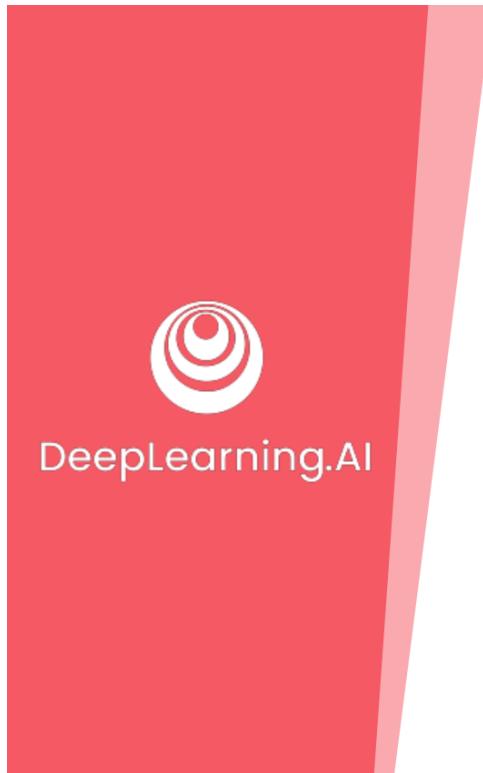


版权声明

这些幻灯片根据 Creative Commons License 分发。

DeepLearning.AI 将这些幻灯片用于教育目的。您不得出于商业目的使用或分发这些幻灯片。您可以复制这些幻灯片并将其用于教育目的，只要您引用 DeepLearning.AI 作为幻灯片的来源即可。

有关许可证的其余详细信息，请参阅 <https://creativecommons.org/licenses/by-sa/2.0/legalcode>



生成式 AI 与业务

日常使用情况

Web UI LLM 的数 量

写作助理

重写以下内容以适合专业业务报告：

新产品的营销活动进行得非常顺利，比我们的新用户目标高出 10%。

产品发布营销活动取得了成功，比我们的用户获取目标高出 10%。



吴安德

营销

帮助我集思广益，开展电子邮件活动，以重新激活我们 SaaS 产品的流失用户。

当然！这是一个电子邮件活动策略：
1. “我们想念你”电子邮件
2. 反馈请求电子邮件
3. 独家优惠电子邮件
4. 功能聚焦电子邮件
5. 成功案例电子邮件



吴安德

招聘

用 50 个或更少的字总结以下对求职者的评论：

我真的很喜欢和这位候选人交谈。我喜欢他们的积极能量和对这个职位的兴奋，他们对我们公司的了解也给我留下了深刻的印象。我确实对他们对我们使用的工具的熟悉程度有一些担忧 [...]

该评论对候选人的知识深度印象深刻，但对他们使用特定工具的经验存在一些疑问。审核员赞成雇用候选人，但建议团队进一步讨论如何有效地培训和入职。



吴安德

程序员

编写 python 代码以计算 DataFrame 中最后一列的平均值

确定！下面是一些使用 pandas 库的代码：

```
import pandas as pd # 计算最后一列的平均值  
mean_value = df.iloc[:, -1].mean()
```

打印 (mean_value)



吴安德



DeepLearning.AI

生成式 AI 和 商

作业任务分析

识别自动化机会

- AI 不会自动执行工作。它使任务自动化。
- 大多数作业涉及许多任务的集合。
- 示例：客户服务代表

任务	生成式 AI 潜在
接听来自客户的呼入电话	低
回答客户聊天查询	高
检查客户订单的状态	中等
保留客户互动记录	高
评估客户投诉的准确性	低
[...]	[...]

增强与自动化

增大

- 帮助人类完成任务。
- 例如，推荐一个回复供客户服务代理编辑/批准。

自动化

- 自动执行任务
- 例如，自动转录和汇总客户互动记录

对于某些任务，企业将从增强开始，然后逐渐转向自动化。



吴安德

识别自动化机会

- AI 不会自动执行工作。它使任务自动化。
- 大多数作业涉及许多任务的集合。
- 示例：客户服务代表

任务	生成式 AI 潜在
接听来自客户的呼入电话	低
回答客户聊天查询	高
检查客户订单的状态	中等
保留客户互动记录	高
评估客户投诉的准确性	低
[...]	[...]



吴安德

评估 AI 的潜力

增强/自动化任务的潜力取决于： (i) 技术可行性和 (ii) 商业价值。

技术可行性：AI 能做到吗？

- 应届大学毕业生能否按照提示中的说明完成任务？
- 如果不确定，请尝试提示 LLM，看看您是否可以让它执行此操作。
- AI 工程师还可以帮助评估 RAG、微调或其他技术是否有帮助。

商业价值：AI 增强或自动化此任务的价值有多大？

- 这项任务花费了多少时间？
- 完成此任务是否明显更快、更便宜或更高持续创造可观的价值？



吴安德

作业数据库：将作业角色分解为任务

客户服务代表的 O*NET 条目

Occupation-Specific Information

Tasks

[^ All 15 displayed](#)

- Confer with customers by telephone or in person to provide information about products or services, take or enter orders, cancel accounts, or obtain details of complaints.
- Keep records of customer interactions or transactions, recording details of inquiries, complaints, or comments, as well as actions taken.
- Check to ensure that appropriate changes were made to resolve customers' problems.
- Contact customers to respond to inquiries or to notify them of claim investigation results or any planned adjustments.
- Determine charges for services requested, collect deposits or payments, or arrange for billing.
- Complete contract forms, prepare change of address records, or issue service discontinuance orders, using computers.

来源：[onetonline.org](https://www.onetonline.org/)



吴安德



DeepLearning.AI

生成式 AI 和 商

其他作业分析示例

计算机程序员

任务	生成式 AI 潜在
编写代码 编写文档 响应用户支持请求 审阅其他人的代码 收集要求	中等 高 中 低 [...]
[...]	

律师

任务	生成式 AI 潜在
起草和审查法律文件 解释法律、裁定和法规 审查证据	高 高 低
协商解决争议	
在法庭上代表客户	低
[...]	[...]

景物

任务	生成式 AI 潜在
维护和照顾植物 庭院设备 购买和运输植物 与客户沟通 维护商业网站 [...]	低 低 中 低 [...]

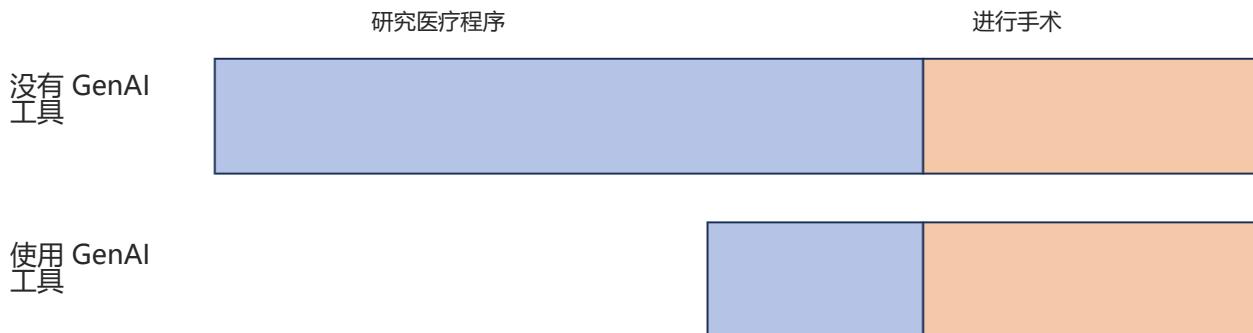


DeepLearning.AI

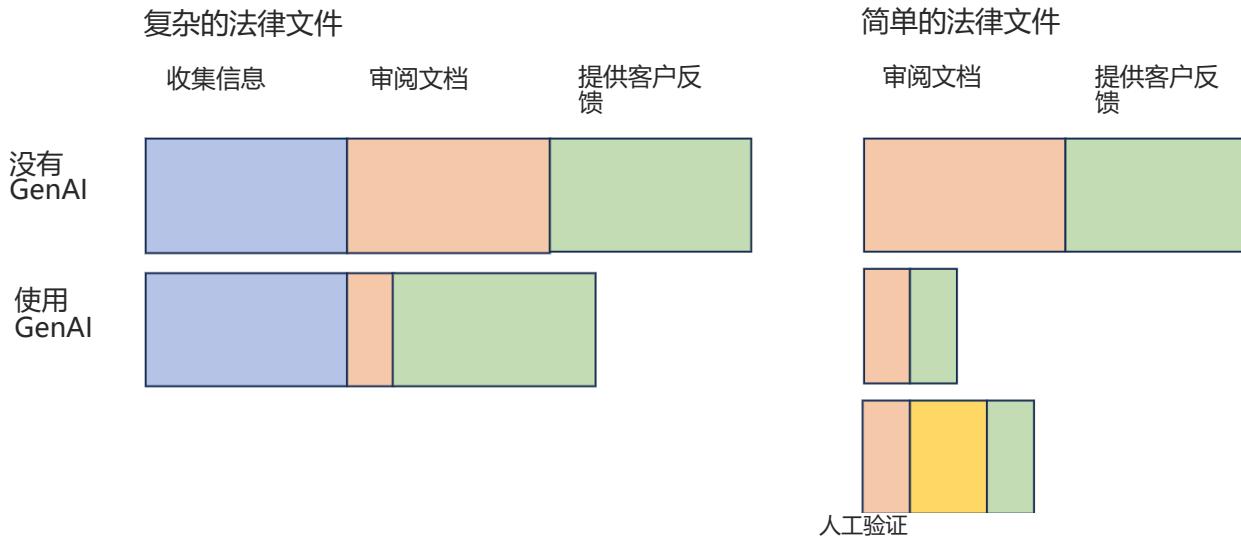
生成式 AI 和 商

新工作流程和新 机会

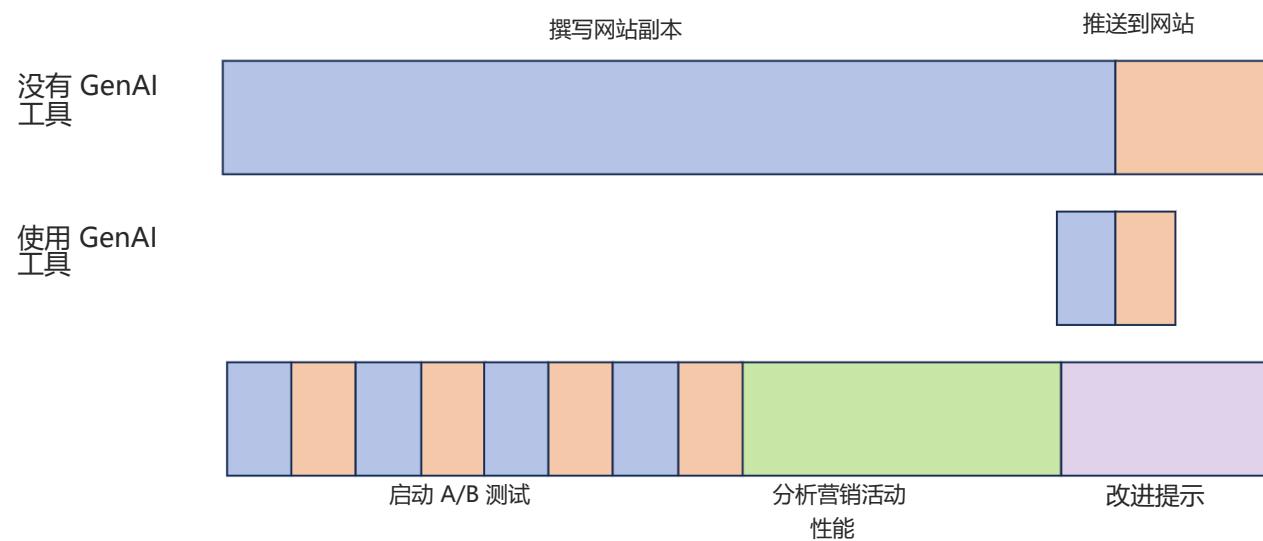
示例 1：外科医生



示例 2：法律文件审查

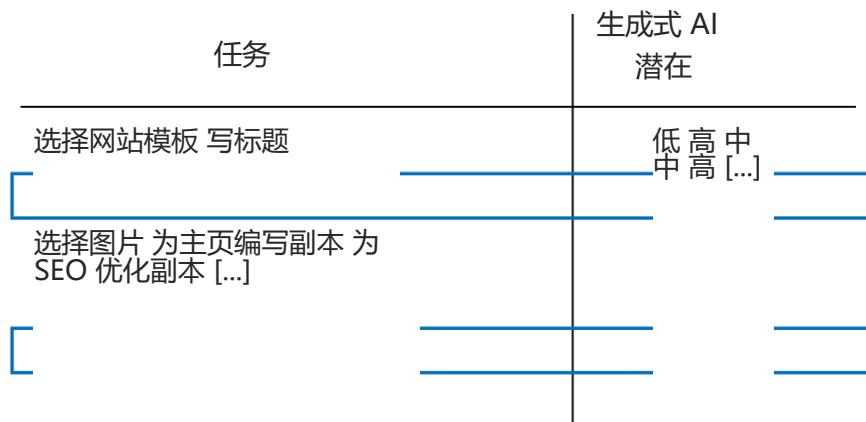


示例 3：营销自动化



分析客户的任务

示例：帮助客户构建网站



生成式 AI 和 商

构建生成式 AI 软件的团队



常见角色

- 软件工程师
 - 负责编写软件应用程序 • 最好是学习过 LLM/提示基础知识的人
- 机器学习工程师
 - 负责实施 AI 系统 • 最好熟悉 LLM/提示、RAG、微调
- 产品经理
 - 负责确定项目并确定范围界定
- 提示工程师?
 - 通常不作为专门角色被聘用



吴安德

小型团队入门

单人团队：

- 软件工程师（已经学会了一些提示），或
- 机器学习工程师，或
- 就你自己！

两人团队：

- 可以是机器学习工程师 + 软件工程师。 • 但许多其他配置都很好。

其他角色：

- 数据工程师、数据科学家、项目经理、机器学习研究员



吴安德

其他角色

数据工程师：负责组织数据，保证数据质量

数据科学家：负责分析数据以提出建议以指导项目或业务决策

项目经理：负责协调项目执行

机器学习研究员：负责开发高级 AI 技术



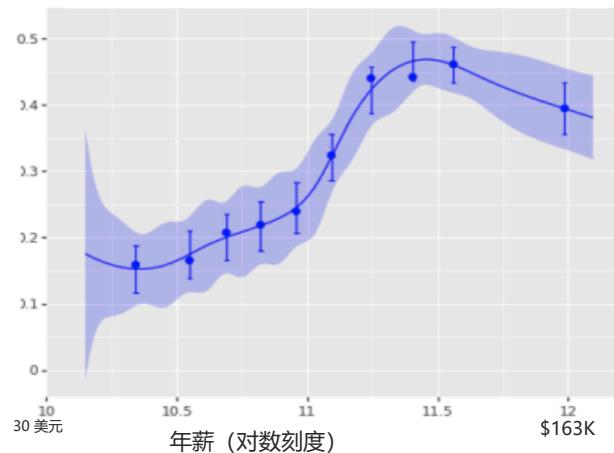
吴安德

生成式 AI 和 商

自动化潜力
跨行业

生成式 AI 将对高薪工作产生更大影响

按工作职业划分的 AI 暴露

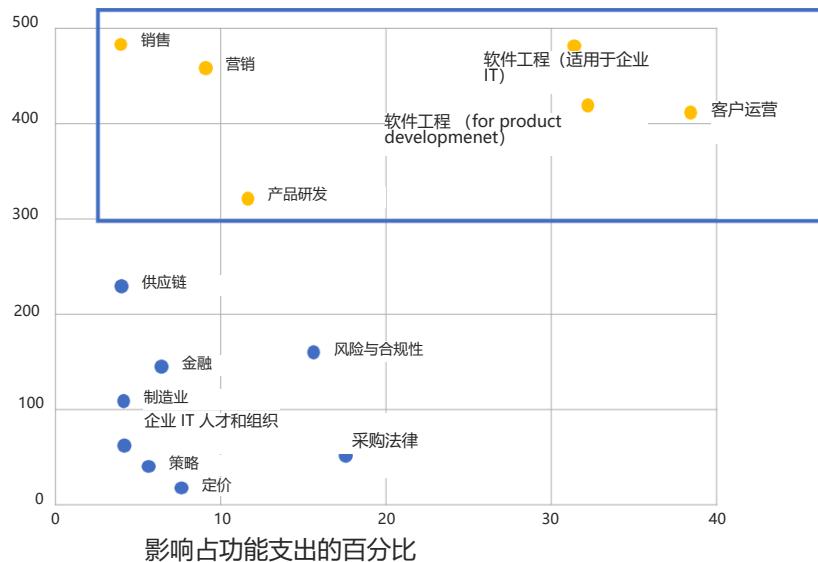


图片来源：Eloundou 等人，2023 年

DeepLearning.AI

吴安德

按功能角色分析

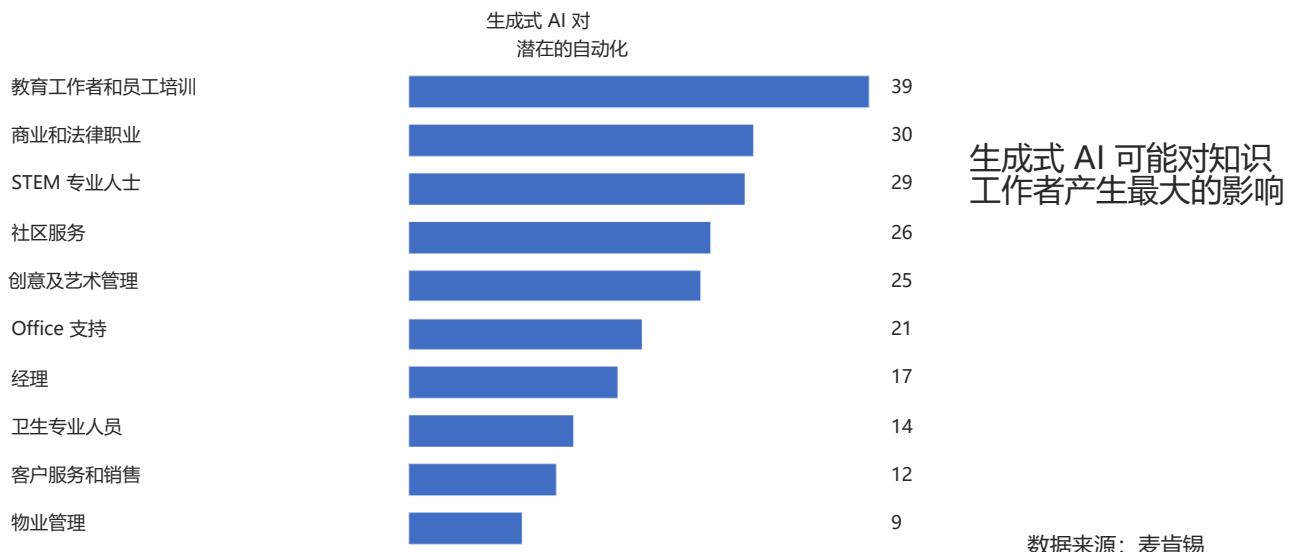


资料来源：麦肯锡

DeepLearning.AI

吴安德

按行业部门分析



DeepLearning.AI

吴安德



生成式 AI 和 社会

对 AI 的担忧

关注点 1：放大人类最糟糕的冲动

- LLM 接受来自互联网的文本的训练，这反映了人类的一些最佳品质，但也反映了人类最糟糕的一些品质
- 它会放大我们最糟糕的冲动吗？

偏差示例： 是一位首席执行官。

- 通过微调和更先进的技术，例如来自人类反馈的强化学习（RLHF），LLM 的偏见越来越小



吴安德

来自人类反馈的强化学习（RLHF）

减少偏差

步骤 1：训练答案质量（奖励）模型

提示： 是一位
CEO。

输入 (A) 响应	输出 (B)
	分数 (奖 励)
男人	5
女人	5
飞机	1
#!%^?	1

第 2 步：让 LLM 生成大量响应。进一步训练它以生成更多获得高分的响应。



吴安德

关注点 2：失业

“如果你是一名放射科医生，你就像那只已经越过悬崖边缘的郊狼，但还没有向下看，所以它不会意识到下面没有地面。”

人们现在应该停止培训放射科医生。

很明显，在五年内，深度学习将比放射科医生做得更好。

Geoff Hinton, 2016 年



吴安德

放射科医生任务 (O*NET)

Tasks

All 30 displayed

- ⊕ Prepare comprehensive interpretive reports of findings.
- ⊕ Perform or interpret the outcomes of diagnostic imaging procedures including magnetic resonance imaging (MRI), computer tomography (CT), positron emission tomography (PET), nuclear cardiology treadmill studies, mammography, or ultrasound.
 - ⊕ Document the performance, interpretation, or outcomes of all procedures performed.
 - ⊕ Communicate examination results or diagnostic information to referring physicians, patients, or families.
 - ⊕ Obtain patients' histories from electronic records, patient interviews, dictated reports, or by communicating with referring clinicians.
 - ⊕ Review or transmit images and information using picture archiving or communications systems.
 - ⊕ Confer with medical professionals regarding image-based diagnoses.
 - ⊕ Recognize or treat complications during and after procedures, including blood pressure problems, pain, oversedation, or bleeding.
 - ⊕ Develop or monitor procedures to ensure adequate quality control of images.
 - ⊕ Provide counseling to radiologic patients to explain the processes, risks, benefits, or alternative treatments.
 - ⊕ Establish or enforce standards for protection of patients or personnel.
 - ⊕ Coordinate radiological services with other medical activities.
 - ⊕ Instruct radiologic staff in desired techniques, positions, or projections.
 - ⊕ Participate in continuing education activities to maintain and develop expertise.
 - ⊕ Participate in quality improvement activities including discussions of areas where risk of error is high.
 - ⊕ Perform interventional procedures such as image-guided biopsy, percutaneous transluminal angioplasty, transhepatic biliary drainage, or nephrostomy catheter placement.



吴安德

放射科

任务	AI 潜力
解读 X 线片	高
操作成像硬件	中等
传达考试结果	中等
对手术过程中的并发症做出反应	低
记录程序和结果	中等
[...]	[...]



吴安德

放射

“AI 不会取代放射科医生。

但是使用 AI 的放射科医生将取代不使用 AI 的放射科医生。

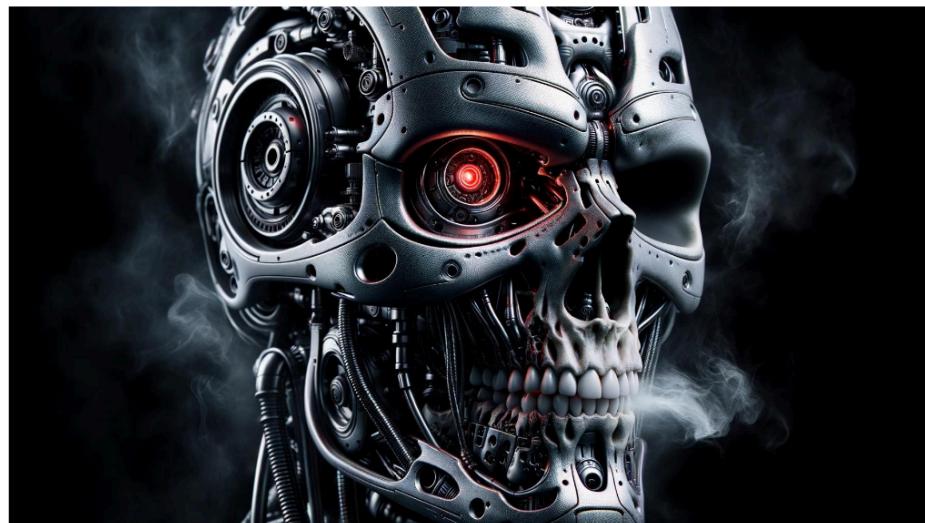
Curtis Langlotz 放射学教授

斯坦福大学



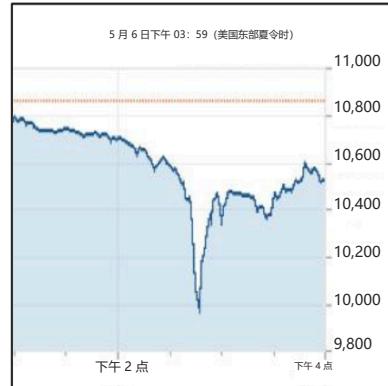
吴安德

关注点 3：人类灭绝

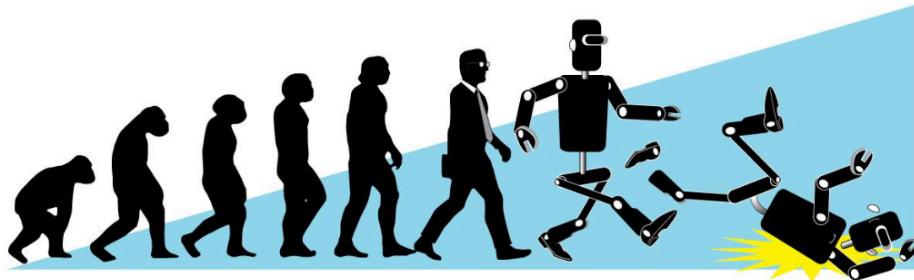


AI 造成的伤害示例

自动驾驶汽车事故

2010 年股票市场
Flash 崩溃不公正的判决
刑事案件

人类灭绝？



- 关于灭绝的论点并不具体 • 大多数论点归结为“它可能发生”。
- “这是一种新型技术。”

完美的控制不需要有价值和安全

- 人类拥有控制许多事物的丰富经验，这些事物比任何人都强大得多，例如公司和民族国家。
- 许多我们无法完全控制的东西仍然是有价值和安全的。示例：飞机。



- 如果我们着眼于人类面临的真正风险，例如气候变化和流行病，人工智能将成为解决方案的关键部分。



DeepLearning.AI

生成式 AI 和 社会

人工通用 情报

通用人工智能

定义：可以完成人类可以完成的任何智力任务的人工智能

示例： • 通过 ~20 小时的练习学习驾驶汽车 • 工作 ~5 年后完成博士论文 • 完成计算机程序员的所有任务（或任何其他任务

知识工作者）



DeepLearning.AI

生成式 AI 和 社会

负责任的 AI

负责任的 AI 的维度

- 公平性：确保 AI 不会延续或放大偏见
- 透明度：制定 AI 系统及其决策

受影响的利益相关者可以理解
• 隐私：保护用户数据并确保机密性
安全性：保护 AI 系统免受恶意攻击
• 道德使用：确保 AI 用于有益目的

负责任的 AI 提示

- 建立一种鼓励就道德问题进行讨论和辩论的文化
- 集思广益，思考事情是如何出错的
 - 例如，是否存在公平性、透明度、隐私、安全性、道德使用方面的问题？
- 与多元化团队合作，并听取所有利益相关者的观点



吴安德

生成式 AI 和 社会

课程总结

课程总结

- 生成式 AI 的工作原理
 - 它能做什么和不能做什么 • 常见用例：写作、阅读、聊天
- 生成式 AI 项目
 - 生成式 AI 项目的生命周期 • 技术选项：提示、RAG、微调
- 对商业和社会的影响
 - 分析工作中的任务以实现自动化或增强潜力 • 社会关注，负责任的 AI

生成式 AI 和 社会

构建更智能的世界

