

深圳大学考试答题纸

(以论文、报告等形式考核专用)
二〇二二~二〇二三学年度第二学期

课程编号	1504550001	课程名称	计算机伦理	主讲教师	贺颖	评分	
学号	2021150122	姓名	郑雨婷	专业年级	2021 级计算机科学与技术高性能班		

教师评语：

一、生成式人工智能技术的发展历程及社会影响 (20 分)：

二、生成式人工智能技术所涉及的伦理问题分析 (30 分)：

三、个人阐述 (40 分)：

四、报告写作 (10 分)：

题目：关于 ChatGPT 等新一代人工智能技术对人类社会的影响

说明:

- 不要删除或修改蓝色标记的文字,也不要删除线框。
- 请在相应的线框内答题,答题时请用 5 号、宋体、黑色文字、单倍行距。

题一(20分)、题二(30分)、题三(40分)、报告格式和总体撰写情况(10分)

一、生成式人工智能技术的发展历程及社会影响(20分)

1.1 生成式人工智能技术的发展简介。(10分)

生成式人工智能是人工智能的一个子领域,是一种基于机器学习和深度学习的技术,旨在模拟人类的创造性思维和逻辑能力,通过学习数据的模式和规律来生成新的内容、图像、音频或文本等。生成式人工智能模型通过观察现有数据集中的样本,学习到其中的模式和特征,并使用这些学习到的知识创造出与原始数据相似但又有创造性的新内容。

新一代生成式人工智能技术中运用了许多具有重要意义的相关技术,例如预训练模型、条件生成模型、自监督学习、风格迁移技术和图像编辑技术等。最近大热的 ChatGPT 的 GPT 代表的就是一种大规模预训练模型。下面具体介绍 GPT 技术背景和发展历程。

GPT(Generative Pre-trained Transformer)技术是一种自然语言处理技术,它通过先前的大规模文本语料库进行预训练,然后使用这些数据来生成新的自然语言文本。该技术的核心是 Transformer 模型,它是一种基于注意力机制的神经网络架构,广泛应用于自然语言处理领域。GPT 技术的发展得益于近年来大量文本数据的积累,以及计算能力的提高,使得预训练模型能够更好地捕捉语言的模式和语义。

2018 年,GPT-1 诞生,这一年也是 NLP(自然语言处理)的预训练模型元年。性能方面,GPT-1 有着一定的泛化能力,能够用于和监督任务无关的 NLP 任务中。虽然 GPT-1 在未经调试的任务上有一些效果,但其泛化能力远低于经过微调的有监督任务,因此 GPT-1 只能算得上一个还算不错的语言理解工具而非对话式 AI。

2019 年,GPT-2 如期而至。在性能方面,除了理解能力外,GPT-2 在生成方面第一次表现出了强大的天赋:阅读摘要、聊天、续写、编故事,甚至生成假新闻、钓鱼邮件或在网上进行角色扮演通通不在话下。在“变得更大”之后,GPT-2 的确展现出了普适而强大的能力,并在多个特定的语言建模任务上实现了彼时的最佳性能。

2020 年,GPT-3 出现,作为一个无监督模型,几乎可以完成自然语言处理的绝大部分任务,例如面向问题的搜索、阅读理解、语义推断、机器翻译、文章生成和自动问答等等。它包含了更多的主题文本,显然优于前代的 GPT-2。

2022 年,OpenAI 发布 ChatGPT。主要针对对话任务进行了优化,增加了对话历史的输入和输出,以及对话策略的控制。ChatGPT 在对话任务上表现出色,可以与人类进行自然而流畅的对话。

2023 年,OpenAI 发布 GPT-4。在随意谈话中,ChatGPT 和 GPT-4 之间的区别是很微妙的。只有当任务的复杂性达到足够的阈值时,差异就出现了,GPT-4 比 ChatGPT 更可靠、更有创意,并且能够处理更细微的指令。为了了解这两种模型之间的差异,OpenAI 在各种基准测试和一些为人类设计的模拟考试上进行了测试,并且取得了非常好的结果。同时 GPT-4 有很强的多模态能力,可以理解图片。

总体而言,生成式人工智能的发展为创造性内容的自动生成提供了新的可能性,并在多个领域展示出巨大的应用潜力。未来,随着技术的不断推进和完善,生成式人工智能有望在创意创新、辅助创作和智能交互等方面发挥越来越重要的作用。

1.2 概述生成式人工智能技术的发展对社会的影响。(10 分)

随着生成式人工智能技术的快速发展，其应用越来越广泛地渗透到我们的生活中。正在对人类社会产生广泛而深远的影响，这些影响中有正面的，也有负面的，有积极的，也有消极的。以 GTP 技术为例，下面列举器发展对社会的影响：

一、就业方面。

GPT 技术的发展对就业市场产生了影响，特别是对软件工程师和其他相关领域。自动化和其他生成式人工智能模型变得越来越复杂，导致某些工作被替代。为了缓解这种情况，软件工程师应该注重提升自己的技能，并转入其他工作市场。

二、提高生产力。

GPT 技术的发展可以帮助人们生成大量符合规范且质量较高的文本，这将对一些行业产生深远影响。例如，在新闻、广告和市场营销等领域，GPT 生成的文章可以帮助记者快速完成稿件撰写、广告公司快速生成广告文案。代码生成是生成式人工智能的另一个令人兴奋的应用，它可以帮助开发人员更快速、更有效地编写代码。通过学习现有代码库，人工智能系统可以生成代码片段甚至整个应用程序，减少软件开发所需的时间和精力。

三、有助于医疗产业发展。

GPT 技术的发展有助于医疗产业发展。在药物研发和材料科学中，生成式人工智能具有带来变革的潜力。通过生成新型分子结构并分析其性质，人工智能技术可以帮助研究人员以前所未有的效率确定有前途的新化合物和材料，为改变人类生活带来希望。

四、促进信息传播。

GPT 技术的发展能够促进信息传播。生成式人工智能可以生成各种语言形式的文本，包括新闻报道、科技论文、小说等，这将促进信息传播和知识分享。事实上，已经有一些新闻机构和媒体开始使用 GPT 模型来自动生成文章。并且，生成式人工智能在机器翻译方面具有广泛应用。它可以帮助消除语言障碍，将信息从一种语言翻译成其他语言，促进跨文化交流和信息共享。

五、信息失真。

GPT 技术的发展可能导致信息失真。如果数据集中存在不平衡的样本或者含有不准确、歧视性或有害的内容，模型可能会学习并生成具有类似偏差和失真的信息。生成式人工智能模型可能在理解和解释语言的复杂性方面存在局限性。这种情况下，模型可能会误解用户输入或产生错误的输出。例如，对于一些含糊或具有多义的问题，模型可能无法正确解释用户意图，进而导致生成的信息含有失真。

六、侵犯隐私。

GPT 技术的发展可能造成隐私侵犯。生成式人工智能需要大量数据进行训练，在这个过程中可能会涉及个人隐私数据的使用和泄露。如果数据处理不当或未经合法授权，可能导致隐私泄露和滥用的风险。

总之，生成式人工智能技术的发展对社会产生了广泛而深远的影响，以上只是其中可列举出来的一部分。我们应该主动接受生成式人工智能并了解其进展，利用其力量来彻底改变产业、重新定义内容创作，并以前所未有的方式重塑我们的生活。同时，我们也需要采取必要的措施来确保其的发展符合社会的利益和价值观。

二、生成式人工智能技术所涉及的伦理问题分析。（30 分）

知识点 1:（10 分）

（1）教材中的对应章节（1 分）：第 7 章第 1 节，128 页。

（2）对教材中对应内容的简要概括（2 分）：知识产权是指公民、法人或者其他组织团体在科学技术方面或文化艺术方面，对创造性的劳动所完成的智力成果享有的专有权利；制止不正当竞争也包括在知识产权中。任何人、组织或单位都不能将未授权的智力劳动成果肆意使用，否则就属于侵犯知识产权。

（3）生成式人工智能技术涉及相同内容（2 分）：生成式人工智能在未经原作者的同意下，就使用他人的作品进行模型训练，这属于侵犯知识产权。除此之外，在训练完成之后还会利用生成式人工智能生成相似的作品，这更属于侵犯知识产权。

（4）关于两者相同之处的简要说明（5 分）：生成式人工智能可能涉及侵犯知识产权的问题。现有的许多文章、图画、音乐等都在知识产权保护的范围之内。由于生成式人工智能可以学习这些作品，如果在生成过程中产生与已存在的作品相似甚至相同的内容，这就会引发版权和知识产权的争议。例如，生成式人工智能可以使用大量的文本数据进行训练，学习并生成新的文章或文件。如果生成的作品与已经存在的作品非常相似甚至一模一样，就涉及到对原作品的侵权。

知识点 2:（10 分）

（1）教材中的对应章节（1 分）：第 8 章第 2 节，154 页。

（2）对教材中对应内容的简要概括（2 分）：公共数据库是指在一个单位、企业或城市、地区甚至一个国家等范围内建立一个信息共享的平台，其中所有的数据库可以共享。但是这些信息技术在实际使用中存在很大问题，很大一部分的公共数据库都存在触犯信息保护法律法规，侵犯个人隐私或人权的问题。

（3）生成式人工智能技术涉及相同内容（2 分）：生成式人工智能是通过训练大规模的文本数据集来生成响应。在模型的训练过程中，使用了包含各种来源的公共数据库，如互联网上的网页、电子书、论文、新闻文章等。这些公共数据库提供了广泛的语言表达方式，帮助模型学习自然语言处理和语义理解。

（4）关于两者相同之处的简要说明（5 分）：生成式人工智能本身并不直接侵犯隐私权，因为它使用的公共数据库按理应该是经过去标识化处理的，其中的个人身份和敏感信息已被删除。然而，在训练模型的过程中使用的公共数据库可能存在侵犯个人隐私和人权的问题。以英国为例，目前有 46 个大型公共数据库，但其中只有 6 个数据库（如电视执照系统等）被认为没有侵犯个人隐私和人权。虽然生成式人工智能模型本身不会直接泄露个人信息，但在使用公共数据库进行模型训练时，需要采取严格的数据处理措施和隐私保护机制，以确保个人隐私的安全。

知识点 3: (10 分)

(1) 教材中的对应章节 (1 分): 第 4 章第 3 节, 65 页

(2) 对教材中对应内容的简要概括 (2 分): 计算机技术职业和所有职业一样, 都应该遵循职业道德规范。计算机技术职业人员应当做到: 避免伤害他人、尊重知识产权、尊重其他人的隐私、保守机密、诚实可信、恪守公正并在行为上无歧视等。

(3) 生成式人工智能技术涉及相同内容 (2 分): 生成式人工智能是计算机领域的一项新兴技术。研究开发或者使用生成式人工智能的人员和其他计算机技术职业人员一样, 遵守计算机技术职业道德。

(4) 关于两者相同之处的简要说明 (5 分): 计算机技术职业人员在应用生成式人工智能时, 应当遵循以下主要职业道德原则:

- ①尊重隐私和个人权益: 确保在使用生成式人工智能的过程中严格遵守隐私政策和相关法律法规, 尊重用户的隐私权利, 并采取适当的措施来保护个人身份信息和敏感数据。
- ②满足法律和道德标准: 在生成式人工智能应用中, 计算机技术职业人员应始终遵守适用的法律法规和行业规范, 不从事任何非法、不道德或伦理上有争议的活动。
- ③提供真实和准确的信息: 确保生成的内容或结果不包含虚假、误导性或欺骗性信息。应避免以任何方式散布错误、误导性或有害的信息, 注重提供准确、可靠的结果。
- ④社会责任和影响评估: 在设计和实施生成式人工智能系统时, 必须考虑其可能带来的社会、经济和伦理影响。评估潜在风险并采取措施来最大限度地降低负面影响, 同时积极推动技术的正向应用、公平性和可持续性。
- ⑤避免歧视和偏见: 生成式人工智能应遵循公正和无偏差原则, 并避免基于种族、性别、年龄、宗教或其他个人属性的歧视行为。在数据收集、训练和验证阶段, 要警惕并纠正任何潜在的偏见。

三、个人阐述。(40 分)

3.1 就生成式人工智能技术对社会的深远影响及应对策略来论述自己的观点和看法。(20 分)

生成式人工智能技术对社会具有诸多影响。对于正面影响, 我们应该充分利用, 以提升生活水平, 对于其负面影响, 我们应当制定相应的应对策略, 尽可能地减小负面影响。

面对生成式人工智能对就业市场产生的影响, 我认为需要辩证地看待。有不少人担心 ChatGPT 或其他生成式人工智能的诞生会逐渐取代某些职业, 例如程序员、会计员、电话客服等, 因为显然计算机的算力和学习能力比人类要强, 因此, 相关从业人员确实要提升自己的专业技能, 以防止被人工智能取代。但也不需过的担忧, 尽管生成式人工智能功能强大, 能够很好地完成某项具体的工作, 但在一个企业公司中, 它无法做到各个部门之间的沟通和对接, 还是需要人类对其进行操作。

面对生成式人工智能对有助于生产力提高的影响, 我们应充分、正确地利用。生成式人工智能无疑能够大大提高工作效率, 提高生产力, 但前提是我们能够合法合理地利用它。因此, 国家政府部门和公司企业应当加强教育与培训, 培养公众对于生成式人工智能技术的理解和认知, 提供相关

教育和培训，帮助人们应对技术发展带来的挑战和机遇。加强教育与培训我们应该正确地生成式人工智能可以帮助人们生成大量符合规范且质量较高的文本，这将对一些行业产生深远影响。例如，在新闻、广告和市场营销等领域，GPT 生成的文章可以帮助记者快速完成稿件撰写、广告公司快速生成广告文案。代码生成是生成式人工智能的另一个令人兴奋的应用，它可以帮助开发人员更快速、更有效地编写代码。通过学习现有代码库，人工智能系统可以生成代码片段甚至整个应用程序，减少软件开发所需的时间和精力。

面对生成式人工智能对医疗产业的影响，在确保技术的安全性和有效性后，可以为医生和患者提供更好的医疗服务。包括以下几点：生成式人工智能可以应用于医学影像的分析和诊断，辅助医生进行早期疾病检测、异常发现和治疗计划制定；生成式人工智能可以自动生成医学报告，从医学记录中提取关键信息并生成结构化的报告，提高工作效率和减少医生的负担；生成式人工智能可以模拟和加速药物设计过程。通过学习药物分子的结构和性质，生成式人工智能可以帮助科学家预测药物的活性、毒副作用和适应症，加速新药的研发过程；生成式人工智能可以结合临床数据、基因组学信息和病历资料，辅助医生进行疾病诊断和治疗方案选择。通过学习大量的医学文献和病例经验，生成式人工智能可以提供个性化的诊断建议和治疗方案；生成式人工智能可以在肿瘤精准治疗中发挥重要作用。通过分析患者的基因组数据、病历信息和临床研究数据，生成式人工智能可以预测肿瘤的复发风险、选取最佳的治疗方案，并帮助制定个性化的治疗策略。

面对生成式人工智能对信息传播产生的影响，我们应仔细甄别信息。生成式人工智能可以生成各种语言形式的文本，包括新闻报道、科技论文、小说等，这将促进信息传播和知识分享。已经有一些新闻机构和媒体开始使用 GPT 模型来自动生成文章。生成式人工智能可以用于生成虚假信息、伪造内容或模仿声音、视频等，因此我们需要保持警惕，确保所接触到的信息来源可靠。要多方核实信息，获取来自权威机构、专业媒体或可信渠道的消息。面对生成式人工智能对信息传播的影响，我们需要保持警惕、加强监管和规范，并推动科学研究与社会共识的形成，以确保生成式人工智能的应用符合法律、道德和社会利益。

面对生成式人工智能造成的信息失真，我们需要保持警惕，多方查证信息，避免被误导或混淆视听。如果数据集中存在不平衡的样本或者含有不准确、歧视性或有害的内容，模型可能会学习并生成具有类似偏差和失真的信息。生成式人工智能模型可能在理解和解释语言的复杂性方面存在局限性。这种情况下，模型可能会误解用户输入或产生错误的输出。例如，对于一些含糊或具有多义的问题，模型可能无法正确解释用户意图，进而导致生成的信息含有失真。因此应该强化法律法规和监管，制定相关法律法规，明确生成式人工智能技术的边界和规范，保护个人隐私和数据安全，同时加强监管和审查制度。

面对生成式人工智能可能导致的隐私泄露风险，我们可以采取以下措施加强个人数据安全意识：选择可信赖的应用和服务供应商；仔细阅读隐私政策，了解数据收集、处理和存储方式；保护个人信息，避免轻易泄露敏感数据；使用强密码，并定期更新软件和操作系统；谨慎选择分享信息的对象和范围，并注意个人隐私设置；避免在公共场合透露敏感信息；运用加密通信工具和虚拟专用网络（VPN）增强数据传输和存储的安全性；定期检查在线账户和应用程序权限，并及时注销或取消不再需要的服务；了解个人信息使用的范围和目的，并要求删除无关信息。这些措施有助于减少隐私泄露风险，提高个人数据的安全性。同时，政府和相关部门应当定相关法律法规，明确生成式人工智能技术的边界和规范，保护个人隐私和数据安全，同时加强监管和审查制度。

总之，应对生成式人工智能技术带来的深远影响，需要全面考虑经济社会因素和伦理法律原则，积极引导和监管技术的发展和运用，推动科技与人类共同进步，实现人工智能的良性发展。

3.2 搜索实际生活中案例，论述新一代人工智能对个人及社会生活带来哪些影响。（20 分）

案例 1:

（1）案例出处（如网址）：<https://zhuanlan.zhihu.com/p/363642747>

（2）案例简述（100-200 字）：为提高配送效率，一些外卖平台借助 AI 算法，平台可以最优化地安排订单，也能给骑手规划最合理的路线。但出于平台、骑手和用户三方效率最大化的目标，AI 算法将所有时间压缩到了极致，为了按时完成配送，骑手们只能用超速去挽回超时这件事。超速、闯红灯、逆行……外卖骑手挑战交通规则的行为是一种逆算法，是骑手们长期在系统算法的控制与规训之下做出的不得已的劳动实践，而这种逆算法的直接后果则是——外卖员遭遇交通事故的数量急剧上升。

（3）案例简析（100-200 字）：这个案例虽然讲的是外卖骑手，但我们可以看出人工智能治理问题不是某一个单一主体可以解决的。我们可以看到，在这个案例里还包括交通规划、收入保障体系不到位等问题，我的结论是——我们需要一个多元主体共治的治理体系。在人工智能普及的时代，也应充分保护劳动者的合法权益，人工智能算法的应用在追求效率和利益之外，不可忽略人性与理性，不能让机器的冰冷碾压人性的光辉。

案例 2:

（1）案例出处（如网址）：<https://zhuanlan.zhihu.com/p/363642747>

（2）案例简述（100-200 字）：“水表”和“老人安全”看似没有关联，但在上海市长宁区江苏路街道“一网统管”平台，隐藏背后的逻辑清晰可见——当独居老人家中超过 12 小时用水不足 0.01 立方米时，系统会判断老人家中有事，并及时向居委会预警。一个小小的设计，没有像摄像头那样不得不关注老人过多的隐私，也无需大量的经济和人力投入，就能实现对老年人的关爱，甚至关键时刻能挽救老人的生命。

（3）案例简析（100-200 字）：AI 技术的落地应用，在一定程度上可有效提升人们生活的安全性，这个案例通过水表对独居老人进行安全预警，体现对独居老人的无感关爱，这对当前促进社会民生改善有一定参考价值。当然，这个案例中要提高水表计数的敏感度和数据采集的及时性与准确性，才能降低误报率。