

OpenAI软件系统开发-漫云科技-行业app软件火爆项目

产品名称	OpenAI软件系统开发-漫云科技-行业app软件火爆项目
公司名称	河南漫云科技有限公司
价格	.00/个
规格参数	漫云科技:区块源码 区块系统:定制开发 区块app系统开发:快速搭建
公司地址	郑东新区升龙广场3号楼A座3202
联系电话	13103827627 13103827627

产品详情

该公司表示，他们提供的AI工具能够提高工作效率、增强创造力，河南漫云科技有限公司成立于2019年，是一家以“践行工匠精神，用心服务客户”为使命，小程序定制开发、行业软件开发、为企业打造属于自己平台的开发解决，小程序定制开发、行业软件开发、为企业打造属于自己平台的开发解决，公司目前已经为大中小型企业提供了一站式的APP开发服务，还提供量身定制的学习体验。然而，他们也认识到这些工具存在潜在风险，因此他们在各个系统层面上努力确保安全性。他们不仅仅在推出新的人工智能系统之前进行严格的测试，还征求外部专家的意见并且通过人工反馈的强化学习等技术来改善模型的表现。同时，他们建立了广泛的安全和监控系统。OpenAI积极与各国接触，探讨的监管形式，以确保创建和发布更加安全的人工智能系统。以下是OpenAI博客文章全文：OpenAI致力于确保强大的AI安全，让尽可能多的人受益。我们知道，我们的AI工具为当今人们提供了许多帮助。世界各地的用户向我们反馈，ChatGPT有助于提高他们的工作效率、增强他们的创造力，并提供量身定制的学习体验。但我们也认识到，与任何技术一样，这些工具也存在真正的风险——因此我们正在努力确保在各个系统层面上的安全性。构建越来越安全的人工智能系统 在推出任何新的人工智能系统之前，我们都会进行严格的测试，征求外部专家的意见，并通过人工反馈的强化学习等技术来改善模型的表现。同时，我们还建立了广泛的安全和监控系统。以我们的模型GPT-4为例，完成训练后，我们在整个公司范围内进行了长达6个月的测试，以确保其在公开发布之前更加安全可靠。我们认为，强大的人工智能系统应该接受严格的安全评估。监管是必要的，以确保这种做法被广泛采纳。因此，我们积极与各国接触，探讨的监管形式。从实际使用中学习以改进保障措施 我们尽力在系统部署之前预防可预见的风险，但实验室中的学习永远有限。我们广泛研究和测试，但无法预测人们会如何使用我们的技术，或滥用它。因此，我们认为从实际使用中学习，是创建和发布越来越安全的人工智能系统的关键组成部分。我们谨慎地将新的人工智能系统逐步发布给人群，并采取实质性的保障措施，并根据我们吸取的教训不断改进。我们提供自己服务和API中*强大的模型，以便开发人员可以直接将技术集成到他们的应用程序中。这使我们能够监控滥用行为并采取行动，同时制定应对措施。这样，我们可以采取实际行动，而不仅仅在理论上想象如何应对。实际使用中的经验也促使我们制定了越来越细化的政策，以应对对人们构成真正风险的行为，同时仍然允许我们的技术以更有益的方式使用。我们认为，社会需要更多时间适应日益强大的人工智能，每个受其影响的人都应该在人工智能的进一步发展方面拥有发言权。迭代部署有助于不同利益相关者更有效

地参与人工智能技术的对话中，而拥有使用这些工具的手经验至关重要。 保护儿童 我们安全工作的重点之一是保护儿童。我们要求人们必须年满18岁，或者在父母同意的情况下年满13岁，才能使用我们的AI工具。目前，我们正在研究验证功能。 我们不允许我们的技术被用于生成仇恨、骚扰、或成人内容等类别。与GPT-3.5相比，我们的模型GPT-4对受限制内容请求做出响应的可能性降低了82%，并且我们已经建立了一个强大的系统来监控滥用情况。GPT-4现在可供ChatGPT Plus订阅者使用，我们希望随着时间的推移让更多人可以使用它。 我们已做出重大努力，以尽量减少我们的模型生成伤害儿童内容的可能性。例如，当用户试图将儿童安全虐待材料上传到我们的图像工具时，我们会阻止它并向国家失踪和受剥削儿童中心报告此事。 除了我们默认的安全护栏外，我们还与非营利性可汗学院等开发人员合作——该学院开发了一个人工智能助手，既可以作为学生的虚拟导师，也可以作为教师的课堂助手——为他们量身定制安全措施。我们还在研究允许开发人员为模型输出设置更严格标准的功能，以更好地支持需要此类功能的开发人员和用户。 尊重隐私 我们的大型语言模型是在广泛的文本语料库上进行训练的，其中包括公开可用的内容、获得授权的内容以及由人类审核人员生成的内容。我们不利用这些数据来销售我们的服务或，也不用它们来建立个人档案。我们只是利用这些数据来让我们的模型更好地为人们提供帮助，例如通过与人们进行更多对话来提高ChatGPT的智能水平。 尽管我们的许多训练数据中包括可以在公共网络上获得的个人信息，但我们希望我们的模型了解的是整个世界，而不是个人。因此，我们致力于在可行的情况下从训练数据集中删除个人信息，微调模型以拒绝个人信息的查询请求，并响应个人从我们的系统中删除其个人信息的请求。这些措施将我们的模型生成包含个人信息的响应的可能性降至。 提高事实准确性 现今的大型语言模型，基于之前的模式和用户输入的文本，可以预测下一个可能要使用的词汇。但在某些情况下，下一个*有可能出现的词汇实际上可能存在事实性错误。 提高事实准确性是OpenAI和许多其他AI研究机构关注的重点之一，我们正在取得进展。通过利用被标记为不正确的ChatGPT输出的用户反馈作为主要数据来源，我们提高了GPT-4的事实准确性。相比于GPT-3.5，GPT-4更可能产生符合事实的内容，提高幅度达到40%。 当用户注册使用该工具时，我们努力做到尽可能透明，以避免ChatGPT可能给出错误回复。然而，我们已经认识到，要进一步降低误解的可能性，并教育公众了解这些AI工具目前的局限性，还有很多工作要做。 持续的研究和参与 我们认为，解决AI安全问题的切实可行方法是投入更多时间和资源来研究有效的缓解和校准技术，并测试它们在现实世界中可能的滥用情况。 重要的是，我们还认为，提高人工智能的安全性和能力应该齐头并进。迄今为止，我们的安全工作来自于我们*强大的模型，因为它们更善于遵循用户的指示，更容易操纵或「引导」。 我们将越来越谨慎地创建和部署功能更强大的模型，并将随着我们的人工智能系统的发展继续加强安全预防措施。 虽然我们等了6个多月才部署GPT-4以更好地了解其功能、优势和风险，但有时可能需要比这更长的时间来提高AI系统的安全性。因此，政策制定者和AI开发商需要确保AI的开发和部署在范围内得到有效监管，这样就没有人会为了取得成功而采取捷径。这是一项艰巨的挑战，需要技术和制度创新，但我们渴望为之做出贡献。 解决安全问题还需要广泛的辩论、实验和参与，包括人工智能系统行为的界限。我们已经并将继续促进利益相关者之间的合作和公开对话，以创建一个安全的人工智能生态系统。