

负责任创新何以可能？ ——以清华大学自动驾驶汽车实验室为例

许 畅，王 楠[✉]

(中国科学院大学 人文学院，北京 100049)

摘 要：“负责任创新”是欧美学者提出的一种新的发展理念，实验室作为研究和创新的重要场所，也是负责任创新的重要实践场所，实验室中的科研人员如何认知和履行负责任创新是非常重要的研究内容。以清华大学自动驾驶汽车实验室为例，通过田野调查的方式，结合负责任创新的基本理论，对该实验室的负责任创新环境、实验室研究人员的负责任创新认知和实践现状以及影响因素等内容进行了调查研究。

关键词：负责任创新；实验室；自动驾驶汽车

中图分类号：U471；B82-0 **文献标识码：**A **文章编号：**1674-4969(2021)01-0035-10

引言

“负责任创新”(又称负责任研究与创新，Responsible Research and Innovation，简称 RRI)是近年来由欧美学者提出的一种新的发展理念，旨在将负责任的理念与研究和创新相结合。根据欧盟“地平线 2020”计划，负责任创新指的是“在研发和创新的整个过程中，把社会各方的人员(研究人员、公民、决策者、商界、第三方组织等)结合在一起，协调各方使其符合欧洲社会的价值观、需求和期望。”^[1]因此，负责任创新的核心内涵是在研究和创新伊始让利益相关者参与进来，创新人员和社会行动者共同履行相关责任，通过创新中各个环节的交互，使得创新成果可持续发展，符合社会需求和道德价值，在伦理层面更为社会所接受。

实验室是研究和创新的重要场所，因此也是负责任创新的重要实践场所。实验室中的科研人员作为研究和创新的直接主体，比一般人掌握更

多的专业知识，具备更高的专业能力，这就使得他们比普通人能更准确、更全面地预测一些知识的未来应用前景，因为他们是科技成果和知识的缔造者^[2]。因此，科研人员对负责任创新的态度和看法，以及他们在实验室中如何做出“负责任创新”的研究和创新行为，是负责任创新研究中的重要内容之一。

自负责任创新概念提出以来，欧美学者尝试着从不同的研究领域展开相关研究，形成了多条发展进路，其中的“中游调节”进路就是对实验室和实验室中的科研人员的负责任创新研究^[3]。“中游调节”是美国学者埃里克·费舍尔(Erik Fisher)和卡尔·米切姆(Carl Mitcham)等人在“社会-技术整合研究”的实践经验中发展出来的，将实验室中的研究过程视为“科技创新治理”这一条河流的“中游”，其“上游”是政策导向和经费投入，“下游”是最终产品、用户和市场选择与监管。“中游调节”指的是在实验室的微观决策层面，

收稿日期：2020-04-15；修回日期：2020-09-14

作者简介：许 畅(1994-)，女，硕士研究生，研究方向为工程技术哲学。

[✉] 王 楠(1979-)，女，博士，副教授，研究方向为工程技术哲学。E-mail: wangnan@ucas.ac.cn(通讯作者)

通过让人文社会科学学者进入实验室、提升科研人员的伦理反思意识等方式来影响科技发展的路径, 以实现负责任的创新^[3]。此外, 荷兰、丹麦、比利时等国家 STS 领域的学者在政府机构的支持下, 也正在开展此类实验室合作的实践研究, 尤其是在纳米技术、合成生物学等有着广阔的应用前景、对社会环境影响巨大、风险尚未明确的研究领域^[4]。

近年来, 随着我国社会的快速转型发展, 负责任创新的发展理念开始进入我国部分学者和决策者的视野^[5], 涌现出一系列对欧美学者研究成果的梳理和介绍。然而, 国内现有研究主要集中在对负责任创新概念的内涵、思想来源、研究进路、欧美相关科技政策等^[3, 6-8], 尚缺乏对实验室和实验室中的科研人员的负责任创新研究。

当前, 自动驾驶作为人工智能最受关注的应用场景之一, 发展态势十分迅猛。特别值得注意的是, 自动驾驶的研究和创新涉及众多利益相关者, 面临着十分复杂的决策情景。因此, 本文选择了清华大学自动驾驶汽车实验室和实验室中的科研人员作为研究对象, 探索“负责任创新”在我国研发实验室中的践行情况。作者结合负责任创新的基本理论, 通过在该实验室近五个月的调研和访谈, 对该实验室的负责任创新环境、实验室研究人员的负责任创新认知和实践现状、影响因素等内容进行了调查研究。

1 清华大学自动驾驶汽车实验室概况

实验室作为科研活动的重要场所, 其人员构成和客观环境都会影响科研人员的研究行为。该实验室是国家重点实验室, 聚焦于自动驾驶汽车的感知、决策、控制等核心技术, 研究方向主要包括: 智能汽车与智能交通系统、人-车-路协同与车联网技术、智能车辆感知、决策与控制、驾驶员行为特性分析与建模等, 涉及自动驾驶汽车技术的多个重要环节。

1.1 人员构成

目前该实验室的研究人员有 42 人, 包括 3 名研究员、1 名副研究员、2 名助理研究员、2 名工程师、6 名博士后、20 名博士/硕士研究生和 8 名本科生。男性 35 人, 占比 83.33%; 女性 7 人, 占比 16.67%。硕士及以上学历为 80.95%。该实验室研究人员的主要研究内容是自动驾驶汽车的感知、决策和控制三个方向。其中自动驾驶汽车感知方向的研究人员有 15 人, 占比 35.71%; 自动驾驶汽车决策方向的研究人员 14 人, 占比 33.33%; 自动驾驶汽车控制方向的研究人员 13 人, 占比 30.95%。大部分科研人员的专业背景是汽车工程, 占比 91.5%, 少部分科研人员的专业背景是应用数学和计算机工程, 占比分别是 6.5% 和 2%。

总体来看, 整个实验室硕士以上学历的占比很高, 在研究内容上三个方向的研究人员分布相对均衡, 大部分科研人员的专业背景是汽车工程, 少部分人员的专业背景是应用数学和计算机相关专业。因此, 整个实验室科研人员的专业背景是一致的, 这使得他们对负责任创新的态度和看法显示出较多的相似性。

1.2 实验室环境

实验室不仅仅是做实验的物理场所, 拉图尔曾把实验室的意义扩展为“知识的生产场所”, 那里不仅包括各种复杂的实验仪器和实验材料, 而且把实验室人员以及实验活动的结果——科学陈述包含在内^[9]。实验室是研发活动的重要场所, 它的环境和布局有时候也会影响科研人员的研究兴趣, 科研人员在办公室讨论的内容并不限于纯粹的科学内容, 还包括个人兴趣爱好、学科的未来发展等。

该实验室是自动驾驶汽车研发的典型实验室, 由三部分构成: 办公室、地下实验室和室外试验场。

公共办公室是整个实验室的核心, 科研人员花费大量时间在办公室写作、讨论问题和日常交谈。每个学生有一个格子间, 主要办公设备是电脑, 学生用于看文献、写代码等。助理研究员在公共办公室内的封闭办公区处理日常事务和进行科研工作。研究员有单独的办公室, 每人一间, 在公共办公室的对面。公共办公室的开放状态有利于科研人员之间的交流, 大家有问题可以随时进行讨论。

地下实验室和室外试验场也是该实验室的重要组成部分, 地下实验室与办公室在同一栋楼, 位于整个楼的地下一层, 地下一层有很多工程实验室, 该实验室就是其中之一。地下实验室主要包括四部分, 分别是智能交通沙盘、车库、仿真实验平台和监控室。科研人员在这里进行自动驾驶汽车的部分功能模拟和操作监控。室外试验场虽然不属于实验室的管辖范围, 但由于它是整个实验中不可或缺的一个重要环节, 所以本研究认为室外测试场应当被视为实验室的一部分。通常情况下, 室外测试会选择在通讯信号好、行人相对较少的地方进行, 通讯信号好有利于地图采集、车辆接收信息, 人员稀少的地方可以最大程度保证车辆和行人的安全。

实验室不再被看作是一种被动的、与实验匹配而建造的制度环境, 而是决定知识生产的环节^[10]。因此, 实验室所具有的开放的交流氛围、先进完善的实验设备, 都为科研人员提供了良好的知识生产的环境。高素质的科研人员和有序的环境为负责任创新的实现提供了基础保障。

2 清华大学自动驾驶汽车实验室科研人员对负责任创新的认知

在科技创新系统中, 科研人员扮演着至关重要的角色, 他们的工作与知识和潜在技术的生产最为接近, 而这些知识和潜在的技术可能会成为创新整个过程的发动机^{[2]15-16}。科研人员作为科学研究的主要参与者, 能比普通人更准确、更全面

地预测技术未来的应用前景, 也有可能最早注意到技术潜在的风险。因此, 科研人员需要具备负责任创新意识, 从而能在科研工作中做出负责任创新的行为。本研究通过调查问卷和访谈的方式, 对实验室研究人员负责任创新的认知情况进行了调查、分析, 涵盖了自动驾驶汽车感知、决策、控制三个研究方向的科研人员。本次调查共发放问卷 42 份, 回收有效问卷 42 份, 有效回收率达 100%。

2.1 对负责任创新概念的认知

(1) 对负责任创新概念的了解。“您是否听说过‘负责任创新’概念。”调查结果: 共有 3 人回答“是”, 分别是 2 名研究员和 1 名助理研究员, 占比 7.14%。这说明该实验室科研人员大多不了解“负责任创新”概念, 但该实验室 4 名中高级职称的科研人员中有 3 人回答“是”, 这一结果可能与他们科研工作的国际化程度较高有关。

(2) 对负责任创新概念的认同。“您是否认同负责任创新的内涵(即科研人员与人文学者、社会学者、政府、公众等通过多方协商的方式确定研究目标和价值, 各参与方有共同协商和讨论的过程, 并依此塑造科技研究的路径和方向, 从而实现科技与社会和谐发展)。”调查结果: 共有 39 人回答“是”, 占比 92.85%, 可见实验室的科研人员基本上都认同负责任创新概念。3 名回答“否”的科研人员在访谈中表示, 他们之所以不认同此概念, 一是担心如果科研掺杂了太多其他因素会使科研变得不纯粹, 他们只想做纯技术开发, 不愿意考虑太多; 二是认为负责任创新的过程过于复杂, 利益相关方太多, 会影响科研成果的落地和研发速度。他们对负责任创新的担忧与其他技术研发实验室的科研人员的看法相似, 费舍尔带领的团队对北美、欧洲和东亚的 11 个国家的 30 个实验室的参与式研究结果显示, 部分科研人员对负责任创新中的公众参与也持怀疑态度^[4]。

2.2 对自动驾驶的社会、伦理问题的关注

(1) 是否应当关注自动驾驶汽车的社会、伦理问题。“您认为科研人员在研究中是否应该关注社会、伦理等问题?” 调查结果: 共有 27 人回答“是”, 占比 64.28%。其中, 感知方向的科研人员有 6 人, 决策方向的科研人员有 13 人, 控制方向的科研人员有 8 人, 在三个研究方向的人数上分别占比 14.29%、30.1%、19.04%。可以看出, 该实验室的大部分科研人员认为应该关注自动驾驶的社会、伦理等问题, 但是由于他们的研究方向存在差异, 对该问题的看法也存在差异。决策方向的研究人员绝大部分认为应该关注此类问题, 控制方向的次之, 感知方向的最少。

(2) 是否考虑过自动驾驶汽车的社会、伦理问题。“您在当前的科研活动中是否考虑了与之相关的社会、伦理等问题?” 调查结果: 共有 29 人回答“否”, 占比 67.44%。其中, 感知方向有 13 人, 决策方向 6 人, 控制方向 10 人, 在这三个研究方向的人数上分别占比 30.94%、14.28%、23.81%。与上一个问题相比, 这个问题的结果让人有点意外。回答“否”的科研人员在访谈中表示, 自动驾驶汽车在现阶段只涉及基础技术, 还未发展到考虑伦理问题的阶段, 当前的主要任务是如何尽快实现这项技术。部分决策方向的科研人员表示, 因为技术决策本身会涉及相关的问题, 所以考虑这类问题的时间会早于其他两个方向的研究人员。

通过以上两个问题可以看出, 该实验室的科研人员因为研究方向不同, 考虑问题的侧重点也不同, 决策方向的研究人员会更多地考虑相关的社会、伦理问题, 而感知和控制方向的研究人员更加注重技术上的实现。

2.3 对跨学科交流问题的看法

从伦理角度有效评估和影响技术创新各个环节, 以保证技术创新成果的可持续性和社会可接

受性是负责任创新的一项重要内容^[11]。欧美国家的 STS 学者曾在纳米技术、合成生物学领域, 进入实验室进行观察和干预, 从而影响了科研发展的方向、促进了相关立法、增加了科研人员思考问题的维度, 使得相关技术更好地落地。本研究按照这一思路对该实验室展开了相关调查。

(1) 对跨学科交流的认识。“您是否愿意让社会学者、人文学家参与到技术的研发中?” 调查结果: 共有 35 人回答“是”, 占比 83.34%; 有 5 人回答“无所谓”, 占比 11.9%; 2 人回答“否”, 占比 4.76%。因此, 该实验室的绝大部分科研人员愿意参与跨学科交流, 只是个别人员有顾虑。回答“否”的科研人员在访谈中表示, 人文学者、社会学家不懂相关的专业知识, 不会在研究中起到积极作用。

(2) 何时开展跨学科交流。“您对社会学者、人文学家参与到技术研发中的时机有何期待?” 调查结果: 感知和控制方向的 22 名科研人员都认为, 应该在技术发展到一定程度的时候再让社会学者、人文学家参与进来; 决策方向的研究人员中, 4 人认为在技术发展到一定程度的时候再让社会学者、人文学家参与进来, 其余 10 人认为, 应该在技术发展的早期让社会学者、人文学家参与进来。这个调查结果表明, 该实验室的科研人员因为研究内容不同, 对人文学者、社会学家参与自动驾驶汽车开发的时机的看法是不同的, 与 2.2 部分的调查结果一致。

上述问题的调查结果表明, 该实验室科研人员对“负责任研创新”概念的了解程度与其职称有关, 这可能与他们的科研工作国际化程度高低有关。尽管他们对这一概念的了解程度不高, 但绝大多数科研人员认同“负责任创新”的内涵。由于研究方向不同, 他们对自动驾驶的社会、伦理问题的关注、及对跨学科交流问题的看法, 出现了明显的差异, 偏重技术方向的研究人员对此类问题的态度比较保守, 远离技术方向的研究人

员对此类问题的态度比较开放。

3 清华大学自动驾驶汽车实验室科研人员负责任创新的实践情况

负责任创新不仅是一个理念,也是一个行为范式,从而可以在实践中解决技术与环境或者社会的关系问题,解决各利益相关者关系的冲突问题等。以欧文为代表的学者围绕该理念构建了负责任创新实践的理论框架,并将该理论框架分为四个维度:预测-反思-协商-反馈。在这个框架中,负责责任的科技创新实践通过多种形式呈现出来,包括技术预见与评估、公众参与和协商、价值敏感设计、跨学科合作等^[12]。该框架以阶段性参与的方式,得到了学术界的普遍认可。本调查按照这个理论框架的四个维度对该实验室负责任创新的实践情况进行了调查分析。

3.1 制度和规范

人类有预知风险的能力,规章制度和行为规范可以预测活动中可能存在的风险,约束人类的行为,从而通过强制的手段规避这些风险。实验室明确的规章制度以及科研人员默认的操作程序可以规范科研人员操作的准则,这直接影响到研究的流程和成果。

该实验室目前没有文本形式的实验室规章和操作指南,但是有一些不成文的规定和操作流程,比如“在实车测试阶段,学生和研究员不能驾驶测试车,测试车只有工程师可以驾驶。”这条规定明确了科研人员的工作范围和工程师的工作职责。因为即使科研人员有驾照、会开车,但是测试中的自动驾驶车辆不同于普通的汽车,很多感知和控制功能仍处于试验阶段,如果操作不当或者系统出现问题,将会面临巨大危险。而工程师熟悉车辆硬件系统,操作技术更加专业,可以很好地控制车辆,应对突发状况,在保护学生、研究员和工程师安全的同时,也最大程度地保证了周围环境(包括公众)的安全。

实验室科研人员除了要遵守实验室内部的规章制度,还要遵守实验室外部(如政府)的规范标准。例如,工信部、公安部和交通运输部于2018年4月联合发布的《智能网联汽车道路测试管理规范(试行)》,对自动驾驶汽车的测试流程进行了规定。该规范要求:自动驾驶汽车必须在建构有各种环境的封闭测试场内按照相关标准进行测试,达到一定的时间和里程经过系统性的测试后,才可以进入开放道路测试^[13]。该实验室在进行自动驾驶车辆测试时,严格遵守相关规定,首先在专门的封闭测试场进行车辆的性能测试,待通过系统性测试后,再在校内开放道路测试,然后才到更加开放的符合开放道路测试标准的社会道路进行性能测试。该规定最大程度地保证了测试人员、测试车辆和周围环境(包括公众)的安全。

自动驾驶汽车技术作为新兴技术,没有现成的研发规范可以依照,而技术的创新又不能任其发展,因此必须在研发过程中不断地进行自我反思、不断地完善,形成试行规范。该实验室的科研人员严格遵守各类操作流程和规章制度,尽量避免自动驾驶汽车在研发中可能发生的事故,最大程度地保证科研人员 and 公众的安全,这是实现负责任创新的具体表现之一。

3.2 研究生负责任创新意识的培养方式

研究生是科学研究与创新的生力军,培养研究生的负责任创新意识是实验室文化的重要体现,同时也是实验室负责任创新实践的表现之一。

(1) 课程设置。课程是培养研究生的重要方式之一。从2018级开始,该实验室的所有研究生都需要选修《科研伦理与学术表达》课程。虽然该课程是全校理工科院系的必修课,但是在授课方式和内容上,各院系根据本学院的研究内容,自行选派老师,单独授课。该实验室所在院系授课内容,主要侧重工程伦理的相关内容,而且十分注重工程伦理理论与具体案例的结合,特别是工程实践中的伦理。在具体的工程实践中,由于

工程活动的复杂性和不确定性, 工程师需要培养良好的工程决策能力, 即把伦理意识和伦理规范具体落实到解决工程实践面临的伦理问题上。比如以“电车难题”为代表的工程伦理困境, 是他们课程中重点讨论的问题之一。另外, 该课程还着重于培养工程师的工程伦理意识和责任感, 引导其掌握工程伦理的基本规范, 提高工程师的工程伦理决策能力, 为研究人员在实践中树立正确的工程伦理观打下基础。

作者分别对上过该课程和没上过该课程的研究生进行了访谈。上过该课程的研究生表示, 他们在科研中会更多地考虑工程伦理的问题, 也会更加关注与自动驾驶汽车有关的新闻时事、社会热点以及相关政策。受访者 A 说: “这个课程最大的好处是让我开始关注和思考这个问题, 因此我在科研过程中会有意识地做出自己认为具有伦理倾向的选择。”受访者 B 说: “是的, 我们每天面对的是一些具体的技术问题, 宏观性思维很重要, 它不只是解决技术问题还包含了一些社会问题。因为我们的技术最后是要落地的, 不可能不和社会接触, 这(自动驾驶汽车)以后就是生活的一部分, 所以社会的、伦理的问题, 与我们开发的技术息息相关, 遇到这些问题该怎么思考、怎么应对, 也是我们必须面对的问题。”因此, 这门课程给予了学生相对完整的工程伦理思维, 他们知道作为科研人员应具有的责任、可能会面对的问题、应该对什么问题进行思考和如何思考。没有上过该课程的研究生认为, 虽然因为各种原因他们可能会关注到与科研有关的伦理、社会问题, 但是他们不具备系统的工程伦理的思维, 对于相关问题的思考是零散、肤浅的。受访者 C 说: “因为我入学比较早, 那时候还没开这个课, 事实上遇到那些问题我也会去想, 但是不太知道从哪入手。”

(2) 导师引导。导师是研究生培养的第一责任人, 不仅培养学生的科研思维和科研能力, 也会促进学生形成工程伦理意识。作者通过对研究生的访谈了解到, 如果他们的导师关注相关的伦

理、社会问题或者对负责任创新有一定的了解, 那么学生就会更加关注这方面的问题。比如某导师参加了跨学科研讨会后, 会在开组会的时候提出这类问题, 引导学生进行思考。一些关注社会、伦理问题的导师还会让本文作者在组会上讲述人文学者对自动驾驶汽车的看法和争论, 并让他的学生进行讨论。不太关注社会、伦理问题的导师, 较多地关注自动驾驶汽车技术的前沿问题, 他的学生一般也不会关注这类问题。由此可见, 实验室研究生的负责任创新的意识形成与导师的引导有着很强的相关性。

3.3 实验过程中对负责任创新的考量

在负责任创新的框架中, 科研人员除了需要考虑研究诚信之外, 还需要更多地考虑其研究工作的广泛的社会影响^[14]¹⁴。对于实验室中的科研人员来说, 他们需要从实验和项目入手, 通过预测-反思-协商-反馈一系列环节, 对创新之后可能会出现的结果进行反思, 做出政策性的调整, 制定解决方案, 改进原有的设计^[15]。接下来, 本文以该实验室的“自动驾驶车辆在交叉路口决策”项目为例, 探究该实验室如何通过预测-反思-协商-反馈实践负责任创新。

(1) 预测。即科研人员在研发的过程中通过学习、沟通, 预测可能会发生的问题、采取的措施和解决方案。

在该项目的初期, 科研人员把决策算法的决策对象分为三类: 行人、非机动车、机动车, 仅根据三者的物理特性来进行区分。后来, 他们通过对三类对象的行为结果进行预测, 发现社会、环境和人等因素会使这些对象的行为结果存在很大差异, 他们决定将决策对象进行更详细的分类。新的分类方式不仅要考虑到区别行人、非机动车和机动车之间的行为差异, 还要区别人与人之间的行为差异, 很多交通事故正是因为当事人性格不同导致其行为方式的差异而引发的, 如果仅将决策对象按照物理状态分为三类, 忽视个体间的

差异, 显然背离了现实交通发生的场景。因此科研人员提出了“每个人生而不同”的理念, 即自动驾驶的决策对象不是电脑模拟的假人, 而是有着不同性格和特征的鲜活的个体, 考量决策是否合格的最终标准应当是面对真实场景时, 自动驾驶汽车对不同个性的人所能做出的正确决策。他们通过对停止-启动加速度曲线、特定姿态、路灯等候时间三个因素的考量, 最终将行人又进一步细分为保守的行人和激进的行人。再通过安全场、势能场、功能场和行为场理论, 对行人行进的风险做出评估, 进而对行人可能出现的行为做出预判。从而自动驾驶汽车会根据对行人的行动轨迹预测的结果, 采取不同的行驶方案, 实现“见风使舵”般的人性化智能决策。

在研发过程中这种有差别的区分可以使得决策结果更符合公众期望, 也更利于技术的顺利市场化, 实现一种前瞻性的预测式研究。

(2) 反思。即科研人员通过预测和实践发现了现有研究中存在的问题, 通过科学共同体内部的反思和沟通, 共同讨论出新的研究路径。

在该项目中, 为了实现“见风使舵”般的人性化智能决策, 要对当前自动驾驶汽车的决策方法进行讨论。当前自动驾驶汽车的决策主要是基于规则与局部路径规划的传统决策方法, 这只能满足简单场景下的驾驶功能, 无法应对更高等级的自动驾驶任务。如何基于客观交通风险与主观驾驶预期, 产生能适应时变的、复杂的交通环境的个性化驾驶决策, 以便更好地适应社会和道德需求是一大挑战。该项目组为了解决这个问题, 首先在科学共同体内部开展沟通交流, 对该问题进行研究、讨论。

他们的工作分为两个步骤: 第一步, 寻找解决问题的方向——召开组会。虽然这个问题的解决只涉及自动驾驶汽车决策方向的研发人员, 但感知和控制方向的人员也参加了组会, 因为他们认为自动驾驶汽车是一个整体, 各方向之间存在

联系, 很多方法也可以共用。在对解决方案的讨论中大家各抒己见。A 提出: 可以基于应用博弈论与机器学习等方法, 增强对交通环境的预测与交互行为的认知。B 认为: 该方法可行, 但需要有相关案例的证明。C 提出: 该方案是基于特定场景的开发, 可能不适用于复杂的道路。组会讨论的意见是, 会后继续查找国外应用博弈论与机器学习的案例。经过查找资料后, 项目组发现这些方法确实如 C 所言是基于特定场景开发, 并且发现其功能相互独立, 不能很好地适用于复杂的交通环境。因此, 项目组决定放弃该方法, 再次召开组会, 讨论新的方法和解决路径。第二次组会上, D 提出: 对人类驾驶学习过程的模拟, 应当采用端到端的方法, 并基于机器学习与大量模拟驾驶数据对车辆进行训练。组会后, 研究人员根据该方法做出训练模型, 进行小范围的组内讨论, 优化模型, 再进行电脑仿真实验, 根据实验结果, 再次优化模型。但是, 经过训练发现该方法的泛化能力较弱, 且难以结合交通规则和驾驶人个性需求。

第二步, 通过讨论确定解决问题的方向。经过几次组会讨论和会后研究, 项目组最终敲定了自动驾驶汽车拟人化决策技术, 这个思路是对复杂道路的交通决策中一种新的尝试。研究人员的思路是: 由于驾驶是人的一种行为, 如果单纯地把驾驶切割为机器的步骤, 很难全面掌握驾驶行为本身, 无法对复杂的交通环境做出个性化的驾驶决策, 而拟人化的决策可以把驾驶行为作为一个整体, 从而确定了解决方案的大方向。

(3) 协商。协商环节指的是科学共同体和共同体之外的其他利益相关者共同协商解决方案, 尽量让更多的利益相关者参与进来。

在这个项目中, 实验室科研人员先对方案进行试验, 再让实验室之外的人员参与试验, 这些人员除了作为试验对象提供试验数据外, 还对试验进行价值敏感监督, 了解实验数据可能导致的

后果, 并对试验本身提出自己的想法。

项目组的研究人员发现, 虽然驾驶员的驾驶决策行为具有自然属性, 驾驶员不同属性直接影响驾驶决策结果, 但驾驶员的“趋利避害”属性决定了他们在驾驶过程中都会追求安全和高效, 因此, 研究人员根据驾驶员追求安全与效率之间平衡的行为特性, 提出了驾驶员的决策机制遵循最小作用量原理, 并基于最小作用量原理提出了智能决策方法。根据这个方法, 项目组选择了 33 名驾驶人, 每人提供 32.5 小时的有效实验数据, 基于驾驶员驾驶决策过程的风险预判, 设计了复杂交通环境下各交通参与者的普适性车辆行驶安全辅助算法。参与数据收集的驾驶员(被测试者)在试验前就被明确告知, 哪些数据将被采集, 是否涉及敏感数据。他们还可以根据自己的驾驶经验对收集的数据提出自己的看法。在这个项目中, 参与数据收集的驾驶员根据自己的驾驶经验和习惯, 对三项数据的收集方法提出了改进的建议, 经过与研究人员的协商后, 项目组改变了对这三项数据的收集方法。

(4) 反馈。即根据各方利益相关者对协商后提出的解决方案的反应, 采取一系列的行为和措施, 使得方案和技术可以更好地落地。

通过预测-反思-协商环节, 项目组显著提升了驾驶安全辅助系统的拟人化水平和道路环境适应性, 具有较强的可行性, 可以驱动车辆控制智能化, 促进智能车辆的安全决策。该技术很快投入市场, 并获得了以下三项主要成果。1) 辅助驾驶系统。研发了基于安全场的拟人化驾驶安全辅助算法和系统, 实现了在商用车、军用车、乘用车上的大规模产业化应用。2) 车路协同系统。通过集成交叉口冲突解析、风险评估和协同控制技术, 研发了多目标交通信号与行驶车辆智能化协同控制系统, 实现了车路协同系统的示范应用, 通行效率提高 7%, 行车安全性提升 5%。3) 无人驾驶系统。提出了适应复杂环境的风险评估算法, 集成了智能车辆风险评估、路径规划、智能控制三

大技术, 研制出无人驾驶专用车辆, 实现了智能驾驶技术的产业化应用。

综上所述, 该实验室的“自动驾驶车辆在交叉路口决策”项目, 通过预测-反思-协商-反馈这些环节, 让各方利益者共同参与方案的研究和制定, 从而在创新过程中发现问题并加以解决、完善, 履行了“负责任创新”的理念, 上述一系列环节是“负责任创新”在该实验室得以实现的基本路径。

4 影响科研人员履行负责任创新的其他因素

除了制定规章制度、对研究生进行培养、导师引导等方式会影响科研人员履行负责任创新, 还有哪些因素会对科研人员履行负责任创新产生影响呢? 作者采用调查问卷的形式对实验室的科研人员进行了调查。

(1) 影响因素。“您认为制约科研人员履行创新责任的主要约束因素是(多选): A、对‘负责任创新’的认知; B、奖励和激励机制; C、技术发展水平; D、科学共同体内部的影响。”调查结果: 95.23%的科研人员选 A; 83.33%的人员选 B; 64.28%和 38.1%的人员分别选择 C 和 D。

(2) 促进机制。“您认为如何才能更好地促进科研人员履行负责任创新?(多选): A、政府引导; B、法律或伦理规范强制; C、科研人员自觉履行; D、社会监督; E、改善现有的激励机制”。调查结果显示: 97.62%的科研人员选 E; 85.71%的人员选 C; 71.43%的人员选 A; 71.43%的人员选 B; 35.71%的人员选 D。

通过对上述 2 个问题的结果分析后发现, 科研人员对于影响负责任创新的因素和机制的看法几乎相同, 没有表现出由于性别、职位、研究内容带来的差异化。可以看出, 科研人员普遍认为奖励与激励机制是影响他们履行负责任创新的最重要因素, 而技术发展水平也是科研人员履行负责任创新的重要影响因素之一。此外, 政府引导、

法律或伦理规范强制、科研人员自觉意识和社会监督机制这些都会影响科研人员履行负责任创新。因此, 尽管实验室在很多影响科研人员履行负责任创新的环节做出了努力, 如内部协调和奖励机制等, 政府和社会等外部因素也会在一定程度上影响实验室和实验室中的科研人员履行负责任创新。

5 结论

本文以清华大学自动驾驶汽车实验室和实验室中的科研人员作为研究对象, 通过田野调查的方式, 结合负责任创新的基本理论, 对该实验室的负责任创新环境、实验室研究人员的负责任创新认知和实践现状以及影响因素等内容进行了调查研究, 得到以下结论。第一, 该实验室所具有的开放的交流氛围、先进完善的实验设备, 都为科研人员提供了良好的知识生产的环境。高素质的科研人员和有序的科研为负责任创新的现实提供了基础保障。第二, 该实验室科研人员尽管对“负责任创新”概念的了解程度不高, 但绝大多数科研人员认同“负责任创新”的内涵。由于研究方向不同, 他们对自动驾驶的社会、伦理问题的关注以及对跨学科交流问题的看法, 出现了明显的差异。第三, 该实验室通过预测-反思-协商-反馈这些环节, 让各方利益者共同参与方案的研究和制定, 从而在创新过程中发现问题并加以解决、完善, 履行了“负责任创新”的理念。第四, 该实验室的科研人员普遍认为奖励与激励机制是影响他们履行负责任创新的最重要因素, 而技术发展水平也是科研人员履行负责任创新的重要影响因素之一。此外, 政府引导、法律或伦理规范强制、科研人员自觉意识和社会监督机制这些都会影响科研人员履行负责任创新。

近年来中国社会快速转型发展, 社会公众、科技共同体、企业和政府针对创新社会责任的态度和行为方面出现了一系列新变化, 例如我国治理实践中提炼出来的“创新、协调、绿色、开放、

共享”五大发展理念以及构建人类命运共同体的思想等重要理论成果, 与负责任创新不无重叠共通之处, 这为负责任创新在中国落地成长创造了有利条件^[5, 6]。然而, 正如负责任创新理念的著名倡导者雷内·冯·尚伯格(René von Schomberg)所提到的^[16], 负责任创新在不同国家的实施会因为国情不同而遇到不同的问题, 因此, 负责任创新这一理念在中国的推行还面临着许多问题, 特别是实验室作为负责任创新的重要实践场所, 如何促进实验室中的科研人员更好地履行负责任创新任重而道远。

参考文献

- [1] 中国-欧盟科技合作促进办公室. 中国科学技术交流中心. 欧盟“地平线”2020计划[EB/OL]. <http://www.cstec.org.cn/infoDetail.html?id=4980&column=1034>.
- [2] 伏佳伟. 科技工作者的伦理责任[D]. 渤海大学, 2014.
- [3] 廖苗. 负责任(研究)与创新的概念辨析和学理脉络[J]. 自然辩证法通讯, 2019, 41(11): 77-86.
- [4] 廖苗. “社会技术整合研究”述评[J]. 自然辩证法通讯, 2017, 39(5): 25-32.
- [5] 赵延东, 廖苗. 负责任研究与创新在中国[J]. 中国软科学, 2017(3): 37-46.
- [6] 梅亮, 陈劲, 李福嘉. 负责任创新内涵辨析与启示[J]. 自然辩证法研究, 2017, 33(2): 49-53.
- [7] 廖苗. 欧洲政策思想史中的“负责任研究与创新”[J]. 科学学研究, 2019, 37(7): 1211-1219, 1276.
- [8] 薛桂波, 赵一秀. 基于“负责任创新”的欧盟科技政策转型及启示[J]. 中国科技论坛, 2017(7): 172-177.
- [9] 布鲁诺·拉图尔, 史蒂夫·伍尔伽. 实验室生活: 科学事实的建构过程[M]. 张伯霖, 刁小英译. 北京: 东方出版社, 2004.
- [10] 阎莉, 康中和. 从实验室研究透视SSK与科学哲学的差异[J]. 科学技术哲学研究, 2013, 30(6): 105-108.
- [11] Traweek S. Beamtimes and Lifetimes[M]. Cambridge: Harvard University Press, 1988: 157.
- [12] Owen R, Stilgoe J, Macnaghten P, et al. A Framework for Responsible Innovation[M]. Responsible Innovation. John Wiley & Sons Ltd, 2013: 27-50.
- [13] 中华人民共和国工业和信息化部, 中华人民共和国公安部, 中华人民共和国交通运输部. 智能网联汽车道路测试管理规范(试行)[Z]. 2018-4-3.
- [14] 廖苗. 负责任创新理念下科研人员的伦理参与责任[J].

洛阳师范学院学报, 2018, 37(6): 13-19.

[15] 陈佳丽. 伦理学视阈下的“负责任创新”研究[D]. 南京林业大学, 2016.

[16] 廖苗, 等. 从“负责任创新”到“开放科学”——雷内·冯·尚伯格访谈录[J]. 长沙理工大学学报(社会科学版), 2020, 35(3): 14-27.

How is Responsible Innovation Possible? Take the Automatic Driving Automobile Laboratory of Tsinghua University as an Example

Xu Chang, Wang Nan✉

(School of Humanities, Chinese Academy of Sciences University, Beijing 100049, China)

Abstract: “Responsible Research and Innovation (RRI)” is a new development concept which is put forward by European and American scholars. As an important place for RRI, the laboratory is also an important practical place for RRI. Therefore, how to recognize and perform RRI is a very important research content. Taking the Autonomous Vehicle Laboratory of Tsinghua University as an example, and based on the RRI theories, this article makes an investigation on the RRI environment of the laboratory, the laboratory researchers’ awareness, the practice status and the influencing factors of RRI.

Key Words: responsible innovation; laboratory; automated vehicles