维度的发展问题，最终实现可持续的发展道路，而人工智能的日益勃兴必然会为可持续发展目标的实现提供助力。人工智能通过对海量数据的处理和分析，有利于确保健康的生活方式，促进各年龄段人群的福祉；而实现包容和公平的优质教育在全球服务内的共享，也离不开人工智能在教育实践中提供多样化的手段和发挥整合的功能。同时，即

使中国的青年在人工智能中对职业表示明显的担忧，但我们也应注意到人工智能对创新性产业的革新作用，由此实现全球范围内持久、包容和可持续的经济增长。全球气候的变化是跨越国界的每一位公民都要面对的挑战，而人工智能的不断发展必将为全球气候治理提供更多的可选择的方案，加深我们对气候和地球系统的理解；在水生生物和陆生生物的保护方面，人工智能在全球观测和预警机制的建立方面也将发挥着不可替代的作用。综上，人工智能在全球范围内的延伸和深入，有利于重振全球治理中各国之间的可持续发展伙伴关系，最终在全球范围内实现正义和和平！

参考文献:

- [1] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).
- [2] Margetts, H., & Dorobantu, C. (2019). Rethink government with AI. *Nature* 568, 163–165.
- [3] Sherlock, C. (2017). The DeepMind debacle demands dialogue on data. *Nature* 547, 259.
- [4] Crawford, K., & Calo, R. (2016). There is a blind spot in AI research. *Nature* 538, 311–313.
- [5] Pandya, J., & WORLD, C. (2019). The Dual-Use Dilemma Of Artificial Intelligence. *Forbes*.
- [6] 国内首起“特斯拉自动驾驶”车祸致死案已有最新进展 [N].<http://news.cctv.com/2018/04/19/ARTILMtSgahcm3W5j3bx87as180419.shtml>
- [7] Davies, Harry (December 11, 2015). “Ted Cruz campaign using firm that harvested data on millions of unwitting Facebook users” . the Guardian. Archived from the original on February 16, 2016.
- [8] Zou, J., & Schiebinger, L. (2018). AI can be sexist and racist — it’ s time to make it fair. *Nature* 559, 324–326.
- [9] Courtland, R. (2018). Bias detectives: the researchers striving to make algorithms fair. *Nature* 558, 357–360.

清华大学战略与安全研究中心

办公地点：清华大学明斋 217

联系电话：010-62771388

电子邮箱：ciss_thu@163.com