

# 致毁知识与科技伦理失灵：科技危机及其引发的智业革命

刘益东

(中国科学院自然科学史研究所, 北京 100190)

**摘要:**以致毁知识不可逆增长为核心的科技危机是人类面临的最大的危机与挑战。化解科技危机,当前既有科技发展的乐观派、认识误区根深蒂固、科技发展模式的内外缺陷等难点,更有科技伦理失灵、“囚车剑魔”(PTSR)四大困境、科技冷战与军备竞赛等制约因素。在知识极易扩散的互联网时代,科技伦理因不能约束所有研发活动而失灵,目前兴起的基因技术伦理、人工智能伦理等尖端科技伦理难以奏效。化解和应对科技危机与科技巨风险,必须大力推进以社会科学与交叉科学崛起为特征的新科学革命、以受控技术崛起为特征的新技术革命、以智库与文创产业崛起为特征的新产业革命及人类社会发展模式大转型。只有这样,才能彻底摆脱修昔底德陷阱,共建人类命运共同体,引领世界的和平发展和永续繁荣。

**关键词:**致毁知识;科技危机;科技伦理失灵;智业革命;科技底线伦理;研究与编导;科技巨风险

中图分类号: N031

文献标识码: A

文章编号: 1008-7699(2018)06-0001-15

这是一个快速发展、急剧变化的时代,关心未来成为人们的迫切需要。汤因比的挑战与应战说是思考未来的有效方法。无论是个人、集体还是人类社会,洞察和预知面临的重大危机与挑战都是头等大事,因为它直接决定处境判断、及时应对和战略抉择,面对什么比拥有什么更重要,这决定了自己的安危和出路。为此本文探讨三项内容,一是人类面临的重大危机与挑战是什么?二是化解重大危机与挑战的难点和危机严峻程度,三是科技危机引发的以新科技革命、新产业革命与社会大转型为核心内容的智业革命,最后是主要结论、展望和启动方案。

## 一、以致毁知识不可逆增长为核心的科技危机是人类面临的重大危机与挑战

人类面临的重大危机与挑战是什么?这是关心和思考社会发展与人类命运的有识之士乐于回答、也必须回答的问题。可以说,以看护人类为己任的思想家、学术大师、专家学者、哲人智者们往往都会展望未来、探究人类面临的危机与挑战。看护人类、为世界体检看病,首先关注的当然是最严重的病患。谁能够最先、最深刻、最准确地发现人类面临的最大的困境、危机与挑战,谁就揭示了关乎人类命运的头等重要的奥秘。这既是最具现实意义的探索,也形成了争相探究重大危机真相的高端智力竞赛。

1945年原子弹爆炸之后,“科技发展下去将发生毁灭性灾难甚至毁灭人类”就逐渐成为流行的观点,甚至在电影银幕上大行其道,但这种看法并没有得到学术界和社会主流的认可,迄今为止悲观派与乐观派仍然各执一词。笔者认为,两派是不对称的,如果悲观派错了,至多只是放缓、延迟科技的发展而已,如

收稿日期: 2018-10-18

基金项目: 中国科学院自然科学史研究所重点培育方向“科技的社会风险”项目支持

作者简介: 刘益东(1961—),男,北京人,中国科学院自然科学史研究所研究员,博士生导师。

果乐观派错了,则将积重难返、万劫不覆,将错失人类自我拯救的最后时机。

为突破悲观派与乐观派相持不下的学术僵局与科技发展乐观论这一主流共识,更为探究“科技发展下去将毁灭人类”这一判断是真是假?笔者于 1999 年提出“致毁知识”概念并以其增长、传播、扩散、应用、禁止、治理等为研究对象开展科技巨风险研究至今 20 年,提出独特的基本假设、核心概念、核心问题、研究思路、分析框架与方法及基本结论,开辟了以科技巨风险为专门研究对象的学术领域,将科技发展下去会发生毁灭性灾难甚至毁灭人类这一老生常谈转化为严谨的学术研究,对其进行了较精确的表述并作出充分的肯定性论证,并给出条件关系。所谓致毁知识,即可用于制造毁灭性武器等导致毁灭性灾难的各种产品或方案的核心原理、核心技术等核心知识,如核裂变知识、链式反应知识、DNA 重组技术和基因编辑技术。核裂变知识出现的 1938 年是致毁知识元年。致毁知识概念的基本内涵非常鲜明,核裂变知识、DNA 重组技术等非常独特,一旦应用破坏力巨大,但是其外延则因不断有新的尖端科技出现而有所变化,存在半致毁、软致毁知识等,而且从不同学科的角度也可以有大同小异的理解。拙著《智业革命》对提出致毁知识概念的合理性、狭义与广义、硬致毁知识与软致毁知识等有详细讨论<sup>[1]158-163</sup>,在此不赘。笔者提出的核心问题包括三个前提和一个问题,三个前提是“尖端科技的正负效应不可抵消、科技知识增长不可逆、知识与应用具有连锁效应”,一个问题是“在科技知识增长的同时,能否阻止其中一类破坏力极大的科技知识——致毁知识——的增长与扩散?能否有选择地生产知识?能否有选择地进行创新?”在这三个前提下,凸显研究该问题的至关重要。尖端科技正负效应不能抵消、不能弥补,即无论尖端科技的正面效应(如核电站、核医学)多大,也无法抵消、无法弥补其负面效应(如核武器灾难、核事故),是“一坏遮百好”,是可以扬长却不能避短;科技知识增长不可逆意味着可以销毁核武器、却不能销毁制造核武器的知识;知识与应用具有连锁效应,R&D 就是连锁的制度化。越是威力巨大的知识就越要付诸应用,意味着尖端科技会被军方制度化地、在法律保护下用于武器研发(往往声称是用于防御),显然科技伦理就会失灵,因为仅对非军用领域进行伦理约束并不能禁止该致毁知识的增长,或早或晚必然会扩散。因此是否接受和发展一种科技,不是由其正面效应决定的,而是由其负面效应能否被化解所决定的。这与目前的科技发展观完全不同。显然用致毁知识概念分析科技风险要简洁有效的多,比如用双刃剑分析容易陷入各执一词,用不可抵消的致毁知识则很容易说清楚是否该禁止;又如关于限制 AI 发展问题,并不需要发展到接近或超过人脑才必须禁止,只要 AI 武器可制造毁灭性灾难,只要利用 AI 可以研究出致毁知识,在此之前的 AI 就必须禁止了,那种“AI 不会超过人脑”的说法不能成为无所顾忌发展 AI 的借口。

笔者研究表明,致毁知识增长与扩散是不可逆和不可抵消的,在目前世界主流科技与经济发展模式下,致毁知识的增长与扩散又是不可阻止的,这意味着人类面临毁灭的危险是不断累加和递增的,达到一定程度必然会发生毁灭性灾难,而且这种不可逆增长的危险累进方式,使得毁灭性灾难发生的几率越来越大,如不及时制止,必然直至爆发。这是科技的生存危机(简称科技危机),也是人类面临的重大危机和挑战。<sup>[2,3]</sup>目前恐怖及个人恐怖主义盛行,企业实验室与创客的知识生产更难以控制,互联网使致毁知识易于扩散,人类处境越发凶险。简言之,人类面临的重大危机和挑战不是好东西(自然资源等)快消耗殆尽,而是坏东西(“三废”、致毁知识等)越积越多快到地球容负容废的极限,<sup>[4]</sup>而且这种自毁进程陷入“动车困境”而难以终止。<sup>[5]</sup>

造成科技危机的原因是多方面的,主要有如下五项。①科技模式自身的缺陷。长期以来的科技发展大多是粗放式发展,科学技术与生俱来的在观念(科学无禁区)、机制(只有关于理论与事实符合与否的纠错机制,没有关于科学发展与应用正确与否的纠错机制)、体制(优先权和专利权制度,眼前利益优先)和管理(奖惩极不对称)方面存在的内在重大缺陷。②科技与社会关系方面的缺陷。现行世界主流科技与经济发展模式是 18 世纪工业革命之后形成的,其核心特征是科技与经济一体化并在市场经济的体制机制下运行发展。正如名著《后工业社会的来临》的作者丹尼尔·贝尔所指出的“通过研究的制度化,把科

学并入现行的经济结构,使之成为商业组织的一个正常部分”。与资本的结合是系统性的,从观念(科学技术的优先权和专利权竞争与资本的眼前利益优先是一拍即合)、组织(R&D、科技园区等)到制度(科技经济一体化、知识经济制度等),这种结合一方面使得科技发展迅猛,另一方面也呈现出短视行为、忽视科技风险、忽视负的外部性的特征,体现出资本主义制度的贪婪和短视。美国著名经济学家瑟罗(Lester C. Thurow)指出当代资本主义具有贪婪、利己主义、生产力和缺乏远见等四大特点,资本主义最基本的原理就是激发人的贪欲,只考虑如何在此时此地获取最大利益。<sup>[6]</sup>③技术理性和商业理性。工业革命在思想观念上的最大变化是不仅使得技术理性和商业理性得到极大发展,而且两者紧密地结合生成一种新型的强势理性,成为主导社会、全面渗透社会的思想力量,笔者将其称为“技术-商业理性”,并以此为基础形成新型的“技术-商业意识形态”,与相应的制度、组织、管理和科技等因素一同建构出“技术-商业社会”或“科技-商业社会”。<sup>[1]100-105</sup>按照李约瑟的概括,西方文化在涉及科学方面有两个特点:“第一是认为科学是认识和理解宇宙的唯一有效途径;第二是认为将科学应用于掠夺性技术从而增加个人财富是天经地义的”。<sup>[7]</sup>④人性中的好奇心、竞争性、占有欲。这些天性在科研体制机制中得到很好的保护和激励,也造成了只顾个人眼前利益。⑤利益驱动的辩护系统和话语系统,使得危机愈演愈烈。贝克(Ulrich Beck)利用“有组织地不负责任”概念进行了深刻的分析,他在《风险时代的生态政治》一书中指出:公司、政策制定者和专家结成的联盟制造了当代社会中的危险,然后又建立一套话语和理由来推卸责任。<sup>[8]</sup>利令智昏,资本主义的缺乏远见,也让一些绝顶聪明的战略思想家目光短浅。例如,布热津斯基提出的著名“奶头乐”战略,要采取麻醉、低成本、半满足的办法解除世界上80%“边缘化”人口的不满,让20%高端人口高枕无忧。他在自己的得意之作《大棋局》中直言不讳地写到:“这世界上没有什么对与错,只有国家利益最大化。”在笔者看来,布热津斯基只知其一不知其二,缺乏对科技及其巨风险的深刻理解使他(也包括基辛格等貌似深谋远虑的人物)根本驾驭不了人类发展这样的宏大主题。从总体上看,资本主义社会何尝不是沉溺于追逐金钱和消费的“奶头乐”生存方式呢,以致目光短浅,丧失了思考人类命运和未来发展能力。国家利益最大化是以人类持续生存为前提的,当代资本主义是科技资本主义,追逐科技与资本及其结合的游戏就是加强版本的“奶头乐”生存方式而已,损人害己,高科技时代的丛林法则、弱肉强食只能玉石俱焚、共同毁灭。综上所述,产生科技危机的原因是观念、文化、组织、制度、结构、系统等多方面的,因此也是制度危机、文化危机、系统性危机。科技在争议中快进,20年的研究与观察不断验证了笔者的判断和预言。

## 二、化解科技危机的十大难点与科技危机的严峻程度

科技危机是人类面临的重大危机与挑战,化解危机、应对挑战,仍然存在许多制约因素和条件,笔者认为至少有如下十大难点。

### 1. 科技发展的乐观派是国内外社会各界的主流

化解科技危机的前提是认清科技危机的存在、严峻和紧迫,但目前的情况是,在科技界几乎全部是乐观的,政府与公众的绝大多数都是乐观的,智库、人文社科领域的学者多数是乐观的或谨慎乐观的,少数是悲观的,基于认真思考和研究真正认识到科技风险严重性的更是很少。例如,无论是短期还是中长期,世界各国的科技规划对于科技风险都几乎不提或只是一笔带过;世界主流智库认为人类面临的最大挑战是气候变化和生态环境问题;即使对有争议、高风险的尖端科技的投入都在不断提高,科技发展势头强劲,这里无需列举统计数据,只要查看近年来各国名目繁多的科技创新计划、项目、经费投入、科技人才竞争、名人名家对未来的展望就能够看出对科技发展的乐观态度和预期。

### 2. 学术研究薄弱、认识混乱

关于科技风险、科技伦理、伦理法律与社会问题(ELSI)等是有相关的专业研究和理论探讨的,包括科技伦理学、科技与社会、科技哲学、社会风险理论、技术批判理论、创新风险理论、知识治理理论、技术评估、负责任创新理论等。但迄今为止缺乏对科技风险的深入系统的研究,除笔者的研究外,至今未见到国内外有关于科技巨风险的专门研究。例如欧洲学者发起的负责任的研究与创新(RRI)思潮与学派,但是他们并不研究最不负责任的创新或不负责任的创新中危害最大的创新。<sup>[9]</sup><sup>229</sup>“风险社会”理论的创始人贝克(UlrichBeck)尽管研究过科技风险,但是他认为人类社会最大的风险是气候变化。那些认为科技风险是最大风险的专家学者,由于没有揭示出导致科技灾难的机制和条件,无论如何渲染科技灾难的严重性,也缺乏说服力。西方的理论分析工具也存在严重缺陷,例如科学知识社会学(SSK)的社会建构论的根本缺陷在于无法解释迄今为止人类取得的公认的现代科学成就(包括众多诺贝尔科学奖成果)都是基于他们所批评的科学观做出来的,要么这些科学成果是错误的,要么SSK的别出心裁的基本主张是多余的。研究薄弱的结果是分析科技风险的主流思维仍然是传统直观的双刃剑模式等,关于科技发展的乐观派与悲观派长期各执一词。

### 3. 认识误区根深蒂固

科技发展的乐观派是主流,对科技巨风险缺乏深入的认识,人们在认识、策略、思维方式、价值观和科学观等方面存在十大误区。<sup>[1]</sup><sup>149-158</sup>这里只提及六个误区。一是在科技风险感知方面,不仅公众普遍迟钝,就是专家学者也远没有给予应有的重视。例如创新理论是一个大的族群,但几乎全是关于科技正面效应的论述,创新风险也指的是财务风险、投资风险而已。二是红水母效应与虚幻的利弊权衡,对科技正负效应的判断十分复杂,在不同条件下会发生转变。笔者提出用“红水母效应”(也可称之为“科技变色龙效应”,以区别已有的“变色龙效应”)表示科技风险的复杂性,它有本身发生变化和对其认识因人因时因地而异这两层含义,它比“双刃剑”更能够反映科技正负效应的复杂性。“黑天鹅”比喻小概率而影响大的危机,“灰犀牛”比喻大概率且影响大的危机,“红水母”比喻概率可大可小、忽大忽小(“小”掩盖“大”)而影响大的危机。水母可通过自产一氧化碳气体的充气 and 放气使自己可大可小,其外表漂亮具有迷惑性,而实际上生性凶猛,适合比喻科技风险。例如克隆猴的出现就是“红水母事件”,它自身风险可大可小,对其风险认识也因人而异,风险小到疾病动物模型使用不当,风险大到实现克隆人并用其研发更有效的基因武器等。增加“红水母”比喻,就用海陆空动物比喻完整覆盖了概率大小、可大可小而影响大的危机。科技正负效应本身的复杂性和广泛性、对其认识的复杂性和差异性,使得用权衡利弊来决定能否发展某些科技是难上加难。三是“双刃剑”思维至今仍然是人们考虑科技风险的主流思维方式,笔者认为它有三个特征:①强调要扬长避短,但是并没有注意到扬长容易、避短难,很难两全其美;②把知识的创造者(铸剑者)与应用者(用剑者)分开,把负面效应的产生归咎于应用者,效果好坏取决于使用知识的人,把人当作解决问题的关键,强调科技伦理重要性就是如此。实际上许多情况下是知识诱使或迫使人来应用,面对科技身不由己,知识与应用具有连锁效应;③不区分普通负面效应和极端负面效应。许多普通负面效应是可以正负抵消的,而尖端科技的正负效应却无法抵消(如核电站与核武器),不能抵消负面效应的就不是双刃剑,而是单刃斧。目前科技专家、决策者和公众往往都是用双刃剑思维来思考科技风险的,如果不突破它并用更有效的新思维取而代之,用新的思维武装头脑,则再精妙深奥的理论也无济于事,仅仅是社科学家之间彼此唱和欣赏的理论是不能解决科技风险问题的。<sup>[10]</sup><sup>149-150</sup>四是笼统的社会控制乐观论与政策语言的一厢情愿。尖端科技的快速发展引起公众担忧,结果在科技政策的表述上出现了“负责任、扬长避短、协调健康发展”等等,但是如果缺乏有效的具体措施和真抓实干的执行力,这种政策表述就容易掩盖实际上的科技粗放式发展,许多高危的尖端科技在扬长避短的主张的掩护下大行其道,而根本不探究能否做到扬长避短,笔者研究的结论是“可以扬长,但不能避短,正负效应不能抵消”。五是辩护性的风险与伦理

研究。随着尖端科技发展,社会各界对其风险的担忧日益加重,关于科技风险与伦理的研究也随之增加,例如伦理、法律与社会问题(ELSI)研究,但是许多研究的前提是先定的,就是确保高科技继续发展,解除公众的担忧和干扰。这种为尖端科技发展保驾护航的辩护式研究已经受到一些学者的批评。例如麦克纳顿(Macnaghten)等人认为人类基因组的 ELSI 研究被锁定在了既定的框架中,使得人们认为技术是既定的、难以改变的,只有接受技术然后再去设想技术可能产生的伦理与社会反应,而缺少真正触动创新过程的机制。<sup>[11]</sup>批评者认为 ELSI 的目标就是为了排除公众干扰,使科学家坦然地认为——我们已经对这些问题思考过了,可以无忧无虑地做研究了。ELSI 是科学家为了“便利”的产物。如里普(Rip)指出的这是一种同心轴结构,ELSI 研究开始于科学计划,然后在其外圈添加一层一层的解释与考量。<sup>[12]</sup>20 年过去了,现在一些尖端科技伦理的研究何尝不是如此呢?六是社会发展必然论与科技发展必然论。目前人们普遍认为科技发展是必然的,是不以人的意志为转移的,根本不能想象可以叫停某些科技发展。实际上能不能叫停与该不该叫停是两回事。如果某些高危的尖端科技确实威胁到人类生存,就必须无条件叫停。世界上只有一个不证自明的前提,就是人类要持续生存,凡是与此抵触的就必须纠正。对待 AI 的态度就是典型的科技发展必然论。AI 是非常另类的科技,其他科技都是“人主科技”,AI 是“自主科技”,通常说技术是人体器官的延伸,AI 不是延伸,是另搞一套,是喧宾夺主,是外来入侵的物种,但是科技与社会的主流却主张大力发展 AI。甚至对 AI 保持警惕并正确指出 AI 具有没完没了的“暗无限”特点的学者,都还要迁就 AI 的发展,仍然主张“让机器人更像人,如此人才能安全存在”<sup>[13]</sup>,可见科技发展必然论多么深入人心。实际上只要发展通用 AI,机器人接近人就会很快超过人,因为机器人有不断升级换代的计算能力、存储能力和网络能力。禁止发展通用 AI,发展非自主的、经过严格筛选的专用 AI,才是对待 AI 的明智之举。此外,预警疲劳与从众心理也使得人们面对越来越多的关于科技风险、科技灾难的言论和影视作品处于麻木状态,不少科幻影片起到了预警疲劳的作用。对科技的盲从导致许多匪夷所思的现象,例如,人们对外来生物物种入侵非常警惕,但是对凶悍之极的机器人和人工智能的到来却欢呼雀跃;区块链技术瓦解政府功能的技术,却得到不少政府政策的支持。

#### 4. 科技发展模式的内外缺陷

在上文分析造成科技危机的原因时已经提到科技与生俱来的内在缺陷与外部缺陷,显然这也是阻碍化解科技危机的难点。特别是科学无禁区、缺乏关于科学发展与应用正确与否的纠错机制、优先权与眼前利益优先结合与管理上的奖惩极不对称。科技与资本结合并在市场经济制度下运行,必然导致急功近利。目前社会危机重重,生态危机、气候变化、经济危机、道德下滑、信任危机、贫富悬殊、急功近利、恐怖主义等等,这些危机放大了科技发展模式的内外缺陷,危机相互加强,使得科技危机更加严峻和凶险。

#### 5. 科技伦理失灵

科技出现风险和负面效应,人们容易想到的就是加强科技伦理研究和实施。比如近几年人工智能发展迅速引起人们担忧和恐慌,人工智能伦理逐成热点,2018 年新学期开始,哈佛、MIT、斯坦福等美国高校开设人工智能伦理、数据科学伦理、技术伦理、机器人伦理等新课。这是对公众因人工智能等新技术感到担忧的回应。长期以来,硅谷的信条是“先做出来,再请求谅解”,伦理不是硅谷顶级科技专家考虑的事,他们一贯将伦理视为技术创新和进步的拦路石。<sup>[14]</sup>科学界同样如此,科学探索无禁区至今仍是科学家所信奉和遵从的原则。国内关于科技伦理、人工智能伦理的研究和课程也不断升温。遗憾的是,这种试图利用科技伦理来治理科技风险、科技负面作用的做法暴露了美国学术界研究的薄弱,根本没有意识到现行的科技伦理并不能解决科技巨风险问题。伦理失灵这个术语早已有之,科学家伦理失灵也有提及,但是限于讨论科学家行为不端(剽窃、造假、署名不合理)问题,笔者从治理科技风险的角度提出“科技伦理失灵”概念,即:科技伦理因不能约束所有科技专家的行为,而使本应禁止的某些科研活动得不到有效禁

止而使得科技伦理失灵,在知识极易扩散的互联网时代和尖端科技领域,科技伦理失灵尤为严重。科技的特点决定了只要不能约束所有人的研发行为(如不能约束军方和黑客),少数人研发的成果同样可以扩散而导致科技伦理失灵。科技伦理失灵有四点涵义:①挂万漏一即失灵。科研的特点是,一项科学成果被发现 1 次和被发现 100 次是一样的,技术发明也如此。因此科技伦理如果有效就必须禁止所有科技专家和所有实验室都不能从事被禁的研究,否则就无效。显然科技伦理不能约束所有科技专家和实验室。这与社会生活不同,社会上大多数人遵循伦理就有好效果,少数人不恪守伦理的危害有限;②连锁效应与易于转化即失灵。即使恪守科技伦理,避免直接做禁止的研究,但是只要做相关的研究,该研究易于发生连锁效应或转化为被禁的研究成果,同样也是科技伦理失灵。例如通常认为基础研究是可以得到伦理学辩护的,但这是错误的,因为基础研究与应用研究具有连锁效应,基础研究一旦突破,人们就会去尝试应用,而且越是破坏力巨大的知识(致毁知识)越是如此;③阳奉阴违即失灵。法律硬约束都难以禁止违法之事,伦理软约束更是难以限制,学术研究很复杂,为阳奉阴违提供了借口和便利。可见科技伦理有效的条件十分苛刻,在目前条件下,尖端科技在法律保护下用于军方武器研发是不可避免的,黑客、创客不受伦理约束,企业研发机构也不会因易于有危害的转化而束手束脚,科技伦理必然失灵;④相互制约式科技伦理失灵。AI 武器、基因武器等尖端科技武器的使用门槛较低,技术容易扩散,又不依赖稀缺原材料,无法维持“AI 武器平衡”“基因武器平衡”,因此希望用类似“核平衡”的相互制约的方式解决科技伦理问题也不可能。“核平衡思维”误导了人们,仿佛相互制约可以避免同归于尽,实际上没有爆发核战争或核误动灾难是有运气和侥幸因素的。即使国家之间有“平衡”的考量,恐怖分子却全然不受此约束。霍金、伊隆·马斯克等上千名知名科学家和人工智能专家于 2015 年签署公开信,指出 AI 武器将在几年内就会出现,这将带来极大风险,它将是继火药和核武器之后的第三次武器革命。公开信警告不要进行 AI 军备竞赛,敦促联合国发布针对“攻击性自动化武器”的禁令。<sup>[15]</sup>显然,这项努力不会奏效,因为无法禁止军方进行防御性 AI 武器的研发(尖端科技武器往往都是以防御为由研发的,实际上它与攻击性武器并没有本质区别)更不能禁止黑客和恐怖分子进行武器研发。在这个碎片化、区块化的时代不可能用统一的伦理加以有效约束,科技伦理必然失灵,试图用科技伦理来解决科技风险问题,是避重就轻,犯了方向性错误(至少是杯水车薪)。早在 2009 年江晓原、刘兵、田松就主编了题为《伦理能不能管科学》的文集<sup>[16]</sup>,众多学者对此问题进行了深入讨论,在此笔者从科技伦理失灵的角度也给出一个回答,就是伦理不能管科学。

#### 6. “囚车剑魔”(PTSR)四大困境

这里把著名的囚徒困境与笔者提出的动车困境、双刃剑困境、魔戒困境合在一起,从人类是否犯重大错误和能否认清、纠正、抵消易犯重大错误的角度来揭示人类面临的根本性困境,简称“囚车剑魔”四大困境。囚徒困境表明吃一堑未必长一智;动车困境表明纠正大错是边运行、边分析、边争论、边纠错,纠错需要迈过“达成正确共识、预期共赢、共同有效行动、具备有关条件”四道坎。纠错条件十分苛刻,在巨大惯性之下纠错更是难上加难,认识到错误、达成共识都未必能够纠正错误,可谓“刹车失灵”。环境问题基本“达成共识”,科技风险问题则远未“达成共识”;双刃剑困境表明正负效应无法抵消、无法弥补,是一“坏”遮百“好”;魔戒困境表明犯重大错误的门槛越来越低,人有理性所以经不住诱惑,尖端科技使小人物和机器人可以犯大错误。四大困境表明人类不断犯大错,大错误又难以纠正、无法抵消、无法弥补,犯大错误的门槛越来越低,小人物和机器人也能犯大错误,这是人类面临的最严峻的四大困境,人类在科技发展与应用问题上所犯的 error 最为严重。<sup>[9]221-230</sup>

#### 7. 裹挟困境与反对无效:不问善恶对错,只问强弱快慢

在竞争、资本、市场的催动下,科技在争议中快速发展,担忧和反对无效、反对者出局。例如奥本海默反对研制氢弹,结果靠边站了,特勒成为氢弹之父。为发展生物技术,美国多次放宽对重组 DNA 的管理,

DNA 双螺旋结构发现者之一沃森将这种专利申请斥为“十分愚蠢的行为”,并因此辞去美国人类基因图谱工程主任之职,但这丝毫没有影响基因图谱工作的迅速进展。对生物技术管理十分严格的德国,也在人才流失、资金流失和经济竞争压力下放宽了限制。谷歌云首席科学家李飞飞带头反对谷歌与军方合作,结果被迫离职。不是没有自觉担负社会责任的科学家,而是他们不起作用。科技与资本结合裹挟众人前行,上行下效,前行后效,无暇顾及善恶对错,必须加入主流,必须趋同,反对者、迟疑者出局。

#### 8. 区块链去中心化、摆脱第三方监管

化解科技危机依靠各方齐心协力,而以加密的分布式记账技术为核心的区块链是去中心化,摆脱第三方的帮助和约束,供求双方自行其是即可,而在尖端科技中设置科研禁区恰恰需要第三方的制约和监督。在区块链中用加密数字货币来结算,例如比特币(BTC)、以太币(ETH)、零币(ZEC)、门罗币(XMR)等。目前常用的比特币广泛用于匿名买卖双方,在“深网”(Deep Web)中进行大量非法交易,只有想不到没有做不到的非法交易五花八门、触目惊心,其中一家名为“丝绸之路”的网站,全球有 100 万顾客,全年营业额在 12 亿美元以上。<sup>[17]</sup>显然,这些非常便于致毁知识的流动和扩散。

#### 9. 科研平台化、全球化、网络化、企业化、创客化、分散化、人工智能化

进入 21 世纪,科技发展提速,科研与创新平台升级换代是重要原因,互联网、物联网、云计算、大数据、人工智能、虚拟技术、数字技术、纳米技术、基因编辑技术等构成的科研与创新平台大幅提升了科研的工作效率与合作效率,实现在全球范围内的高效合作。由于拥有数据、薪酬和拔尖人才优势,企业的科研实力快速攀升,在科研布局中的地位日益突出。借助互联网科研与创新平台,创客的科研能力也显著提升。与以往相比,科研呈现出分散化趋势。以往对于大学和国立机构实验室进行控制管理都很困难,在这种情况和趋势下,要管理企业的研发和创客的研发更是难上加难。人工智能可以从事科研工作,甚至可以自主进行。据《Nature》网站报道,最近,研究者借助 AI 技术发现了近 6000 种新病毒,这一工作已在 2018 年 3 月 15 日由美国能源部(DOE)组织的一场会议中展示,AI 成为了一种探索发现数量巨大、种类繁多的病毒的新工具。<sup>[18]</sup>近年来,对强 AI 争议不断,但并没有减少对其研究<sup>[19]</sup>,纳米、生物、信息和认知科学的会聚技术(NBIC)同样可以因为计算机和人工智能而形成自主发展。最近量子计算机的突破,不仅可以大幅提高人工智能技术,而且与合成生物学、基因编辑技术的结合将产生爆发式发展,致毁知识将随之快速增长。

#### 10. 冷战、科技冷战与军备竞赛

众所周知,近年来国际形势严峻,如 2018 年 5—6 月号的美国《国家利益》杂志发表的封面文章“美国对阵中俄:第二次冷战到来”,指出从政治、外交、军事、经济等角度分析论证了目前世界面临第二次冷战到来的现实。<sup>[20]</sup>类似的观点很多,显然这个判断是符合事实的。进入 21 世纪,人类社会发生了巨大变化,尤其是科技在争议中迅速发展,科技力量巨大,越来越难以驾驭。但是面对新形势,传统思维、传统方法却占据许多人的头脑,加剧了大国之间的对抗乃至陷入修昔底德陷阱和囚徒困境,冷战、最危险的是其中的科技冷战与军备竞赛,尖端科技与军工科技的恶性竞争将产生越来越多的致毁知识和破坏力越来越大、越来越便捷的致毁武器。只要对比 20 世纪 30—40 年代的研发平台与当今的研发平台,就知道尖端科技与军工科技的竞争不仅是极度危险的,也是人类根本承受不起的。在过去那样简陋的研发平台(只有机电计算机)上都能够研制出原子弹,当今研发平台则包括互联网、物联网、云计算、大数据、人工智能、基因编辑技术、合成生物学、纳米技术等高科技,在其上的研发竞争必然导致军备竞赛的升级换代,致毁知识不断涌现,不仅国家之间的战争还有极易发生的恐怖袭击和个人恐怖袭击,把人类推向自我毁灭的深渊。<sup>[21]</sup>

上述分析表明科技危机异常严峻且化解困难重重。笔者把危机的严峻程度分为六个要素:①危机本

身的危害与凶险程度;②危机相关因素范围与复杂程度;③对其认知和共识的难度;④化解危机的难度;⑤紧迫程度;⑥发生的概率,且用 1 星到 5 星由低到高表示级别,那么科技危机在这六个方面都是最高的 5 星(见表 1)。目前科技已经渗透到社会的方方面面,因此“危机相关因素范围与复杂程度”是最高的。如果是小行星撞地球,尽管危害程度是最高,但是相关因素却不多。以致毁知识不可逆增长为核心的科技危机是人类有史以来最大的危机和挑战。

表 1 科技危机的严峻程度

科技危机	危机本身的危害 与凶险程度	危机相关因素范围 与复杂程度	对危机认知和 共识的难度	化解危机的难度	紧迫程度	发生概率
危机严峻程度	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★	★★★★★

### 三、科技危机引发新科技革命、新产业革命与社会大转型

危机引发革命,科技危机引发科学革命、技术革命,进而引发产业革命。这里的科技危机引发科技革命,显然与以往的库恩所说的科学危机引发科学革命并不相同,但是从根本上看又有相通之处:都是不同层面的生存危机引发的变革图存。新科技革命、新产业革命与社会大转型构成智业革命,人类继农业社会、工业社会之后进入智业社会,<sup>[1]201-245</sup> 农业社会缺乏知识,工业社会利用知识,智业社会控制知识。

#### 1. 以社会科学与交叉科学崛起为特征的新科学革命

哥白尼、伽利略、牛顿、爱因斯坦等以往的科学革命都是在“现象—分析”二维平面内完成的,而这次科学革命则是引入价值维和一套新的纠错机制,以纠正科学不合理发展与应用的错,实现从“平面”科学到“立体”科学的转型,这是科学史上第一次把知识体系的变革与人类的生死存亡关联在一起。科学包括自然科学与社会科学,这已经成为学术界的共识,“立体”科学的确立包括文理交融和科学重心转移、社会科学交叉科学崛起。或者说,社会科学交叉科学崛起才能够把价值维确立,使科学从平面到立体,建立以人为本的新科学。<sup>[22]</sup> 揭示人类面临的最大危机和挑战的严峻性与紧迫性,必然产生培养人们反思意识、风险意识与未来意识的强烈需求。刘孝廷教授强调指出:“如果不反思发展的前提、发展的动机和目的、发展的手段等,盲目发展的后果是不堪设想的。”<sup>[23]</sup> 对科技发展的反思正是如此,这些都成为社会科学崛起的契机。目前社会科学崛起存在内外两方面的有利因素,内在原因是一方面社会科学本身发展迟缓,例如,社会科学中大部分方法、理论都是在个人视角、集体(包括国家)视角下做出的,而较少人类视角的方法和理论;又如,经济学被认为是社会科学中科学程度最高的学科,但是经济学中关于技术的研究非常薄弱,要么把技术视为一个函数、黑箱,要么不考虑技术的负面作用。创新理论是很大的理论族群,各种创新理论都把科技视为推动经济发展的因素,而不考虑科技的风险和负面作用(或者只考虑投资风险、财务风险),知识经济学则不考虑知识的负面作用、特别是没有考虑知识中存在一类破坏力极大的知识(致毁知识)的作用,内生经济增长理论强调知识这类带有外部性的生产要素可以导致收益递增,使得经济增长可以长期持续下去甚至增长可以越来越快,这显然是错误的。因此社会科学有巨大的发展空间。另一方面是新的研究工具和研究方法纷纷出现,如互联网、物联网、大数据、云计算,虚拟科技、仿真平台、实验经济学、复杂性科学研究方法、计算社会科学(社会计算)、平行管理、政策仿真等,在研究范式上也出现了可研究未来的“虚拟科学”这一新的研究范式。<sup>[24]7-13</sup> 外在原因是社会对社会科学研究的需求越来越大,从 20 世纪 90 年代开始的人类基因组计划,就要求用该计划经费的 3—5% 做 ELSI 研究,此后这已经



在国际上成为惯例,虽然 ELSI 研究还存在不足,但是对社科研究的需求非常明显。从总体看,不管是社会危机还是科技危机,这些都急需更丰富更有效的社会科学来解决。科技危机的严峻性与紧迫性也表明,化解它需要社会科学与交叉科学,需要未来学、战略学、科技与社会、社会学、经济学、管理学、历史学等和理工科各学科通力合作,进行跨学科、交叉学科的综合研究与政策推进。STS 参与进路(STS 研究者的角色从外在于“科学技术”的观察者到内在于‘科学知识生产’的参与者的转变)<sup>[25]</sup>等已经做出宝贵的探索。基于长期的思考和研究,笔者拟从新研究范式、评价方法、方向管理与未来意识、社科研发与智库及文创产业结合、科技底线伦理、可持续创新与发展理论等几个方面作简要论述,它们是社科与交叉科学崛起的部分重要内容。

(1)研究未来的虚拟科学:科学研究的第五范式。把世界划分为现实世界和虚拟现实世界,借鉴吉姆·格雷的四种研究范式说,笔者提出“虚拟科学”作为科学研究的第五范式,并对这一新的研究范式的定义、要素、功能及其与其他范式的关系进行了探讨。“虚拟科学”比“数据科学”更能反映出当代科学研究的最新方式与最新进展,使得科学研究终于完成了对已经发生和尚未发生的事物与现象研究的全覆盖。虚拟科学使想象力成为受控的想象力、精确的想象力,实现对想象与情景的有效描述、预测、控制和评价,可以预演、预测科学研究与传播、应用的每个环节及其带来的风险,从而在立项之初或研究伊始就提前终止有潜在风险的研究项目。创新理论直到虚拟科学的出现与应用才真正实现全方位、全工况、全覆盖、全面考虑科技正负效应的创新研究。<sup>[24]7-13</sup>

(2)开放式评价:社科与交叉科学的有效评价方法。社会科学发展迟缓的原因是多方面的,社会科学要解决的问题更复杂、有自反性、难以量化,评价的主观性更强,同行相轻更严重,原创成果更难及时胜出。笔者提出开放式评价法使社科评价更高效合理,其策略是公开展示、规范确认、抓大放小、集中评价代表作。学术评价不仅要分类也要分级,“较好”标准是“数数”(论文、基金、获奖、头衔等数量等级)、是综合考虑,但这不是“拔尖”的标准,如同顶尖大学的博士可保证其“较好”,是否“拔尖”则需看其代表作。对于拔尖人才只有通过其代表作水平来衡量,拔尖人才不需要加权综合考虑,用代表作的水平高度足以衡量。抓大放小就是甄别拔尖人才,令其成为学术带头人,并由其自己再选择团队成员。拔尖人才的学术水平不是由其成果的数量和质量决定的(中等成果的质量再高也是中等水平成果),而是由其代表作的水平高度决定的。凡在细分领域做出突破者,均可具备标志点,并作为其学术招牌。学术招牌可以表示其核心贡献、学术前沿地位和其他学者研究的参照点,其信息量和含金量都比笼统的人才帽子要高得多。对突破点四要素的提炼与概括,有利于今后用人工智能进行评价,如“AI+代表作”或“互联网+AI+代表作”(IAI+代表作),也有“互联网+人机智能+代表作”,基于同行挑颠覆性错误的 AI 评价与推荐法(简称“基于同行挑错的 AI 评价法”)将成为重要评价方法,甚至可能取代传统的同行评议。开放式评价把以往的同行人承认作为唯一标准,改变成为同行承认、用户选择、检索推荐、AI 推荐、社会评价等多元标准,把创新力度、启发性和有用性也作为重要标准,避免同行相轻,让创新成果和创新人才及时胜出。<sup>[26]2-10, [10]145-150</sup> 细分领域很多,突破性成果有大有小。原创性地提出好问题属于方向型突破性成果,原因有三:①原创性好问题可能意味着开辟一个研究的新方向、新领域;②好问题三要素是新颖、重要和可研究性,好问题往往是解释性问题,其中蕴含了新观察、新理解和条件关系,重要性包括挑战性、紧迫性和扼喉性(卡脖子);③在教育 and 学术领域,非常强调“问题意识”“问题导向”,所有学者都知道提出好问题至关重要甚至头等重要,而好问题是稀缺的,提出好问题竞争激烈,是竞争点和比较点。提出问题有时比解决问题更重要(爱因斯坦语),提出正确的问题往往等于解决了问题的大半(海森堡语)。因此,提出好问题也是做出了突破,是学者拥有创造力的标志,是拔尖人才的试金石。用能否提出好问题来甄别拔尖人才,简洁有效,一目了然,难以假冒。好问题具有难得易懂的特点,提出很难,提出之后说出来往往容易理

解,那种把已有的问题进行变形或简单组合,以旧充新,是无法冒充原创问题的。原创具有新颖与增长两个特点,缺一不可。原创不仅是从 0 到 1,而是从 0 到 1 再到 N,如果仅仅是从 0 到 1,那么许多新观点、新东西都是原创了。标志性成果往往就是突破性成果,开放式评价与“互联网+代表作”“互联网+AI+代表作”的实施,可实现客观公正及时地评价学者水平,破除唯学历、唯奖项、唯职称头衔、唯论文数量及刊物级别,将真正做到以成果论英雄,这不仅是科研管理的改革,更是一场科学革命和教育革命,因为它让真正有创造力的人才摆脱条条框框的束缚,充分释放创新活力,成就自己、服务社会。

(3)方向管理与未来意识培养。克服急功近利,培养高瞻远瞩和未来意识,要从对工作学习有切实帮助入手,战略方向管理就是这样一新兴的实用学科。探讨未来发展首先应该探讨未来发展的方向,目前的研究还比较薄弱。例如作为战略规划中最常用的方法,SWOT 分析法就缺少对发展方向的考虑,为此笔者提出 Challenge(挑战)、Opportunity(机会)、Adaptation(适配)和 Incompatible(不兼容)方向分析法(简称“COAI 方向分析法”)。战略方向主要有四类:①趋势外推;②改弦易辙;③突发突变;④前面两种或三种的组合,突发突变是短时间的变化,关键是研判总的未来发展究竟是趋势外推还是改弦易辙。<sup>[27]7-9</sup> COAI 方向分析法是通过洞悉面临的挑战来决定方向的,如果挑战力度不大,现行发展模式可以化解挑战,则是趋势外推,否则因挑战力度大以致现行发展模式无法应对而不得不改弦易辙。以发展模式对机会和挑战是否适配有四种情形,如表 2 所示。COAI 方向分析法可以解决战略方向问题,可以与 SOWT 态势分析法联合使用,更好地解决决策与战略规划问题。

(4)社科研发的崛起:研究与编导(R&WD)及其与智库及文创产业的结合。目前所说的研发(R&D)主要指自然科学技术,有时在社会科学领域也如此表达,例如在统计社会科学的研究与发展经费方面。自然科学与社会科学多有不同,研究与发展也应该遵循各自的规律和特点而有所区别。按照百科辞典的解释,研究与发展(R&D)

表 2 COAI 方向分析法

适配	主动趋势外推	增强趋势外推
不适配	诱致性改弦易辙	强制性改弦易辙
	机会	挑战

是指为了增进知识,以及利用这些知识去开创新的用途而进行的系统的创造性工作,它包括基础研究、应用研究和实验发展。笔者提出“研究与编导”(R&WD)作为社科研发和部分的交叉科学研究,它是指为了研究问题、增进知识及其应用,而进行的包括从学术研究到剧本创编再到导演实现的系统的创造性工作,包括学术研究、传播或应用的剧本设计和导演实现,三者之间有连锁关系。<sup>[28]</sup>“研究与编导”有两层含义。一是狭义的。编剧和导演就是影视和舞台剧等通常意义,这时强调编剧和导演要基于学术研究,比如电影《指环王》的知识含量高,其编剧是牛津大学中世纪英国文学教授,对于知识密集型影视作品,如科学预演片、政论片、研究型纪录片和研究型故事片,其编导应该升级为“研究与编导”。二是广义的。研究是发现问题、解决问题,剧本是解决方案,导演是对其实施、反馈和评价。广义的“研究与编导”就是社科领域的研究与发展。把“发展”替换成“编导”的理由是社会科学成果的传播与应用往往就是观念、创意、设计、规划、策划等知识产品(而不是自然科学成果应用的实体产品),最适合编剧和导演来承担,而且还可以用新媒体方式进行传播,传播者和接受者在编码、解释、接受、反馈等信息传播与交流过程中,互动、循环、体验、相互影响。类似体验经济,笔者提出“体验学术”(或“体验科学”“体验科技”“体验知识”)概念,它是通过反映和体验研究与传播或应用的过程来增加学术附加值的活动,拉图尔(Latour)的名言“用眼睛和手进行思考”可以作为体验学术的一个注脚,“研究与编导”有助于实现体验学术。学术演讲、朗读等可视为“体验学术”或包含“体验学术”成分,在知识社会、智识社会中,“体验学术”应该与“体验经济”并驾齐驱。“研究与编导”可以有效解决智库与传媒(TT&M)结合问题,使之发挥更大作用,若创建“智库与传媒学院”,“研究与编导”将更能彰显重要作用,两者相辅相成,都会得到快速发展。社会科学崛起不仅需要其

本身的方法、理论、范式的升级换代,也需要与产业结合,以解决经费和经济回报问题。一百多年前,德国的工业实验室和美国的 R&D 机构的出现,较好解决了自然科学与企业结合的问题,实现了自然科学崛起。同理,“研究与编导”要解决社会科学与智库和文化创意企业的结合,通过研究产生创意,通过研究实现编剧和导演,而不仅仅依靠个人灵感和经验。在计算机辅助、AI 辅助编剧、导演的时代,学者成为编导比编导成为学者容易,只要善用计算机和 AI 等技术辅助系统和摄制组专家辅助系统,可在较短时间内成为“研究与编导”,研究与编导一体化让更多的专家学者掌握编导技能,传统意义的编导则应注重学术、注重提升研究能力,都成为研究与编导的复合型人才,成为“兼做编导的学者”“编导型学者”和“学者型编导”,把自己和他人的学术思想进行有效传播、更新观念,造福社会。社会科学崛起还在于其重要的更新观念的作用。观念对人的行为影响巨大,固守旧观念、旧思维严重阻碍进步、错失转型的良机。但是,更新的重要性远远没有得到应有的重视,更新应该和创新同等重要,创新理论众多,而更新理论却很少。笔者认为更新理论至少有三个原则,应该替代、能够替代、限时替代,在规定时限内完成更新至关重要,迄今为止,限时原则被忽略了。<sup>[26]2-10</sup>

(5)科技底线伦理:基于科技总体安全观的新科技伦理。为应对科技伦理失灵,笔者提出科技底线伦理作为新的科技伦理,它基于科技总体安全观,科技的总体安全包括底线安全、人类安全、可持续安全。底线安全就是即使军事应用、恶意应用和滥用也能避免或也能承受或也能化解的科技才是可以接受的科技,一项科技是否应该发展首先不取决于其正面效应,而是首先取决于其负面效应的严重性,取决于社会能否承受、化解、抵消该科技的负面效应。<sup>[9]228</sup>这就关乎科技底线伦理,科技底线伦理是为了实现科技底线安全与可持续创新所规定的科技工作者及其共同体应恪守的价值观念、社会责任和行为规范。科技底线伦理有四项基本原则:①底线安全的价值观,与科学无禁区相反,要确立科研禁忌的价值观,凡是预知和应该预知科研工作有可能产生危害人类安全的结果,就坚决停止,哪怕因此丧失优先权。如《中国科学院关于科学理念的宣言》所述“要求科学工作者更加自觉地规避科学技术的负面影响,承担起对科学技术后果评估的责任,包括:对自己工作的一切可能后果进行检验和评估;一旦发现弊端或危险,应改变甚至中断自己的工作;如果不能独自做出抉择,应暂缓或中止相关研究,及时向社会报警。”<sup>[29]</sup>价值观的意义在于遏制自发的逐利行为,底线安全的价值观就是遏制无视后果的追求优先权和利益的行为;②科学研究底线安全原则,基础研究等科研要确保致毁知识零增长,例如,即使人工智能不超过人类,只要人工智能发展下去要产生致毁知识,就必须叫停;③技术开发底线安全原则,以知识扩散不可避免为前提制定措施,例如,只要军方和黑客开发人工智能武器(即使以防御为目的),制造人工智能武器的知识就迟早会扩散到社会,为恐怖分子等所利用,因此禁止军方和黑客研发人工智能武器,如果无法禁止,则科技界不能继续相关研究以避免被军方和黑客及恐怖分子所利用;④全体持续底线安全原则,尖端科技是扬长不能避短、正负效应不能抵消,需要全体研究者长期恪守底线,独善其身是不能解决问题的。例如人工智能顶尖学者李飞飞在斯坦福大学创建了以人为本的人工智能研究院(Stanford Human-Centered AI Institute, HAI),非常令人敬佩,但是一万个天使抵消不了一个魔鬼,所以治理 AI 巨风险既要尽量让更多的人工智能专家做 HAI,更要禁止开展强 AI、通用 AI 的研究。在设置高风险尖端科技禁区时,本着宁可错禁、不可错放的原则,本着不安全推定法则保证底线安全;在无法证明一项科技是底线安全的,那它就是底线不安全的,为谨慎起见,需要叫停。科技底线伦理也是一种评价标准,例如对 1975 年在美国加利福尼亚州举办的阿西洛马会议(Asilomar Conference)至今评价不一,从科技底线伦理的角度就很容易做出评判:该会议是失败的。因为它实际上在科学共同体内部就没有阻止危险的生物技术出现,比如出现了被称之为潜在的大规模杀伤性武器的基因编辑技术,它更无法阻止以防御为名的基因武器的研制。

(6)可持续创新与发展(SID)理论。通过科技巨风险研究,首先揭示出何为最不可持续的创新、何为

最危险的粗放式创新,以此为前提和治理对象,笔者提出可持续创新与发展(SID)理论加以应对,它是创新理论的反向的全新开拓,是可持续发展理论的升级换代。<sup>[27]</sup><sup>[13-14]</sup>通过与负责任研究与创新(RRI)理论的对比(见表 3),可以突出 SID 的特点和优势。负责任创新虽然是学术热门,但是也受到不少批评,如英国诺丁汉大学的布丽奇特·奈丽诗批评负责任创新炒作大于实质、内容空洞、缺乏实质性内容。<sup>[30]</sup>以强调各方负责任可能是 RRI 的一个方向性的错误,因为科技危机如此严峻不是依靠科研主体和创新主体负责任所能够解决的,因为不可能所有相关主体都能负责任<sup>[9]</sup><sup>[229]</sup>。以往对待科学有辩护和批判两种取向,刘大椿教授提出了科学审度观,把审度作为对待科学的第三种取向,它用多元、理性、宽容的观点看待科学。<sup>[31]</sup>科学审度观非常重要,它扩大了思考科技与创新问题的空间,受此启发,可持续创新与发展理论将继续对以往的科技创新理论和负责任创新理论进行继承和扬弃,使创新理论升级换代。

表 3 可持续创新与发展理论与负责任研究与创新理论比较

可持续创新与发展(SID)理论	负责任研究与创新(RRI)理论
以可持续创新的反面——粗放式创新——为研究重点,并以治理最危险的粗放式创新为基础和前提	没有以负责任创新的反面——不负责任的创新——为研究重点(尽管也有所提及),更没有研究最不负责任的创新,没有研究不负责任的创新中危害最大的创新
首倡科技底线伦理,明确“底线性”:遵守安全规范是不够的,需要确保致毁知识零增长。发展底线安全的可持续创新	默认现行的科技伦理,只要遵守安全规范即可,没有注意到科技伦理失灵问题,没有明确“底线性”
明确“时限性”:在规定时间内解决问题	没有明确“时限性”,忽视了限时原则
明确增加新的创新主体:人文社科、智库及其与文创产业结合的专家学者,形成新主导产业	没有明确增加新的创新主体
以发现科技危机为人类最大危机与挑战入手,揭示科技危机的严峻性、复杂性和紧迫性,指出通过科学革命、技术革命、产业革命,实现科学技术重心转移、社会科学崛起,创建可持续创新的新科学、新技术、新产业,以知识分子为核心的公共利益集团崛起来化解危机!	强调解决问题的关键是让创新主体(科学家、科技专家、企业家、企业等机构)负起责任来,强调“全责任”,提升道德水平,强调人权、强调对弱势群体的保护。强调多学科交叉、跨领域合作是负责任创新的本质特征
特点是揭示最大危机,确定挑战的严峻性、紧迫性和艰巨性,在此前提之下研究政策的制定和实施	政策实践经验丰富,但是由于没有认识到挑战的严峻性、紧迫性和艰巨性,而使得政策不能面对真正的问题

## 2. 以受控技术崛起为特征的新技术革命

新技术革命是从粗放式创新转变为可持续创新,以往的技术革命是人类对物质、能量、信息的利用与控制能力的跃升,技术创新大多也是粗放式创新,粗放式创新是指只考虑或只主要考虑经济收益的创新,不考虑或不优先考虑创新风险和负面效应(考虑的也是投资的财务风险),不考虑外部性,以牺牲安全为代价的创新<sup>[4]</sup><sup>[238]</sup>。即将发生的新技术革命则体现为对技术本身的控制能力的跃升,目的是实现可持续安全与可持续创新。受控技术主要指技术种类、技术进程、技术应用是受控的,利用上述的科技总体安全观和科技底线伦理为指导。

## 3. 以智库与文创产业崛起为特征的新产业革命

科学革命与技术革命必然引发产业革命。人们常说的四次产业革命(机械化、电气化、信息化、智能化)是传统意义上的技术革命及其产业化的产物,都是 18 世纪工业革命形成的市场经济与科技经济一体化的基本框架内的后续系列,共同之处都是以满足人们眼前需求为主的生产与营销,而即将发生的新产业革命则是转型为满足眼前需求与长远需求并重的可持续创新与发展的产业体系,其研发体系是自然科学研发与社科研发,包括研究与编导等新型研发体系。与以往的产业革命的主体是理工科科技专家

和科技企业家不同，这次新产业革命的主体增加了社科学者、智库专家和智库与文创企业家。高风险的高科技产业将叫停或缓行，底线安全的科技产业将继续发展，智库与文创产业将崛起为主导产业，这场新产业革命的力度空前，远超互联网与人工智能带来的产业冲击。<sup>[32]</sup>这与目前国内外主流共识——依然以理工科科技专家和科技创新为主导的第四次工业革命等趋势外推——完全不同。

#### 4. 人类社会发展模式首次大转型

现行市场经济与科技经济一体化模式因不能克服其内生的致毁知识的增长而必须改弦易辙。从人类活动的根本法则转变的角度看，笔者认为这是人类有史以来的首次大转型：从眼前利益优先、丛林法则至上转变为可持续发展优先、安全法则至上。1945年原子弹爆炸是人类历史上的分水岭，把人类社会的发展分为两大阶段，此前是“车到山前必有路”式的摸索前行，思考未来的思想家、哲人学者是在做锦上添花的事情；1945年之后是“车到山前没有路，出路必须提前谋”，思考未来、研究与规划未来则应该成为雪中送炭的必要工作，未来学、战略学等应该成为显学。但实际上人类社会的反应有所延迟，至今人类总体上还是延续有史以来养成的眼前利益优先、个体利益优先和走一步算一步的试错习惯。人类早已具有毁灭自己的武器和越来越难以驾驭的科技力量，如果任由发展，不约束自己、不认真规划发展道路，根本无法得过且过。当科技危机这一最大危机被揭示出来的今天，这一转变已经凸显，要满足社会关于可持续创新与发展的需求，就要像当年满足社会公共服务需求而建立政府一样，社会要建立以知识分子为核心的新型学界与智库系统，从“市场-政府”二元社会发展为“市场-政府-学界与智库”三元社会，学界与智库不仅满足社会当下需要，更要满足社会可持续创新与发展需要，为此要建立新型的“贡献—分配”系统，主要对当下需求敏感的市场系统进行升级换代，实现对满足当下需要与满足长远需要的激励均衡，改变长期以来对满足人们眼前需要的商人企业家和科技专家过度激励而对满足人类社会长远发展需要的专家学者和机构却严重激励不足的局面，这是史无前例的巨大变革，是化解科技危机等众多深刻危机的根本措施，是智业革命的关键。<sup>[1]201-245</sup>

## 四、结语

综上所述，重视科技的底线安全极为必要。以人类视角考虑，目前人类拥有的科技已经很多，如果善用已经可以给人类相当多的便利和福祉，冒着越来越大的风险去发展尖端科技是否值得？至少发展那些高危的尖端科技肯定得不偿失。当然，如果用个人（科学家等）视角、集体（包括企业和国家等）视角考虑，则非常需要发展科技。何种视角更合理，应以符合可持续发展的长远标准来衡量。“挑战—应战”是思考未来的有效方法，无论是个人、集体还是人类社会，洞察和预知面临的重大危机与挑战是最为必要的头等大事，笔者提出“致毁知识”概念并以其为研究对象开展科技巨风险研究至今20年，做出了上述一系列新发现和新结论。

笔者对科技巨风险的研究表明，科学（也就是西方科学）存在着与生俱来的内在缺陷，随着科技知识力量日益增强，内在缺陷越发成为阻碍科技继续发展的制约因素，长期以来的科技发展与科技创新大多是粗放式发展与粗放式创新，18世纪工业革命以来形成的西方市场经济制度及其科技经济一体化发展模式，既高度依赖科技知识增长与科技创新，又无法阻止其内生的自毁因素——致毁知识——的增长、积累和扩散，必将难以为继、自我颠覆，这是最大的颠覆，科技危机引发的新科技革命与新产业革命也许能够成为最大的颠覆性创新。如果认清这一挑战的严峻性和紧迫性，认识到最大的威胁不是来自其他大国而是来自掌握致毁知识的恐怖分子和极端分子，则完全可以摆脱军备竞赛的囚徒困境，世界各国就应该齐心协力地去应对，而不要延续传统的丛林法则、冷战思维与零和博弈，高科技时代的丛林法则和军备竞赛必然导致玉石俱焚、共同毁灭。

从国家发展与竞争的角度看,科技发展的乐观派与悲观派是不对称的,目前国际国内的主流看法是乐观的,如果乐观派正确,未来发展就是乐观地趋势外推:从信息科技时代到生物科技时代、智能科技时代,从第三次工业革命到第四次工业革命等等,则我国将在相当长的一段时间内处于追赶西方发达国家的状态之中;反之,如果笔者的研判正确或基本正确,则局面完全不同,西方发达国家因路径依赖、积重难返而难以及时转变,我国却具有后发优势,另辟蹊径,应采取换道超车战略,率先发动上述的新科技革命与新产业革命,开创可持续创新与发展的新型道路,新科技与新产业用追求“更安全、更受控、更环保、更永续发展”来超越和替代现行道路上的“更快、更高、更强、更赚钱”的竞争,不比速度、强势、赚钱,比安全、稳健、公正和可持续创新。<sup>[32]</sup>正如一位诺贝尔奖得主所说,将来引领全世界往前迈进的国家,一定是为人类永续发展做出最大贡献的国家,这些国家不一定是目前的发达国家,也可能是目前正在迎头赶上的发展中国家。我国率先认清人类面临的重大危机和挑战,做出及时有效的应对,完全可以彻底摆脱修昔底德陷阱,共建人类命运共同体,引领世界的和平发展和永续繁荣。

智业革命的启动方案。第一步是快速提升我国科技实力与智库实力,解决扼喉技术难题,引领国际思想市场,用笔者提出的“互联网+代表作”方法盘点普查科研和智库领域的突破性成果以发掘拔尖人才,让大材小用的现有的拔尖人才立刻胜出并充分发挥作用,此法简单易行,立竿见影。当务之急是要盘点比较国内外关于人类命运和社会未来发展重大问题及对策研究领域的突破性成果。<sup>[33]</sup>常格不破,人才难得(包拯语),以突破论英雄,用“互联网+代表作”有效落实代表作评价制度,不拘一格降人才,可快速提升我国科研与智库水平,以率先发动科技革命、产业革命乃至智业革命。

#### 参考文献:

- [1]刘益东. 智业革命:致毁知识不可逆增长逼迫下的科技转型产业转型与社会转型[M]. 北京:当代中国出版社,2007.
- [2]刘益东. 科学的目的是追求真理吗? [C]//宋正海. 边缘地带——来自学术前沿的报告. 北京:学苑出版社,1999:17-28.
- [3]刘益东. 试论科学技术知识增长的失控(上)[J]. 自然辩证法研究,2002(4):40-42.
- [4]刘益东. 试论粗放式创新、致毁创新、可持续创新理论与知识安全学[J]. 科技资讯,2014(25):238-242.
- [5]刘益东. 对不准原理与动车困境:人类已经丧失纠正重大错误的能力[J]. 未来与发展,2011(12):2-6.
- [6]莱斯特·瑟罗. 经济探险[M]. 上海:上海远东出版社,1999:29.
- [7]刘钝,王杨宗. 中国科学与科学革命:李约瑟难题及其相关问题研究论著选[C]. 沈阳:辽宁教育出版社,2002:22.
- [8]BECK U. Ecological politics in an age of risk[M]. Cambridge:Polity Press,1995:63-71.
- [9]刘益东. 挑战与机遇:人类面临的四大困境与最大危机及其引发的科技革命[J]. 科技创新导报,2016(35).
- [10]刘益东. 外行评价何以可能——基于开放式评价的分析[J]. 河南大学学报(社会科学版),2016(9).
- [11]MACNAGHTEN P M, KEARNES M B, WYNNE B. Nanotechnology, governance, and public deliberation: what role for the social sciences? [J]. Science communication, 2005, 27(2):1-24.
- [12]RIP A. Future of ELSA[J]. Embo reports, 2009, 10(7):666-670.
- [13]蔡恒进. 超级智能不可承受之重——暗无限及其风险规避[J]. 山东科技大学学报(社会科学版), 2018(2):9-15.
- [14]王鹤峰. 美国高校将人工智能伦理引入课堂[EB/OL]. [2018-09-25]. <http://www.ccidnet.com/2018/0723/10408573.shtml>.
- [15]刘园园. 霍金等千名专家敦促禁止“杀人机器人”[N]. 科技日报, 2015-08-04(2).
- [16]江晓原,刘兵,田松. 伦理能不能管科学[C]. 上海:华东师范大学出版社,2009.
- [17]长侠. 隐匿之地:深网 TOR 揭秘[EB/OL]. [2014-01-13]. <http://www.8btc.com/deep-web>.
- [18]MAXMEN A. 机器学习助力医疗,通过数据分析发现近 6000 种新病毒[EB/OL]. [2018-10-11]. [http://tech.ifeng.com/a/20180321/44914699\\_0.shtml](http://tech.ifeng.com/a/20180321/44914699_0.shtml).
- [19]王彦雨. 学界关于“超级 AI”的论争及其实现的可能路径[J]. 未来与发展, 2017(8):30.
- [20]LIND M. America vs. Russia and China; welcome to Cold War II[EB/OL]. [2018-09-09]. <https://nationalinterest.org/feature/america-vs-russia-china-welcome-cold-war-ii-25382>.
- [21]刘益东. 知识主义社会的理论探索——评李喜先新著《论知识主义社会》[J]. 民主与科学, 2018(2):72.

- [22]刘益东.人类面临的最大挑战与科学转型[J].自然辩证法研究,2000(4):50-55.
- [23]刘孝廷.超越技术与进步——从核风险看人类发展文化的取向[J].山东科技大学学报(社会科学版),2011(10):7-10.
- [24]刘益东.虚拟科学:科学研究的第五范式[J].科技创新导报,2015(9).
- [25]高璐,李正凤.构建STS参与进路:从人类基因组计划到英国基因组学网络[J].科学与社会,2014(1):65-79.
- [26]刘益东.开放式评价与前沿学者负责制:胜出机制变革引发的云科学革命[J].未来与发展,2013(12).
- [27]刘益东.科技巨风险与可持续创新及发展研究导论——以致毁知识为中心的战略研究与开拓[J].未来与发展,2017(12).
- [28]刘益东.影视未来学与科学预演影片初论[J].科技创新导报,2012(32):247-250.
- [29]中国科学院.中国科学院关于科学理念的宣言[EB/OL].[2007-02-27].<http://www.cas.cn/>.
- [30]刘战雄.负责任创新研究综述:背景、现状与趋势[J].科技进步与对策,2015(11):157.
- [31]刘大椿,等.审度:马克思科学技术观与当代科学技术论研究[M].北京:中国人民大学出版社,2017:1-6.
- [32]刘益东.科技危机引发新科技革命和新产业革命[N].社会科学报,2018-09-13(5).
- [33]刘益东.开放式评价与前沿学者负责制[N].中国社会科学报,2016-12-20(1).

## Ruin-causing Knowledge and the Failure of the Ethical Governance on Science and Technology: Science and Technology Crisis and Their Intelligence Industry Revolution

LIU Yidong

*(Institute for the History of Natural Sciences, Chinese Academy of Sciences, Beijing 100190, China)*

**Abstract:** The biggest challenge that human beings face now is the science and technology (S&T) crisis brought by ruin-causing knowledge, which is born from the extensive development of S&T. There are a variety of major barriers when resolving the issue of S&T crisis, such as, the optimism of S&T development, deep-rooted misunderstandings of social functions of S&T, the internal and external defects of their current development mode, failure of ethical governance on S&T, PTSR predicaments, technology cold war and arms race. In the Internet era when knowledge can be easily diffused, ethical governance on S&T always fails as it cannot successfully restrict all R&D activities within reasonable limits, especially for the advanced S&T, such as artificial intelligence. In order to deal with S&T crisis and their huge risk, it is essential to vigorously promote the new scientific revolution featured by the rise of social sciences and interdisciplinary sciences; to greatly push forward the new technological revolution featured by the spring up of controlled technology; to advance the new industrial revolution characterized by the rise of think tanks and cultural and creative industries; and to carry forward the transformation of social development of human beings. Only in this way can humans get rid of the Thucydides Trap successfully, jointly build a community with shared future, and lead the whole world to a mode of peaceful development and sustainable prosperity.

**Key words:** ruin-causing knowledge; science and technology crisis; failure of ethical governance on science and technology; intelligence industry revolution; bottom line of science and technology ethics; research and direct; huge risk of science and technology

(责任编辑:黄仕军)