

许成钢：人工智能、工业革命与制度

当代经济学基金会 3 天前

随着大数据和人工智能的发展，国内相关人士引发了一场关于是否可以重振计划经济的争论。许成钢的文章从大数据产生的制度基础、数据和机器学习的性质以及人类有限理性入手，对这一争论给出了否定的回答。首先，从制度角度来看，大数据来自市场，如果消灭了市场，不仅数据无从产生，还会严重损害激励机制，导致创新无法产生。其次，从大数据本身来看，有硬数据和软数据的区别，软数据是不可传递无法度量的，这就导致大数据不可能完备；从机器学习的角度看，有热识别和冷识别之分，而热识别是机器无法学习的。再次，人类的理性是有限的，因此人的认识和推理是片面的、不完备的，甚至可能是矛盾的，这就决定了，模仿人类进行推理和规划的机器人也不可能比自然人做得更好。由此看来，建立一个前所未有的新式计划经济是不可能的。本推送共计 10500 字，有删节，全文刊载于《比较》第 95 辑。财新网全文链接为：<http://opinion.caixin.com/2018-06-04/101262161.html>

2017 年 7 月 8 日，国务院发布《新一代人工智能发展规划》。9 月初，俄罗斯总统普京在谈及人工智能时表示，人工智能技术发展引发了“难以预测的巨大机遇和威胁”。“谁能成为这一领域的领先者，谁就是未来世界的统治者。”在普京的眼里，人工智能不仅是重大战略机遇，而且是俄罗斯存亡的关键。

迄今为止，在人类历史上，所有技术革命，所有创新性、革命性技术的产生和重大发展，都是从市场上产生出来的，都是以私有制为基础的企业家们创造出来的，都是市场优胜劣汰的最后结果。大力推动人工智能，这个大方向是非常正确的，但是我们要关心的是怎么做：是用计划的方法做吗？历史的教训和经济学的道理告诉我们，计划的方法做不出来。计划的方法充其量就是跟在市场的后面走，自己走不远。

此外，在关于人工智能和制度的辩论中，有一种重要观点：人工智能可能是新一代计划经济的技术基础。例如，人工智能很重要的一个应用是在金融领域。如果机器可以决定投资，那是否意味着部分计划经济？这涉及一个非常重要的基本问题：机器收集和处理的信息来源是什么？计划什么是有效率的？计划什么是低效率的？计划的界限在哪里？

下面我首先要解释，中国发展人工智能面临的最重要问题是制度，而不是一揽子技术问题。因为国民经济的整体状态的重要性超过一两项技术是不是全球领先。为了讨论这个问题，让我们先来看一些数据。

2016 年麦肯锡做了一次调研，发现中国的劳动生产率仅为经合组织（OECD）国家平均值的 15%—30%，也就是说中国的劳动生产率比发达国家低很多倍。这意味着中国普遍较为落后。这不是在最前沿的个别领域上领先能自动解决的问题。

以上我们看到的是劳动生产率低下。在比较完善的市场经济中，低劳动生产率伴随着低劳动力成本。但是，许多文献表明，中国的劳动力成本反常地高。经济学人智库的分析显示，2016 年中国单位劳动力成本高于美国和西欧。牛津经济研究院的调研结果表明，2016 年中国劳动力成本仅比美国低 4%。另一个计算方式是按照单位劳动生产率，也就是说，如果我们在中国和美国制造同样的产品，中国的劳动力成本比美国贵。

劳动力成本非常贵，是不是意味着中国的工人拿钱太多了呢？并非如此。中国的家庭收入只占中国 GDP 的三分之一略强，为世界最低之一。世界上大多数其他国家，家庭收入占本国 GDP 的比例在一半以上。换言之，中国的工人并没有拿钱太多。

劳动力这么贵，工人却没多拿钱，成本压在企业身上，企业也没拿到，因此是国家拿走了。政府的财政收入增长率连续二十多年超过 GDP 增长率，如今仍然如此。所以，政府收入占本国 GDP 的份额是世界最高之一，这导致劳动者和企业承担了巨大的制度成本。

此外，在中国，由于制度原因，最盈利的行业是银行、房地产，而不是与研发密切相关的行业。作为对比，在美国，最盈利的行业是与研发密切相关的行业。

因此，不解决制度问题，仅仅把力量集中在最前沿技术的所谓产业升级上，即便在个别产品个别行业得到产业升级的结果，也解决不了普遍的经济问题。而且，即便是专门讨论产业升级，在不解决制度问题的前提下，也只能解决个别产品问题。

工业革命的负面教训

一个流行的看法是，大数据和人工智能在引领新一轮工业革命。这一看法是有道理的。在人类历史上，国家之间拉开巨大的距离产生于工业革命。新一轮工业革命会进一步拉开国家之间的距离。

但是，工业革命的影响并非在所有方面都是正面的。我们也需要非常冷静地回顾过去历次工业革命带来的负面教训，这些教训都来自对新技术的滥用和高估。与过去的工业革命相关的最大负面教训，就是大规模的国有制和中央计划经济。这个想法最早产生于两次工业革命的过程。一些极左翼知识分子误认为人类有能力认识一切，误认为人类有能力掌控社会所有的方方面面，进而误认为人类有能力知道所有人的福利是什么，有能力安排所有人的福利，有能力帮助人类安排所有的技术变化和生产。因此，以私有制为基础的市场经济最终也会必须被以国有制为基础的中央计划取代。以国有制为基础的中央计划经济影响了全球三分之一的人口。虽然过去 80 年里，技术和市场的运作发生了翻天覆地的变化，哈耶克在 20 世纪三四十年的相关讨论仍然有效。

与工业革命相伴随的对技术的滥用和误解还有另一个重大负面作用，那就是对环境的破坏。举两个例子：第一个例子是巨大的水利工程。人们以为自己有能力规划河道、湖泊、陆地，可以通过建坝的方式建造前所未有的巨大人工湖，直到发生一系列灾难性后果，才意识到有大量事情是无法计划和规划的。二三十年前，国与国之间达成共识，对建造大坝极其谨慎，原则上不支持建造巨型大坝。任何大坝在建设之前，第一重要的是评价它们对环境的影响。

第二个例子是，碳排放导致全球变暖。这是过去历次工业革命、技术发展、经济发展完全没有意识到的问题。

现在大数据和人工智能来了，我们应当吸取以前的教训，关注滥用大数据、滥用人智能带来的危险。当前，有一些科学家已经发出警告，但是他们大部分讨论的是基本的道德问

题，机器人会不会残杀人类，我更想强调的是另外一类问题：技术与制度的关系，市场和计划之间的关系。

当人们误认为科学技术万能，误认为某种万能的技术可以取代甚至颠覆人类文明千年积累下来的制度和人们对制度的认识，灾难就会接踵而至。比如一些政府或者垄断性的大公司，试图利用他们掌握的大数据控制社会、取代市场，这不仅会从根本上阻碍技术进步，更可能给社会带来巨大灾难。

大数据、人工智能作为新兴工业革命的基础

在没有能力处理的情况下，大数据本身并没有太大的重要性，其重要性在于它是一部分人工智能发展的基础。除了中国、美国以及俄罗斯政府外，世界上几乎所有的发达国家也都高度关注人工智能的发展，其原因就是现在各国政府已经清楚地意识到，人工智能是一场全新的、目前处在起步阶段的第四次工业革命的基础或者核心部分。这场工业革命引发的后果包括无人工厂、无人服务，可能是前所未有的超高效率，同时也可能引发大量失业。

在前三次工业革命中，人类最重要的资源是原材料，是能源，也就是地球上的自然资源。随着人工智能的发展，大数据开始变成了一种基本资源。一大类的人工智能用数据训练机器。因此，有数据才有训练，才有智能；没有数据就没有训练，就没有智能。数据越多，量越大，资源也就越大。于是，大数据逐渐成为和原材料、能源并列的基本资源，但是这一资源本质上不同于其他的基本资源。在此之前的资源大多是天然的，而大数据是人造的，自然资源越用越少，数据却越用越多。

人造的资源就意味着这一资源的性质、质量和数量都取决于其被制造的环境，尤其是相关的制度安排。例如，在许多意见不能表达的制度下，大数据的性质就是这些表达的缺失。而在意见能自由表达的制度下，收集到的就是所有意见表达的数据。比如“大跃进”时期大饥荒产生的一个基本原因就是信息无法传递，这不是技术问题，而是制度问题。所以，大数据不仅受制度约束，更取决于制度。

在人工智能中，所谓智能的核心部分是算法。早在 20 世纪 50 年代，科学家就开始了这方面的探索。“人工智能”这个词、算法和其他想法，甚至一些指导性意见是几个创始人在 1956 年的一次会议上确定下来的。其中，重要的奠基人之一是经济学家赫伯特·西蒙教授（1916—2001），1978 年诺贝尔经济学奖得主，他既是经济学教授，也是计算机和心理学教授。确切地说，人工智能从诞生之日起，就是这三个领域的联合产物。

在各种算法当中，如今应用方面发展得最突出的，是所谓的人工神经网络模型。这种模型使机器可以在人的指导下学习，包括所谓的深度学习。人工智能另外一个普遍使用和探索的方法是统计算法。无论使用深度学习还是统计算法，都必须有大量的数据。这就是为什么大数据是人工智能的基础。

人工智能的第二个基础是计算能力。在过去的半个世纪里，计算速度和存储能力基本上每两年提高一倍。现在，超强的计算能力使得无论基于哪一种算法的人工智能，在某些领域里机器都可以超过人类，而且远远超过人类。

以上概括了为什么人工智能的基础是大数据。那么，大数据本身的技术基础是什么？这是一个关键问题，因为弄清楚大数据的技术基础，能帮助我们了解人工智能可以做什么、不可以做什么的界限。

大数据的核心是收集、传输、存储和处理所有可度量的数据。度量的技术基础是传感器（包括录音机、摄像机等）和移动设备。通过这些设备检测到具体数据，通过互联网、物联网传送和集中，形成大数据。这里，传感器的“可度量”是大数据的关键所在，它决定了人工智能可以做什么、不可以做什么。另一类大数据是利用历史上积累的大量文献，其中包括各个学科积累的文献，比如图书馆里的文字、图形、音频、视频记录，供机器学习分析。

所有可以通过传感器和移动设备度量的数据，都可以收集、传输和存储，但可度量是非常重要的技术条件。任何不可度量的对象，都无法变成数据，机器也就无法处理。

在应用方面，目前人们认为最有前途的是所谓深度学习，用大数据训练机器，让它产生识别、推理和规划的能力。实际上，从开始创造人工智能学科起，人工智能就包括规划。在英语里，规划和计划是同一个单词 **planning**。从事工程的人和其他非经济学者，往往会把这两个概念混作一团，把规划的概念混作为计划经济。

人工智能就是用机器模仿人的推理和规划，在科学领域里，研究人的推理和规划的学科叫作决策理论（**decision theory**），它在很大程度上是经济学的一个部分。从这个角度讲，人工智能和经济学讨论的对象是相似的。但是，经济学讨论的是人的推理和规划，人工智能讨论的是让机器模仿人的推理和规划。

人工智能和大数据的科学技术基础

所谓深度学习是一种算法。这种算法从刚一产生就和经济学的决策理论紧密联系在一起。首先，设计者要为人工智能或者机器人规定其目标，在经济学和人工智能里都称之为目标函数，也就是机器人要干什么，“活”在世上图什么。从分析的角度看，在追求其自身的“目标”方面，机器人与经济学家讨论的自然人的行为相似，都是在一系列约束条件下，寻找并执行达到其目标最大化的途径。

为了认清人工智能，我们必须区分人的智能和人工智能。我们必须清楚地知道，关于人的智能，经济学家对此的认识极其有限。这本身就决定人工智能的界限。经济学家可以把人的目标抽象地写成某个函数，但并不真的知道人的真实具体目标是什么，什么因素影响其目标。比如，人的目标是为了愉快或者满意。但是否满意，多么满意，什么因素影响了满意的程度，这些只有每个人自己知道，经济学家不知道，任何计划者也不知道。因此，对于经济问题，人们通过市场可以找到使自己满意的安排。相反，在没有市场的计划经济里，由于无法得知人们的基本信息，也就无法设计出使人们满意的结果。因此，哈耶克的推理完全没有过时。对于任何能运作的机器人，设计者必须确切规定机器人的目标。表面上看人们对机器人的目标知道得很确切，实际上，这个“知道得很确切”本身就意味着机器人与自然人的基本差距。

此外，人的智能产生于人的生理和心理感知，以及人收集的信息。早在 20 世纪 50 年代，赫伯特·西蒙教授在讨论人工智能时，就提出了“识别”这个概念，它是今天人工智能的核

心。早在那个时代，人们就已经把识别区分为“冷识别”和“热识别”，并有过激烈的辩论。所谓“冷识别”就是机器能够识别的，而所谓“热识别”是人带着感情的识别，是机器学不来的。另一对相关的基本概念是硬数据和软数据。所有可度量可传递的数据就是硬数据。和硬数据对立的是软数据，即无法用机器度量、无法传递的数据。当我们讨论人工智能是基于大数据训练出来的时候，必须要知道存在热识别问题和软数据问题。

另一个问题是，人的智能里有一个基本部分——直觉。直觉是基于人对硬数据和软数据、冷识别和热识别的综合，以及经济学家无法确切描述的目标而产生出来的一种高度抽象的、非逻辑的理解。它依赖的数据不但不可度量、不可传递，而且如爱因斯坦所述，直觉本身是无法描述的。

可见，人的智能和人工智能之间有一道鸿沟：人工智能的基础是人造的确切目标，依赖的是可度量、可描述、可传递的数据。但生物科学告诉我们，有一系列的人的生理感知是无法度量的，比如嗅觉、味觉、性欲，等等。在没有数据的领域，人工智能的界限清楚可见。

下面让我举一个简单的例子，嗅觉。酒是一个大产业，好的品酒师很贵。而且，可靠的品酒需要请一群品酒师。为什么？因为我们无法度量每个人的舌头对酒的感觉是什么。生物科学的原理告诉我们，嗅觉无法度量，即传感器做不出来。跟嗅觉的道理相似，气味也是无法度量的，也无法制造气味的传感器。此外，人的最基本欲望之一——性欲，在科学上，其程度也是不可度量的。所以，科学决定了人工智能或机器人只能是没有嗅觉、味觉，没有性欲的，也就是说，一定没有人的最基本的欲望和最基本的感觉。

除了上述的一部分直接的生理内容，还有大量的心理感觉，它们更复杂，而且都不能度量，比如喜爱、喜悦、厌烦、痛苦、抑郁、贪婪、野心，从好到坏，从有趣到无趣，等等。在生理和心理内容不能度量的情况下，机器人只能通过收集的数据，在某些方面模仿或推测人在一定场景下的偏好，这与人的真实偏好存在质的差别。

如上所述，在无法靠度量得到人类偏好的原始数据时，对人工智能的社会训练通常使用的方法是借助文字刻画某些方面的数据的内容，或者观察某个行为，比如把人们在各种情况下的行为录下来，然后用机器识别各种各样的人类行为。运用这样的方法，今天我们看到，在一些具体领域里，机器人呈现出令人吃惊的模仿和推测能力。但是这些“学习”来的模仿和推测是有局限的、静态的，学习的是过去。只要场景变化，这些模仿和推测都会失效。

总之，深度学习最终很难涉及人类智能的原始基本要素，因为人类大量的基本感知无法度量。而没有这些基本感知，机器就不可能通过学习产生和人相似的感觉。到目前为止，设计任何人工智能的机器设备或者机器人，第一重要的是它的目的是什么。由于上述原因，机器人的目的无法通过深度学习获得，而只能由设计它的人分配给它。这个目的在技术上被称为目标函数。

换言之，迄今为止机器人只能是由设计者为它设置有限范围的目标函数，而设置的目标函数必须定义很窄，而且限于短期和静态。经济学和社会科学告诉我们，人们在科学上无法确切描述人的普遍目标。因为人类无法确知面对不同事情不同场景时的目标变化。于是，人们无法为机器规定普遍的目标函数。同时，因为机器人没有人的基本感知。于是，人们也无法指导人工智能，通过学习产生与人相似的普遍的目标函数。总之，设计者分配给机器的目

标函数不是人的普遍目标，也不能代替人的目标函数。人们可能会以为，人工智能或机器人可以读报纸，读杂志，读手机上的微信，并据此进行推测。但是，大数据和人工智能能够达到的水平是受到其依赖的科学原理限制的。哪怕是在一个完全自由、开放、安全的社会里，人类的文字表达和语言表达也只是人类复杂的方方面面感觉的一个部分。从文字、语言、图像的大数据里推断出来的内容，是片面的，有歧义的，依赖于特定场景的。

人工智能创始人、诺贝尔经济学奖得主西蒙当年提出了“有限理性”这个概念。这是至今影响经济学发展的最基本概念之一。“有限理性”承认人的认识和推理的片面性、不完备性，甚至自相矛盾。有限理性决定了机器人执行人分配给它的目标函数，不可能比自然人做得更好。

总之，由于感知的不可度量性，我们没有办法训练机器人本身产生它自身的偏好或动态的目标函数。同时由于有限理性，我们也无法为机器设定普遍的目标函数。所以，深度学习的人工智能只限于模仿人在已知环境（场景）里的行为。因为它赖以学习的是已知环境中收集的数据，一旦脱离了训练它的环境，没有原始动力的人工智能就丧失了基本能力。

人工智能革命将集中在可明确规定目标的任务方面

一旦有了清楚规定的目标，通过机器学习，人工智能的计算和存储能力可以大大超过人类。比如下棋，无论多复杂的棋，目标都很简单，就是赢对方。与此相似，开车、开飞机、开船，目标都同样简单，就是安全到达目的地。为此，需要知道目的地，需要知道路线，不能撞人，等等。还有，根据规定的目标，设计机器人制造产品，包括成体系地制造产品，达到制造业的全面自动化。当工厂里只需要很少的人或者没有人时，经济学和管理学讨论的企业内部的激励机制问题就消失了，因为激励机制问题是人的问题，而机器没有自己的激励机制问题。就此而言，这会是一场真正的革命。

在一些领域实现部分或全面自动化时，一些职业可能被人工智能代替。其中包括助教、助研、金融分析助理、医生助理、律师助理、军事助理，等等。最近几年美国的许多机构做过许多有趣的实验。例如，在美国很好的大学里，IBM 的机器人“沃森”（Watson）与真人一起做同一门课程的经济学助教。在只用网上辅导和回答问题的条件下，一个学期之后，学生分不出哪个是真人助教哪个是机器人助教。机器人助教与真人助教相似，可以聊天、开玩笑、解题等，但是面对真正的挑战性问题时，它必须依赖教授决定。也就是说，大主意还需要人拿。

依据已经取得的成就，最乐观的预测是人工智能最终可能基本可以自动进行规划和执行企业与军事方面的多数任务。其中企业方面包括制造业、服务业、物流业，等等。在军事方面，俄罗斯总统普京曾说将来的空战将被无人的机器统治。他说的可能是对的，因为空战的目标具体、简单，与下棋极其相似。

但是，人工智能不能做的是什么呢？国民经济以及超出企业的经济任务它不能做。原因在于，无人企业的目标简单，就是利润。但是，国民经济的目标函数是国民福利。利润与国民福利有本质不同。与之相似，军事任务的目标简单，就是打胜仗；而经济任务的目标是追求国民福利，这两者本质上不同。

国民经济不仅仅是工厂，也不仅仅是提供服务。国民经济是满足每一个人的需求。经济学上讲的术语福利，指的是社会上所有个人的利益的总和。上文已论证，人工智能的机器人无法知道社会上所有人的利益是怎么回事。所以，面对人和人之间的关系时，需要人来解决。人可以用机器帮助分析，但机器不能解决人之间的问题。

有人说现在大数据如此发达，它把每一个人的情况都摸透了，每一个消费者的情况都摸透了，用如此发达的大数据是不是就可以帮助我们知道社会福利函数，从而建立一个前所未有的新式计划经济？既然摸透了所有人的利益，可以做出完美的计划，市场经济的各种弊病也就能解决了。但是，这样的论述严重地自相矛盾。

消灭市场就破坏了人工智能计划经济赖以工作的基础

上文在讨论人工智能的时候，已经谈到，使人工智能可以工作的第一步是收集和处理数据。数据从哪里来？是从市场上收集来的。如果把市场消灭了，数据来源也就不复存在了。如果从今天起，把全世界的技术冻结，只重复生产以前生产过的产品，在这个背景下，用计划经济解决生产和分配问题，是可以做到的。但是，我们现在讨论的是技术不断变化的世界，是不断有新产品发明的社会。在此条件下，如果机器人取代了市场，就意味着社会从此进入了静态。任何高度集权的机器都不可能跳过市场，仅凭人工智能掌握消费者对新产品的需求行为。也就是说，消灭市场的那一天就是消灭数据的那一天。这意味着，试图用人工智能建立计划经济与人工智能得以发展的基础自相矛盾。

所谓高度集权的机器是指搜集一切数据的政府或大公司，比如阿里巴巴、腾讯收集的消费者在市场上的选择行为数据。在市场制度下，基于这些数据的人工智能将大显身手，会帮助市场有效运作。但必须注意的是，在市场收集到的消费者行为数据离开市场即失效。消灭了市场，这些数据就不再反映消费者需求，除非技术、商业都不再变化。因此消灭或削弱市场就破坏了人工智能赖以工作的基础。人的需求产生于人的生理和心理感知。实际上，任何一个人并不事先知道自己对已有产品的需求，而是见到它以后才有反应。人的需求是高度动态和高度不确定的。因此，不可能单纯依赖收集到的大量数据，计算出将来的产品，甚至是不存在的产品的需求。离开了市场的计划，必然失败。

更重要的是，如果用人工智能的计划经济取代市场，这个经济无法解决创新问题。创新涉及严重的激励机制问题。发明者、创新者往往都是缺乏资金的人，他们的资金来自市场上的风险投资。为什么是风险投资，而不是中央计划，不是政府，不是银行，也不是证券市场？因为没有人能知道那些人的发明和创新意味着什么，没有人能承担创新带来的风险。风险资本是市场制度的一部分。此外，做风险投资的人有高度的直觉，是有观察力、有经验、有判断力的天才人物。

是否有可能训练一个机器人，通过深度学习，达到比风险投资人还要厉害的程度？这是不可能的。为什么？因为投资人做判断的时候，有大量直觉产生于人的心理，包括他自己的心理和他对企业家心理的判断。由于这类数据也属于上文提到的“软数据”，而软数据的意思就是有信息、没数据、不能表达，所以这些数据无法度量。

上文我们已经讨论过，人类的智力活动涉及大量直觉。在任何领域中，最有创造性的内容都是直觉。一个人是不是天才，关键在于有没有直觉。技术能力是可以教授、训练出来的；

而直觉无法度量，无法训练教授，也无法传递。正是因为直觉无法度量，没有数据，不能表达，所以我们无法训练机器人产生直觉。试图跳过市场，想象将来的技术高度发达，可以取代市场，这是自相矛盾，根本不可能的。

离开了私有产权、市场、自由的环境，单靠人工智能无法解决创新中的激励机制问题。任何高度集权的机器都不可能跳过市场，仅凭人工智能解决创新创业中的投资问题。因为创新创业的想法产生于个人的创造性、创造能力、心理和生理感知，具有高度动态和不确定的特点，而且核心的内容不可度量（其中只有极其有限的方面可以度量，可以收集数据，比如风投、私募收集的数据，创新创业者在市场上的表现，等等）。

制度对创新的影响，在人工智能领域与在所有的创新领域一样。全球人工智能领域的人才集中产生在私有产权、市场、自由的环境中。领英发布的《全球人工智能领域人才报告》显示，截至 2017 年第一季度，全球人工智能领域的技术人才 190 万，其中中国 5 万（10 年从业者占 38.7%），美国 85 万（10 年从业者占 71.5%）。

制度深刻影响大数据、人工智能的发展

人工智能可以发展到什么程度，取决于传感器、算法和数据。但是，数据本身深受制度影响。如果制度能够充分保护人的言论自由和隐私权，人们的表达就会比较充分。但是，一旦人们知道他的隐私权不受保护，表达就会留有余地或者会自我保护。于是大量数据是得不到的，即使能得到也是不真实的。于是，同样的技术在不同场景得到的数据可以大不相同。

侵犯隐私权的主体有两大类：一类是政府，一类是具有垄断地位的大公司。我们知道美国联邦政府出于反恐的需要从谷歌和苹果拿数据，但这两家公司坚决抵制。如果站在反恐的一面，你会认为这两家公司没有道理，但是它们坚持的是最基本原则，关系到全社会对使用数据和信息的态度。站在公司的角度，能够保护个人隐私的是好公司，将来这个行业才有可能发展。从全社会的角度来说，只有保护隐私，整个社会的运作才是正常的，这是整个社会的底线。如果政府控制了数据，那么后果就是政府与公民的关系、政府与企业的关系、政府与社会的关系全部变成另外一种情况。2018 年 3 月下旬媒体披露，英国剑桥分析公司在 2016 年通过脸书盗窃和滥用五千多万个人的数据，有效地帮助了美国总统特朗普成功竞选，这一丑闻揭示大数据、人工智能是把双刃剑，可以对社会造成巨大破坏。这和过去的技术发展是把双刃剑的道理是一样的。大数据的发展不一定在所有方面都是有利于社会的，取决于谁在用，用在什么地方，用来做什么，怎么用，等等。

如果政府掌握了数据，利用它控制社会，很可能损害社会福利。另一面，如果具有垄断性质的企业控制了数据，不受限制地利用数据，也可能对社会造成巨大的伤害。

现在欧盟国家正在试图起诉并限制谷歌和脸书，原因就是担心这种垄断性质的企业拿到数据之后，会不会造成重大负面后果。因为人工智能还在发展中，如果不未雨绸缪，一旦人工智能在数据的基础上发展出对社会的重大负面应用，那么整个社会经济的竞争状态将发生变化。如果人工智能的能力集中在一两个有垄断能力的企业手里，整个社会的状态就会改变。市场经济的竞争也就不再是通常意义上的竞争了。

另一方面，人工智能的发展也深受制度的制约。人工智能可以走多远，也严重受到制度

的制约。人工智能的发展会引起新一轮大规模的自动化，比如无人工厂、无人车队、无人飞机队。但是，它们会不会在社会中普遍出现？这不仅是一个纯技术问题，也是一个社会问题，更是一个制度问题。因为当它出现之后，会淘汰一大批企业，有大量的人会失业。

比尔·盖茨曾经做过有关讨论，他提出了一个有趣的想法，向机器人征税。他说把每一个机器人当成一个人，你雇用了机器人，机器人就需要付所得税，用所得税补贴给失业工人。作为一个原则，他是对的。不补贴失业者，自动化一定会引起社会冲突，导致社会混乱。

自动化不但不一定能改进社会福利，而且可能破坏社会福利。大量的人失业，严重的贫富不均，这种不平等会导致尖锐的社会矛盾。而这种不平等是因为在技术发展之后，制度没有相应地跟上。

另一方面，制度影响国家的基础科学、社会科学、人文科学。如果一个国家没有自由，基础科学发展会受到阻碍。在这种情况下，数据收集会受到约束或扭曲，科学探索包括人工智能的研究也都会受到约束。因此，任何对科学研究和教学自由的限制都会压制人工智能的发展。

政府有可能按照他自己的目的，支持人工智能发展的某一面，压制人工智能发展的另一面。政府以为这样做对人工智能有好处，但是长远来看，会严重影响人工智能的均衡发展。因为许多问题只有均衡发展的时候，才能最终得到解决。

今天中国出现的不平衡加剧，就是和政府有意支持某个学科，压制其他学科直接相关的。经济学就是一个好的例子。在改革开放早期，我们这一代人去美国留学的时候，中国实质上没有作为社会科学的经济学科。我们这一代经济学者，基本上都是工程、物理、数学背景的。过去压制社会科学的做法，使得中国的政治学、经济学、社会学很落后。现在这些学科追上了一些，但仍然很落后。

总之，制度影响大数据和人工智能的发展，同时大数据和人工智能自身的发展又影响制度。当人工智能普遍应用之后，全社会的生产方式会发生变化，资源配置也会发生变化，这反过来会影响制度。人工智能会影响全社会的效率，影响产业结构，由此影响制度。因此，任何国家，如果在大数据和人工智能方面落后，也会导致它的制度跟着更落后。