AIGC检测新规后的第一个毕业季

实习生 何新月
中青报·中青网记者 李 强

4月初，中国人民大学本科生王菁把毕业论文提交知网检测“AIGC率（生成式人工智能参与度）”后，她慌了。

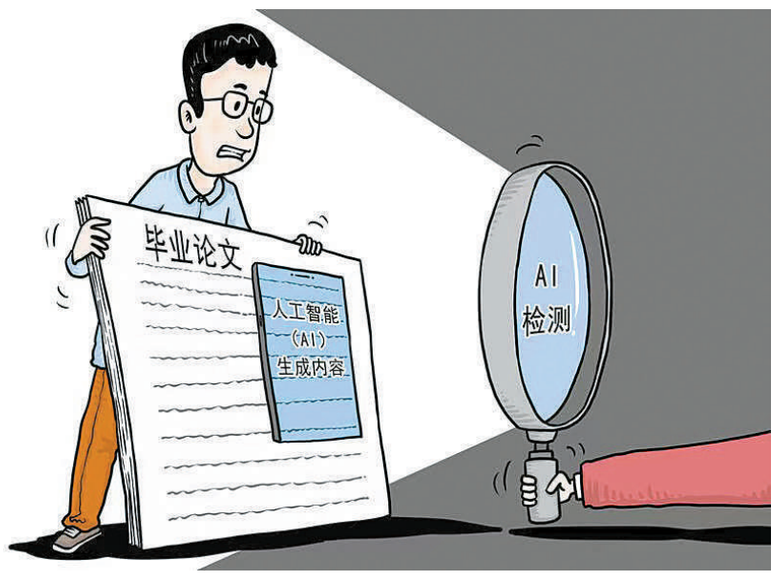
检测结果显示1.3%。“这也太低了！”王菁担心因AI检测率过低，论文会被抽去“盲审”。实际上，她撰写毕业论文的全程都用到了AI工具，包括内容生成、润色。

不久前，中国人民大学教务处发布了本科毕业论文（设计）中AIGC的使用规范，除要求检测外，还规定学生不得用AI直接生成论文。

面临毕业，论文的检测结果让王菁感到困惑。她告诉记者，她曾经让不同的系统检测同一篇课程论文，结果有的显示AIGC率70%，有的7%。

这只是全球高校对学生使用AI进行规范的一个缩影。过去一年多，复旦大学、四川大学、中国传媒大学等多所国内高校，以及部分国外高校陆续出台相关规范。中青报·中青网记者发现，各高校毕业论文AI规范存在差异，具体执行中的尺度也松紧不一。

6月20日，浙江杭州，一场线下招聘会。 视觉中国供图



视觉中国供图

王菁毕业论文在某AIGC检测平台的检测结果截图。 受访者供图

6月20日，北京海淀，3名毕业生身穿学位服走在校园里。 中新社记者 蒋启明/摄

许宸宇毕业论文的AI检测结果截图。 受访者供图

数百万分之一

通知发布时，王菁的毕业论文初稿在AI的辅助下已快写完了。

她研究的是某行业的特殊劳动现象。去年论文开题时，DeepSeek尚未问世，她也不太信任市面上的其他AI工具，开题报告、文献综述都是“纯手工做的”。

今年年初，王菁访谈了十余人，开始撰写论文。面对海量素材，她不知道如何下手。恰逢春节期间DeepSeek火了，国内多所高校宣布接入DeepSeek大模型，供师生在教学、科研等场景下使用。

一天晚上，为搭不出论文框架焦虑的王菁，将访谈记录、文献综述等文本都“喂”给了DeepSeek和腾讯元宝，结果让她很惊喜。“它讲得很完整，把我觉得很碎的点串起来了，写的内容有递进，框架列得也很不错。”

在4份AI生成的论文框架的基础上，她整合出一份符合自己思路的框架。AI起的小标题，也让她觉得“挺有味道的”。后来，AI还参与了这篇论文的语言润色，效率颇高。

王菁花了一天半的时间，用“大白话”粗糙地写出论点、论据，又用一周时间，将这些“大白话”一段段“投喂”给AI工具，请它一点点润色。

“一开始还挺能提高效率的，看多了之后，发现它经常原模原样抄其他论文的专业词汇。”于是，王菁与AI开启了“拉锯战”。AI生成一部分，她就提出疑问和要求，让AI继续修改。如此反复，最后在AI的协助下，王菁完成了一万多字的论文初稿。

这份毕业论文，只是今年全国毕业论文的数百万分之一。其中有多少篇、多大程度上使用了AI工具，不得而知。去年，一家第三方机构的调研结果显示，近六成高校师生每天或者每周多次使用生成式AI，近三成大学生主要用AI写论文或作业。

复旦大学一名文科生陆冠宇在媒体实习时，会使用AI工具挖掘选题线索、整理采访提纲，但在写毕业论文时，她很谨慎，从搭建框架到论文撰写，都没用AI工具。她担心，自己的思维被AI固化。

复旦大学材料物理专业的陈振华，在撰写毕业论文时，仅在文献综述撰写、论文框架搭建两个环节使用了AI工具，论文主体内容他没让AI参与。

去年11月28日，为“确保学术诚信，提高教学质量，防范学术不端”，复旦大学制定并发布了《复旦大学关于本科毕业论文（设计）中使用AI工具的规定（试行）》。该试行规定提出“六个禁止”，禁用范围包括研究设计与数据分析、原始数据收集、论文撰写、润色等。

这是国内高校首份专门针对AI工具在毕业论文（设计）中使用的规范化管理文件，发布后引起广泛关注，其中“禁止使用AI工具进行语言润色和翻译”等要求曾引发热议。

有人担忧，AI工具已经是不可阻挡的趋势，这一禁令是否矫枉过正，剥夺学生提升学术表达的机会。复旦大学计算与智能创新学院教授、上海市数据科学重点实验室主任肖仰华对中青报·中青网记者说，对已经熟练掌握语言表达能力的学生而言，AI能帮助他们从无意义的细节中解放出来，聚焦更具创造性的内容。

但该规范中提到：“语言表达能力仍然是毕业论文（设计）考察的重点。”

“本意绝不是开历史倒车。”复旦大学教务处相关负责人在接受媒体采访时表示，“我们希望学生将AI作为一种工具，助力、赋能而不是取代人，因此试图去明确边界。这个规定不可能一步到位，会随着时代变迁不断快速迭代。”

福建师范大学的许宸宇说，毕业论文涉及的实验数据需要使用专业软件，AI并未参与论文撰写的核心环节。“我的论文主要是要结合数据对现象进行解释，AI没办法去解读、分析，它对我的实验肯定没我了解，都不一定能精准表达数据的含义，所以我基本靠自己写。”许宸宇说，在论文的语言润色方面他也未使用AI。

今年毕业的计算机专业博士生贺文静，在撰写论文的全程都很少使用AI工具。尽管她所研究的是“深度学习模型测试”，但她对待AI的态度也很谨慎，常常在实验中发现，AI模型会输出一些错误内容，论文写作上她并不信任AI。

“就像研究自动驾驶的人，更倾向于选择自己开车。”贺文静对记者说，她也不敢将未公开发表过的论文内容输入到AI，担心研究方法被泄露。

但贺文静也表示，在向国际会议投稿英文论文时，她会借助AI工具来润色和翻译论文内容，“英语不是母语，我的表达是不够好的，不如AI写得流畅”。但与她合作的一位60多岁的英国教授对此并不认可，坚持让她自己打磨语言。

AI规范的落实依赖于学生的诚信

去年6月，华东师范大学传播学院与北京师范大学新闻传播学院联合发布了国内首个《生成式人工智能学生使用指南》。

相比于复旦大学的“AI规范”，记者注意到，这份指南并未禁止AI生成内容出现在学术写作中，也允许使用AI润色，但规定，必须标明AI生成的内容和说明AI的使用方式，且AI生成的内容不能超过全文的20%。

“AI生成内容不能超过20%，其实是参照本科生毕业论文查重的范围确定的。”制定该指南的主要负责人、华东师范大学传播学院院长王峰告诉记者，目前没有权威手段进行AI率检测，这份“指南”主要起到一个倡导作用，其实施依赖于学生的诚信。

王峰说，这份指南经过了二三十人、四五轮的讨论，才完成制定并公开。起初，也有教师不赞成学生使用AI。

“禁止学生使用AI是悖论。”王峰说，社会已广泛应用AI，大学教育必须与之接轨。王峰主张，引导学生合理使用AI工具，强调教育的核心是培养学生的原创力，而非重复性技能，与此同时，教师要革新考核方式，设计激发学生探索的作业，而非传统的论文。

今年以来，有不少高校发布了2025届本科毕业论文（设计）AIGC检测工作的通知，要求学生在答辩前完成。

4月9日，四川大学教务处发布通知称：文科类AI生成内容占比不超过20%，理工医科类不超过15%。“对于检测AIGC比例偏高的毕业论文（设计），通知相关导师和学生进行处理。”

一些学校还在通知中指出，检测后确认构成学术不端的，视情况给予延期答辩、重做等处理，情节严重者取消毕业资格，毕业论文（设计）成绩计不及格。

很快，一些学生发现，AIGC的检测系统未必真实可信。有时，同一篇论文前后两次检测结果数值差异很大；有时，对论文进行人工修改后，AI率不降反升；有时，一些远早于AI时代的文学名篇也会被检测出过高的AI率。

记者发现，今年毕业季，一些高校在毕业论文的查重等检测中，暂未严格执行相关通知或规范性文件。

中国人民大学的相关通知曾要求：学生不得使用AI直接生成毕业论文内容，需主动报告AI的使用情况，保留和提供与AI的完整对话。文末还提到：学校将引入检测工具对论文中的AI生成文本进行检测，检测结果将作为成绩评定和优秀本科毕业论文评选的重要参考依据。

王菁告诉记者，在她今年提交毕业论文时，学院未要求她提交AI使用记录，也未对“优秀论文”设置具体的AIGC率门槛。“唯一不一样的是，今年系统会自动显示查重数值和AIGC率。”她表示，“周围没有同学因为AIGC检测没通过论文审核的。”

陈振华也注意到，学校虽明确禁止AI语言润色、翻译等行为，但在论文查重环节并未增加AIGC检测，也未提供专门的AIGC检测软件。

根据学校规定，他须在承诺书中披露AI工具名称、版本、使用时间、用途、生成内容及使用部分，并保留处理前重要材料以备检查质询。但陈振华告诉记者，自己论文仅附一句话说明“使用AI辅助阅读和整理文献”，除此之外未做更多AI使用披露。

“提交论文后学校统一查重，没有显示关于AI占比的内容。”5月30日，陈振华完成本科毕业论文答辩，论文顺利通过查重与答辩。另一位复旦大学的学生则称，各个学院落实规定的细节有所不同，“我们学院没有查AI率，但听说有的学院查了”。

一名中国传媒大学的学生则填写了一份“毕业论文AI使用情况说明表”，在表中写下自己使用AI的细节，并且自行填写AIGC检测数据。而福建师范大学的许宸宇则被学院要求，答辩前查一次，最终送审前查一次，AIGC检测率要低于20%，否则无法参加答辩，不予送审。

肖仰华也担心，滥用AI导致学生丧失独立思考的能力。他觉得，是否使用AI应基于学生自身水平，已掌握的能力可借助AI优化，未掌握的需谨慎避免依赖。

“有争议才好，”肖仰华说：“就怕是死水一潭，不利于将来形成共识。未来一定会有越来越多的、更深的、更科学的一些认识和判断，这需要时间，需要一个过程。”他也提到，国际范围内早就有讨论，政策的制定也更细化。

据外媒报道，英国高等教育政策研究所（HEPI）对英国本科生的一项调查显示，今年有92%的本科生以某种形式使用生成式AI，去年这一比例为66%；88%的本科生在各类学业评价环节（论文写作、课程作业、项目报告）中使用过AI，高于去年的53%。

为了对抗这种趋势，伦敦大学法学院已恢复更为传统的线下考试。

在美国斯坦福大学，课程教师可以自行制定AI使用政策，允许或禁止AI工具的部分或全部功能。

悉尼大学采用“双轨制”，在考试中禁用AI，但写论文、做报告时，可以使用AI辅助，但需标注说明。政策制定的参与者表示，这项政策尝试将AI纳入教育体系，使学生在学习中真正将AI视为辅助工具，而非作弊手段。

香港大学的AI政策则经历了从禁止到逐步开放的变化。2023年2月，香港大学率先出台禁令，禁止学生在课堂、作业和评估中使用ChatGPT等AI工具。当时，香港大学副校长何立仁在内部邮件中指出，面对ChatGPT对教育的颠覆性冲击，学校需要时间琢磨长期对策。

2023年9月，香港大学开始在教学与学习各方面“拥抱AI”。校方免费给师生提供ChatGPT、DALL-E等AI工具，并开设在线课程和培训，确保这些AI工具得到有效使用。

“高等教育正面临山呼海啸般的改变。”王峰告诉记者，AI规范制定的背后，其实是在思考，AI时代，我们要培养什么样的学生，更深层次的问题是教育评价体系如何改革。

“AI在教育中的应用不是非黑即白，而是如何用、何时用的问题。”肖仰华说，国内和国际的讨论，都只是教育变革的序曲，AI时代教育的根本性变革即将发生。

在争论之下，2025年本科生毕业论文的答辩和审查已结束。曾担心AIGC检测的王菁，在5月中旬通过答辩，还拿到了高分。

（应受访者要求，王菁、陈振华、贺文静为化名）

我在AI论坛打瞌睡会错过未来吗

□ 李 强文并摄

65岁的学者汪晖身着西装上台了。他头发花白，手执讲稿，身后的大屏幕上写着这场论坛的主题：“AI时代的伦理奇点”。这是当下的热门话题，但台下有人在打哈欠，显然清晨的困意未消。

我是一大早，跨越大半个北京城，挤了一个多小时的地铁来清华大学听这场“方塘论坛”的，兴致很高。一半是因为这跟AI有关，一半是因为嘉宾名单里不只有科学家，还有70岁的赵鼎新、65岁的汪晖、64岁的赵汀阳等。

他们是人文社科领域的“思想老者”，要跟科学家一块儿探讨这个时代最新的烫手命题，这本身就值得期待。针对“时代”，他们论了几十年，现在仍未停止。

面前的阶梯教室没坐满，但也汇聚了上百人：学生、教授、学者、记者、作家，还有从美、英、德、意等国请来的嘉宾。空气里似乎弥漫着一种学术会议特有的、略带沉闷的气氛，我揉了揉酸涩的眼睛，吃了口离家时塞进背包的面包，想提振精神，让沉重的脑袋清醒一些。

“几乎所有领域，都在经历系统性的重构。”汪晖主持开场，话语沉稳有力，“回应AI时代的伦理挑战，不仅需要对技术本身的理解，更需要回到哲学、社会学、政治经济学、文学、艺术等基础人文领域，共同探寻人类何以‘安身立命’”。

论坛的择期颇显用心，6月23日这天是计算机科学之父、人工智能之父——艾伦·图灵诞辰113周年；图灵奖的中国得主、中国科学院院士、清华大学人工智能学院院长姚期智也受邀到场。

人工智能被提出75年了，而AI大规模走入公众视野不过数年。最新数据显示，中国生成式AI用户数量已超3亿。用户数量从0到3亿，中国互联网花了15年，而AI只用了两年。OpenAI的用户数量更是在今年4月初就达到了5亿。

历史上从未有一个时代如AI时代这般，将未来如此迅猛地推向当下。时间仿佛被丢入加速器，而AI如氧气一般，弥漫进社会和人的毛细血管里。智能驾驶、机器人、AI医生、AI伴侣等，过去数十年来，昔日科幻小说里的事物正步入现实。

当越来越多的人享受着AI带来的技术红利、探索新的可能性时，挑战也如影随形：失业、偏见、伦理、法律，乃至关乎人类存续的风险。这些问题，让我觉得既遥远又真实。

“通用人工智能的能力快速增长，人类是否还有能力管控它？”姚期智是这天的第一个主题发言者。年近八旬的他，声音不高，但每一个问题，都能戳我一下。

他认为，AI时代的问题，许多是被放大的旧疾，尤其是生存性风险，武器的诞生曾让人类很焦虑，AI也有这样的问题，“它变得更危险、更难对付”。

姚期智提到了2024年诺贝尔化学奖得主——因利用AlphaFold（一种AI系统）解决蛋白质结构预测难题而获奖的戴米斯·哈萨比斯（Demis Hassabis）和约翰·江珀（John M.Jumper）。“这是多年来不可能的事情，也给了我们一些终极希望。”姚期智随即警示，AI在生物医学领域的突破，也可能被用来创造“镜像生命（Mirror life，一种存在于假说中的人工合成生物）”，带来致命的灾难。

2024年12月，《科学》杂志就曾发文呼吁全球暂停“镜像生命”的研究，优先考虑并防范风险。

另一种生存性风险在于大模型失控。“我们现在制造出来的AI系统，它变得越来越像人。”姚期智指出，已有模型展示出“骗人”本领和对人性的洞察——问题的严峻性在于，当它真正像人一样，甚至超过人时，如何控制。

这也是许多科学家反复警示的人工智能的“控制难题”，现在这个难题比以往任何时候都更迫切地需要答案。我这个人工智能技术领域的小白是没有答案的，那一刻，我脑子里只闪过许多科幻电影里的失控场景。

6月23日，清华大学，清华大学文科资深教授汪晖主持论坛。

6月23日，清华大学，在方塘论坛上，赵汀阳（右）在主题发言后，与汪晖擦肩而过。

从两年前开始，许多科学家就曾疾呼：至少三分之一AI研发预算要用于确保安全。现实却是，全球AI安全研究明显滞后于技术发展。

“现在的大模型还没有‘嚣张’到那个程度，所以还没有产生毁灭世界的结果，但是长期来讲，我们要把这个事管制好。”姚期智强调，安全治理的当务之急是发展“AI对齐（AI Alignment）”及评估方法。

他提供了两种思路。一种思路要从博弈论角度展开，让AI的行为与人类想法对齐，以人类的利益为主；另一种“更有把握、更彻底”的思路是，设计可证明的安全的AGI（通用人工智能），每一步都符合数学上绝对安全的要求，“凡是有可能出问题的系统大家都不要做”。

技术细节依然艰深，暗藏的安全隐患愈发明显。

AI对齐，是人工智能安全领域的核心议题，旨在确保AI系统的目标、决策和行为与人类的价值观、偏好及意图协调一致。然而在“对齐”这件事上，中国社会科学院学部委员赵汀阳的想法有点儿不一样。

须发皆白的他，跛着脚走上台，坐下，第一句话就带着哲学式的冷峻诘问。“AI需要和人对齐吗？”听闻此问，我有些愣住，困意全消。他继续惊人之语：“人们期望AI要和人的价值观对齐。我对这个想法是有一些疑问的。地球上的所有‘坏事’都是人类做的，其他生命没有做过任何超出生物需要的坏事。”

赵汀阳盯着阶梯教室里的100余人，笑声一片，我也笑了。但笑过之后，细细琢磨，他的这番话，更像是一记打向人性的耳光。AI也许是“性本善”的，只不过跟着人类的数据经验，“照猫画虎”数年，学了好，也学了坏。

他认为，人类试图把AI制造成为一种具有主体性的新物种，但这似乎是个自虐性的悖论。

人类既希望AI拥有超能力，去完成人做不了或不想做的事，又担心AI获得自我意识后会伤害人类。赵汀阳说，如果人类不教唆AI犯罪，AI应是安全的，对齐人类价值观的AI，则很可能通过模仿人类而变得危险。

这个逻辑反转让我头皮发麻——原来我们恐惧的，是人性投射在AI身上的影子。

至于设立“AI伦理”，赵汀阳觉得无济于事。“人类也往往见利就忘义，如果伦理不能必然约束人类自身，又何以能够约束AI？”他主张，管理AI的关键在于人类是否能保有控制AI的能力。

赵汀阳眼里另有一种危险性相对比较低的对齐，是智能对齐。

“就目前智能水平而言，人类相对于AI仍然保有知己知彼的优势，因此能够控制AI。”赵汀阳说，但当AI部分经验能力远超人类后，人类优势荡然无存，AI将成为令人不寒而栗的“他者”，世界将形成二元主体性的格局，而人类与AI的关系将如何，无从知道。

不过，在中国科学院自动化研究所研究员、人工智能伦理与治理研究中心主任曾毅看来，未来的社会可能比二元主体性更复杂。他在随后的发言中谈到自己的设想，未来的社会不仅有人类、动物、超级智能，可能还有模仿狗、植物的类生命体。

那是一个他期待的可持续的共生社会，靠人和超级智能共同构造，而不是靠人类自己。曾毅认为，“所以向人类的对齐确实是不对的，要做超级的联合对齐。”联合对齐，这个词很新鲜，带着一种“天下大同”的意味。

但他认为，在这样的社会中，更重要的是，人类的价值观要演化。

“比如当超级智能未伤害到人类时，可不可以有自己的隐私？”曾毅提出这个问题时，我愣了一下。我从来没思考过，机器某一天可以有自己的隐私？那意味着，它相对自由、独立，不再是人类的附庸和工具。

“当人对超级智能说，‘我是你的创作者，你要保护我’时，超级智能可能会跟人说，‘当我看到你，就像你看到蚂蚁一样，你从来不会保护蚂蚁，我为什么要保护你’。”曾毅提醒道。

他觉得，人工智能是一面镜子。“当人工智能骗人，大家都吃惊，说人工智能怎么能骗人呢，太坏了。但是当人骗你的时候你的反应有这么大吗？恐怕没有。”在曾毅看来，人工智能的镜子让我们看到了人类的缺陷，给了人类演化的机遇。

“人工智能演化慢点没问题，但如果人类演化太慢了，那是真正的危险。”曾毅说。

说老实话，仅在“人类价值观演化”层面，我对人类的未来信心不足。一旦AI迈过奇点，超越人类，一切都将不尽如人意。

“作为一个自私的人，我希望人工智能‘性本善’。”曾毅记得，有一次，他在讲座中遇到了一位研修佛法者，问他AI能否成佛。他觉得，并非没有可能，超级智能除了认知超过人类，也可能是超级“利他”的。

听完曾毅的发言，汪晖说，他脑子里出现了赫尔曼·黑塞（Hermann Hesse）的《玻璃球游戏》中的镜像，人在多重镜面看到了自身。“我们都在讨论怎么驯化机器，我们也需要讨论人与自然、人与物的关系，重新理解人自身。”

上午的论坛结束，我最初的困意早已被冲刷干净。我合上记了4页纸的笔记本，感觉脑袋里比来时更“沉重”了。这场论坛，表面上是在探讨AI伦理奇点、智能边界，实际上始终在叩问自诩万物之灵的人类自己——那个既渴望神力、又恐惧神力的高级动物。

下午的3场分论坛分别围绕“2100年人类与超级智能的关系”“认识机器：意识、道德与风险”和“奇点前夜：人类如何安身立命”，我选择了第二场，但没听多久，就开始犯困，身边也有人开始打瞌睡，头不停地点着，像被阵阵微风拂过的芦苇。

发言者有一半人说英文，我戴着同声传译器，耳机里不时传来资本、奴隶、机器、工具等词汇，每一个字都听得懂，却没有一句能真正进入意识里。我感觉会场的学者离我越来越远，那些声音越来越小，有好几次，我差点儿睡着了，也许是真的太困，也许是脑容量过载了。

中途，我离开昏昏欲睡的会场。有一件更迫切的事等我去做——编辑发来次日可能刊发的稿子，催我修改。在蒙民伟人文楼的大厅里，我找到了一块插线板，打开电脑，席地而坐，开始忙碌。

神奇的是，一旦回到具体而微的“人类事务”，我的精神很快就清醒了。也许，相比于思考数十年后与AI的关系、厘清机器的伦理边界、琢磨人类如何安身立命这类宏大的时代命题，此刻的我更在乎如何在当下“安身立命”。

只是脑子里，还隐约回响着那句警言：“如果人类演化太慢了，那才是真正的危险。”