

ARTIFICIAL INTELLIGENCE AND LIFE IN 2030

ONE HUNDRED YEAR STUDY ON ARTIFICIAL INTELLIGENCE | REPORT OF THE 2015 STUDY PANEL | SEPTEMBER 2016

本篇报告由机器之心独家编译，供大家学习所用。
欢迎关注机器之心公众号。



斯坦福人工智能百年研究：2030年的人工智能与人类生活

2014 年秋季，人工智能百年研究（One Hundred Year Study）项目启动，这是一项对人工智能领域及其对人类、社区、社会影响的长期学术研究。这项研究包含使用人工智能计算系统的科学、工程和应用实现。监督该「百年研究」的常务委员会（Standing Committee）组建了一个研究小组（Study Panel）来每五年评估一次人工智能所处的状态——这是本项目的核心活动。

本研究小组要回顾从上次报告到现在这段时间人工智能的进展，展望未来潜在的进展并且描述这些进展对于技术、社会的挑战与机遇，涉及的领域包括：道德伦理、经济以及与人类认知兼容的系统设计等等。

「百年研究」定期进行专家回顾的首要目标是：提供一个随着人工智能领域发展的关于人工智能及其影响的收集性的和连通的集合。这些研究希望能在人工智能领域的研究、发展以及系统设计方面、以及在帮助确保那些系统能广泛地有益于个人和社会的项目与政策上提供专业推断上的方向指南及综合评估。

「百年研究」是仿照一个早期的被非正式地称为「AAAI Asilomar Study」的项目。在 2008 到 2009 年期间，当时的美国人工智能协会（AAAI）的主席 Eric Horvitz 聚集了一群来自这个领域不同机构与方向的人工智能专家以及认知科学、哲学与法律相关的学者。

参与者在分布的小组中研究近期人工智能的发展、长期的发展前途以及法律与道德上的影响，之后他们在 Asilomar 一起开了三天的会议，分享并讨论了他们的发现。会议给出了其中密集讨论的内容的一份短的书面报告，并且由参与者与其他同事在后续的讨论中充实；这份报告在领域内以及之外引起了广泛的关注与争论。

Asilomar 会议的巨大影响以及人工智能的重要进展（包括人工智能算法与技术在全球范围内开始进入到日常生活中）刺激产生了对人工智能及其对人们与社会的影响进行长期反复研究的想法。这个百年研究项目随后在一所大学中开始，这是为了延伸深度思考以及跨学科的学术研究，激励创新并且给政府机构与产业提供明智的建议。

这篇报告是计划持续至少 100 年的研究系列中的第一篇。常务委员会在 2015 年的暑期成立了一个研究小组来负责组建现在这个初始的研究小组，并任命了得克萨斯大学奥斯汀分校的教授 Peter Stone 担任该小组的主席。这个包含了 17 名成员的研究小组由人工智能学术界、公司实验室以及产业界的专家与了解人工智能的法律、政治科学、政治以及经济方面的学者组成，并于 2015 年秋季中期启动。

参与者代表着不同的专业、地区、性别以及职业阶段。常务委员会广泛讨论了 Study Panel 相应的责任，包括人工智能最近的发展与在工作、环境、运输、公共安全、医疗、社区参与以及政府的潜在社会影响。委员会考虑多种聚焦研究的方式，包括调查子领域及其状态、研究特定的技术（例如机器学习与自然语言处理）以及研究特定的应用领域（例如医疗与运输运输）。

委员会最终选择了「2030 年的人工智能与生活（AI and Life in 2030）」为主题以强调人工智能的各种用途与影响的发生不是独立于彼此，也不独立于其他许多社会和技术上的发展。意识到了城市在大多数人类生活中的核心作用之后，我们将专注重点缩小到大多数人居的大都市。

常务委员会将进一步将专注重点缩小到具有很好可变性的城市环境以及世界性文化的典型的北美城市，从而限制研究小组的工作量。常务委员会期望该研究中的预测、评估和预先的指导将具有广泛的全球性的实质作用；另外，委员会正在制定将这个项目扩展到国际上的未来研究计划。

既然做出了关注北美城市生活的决定，那么军事项目就不在这篇最初的报告范围中。这不是忽视细致监控与审议人工智能在国防与战争（包括潜在地使发展和部署不平衡）上产生的影响的重要性。这篇报告计划展现给四种读者。

对于公众，本报告的目的在于提供一个可获取的、在科学和技术上准确的对于人工智能当前状态及其潜力的描述。对于产业界，本报告描述了相关技术、法律以及道德上的挑战，可能帮助指引资源再分配。本报告也可以帮助地区、国家、国际政府更好地规划人工智能在政府中的应用。最后，本报告能帮助人工智能的研究者以及他们的机构和资助者设置优先级，以及考虑人工智能研究及其应用所产生的道德与法律上的争议。

考虑到「百年研究」在人工智能上的独特本质，我们期望未来的常务委员会与研究小组，以及研究科学家、政策专家、私人与公共部门的领导以及普通民众，在对人工智能的未来进行评估时能想到该报告。不管是否成功准确预测了人工智能的发展轨迹和影响，我们都希望这个系列中这一份初次成果是有用的。

STUDY PANEL

Peter Stone, University of Texas at Austin, Chair
Rodney Brooks, Rethink Robotics
Erik Brynjolfsson, Massachusetts Institute of Technology
Ryan Calo, University of Washington
Oren Etzioni, Allen Institute for AI
Greg Hager, Johns Hopkins University
Julia Hirschberg, Columbia University
Shivaram Kalyanakrishnan, Indian Institute of Technology Bombay
Ece Kamar, Microsoft Research
Sarit Kraus, Bar Ilan University
Kevin Leyton-Brown, University of British Columbia
David Parkes, Harvard University
William Press, University of Texas at Austin
AnnaLee (Anno) Saxenian, University of California, Berkeley
Julie Shah, Massachusetts Institute of Technology
Milind Tambe, University of Southern California
Astro Teller, X

目 录

概述	6
人工智能接下来的研究？	9
现在与未来的人工智能政策	10
第一部分：什么是人工智能？	11
定义人工智能	11
人工智能研究趋势	13
第二部分：人工智能在各领域的应用	17
交通	17
家庭/服务机器人	21
真空吸尘器	21
医疗	22
教育	26
低资源社区	29
就业和劳资	31
娱乐	33
第三部分：人工智能公共政策的前景与建议	34
如今以及未来的人工智能政策	34
附录 I：人工智能简史	41

概述

电影和小说中描绘的人工智能主宰的可怕未来塑造了大众对人工智能的想象，但这些都是虚构的。在现实中，人工智能已经在改变我们的日常生活了，而且基本上都是在改善人类健康、安全和提升生产力等好的方面。

和电影不一样，现在并没有什么超人类的机器人正待到来，这甚至完全不可能。而且尽管滥用人工智能技术的可能性应该得到承认和解决，但其中它们更大的可能性是让驾驶更安全、帮助孩子学习和扩展及增强人类的生活。

事实上，有益的人工智能应用已经在学校、家庭 and 医院里开始了加速增长。大多数研究型大学也划出了单独的人工智能研究科系，苹果、Facebook、谷歌、IBM 和微软等科技公司也在探索人工智能应用上投入了巨资——他们将其视为未来发展的关键。即使好莱坞在将他们的反乌托邦人工智能幻想故事搬上荧幕时也用到人工智能技术。

基于计算机的视觉、语音识别和自然语言处理方面的创新是这些变化的驱动力，同时也将推动相关领域的科学和技术的发展。人工智能也正在改变人们与技术的交互方式。许多人已经习惯上了在他们的智能手机上进行触控和对话。

随着人工智能系统学习对个人特点和目标的适应，人机关系的未来将会变得越来越微妙、流畅和个性化。这些人工智能应用将帮助监控人们的生活状态、警告人们前面的风险以及提供人们想要的或需要的服务。

比如，在北美一座城市的未来 15 年里——正是这份报告的时间框架和范围——人工智能应用将很有可能将交通交由自动驾驶汽车接管，实现人与货物的实时接取和运送。光这一项就足以改变城市的整体面貌，因为再也不会再有交通堵塞和停车困难的难题了。

本研究刻意重点关注了典型的北美城市，旨在突出人工智能对城市中居住的数百万居民的日常生活的影响。研究小组还进一步将调查范围缩小到了 8 个人工智能可能将带来最大影响的领域：交通（transportation）、医疗（healthcare）、教育（education）、低资源社区（low-resource communities）、公共安全（public safety and security）、就业和工作场所（employment and workplace）、家庭/服务机器人（home/service robots）和娱乐（entertainment）。

尽管人工智能技术来自于常见的研究资源，但这些技术已经影响并将继续影响这些不同的领域。其中每个领域都面临着各种不同的与人工智能相关的难题，包括难以创造安全可靠的感知和作业硬件（交通和服务机器人）、难以获得公众的信任（低资源社区和公共安全）、难以克服对人类被边缘化的恐惧（就业和工作场所）、以及还存在减少人际交往的风险（娱乐）。

其中一些领域主要是在商业方面，如交通和医疗，而另一些则更面向消费者，例如娱乐和家庭服务机器人。一些还是跨领域的，例如就业/工作场所和低资源社区。

在每一个领域，即使人工智能在不断带来重要的好处，但同时也带来了道德和社会上的问题，其中包括隐私问题。机器人和其它人工智能已经在一些机构里面替代人类工作了。作为一个社会，我们现在正处在一个确定如何部署基于人工智能的技术的紧要关头，从而提振而不是阻碍民主的价值观——例如自由、平等和透明。

对于个人而言，我们生活的质量和我们的贡献所获得的价值很有可能会逐渐发生转变，但这些改变会很明显。在接下来的几年中，人工智能研究、系统开发和社会与监管框架将会塑造人工智能的好处与其代价和风险之间的权衡，以及这些好处如何散布的方式。

人工智能的准确和精细的未来图景——一个与其流行的描述相竞争的图景——一开始就受到了难以精确定义「人工智能」这一问题的阻碍。在研究小组所考虑过的所有定义方法中，没有一个认为现在存在某种「通用的」人工智能。

尽管使用了共同的研究和技术，但人工智能系统都专注于完成特定的任务，而且其中每一种应用都需要数年的专注研究和精心独特的构建。这也造成了这 8 个领域内进展的不平衡。

交通是其中一个典型的例子——其中一些关键技术让人工智能以惊人的速度得到了广泛的采用。自动化交通很快就将无处不在，并将成为大多数人与嵌入了人工智能系统的物理实体的第一次经历，并将极大地影响公众对人工智能的感知。随着汽车自己变成了比人类更好的司机，城市居民的私家车将会更少，也将住得离工作场地更远，花费时间的方式也将发生改变，从而带来一种全新的城市组织形式。

在 2030 年一座典型的北美城市，嵌入了人工智能应用的物理实体将不仅限于客车，还将包含货车、飞行器和个人机器人。安全和可靠硬件的改进将会刺激未来十五年内的创新，也将带来更好的家庭/服务机器人——它们已经进入了人们的家庭，目前主要是以真空吸尘器的形式。

更好的芯片、低成本 3D 传感器、基于云的机器学习和语音理解的进步将提升未来的机器人服务和它们与人类的交互。专用机器人将能够递送包裹、清洁办公室和增强安全性。但在可预见的未来里，技术限制和可靠的机械设备的高成本将继续将商业机会限制在定义狭窄的应用范围内。

在医疗保健方面，在收集用于个人监控设备和移动应用的有用数据上已经出现了巨大的飞跃，这些数据来自医疗机构的电子病历（EHR），以及小范围的用于辅助医疗手术的手术机器人和支持医院运营的服务机器人。未来几年，基于人工智能的应用将能够改善数百万人的健康状况和生活质量。

尽管临床应用从实验室到真实世界的过程一直都较为缓慢，但有希望的迹象表明创新的步伐将得到提升。通过共享数据的激励和机制以及去除无理的政策、法规和商业障碍，可以促进医疗保健领域的进步。对于许多应用，人工智能系统必须紧密地与医护人员和病人密切合作才能获得他们的信任。改善智能机器与医护人员、病人和病人家庭的自然交互方式是很关键的。

为人类和有前途的人工智能技术带来更流畅的交互同样也是教育领域的一大关键难题，这个领域在这段时间也出现了一些可观的进展。尽管素质教育总是需要人类教师的活跃参与，但人工智能有望改善所有层次的教育，尤其是通过大规模提供个性化教育。

交互式机器导师现在已经被用于根据匹配的学生教授科学、数学、语言和其他学科。自然语言处理、机器学习和众包也已经极大地促进了在线学习，让教育水平更高的教师可以将他们的教室扩大很多倍，同时还能应对单个学生的学习需求和风格。在接下来的十五年的典型北美城市里，这些技术在教室和家庭里的应用将很有可能得到极大的扩展，只要它们能够有意义地和面对面的学习进行整合。

除了教育方面，在协助低资源社区上，人工智能方法也存在很多机会——通过为各种社会问题提供缓解方案和解决方案。传统上，投资者对缺乏商业应用的人工智能研究投资不足。有了针对性的激励和资助重点，人工智能可以帮助解决低资源社区的需求，这方面一些新兴的工作是很有希望的。

比如，使用数据挖掘和机器学习，人工智能可以创造预测模型来帮助政府机关解决防止儿童铅中毒和高效食物分配等问题。这些萌芽中的工作认为还有更多的事情要做，尤其是当机构和组织也能参与进来并与这些社区建立起信任时。

获取公众信任也是公共安全领域内人工智能应用的一大难题。北美城市和联邦机构已经开始在边境管理和执法部署人工智能技术。到 2030 年，他们将在很大程度上依赖于它们，包括用于侦察的改进过的相机和无人机、用于检测金融诈骗的算法和预测性警务。

后者带来了对于无辜的人进行无理的监控的问题，我们必须小心谨慎不要给系统引入人为偏见，并保护公民自由。部署良好的人工智能预测工具有望给数据和推断带来新形式的透明，并可能被应用于检测、删除和减少人类偏见，而不是增强它。

在人工智能对就业和工作场所的影响上的社会和政治决策也是类似，比如需要安全网来保护人们免受经济结构变化的影响。人工智能有望在一些特定类型的工作中取代人类，比如驾驶汽车或卡车。但是，在许多领域，人工智能在短期内很有可能将取代很多任务，但不能取代工作岗位，而且还将创造新类型的工作。

但这些将会出现的新型工作目前还比将可能失去的工作更难预想。人工智能还将降低许多货物和服务的成本，有效地使每个人都过得更好。更长期来看，人工智能可能会被视为一种非常不同的财富创造机制，其中每个人都应该从全世界的人工智能所创造的财富中分一杯羹。对于人工智能技术的经济成果的分配方式，不久之后已经就应该会有相关的社会争论出现了。

娱乐已经被社交网络和其它用于共享和浏览博客、视频和照片的平台转变了；这些平台依赖于 NLP、信息检索、图像处理、众包和机器学习等领域内正被活跃开发的技术。一些传统的娱乐资源已经开始拥抱人工智能，包括谱曲、创作舞台表演、甚至根据自然语言文本生成 3D 场景。

人们对人工智能驱动的娱乐的热情让人惊叹。和人工智能的许多方面一样，在技术对社交的替代或增强方面也一直存在着争议。人工智能能让娱乐越来越交互式、越来越个性化、越来越有参与感。应该引导一些研究来理解如何利用这些性质为个人和社会利益服务。

人工智能接下来的研究？

助力人工智能革命的研究也在快速发展。其中最重要是机器学习的成熟，它受到了数字经济崛起的部分影响。数字经济和机器学习都提供并利用大量的数据。其他因素包括云计算资源的崛起，以及消费者对语音识别和导航支持这样的技术服务的需求。

机器学习由于人工神经网络的成功经验前进了一大步，如今能够使用大型数据集和大规模计算对其进行训练。这一方法也就是我们所熟知的深度学习。信息处理算法的性能也因为传感、感知、目标识别等这些基础操作的硬件技术的重大进步，而取得了跳跃性进展。

数据驱动产品的新型平台和市场以及经济激励发现新的产品和服务，也刺激着研究的发展，如今，随着它称为了社会的核心力量，人工智能领域转向了建立能与人高效合作的、具有更广泛人类意识的智能系统，包括开发人教机器人的交互与可延展方式的创造性方法。

这些趋势引发了如今人工智能基础方法与应用领域研究的热门：

大规模机器学习涉及到学习算法的设计，延展已有的算法，在极大型的数据集上做研究。

深度学习是一类学习方法，促进了在图像、视频标记和运动中的目标识别，并且在其他感知领域都有重大影响，比如音频、语音和自然语言处理。

强化学习是一个框架，能将机器学习的重点从模式识别转变到经验驱动的序列决策上。它有望推动人工智能应用的前进，在现实世界中做出决策。虽然在过去数十年，强化学习限于学术界，但如今我们也看到了一些实际的成果。

机器人目前主要涉及到如何训练机器人以一种更泛型的、预测性的方式与周围世界进行交互，如何在交互环境中促进对目标的操控，以及如何与人进行交互。机器人的进步将依靠相应的技术进步，从而改进计算机视觉和其他形式的机器感知的可靠性与普遍性。

计算机视觉是目前机器感知中最突出的形式。它是受到深度学习崛起影响最大的人工智能子领域，实现了史上第一次计算机能够比人类更好的完成视觉任务。如今更多的研究集中在图片和视频的自动化文字描述上。

自然语言处理通常与自动语音识别相结合，快速成为有大型数据集的普遍使用的口语语言上的一种商品。目前的研究转向了开发能够与人通过对话交互的精致、可用的系统，而不仅是对程式化的问答做回应。在多种不同语言间进行机器翻译也已经取得了重大进展，在不远的将来将有更多的实时的个人到个人交流。

协作系统研究主要是研究帮助开发能与其他系统或人类协作的自动化系统的模型和算法。众包和人类计算研究是研究增强计算系统的方法，增强方式是通过自动调用人类经验，解决计算机不能单独解决的问题。

算法博弈论（Algorithmic game theory）和**计算社会选择**（computational social choice）吸引了人工智能经济和社会计算领域的注意，比如系统如何处理潜在的不恰当激励机制，包括自私的人类参与者

或公司以及代表他们的自动化人工智能代理。

物联网（IoT）研究专注于将大量设备（包括家电、汽车、建筑、摄像头）连接起来，收集并共享各自丰富的传感信息从而达到智能目标。

神经形态计算是一系列寻求模拟生物神经网络从而改进计算机系统硬件效率和稳健性的技术，它取代了用独立的输入/输出、处理指令、存储模块的旧方法。

现在与未来的人工智能政策

人工智能应用是否成功的标准是它们为人类创造的价值。在此角度看，设计人工智能应用应该使得人们成功的理解人工智能系统，参与到其使用中并建立对它们的信任。公共政策应该帮助社会轻松采纳人工智能应用、扩展其收益，并缓和它们的难免的错误和失败。

对人工智能如何开展的争论包括对隐私保护和人工智能收益公平分配的忧虑，这种忧虑是该被鼓励的。在人工智能技术如此快的实现速度之下，也伴随着对其应用的担心。研究小组建议所有的政府律师掌握人工智能技术知识。

此外，要通过移除障碍并增加私营和公共支持，鼓励在人工智能系统公平性、安全性、隐私和社会化应用上的研究。目前在美国，至少有 16 个独立的经济相关的政府部门涉及到人工智能。人工智能研究，特别是其应用的快速进展，需要这些部门的专家开发新的法律、政策概念与隐喻（metaphor）。

当自动驾驶汽车发生车祸、医疗设备事故时谁该负责？如何避免人工智能应用宣传种族歧视或者金融诈骗？谁该得益于人工智能技术带来的效率？对技术过时的人群应该提供什么样的保护？随着人们在产业流程和消费产品中更广、更深的融入人工智能，更好的实践需要被铺开，也须要有更适合的政府制度。

研究小组没有考虑近期人工智能系统会自动选择对人造成伤害，可能的是人们使用基于人工智能的系统做好的以及坏的目的。而且尽管人工智能算法可能比人类做出一些较少偏见的决策，保证用于人工智能决策的数据没有偏见仍是一项技术挑战，这些偏见可能导致基于种族、性别以及其他因素的歧视。

面临人工智能技术产生的重大变革，逼迫生成更多、更严格的条例也可能使不可避免的。错误理解什么是人工智能可能刺激人们反对有益于每个人的这些技术。不合适的条例可能造成悲剧。信息贫乏的条例会扼杀创新，或者将其转移到其他司法权可能是达不到预期目标的。

幸运的是，领导产生如今数字技术成功条例的原则为我们提供了起点。在私营条例中，宽泛的法律规定配合强硬的透明度需求和有意义的强制措施（而非严格的控制）能够鼓励公司开发流程，鼓励职业人员实施隐私控制，与外部利益相关者结合，并将实践适应于技术发展。这反过来也能支持专业贸易协会和标准委员会在开展最好的实践上的发展。

在人工智能中，监控着也能加强内部和外部问责、透明度、专业度之间活动的良性循环。我们急需一

场重要的、有力的、信息丰富的讨论会，讨论如何以一种丰富我们生活与社会的方式最好的控制人工智能，同时鼓励该领域的创造性。

由于高计算能力和大规模数据是很多人工智能技术的燃料且不平等的分布于社会，这使得人工智能技术把机遇不平等性变得更宽。这些技术将促进有机会接触数据和计算能力的那群人的能力和效率。在政策是否培养民主价值观、公平共享人工智能收益，或者是否将力量和收益聚集于少数幸运的人手中，我们要对政策进行评估。

就像本报告记录的那样，人工智能相关的重大进展在过去 15 年中已经对北美城市造成了影响，而且在接下来的 15 年中将会发生更多大量的进展。近期的进步主要是由于互联网带来的大型数据集的增长和分析、传感技术的进步，以及深度学习的应用。

在接下来几年，随着公众鼓励在交通、医疗这些领域的人工智能新应用，它们肯定会以一种建立可靠的、理解的方式被引入，并且还要尊重人权与公民权利。在鼓励创新的同时，政策和流程应该解决道德、隐私和安全影响，还应该保证人工智能技术的收益公平、广泛的分配。

如果人工智能研究与应用到 2030 年以及更远的未来在北美城市生活发挥积极的影响，做上面提到的事情是很重要的。

第一部分：什么是人工智能？

本节介绍了研究人员和从业者如何定义「人工智能」以及目前正在蓬勃发展的人工智能研究和应用领域。它提出了人工智能是什么和不是什么的定义，并介绍了一些当前人工智能研究的「热点」领域。

本节为第二部分的内容奠定了基础，第二部分阐述了人工智能在八个领域和在第三部分中的影响与未来，第三部分介绍了涉及人工智能设计和公共政策的问题，并提出在保护民主价值的同时如何鼓励人工智能创新的建议。

定义人工智能

奇怪的是，人工智能缺乏一个精确的、被普遍接受的定义，这或许有助于该领域的加速成长、繁荣以及前进。虽然人工智能的从业者、研究人员和开发人员由一种粗略的方向感和一个「与它相处」的命令所引导，人工智能的定义仍然很重要，而 Nils J. Nilsson 就提供了一个有用的定义：「人工智能就是致力于让机器变得智能的活动，而智能就是使实体在其环境中具有远见地、适当地实现功能性的能力。」

从这个角度来看，对人工智能的表征取决于个人愿意「适当地」并「有远见地」为功能性提供合成软件 and 硬件的信用。一个简单的电子计算器比人类大脑进行的计算要快得多，而且几乎从来不出错。

电子计算器智能吗？像 Nilsson 一样，研究小组以一种宽泛的视角来看待此问题，认为智力取决于一个多维频谱。根据这一观点，算术计算器和人脑之间的区别不是某一类，而是规模、速度、自主性和

通用性的区别。

同样的因素可以用来评估智能的其他各例——智能语音识别软件、动物大脑、汽车巡航控制系统、围棋程序、自动调温器——并将它们放置在频谱中的适当位置。虽然我们的宽泛解释把计算器列在了智能频谱中，但是如此简单的设备与今天的人工智能相比几乎没有相似之处。

从这个角度看，对人工智能的表征取决于个人愿意「适当地」并「有远见地」为功能提供合成软件和硬件的信用。一个简单的电子计算器比人脑计算快得多而且几乎从不出错。

人工智能的边界已经远远走在前面，而计算器可以实现的功能只是当下的智能手机的百万分之一。目前人工智能开发人员正在改进、推广和扩大从当下的智能手机中所建立起来的智能。事实上人工智能领域是一个不断努力推动机器智能向前发展的过程。

具有讽刺意味的是，人工智能正在遭受失去话语权的长期灾难，最终不可避免地会被拉到边界内，即一个被称为「人工智能效应（AI effect）」或「奇怪悖论（odd paradox）」的重复模式——人工智能将一种新技术带到了普通大众中去，人们习惯了这种技术，它便不再被认为是人工智能，然后更新的技术出现了。

同样的模式将在未来继续下去。人工智能并没有「交付」一个惊雷般改变生活的产品。相反人工智能技术以一个连续的、进步的方式正在继续更好的发展。

人类尺度

值得注意的是，作为频谱的智能表征并不授予人类大脑以特殊地位。但到目前为止，即使人类拥有着「推理、实现目标、理解和生成语言、感知和响应感官输入、证明数学定理、玩具有挑战性的游戏、合成和总结信息、创造艺术和音乐甚至撰写历史」的能力，人类智力在纯粹的多功能性上与生物的和人工的世界并不相匹配。

这使得人类智能成为检测人工智能进步的一个自然选择。它甚至可能是作为一个经验法则被提出来，认为计算机能够执行的、以及人们曾经实现了的任何活动，都应该算作智力的一个实例。但是对于任何人类能力的匹配只是一个充分条件而非必要条件。至少在速度上，已经存在许多超越人类智力的系统，比如对机场成千上万的航班进行每日往返行程的调度。

对人工智能在国际象棋游戏中打败人类玩家的长期探索以及最终的成功为人类和机器智能之间的比较提供了一个高知名度的实例。人们已经被国际象棋吸引了数百年。

当建立计算机的可能性即将来临时，被许多人认为是计算机科学之父的 阿兰图灵「提到了以国际象棋作为范例来展示智力的计算机概念。」没有对强大的计算机的访问，「图灵玩了一个模拟计算机的游戏，他每走一步棋大约花半个小时。」

但只是在六十和七十年代，由来自卡耐基-梅隆大学、斯坦福、麻省理工学院、莫斯科的理论和实验物理研究所以及西北大学的小组所贡献的一长串改进之后，国际象棋游戏程序开始取得成果。终极推动是来自 IBM 的一个长期项目，Deep Blue 程序于 1997 年以 3.5-2.5 的成绩打败了国际象棋冠军

Garry Kasparov。

奇怪的是，在人工智能赶上了它的难以捉摸的目标之后没多久，Deep Blue 被描绘成一个「蛮力方法 (brute force methods)」集合，而不是「真正的智能 (real intelligence)。」事实上，IBM 关于 Deep Blue 的后续出版物提供了有关其搜索和评价程序的丰富内容，完全没有提及「智能的 (intelligent)」一词！Deep Blue 是否智能？又一次，边界移动了。

一个可操作的定义

人工智能也可由其研究人员的工作来定义。本报告主要将人工智能视为通过合成智能以研究智能特性的计算机科学的一个分支。虽然人工智能的出现是依靠着硬件计算资源的快速发展，但是这里对于软件的关注反映出人工智能界的一个趋势。尽管，最近在为基于神经网络计算建立定制硬件上取得的进展已经促进了人工智能硬件和软件之间的严格耦合。

「智能 (intelligence)」仍然是一个复杂现象，其不同方面已经吸引了几个不同研究领域的注意，包括心理学、经济学、神经科学、生物学、工程学、统计学和语言学。人工智能领域自然是得益于所有这些相关领域的进步。比如人工神经网络，如今已成为若干基于人工智能的解决方案的中心，它最初是受生物神经元中有关信息流的思想所启发。

人工智能研究趋势

直到本世纪初，人工智能的吸引点主要在于它所传递的承诺，但在过去的十五年里，大多这样的承诺已经得到兑现。人工智能技术已经充斥了我们的生活。当它们成为了社会的一股中心力量时，该领域正在从仅仅建立智能系统，转向了建立有人类意识的、值得信赖的智能系统。

几个因素加速了人工智能革命。其中最重要的是机器学习的成熟，部分由云计算资源和广泛普及的、基于 Web 的数据收集所支持。机器学习已经被「深度学习 (deep learning)」急剧地向前推进了，后者是一种利用被称作反向传播的方法所训练的适应性人工神经网络的一种形式。

信息处理算法的这种性能飞跃一直伴随着用于基本操作的硬件技术的显著进步，比如感觉、感知和目标识别。数据驱动型产品的新平台和新市场，以及发现新产品和新市场的经济激励机制，也都促进了人工智能驱动型技术的问世。

所有这些趋势都推动着下文中所描述的「热门」研究领域。这种编辑只是想要通过某个或另一个度量标准来反映目前比其他领域得到更大关注的领域。它们不一定比其他领域更重要或更有价值。事实上目前的一些「热门」领域在过去几年中并不怎么流行，而其他领域可能在未来会以类似的方式重新出现。

大规模机器学习

许多机器学习的基本问题（如监督和非监督学习）是很好理解的。目前努力的一个重点是将现有算法扩展到更庞大的数据集上。例如鉴于传统方法能够负担得起若干遍数据集的处理，现代方法是为单次处理所设计；某些情况只认同非线性方法（那些只关注一部分数据的方法）。

深度学习

成功训练卷积神经网络的能力非常有益于计算机视觉领域，比如目标识别、视频标签、行为识别和几个相关变体的应用。深度学习也在大举进军感知方面的其他领域，如音频、语音和自然语言处理。

强化学习

鉴于传统机器学习主要关注于模式挖掘，强化学习将重点转移到决策中，这种技术将有助于促进人工智能在现实世界中更深入地进入相关研究和实践领域。作为一种经验驱动型的序贯决策框架，强化学习已经存在了几十年，但是这个方法在实践中没有取得很大成功，主要是由于表征和缩放的问题。然而深度学习的出现为强化学习提供了「一贴强心剂」。

由谷歌 DeepMind 开发的计算机程序 AlphaGo 在五次对抗比赛中击败了人类围棋冠军，它最近所取得的成功在很大程度上要归功于强化学习。AlphaGo 是通过使用一个人类专家数据库来初始化一个自动代理的方法被训练的，但随后提炼的方法是通过大量地自我对抗游戏以及应用强化学习。

机器人

至少在静态环境中，机器人导航在很大程度上被解决了。目前的努力是在考虑如何训练机器人以泛型的、预测性的方式与周围世界进行交互。互动环境中产生的一个自然要求是操纵，这是当下所感兴趣的另一个话题。

深度学习革命只是刚开始影响机器人，这在很大程度上是因为要获得大的标记数据集还很困难，这些数据集已推动了其他基于学习的人工智能领域。

免去了标记数据需求的强化学习可能会有助于弥合这一差距，但是它要求系统在没有错误地伤害自己或其他系统的情况下能够安全地探索出一个政策空间。在可信赖的机器感知方面的进步，包括计算机视觉、力和触觉感知，其中大部分将由机器学习驱动，它们将继续成为推进机器人能力的关键。

计算机视觉

计算机视觉是目前最突出的机器感知形式。它是受深度学习的兴起影响最大的人工智能子领域。直到几年前，支持向量机还是大多视觉分类任务所选择的方法。但是特别是在 GPU 中的大规模计算的汇合，使得更大数据集的可获得性（尤其是通过互联网）以及神经网络算法的改进导致了基准任务中能的显著提高（比如 ImageNet 中的分类器）。计算机首次能够比人类更好地执行一些（狭义定义的）视觉分类任务。目前的研究多是关注于为图像和视频自动添加字幕。

自然语言处理

自然语言处理是另一个通常与自动语音识别一同被当做非常活跃的机器感知领域。它很快成为一种拥有大数据集的主流语言商品。谷歌宣布目前其 20% 的手机查询都是通过语音进行的，并且最近的演示已经证明了实时翻译的可能性。现在研究正在转向发展精致而能干的系统，这些系统能够通过对话

而不只是响应程式化的要求来与人互动。

协同系统

协同系统方面进行的是对模型和算法的研究，用以帮助开发能够与其他系统和人类协同工作的自主系统。该研究依赖于开发正式的协作模型，并学习让系统成为有效合作伙伴所需的能力。能够利用人类和机器的互补优势的应用正吸引到越来越多的兴趣——对人类来说可以帮助人工智能系统克服其局限性，对代理来说可以扩大人类的能力和活动。

众包和人类计算

在完成许多任务方面由于人类的能力是优于自动化方法的，因而在众包和人类计算方面，通过利用人类智力来解决那些计算机无法单独解决好的问题，该领域研究调查了增强计算机系统的方法，这项研究的提出仅仅是在大约 15 年前，现在它已经在人工智能领域确立了自己的存在。最有名的众包例子是维基百科，它是一个由网络公民维护和更新的知识库，并且在规模上和深度上远远超越了传统编译的信息源，比如百科全书和词典。

众包专注于设计出创新的方式来利用人类智力。Citizen 科学平台激发志愿者去解决科学问题，而诸如亚马逊的 Mechanical Turk 等有偿众包平台，则提供对所需要的人类智力的自动访问。通过短时间内收集大量标记训练数据和/或人机交互数据，该领域的工作促进了人工智能的其它分支学科的进步，包括计算机视觉和自然语言处理。基于人类和机器的不同能力和成本，目前的研究成果探索出了它们之间理想的任务分离。

算法博弈理论与 (基于) 计算机 (统计技术的) 社会选择

包括激励结构、人工智能的经济和社会计算维度吸引到了新的关注。自 20 世纪 80 年代初以来，分布式人工智能和多代理 (multi-agent) 系统就已经被研究了，于 20 世纪 90 年代末开始有显著起色，并由互联网所加速。一个自然的要求是系统能够处理潜在的不恰当激励，包括自己所感兴趣的人类参加者或公司，以及自动化的、基于人工智能的、代表它们的代理。

备受关注的主题包括计算机制设计 (computational mechanism design) (一种激励设计的经济理论，它寻求激励兼容的系统，其中输入会被如实报告)、(基于) 计算机 (统计技术的) 社会选择 (computational social choice) (一种有关如何为替代品排列顺序的理论)、激励对齐信息获取 (incentive aligned information elicitation) (预测市场、评分规则、同行预测) 和算法博弈理论 (algorithmic game theory) (市场、网络游戏和室内游戏的平衡，比如 poker——它在近几年通过抽象技术和无遗憾学习 (no-regret learning) 已经取得了显著的进步)。

物联网 (IoT)

越来越多的研究机构致力于这样一个想法：一系列设备可以相互连接以收集和分享它们的感官信息。这些设备可以包括家电、汽车、建筑、相机和其他东西。虽然这只是一个技术和无线网络连接设备的

问题，人工智能可以为了智能的、有用的目的去处理和使用所产生的大量数据。目前这些设备使用的是令人眼花缭乱的各种不兼容的通信协议。人工智能可以帮助克服这个「巴别塔」。

神经形态计算

传统计算机执行计算的冯诺依曼模型，它分离了输入/输出、指令处理和存储器模块。随着深度神经网络在一系列任务中的成功，制造商正在积极追求计算的替代模型——特别是那些受到生物神经网络所启发的——为了提高硬件的效率和计算系统的稳定性的模型。

目前这种「神经形态的（neuromorphic）」计算机尚未清楚地显示出巨大成功，而是刚开始有望实现商业化。但可能它们在不久的将来会变成寻常事物（即使仅作为冯诺依曼所增加的兄弟姐妹们）。深度神经网络在应用景观中已经激起了异常波动。当这些网络可以在专门的神经形态硬件上被训练和被执行，而不是像今天这样在标准的冯诺依曼结构中被模拟时，一个更大的波动可能会到来。

总体趋势以及人工智能研究的未来

数据驱动型范式的巨大成功取代了传统的人工智能范式。诸如定理证明、基于逻辑的知识表征与推理，这些程序获得的关注度在降低，部分原因是与现实世界基础相连接的持续挑战。规划（Planning）在七十和八十年代是人工智能研究的一根支柱，也受到了后期较少的关注，部分原因是它强烈依赖于建模假设，难以在实际的应用中得到满足。

基于模型的方法——比如视觉方面基于物理的方法和机器人技术中的传统控制与制图——已经有很大一部分让位于通过检测手边任务的动作结果来实现闭环的数据驱动型方法。即使最近非常受欢迎的贝叶斯推理和图形模式似乎也正在失宠，被数据和深度学习显著成果的洪流所淹没。

研究小组预计在接下来的十五年中，会有更多关注集中在针对人类意识系统的开发上，这意味着它们是明确按照要与之互动的人类特点来进行建模与设计的。很多人的兴趣点在于试图找到新的、创造性的方法来开发互动和可扩展的方式来教机器人。

此外在考虑社会和经济维度的人工智能时，物联网型的系统——设备和云——正变得越来越受欢迎。在未来的几年中，对人类安全的、新的感知/目标识别能力和机器人平台将会增加，以及数据驱动型产品数量与其市场规模将会变大。

研究小组还预计当从业者意识到纯粹的端到端深度学习方法的不可避免的局限性时，会重新出现一些人工智能的传统形式。我们不鼓励年轻的研究人员重新发明理论，而是在人工智能领域以及相关领域（比如控制理论、认知科学和心理学）的第一个五十年期间，保持对于该领域多方面显著进展的觉察。

第二部分：人工智能在各领域的应用

虽然人工智能的很多研究和应用会基于一些通用技术，比如说机器学习，但在不同的经济和社会部门还是会有所区别。我们称之为不同的领域（domain），接下来的这部分将介绍人工智能研究和应用的不同类型，以及影响和挑战，主要有八个方面：交通、家庭服务机器人、医疗健康、教育、低资源社区、公共安全、工作和就业、娱乐。

基于这些分析，我们还预测了一个有代表性的北美城市在未来 15 年的趋势。与人工智能的流行文化中的典型叙述不同，我们寻求提供一个平衡的观点来分析，人工智能是如何开始影响我们日常生活的，以及从现在到 2030 年，这些影响将如何发展。

交通

交通可能会成为首批几个特定应用领域之一，在这些领域，大众需要对人工智能系统在执行危险任务中的可靠性和安全性加以信任。自动化交通会很快司空见惯，大多数人在嵌入人工智能系统的实体交通工作的首次体验将强有力的影响公众对人工智能的感知。

一旦硬件做到足够安全和稳健，它进入日常生活的速度将使公众所惊讶，这可能需要时间进行调整。随着汽车逐渐取代人类成为更好的司机，城市居民拥有的车辆将减少，住所和工作地点的距离也会更远，花费的时间也会发生变化。这会导致出现一种全新的城市组织架构。再者，2030 年在典型的北美城市中，人工智能带来的改变将不仅限于车辆，还会影响各种飞行器和个人机器人，这将引发对社会、道德和政策等议题的讨论。

一些关键技术已经催化了人工智能在交通领域的广泛应用。与 2000 年相比，智能手机的普及，各类传感器成本的下降和性能的提升，使得我们今天能获取的关于个人和群体交通数据的规模和多样性获得了惊人的增长。如果没有这些数据、连接、实时传感和交通预测等应用、路线规划，共享乘车和自动驾驶将不可能实现。

智能汽车

2001 年，GPS 被应用于私家车用于定位，目前已经成为交通系统的基础部件。GPS 在辅助司机进行驾驶时也将关于交通模式的大量数据提供给科技公司和城市部门。拥有 GPS 功能的智能手机的广泛应用也进一步增强了链接以及位置信息的共享。

当前的车辆还配备了各种传感设备。根据估算，美国平均每一辆车上装有 70 个传感器，包括陀螺仪、加速器、亿光环境光源传感器和湿度传感器。对于汽车来说，传感器已经不是什么新事物了。2000 年之前制造的汽车已经装配了能够获取自身速度、加速和轮胎状况的传感器。

这些车辆还拥有一些功能，能够把实时传感和决策（例如 ABS）、气囊控制、电气牵引系统和控制和电子稳定程序结合起来。自从 2003 年以来，一些自动化功能就已经被引入了商用车，具体功能见下表。

Context	Automated Functionality	Release Date
Parking	Intelligent Parking Assist System	Since 2003 ²³
Parking	Summon	Since 2016 ²⁴
Arterial & Highway	Lane departure system	Since 2004 in North America ²⁵
Arterial & Highway	Adaptive cruise control	Since 2005 in North America ²⁶
Highway	Blind spot monitoring	2007 ²⁷
Highway	Lane changing	2015 ²⁸

这些功能能够协助司机，甚至完全取代司机的某些活动来增强安全性和舒适性。现在的汽车能够自动泊车，在高速路上可以自适应巡航控制，变换车道时可以提示驾驶员盲区中出现的物体信息。视觉和雷达技术被引入到碰撞预防系统。深度学习也已经应用在了这些方向，通过对周围环境的物体检测和声音识别来提升自动驾驶汽车的性能。

不远的未来，在用于驾驶的功能方面，感知算法将超过人类的水平。包括视觉在内的自动化感知，在处理是被和跟踪等任务时，已经接近人类水平。

除了感知方面的进步，随之出现的还有算法的进一步提升所带来的推理和规划能力。近期一份报告预测，自动驾驶汽车将在 2020 年得到广泛应用。而自动驾驶功能的应用也将不局限于个人交通。

我们将看到自动驾驶汽车和远程操控的运载车辆、飞行器、和自动驾驶卡车。基于用户共享的交通服务也将充分利用自动驾驶汽车。此外，机器人技术的进步也将更有利于其他类型自动驾驶设备的创造和应用，包括机器人、无人机等。

现在还不明确的一点是，自动驾驶汽车需要发展到何种程度才能引起大众的广泛接受。半自动驾驶汽车中的人车合作与人类司机认知符合的潜在意义都还未能完全理解。

但如果未来的自动驾驶汽车以可预测的速度走向应用，且超越人类的驾驶水平，那势必将带来重要的社会变化。自动驾驶汽车消除了美国交通伤亡的一大原因，可以延长人们的平均寿命。

美国一个通勤者的单程平均驾驶时间是 25 分钟。当有了自动驾驶技术之后，人们可以在通勤中有更多的时间来工作和休闲。自动驾驶汽车带来的舒适度的提升和认知负担的减少，再加上共享交通，将影响人们选择居住地点。

自动驾驶汽车和点到点交通服务可能消除对私人汽车的需要。对整体汽车产业的影响难以预测。空汽车以及人们日益增加的旅行意愿可能导致更多的总驾驶里程。

此外，共享自动驾驶汽车（人们使用汽车作为服务而非自己拥有汽车）可能减少总里程，特别是如果与好的结构合理的激励机制结合起来，比如通行税或者折扣，就能铺开旅行需求、拼车、减少拥堵。

共享交通的实现可能取代对公共交通的需求，或者公共交通可能改变形式，迈向个人快速公交系统。这已经在 4 个城市中实现了，它们使用小容量汽车在许多站台间进行按需运输与点到点运输。

随着自动驾驶汽车变得更加普遍，安全上的问题将会出现，包括如何在发布前保证该技术是安全的、在不同的路况下有适当测试的。自动驾驶汽车和连接交通基础设施将为黑客创造一个新的犯罪地点。当死亡不可避免时，汽车编程选择牺牲什么人也会引发到的问题，特别是当汽车需要瞬间做出选择的时候。

美国大部分州的法律系统没有覆盖自动驾驶汽车的条例。到 2016 年，美国的 4 个州（内华达、佛罗里达、加州、密歇根）、加拿大的安大略、英国、法国、瑞典都已经通过了在公共道路测试自动驾驶汽车的条例。甚至这些法律都没有解决自动驾驶汽车和半自动驾驶汽车责任和事故问责的问题。

交通规划

到 2005 年，城市开始投资交通基础设施以为汽车和行人交通开发传感能力。目前已经使用的传感器包括感应线圈、视频摄像头、远程交通微波传感器、雷达和 GPS。比如，2013 年纽约市开始使用微博传感器、摄像机网络和流量计数器来检测城市中的车辆交通。

城市使用人工智能方法通过多种方式来优化服务，比如公交车和地铁规划、跟踪交通状况以动态调整限速或在高速路、桥梁和 HOV 车道上采用智能定价。通过道路网中的传感器和相机，他们可以优化交通等定时以改善交通流和帮助自动执法。这些动态策略的目标是更好地利用交通网中受限的资源，而且也因为数据的可用性和个体的广泛连通性而成为了可能。

在 2000 年代以前，交通规划者被迫依赖受限于特定日期或时间的静态定价策略来管理需求。随着动态定价策略的采用，这会带来涉及到公共利益公平分配的新问题，因为高需求环境的市场条件可能会让公共部分无法使用这些服务。

大规模数据的可用性也让交通成为了机器学习的一个完美领域。自 2006 年以来，Mapquest、谷歌地图和 Bing 地图等应用已经广泛地被公众用来规划旅行线路、使用公交系统、接收关于交通状况的实时信息和预测、以及寻找某个地点周围的服务。优化搜索算法已经被用于汽车和行人去往给定目的地的路径规划。

尽管有这些进步，传感和优化技术在城市基础设施上的广泛应用却慢于这些技术在单独的车辆或人上的应用。尽管已经有个别城市已经部署了传感和优化应用，但目前却还没有传感基础设施和人工智能技术的标准可用。

基础设施需要成本，各个城市的重点也不一样，而且涉及到人工智能的各方之间协调的高额成本对城市基础设施的影响也越来越大。个体行动、他们的偏好和他们的目标的准确预测模型会随着更大的数据可用性而涌现。相关的道德问题将在本报告的第三节进行讨论。

美国交通运输部在 2016 年发布了一份提案，要求中等规模的城市想象用于交通的智慧城市基础设施。这一倡议计划向展示了技术和数据可以如何被用于重构人与货物的流动方式的城市奖励 4000 万美元。

一个愿景是实现车联网，其可以通过车对车的通信实现高水平的驾驶安全。如果这一愿景成为了现实，我们预计在多代理协调、协作和规划上的进步将极大地影响未来的汽车发展，并在打造更加可靠更加高效的交通运输系统上发挥作用。

机器人也可能会通过运输个人和包裹来参与到交通中（比如，Segway 机器人）。对于货物运输，人们对无人机的兴趣越来越大，亚马逊现在就正在测试使用无人机的快递系统，尽管在合适的的规则和监管方面还存在着一些问题。

传感能力的提升、无人机的采用和互连的交通基础设施也将引发对个人隐私和私密数据安全的担忧。未来几年，这些问题和相关的交通运输问题也将需要得到解决——不管是通过部分行业内先发制人的行动还是在法律的框架内。正如第三节政策讨论中所指出的那样，这些行动的效果将会影响交通运输行业内人工智能相关进展的步伐和范围。

即时交通

Uber 和 Lyft 等即时交通服务已经涌现成为了传感、连接和人工智能的另一项关键应用，这些技术可以使用算法根据位置和合适度（声誉模型）来匹配司机和乘客。

通过动态定价，这些服务可以通过支付意愿进行配给，动态定价还有利于估计司机数量的增长，这已经成为了城市交通的一种流行的方法。随着它们的快速发展，一些政策和法律问题也随之出现了，比如和已有的出租车服务竞争以及对缺乏监管和安全的担忧。按需交通服务似乎很有可能成为自动驾驶汽车的一个主要推动力。

拼车和驾乘共享一直以来都被视为有希望缓解交通拥堵的方法，而且还能更好地利用个人交通资源。Zimride 和 Nuride 等服务将具有类似路线的人聚集到一起进行联合旅行。但拼车的方法难以获得大规模的推动力。

人机交互

几十年来，人们一直在想象非常不同的、未来感十足的交通载具。尽管未来的车辆将会更加智能，无人机也将得到广泛应用，但在 2030 年之前我们不太可能会大规模地用上和现在已有的交通工具在外观和功能上大相径庭的交通载具。我们的研究小组并不指望可以海陆空穿行的无人机或四轴飞行器在这个时间范围内变成常用的交通运输方式（尽管现在已有原型机出现了）。

我们预计人类会成为自动驾驶汽车和无人机的训练、执行和评估过程中的合作伙伴。这种伙伴关系可能是物理地，也可能是虚拟的。我们预计算法的进步将促进机器从人类输入中的学习。我们还预计会出现建模人类注意机制、以及支持人机之间的通信和协作的模型和算法。这是未来汽车发展的一个组成部分。

家庭/服务机器人

过去十五年中，机器人已经进入了人们的家庭。但应用种类的增长慢得让人失望，与此同时，日益复杂的人工智能也被部署到了已有的应用之中。人工智能的进步常常从机械的革新中获取灵感，而这反过来又带来了新的人工智能技术。

未来十五年，在典型的北美城市里，机械和人工智能技术的共同进步将有望增加家用机器人的使用和应用的可靠性和安全性。特定用途的机器人将被用于快递、清洁办公室和强化安全，但在可预见的未来内，技术限制和可靠机械设备的高成本将继续限制狭窄领域内应用的商业机会。至于自动驾驶汽车和其它新型的交通机器，创造可靠的、成熟的硬件的难度不应该被低估。

真空吸尘器

经过许多年的开发之后，真空清洁机器人 Electrolux Trilobite 在 2001 年成为了第一款商用家庭机器人。它有一个简单的控制系统，可以做到避障和一些导航。一年之后，iRobot 推出了 Roomba，它的价格只有 Trilobite 的十分之一，只有 512 字节的 RAM，运行着一个基于行为的控制器。其所能做到的最智能的事情是避免从楼梯上跌落。自那时起，已有 1600 万台 Roomba 被部署到了世界各地，现在也已有了其它一些竞争品牌。

随着低成本嵌入式处理器的处理能力和 RAM 容量在 2000 年时令人沮丧的状态的基础上实现的提升，这些机器人的人工智能能力也取得了显著的提升。简单的导航、自充电和处理垃圾箱装满的行为等已经得到了解决，接着又解决了电线和地毯流苏的问题，这些是通过机械改进和基于传感器的感知的结合实现的。

最近，全 VSLAM（视觉同步定位和测绘，一种已经存在了 20 年的人工智能技术）的加入让机器人可以构建它们所清洁的房屋的全 3D 世界模型，从而让他们在他们的清洁范围内更加高效。

人们早期预测过的一些家用机器人应用还没有实现。机器人真空吸尘器也仅限于局部的平坦区域，而真正的家里有大量的单级阶梯甚至楼梯；在机器人在真实家庭中的运动上的研究还非常少。硬件平台的构建仍然很困难，人们只对很少的应用有足够的购买意愿。用于图像标注和 3D 对象识别等功能的感知算法尽管在人工智能会议上很常见，但离产品化还仍有几年的时间。

家庭机器人 2030

尽管家庭机器人到来的速度缓慢，但仍有迹象表明在接下来的 15 年内会发生改变。像 Amazon Robotics 和 Uber 这样的公司正在使用各种聚合技术开发大规模经济。也包括：

模块内系统（System in Module, SiM）和许多片上系统（System on Chip, SoC）的子系统，如今都被手机芯片制造商拒之门外（高通的 Snap Dragon，三星的 Artik 等）。

这些系统至少要比十年前的超级计算机更好，它们有着 8 ~ 64 核，配备密码学的专门化硅片、摄像头

驱动、附加的 DSPs，以及针对特定感知算法的硬硅片。这意味着低成本设备将能够支持比我们 15 年前能够想象到的更多的机载人工智能。

云（其他人的计算机）将使得家庭机器人上的新软件的快速迭代成为可能，共享更多的不同家庭内收集的数据集，这反过来也能供给基于云的机器学习，然后进一步改进已经部署了的机器人。

由深度学习带来的在语音理解和图像标记上的巨大进步将增强机器人在家庭中与人的交互。

低成本的 3D 传感器受到游戏平台的驱动，全球数千名研究人员已经开始在 3D 感知算法上的研究，这也将加速家庭和服务机器人的开发和采用。

在过去的 3 年中，低成本的安全机械臂在全球已经被数百个研究实验室引入，激起一波在家用机器人上的新型研究，这个日期可能是 2025 年。全球超过 6 家创业公司正在开发基于人工智能的家庭机器人，如今主要集中于社交互动。最后，新的道德和隐私问题可能会浮出表面。

医疗

对人工智能而言，医疗领域一直被视为一个很有前景的应用领域。基于人工智能的应用在接下来的几年能够为千百万人改进健康结果和生活质量，但这是在它们被医生、护士、病人所信任，政策、条例和商业障碍被移除的情况下。主要的应用包括临床决策支持、病人监控、辅导、在外科手术或者病人看护中的自动化设备、医疗系统的管理。

近期的成功，比如挖掘社交媒体数据推断潜在的健康风险、机器学习预测风险中的病人、机器人支持外科手术，已经为人工智能在医疗领域的应用扩展出了极大的应用可能。与医学专家和病人的交互方法的改进将会是一大挑战。

至于其他领域，数据是一个关键点。在从个人监护设备和手机 App 上、临床电子数据记录上收集有用的数据方面，我们已经取得了巨大的进展，从协助医疗流程和医院运行的机器人那里收集的数据可能较少一些。但使用这些数据帮助个体病人和群体病人进行更精细的针对和治疗已经被证明极其的困难。

研究和部署人工智能应用已经被过时的条例和激励机制拉扯后腿。在这样大型的、复杂的系统中，贫乏的人机交互方法和固有的难题以及部署技术的风险也阻碍了人工智能在医疗的实现。减少或者移除这些障碍，结合目前的创新，有潜力在接下来几年为千百万人极大的改进健康结果和生活质量。

临床应用

十几年来，人工智能临床助手的观点已经成为了陈词滥调。尽管在医疗领域已经有了成功的人工智能相关技术的试点，目前医疗输送系统仍非常不幸的因结构病态难以吸收、应用新进展。由平价医疗法案（Affordable Care Act）提供的激励机制已经加快了 EHR 在临床实践中的渗透，但执行力度缺乏，破坏了临床医师对 EHR 作用的信心。

一小撮公司控制着 EHR 市场，用户界面普遍认为不合格，包括医师通常不理睬恼人的弹出窗口。使

来自 EHR 数据的新型分析技术的潜力，包括人工智能，就由于这些原因和其他的条例与架构障碍而难以被意识到。

展望下一个 15 年，人工智能在进步，如果配合足够的数据与有的放矢的系统，这极有可能改变分配给临床医师的认知任务。医师如今还是老一套的从病人那里听取病症描述，在脑海中关联与已知疾病的临床表征的模式。

有了自动化助手，医师反而能够监督这一流程，应用自己的经验和直觉指导输入流程，并评估机器智能的输出。医师文字版本的实地经验仍然很重要。一个主要挑战是将人类方面的医疗与自动化推理流程最佳结合。

为了在未来取得进步，临床医师必须从开始就参与其中，从而保证系统能很好建立且值得信任。新一代的熟悉技术的医师已经在手机设备上经常使用专门化的 App 了。同时，初级保健医师的工作量已经增加到了寻求帮助的点。

因此，开发新机器学习方法的机会，通过自动挖掘科学文献创造推断架构模式的机会，以及通过支持自由形式对话的认知助手的机会，都是无比巨大的。鉴于这些进展并未受到条例、法律、社会障碍的阻碍，对医疗价值的巨大改进在我们的掌握之中。

医疗分析

在人口层次上，人工智能从数百万病人临床记录挖掘结果的能力能使得精细的、更私人化的诊断和治疗成为可能。随着完全的、一生仅一次的基因序列成为病人常规，基因型-表型连接的自动发现也将成为可能。一个相关的（可能较早）能力将会是发现「像我的病人一样的病人」，从而基于分析类似群体做治疗决策。

传统的与非传统的医疗数据，因社交平台得以加强，可能导致自定义亚人口的出现，每个人受到周围有自动建议和监控系统的医疗提供商生态系统的管理。随着医疗流程和数百万个体的终身医疗记录变得可用，这些发展有潜力从根本上转变医疗输送系统。

类似地，可穿戴设备对个人环境数据的自动获取将扩展个人医疗。随着供应商发现吸引大量人口（例如，共享医疗）的发现，然后创造出能挖掘产生个人分析与建议的大规模人工数据，这些活动将具有更大的商业化可行性。

不幸的是，FDA 在认可创造性诊断软件上进展缓慢，在加速创新上也有遗留障碍。HIPAA（Health Insurance Portability and Accountability Act）法案要求保护病人隐私，这导致了病人数据流应用人工智能技术的法律障碍。已批准药物的意外副作用可能要比今天出现的更快、更严厉，但分析药物相互作用的手机 APP 可能受阻于应用病人病例信息。

更普遍的是，医疗领域的人工智能研究与创新因缺乏被普遍接受的隐私保护方法与标准而受阻。FDA 认可创新软件的速度缓慢，部分是因为不能清楚理解这些系统的成本/收益权衡。如果监管者（主要是 FDA）认识到有效的上市后报告是一个可靠的对安全风险的对冲，对新的治疗与干预措施的最初认可

可能成为可能。

自动图像解析也是研究了数十年的课程。对大型弱标记图像的解析进展是爆发性，比如来自网页的大型图片档案。乍一看，你会很惊讶几乎没有变革类似于医疗图像解析的。大部分生成医疗图像的方式（CT、MR、超声波）都是电子化的，图像都是存档的，而且也有大型的公司有专注于图像的内部研发机构（比如，西门子、飞利浦、GE）。

但目前为止，还有数个障碍限制了进展。在过去十年中，大部分医院的图像档案都电子化了。更重要的问题是不知道图片中的内容（图片中是肝脏还是肾脏？），更不要说做出精细判断了（肝脏中的小黑点表明潜在的肿瘤？）。

严厉的条例管理着这些高风险判断。即使有前沿的技术，放射科医师仍然倾向于查看图像，所以价值定位还并不引人注目。同样，医疗条例阻碍了机构间的数据联合。因此，只有大型的整合医疗公司才能够解决这些问题，比如 **Kaiser Permanente**。

自动/增强图像解析的发展开始加快。接下来的 15 年内可能不会带来全自动的放射学，但最初的图像分类或者下一层的图像检查可能提高医疗图像生成的速度的成本高效益。当与电子病人记录系统结合的时候，大规模的机器学习技术可能应用于医疗图像数据。

例如，多专业医疗系统有百万病人扫描档案，每个都有相关的放射报告，而且大部分有相关的病人记录。已经有论文表明通过在这些数据上训练深度神经网络，能产生基础的放射性发现，有很高的可靠性。

医疗机器人

15 年前，医疗机器人大多是科幻小说内的形象。一家从 IBM 衍生出的名为 **Robodoc** 的公司为矫形外科手术开发机器人系统。技术是可用的，但公司难以商业化最终关闭，技术也被收购。尽管，近期外科机器人的研究与实际使用也开始爆发。

在 2000 年，**Intuitive Surgical** 介绍了达芬奇系统，这是一个最初市场化的新技术，支持最低程度的心脏搭桥手术，然后在前列腺癌的治疗上也获得大量的市场吸引力，并在 2003 年与其主要的竞争对手 **Computer Motion** 合并。

医院运行中的智能自动化方面的成果较少，它不像外科机器人。20 年前，一家名为 **HelpMate** 的公司为医院输送（比如食物和医疗记录的输送）创造了一个机器人，但最终破产。

最近，**Aethon** 为基础输送引入了 **TUG** 机器人，但目前为止少有医院投资该技术。然而，其他服务产业的机器人，比如酒店与仓库，包括 **Amazon Robotics**（前身为 **Kiva**）的机器人都证明这些技术至少在大规模环境中是可用且具有成本效益的。此外，这些技术最终可能在医疗领域产生额外的创新。

往前看，医疗领域出现的很多任务将增强而非完全的自动化。例如，机器人可能向正确的屋子输送物品，但然后就需要人进行选择，并放到最终位置。一旦一位病人站到步行器上，陪伴病人在走廊走动可能就相对简单了（尽管面对术后或者老年病人，这个任务更琐碎，特别是在走廊有大量器械和其他

人的情况下)。

在手术针准确放置的情况下，进行缝合对机器人而言也很简单。这表明很多未来的系统将在人与机器之间亲密交互，而且需要在两者间建立协作的技术。

自动化的成长将产生对医疗流程的新洞见。历史上，机器人不是非常受数据驱动或是面向数据的科学。随着（半）自动渗透到医疗领域，这正在发生改变。随着新型外科、输送、以及病人看护平台变成线上，量化以及预测分析开始建立到来自这些平台的数据之上。

这些数据将被用于评估性能质量、识别缺陷、错误或者潜在的最优方式，而且将被作为反馈改进新能。简言之，这些平台将促进在做什么与结果档案间建立连接，使得真正的「闭环」医疗成为可能。

移动健康

目前为止，医疗上的循证分析依靠传统的医疗数据，主要是上面讨论过的电子医疗记录。在临床环境，有希望带入新数据。例如，在人工智能代理的帮助下，Tele-Language 使得医师能够同时与数个病人进行语言治疗。而且从病人手机被动收集的数据中提取行为模式并发出警报的 Lifegraph 已被以色列的精神病医师采纳，用于检测病人苦恼行为的早期迹象。

向前看，受到移动计算革命的驱动，自然环境下生物统计学的惊人成长以及平台的爆发和应用是一个有希望且未曾意料到的趋势。大量的手机 App 如今在提供信息、引入行为修改，或识别「像我一样」的群体。这些结合更专门化的运动追踪设备（比如 Fitbit）趋势的兴起，加上家庭环境与健康追踪设备之间的（内部）连通性的兴起，已经创造出了一个惊人的新的创新领域。

通过结合社交和医疗数据，一些医疗 App 能进行数据挖掘、学习与预测，尽管他们的预测相对基础。数据的聚合与功能性的跨应用可能将刺激出新的甚至更显著的产品，比如不仅建议锻炼日程甚至建议最佳锻炼时间的 App，并且还能在这些锻炼日程提供辅导。

老年看护

在接下来的 15 年中，美国的老年人数量的增长将超过 50%。美国国家统计局表示在接下来 10 年内，家庭健康助手将增长 38%。尽管该领域有广泛的机会（基础的深灰指出、交互以及通信设备，家庭健康监控，各种简单的户内身体助手，比如步行器、备餐助手），在过去的 15 年中发生过很少。

但即将到来的时代转变将伴随着老年人对技术接受度上的改变。目前出生于 1946 年的 70 岁的老人可能在中年或者更晚的年纪才体验到一些形式的个性化 IT，然而如今 50 岁的老人对科技更亲近也更熟悉。最终，对支持身体、情感、社交以及心理健康的可用的或者成熟的技术，老人将有更大的兴趣，也将有更大的市场。以下是一些可能存在的例子：

生活质量与独立性：

- 自动化交通将为老人提供持续的独立性以及更宽的社交界限。

- 共享信息将帮助家庭与出门在外的家庭保持联系，预测分析可能被用于推动家庭向积极行为发展，比如提醒往家打电话。
- 家庭中的智能设备将在日常生活活动上提供帮助，比如做饭。如果机器人操作能力有足够的改进，也能帮助老人穿衣服、如厕。

健康保健

- 监控运动与活动的手机应用配合社交平台将能够推荐保持心理与身体健康的建议。
- 室内健康健康和健康信息的接入将能够检测心态与行为的变化，并做出护理提醒。
- 个性化健康管理将帮助消除与多合并病症和/或治疗作用之间关联的复杂性。

治疗与设备

- 更好的听从助手与视觉辅助设备的建议将消除听觉与视觉障碍上的影响，提高安全性以及与社会沟通。
- 个性化康复与家庭内治疗将减少去医院以及护理中心的需要。
- 身体辅助设备（智能步行器、轮椅、外骨骼）将拓展身体有限制的人的活动范围。

研究小组期望低成本传感技术的爆发能为家庭内的老人提供大量的能力。原则上，社会中实际存在（指人）以及有简单身体能力（例如，有基础交流能力的移动机器人）应该为新的创新提供一个平台。然而，如此做需要融合人工智能的多个领域（自然语言处理、推理、学习、感知、机器人），从而创造一个老年人可用的、合用的系统。

这些创新将引发关乎多个圈子的隐私问题，包括朋友、家庭以及护理员。

教育

在过去的十五年间，教育界见证了为数众多的人工智能科技的进步。诸如 K-12 线上教育以及大学配套设备等等应用已经被教育家和学习者广泛利用。尽管素质教育还是需要人类教师的活跃参与，但人工智能在所有层面上都带来了强化教育的希望，尤其是大规模定制化教育。如何找到通过人工智能技术来最优化整合人类互动与面对面学习将是一个关键性的挑战，这一点医疗行业也是如此。

机器人早已经成为了广受欢迎的教育设备，最早可以追溯到 1980 年 MIT Media Lab 所研制出的 Lego Mindstorms。智能辅导系统（ITS）也成为了针对科学、数学、语言学以及其他学科相匹配的学生互动导师。

自然语言处理，尤其是在与机器学习和众包结合以后，有力推进了线上学习，并让教师可以在扩大教室规模的同时还能做到解决个体学生的学习需求与风格。大型线上学习的系统所得的数据已经为学习分析产生了迅速增长的动力。

但是，学院与大学采用人工智能技术的步伐依然很缓慢，主要是由于资金的缺乏，以及其可以帮助学生达成学习目标的有力证据。一个典型美国北部城市的未来五十年，智能导师与其他人工智能技术帮助教师在课堂或家中工作的规模很有可能会显著扩大，因为意愿学习是基于虚拟现实的应用。但是计算机为基础的学习系统将无法完全替代学校里的教师们。

教育机器人

今天，K-12 线上教育学校能够利用更加复杂和多样化的设备，这些都来自于许多创造了可在多种语言中编程的新型传感技术的公司。Ozobot 是一个机器人，教孩子们编代码以及逻辑分析去让它跳舞或根据颜色代码模型玩游戏。Cubelets 帮助教孩子们通过组装有着不同功能的机器人模块以进行思考、行动或感知来学习逻辑思维。Wonder Workshop 的 Dash 和 Dot 则有着广泛的编程能力。

八岁以及以上的孩子们可以利用一个虚拟编程语言 Blockly 创造简单的动作，或者通过 C 语言或 Java 制造一些 IOS 或安卓的应用。PLEO 机器人是一个机器人宠物，帮助孩子们通过让机器人根据环境的不同方面做出反应反应学习生物学。然而，尽管有些机器人的确参与其中并带来乐趣，但是为了让它们更加普遍被利用，我们还需要一些有着足够说服性的证据以证明它们可以提高学生们的学术表现。

智能辅导系统 (ITS) 与线上学习

ITS 由一些诸如 Why-2 Atlas 的研究实验室项目发展而来，后者帮助人-机对话在这个时代的早期解决物理问题。ITS 从实验室的试验阶段到实际应用的迅速转变十分惊人，但同时也是我们乐于见到的。可下载的软件与线上系统，例如 Carnegie Speech 或 Duolingo 提供了利用自动语音识别和神经语言程序学技术来识别语言中的错误并帮助用户修改的外语训练。

辅导系统，例如 Carnegie Cognitive Tutor 已经在一些美国高校中有所应用，帮助学生们学习数学。其他 ITS 也针对在图像、电路、医疗诊断、计算机知识和编程、遗传学、化学方面的训练进行了发展。认知辅导师软件可以模仿一个优秀的人类辅导师，例如在学生遇到某些数学问题时提供暗示等。基于所要求的暗示与提供的答案，辅导师可以提供详细具体的反馈。

这种应用也在向更高的教育扩展。一个叫做 SHERLOCK 的 ITS 正在开始用来教空军技师诊断飞行器中的电路系统障碍。南加州大学的信息科学学院已经研发了更多先进的，基于 avatar 的训练模块帮助训练被送往国际岗位的军队人员与有着不同文化背景的人们进行交流。对于个性化辅导的新算法，例如 Bayesian Knowledge Tracing，让个体掌握学习和解决问题的次序。

大部分的惊喜来自于 MOOCs 的爆炸式发展与其他各种线上教育的模型——包括像维基百科和 Khan Academy 以及复杂的学习管理系统，建立于同步或异步教育和适应性学习工具。自从上世纪九十年代后期，例如像 Educational Testing Service 和 Pearson 已经在研发自动化的自然语言程序学估测工具，在标准测试中联合评估论文。许多 MOOCs 课程非常受欢迎，还有 EdX、Coursera 以及 Udacity 都在利用自然语言处理、机器学习与众包技术来评估较短的答案与论文中的错误以及编程作业。支撑专业生级别的教育和终生学习的线上教育系统也在迅速扩张。这些系统都具有很大的潜力因为面对面的交流需要对于专家和跳槽者并不是非常重要。尽管他们不是人工智能系统和应用界的领导，但是他们会成为这些技术在经历测试和强化过程中的早期采用者。

可以认为人工智能是教师的秘密武器，特别是在高等教育中，可使用人工智能大幅度扩展他们的课堂级别。为了不断测试不同年级的学生，自动生成问题也是有可能，比如评估单词量的问题、多项选择

问题，使用到了 WordNet、Wikipedia、和线上知识本体这样的资源。随着线上课程的爆发，这些技术无疑将被线上教育采纳。尽管这些系统在教育系统中的长期影响还不明显，人工智能社区在很短时间内学习到了很多东西。

学习分析

从 MOOC 到 Khan Academy 这些大规模线上学习系统以及较小的线上项目收集到的数据集，以及促进了学习分析领域的快速发展。线上资源不仅对普遍传播有里，也是进行数据收集以及设备实验的天然工具，它有助于科学发现并规模化的改进学习质量。Society for Learning Analytics Research (SOLAR) 这样的公司、Learning Analytics and Knowledge Conference 和 Learning at Scale Conference (L@S) 这样的大会的出现无一不表明这一趋势。该社区也使用机器学习、自然语言处理以及其他技术分析学生的参与度、行为和结果。

目前的计划想要建模学生共同的错误概念，预测那个学生有失败的风险，并提供与学习结果紧密关联的实时反馈。近期的的工作也主要专注于了解关于阅读理解、书写、知识掌握、记忆和应用到教育实践（通过开发并测试这些教育技术）中的认知流程。

挑战和机遇

可以预见，人工智能技术将在学校、学院和大学有越来越先进的应用。其不可用的大部分原因是这些机构中金融资源以及数据的缺乏，从而无法有效地利用这项技术。这些问题正在得到解决，尽管很缓慢——通过私人基金和众多项目在暑期项目中培训（主要是）中学教师。就像在人工智能的其它领域中一样，过于 MOOC 的能力的过度炒作和承诺意味着预期往往会超出实际。圣何塞州立大学等特定机构的与 Udacity 的实验经历能让我们对新型教育技术的可能性有更清醒的认识。

在未来十五年，人类教师很可能会在人工智能技术的协助下实现更好的人类交互——不管是在教室里还是在家里。本研究小组预计会出现更一般化的而且更先进的虚拟现实场景，让学习可以沉浸式地学习他们将要学习的科目。这个方向已经有人工智能研究者和人文与社会科学研究者增加协作所得到的一些进展，例如斯坦福大学的 Galileo Correspondence Project 和哥伦比亚大学的 Making and Knowing Project。这些跨学科项目将历史文献和虚拟现实应用结合起来探索交互式考古。虚拟现实技术已经在生物学、解剖学、地质学和天文学等自然科学领域得到了应用，从而让学生可以与在真实环境很难触及的环境和物体进行交互。这种对过去世界和幻想世界的再创造将会在艺术和一些科学的学习中变得越来越流行。

人工智能技术将逐渐模糊正式的课堂教育与自学的个人学习之间的界限。比如，自适应学习系统 (Adaptive learning systems) 将会变成高等教育的核心，因为这可以在控制成本的同时教授大量学生，同时还可以让学生更快速地更换学校。尽管正式教育不会消失，但本研究小组相信 MOOC 和其它形式的在线教育将会以一种混合的教室体验成为所有阶段的学习的一部分——从幼儿园到大学。这一发展将促进越来越定制化的学习方法，其中学生可以使用教育技术按照对他们最好的节奏进行学习。在线教育系统将会和学生一起学习，从而帮助改进我们对学习流程的理解。而对学习的分析又会反过来加速个性化教育工具的发展。

目前从硬皮书到数字的音频媒体和文本的转变很可能将遍及整个教育领域。数字阅读设备也将变得更加聪明，让学生可以轻松获取他们所学科目的额外信息。机器翻译技术也将使不同教学材料之间的翻译更见简单，同时还能保证相当程度的准确性，就像现在翻译技术手册一样。目前仅依赖于人类译者的教科书翻译服务将会逐渐整合自动方法以提升速度和降低学校系统的服务成本。

在线学习系统也将给成年人和专业人士带来机会，让他们可以强化自己的知识和技能（或重新涉足新的领域），更好地适应这个快速演进的世界。这将会包括完全在线专业学位以及基于在线课程的专业认证的扩张。

更广大的社会成果

自广大人民难以获得教育的国家，如果这些群体有可以获取在线教育的工具，那么在线资源将会产生重要的积极影响。在线教育资源的发展应该能让支持国际教育项目的基金会可以通过提供工具和相对简单的使用培训来更轻松地提供素质教育。比如说，针对 iPad 开发出了大量的、且大部分免费的教育应用。

在消极的一面，现在学生已有把自己的社会接触限制在电子设备上的趋势了，他们在网络程序的互动上花费了大量时间，却没有进行社会接触。如果教育也越来越多地通过网络进行，那么在学生的社会发展阶段缺乏与同龄人有规律的面对面接触会带来怎样的影响呢？特定的技术已经表明这会产生在神经方面的影响。另一方面，自闭症儿童已经开始从与人工智能系统的互动中受益了。

低资源社区

人工智能存在许多机会去改善生活于一个典型北美城市的低资源社区中的人民生活状况——事实上在某些情况下已经有所改变。了解这些人工智能的直接贡献也可能会激发对于发展中国家最为贫穷的地段的潜在贡献。在人工智能的数据收集过程中并没有对这个人群的显著关注，而且传统上人工智能资助者在缺乏商业应用的研究中表现得投资乏力。

有了有针对性的激励和资金优先次序，人工智能技术可以帮助解决低资源社区的需求。萌芽中的努力是有希望的。人工智能可能会有助于对抗失业和其他社会问题带来的恐惧，它或许会提供缓解措施和解决方案，特别是通过受影响的社区以与其建立信任的方式来实现。

机器学习，数据挖掘方法

在「为了实现良好社会的数据科学」大旗下，人工智能被用来建立预测模型，以帮助政府机构更有效地利用有限预算来解决诸如铅中毒的问题，由于在密歇根的弗林特正在上演一桩事件，它成为了新闻中一个主要的公共健康问题。孩子们可能会被测试出血铅水平的升高，但不幸的是这意味着问题只在他们已经中毒之后才被检测到。

使用预测模型的许多努力正在进行中以辅助政府机构优先处理处于危险中的儿童，包括那些可能尚未发病的儿童。同样，伊利诺斯州的人力服务部门（IDHS）为了最大限度地影响产前护理，也使用预测模型来识别有着不良出生结果风险的孕妇。辛辛那提市使用它们来积极地进行识别并将检查人员部署到有代码违规风险的物业中去。

调度、规划

任务分配调度和规划技术已被许多不同组织所运用，用以在食物变质之前把它们分发给食物银行、社区中心和个人，这些食物来自于那些过剩机构，比如餐厅。

运用社交网络与影响最大化进行推理

社会网络可以被利用来创建早期成本较低的、涉及到大量人群的干预措施。例如人工智能可以帮助传播健康相关的信息。洛杉矶有 5000 多个无家可归的年轻人（在 13 到 24 岁之间）。个人干预是困难和昂贵的，而且年轻人对权威的不信任决定了关键信息最好是通过同辈的领导者来进行传播。

人工智能程序或许能够利用无家可归青年的社交网络来战略性的选择同辈领导人去传播与健康有关的信息，比如如何避免艾滋病的传播。这种网络的动态、不确定的特点为人工智能的研究带来了挑战。

同时，也必须注意防止人工智能系统去复制歧视行为，比如机器学习通过非法种族指针去识别人群，或是通过高度相关的替代因素，比如邮政编码。但如果部署得非常谨慎，对人工智能的更多依赖可能会导致歧视的整体减少，因为人工智能程序本质上比人类更容易被审计。

公共安全与防护

城市已经为公共安全和防护部署人工智能技术了。到 2030 年，典型的北美城市将在很大程度上依赖它们。这些措施包括可以检测到指向一个潜在犯罪的异常现象的监控摄像机、无人机和预测警务应用。与大多数问题一样，好处与风险并存。

获得公众信任是至关重要的。虽然会存在一些合理的担心，即与人工智能合作的警务可能会在某些情况下变得霸道或是无处不在，而相反的情况也是可能的。人工智能可能使警务变得更有针对性并只在需要时被使用。而且假设经过仔细的部署，人工智能也可能有助于消除一些人类决策中固有的偏见。

对于人工智能分析学更成功的一个应用是检测白领犯罪，比如信用卡诈骗罪。网络安全（包括垃圾邮件）是一个被广泛关注的问题，而机器学习也对其有所影响。

人工智能工具也可能被证明有助于警察管理犯罪现场或是搜索和救援活动，它可以帮助指挥官排列任务的优先次序以及分配资源，尽管这些工具还没有为这些活动的自动化做好准备。在一般的机器学习尤其是在转换学习中的改进——在新情境中基于与过去情况的相似性而加快学习——可能有利于这样的系统。

相较于防止犯罪的发生，如今几乎部署在世界各地的摄像机更倾向于帮助解决犯罪问题。这是由于从视频中识别事件像素太低以及缺乏人力来查看海量的视频流。随着人工智能对这一领域的改进，它将更好地辅助犯罪预防和起诉，通过更精确的事件分类和有效的视频自动处理来检测异常现象——包括潜在的警察渎职证据。这些改进可能会导致更广泛的监督。一些城市已经增加了用于监视目的的无人驾驶飞机，以及警用无人机来维护港口、机场、沿海地区、航道的安全，工业设施可能会增多，这增加了对隐私、安全等问题的关注。

纽约警察局的 **CompStat** 是第一个指向预测警务的工具，并且许多警察部门目前都在使用它。在犯罪于何时何地更可能发生以及谁可能是罪犯方面，机器学习显著地增强了这种预测能力。正如电影 **Minority Report** 所戏剧化地描述得那样，预测警务工具引发了那些被无理针对的无辜群众挥之不去的恐惧。然而部署得好的人工智能预测工具有可能消除或减少人类的偏见，而不是加强偏见，而且研究和资源应当被用以确保这种效果。

人工智能技术可被用以开发智能模拟，用来培训执法人员之间的合作。而来自不同国家的国际犯罪组织和恐怖分子会相互勾结，那么来自不同国家的警察部队联手对抗他们时就仍有困难。培训国际的执法人员团队之间的协同工作是非常具有挑战性的。欧盟通过 **Horizon 2020** 项目，目前支持诸如 **LawTrain** 项目的尝试。下一步将通过提供支持这种合作的工具继而从模拟向实际调查迈进。

用于扫描推特和其他信息给料来发现某些类型的事件以及它们影响安全的方式，这种工具确实存在。比如人工智能可以在社交网络分析中帮助防止那些由 **ISIS** 或其他暴力团体所激发的风险。试图从社交媒体中检测出破坏性事件的计划，以及在大型聚会中用于分析安全性的监控活动，执法机构对于它们越来越感兴趣。用来确定如何控制人群的人群模拟是很重要的工作。与此同时，对于执法机构过度滥用工具而侵犯人们的隐私的潜在可能性，人们提出了合理的关切。

美国运输安全管理局（**TSA**）、海岸警卫队和其他许多当前依赖人工智能的安全机构，可能会增加这种依赖性以确保效率和有效性的显著改进。人工智能技术——视觉、语音分析、步态分析——可以帮助面试官、审讯人员和保安来检测可能的欺骗和犯罪行为。例如目前 **TSA** 在全国范围内有一个雄心勃勃的项目来重新规划机场安全。该系统叫做 **DARMS**，是基于个人的风险分类和所将乘坐的航班，通过依靠个人信息来定制安全，从而提高机场的效率和有效性。该项目的未来愿景是一个隧道，当人们步行通过它时可以检查人们的安全。再次强调，此技术的开发者应通过使用反映了之前偏见的数据集来小心地避免建筑偏差（例如有关人们的风险水平类别）。

就业和劳资

尽管人工智能很有可能会对典型北美城市的就业和工作场所产生深远的影响，但对当前的影响我们目前还难以作出评估——是积极的还是消极的。在过去十五年，由于经济衰退和日益的全球化，尤其是中国参与到了世界经济中，就业状况已经发生了改变，非人工智能的数字技术也发生了很大的变化。自 1990 年代以来，美国经历了生产率和 **GDP** 的连续增长，但平均收入却停滞不前，就业人口比率也已经下降。

有一些数字技术有重大影响（好的影响或坏的影响）的行业的显著案例，而在一些其它的行业，自动化将很有可能能在不久的将来发生重大的改变。许多这些改变已经得到了「例行的」数字技术的推动，其中包括企业资源规划、网络化、信息处理和搜索。理解这些改变应该能为人工智能影响未来劳动力需求的方式（包括技能需求的改变）提供见解。

到目前为止，数字技术已经给中等技能的工作（比如旅行代理）带来了更大的影响，而不是非常低技能或非常高技能的工作。另一方面，数字系统所能完成的任务的范围正随着人工智能的演进而提升，这很可能会逐渐增大所谓的「例行任务」的范围。人工智能也正向高端的领域蔓延，包括一些机器之前无法执行的专业服务。

为了获得成功，人工智能创新将需要克服可以理解的人们对被边缘化的恐惧。在短期内，人工智能很有可能取代任务，而非工作，同时还将会创造新类型的工作。但新类型的工作比将可能失去的已有工作更难以想象。就业领域的变化通常是渐进的，不会出现剧烈的过渡。

随着人工智能进入工作场所，这很有可能是一个持续的趋势。影响的范围也将扩大，从少量的替代或增强到完全的替代。比如说，尽管大部分律师的工作还没被自动化，但人工智能在法律信息提取和主题建模方面的应用已经自动化了一部分第一年工作的律师新人的工作。在不远的将来，包括放射科医生到卡车司机到园丁等许多类型的工作都可能会受到影响。

人工智能也可能会影响工作场所的大小和位置。许多组织和机构很庞大的原因是他们所执行的功能只能通过增加人力来扩大规模，要么是「横向」扩展地理区域，要么是「纵向」增多管理层级。随着人工智能对许多功能的接管，扩展不再意味着会带来大型的组织。

许多人已经指出一些知名的互联网公司只有很少数量的员工，但其它公司并不是这样。人类企业可能存在一个自然的规模大小，在这样的企业中，CEO 能够认识公司里的每一个人。通过将创造有效地外包给人工智能驱动的劳动力市场，企业会倾向于自然的大小。

人工智能也将创造工作，特别是在某些行业中，通过使某些特定任务更重要，以及通过产生新的交互模型创造新类型的工作。复杂的信息系统可被用于创造新的市场，这往往会带来降低门槛和增加参与的影响——从应用商店到 AirBnB 再到 taskrabbit。人工智能界有一个活跃的研究社区在研究创造新市场和使已有市场更高效地运作的进一步的方式。

尽管工作本身有内在的价值，但大部分人工作是为了购买他们看重的商品和服务。因为人工智能系统可以执行之前需要人力的工作，因此它们可以导致许多商品和服务的成本下降，实实在在地让每个人都更富有。当正如当前的政治辩论中所给出的例子一样，失业对人们的影响比对散布的经济效益的影响更显著——尤其是那些直接受其影响的人；而不幸的是，人工智能常常被视作是工作的威胁，而不是生活水平的提升。

人们甚至在某些方面存在恐惧——害怕人工智能会在短短一代人的时间内迅速取代所有的人类工作，包括那些需要认知和涉及到判断的工作。这种突变是不太可能发生的，但人工智能会逐渐侵入几乎所有就业领域，这需要在计算机可以接管的工作上替换掉人力。

人工智能对认知型人类工作的经济影响将类似于自动化和机器人在制造业工作上对人类的影响。许多中年工人失去了工厂里的高薪工作以及伴随这个工作的家庭和社会中的社会经济地位。长期来看，一个对劳动力的更大影响是失去高薪的「认知型」工作。

随着劳动力在生产部门的重要性的下降（与拥有知识资本相比），大多数市民可能会发现他们的工作的价值不足以为一种社会可以接受的生活标准买单。这些变化将需要政治上的，而非单纯经济上的响应——需要考虑应该配置怎样的社会安全网来保护人们免受经济的大规模结构性转变的影响。如果缺少了缓解政策，这些转变的一小群受益者将成为社会的上层。

短期来看，教育、再训练和发明新的商品和服务可以减轻这些影响。更长期来看，目前的社会安全网可能需要进化成更好的服务于每个人的社会服务，例如医疗和教育或有保障的基本收入。事实上，瑞

士和芬兰等国家已经在积极地考虑这些措施了。

人工智能可能会被认为是一种财富创造的完全不同的机制，每个人都应该从全世界人工智能所生产的财富中分得一部分。对于人工智能技术所创造的经济成果的分配方式，相信不久之后就会开始出现社会争议了。因为传统社会中由孩子支持他们年老的父母，也许我们的人工智能「孩子」也应该支持我们——它们的智能的「父母」。

娱乐

随着过去十五年互联网的爆发式增长，很少有人能想象没有它的生活。在人工智能的驱动下，互联网已经将用户生成的内容作为了信息和娱乐的一个可行的来源。**Facebook** 这样的社交网络现在几乎已经无处不在，而且它们也成为了社会互动和娱乐的个性化渠道——有时候会损害人际交往。**WhatsApp** 和 **Snapchat** 等应用可以让智能手机用户与同伴保持「接触」和分享娱乐和信息源。

在《第二人生》这样的在线社区和《魔兽世界》这样的角色扮演游戏中，人们想象在虚拟世界中有一个虚拟的存在。亚马逊 **Kindle** 这样的专用设备已经重新定义了打发时间的要领。现在只需手指点点划划几下，就可以浏览和获取书籍了；一个口袋大小的设备就可以存储成千上万本书，而阅读体验基本上可手持的纸质书差不多。

现在我们有共享和浏览博客、视频、照片和专题讨论的可信平台，此外还有各种各样用户生成的内容。为了在互联网的规模上运行，这些平台必须依赖现在正被积极开发的技术，其中包括自然语言处理、信息检索、图像处理、众包和机器学习。比如，现在已经开发出了协同过滤（**collaborative filtering**）这样的算法，它可以基于用户的人口统计学细节和浏览历史推荐相关的电影、歌曲或文章。

为了跟上时代的步伐，传统的娱乐资源也已经开始拥抱人工智能。正如书和电影《点球成金》中给出的例子，职业运动现在已经转向了密集的量化分析。除了总体表现统计，赛场上的信号也可以使用先进的传感器和相机进行监控。用于谱曲和识别音轨的软件已经面世。

来自计算机视觉和 **NLP** 的技术已被用于创建舞台表演。即使非专业用户也可以在 **WordsEye** 等平台上练习自己的创造力，这个应用可以根据自然语言文本自动生成 **3D** 场景。人工智能也已经被用于协助艺术品的历史搜索，并在文体学（**stylometry**）得到了广泛的应用，最近还被用在了绘画分析上。

人类对人工智能所驱动的娱乐的热情是很令人惊讶的，但也有人担心这会导致人与人之间的人际交互减少。少数人预言说人们会因为花在屏幕上花费了太多时间而不再与人互动。孩子们常常更愿意在家里快乐地玩他们的设备，而不愿意出去和他们的朋友玩耍。人工智能会使娱乐更加交互式，更加个性化和更有参与感。应该引导一些研究来理解如何利用这些性质为个人和社会利益服务。

想象未来

娱乐任何形式的成功都由作为接收者的个人和社会群体决定。人们觉得有吸引力的娱乐模式是多种多样的，而且还会随时间变化。因此很难准确预测未来十五年娱乐形式会如何演变。尽管如此，目前的趋势表明未来的娱乐形式很可能至少会包含一些特征。

到目前为止，信息革命主要是以软件的形式展开。但是，随着可用的传感器和设备越来越便宜，娱乐系统中也有望出现在硬件上的更大创新。虚拟现实和触觉应用可能会进入我们的客厅——一个个性化伴侣机器人已经处在开发中了。

伴随着自动语音识别领域的进步，我们的研究小组预计与机器人和其它娱乐系统的交互将会变成以对话为基础的形式，也许一开始还会有所限制，但会渐渐更加人性化。同样地，预计交互系统还将发展出新的特点，比如情绪、共情和对每日时间等环境节奏的适应性。

今天，一位有摄影机和可用的软件工具的业务爱好者也可以做出一部相对好的电影。未来，更先进的工具和应用将会使高质量内容的产生更加容易，比如创作音乐或使用虚拟角色编排舞蹈。娱乐产品的创造和创作将受益于 ASR、配音和机器翻译等技术的进步，这将使得内容可以低成本地针对不同的用户进行定制。这种人工智能创造的媒体的民主化和扩散将使得人们的娱乐口味更难以预料，现在已有的动态变化还将继续演进。

随着内容传播的日益数字化和用户偏好和使用特点的大量数据被记录下来，媒体生产者将能够为日益特定细分的人群提供微分析和微服务的内容——一直到能为个人定制。有人相信这种情况会导致传媒集团涌现成为「老大哥」，他们将有能力控制特定个体所接触的思想和网络体验。广大社会是否能发展出阻止它们涌现的措施还有待观察。这个主题和其它与人工智能有关的政策方面的话题将在下一节进行更详细的讨论。

第三部分：人工智能公共政策的前景与建议

人工智能应用的目标必须是对社会有价值。我们的政策建议也会遵循这个目标，而且即便这个报告主要关注的是 2030 年的北美城市，建议依然广泛适用于其他城市，同时不受时间限制。一些提升解读和人工智能系统能力并参与其使用的策略可以帮助建立信任，同时防止重大失败。

在增强和提升人类能力和互动时需要小心，还有避免对不同社会阶层的歧视。要强调多做鼓励这个方向以及沟通公共政策探讨的研究。鉴于美国目前的产业监管，需要新的或重组的法律和政策来应对人工智能可能带来的广泛影响。

政策不需要更多也不要更严，而是应该鼓励有用的创新，生成并转化专业知识，并广泛促进企业与公民对解决这些技术带来的关键社会问题的责任感。长期来看，人工智能将会带来新财富，整个社会也要探讨如何分配人工智能技术带来的经济成果的分配问题。

如今以及未来的人工智能政策

纵观历史，人类改造也适应过很多新技术。这篇报告提出人工智能技术开发和应用的推进需要循序渐进——不是单靠技术本身的瞬间意外的跳跃发展——而且我们会以当下现有的技术与环境为基础，让整个适应过程更容易一些。

另一方面，技术、计算机能力或者数据可用性上的小提升也可能会偶尔带来全新的、革新规则的应用。

人工智能应用成功与否要看它能为人类生活创造多少价值。未来，是否能轻松地使用和适应人工智能应用，也很大程度地决定了人工智能的成功与否。

相反，由于人工智能应用容易出错和失败，所以其成功与否的标志也包括用户如何感知和容忍这些应用的缺陷。随着人工智能逐渐渗入日常生活并应用到越来越多的重要任务中，系统错误可能会导致用户的反弹，并带来信任问题。

例如，即便自动驾驶汽车造成的事故要比人类少很多，但总是会吸引更多的关注。在提升人类理解人工智能系统和决策（诸如明确解释那些决定）和促进人们使用应用上需要策略，这些策略有助于建立信任并防止重大失败。

同样的，开发者应该帮助维护人们对人工智能应用的期望，这会影响他们使用的幸福感和满意度。如果一个系统承诺的某些功能没能实现，会减少人们对系统的信任度，并降低他们未来使用该系统的意愿。

另一个重要的考虑是接管重要任务的人工智能系统会如何影响人类的交际和各方面的能力。由于机器在某些任务上的表现相当于超人，人类的这些能力可能会逐渐消逝。比如，计算器引入课堂后，孩子们的基本算术运算能力降低了。所以，人类和人工智能系统的能力应该是互补的。人类专注于机器做不好的任务，包括复杂的推理和创造性的表达。

然而，孩子们与人工智能应用接触的越来越多，比如与智能手机上的个人助理互动，跟主题公园里的虚拟动画人物玩耍。较早地接触会促进孩子们与人工智能应用的互动，这也会成他们日常生活中自然的一部分。但是，这也会造成年轻一代与上一代人在感知人工智能对社会的影响时产生代沟。

如果人工智能技术（包括驱动这些技术的高性能计算和大规模数据）无法让社会各阶层间都平等受用，那么同样能加剧现有的机会不平等。接触到这些技术的人在效率和能力上都会有提升。一个人如果有了精确的机器翻译技术，那么他就能更好的使用不同语言的学习资源。类似的还有，如果语音翻译技术只能用在英语中，那么不懂英语的人就会处于劣势。

进一步，人工智能应用及其依赖的数据也许会反应出设计者和使用者的偏见，因为数据是他们决定的。这些偏见会加深现有的社会偏见，并无法让各个社会亚群体平等的受益人工智能带来的好处。

例如，一些语音识别技术不能很好的适用于女性和带有口音的人。随着人工智能逐渐渗入到重要的应用中，这些偏见会让社会各团体间的不平等问题浮出水面。另一方面，与记录在案的人类决策偏见相比，基于人工智能的决策工具有望能极大减少重要决策中的偏见，比如对借钱的人或进过监狱的人的偏见。

私人也越来越担心带有人工智能功能的监控，尤其是在监控设施普及的城市。**Sousveillance**，一种通过便携设备记录人参与活动的监控也越来越多。

既然对偏见和隐私的看法是基于个人与社会的道德和价值判断，所以关解决这些问题的争论也会多起来，争论一多，解决方案短时间内也就出不来了。由于人工智能正在产生巨大的价值，关于如何分享人工智能经济成果的争议也会多起来——尤其是驱动应用的人工智能专业知识和基础数据集中在一小批大公司手上。

为了帮助解决个人和社会对快速发展的人工智能技术产生的忧虑，该研究小组提供了三个一般性政策建议。

1. 在所有层级的政府内，制定一个积累人工智能技术专业知识的程序。有效的监管需要更多的能理解并能分析人工智能技术、程序目标以及整体社会价值之间互动的专家。

缺少足够的安全或其他指标方面的专业技术知识，国家或地方政府官员或许或拒绝批准一个非常有前途的应用。或者缺少足够训练的政府官员可能只会简单采纳行业技术专家的说法，批准一个未经充分审查的敏感的应用进入市场。不理解人工智能系统如何与人工行为和社会价值互动，官员们会从错误的角度来评估人工智能对项目目标的影响。

2. 为研究人工智能的平等、安全、隐私和对社会的影响扫清感知到的和实际的障碍。

在一些相关的联邦法律中，如计算机欺诈和滥用法案（Computer Fraud and Abuse Act）和数字千年版权法的反规避条款（the anti-circumvention provision of the Digital Millennium Copyright Act），涉及专有的人工智能系统可能被如何逆向工程以及被学者、记者和其他研究人员评价的内容还很模糊。当人工智能系统带来了一些实质性后果需要被审查和追究责任时，这些法律的研究就非常重要了。

3. 为人工智能社会影响的跨学科研究提供公共和私人资金支持。

从整个社会来看，我们对人工智能技术的社会影响的研究投入不足。资金要投给那些能够从多角度分析人工智能的跨学科团队，研究范围从智能的基础研究到评估安全、隐私和其他人工智能影响的方法。一下是具体问题：

当一辆自动驾驶汽车或智能医疗设备出现失误时，应该由谁来负责？如何防止人工智能应用产生非法歧视？谁来享有人工智能技术带来的效率提升的成果，以及对于那些技能被淘汰的人应该采取什么样的保护？

随着人工智能被越来越广泛和深入地整合到工业和消费产品中，一些领域中需要调整现有的建立监管制度以适应人工智能创新，或者在某些情况下，根据广泛接受的目标和原则，从根本上重新配置监管制度。

在美国，已经通过各种机构将监管具体到各个行业。在设备中使用人工智能实现医疗诊断和治疗由食品药品监督管理局（FDA）监管，包括定义产品类型和指定产生方法，还有软件工程的标准。无人机在管制空域中的使用由美国联邦航空局（FAA）监管。面向消费者的人工智能系统将由联邦贸易委员会（FTC）监管。金融市场使用的人工智能技术，如高频交易，由证券交易委员会（SEC）监管。

除了针对具体行业制定监管的方法外，「重要基础设施」中定义模糊和广泛的监管类别可能适用于人工智能应用。

奥巴马政府的总统政策指令（PPD）将「重要基础设施」广泛地定义为「对美国至关重要的实体或虚

拟的资产、系统和网络，缺少或破坏这些设施将会对安保、国家经济安全、国家公共卫生或安全或其任何组织都会造成破坏性影响。」今天，一家企业仅通过广泛定义的手段就可以完全不受联邦的监管。但是，联邦政策的大趋势是将监管具体到涉及国民经济的 16 个行业。

涉及到人工智能，重要的基础设施主要是由终端应用来定义的，而不是技术活实际生产人工智能软件的环节来定义的。

像谷歌、Facebook 和亚马逊这样的软件公司已经在积极游说监管机构，避免自家公司被认定为对国民经济关键的公司，声称这将打开监管大门，将不可避免地导致他们为加速产品开发周期而妥协创新能力。无论如何，随着公司在创建，运作，并维持重要设施中使用人工智能，他们会逐渐重视对软件的监管。

一些现有的软件安全监管制度（例如，FDA 对帮助医生诊疗疾病的软件的管理）要求特定软件工程达到能实现开发者水平上的实践。然而，组装现代软件系统的部件常常来自多家提供商，这些部件都有相对独立的应用。让这些所有开发者都遵循最重要的罕见应用的标准是不可行的或者不可取的。

允许在安全关键应用中不受管制地使用这些部件也是不可取的。从概念上和实践上权衡促进创新和调节安全比较困难。最低限度是，监管机构需要更多的专业知识来了解标准的含义以及研究人员、政府和行业实施的措施。

政策和法律上的考虑

虽然综合研究人工智能（AI）与法律互动方式超过了这篇报告最初的内容范围，似乎是很清楚的是，作为一种变革性的技术，人工智能已具备挑战任何数量的短期、中期和长期的法律假设。

正是法律和政策将如何适应人工智能的进步——人工智能将如何适应法律和政策所体现的价值观——取决于各种社会、文化、经济和其他影响因素，并且根据不同的司法管辖权，可能会有不同的变化。

美国法律是普通法、联邦、州和地方法规和条例以及监管法规的混合体，监管法规也许与人工智能关系最大。在具体的实施中，人工智能可能牵涉到这些法律的每一个来源。例如，内华达州通过了一项法律，广泛允许自动驾驶汽车，并给出了内华达州机动车辆的工艺要求。

同时，美国国家公路交通安全管理局（National Highway Transportation Safety Administration）规定，自动驾驶车的驾驶员是该车的系统，而不是车里的人。一些汽车设计通过将自动模式设定为只有手在方向盘上（至少每一次都是这样）来回避这个问题，让人掌握对车最终的控制和责任。然而，特斯拉的采用这一策略并没有防止 2016 年发生的第一起涉及自动驾驶汽车的交通事故。而且由于大多数人都是第一次体验到实体的人工智能工具，自动交通会强烈影响公众对人工智能的感知。

当然无人驾驶车只是人工智能在服务、产品和其他情境中应用的众多例子之一。将人工智能技术引入税务咨询将人工智能技术引入税务咨询、股票市场的自动交易，或者自动生成医疗诊断中，其法律效力也会随着监管者和具体应用的不同而不同。当下非技术性领域，包括在预测警务，非歧视性贷款，医疗应用如老年护理和药物输送，以及儿童互动系统的设计（例如，自动教学系统设计中也对平衡处理进化 vs. 智能设计的方面制定了相关法规）和互动娱乐方面也对人工智能应用做出了规定。

鉴于目前美国行政法结构，短期内制定出全面的人工智能政策法规似乎不太可能。但是，可以根据人工智能在各种情境中可能出现的法律和政策问题，广泛列出多个类别。

隐私

个人的私密信息可能会被由人工智能做出的决策和预测泄露出来。虽然人工智能牵连到隐私的方式反映的是计算机和互联网技术，但其他问题或许是人工智能独有的。例如，人工智能基于之前模式预测未来行为的潜力也提出一些具有挑战的问题。一些公司已经在使用机器学习预测信用风险。美国利用复杂算法根据罪犯的细节信息预测累犯的可能性，通过此举来考虑假释的可能性。在这些情况下，确保如种族和性取向这类因素不在人工智能做决策时被考虑进去是一个技术挑战。即使当这些特征没有直接提供给算法时，它们或许依然会与一些如邮政编码这类看似无害的特征强烈的关联起来。无论如何，只有经过仔细的设计，测试，和部署，人工智能算法产生的偏见或许才能比一个典型的人少一些。

与人工智能关联的拟人化界面也出现了隐私的问题。社会科学研究发现人们天生会把拟人化技术当成人类看待。在一项研究中，当计算机第一次说出自己何时被生产出来的时候，受试者也更可能回答出自己的出生日期。在另一个研究中，他们在面对一个拟人化的界面时会跳过敏感的问题。基本层面上的问题是：当社会化的机器人「活在」我们的屋子、汽车、办公室、医院病房和手机中时，人类还会继续享受孤独吗？

创新政策

早期关系到责任和话语的法律和政策决策帮助确保互联网中的商业活力。相比之下，今天的软件产业受到了从企业决策的影响，从开放和自由的软件到更加积极追求知识产权保护，导致了专利丛生。在激励人工智能创新并促进合作与保护第三方不受到伤害之间做出适当的权衡将是核心的挑战。

责任（民事）

由于人工智能有组织地直接影响世界，甚至产生实质性的影响，由人工智能造成的伤害会加倍凸显。人工智能可能会做出设计者意料外的行为，这种情况给侵权法中的普遍假设带来了挑战，法院只会判定补偿可预见的伤害。甚至在以公平和效率为准则能更好的判定责任时，法院也许都将责任判给人类当事人。或者，法院可能拒绝追究责任，因为被告在法庭前没有，并且不能预见的人工智能伤害。责任将默认地落在无辜的受害者身上。当人类当事人对机器行为负的责任越来越少时，产品责任的角色作用——还有属于生产这些产品的公司的责任——可能会增加。

责任（刑事）

如果侵权法认为伤害可以预见，那么刑法会可能会进一步认定这些伤害是蓄意的。美国法律尤其重视犯罪动机的概念——作案的想法。随着人工智能应用参与到了原本是由人类完成的任务中，也就可能构成犯罪，法院和其他法律行为人将不得不困惑：应该基于什么理论让谁对所发生的事情负责？

代理

上面的问题带来了这样的问题：在哪种情况下，人工智能系统可以作为一个人或企业的代理进行运作？美国、加拿大等地方的监管机构将条件设置成当软件可以进入有约束性的合同时。人工智能依法开展的突出活动越多，对法律下的代理原则的挑战就越大。

认证

「人工智能」的概念正是意味着对人类技能和智慧的取代。而且在许多情况下（从驾驶到执行手术到法律案件），一个人类必须在获得一些证书或认证之后才能执行一项特定的任务。因此，法律和政策将不得不一一而且已经一一和如何确定人工智能系统的能力做斗争。比如说，想象一家研发可以自动摘除阑尾的手术平台的机器人公司，或者一家编写可以提供法律咨询应用的法律公司。今天，在法律方面，我们还不清楚人工智能系统应该通过怎样的医学或法律要求，更不要说它们应该再那里去获得这样的认证了。

劳动力

随着人工智能对人类的替代，一些工作将会消失，一些工作又会被创造。其对工作的净影响目前还是未知数，但劳动力市场不太可能会给所有人都带来均等的好处。对特定类型技能或能力的要求将很有可能大幅下降，从而给具备这些技能的人的就业和工资水平带来极大的影响。尽管人工智能对收入的水平和收入分配带来的最终影响是不可避免的，但它们极大地依赖于政府政策、企业所选择的组织工作的方式、以及个人对学习新技能和寻找新类型的工作和收入机会的选择。发现人工智能的发展导致了他们的工作被改变或被终结的人可能会诉诸立法机关和法院。这也许就是 Littler Mendelson 律师事务所（可能是世界上最大的就业法律公司）有一整个处理机器人和人工智能案子的团队的原因。

税务

联邦、州和地方收入来源可能会受到影响。使用人工智能来完成任务可以更快更精确——并能避开就业税。因此，人工智能可以会越来越多地从工资和收入的投资转向资本支出。根据国家对工资和所得税的依赖，这种转变可能是不稳定的。人工智能还可能表现出与人类不同的「习惯」，造成仍然更少的收入来源。由于自动驾驶汽车的程序使它们不能违法，许多市政当局依赖超速和停车罚款的情况将不得不寻找替代。因此，在人工智能的进步开始起效时，想要平衡预算的政府机构可能会通过立法减缓或改变这项技术的进程。

政治

人工智能已经被政治人物以 Bot 的形式用到了操纵和定向的「robocalls（机器人电话）」与社交媒体平台上，以便影响选举。他们能够借此协调抗议或预测抗议，以及通过更精确地定位谁在何时说了什么来推进更大的透明。因此，思考人工智能的行政和监管法律可以为促进更大的民主参与而进行设计，或者说，如果居心不良的话，也可以减少民主参与。

这个列表并不是详尽的，主要集中于美国的国内政策，而没有涉及很多人工智能很有可能涉足的领域。一个可以得到的教训涉及到人工智能以一种背景明确的方式被管理和人工智能技术在社会的多个行业和领域的更广泛的思考之间没有联系。可以在多种背景中积累专业知识和设置人工智能标准可能很有诱惑力。本研究小组的共识是对广义上「人工智能」的监管尝试可能会被误导，因为目前还不存在对人工智能的明确定义（它不是任何东西），而其中的风险和所要考虑的问题在不同的领域也非常不同。

相反，政策制定者应该认识到应该随机应变，不同的时间不同的行业需要对使用人工智能构建的或以某种方式整合了人工智能的软件进行不同的明确的、合适的监管。政府将需要专业知识来审议公共和私人行业开发的标准和技术，以及起草必需的法规。

未来的指导原则

面对人工智能技术将带来的深刻变化，要求「更多」和「更强硬」的监管的压力是不可避免的。对人工智能是什么和不是什么的误解（尤其在这个恐慌易于散布的背景下）可能引发对有益于所有人的技术的反对。那将会是一个悲剧性的错误。扼杀创新或将创新转移到它处的监管方法同样也只会适得其反。

幸运的是，引导当前数字技术的成功监管原则可以给我们带来指导。比如，一项最近公布的多年研究对比了欧洲四个国家和美国的隐私监管，其结果却很反直觉。西班牙和法国这样的有严格的详细法规的国家在企业内部孕育出了一种「合规心态（compliance mentality）」，其影响是抑制创新和强大的隐私保护。

这些公司并不将隐私保护看作是内部责任，也不会拿出专门的员工来促进其业务或制造流程中的隐私保护，也不会参与必需范围之外的隐私倡议或学术研究；这些公司只是将隐私看作是一项要满足规范的行为。他们关注的重点是避免罚款或惩罚，而非主动设计技术和采纳实际技术来保护隐私。

相对地，美国和德国的监管环境是模糊的目标和强硬的透明度要求和有意义的执法的结合，从而在促进公司将隐私看作是他们的责任上做得更加成功。广泛的法律授权鼓励企业发展执行隐私控制的专业人员和流程、参与到外部的利益相关者中并采用他们的做法以实现技术进步。对更大的透明度的要求使民间社会团队和媒体可以变成法庭上和法庭外的公共舆论中的可靠执法者，从而使得隐私问题在公司董事会上更加突出，这又能让他们进一步投资隐私保护。

在人工智能领域也是一样，监管者可以强化涉及内部和外部责任、透明度和专业化的良性循环，而不是定义狭窄的法规。随着人工智能与城市的整合，它将继续挑战对隐私和责任等价值的已有保护。和其它技术一样，人工智能也可以被用于好的或恶意的目的。

这份报告试图同时强调这两方面的可能性。我们急切地需要一场重要的辩论：如何最好地引导人工智能以使之丰富我们的生活和社会，同时还能鼓励这一领域的创新。应该对政策进行评估，看其是否能促进人工智能所带来的益处的发展和平等共享，还是说会将力量和财富集中到少数权贵的手里。而因为我们并不能完美清晰地预测未来的人工智能技术及其所将带来的影响，所以相关政策一定要根据出现的社会难题和线索不断地重新评估。

截至本报告发布时，重要的人工智能相关的进展已经在过去十五年内给北美的城市造成了影响，而未来十五年还将有更大幅度地发展发生。最近的进展很大程度是由于互联网所带来的大型数据集的增长和分析、传感技术的进步和最近的「深度学习」的应用。

未来几年，随着公众在交通和医疗等领域内与人工智能应用的遭遇，它们必须以一种能构建信任和理解的方式引入，同时还要尊重人权和公民权利。在鼓励创新的同时，政策和流程也应该解决得到、隐私和安全方面的影响，而且应该确保人工智能所带来的好处能得到广泛而公正的分配。如果人工智能研究及其应用将会给 2030 年及以后的北美城市生活带来积极的影响，那么这样做就是非常关键的。

附录 I：人工智能简史

本附录主要基于尼尔森的著作撰写而成，该著作从目前流行的角度介绍了数据密集型方法（data intensive methods）和大数据。然而重要的是，这本书并没有证明书中能够解决所有问题。要完整全面地展现该领域的历史，这已经超出了本文的范围。

人工智能领域的研究正式诞生于 1956 年，在 John McCarthy 组织的达特茅斯暑期研究项目的研讨会上提出。当时提出此概念的原因是为了探究机器可以在哪些方面模仿人类智能——这一核心思想一直推动着人工智能领域向前发展。

在与 Marvin Minsky, Nathaniel Rochester, 和 Claude Shannon 共同起草的研讨会提议文件中，McCarthy 因为第一次使用术语「人工智能（artificial intelligence）」而声名大噪。紧接着很多后来者也将自己的著名项目归于人工智能的旗帜之下，这其中包括了 Arthur Samuel, Oliver Selfridge, Ray Solomonoff, Allen Newell, 和 Herbert Simon。

虽然达特茅斯探讨会为这一领域开创了一个有统一身份的专门研究社区，但很多具有人工智能特征的技术想法其实早已存在。在十八世纪，托马斯·贝叶斯（Thomas Bayes）为推理事件的概率提供了一个计算框架。

在十九世纪，乔治·布尔（George Boole）证明逻辑推理——可以追溯到亚里士多德——可以像求解方程组那样被系统地执行。到 20 世纪之交，在实验科学的进步导致了统计学的产生，这使得推断可以严格的从数据中得出。

建造工程并让机器执行指令序列的想法，在一些先锋如查尔斯·巴贝奇（Charles Babbage）的头脑中形成，并于 20 世纪 50 年代成熟，最后导致了第一台电子计算机的问世。可以感知和自主行动的第一代机器人，也是在那个时候建造的。

支撑计算机科学最具影响力的思想来自阿兰·图灵，他提出了通用计算机概念。图灵的经典论文「计算机和智能（Computing Machinery and Intelligence）」，设想建造计算机用于模拟人类智能的可能性，并探讨了许多与现在人工智能有关的话题，包括怎样测试人工智能以及机器怎样自主学习。

虽然这些想法启发了人工智能，图灵却没有将他的想法转化为行动所需的计算资源。在探索人工智能的道路上，几个重点研究领域在 50 年代和 70 年代之间出现了。

Newell 和 Simon 率先进军启发式搜索领域（heuristic search），这是一种在大型组合空间寻求解决办法的有效程序。特别的是，他们将这种思想应用于构建数学定理的证明，方法是——先通过他们的逻辑理论家（Logic Theorist）程序，然后使用普遍问题解决方案（General Problem Solver）。

在计算机视觉的领域，Selfridge 和他的同事们做的早期的字符识别工作，为如面部识别等更复杂的应

奠定了基础。到了六十年代后期，自然语言处理方面的工作也相继展开。斯坦福国际研究院（SRI）建造的，一个叫「Shakey」的有轮机器人，标志着在移动机器人领域探索的开始。

Samuel 的跳棋游戏程序能够通过自我对弈提高下棋水平，这是应用机器学习系统的早期实例之一。一个基于生物神经元的计算模型——Rosenblatt 的感知器（Perceptron）成为人工神经网络领域的基础。Feigenbaum 和其他人主张建立专家系统，即为特定领域——例如化学和医疗诊断领域——量身定制的知识资料库。

早期概念上的进步假定了一个符号体系（symbolic system），并可以在此基础上推理和演化。但是到了 80 年代，尽管符号体系在发展不同人工智能领域方面充满前途，它仍然没有值得夸耀的显著实际成就。理论与实践之间的这种差距，部分源于人工智能社区对可直接获取环境信号和数据的基础物理系统的重视不足。

同时，过分强调布尔逻辑（真/假），以及忽视量化不确定性的需要，也是原因之一。在 80 年代中期，由于公众对人工智能的兴趣降低和研究经费的短缺，这个领域被迫承认这些不足之处。Nilsson 将这个时期成为「人工智能的冬天」。

在九十年代，出现了一种被称为「良好的老式人工智能（Good Old- Fashioned AI）」的思想复苏，但是，这种急需的复古思潮作为一个端到端的方法来构建智能化系统是不够的。

相反，智能系统需要从头开始构建，在任何时候都可以解决手头的任务，尽管熟练程度不同。技术进步也使得建立一个由真实世界数据驱动的任务变的可行。更便宜、更可靠的硬件也使得具有传感和驱动的机器人更容易被制造。

此外，互联网具有收集海量数据，以及计算、存储并处理这些数据的能力，使人们能应用统计技术从数据中获得解决方案。正如我们在第二部分详细介绍的那样，这些发展历程使得人工智能能够在过去的二十年中出现，并如今对我们的日常生活产生了深远的影响。

总之，下面列举了人工智能的传统子领域。正如在第二部分描述的那样，他们中的一些因为各种原因比其他的要火热。但是既不能忽视其他方面历史重要性，也不能说他们在未来不可能成为热门领域。

搜索与规划处理目标导向行为的推断。搜索扮演着关键角色，例如，在国际象棋程序中（比如深蓝），决定了如何移动（行为）可以最后获得胜利（目标）。

知识表征与推断的领域需要将信息（往往是大量信息）处理成结构化的形式，便于可靠高效地查询。IBM 的 Waston 项目以一个可以组织、索引、以及获取不同信源信息的有效方案，在 2011 年打败了人类竞争者，赢得了 Jeopardy 挑战赛。

机器学习是一个能使系统通过观察相关数据自动改善其在一个任务中的表现的范式。机器学习在过去十年确实成为了人工智能浪潮中的关键贡献者，从搜索和产品推荐引擎，到语音识别、欺诈检测、图像理解系统以及无数曾经依赖于人类技巧与判断的任务。这些任务的自动化扩大了电子商业这类服务行业的规模。

构建越来越多的智能系统之后，就很自然地需要考虑这样一个问题：这样的系统如何和其他系统互动。多代理系统（Multi-Agent System）领域考虑了这个问题，它在网上市场和运输系统中变得日益重要起来。

从早期开始，人工智能设计并构建出能够融入真实世界的系统。机器人领域研究感知和行动的基础层面——特别是二者的整合，这些研究能使机器人做出有效的动作。自从机器人和其他计算机系统与人类分享真实世界以来，人机交互（Human Robot Interaction）专业也在最近几十年变得重要起来。

机器感知总是在人工智能中扮演核心角色，部分用于开发机器人，但也作为一门完整独立的研究领域。最普遍的感知方式研究是计算机视觉与自然语言处理，它们都有着庞大的、充满生机的社区。

如今人工智能的其他几个焦点领域是互联网发展的结果。社交网络分析研究邻里关系影响个体与社群行为的效果。众包是另一个解决问题的创新技巧，依靠人类（往往成千上万的人）的智能解决困难的计算问题。

即使将人工智能划分成几个子领域，能使深度技术沿着几条不同的方向前进，在任何合理尺度上合成智能还是需要整合许多不同的想法。例如，最近打败人类围棋冠军的 AlphaGo 使用了多种机器学习的算法来训练自己，同时也在下棋的过程使用复杂巧妙的搜索步骤。

