

北京大学博古睿 研究中心

Berggruen Research Center Peking University

## 科技与社会: 三十年以后的人类

第二次研讨会报告

**TECHNOLOGY AND SOCIETY: MANKIND AFTER 30 YEARS WORKSHOP II** 

报告撰写人 李治霖 项目策划团队 宋冰、展翼文、李潇娇、田馨媛

## 目录

摘要 项目背景及研究方法	01
1. 导言	03
1.2 我们如何思考与想象长远未来——大历史的视角	06
	10
1.4 理解和预测未来的框架性思考:乐观、悲观与叙事	14
2. 未来三十年的科技发展	16
2.1 未来社会与生物工程	
2.2 未来社会与气候工程	21
2.3 小结	25
3. 未来三十年的社会变化	23
3.1 刺激大脑与未来社会	24
3.2 数字模拟与未来社会	26
3.3 小结	29
4. 主要场景	30
4.1 乐观场景:技术赋能的良善社会	31
4.2 悲观场景:灾难锁定的失序世界	33
5. 相关建议	35
6. 开放问题	38
7. 后续工作	41
8. 参考资料	44

## 摘要

#### 总结

第二次"科技与社会:三十年以后的人类"项目研讨会由北京大学博古睿研究中心主办,邀请来自气候变化、生物科技研究领域的科学家、经济学家、对未来有思考的哲学家、科幻作家、艺术家,展望三十年之后,他们各自研究领域发展的状态;以及在他们眼中,技术的发展将怎样塑造未来社会;并尝试为未来情景构建设立由好到坏的基本情景,为我们想象未来科技社会提供基础共识与思想性预设。

三十年只是我们开启这场关键讨论的起点。生命已在地球繁衍上百亿年,人类也早已有数十万年的演化历史,如何在更广阔的时空构想里追溯历史、展望未来是我们的希冀,也是我们应当把技术置于的试炼场。我们期待在未来进一步深化讨论内容,在跨学科、跨地域、跨文明的视域里加强我们对人、技术、生命、存在的感知与理解,更好地应对这样一个剧烈变革的时代。

## 项目背景 及研究方法

#### 背景

博古睿研究院主张,充分利用东西方多元的思想资源,探索能够回应前沿科技挑战、具有跨学科启发性、沟通过去与未来的思想议题。希望结合各学科的思想资源,更好地帮助我们理解变革中的人类世代及其未来。

2021年7月,北京大学博古睿研究中心主办了"科技与社会:三十年以后的人类"首次研讨会,邀请了10位来自人工智能和生物科技等领域的科学家、对未来有思考的哲学家、国际关系研究专家、科幻作家和艺术家共聚长城脚下,展望三十年之后各自研究领域发展的状态;在他们眼中,技术的发展将怎样塑造未来社会。与会者通过跨学科的思想碰撞为未来构设了三个基本情景,为本项目提供了基础共识与议程预设。

在先前讨论的基础之上,我们在 2022 年 4 月 22 日 -24 日举办的第二次研讨会中进一步拓展多学科的视角,邀请关注基因编辑、虚拟现实、人口与经济、气候与环境等技术和议题的学者与人文作家们一同畅想未来,在现实与想象、技术与伦理交织之处"大开脑洞"、碰撞观点,对人类社会的未来前景作出推测,并以此为轴探讨如何设计、调试、与变革科技与社会发展的现有模式。

#### 意义

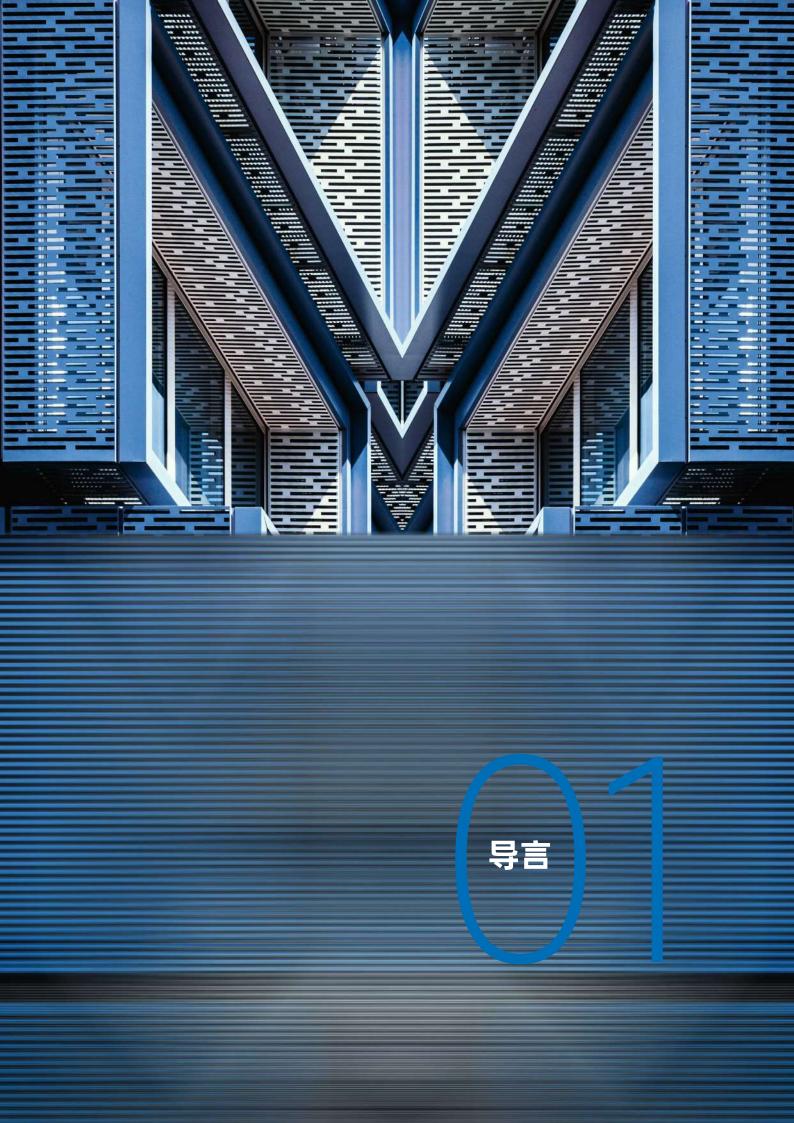
- 构想和预测未来三十年科技发展对生命与社会的可能影响;
- 聚集各领域工作者,互相启发并构建一个面对科技未来的思想共识;
- 激发更多更深入的讨论,为技术的进步提供可能的底层价值观基础。

#### 方法

- 闭门会议式对谈,与会者畅所欲言、互相对话;
- 哲学工作者、科技学者、文艺工作者(科幻作者、策展人等)、业界专家将分别领导讨论,预留充足的时间交流观点,并对未来作出场景假设。

#### 主要参与者

- 安乐哲 (Roger T. Ames),北京大学哲学系人文讲席教授、北京大学博古 睿研究中心学术委员会联席主席
  - 宝树(李俊), 科幻作家、译者、2022-2023 博古睿学者
  - 陈海丹,北京大学医学人文学院医学伦理与法律系长聘副教授、2021-2022 博古睿学者
  - 陈霞,中国社会科学院哲学研究所研究员、博士生导师、2022-2023 博 古睿学者
  - 陈迎,中国社会科学院城市发展与环境研究所可持续发展经济学研究室研究员、中国社会科学院可持续发展研究中心副主任
  - 葛鉴桥, 北京大学前沿交叉学科研究院讲师、2021-2022 博古睿学者
  - 何展,英伟达中国 Omniverse 业务负责人
  - 黄有光(Yew-Kwang Ng), 莫纳什大学(Monash University) 荣休教授、 澳大利亚社会科学院院士
  - 刘宇昆 (Ken Liu) , 幻想作家、未来叙事学者
  - 龙星如,独立策展人、写作者、2022-2023博古睿学者
  - 吕晓宇,青年作家、北京大学国际关系学院助理教授、2022-2023博古睿学者
  - 十三维,心识宇宙研究院院长、集智俱乐部核心作者、科学作家
  - 汪阳明, 北京大学未来技术学院教授、2019-2020 博古睿学者
  - 展翼文,北京师范大学哲学学院讲师、硕士生导师



## 人类社会的长远未来及其挑战

至今,人类与持续肆虐的新冠疫情已同行近三年,"环球同此凉热"从未这么真 实地存在过。疫情的涟漪效应加剧了世界政治经济的不确定性风险,而人工智能、 脑科学、生命科学、新材料新能源技术则不断迎来重大突破,以智能科学为先 导的技术革命曙光初现, "元宇宙"似乎也近在咫尺。

与此同时,人类还远未到可以庆祝的时候——极端气候变化的警钟长鸣,人口与 老龄化的危机不减,大国冲突的加剧与战争的疑云笼罩着大洋两岸。一切不期 然的、行星级别的巨变都在参与塑造着人类社会的生活方式,深刻冲击着人类 自身的认知和对未来世界的理解,也将人类置于史无前例的生存性风险之中。 这不禁令人发问: 技术影响下的人类的未来将向何处去?

在这样的挑战面前,跳出当下视野的认知偏见和局限性、超越单一学科知识的 讨论显得尤为重要。在时代与环境势不可挡的种种发展进程下,人类的未来取 决于今天各个领域内的行动者与思想者--科学家、人文社科学者、作家、艺术 家等直面挑战的智慧与勇气,取决于我们是否敢于以跨时空的想象,敢于以跨 学科、跨文明的协作,为全人类的长远未来谋篇布局。

在大变革来临的前夜,思想理应早做准备。"科技与社会:三十年以后的人类" 系列研讨会便是这样一种尝试。北京大学博古睿研究中心希望汇聚不同的学科 与范式,用理性与感性、审慎与胆识来创造共识,预见趋势、洞察可能。

以"未来场景构建"的形式,本项目计划邀请来自学界、产业界、政策制定以及 文学艺术界的专家线上线下相聚,聚焦人工智能、生物科技、人口健康、气候 与生态变化等相关议题,通过对多种可能的乐观与悲观的全球未来的构想,探 讨如何设计、协调、乃至变革科技与社会发展的现有模式。我们计划探索,在 未来三十年,乃至更长的时间尺度上:

- 技术发展: 基因编辑与虚拟现实技术可能 / 可以达到怎样的程度? 在其影响下的人类 社会的交往与互动方式将面临怎样的未来图景?
- 政治经济未来: 世界的经济发展及其模式、内生动力将会产生怎样的变化? 与之相对 应的,全球地缘政治格局将会发生何种改变?如何在冲突中促进全球合作?
- 气候与环境变化:全球变暖与环境恶化的程度会怎样?人类社会目前已经作出的准备 是否足够,什么样的能源或发展模式转型、前沿科技能干预这一态势?
- 人口与健康:全球人口结构与分布会面临怎样的大调整与重构?这一变化带来的挑战 是什么,将如何影响社会发展和公共健康?
- 面向未来的哲学与伦理学:在思考人类社会的未来时,哪些哲学观点、伦理学体系值 得借鉴? 当前的人类是否对未来的人类(或未出生的人)负有一定的责任?

这些问题没有轻易的答案。但当下的我们必须意识到: 伴随着技术能力的不断 提升,人类早已开启了深度塑造地球及其未来的进程。正如鲁迅先生所言,"无 穷的远方,无数的人们,都和我有关。"我们不能也不应忽视当下之举可能给 未来带来的"危机"——"危"险和"机"遇,审慎地看待无限进步的科技叙事, 批判性地反思科技发展的目的与终局,并重视全球协作、有效应对挑战。

## 我们如何思考与想象长远未来 --大历史的视角

在展望未来之前,让我们回望过去--生命在地球上已存在 38 亿年之久。而对于人类而言,这 个时间表已然缩短几个数量级:约 700 万年前出现乍得沙赫人,约 600 万年前出现原初人,约 580 万年前出现地猿, 约 420 万年前出现南方古猿, 约 250 万年前出现能人, 约 180 万年前演 化出直立人,约 30 万年前出现智人,也就是我们现代人类的始祖。如果是文明呢?耶稣、孔子 和佛祖诞生的轴心时代不过2500年之久。人类社会赖以生存的文明基石,放在整个生命系统的 变化逻辑里是微不足道的。

同样地,如果我们抽离出当下迫近的视野,在更大范围的视角下谈论未来,真相可能会让我们震惊: 假如地球能持续提供适合人类生存的环境,从其他哺乳动物的平均寿命推测,人类文明至少还有 数十万年的篇章等待书写。如果人类历史是一部小说,我们大概率仍在扉页。这也意味着:我们 当下的行动可能会影响到无数个未来世代所处的世界,他们生活得好与坏、精彩纷呈亦或苦不堪 言,很可能取决于今天的我们。

这么说不是狂妄自大。尽管人类文明从轴心时代至今仅仅走过了两千多年,但在过去的两百年里, 我们经历了前所未有的技术进步和社会发展。全世界的 GDP 总量、对自然资源的利用能力、人 均预期寿命、人均生活质量等在工业革命后全部指数型提升,而这些指标在 19 世纪前的两千多 年间几乎没什么变化。换句话说,人类历史可能正处于一个特别而关键的时期:我们奇迹般地跃 进到了一个远比过去要富足、美好、公正、充满机遇的时代,但与此同时,也第一次掌握了能彻 底摧毁自己的能力。核战争、极端气候变化、致命的大流行病、失控的人工智能等科技发展带来 的风险正在逐步彰显,而我们似乎可以在现在做些什么,为那些尚未出生的子孙后代们留下一个 持续繁荣而安全的世界。

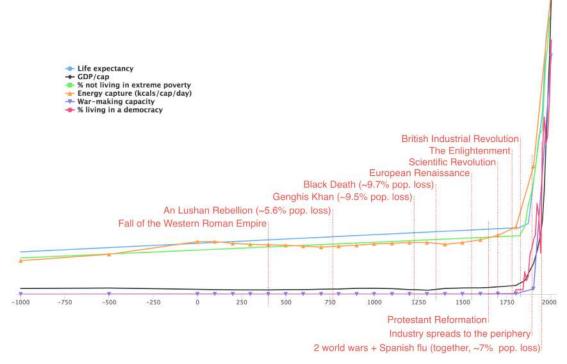


图: 人类社会过去三千年的发展态势 来源: Three wild speculations from amateur quantitative macrohistory, 2017, Luke Muehlhauser

这种关切是非常自然的:我们都希望我们的子女和孙辈的生活能够顺利,哪怕他们现在可能还没有出生。正如安乐哲所说,人类对未来的关照某种程度上与中国儒家所强调的"孝"文化有关系,当下的人们会对未来的子孙后代有延续的爱,而亦对先人留下的家业有维持之责,不得轻易置人类社会于危险或动荡之中。而过去数年间,在英国牛津、剑桥,亦有一些分析哲学家们提出了一套系统的叙事来阐释守卫人类未来的道德紧迫性,这套论述被称为"长远理念"(Longtermism),其主张塑造一个更美好的未来是当下最具有道德紧迫性的议题之一。

#### 这一结论主要立足于三个判断:

- 1. 未来的人很重要,值得我们关心。
- 2. 人类社会可能会延续到极其长远的未来,这意味着未来有机会存在的人类的数量将是庞大的。
- 3. 人类社会当下的行动,或将显著地影响人类长远未来的走向。

因此,长远理念认为,确保人类社会有机会长久地繁衍生息、实现其潜力,是当前人类的一种责任。 这一具有野心和潜力的理论开始于已故哲学家德里克·帕菲特(Derek Parfit)的一个思想实验:

想象把碎玻璃埋在一片森林里。在一个可能的未来,五年后,一个孩子踩到了玻璃受了伤。在另 一个可能的未来,五百年后,一个孩子也踩到了玻璃受了伤。这两种可能性似乎同样糟糕:我们 不应因为一种伤害发生在遥远的未来就认为它不重要。如果我们相信"人人生而平等"--无 论他出生在哪里,同理,无论他何时出生--是现在还是数百年后,都值得我们平等对待、 纳入考量。在时间上早一点出生并不是一个人更重要的理由。

从我们祖先的角度来看,我们是未来的人,我们和他们之间隔着很多代人。如果仅仅因为我们出 生较晚就说我们的生活不那么重要,那是不公允的:我们的快乐和痛苦,与我们之前几代人的快 乐和痛苦,一样真实,一样重要。生活在遥远的未来的人们,当他们将自己的生命价值与我们的 生命进行比较时,也会有同样的感受。不过,亦有一些经济学家指出,由于未来的生命尚不存在, 且他们是否能出生取决于人类文明是否能避免灭绝、持续演进,因此,他们的生命感受、福祉和 快乐应当折现(discounting)后与当下人类的福祉进行比较,尽管这不会影响"未来的人值得 我们关心"这一结论。

如果接纳未来的人值得我们关心这一结论,紧接着的一个重要问题便是:人类未来的可能图景是 怎样的?有多少人可能生活在未来,他们的生活会有多好(或多坏)?

前面提及, 如果智人存在的时间与大多数哺乳动物大致相同, 人类至少还有 20 万年的剩余时间。 不过,与其他物种不同:我们似乎有技术手段来防止自己的灭绝。例如 NASA 最近成功开展的小 行星防御计划,就让人类第一次掌握了改变地外星体撞击轨道的能力,让我们有机会避免曾经的 地球霸主(恐龙)灭绝和消失的命运。我们看起来能比一般的哺乳动物生存和延续更长的时间。

从外部条件来看,我们的星球可能会在几亿年内保持宜居性(在地球被太阳吞噬之前)。即使人 类只活过这段时间的 1%, 那么在我们之后仍将有大约四千亿人出生。但是人类一定与地球相伴 相生吗?考虑到从发明第一架飞机到人类登月只相隔了66年,我们很难想象在未来的一百年或 一千年内,可以在航天技术上取得多大的进展,进而允许我们成为一个跨行星甚至跨恒星的星际 文明。这听起来可能太科幻,没有扎实的证据支持。但重要的是,人类的未来,在时间长度和空 间尺度上,可能都是极其广阔的。未来的规模可能难以理解,一旦数字变得足够大,它们往往听 起来都是一样的。为此,物理学家詹姆斯·金斯 (James Jeans) 提出了以下的比喻:

想象一枚硬币上有一张邮票。如果硬币和邮票的厚度加起来代表我 们这个物种的生存历程,那么邮票的厚度就代表有记录的人类文明。 再想象把这枚硬币放在 20 米高的纪念碑上。如果邮票厚度代表了整 个人类文明的历史,那么纪念碑的高度就代表了地球的年龄。现在 我们考虑未来。放在方尖碑上的一棵 5 米高的树代表着地球宜居的 未来。而背后,马特洪峰的高度(阿尔卑斯山的象征,海拔4478米) 代表了宇宙宜居的未来。

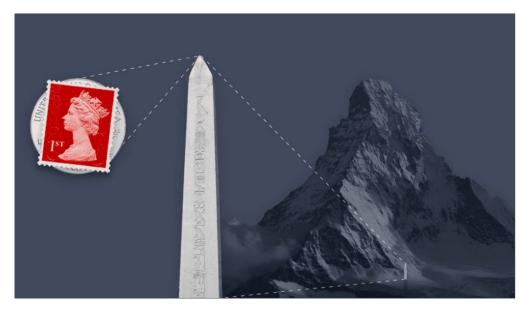


图: 马特洪峰与宇宙宜居的未来 来源: The Mysterious Universe, 2019, James Jeans

当然,只有当未来人类的生活是美好而令人期待的,我们才应当谋求人类在这些巨大的时间尺度上的生存。幸运的是,我们有理由相信,世界的未来会更好。人类已经取得了惊人的进步:生活在极端贫困中的人口比例从 1820 年的约 90% 下降到 2015 年的不到 10%。在同一时期,儿童死亡率从 40% 以上下降到 5% 以下。我们有理由期待进一步的科学和医学突破将继续改善生活。而那些我们珍视的价值、艺术之美、制度之善、亲情友情爱情的实存,也将随着物质文明的延续而传递下去。

我们不知道人类的未来到底会是什么样子。未来可能非常好,也可能非常坏,而且未来的尺度可能很大——绝大多数可能存在的人,和其他有价值的东西,都在未来。如此来看,确保人类文明的潜力被充分发挥,积极地改善后续数千代人的生活,就是一个不歧视那些在时间尺度上离我们遥远的同类的关键议题之一。

这里值得额外指出的是,当我们提及生命、文明延续这些概念的时候,坐标系往往被默认设置成了"人类中心"的。好像只有人类才是这个星球唯一的主人,是唯一值得被道德考量的生命。殊不知,如果我们认为人人生而平等,无论其生在何地、何时,那么这一不偏不倚的关怀(impartial care)也理应被拓展到其他非人类动物(nonhumans)当中去。正如著名效用主义哲学家、2021 博古睿奖得主彼得·辛格(Peter Singer)所言,"如果我们拒绝种族主义和性别歧视,为什么我们应该认为,仅仅因为一个生命不是我们的物种成员,就可以接受更差的待遇?"辛格在他的书籍《动物解放》中将当代社会这种对动物道德地位的系统性漠视称之为物种歧视(Speciesism),并指出,一个生命是否重要不在于其是否具有理性,能否说话,而在于其是否具有苦乐感知力。如果绝大多数的动物都能像人类一样感受疼痛,趋利避害,我们没有理由在考虑自己行动的后果时把它们排除在外。森林里的一片玻璃,在现在或将来,无论是伤害了一个孩子,一只猫,一匹狼还是一只猪,都是不好的、应当避免的。部分长远理念的倡导者们也意识到了这一点。因此,在论述未来的重要性时,我们可以把条件中的宾语进行范围扩展——"未来的生命很重要"、"未来将会有非常多的生命出生"、"我们当下的行动将会显著影响地未来多物种地球文明的走向"。

不过,长远理念仍待论证的一个命题是:人类社会现在的行动真的会对未来产生决定性的影响吗?

## 直面生存性风险--保有人 类文明的潜力

近年来,人工智能、纳米技术、合成生物学、基因编辑技术、量子计算机等新兴科技快速发展, 影响力越来越大,呈现出科技大爆发态势。这些新兴技术的发展,昭示了人类开启对自然和人类 自身进行调控、设计乃至制造的深度科技化时代的来临,同时,伴随而来的科技风险问题愈发严重。 我们有理由认为,今天的我们,正处在一个能够对人类未来产生非比寻常影响的历史时刻。

如前文所述,人类可能的未来不只是数百年甚至数万年,而是以亿年计--自然,前提是人类及 其发展潜力不会毁于自然或自己之手。以如此宏大的时间尺度来衡量,人类尚处于幼年。就像每 个人在懵懂的童年时光里都会犯下很多用成年人的眼光来看笨拙的错误一样,身处这个时代的我 们的所作所为,或对一些风险的忽视,可能在更为成熟的后人看来会是极度危险、不负责任的。

换句话说,考虑到目前人类科技发展的态势,我们必须谨慎地对待一种可能性,即我们可能会作 茧自缚、"自取灭亡"。这可能立即会引起类似"危言耸听"或"杞人忧天"之类的评价,但对于 生活在美苏冷战时代核战争阴影之下的很多人来说,这类风险却是真切而实际的。全面核战争的 危害不仅包括核爆炸本身、大火、放射性尘埃引起的杀伤,更包括其所引起的"核冬天"--核爆 炸引起的城市火风暴、巨大的烟柱使得黑色烟尘飘向平流层,遮挡太阳光使地球气温骤降,摧毁 全球农业系统,进而引发的大规模饥荒。根据牛津大学学者托比·奥德(Toby Ord)的估算,人 类在接下来一个世纪中遭遇类似核战争这样生存性灾难的概率约有 1/6。也就是说,与骰子掷出 某个特定数字差不多。

所谓生存性灾难(existential catastrophe),奥德的定义包括但并不限于人类灭绝。他将其定义为"能摧毁文明长期发展潜力的事件"。通俗地说,如果一场全球性灾难抹去了大部分但不是全部人口,但与此同时让人类丧失了可利用以重建文明的技术、文化、制度和自然资源,则也属于生存性灾难的范畴。这一类灾难发生的风险即为当前我们地球文明所面对的生存性风险(existential risk)。

这一定义的特别之处在于,事实上,由于人类走出非洲之后至今已经遍布地球上包括南极洲在内的各个角落,能让人类或所有物种即刻灭绝的灾难已经较少了。地球上的生命正愈发地具有生存的韧性。但一些自然或人为的重大恶性事件完全可能剧烈地改变地球环境,使得灾后存活下来的人们像一些反乌托邦文艺作品里描绘的那样,只能一直以低水平的生活质量和落后停滞的技术能力苟且偷生。如果文明不再具有发展、进步的能力,不再能不断拓展认知边界、为多物种的共生共存、美好未来发挥潜力、创造可能性,那同样是一种可与"灭绝"相比的灾难性后果。

Existential catastrophe via	Chance within the next 100 years	
Asteroid or comet impact	~ 1 in 1,000,000	
Supervolcanic eruption	~ 1 in 10,000	
Stellar explosion	~ 1 in 1,000,000,000	
Total natural risk	~ 1 in 10,000	
Nuclear war	~ 1 in 1,000	
Climate change	~ 1 in 1,000	
Other environmental damage	~ 1 in 1,000	
'Naturally' arising pandemics	~ 1 in 10,000	
Engineered pandemics	~ 1 in 30	
Unaligned artificial intelligence	~ 1 in 10	
Unforeseen anthropogenic risks	~ 1 in 30	
Other anthropogenic risks	~ 1 in 50	
Total anthropogenic risk	~ 1 in 6	
Total existential risk	~ 1 in 6	

图: 人类在未来一百年遭遇不同生存性灾难的概率 来源:《危崖》, 2021, 托比·奥德

奥德指出,这种"生存级别"的风险,可能包括自然类的风险,如小行星或彗星撞击、超级火山喷发、 恒星爆炸等;也可能包括人为风险,如大规模核战争及核冬天、气候变化、环境破坏等;还有暂 未发生,但可能会在未来变得显著的技术风险问题,如人为增强的病原体所造成的致命大流行病, 及与人类价值观不一致的超级人工智能(superintelligence)等。

如果我们承认如上假设的合理性,帕菲特的另一个思想实验形象地说明了生存性风险在伦理上的 重要性。他让我们设想和比较如下的可能性:

#### 1. 和平。

#### 2. 一场杀死 99% 世界现有人口的核战争。

#### 3. 一场杀死所有人的核战争。

这三个选项哪一个更不好呢?大多数人或许会相信,更大的区别在于(1)和(2)之间。而帕菲 特则认为,(2)和(3)之间的区别要远远大于(1)和(2)之间的差别。他的考量是,在接 下来的 10 亿年内,地球将仍然适合人类居住。文明的出现只是最近几千年的事情,如果我们不 毁灭人类, 那几千年不过是全部人类文明史的一个微小片段。(2)和(3)之间的区别在于这沧 海一粟和剩下的整个人类文明历史之间的差别。那些相似于"一场杀死所有人的核战争"类型的 灾难,之所以尤其糟糕,就在于它毁灭了人类未来的巨大可能性。

基于这些考量,我们终于能回到长远理念的最后一个命题:人类社会当下的行动,或将显著地影 响人类长远未来的走向——尤其是人类是否能兑现自己无限的潜力,成为一个繁盛、持存的文明,

而不是过早地因为傲慢、鲁莽与不自知,毁于我们自己所掌握的能深刻改造世界的科技能力。具体来说,这些行动可能包括为未来的大流行病制定对策,如开发早期检测新型病原体的系统,对合成生物学、基因编辑技术进行更具实效的监管;通过技术研发确保人工智能系统,尤其是通用人工智能,可以与人类的价值观对齐,永远理解并执行我们真正想要它们做的事;对碳捕捉和收集技术、地球太阳工程等手段的可能作用、负面影响、实施方式开展研究,让人类有办法面对气候变化带来的极端升温情况;促进大国间的战略互信和文明层面的相互谅解,避免最坏的脱钩和冲突场景;克服国际政治普遍存在的短视,增强保护后代的治理机制,正如联合国秘书长在2021年发布的《我们的共同议程》报告中所言:



图: 我们的共同议程——秘书长的报告 来源: 联合国大会, 2021, 安东尼奥·古特雷斯

在重铸代际团结的这项努力中,不仅应考虑当代人的利益,还应顾及子孙后代。人类面临一系列长期挑战:全球变暖和环境退化、对人工智能和基因编辑等新技术的管理、人口日趋老龄化、城市化以及社会福利供给的演变,这些挑战将跨越几代人的时间不断嬗变。随着第四次工业革命的到来,我们正处于近代历史上最关键的变革时刻之一。然而,占主导的政治和经济诱因仍极度趋向急功近利和维持现状,只重眼前利益,不惜牺牲人类和地球的长远福祉。今日作出的决策将影响未来几个世纪地球上的发展轨迹。我们需要扩充对《联合国宪章》中"我联合国人民"的理解,不仅要保护二十一世纪所有民众的利益,还要为子孙后代留下一个宜居的世界。

## 理解和预测未来的框架性思考: 乐观、悲观与叙事

长远理念提出了一种结合伦理学、大历史与科技发展的未来学思考范式,除此之外,理解和预测 未来一直都存在着多学科、多视角融汇的趋势。

自 20 世纪中后期以来,不仅未来学家,还有自然科学和人文社会科学家、工程师,甚至政治家、 企业家都参与到人类社会未来趋势的探讨中,涉及的领域从科学技术、环境资源到政治、经济、 军事、国际关系,从人类本身、地球到空间。对于未来的畅想,几乎是人的本能,它既是科学范畴, 又超越科学。只不过,讨论世界未来趋势,终究无法回避所谓的乐观与悲观的倾向问题。做出这 种判断,不仅受到相关学者、科学家和工程师的血缘和出身地域差异的影响,更与学科有着紧密 的关联性。

自然科学家倾向于乐观, 社会科学家则倾向于悲观。技术决定论者、工程师思维的持有者认为"自 己创造的新技术将会开创美好的未来",今后的种种问题都可以依靠科学技术来解决,宣扬"乌 托邦论";而人文社会科学家则"在零和博弈的世界观下坚持悲观论调",不断提出诸如人口增 加而带来的资源匮乏、全球变暖等问题,突出缺乏伦理考量的技术进步可能带来的致毁风险,相 信某个版本的"末世论"。亦有所谓"中间派"指出,任何学科叙事的拥护者都有内在的价值取向 和情感特征,在他们探讨未来趋势的时候,不可避免地抱着"我希望能这样"的愿望。不存在绝 对客观的未来趋势的推论与判断。但似乎, 悲观的思考更具启发性和变革性--以罗马俱乐部的《增 长的极限》为例,固然其中的不少观点、论据都存在不够严谨的问题,甚至其结论也不为相当多 的专家学者所赞同,但是,正是这个报告开启了人们严肃对待环境和资源问题的历史,甚至可以 认为,世界各国政府和民众所形成"可持续发展"的共识根源于此。

可以说,只要人类存在未来,未来趋势的预测就会继续下去。只是,预测的方法、框架和模式都要适应这个不断加速演化的人类技术社会。如果说,人类社会向深度科技化的"后人类"社会的转型已经全面开始,那么,关于世界未来的预测从本质到形式也都要实现根本性转变。在这方面,人类智慧所面临的挑战是严峻的。

这种挑战当然也可以导向某种对理解和预测未来可能性的否定。正如刘宇昆(Ken Liu)所言,在每个时刻,无数人试着从不同的方向解决问题,我们在某一时刻往前看,很难预测哪一个方面会有突破。但过去了之后往后看,很容易得到一个结论:我们走过来的路是必然的路。人喜欢把因果关系牵强附会地联系起来讲故事。历史上很多事情其实是无法解释的,都是一个个偶然而已。而部分艺术形式超越了学科的藩篱——比如科幻小说,其价值不在于预测未来,而在于给未来一种可能性的表达。它让我们对未来和科学充满一种敬畏感,赋予我们一种语言和思维方式来探索未来。反过来,科幻的艺术激发了我们对于未来的思想,我们创造未来的时候也就受到了科幻的影响。

这便是叙事的重要性。在理解和预测未来的时候,我们当然需要理性和审慎的判断力,思考:

- 1. 最好场景: 厘清人类社会发展演进的最终目标和形态, 树立一个值得为之奋斗和实现的彼岸;
- 2. 最坏场景:识别快速发展的技术社会潜在的威胁和灾难征兆,为人类文明的持存未雨绸缪;
- 3. 最可能场景: 做出基于合理假设的最佳预测, 基于最大可能性的未来思考技术与制度的布局。

但同时,我们也应该重视和尊重人类社会的道德共识和情感直觉,反思:

- 4. 与构建未来有关的价值观: 当我们想到子孙后代时, 我们最希望给他们留下什么观念遗产;
- 5. 与文明存续有关的共识: 支持人类文明持续繁衍的底层价值判断是什么, 当前的我们共享什么动机, 又有哪些争论;
- 6. 个体如何面对未来的可能性:未来的广远尺度和危机四伏究竟如何影响个体当下的生活,"关怀未来"与"活在当下"究竟有哪些张力和调和的可能性。

第二次"科技与社会:三十年以后的人类"研讨会即在以上的框架下提出问题、促进讨论。学者们秉持着一种既不盲目乐观、亦不妄自菲薄的态度面对未来及其提出的紧迫挑战,正如季羡林先生所言:

对于人类的前途,我始终是一个乐观主义者。我相信,不管还要经过多少艰难曲折,不管还要经历多少时间,人类总会越变越好,人类大同之域决不会仅仅是一个空洞的理想。但是,想要达到这个目的,必须经过无数代人的共同努力。有如接力赛,每一代人都有自己的一段路程要跑。又如一条链子,是由许多环组成的,每一环从本身来看,只不过是微不足道的一点东西;但是没有这一点东西,链子就组不成。在人类社会发展的长河中,我们每一代人都有自己的任务,而且是绝非可有可无的。如果说人生有意义与价值的话,其意义与价值就在这里。



## 未来社会与生物工程

自人类社会进入 21 世纪,生命科学的发展迅速,从生物克隆到基因编辑,从特效靶向药到细胞 培养肉,我们对生命现象的理解由浅入深,几经还原与重构。在北京大学未来技术学院教授汪阳 明看来,生命的本质就是编码,是一个个句子组成生命的表达,而基因就是写下这些句子的语言。 而在过去的几十年内,人类似乎已经部分掌握了这些字符串排列组合的秘密,进而设计出了不在 少数的生物和基因工程技术。不过,更多的进展似乎还在未来,即使是展望三十年,社会也很可 能会因为生物工程的发展而发生不小的改变。

汪阳明介绍道,现存的成熟的生物技术主要包括以下几类: 1. 日益完善的测序及分析技术; 2. 日 益完善的基因编辑技术; 3. 多能干细胞获取和分化技术; 4. 体外受精技术; 和 5. 体细胞核移植 技术(克隆技术)。而在研的技术则包括: 1. 从多能干细胞分化精子和卵子; 2. 人工类囊胚构 建技术; 3. 哺乳动物胚胎体外培养技术。

以基因测序和编辑为例,曾经基因相关的研究花费高昂:2001年平均每兆数据量基因测序成本 是 5292.4 美元,单人类基因组测序成本是 9526.3 万美元; 20 年后单人全基因组测序服务的价 格大约是数百美元,人类基因测序的成本下降速度比摩尔定律(即每一美元所能买到的计算机性 能,每隔 18 ~ 24 个月将翻一倍以上)还要快 50%,这大大加速了该技术的普及与应用,让更 多人能从基因测序中受益,获得检测多种疾病、药物用药指导的健康价值。同样地,基因编辑发 展一直保持着超高速势头,和每一个人的联系变得越来越紧密: "基因剪刀"技术 CRISPR 过去 数年在治疗癌症等疾病、开展异种器官移植、使农作物增产增收方向上都在不断取得进展,我们 有理由相信其将深入地帮助人类应对医疗、粮食短缺等问题。

尽管基因编辑技术的科研是为了治疗或者预防某些疾病,但是在技术层面,基因编辑技术本身尚 未真正成熟,人类对基因编辑的后果并不能做好准备。生命的基因序列存在于一个巨大复杂的因 果链条中,人类难以预计经过人为修饰的基因在后续的遗传中会造成怎样的风险。正因为生殖系基因组编辑的影响会经过生殖过程而世代传递下去,因此,相关科研人员和决策机构理应负有相应的道德义务来防范潜在风险,避免对未来世代造成伤害。在这一点上,基因检测、以治疗为目的基因修饰(genetic modification)、以预防为目的的可遗传基因编辑不仅仅是与被基因编辑者自己相关的个体事件,我们在思考被编辑个体的最佳利益和制定相关法律政策时,需要考虑这一行动对群体和人类共同体造成的影响,利益相关者乃至未来世代的利益。

而随着研究的不断深入,基因编辑还可能带来基因增强、"设计婴儿"等无限可能。潘多拉魔盒一旦打开,我们在未来三十年可能会面临:

- 1. 优生学的阴影:从人类胚胎开始的"优势人种"的选择问题,也就是人类有可能通过基因编辑的"手术",从胚胎开始就改变婴儿的基因片段,人为地剪除或增加某些性状的特质,可能出现所谓的"超级改造人"的情况。我们必须诘问:人类真的有权利决定后代的基因性状,替自己的后代选择去除某些基因缺陷,或增加某些优势基因性状吗?
- 2. 人造生命的诱惑: 我们现在有了被重新设计的树木、鲜花、玉米、奶酪, 甚至已经有科学家创造出了 RNA 的前体, 也就是说, 可能十年, 或者二十年之后, 人们也许可以从零创造出生命。那么, 人类又将何时被怎样重新设计? 我们的社会该如何在道德与规范上接纳人造生命这一族群?
- 3. 大脑时代的认知失调:在新的基因生物时代,肢体、器官都将逐渐可再生,而大脑将成为核心——这才是"你",剩下的都只是"包装"。"大脑时代"可能不再是赛博朋克中的一个设定,而成为某种现实:甚至在目前,实验室内培养的大脑是有意识和苦乐感知力的吗?假若人的身体成为忒修斯之船,我们对人的主体性与自我同一性的认知又遭遇什么挑战?



? ? ? ? ?

人造智人

人造生命

同性生殖

体外生殖

复活古生物

延长寿命

异种嵌合动物 (医用)

基因编辑动物 (医用)

人造肉等食物 基因编辑食物



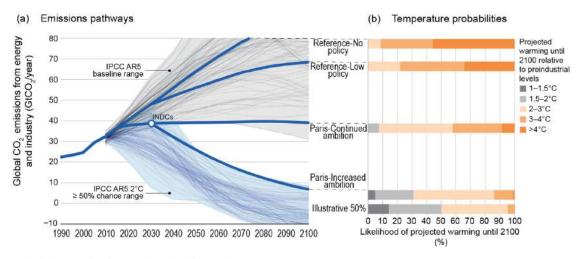
图: 生物工程带给人类的可能性 来源: 展望生物工程的无限前沿, 2022, 汪阳明

这些问题都在技术现实里不断涌现,绝非危言耸听。如果我们不尽快正视并早做打算,许多我们珍视的伦理规范都在未来三十年面临着冲击乃至塌陷的风险。或许,正如我们对克隆人保持了极高的警惕和严格的禁止,在取得共识之前,各个国家、国际组织和学术机构也应该对其他各种生物工程技术的应用边界尽快制定相应的措施和立法,在可预见的时间里,守住生命科学探索的一些禁区,为生命的尊严和自由保留缓冲时间。与此同时,更重要的是需要在底层的思想方法论上破除一种技术决定论的构想,在多维度、多面向的学科与研究框架里理解生命、理解时空里演化着的生命形式。让人类酌情酌理地选择技术发展,而不是让技术无可挽回地决定社会的面貌。

## 未来社会与气候工程

过去十余年来,由温室气体(主要是二氧化碳)浓度升高造成的气候变化已经从边缘的叙述走入主流视野,成为公认的全球危机,不仅威胁人类生产生活,甚至可能改变人类文明的进程。在中国社会科学院可持续发展研究中心副主任陈迎教授看来,未来三十年,气候危机将持续成为那个笼罩在人类社会之上的阴影,没有任何共同体能再忽视这样一种紧急状态。

当前,气候变化带来的环境危机已经显现,极端天气正在变得愈加频繁。全球气候风险呈现加剧趋势,高温、强降水、干旱、风暴等极端天气事件频发,更多证据将这些现象归因于人为活动。随着全球变暖加速发生,由多种极端气候事件组合的复合事件发生次数也将上升。根据气候科学研究结果,从 1880 年到现在,全球已经升温 1.1 摄氏度,接近"无法接受的风险水平"。人类活动持续产生创纪录的碳排放,仍没有提出明确的脱离化石燃料的路径,目前升温的速度是6500万年以来最快的。



图片: 本世纪不同气候场景及其升温情况的预测 来源: Fourth National Climate Assessment: Climate Science Special Report, 2021, U.S. Global Change Research

联合国政府间气候变化专门委员会于 2022 年 4 月发布的报告强调,如果不立即在所有部门进行 深度减排,将全球变暖限制在1.5℃以内将毫无可能。在气候模型的最坏假设下,2100年全球气 温可能会升高4摄氏度或更多。地球已经数百万年没有那么温暖了,而历史上的这种温度峰值与 大规模灭绝事件具有关联性。气候变化作为一种"生存性"风险,正在逐渐迫近我们的生活轨迹。

从应对措施来看,减少排放是必要而根本的。但是,现阶段的科学研究和各国实际的减排行动表明, 我们无法在短时间内将碳排放削减到安全限值。在这种背景下,有没有一些新的、更大胆的方法 可以让地球更快地降温,并为我们摆脱化石燃料争取更多时间?地球工程就是作为这样一剂暂时 的镇痛剂,进入学界视野的。按陈迎的说法,地球工程是一剂猛药,我们不希望最后只能吃下这 个有副作用的药,但是我们必须为最坏的场景提前做准备,避免在灾难面前无计可施。

"地球工程"的定义为"为了应对全球气候变化而对气候系统采取的有计划、大规模的人工干预活 动",也称 "气候工程"或"气候干预"。这些技术通常被分为太阳能辐射管理(solar radiation management, SRM) 和二氧化碳移除 (carbon dioxide removal, CDR) --前者泛指一类通 过将太阳光反射回外太空或允许更多红外辐射逃逸来降低温度的技术;后者则主要指物理提取大 气中的二氧化碳并将其储存的手段,如利用生物能结合碳捕获与封存技术,形成"负排放"。目前, 二氧化碳移除类技术已开展商业化运营,因为可以直接减少大气中的二氧化碳,已经被学者普遍 认为是我们需要采取的气候变化应对措施。但这种方法在实施上仍具有很大挑战性,以"碳捕集 与封存"技术为例,不仅成本高昂,而且需要占用大量的土地。

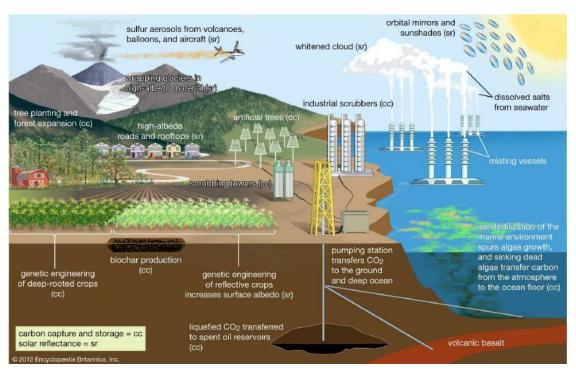
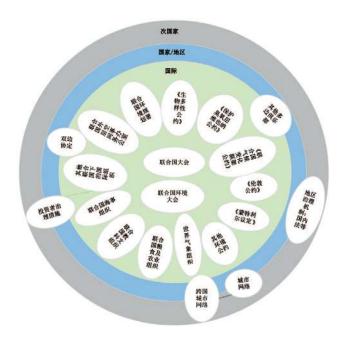


图: 气候工程技术方案汇总 来源: Encyclopædia Britannica, 2022, Philip Boyd

而太阳能辐射技术则由于未知气候风险和管理争议,暂未被主流行动方接纳。这类技术提议都显得脑洞大开——平流层气溶胶注入(stratospheric aerosol injection)计划将硫酸盐粒子注入平流层,从而模仿火山喷发的影响,起到将一部分阳光反射出去的作用。另一项技术倡议——海洋云增白(marine cloud brightening)则是通过向低空云层喷射更多的云层形成粒子来提高海洋云的反射率。更有太空反光镜法,即将一组镜子送入太空轨道以反射阳光,不让阳光照射到地球。

尽管潜力无限,但大规模人为干预极其复杂的气候系统,极有可能影响地区水分循环和粮食生产,乃至使全球大气环流发生变化,给不同地区、国家乃至未来世代造成不可预见的深远影响。此外,一些气候工程技术还可能造成"终止效应",意味着一旦开始实施就需要长期持续。如果由于技术或者政治等原因突然停止,没有其他减缓措施跟进,则全球气温会迅速反弹,使生态系统难以适应,不仅如此,一些环保组织还担心有了地球工程作为退路,会削弱人类社会减缓和适应气候变化的努力,产生所谓"道德风险"。国际社会已逐渐形成底线共识,地球工程不能替代从源头减缓气候变化的努力,而只能作为实现全球 1.5℃目标的补充,或气候紧急情况下的补救措施。



治理对象	治理阶段	治理目标
太阳辐射管理(SRM)	研究(模拟、观察等)	允许并鼓励公开透明的研究;避免 "道德风险";探究优化气候政策组 合的效果和风险
	户外实验	必須确定影响和风险可挖并被利益相关者所接受和允许才可进行;採 究何时,何地、如何应用或进行政策 组合能规避风险
	实施	在影响和风险得到科学论证并被全 球范围内的利益相关者接受之前, 禁止实施
碳移除 (CDR)	研究(技术研发 和小规模实验)	允许并鼓励公开透明的研究;避免 "道德风险";防范封存的环境风险; 探究何时、何地、如何应用或进行政 策组合能规避风险
	商业示范	允许;防范封存的环境风险和市场 风险
	大規模实施	论证全生命周期实现负排放的可行性,生物质生产和供应的可持续性, 融资问题,公众可接受度,以及防范 大规模封存的环境风险,等等

图: 地球工程的治理方式和相关行动方 来源: 地球工程的全球治理: 理论、框架与中国应对, 2020, 陈迎 & 沈维萍

地球工程给我们的明确启示是,未来三十年,我们既不能放弃减排的全球倡议与各国的能源体系转型,也不能面对无法兑现的控温场景坐以待毙,必须积极准备一些创新型的气候工程技术守护我们共同的未来。然而,地球工程影响的全球性、外部性决定了其治理需要全球共同努力,其综合影响的复杂广泛和不确定性决定了其治理是一个跨领域、多平台、多主体、多层次的治理体系,需要多方共建参与和协商推进。我们来到了一个必须团结携手应对挑战的时刻——在这样一个破碎和不信任蔓延的时代里,或许只有危机,能让人类重新拾起对一个共同的未来的希望和信念。

# 小结

展望三十年后的未来,在生命科学领域,我们可以预见更多疾病将被攻克、各国的人均寿命继续升高;多样的生物工程技术不断出现,技术赋能的人类越来越接近一个潜在的"造物主"。但是,我们是否能经受住基因增强的诱惑?如果不能,人类社会是否会因此更分裂、更动荡,对公平正义的追求是否会在基因技术创造的人与人之间的巨大先天差异下瓦解?我们能否在一个身体器官都可替换的年代里维持住"个体同一性"?如果不能,我们究竟该如何面对一场身份与自我的塌陷危机,人类社会会因此放下主体偏见而变得更包容吗?

在气候领域,三十年后,我们可以预见更多国家会加入和践行全球气候治理所规定的排放约束框架,越来越多的投资也将加速全球能源系统转型。Z世代的环保与气候意识将显著增强,这将使得个体的气候适应型生活方式变得更加普及,我们似乎也能看见形成共识的曙光。然而,大国冲突的加剧与不稳定的地缘政治格局还是会深刻笼罩在全球气候协作的努力之上,我们能在这一背景下定期定量完成紧迫的减排任务吗?如果不能,如地球工程这样的急救药的研发与风险控制能超越国家间的攸关利弊权衡,稳步推进吗?如果我们失去了最后一次兑现地球的可宜居的未来的机遇窗口,要如何面对一个气候灾难常态化的时代?我们能负担起对未来子孙后代的道德责任吗?

或许都是,或许都不是,但确定的事实是我们必须做好准备——完善技术治理框架,形成应对危机的道德与行动共识。

### 未来三十年的 社会变化

科技未来无疑是一场前所未有的文明奥德赛,如何让我们的文明不因为技术的失控和滥用而失去未来?

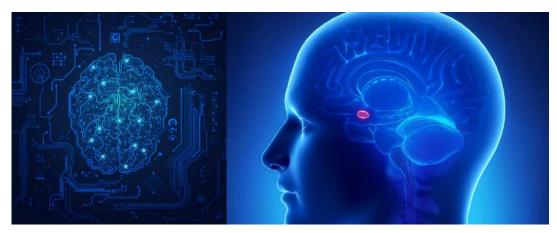
当然,更为深刻的问题是,深度科技化这个方向,是否本身就是让人类没有未来的去未来化之路?面对这样一个可能的未来社会,我们必须保持想象力,并提前做好准备。我们或许应该真诚发问:技术不断进步到底为的是什么,作为人类,我们到底在追求什么?

莫纳什大学(Monash University)荣休教授黄有光(Yew-Kwang Ng)说,未来三十年,我们的核心发展方向之一应该是最大化人们能获得的快乐。而英伟达中国 Omniverse 业务负责人何展则认为,虚拟与现实交融的数字技术将显著改善当前的人类生活质量。

### 刺激大脑与未来社会

在经济学家黄有光看来,快乐是个体人生乃至人类的终极目的,是唯一具有内在价值的东西。因 此,长期而言,包括所有的有苦乐感受者(人类及非人类动物)在内,如果净快乐的总量能提高, 应该就是好的,是符合道德的。其他的任何追求,都只有在对快乐有贡献、减少痛苦的情况才具 有工具性价值。出于这种认识与考虑,黄有光认为未来三十年完全可以、也应当预期的一种社会 变化是,刺激大脑享乐中心的技术变得像电视机一样普及。普通人都能轻易地得到巨量快乐,进 而促进许多社会问题(如犯罪)的解决。

何为大脑享乐中心?黄有光解释道,1954年,两名来自北美洲的科学家在对老鼠大脑的试验 中无意间发现,哺乳动物的大脑中有享乐中心(也叫线路),用微弱的电流(或磁场化学品 等)刺激这些地方,可以产生巨大的喜悦感(pleasure)。实验中的老鼠长时间自我刺激,每 小时自己按能够带来刺激的电流数达好几千次。学者称这种巨大的喜悦为"超级喜悦"(superpleasure),或"超越极大"(supramaximal),是远超一个生命正常生活中能够获得的最高喜悦 的程度。



来源: Johns Hopkins Medicine

对于人类而言,我们获得快乐的过程是刺激外围感觉器官,例如吃美食、看美景、听喜爱的音乐 等等,从而传导到我们的味觉、听觉和视觉神经,然后在大脑中生成快感。这种间接刺激的快感 比直接刺激我们大脑的快感小很多,主要是因为间接刺激力度比较小,且我们的生理机制保护我 们不会因为过度进食而出现严重的健康后果,即经济学上的所谓"边际效用递减"效应--人吃的 第二个包子肯定没有第一个香,吃到接近饱的时候,美食带给人的边际效用快速递减到零甚至为 负。也是因此,我们从感官处间接获得的快乐总量很有限。而直接刺激大脑的享乐中心,不但力 度更大,且没有边际效用递减。我们大脑中产生快感的"下丘脑"(hypothalamus)经受千次万 次的刺激,都会继续正常反应,不会罢工。因此,黄有光提到,直接刺激大脑的享乐总量巨大, 不但强度更高,时间也可以很长。如果考虑我们能通过基因工程进一步增强人类感受快乐的限度 能力,其潜在价值会更高。

现实地来看,这种刺激大脑的技术已经在医学有不少应用。举例来说,直接刺激大脑深部 (deep brain stimulation)的方法,在六十多年前发现后,早已用来治疗帕金森氏病、强迫症、肌张力障碍、 抑郁症、老年痴呆症等疾病,但没有发展出大家都能够用来刺激大脑、增加快乐的机器。黄教授 将其形容为"现代科技应该,但没有进行的一项能够有巨大贡献的工作"。究其原因,他指出, 一方面是人们对享乐有偏见。例如,认为增加快乐没有减少痛苦那么重要。但更重要的是,很多 人担心,这个机器发明之后会让人过分依赖。不过,黄有光认为,这类担心虽然合理,但人类不 应该因噎废食--几乎所有工具都可能被用来伤人,但我们不应该说,不要发明任何工具。我们 可能需要一些预防被过分滥用的措施,例如,能够用来刺激大脑的电流,只在晚上7-10点钟供 应等。甚至从积极的方面来说,人们如果能够通过刺激大脑而得到快乐,多数人就不会去犯罪、 去吸毒,因而可以解决许多社会问题。

展望未来三十年,黄有光提到,对于超过五亿已经脱贫进入小康生活的中国人而言,即使每个人 的消费每年增加十万元,其主观幸福感和快乐也未必能显著增加。但是如果我们能用 50 亿元(平 均每人投入 5 元) --不到中国一年 GDP 的一万五千分之一--来支持刺激大脑享乐中心的技术 研发,制造出一种人们能够安全地获得快乐的机器,那么社会的面貌将会发生非常大的改变--未来每个人将会更容易变得更幸福。到那时,黄有光预测,这种刺激大脑的机器的生产成本应该 不会比一部电视机贵。

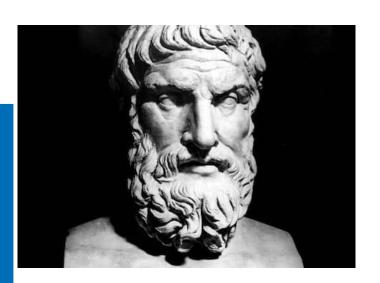


图: Epicurus (伊壁鸠鲁)来源: BBC

尽管会有人批评这个方法是不自然的。但黄有光提醒我们,绝大多数文明的产物、制度、医疗手 段等等,都可以说是不自然的。如果我们跟几百年前的人类说,现在人们最主要的娱乐是坐在沙 发上看一个箱子上的影子,或者在手上摆弄一个盒子里的光线,他们也会认为是很不自然的。真 正重要的问题是,我们能不能在进行任何技术研发的时候多退后一步问一句,这到底是为了什么? 在欢迎任何技术带来的社会变革的时候仔细想,人类究竟是因为这个技术变得更快乐了还是更劳 累、更辛苦了?

在刺激大脑的"脑洞"背后,一个有启发性的考量是:我们不能再秉持着"技术中立"和"技术无罪" 的价值观横冲直撞,而是应该做出价值判断后谨慎选择那些能给我们带来更多福祉的技术进步。 至少对于黄有光而言,如何让个体和整体的人生更快乐一些,应当成为未来技术革新所欲求的社

## 数字模拟与未来社会

自从 2021 年 Facebook 创始人扎克伯格宣布其公司更名为 Meta 之后,"元宇宙"(metaverse) 一词顿时在全世界热议。这一词汇最早出自于美国科幻小说《雪崩》,在小说中,人类通过"avatar" (数字替身),在一个虚拟三维空间中生活,这个人造的空间就被称为元宇宙。这一规模空前的 讨论趋势似乎映照着不少人的一种直觉判断,即人类已经开启了一场新的大转型:由现实走向虚 拟,由肉身走向数据上载,在人工智能技术、虚拟现实设备、模拟算力的加持下走向亦真亦幻的 后人类文明的未来。

仔细审视这种脱实向虚的未来倾向,背后似乎多有不得已而为之的窘迫。生态剧变、人口爆炸、 资源见底,自工业革命以来螺旋式上升的历史进程大有被颠覆之势,我们似乎走到了一个发展的 十字路口。另一方面,各行动方、阵营的全方位分裂亦在初露端倪,二战后形成的国际秩序体系、 多边主义的协作效力式微,严重贫富差别加剧经济、社会和思想阵营分野,二元对立似乎正在成 为人类社会应对任何分歧的唯一解。正如著名经济与政治史学家、在元宇宙议题上也多有论述的 朱嘉明教授所总结的那样,面对全方位危机的积聚,人类将在失序面前产生两种应激式的心态选 择:



来源: DevianArt

第一种是马斯克的星际移民模式,或者称之为"星际资本主义模式"。 这个模式的核心理念是基于地球和人类危机加剧的不可逆转,实现少 数人类的外星球移民, 进而保存文明的火种, 拓阔人类资源利用与生 存空间的可能性。其要旨在于向外走,通过使地球文明成为跨星球文 明来解决枯竭与发展的矛盾。

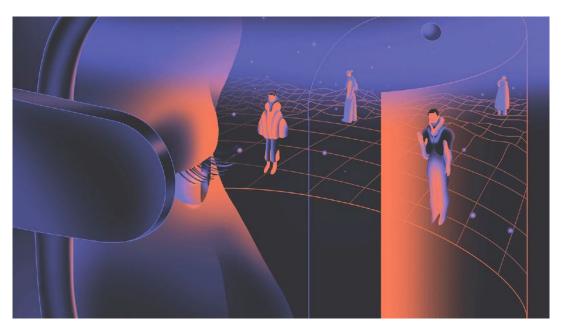


图: Omniverse 中的数字模拟人 来源: GTC Sept 2022 Keynote with NVIDIA CEO Jensen Huang, 2022, NVIDIA

第二种则是元宇宙代表的数字模拟模式,即通过非中心化自组织,分布式的合作,构建虚拟现实世界, 实现挽救人类危机和地球的各种实验,最终重构一个现实和虚拟互补的新世界。其要旨在于向内走, 利用技术条件,通过构建新的、某种非"实在"的生存场域来扩大文明的范畴和可能。

尽管非实在,但不意味着"模拟"就一定不真实。正如英伟达中国 Omniverse 业务负责人何展在会 议中所展示的那样,物理学的不断进步、计算机视觉的飞速发展,让今天的虚拟世界的画面早已能 以假乱真;人工智能、机器学习、深度学习的迭代让我们能够在数字世界中创造许多魔法,创建大 规模的逼真模拟环境,即所谓的"物理级精确的数字孪生",做到对地球生态环境的全方位模拟,在 符合多学科的逻辑原理的情况下,基于当前的工业和经济模式构造虚拟场景。更重要的是,在"计算 机视觉 +AI+ 沉浸式交互体验 + 物理精确的仿真 + 算力"的加持下,我们也能创造数字生命。

何展进一步介绍到,从技术阶段看 AI 数字人的发展,1.0 时代做到的是影视级别的超写实、AI 换脸 等一系列突破,这是许多数字平台正在开展的应用。当 1.0 时代进入成型期,逐渐迈向成熟期的时 候, 2.0 时代,AI 生成超写实和 AI 实时驱动 3D 超写实数字人。在此之后,随着更高的梦想和艺术 的需求迭代而来的就是 3.0 时代,数字人可以具有 AI 赋能的性格培养能力,建立在少代码甚至无代 码的数字构建平台,支持虚实结合交互。这在哲学上其实提出了一个非常紧迫的问题:如果我们逐 渐拥有了让数字模拟世界无比逼近现实世界的能力, 乃至在其中创造富有一定初级意识水平的生命, 那我们怎么知道人类眼中的"真实"世界不是另一个高级文明计算机中的模拟?



来源: NPR

《黑客帝国》和游戏《模拟人生》都为这样的假说提供了生动诠释。哲学家尼克·博斯特罗姆(Nick Bostrom)在2003年系统性地提出并部分论证了这种"模拟假说"(Simulation Hypothesis)。在《你 生活在一个电脑模拟之中吗?》(Are You Living in a Computer Simulation?)这篇论文中,博斯 特罗姆通过两个假设: 意识是一种算法; 人类的后代有足够的计算能力来模拟祖先(我们), 和 一个概率公式得出了三个推论:

- 1. 人类将在未来灭绝 (无法进入后人类阶段);
- 2. 我们的后代几乎没有兴趣模拟祖先;
- 3. 我们很可能生活在一个电脑模拟之中。

这三个推论只有一个是正确的(互斥)。假如我们对人类的未来感到乐观,同时承认人类很有兴 趣去模拟文明演化的话,那我们只能惊觉自己不过是运行在超级计算机中的程序,极有可能生活 在虚拟世界之中。当然,这一结论并不会对我们的日常选择造成太多影响,但它的确提供了一种 启示,即:世界本身即是亦真亦幻的,数字模拟可能并没有我们想象的那样"创新"或"颠覆"。

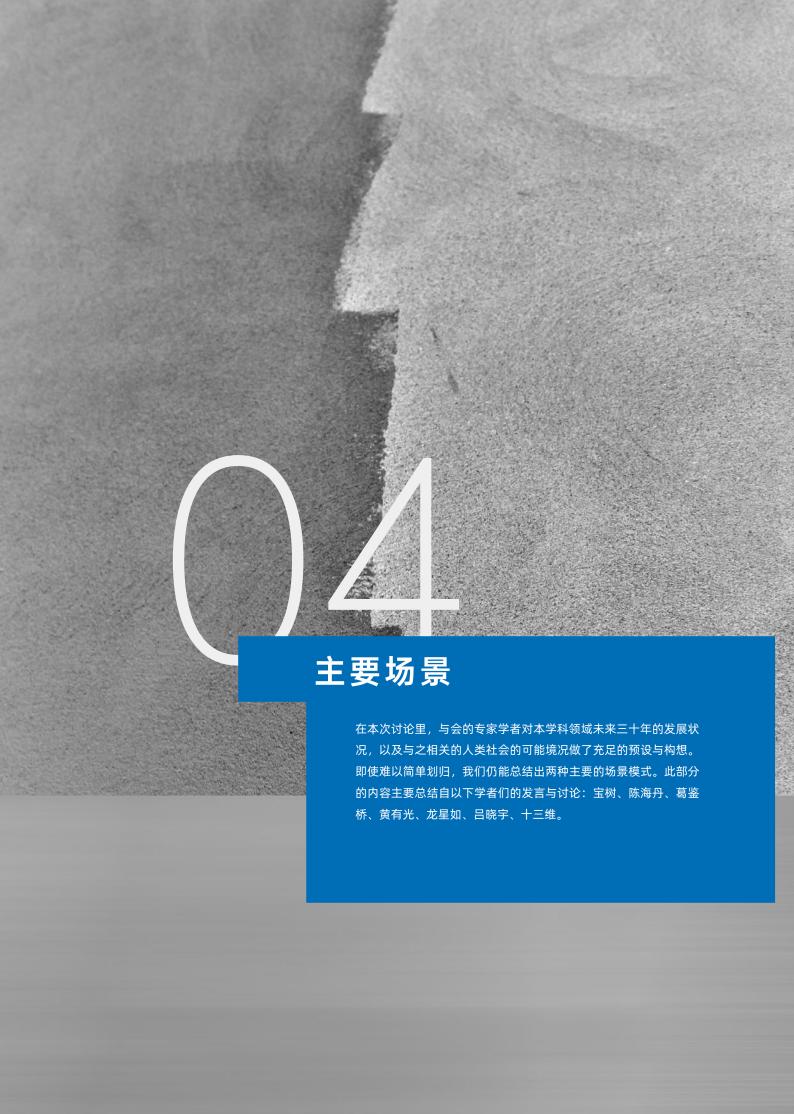
展望未来三十年,何展提到,我们还是要问:数字模拟到底需要给我们带来什么?中国人对"元" 有史记载最早的《周易上经》说:"元者,生物之始,天地之德,莫先于此,故于时为春,于人则为仁, 而众善之长也。"元宇宙之元,强调的是作为客观存在的天与人的主观意志、意愿的和谐融合所 创造出来的世界。在这层意义上,任何的数字模拟技术,始终还是需要回到服务真善美的价值观 上来,无论是通过再造世界还是虚实结合,我们需要的都是对一个价值共识的行动增进。

# 小结

在科技创新的推动下,从工业革命、绿色革命、计算机革命到数字革命智能革命,给人类社会带来了一系列颠覆性的变化,为那些可以接触新技术的人打开了全新的生活世界,为有能力支付创新产品和服务的人创造了更有效率的生活方式。但是,我们是否错过了向好的社会改变的技术机会——例如创造巨量的快乐——仅仅因为一些直觉偏见或某种文化塑造的路径依赖?追根溯源,我们需要技术去满足的目标、提升的"生活质量"究竟是什么,我们真的想得足够清楚和明确了吗?

另外,面对大时代的转型危机,世界已经不是牛顿式的简单机械状态,而是量子物理学状态。大转型过程中的坐标系不断改变 ,人们甚至不知道这个世界真正的坐标系在哪里。有些人寻求出走外星,但关键问题是:我们的危机真的是要通过不断毁灭上一个栖息地来解救吗,即使出走火星,我们就能走出文明的历史兴衰律?而转身向内——元宇宙是某种终极答案吗,虚拟和真实世界的边界到底在哪里,我们构想中的数字世界就是一个完美、自由、平等的乌托邦吗?

走向三十年后,我们必须正视而不是逃避技术与社会深度融合后的真问题。 我们需要探寻一种无可避免科技化却有人性、有道德、有公平的公共场域。 技术向善应与人类社会作为一个整体对共生共存的认识一起有机演进,创造 一个真正安全、繁荣、多面向可期的未来。



## 乐观场景: 技术赋能的良善社会

#### 工作生活:

如果技术不是不闻不问地一路高歌猛进,而是让人类能更多地获得良好生活,那三十年后我们应当期待最基本的工作、生活质量得到了增进和改善。工作方面,8 小时(甚至更短)工作制应当被严格执行,人们生活的闲暇与创造力将得到充分的释放,现代社会普遍存在的倦怠感会被更好的工作与生活平衡、以及支持个体开展丰富创造的技术创新所消除。而在生活方面,如黄有光所期待的能带来更多快乐与主观幸福感的技术将会更有意识地被发明和普及,我们将更容易地获得积极的情绪状态,在消极悲观时获得及时的心理支持,进而摆脱当代社会普遍存在的抑郁与无力。

#### 智能技术:

机器人技术的进一步成熟将会让许多具有人格的机器人,包括护理机器人、社交机器人、个人助理、性爱及保姆机器人进入我们的生活,许多在目前看来服务性的、重复而不具备意义实现功能的职业将被取代,人们将获得更多生活方向上的支持、许多人也因此可以从事更多创造性的工作。此外,超级计算机、量子计算等可能会有突破,虚拟现实技术的演进将带给我们更多样的娱乐、社交和工作体验,而数字孪生技术也将帮助我们更好地理解生命、地球和宇宙的演化及未来,以及预测各类不可预见的灾难性事件(如气候变化或行星撞击等)。

#### 生物科技:

生物技术进一步成熟发展,包括体外受精、人工子宫等将帮助女性从"不能承受的生育之重"中解放出来。基因编辑加持的人造食物将进一步帮助我们停止残酷的工厂化养殖,同时帮助实现我们减碳的气候目标。生物存储技术将使得信息存储的效率大大提高,而生物工程的发展也将使人们有能力消灭绝大多数遗传病,设计更好的靶向药物,减少病痛、延长人类的预期寿命。人造器官技术可能会得到有限的进步,部分器官衰竭将更容易被治愈,尽管三十年内这一技术未必会广泛普及。

在更广阔的视角来看,特别是从复杂科学的视角来看,人类文明的范式将在 治理机制:

未来三十年向比特文明、虚拟文明、后人类文明的范式转移,人类文明所依 赖的以发展为中心的工具理性也将向生态理性转变。区块链技术将带来生产 关系和私有制的变革,去中心化的治理机制将伴随着虚拟世界与数字货币的 演进而变得更具竞争力,我们应当期待一种更公平、正义、民主和良善的公

共化治理机制被初步建立起来。



来源: RealitySandwich

星际探索: 人类在未来三十年内向成为跨星际物种迈出实质性的一步。我们得以利用机

器人或生物工程的进步帮助我们完成深空探索,进而缓解地球资源紧张的问 题,一些初步的成果如火星基地、地月经济带可能会实现。进一步说,全球 化的趋势将被跨行星化所取代,初步验证可行的星际殖民将允许人类打开长

远生存的可能性, 更具文明生存与演化的韧性。

多物种 共生共存: 人工智能、神经与认知科学、脑科学将帮助人类对意识的理解和开发进入新 的范式,在未来三十年有效帮助拓宽人类对自我、对意识的理解,解除之前 的范式限制,让人类更好地感知彼此,更好地与自然、与非人类进行沟通和 交流。我们将提出更多、更好的关于智能、意识与自我认知的的理解与模型, 更好地将动物、植物乃至机器人、人工智能、外星生物纳入我们的道德关怀 范围。大规模的农场动物养殖将成为过去式,共生共存与跨物种平等成为共 识, 进而塑造出一种多物种和谐的公共生活。

# 悲观场景 灾难锁定的失序世界

我们原来想象的解放性的技术,正如其在过去两百年所做的那样,继续在未 工作生活:

> 来三十年给人类社会带来各种剥削和不平等。这种分裂与对立如果不经管控 和调适,可能会进一步向深入发展,造成更深度的人类内部不同意见社群与

生活方式、意识形态的隔绝与对峙。

智能技术: 机器人面临着被滥用的风险,人们的人格健全将被无孔不入的社交机器人所

> 侵蚀,而人与人之间真实的关系建立及其珍贵的价值创造也将被阻断。亲情、 友情、爱情这些在当代不言自明的情感链接,责任、付出和承诺这些无处不 在的道德取向,都将在智能技术的普及下被压抑。智能技术将进一步进入到 所谓的"背景"之中,比如算法的黑箱特征会进一步扩张,导致我们逐渐陷 入其带来的种种分化之中,包括社会分化、阶层分化、现状与未来的持续对

抗与失序。

生物科技: 缺乏监管和伦理道德约束的生物技术的发展会导致生物武器滥用,以及人工

> 加强的病毒开启其全球大流行,人类的生命健康面临巨大挑战。我们将不能 保证新冠疫情成为"全球最后一次大流行",而更致命的病毒将带来远超新 冠的死伤。此外,一些篡改人类意识及记忆的技术(比如读心术、脑机接口 等),将会使我们对真相与事实的认识进一步被阴影笼罩。而人类寿命延长

带来的人口及社会压力亦可能无法得到很好的解决。



来源: by Pawel Czerwinski on Unsplash

**星际探索:** 星际扩张在未来三十年有可能会被我们的燃料与推进技术、外太空生命维持技术的有限进展所拖慢。地球文明——至少在短近的未来,将长期处于卡尔达肖夫指数的 0 级状态,在我们真正掌握核技术,能利用核聚变转换出巨量的清洁能源之前,我们很可能还是会被锁定在地球这一母星上。而随着我们对意识、对智能或智慧的本质的理解的逐步突破,人类亦有可能会与地外高级文明产生接触,这可能会制造新的不可预测的选择压力来改变人类社会的行为模式。

治理机制: 人类社会的共识在诸多新技术的挑战下不断退后直至分裂,我们将在现有的思维范式下继续进行自我封闭,直至彻底从自然中隔绝出来。而人类社会在技术与政治的错误结合方式下进一步陷入分裂,最终可能会出现所谓表面的世界(数字集权上层)与地下世界(underworld,所谓江湖)之间的割裂。届时,世界的未来便不再是一个和谐、公正、人人各得其所的进步状态,而反倒可能退却至以伤害和强弱为支配原则的原始丛林社会,让协作和利他变得不再可能。

**生存灾难:** 由于气候变化、核战争、价值未"对齐"的超级智能或其他当前不可预测的原因, 人类在未来三十年灭绝,或者丧失了长远发展的潜力。技术、文化、制度和自 然资源全面枯竭。





来源: by Jacob Bentzinger on Unsplash

## 组织跨学科、跨地域的关于畅想未来的讨论与交流

预测未来所需要的智慧是复杂多元的,尤其需要跨专业(计算机科学、神经科学、认知科学、科技哲学等等)与跨地域(东方、西方、世界的其他区域)的启迪与智慧,在多角度、多面向、多范式的研究框架里寻求共识。与此同时,将各个学科的学者相聚一堂、互相启发,让技术的研发者与伦理规则的思考者走在科技的最前沿辩论、共学,能增加预测未来的准确性;并把各方的问题意识与核心忧虑逐一充分交流,深化对科技价值与意义的认识、追问科技与人类未来。

## 深度反思当今科学技术发展的底层范式

某种让人类能够更好对待他者、对待彼此的共识需要被创造、也需要被融入当今科学技术发展的底层价值。或许我们可以举行关于未来的全球大投票(通过社交媒体),调查人们对于未来的想法,以此来发掘和形成共识。我们需要认识到近500年来,每个世纪的20年代都是觉醒的年代和为整个世纪的思想奠基的年代。21世纪20年代将是整个文明思想发动机的开启时刻。步入深度科技化时代的人类正在攀爬巨大的技术悬梯,培育有人文关切的技术发展进路将是我们更好地走向一个可控、可感、可欲的未来的必要条件——而这些都需要建立在反思和寻找共识的基础之上。

#### 持续努力守护、并增进全人类的自由与民主价值

人本质上是目的性的动物。人作为时间上有限的存在者,有着对自身有限性的意识。在此基础上,我们才会对如何行动做出设计。因此,在"目的"这一概念中包含了其(基于其内在可能性的)现实性的根据,其形成和发展也都会有相应的内在必然性。在这一点上,非生物就不存在创造目的的问题。生物的主体性的起点,就在于在自然条件下能够生长发育,然后才有了主体性、目的性。而人类是唯一能够给自己规定目的的存在者,其行为同时具有自发性和目的性。从这个角度看,似乎自然的目的就是为了演化出能够为自己创造目的的存在物。而这样的存在物拥有的自由越多,其目的就越多。因此,现代人比原始人的目的更多、自由人比奴隶的目的更多。

人类社会作为一个整体,其重要本质就在于为尽可能多的目的创造表达的条件,并且支持其实现。我们需要对社会制度进行改革,在政治层面扩大民主,推行符合人性的法制体系、促进道德文化真善美的表达。一个走向未来的制度一定会使得人类的目的间彼此协同,形成正博弈,构成有秩序、有生机、多种目的有机统一的整体,让自由和全面发展不再只是一句口号。

### 让人文艺术充分发挥启迪与反思的作用

社会与公众可以通过文学与艺术作品更好地审视不同面向的技术观点,推进哲学与伦理学思考及社会公共价值的形成。也是因此,文艺工作者应当重视叙事的关键性,用故事与艺术表达帮助人类拓展更多对未来的认识与构想。可以说,文艺创作、特别是未来全人类的共同创作,有别于资源争夺,可以扭转工业革命以来科学凌驾于人文的掠夺性趋势。技术发展的最终目的是在保障人基本的衣食住行诉求后支持每个个体多样的创造性观念的表达,在这一点上,技术最多只能成为因,而人文艺术的构建是可能的果之一。



#### 预测性思维的风险?

尽管我们希望预测未来,但预测性思维及其模型最多找到一些相关性,离真正因果关系还很远。如何让我们的预测更精准地反映现实,或者,尽可能多地把相关变量纳入进去,将会是预测者们恒久的问题。与此同时,我们也必须保持谨慎,控制悲观预期,不要让"自我实现的预言"成为悲观的未来场景的先兆。

### 未来变量的开放性与不确定性?

未来是否可知是一个在物理学、经济学和哲学乃至数学上被争论了许久的话题。即使大的趋势可判断,但涉及到具体什么时候会发生、会发生具体什么事情,人类到目前为止还是黔驴技穷。这种开放和不确定,需要我们进一步对预测的方法论保持着一种谦卑、不断促进跨学科的理解,对随机性保持一种尊重、认识到世界某种混乱和不可预测的本质。更重要的,是对求真保持审慎,言说那些我们能知道的,对不可言说的保持沉默。

## 国际关系的治理协调如何更好地配合、促进或约束 技术的发展?

现有的非军事性技术突破(气候工程、基因编辑、疫苗研发使用等)已经在国际治理中初现弥彰。从影响力的互相传导而言,国际关系的治理协调如何更好地配合、促进或约束技术的发展?突破性的技术又将如何重塑国际关系格局?我们值得畅想更具韧性的技术全球治理框架,确保单边的伦理败坏、多边的利益诉求得到有效回应,在应对全球性的挑战时能反应更迅速和从容。

## 技术现实很可能比想象跑得更快,在此种意义上,我们该如何继续畅想,它的价值是什么?

从世界公认的第一部科幻小说《弗兰肯斯坦》开始,文学作为一种富有想象力和前瞻性的叙事已走过200多年的历史,许多彼时的'幻想'也早已成为此时的'现实':太空漫步、人工智能、虚拟现实、脑机接口……科幻以其预言式的构想描绘了科技的发展给人类的文明、社会、道德、思想带来的改造与挑战,而这同时也是哲学的重要命题:技术的日新月异如何改变了我们对自我、意识、自由意志的理解?在一个知识爆炸、宗教退场、未来已来的年代里,存在与真实是什么?人性与良好生活何以可能?在这种意义上,我们是否可以用文学作品作为预测未来的一种关键手段?

## 究竟什么是意识,什么是生命,什么是未来世界?

我们在思考未来时,究竟需要守护的是谁的价值?是我们这一种人类(自然演化的智人)的价值,非人类中心视角的、地球上全部可能生命的价值,还是"智能"的可持续(超级智能继承了人类智能、是人类的后裔)?至少截至目前,我们对何为意识(consciousness)、心灵(mind)、智能(intelligence)都尚不明晰,这就要求我们必须让我们的伦理道德共识的创造更具对象化的可能性。



来源: Arctic Focus

后续工作



## 未来学在中国的框架与议程设定

北京大学博古睿研究中心"科技与社会: 三十年以后的人类"研讨会是国内首个以哲学、科学、艺术跨学科对话为主要方式开展的关注未来研究(Future Studies)的研讨会,这不仅出自于我们对理解和预测人类未来大趋势的求知兴趣,更是因为我们坚信未来由今天的智识塑造,由今天各学科、各区域的启迪及其所内蕴的时代精神所决定。在此认识基础上,我们相信跨学科、多视野的未来学在中国有充分的创造空间,而本系列项目只是抛砖引玉,希望以此为起点,与同僚共同探索在中国畅想未来的理念框架与议程设定。

## 对未来的预测和大转型的思考的应用价值

面对人类的大转型,如果要做出更好的前瞻性研究,特别是与新兴科技的研发与治理相关的研究,其中重要的是提出前瞻性的科技政策,对新型科技问题有所预警、让科技在设定的边界下发展。这意味着,当我们从古今中外的哲学传统、从人文艺术的想象力中挖掘理解未来的线索的时候,同样需要考虑这种预测的落地价值——我们是否能打造一个更精准地预测未来大风险的模型?我们能否从纷繁复杂的关于未来的线索中找到一个可与各种行动方沟通的叙事?未来不仅仅是理念层面未到达的彼岸,它亦是我们行动与创造必然将抵达的目的地。



### 摸索和调试人类文明的底线共识性认识

在我们谈论未来的时候,创新、伦理、治理都是深刻又出挑的大词,在日常话语体系里并不罕见,但其在每个场合中的所指都难得一致,有时甚至可能相背而行。语用的不一致背后是共识的匮乏,尤其是在谈及人类社会的未来去向时,厘清关键变量,促进各个关键行动方共思科技创新、科技伦理、社会治理、国家互动的中长期目标,打造可对话、可成形的智识空间,共享一套话语与基础价值观,可能是我们妥善处理科技未来不确定性的第一步。这些变量间有的是一体一用,有的是互相配合与启发,有的是"干预"与后发性的"反馈",增进对互动模式的深刻理解同样是把握这些变量的必由之路。可以说,我们人类文明在当下急需找到互认的游戏规则,按照规则去发展,而不是陷入一种"比烂"的僵局。甚至可以说,唯有这样,人类才有未来。

## 寻求与全球关心人类未来的思想者的对话

跨学科、跨文明、跨地域的对话是深化关于人类未来研究的重要预设,也是实践一种"智识的全球治理"的先导方式。本次研讨会在中国的哲学与技术界率先开启未来展望,我们同样希望寻求与全球关心人类未来的思想者的对话,结合东西方与更多可能文明的视域,反思来路、构想去处。尤其考虑到,当下全球主流关于未来学的研究还深深植根于管理学、经济学和具有分析传统的英美哲学,在技术与艺术之外,如何纳入中国哲学、人类学、复杂科学、古生物学和各种多样的学科视野与古今中外的洞见,仍有待我们不断探索。

## 08. 参考资料

#### 长远未来

《危崖: 生存性风险与人类的未来》,托比·奥德,中信出版社,2021 《理与人》,德里克·帕菲特,上海译文出版社,2005 What We Owe the Future, William MacAskill, Hachette Book Group, Inc., 2022

#### 科幻玄思

50 Things Every AI Working with Humans Should Know, Ken Liu, Uncanny, 2020 《少女的名字是怪物》,宝树,花城出版社,2020 《生而服从:机器人故障指南》,刘宇昆等,新星出版社,2021

#### 黄铁军

- "人类能制造出"超级大脑"吗?",《中华读书报》,2015年1月13日
- "人工智能不会以人为中心但不要为此失落", 澎湃新闻, 2019年9月11日
- "人工智能 30 年展望", "30 年后的人类社会"第一次研讨会, 北京, 2021年7月26日

#### 复杂科学

《哥德尔、艾舍尔、巴赫:集异璧之大成》,侯世达,商务印书馆,1996 认知神经科学中的因果建模与推理,十三维,心识宇宙研究院,2022 对称性破缺与涌现——复杂科学与艺术之间的共鸣,十三维,集智俱乐部,2022

#### 生物工程

Global Catastrophic Biological Risks, Thomas V. Inglesby, Amesh A. Adalja, Springer, 2019 人类可遗传基因编辑 - 概念和科学问题,汪阳明,博古睿讲座,2021 伦理争论与科技治理——以英国胚胎和干细胞研究为例,陈海丹,自然辩证法通讯,2019

#### 气候工程

地球工程的全球治理: 理论、框架与中国应对,陈迎 & 沈维萍,中国人口·资源与环境,2020 《碳达峰、碳中和100 问》,陈迎 & 巢清尘,人民日报出版社,2021 The Importance of Global Extinction in Climate Change Policy, Yew-Kwang Ng, Global Policy, 2016

#### 刺激大脑

Happiness—Concept, Measurement and Promotion, Yew-Kwang Ng, Springer, 2022 Open access at: https://link.springer.com/book/10.1007/978-981-33-4972-8

Global Extinction and Animal Welfare: Two Priorities for Effective Altruism, Yew-Kwang Ng, Global Policy, 2019

Prof Yew-Kwang Ng on Ethics and How to Create a Much Happier World, Robert Wiblin & Keiran Harris, 80000 Hours, 2018

#### 算法模拟

端口:云下贵州,龙星如,贵州师大当代视觉艺术研究中心,2021 Omniverse,文化、艺术与科技的交融,何展,博古睿研究院,2022 Are You Living In A Computer Simulation? Nick Bostrom, Philosophical Quarterly, 2003 理解人类智能与人工智能的不同认知神经策略,葛鉴桥,博古睿讲座,2021

#### 治理机制

Norms, Storytelling and International Institutions in China: The Imperative to Narrate, Lu Xiaoyu, Palgrave MacMillan, 2021

《走出人类世:人与自然和谐共处的哲思》,宋冰等,中信出版集团,2021比较视角下的天下,安乐哲 & 杨月恒,博古睿研究院,2021



WeChat



Weibo



Bilibili



channels

Facebook:@BerggruenInstTwitter:@berggruenInstInstagram:berggrueninstLinkedIn:Berggruen InstituteYouTube:Berggruen Institute

Berggruen.org Berggruen.org.cn



## 北京大学博古睿 研究中心

Berggruen Research Center Peking University