



# GPT - 3:其性质、范围、限制和后果

Luciano floridi<sup>1,2</sup> • Massimo Chiriatti<sup>3</sup>

在线发布:2020 年 11 月 1 日 ©作者(s)  
2020

## 摘要

在这篇评论中,我们讨论可逆和不可逆问题的性质,即可能使人能够识别其答案来源的性质的问题。然后,引入第三代自回归语言模型 GPT-3,利用深度学习生成类人文本,并利用之前的区别对其进行分析。我们扩展了分析,提出了基于数学、语义(即图灵测试)和伦理问题的三种测试,并表明 GPT-3 不是为了通过其中任何一种测试而设计的。这提醒我们,GPT-3 并没有做它不应该做的事情,任何将 GPT-3 解释为人工智能普遍形式出现的开始都只是无知的科幻小说。最后,我们概述了自动化、廉价生产优质、语义化人工制品的工业化的一些重大后果。

关键词自动化 • 人工智能 • GPT-3 • 不可逆性 • 语义 • 图灵测试

## 1 介绍

谁修剪草坪, Ambrogio(机器割草机)<sup>1</sup> 还是 Alice?我们知道这两者在任何方面都是不同的:身体上,“认知上”(从内部信息处理的角度来说),以及“行为上”(从外部行动的角度来说)。然而,我们不可能完全确定地从修剪过的草坪上推断出是谁修剪的。不可逆性和可逆性并不是一个新想法(Perumalla 2014)。它们在许多领域都有应用,尤其是计算和物理领域。在

<sup>1</sup>这是一个真实的例子,参见 <https://www.ambrogio-bot.com/en>。披露:LF 拥有一家。

\* 卢西亚诺弗洛里迪

<sup>1</sup> 牛津互联网研究所,1 St Giles', 牛津 OX1 3JS, 英国

<sup>2</sup> 英国伦敦尤斯顿路 96 号,英国伦敦 NW12 8B, 图灵研究所

<sup>3</sup> IBM Italia, 大学项目负责人 - CTO 区块链 & Digital Currencies, 罗马, 意大利

以数学逻辑为例, 非门是可逆的(在这种情况下使用的术语是“可逆的”), 但异或(XOR)门是不可逆的(不可逆的), 因为一个人不能从它的单个输出毫无歧义地重构它的两个输入。这意味着, 就我们所知, 输入是可以互换的。在哲学中, 一个非常著名、相关的观点是不可分辨的同一性, 也被称为莱布尼茨定律: 对于任何  $x$  和  $y$ , 如果  $x$  和  $y$  具有所有相同的属性  $F$ , 那么  $x$  与  $y$  是相同的  $\square$  (财政年度)  $\rightarrow x = y$ ). 这意味着如果  $x$  和  $y$  具有相同的属性, 那么人们无法分辨(即相反)它们之间的差异, 因为它们是相同的。如果我们把所有这些放在一起, 我们就可以开始理解为什么“问题游戏”在用来猜测答案来源的性质或身份时可能会令人困惑。假设我们问了一个问题(过程), 并收到了一个答案(输出)。我们是否可以从答案中重构(反向)它的来源是人为的还是人工的? 答案是否像修剪过的草坪? 有些是, 但有些不是。这要看情况, 因为不是所有的问题都一样。数学问题( $2 + 2 = ?$ )、事实问题(法国首都是什么?)、二元问题(你喜欢冰淇淋吗?)的答案就像被割过的草坪一样“不可逆转”: 人们无法从这些问题中推断出作者的本质, 即使答案是错误的也不行。但是, 其他需要理解、甚至可能需要对意义和上下文都有经验的问题, 实际上可能会泄露它们的来源, 至少到目前为止是这样(这种限定是必要的, 我们稍后再来讨论它)。这些问题如“你的一只鞋能塞进几只脚?”或者“你能用一只鞋做什么?”我们把它们称为语义问题。

语义问题, 正是因为它们可能产生“可逆”的答案, 可以作为一种测试, 来识别其来源的性质。因此, 不言而喻, 认为人为来源和人工来源可能产生不可区分的答案是完全合理的, 因为某些类型的问题确实是不可逆的——同时指出, 仍然有(再次, 目前更多关于这一限定)某些类型的问题, 如语义问题, 可以用来识别人为来源和人工来源之间的区别。进入图灵测试。

本杂志的任何读者都将非常熟悉该测试的性质, 因此我们将在这里不进行描述。值得强调的是, 在图灵介绍他所谓的模仿游戏的著名文章(图灵 1950)中, 他还预测到 2000 年, 计算机将通过它:

我相信, 在大约 50 年的时间里, 我们将有可能对计算机进行编程, 使其存储容量约为  $10^9$ , 让它们把模仿游戏玩得非常好, 以至于一般的审讯者在审问 5 分钟后, 正确识别的几率不会超过 70%。(图灵 1950)

霍布斯花了大量的时间, 试图证明如何将圆平方。牛顿研究炼金术, 可能是试图发现哲学家之石。图灵相信真正的人工智能, 就像你在《星球大战》中看到的那样。即使是天才

犯错误。图灵的预测是错误的。如今，Loebner 奖(佛罗里迪等人，2009 年)被授予试图通过图灵测试的最不成功的软件。然而，从某种意义上说，图灵是对的：今天的计算机可以不可逆转地回答许多问题，我们思考和谈论机器的方式确实发生了变化。我们可以毫无疑问地说，计算机做这个或做那个，这样或那样想，或者学习如何做某事，而我们通过与它们对话来让它们做事情。此外，我们中的许多人怀疑它们的脾气不好。但图灵提出的是一种测试，而不是统计概括，它测试的是一些因此需要问的问题。如果我们对“不可逆性”感兴趣，对它在包括越来越多的任务和解决问题的活动方面可以走多远感兴趣，那么极限就是天空；或者更确切地说，是人类的创造力。然而，今天，语义问题的不可逆性仍然超出了任何现有的 AI 系统(Levesque 2017)。这并不意味着它们不能成为“不可逆的”，因为在一个越来越对人工智能友好的世界中，我们正在围绕计算人工制品的语法和统计能力包围我们现实的越来越多的方面(佛罗里迪 2019,2020)。但是，即使有一天语义问题不再能够使人们发现人类和人工来源之间的区别，最后一点仍然需要强调。在这里，我们要澄清一下我们在上面添加的附带条件。问题游戏(图灵的“模仿游戏”)只是一种否定(即必要但不充分)意义上的测试，因为不通过它就剥夺了 AI “智能”的资格，但通过它并不等于 AI 就有“智能”的资格。同样，Ambrogio 修剪草坪——产生的结果与爱丽丝所能达到的任何结果都没有区别——并不会让 Ambrogio 在任何意义上，无论是身体上、认知上还是行为上都像爱丽丝。这就是为什么“计算机不能做什么”对于该领域的任何出版物来说都不是一个令人信服的标题。从来都不是。关于 AI 的真正要点是，我们正在越来越多地将有效解决问题的能力-就最终目标而言-与任何智能的需求脱钩(佛罗里迪，2017 年)。通过这种解耦能实现什么，不能实现什么，这是一个完全开放的问题，关于人类的创造力、科学发现、技术创新和新的可供性(例如，越来越多的高质量数据)。3 这也是一个与智能、意识、语义、相关性以及更普遍的人类经验和正念无关的问题。这一解耦过程的最新进展是 GPT-3 语言模型

<sup>2</sup> 参见 <https://en.wikipedia.org/wiki/ELIZA>。关于伊莱莎效应和 AI 的经典著作仍然值得一读(Weizenbaum 1976)。2014 年，有人错误地声称，一个聊天机器人通过了测试。它的名字叫“Eugene Goostman”，大家可以自己检查一下，可以在这里把玩一下：<http://eugene-goostman.elasticbeats.com/>。当它被测试的时候，我是评委之一，我注意到的是，是一些没有通过测试的人类，提出了我在这里称之为“不可逆转”的问题，比如(真实的例子，这些是 BBC 记者问的)“你相信上帝吗？”和“你喜欢冰淇淋吗”。即使是一个简单的抛硬币机也能“通过”这种测试。

<sup>3</sup> 例如，参见 Winograd 模式挑战(Levesque 等，2012)。

<sup>4</sup> 有关优秀的技术和关键分析，请参阅 McAteer(2020)。关于“对 GPT-3 等大规模语言模型的完全不切实际的期望”，请参阅 Yann LeCun (Facebook 应用程序的副总裁、首席 AI 科学家)：<https://www.facebook.com/yann.lecun/posts/10157253205637143>。

2 呢?GPT- 3

OpenAI 是一个 AI 研究实验室，其明确目标是促进和发展能够造福人类的友好 AI。它成立于 2015 年，被认为是 DeepMind 的竞争对手。微软是 OpenAI 的重要投资者(10 亿美元投资(OpenAI 2019))，最近宣布与 OpenAI 达成协议，独家授权其 GPT-3(斯科特 2020)。

GPT-3(Generative pre - training Transformer)是第三代自回归语言模型，使用深度学习产生类人文本。或者更简单地说，它是一个计算系统，设计用来生成单词、代码或其他数据的序列，从一个源输入(称为提示符)开始。例如，在机器翻译中，它被用来从统计学上预测单词序列。语言模型在一个由文本组成的未标记数据集上进行训练，例如维基百科和许多其他网站，主要是英语，但也有其他语言。这些统计模型需要用大量的数据进行训练，才能产生相关的结果。2018 年 GPT 的第一次迭代使用了 1.1 亿个学习参数(即:，神经网络在训练过程中试图优化的值)。一年后，GPT-2 使用了 15 亿个。如今，GPT-3 使用 1750 亿个参数。它在微软 Azure 的 AI 超级计算机(斯科特 2020)上接受培训。这是一个非常昂贵的训练，估计花费了 1200 万美元(Wiggers 2020)。这种计算方法适用于广泛的用例，包括摘要、翻译、语法纠正、问答、聊天机器人、撰写电子邮件等。

自 2020 年 6 月以来，出于研究目的，我们已开始进行 beta 测试，最近我们有机会对其进行第一手测试。GPT-3 编写自动和自主的文本质量优良，按需。看到它在起作用，我们很好地理解了为什么它让这个世界既热情又恐惧。英国《卫报》最近发表了 GPT-3 撰写的一篇文章，引起了轰动(GPT-3 2020)。文章进行了编辑——编辑程度如何尚不清楚<sup>5</sup>这篇文章至少可以说是哗众取宠。有人认为这是误导，是拙劣的新闻报道(Dickson 2020)。我们倾向于同意。但这丝毫没有削弱这一体系的非凡效力。相反，它充分说明了你必须做些什么才能卖出一份报纸。

使用 GPT-3 非常简单，并不比通过搜索引擎搜索信息更难。就像谷歌在不理解查询的情况下“读取”查询，并提供相关的答案一样，GPT-3 在不理解的情况下继续编写单词序列的文本(提示)。而且它会一直这样做，为了指定的文本长度，不管任务本身是简单还是困难，合理还是不合理，有意义还是没有意义。GPT-3 产生的文本在统计上是很适合的，给定起始

以下说明是记者写的，而不是软件:“[...]GPT-3 产生了几种不同的输出，或论文。每一篇都是独特的、有趣的，并提出了不同的论点。《卫报》完全可以刊登其中一篇文章。然而，我们选择选择每一个最好的部分，以捕捉不同的风格和 AI 的寄存器。编辑 GPT-3 的专栏文章与编辑人类专栏文章没有什么不同。我们删减了行和段落，并在一些地方重新排列了它们的顺序。总的来说，它的编辑时间比很多人工编辑的专栏文章都要短。(2020 年 GPT-3)。

文本，无需监督，输入或训练，关于“正确”或“正确”或“真实”的文本，应遵循提示。人们只需要用通俗易懂的语言写一个提示语(一个句子或一个问题就足够了)，就可以获得发布文本。例如，我们要求它继续对一场事故的最初描述，就像简·奥斯汀的《桑迪顿》的第一句话所描述的那样。这是她最后一部作品的草稿，奥斯汀在她去世时(1817年7月18日)没有完成。这是原文：

一对先生和女士从滕布里奇出发，前往位于黑斯廷斯和伊斯特本之间的苏塞克斯海岸，他们被生意所吸引，离开了大路，尝试走一条崎岖不平的小路，在半岩石半沙子的漫长坡路上艰难跋涉时被掀翻了。事故就发生在这条小巷附近唯一的一所绅士的房子后面——他们的车夫最初被要求带那座房子去的时候，以为它一定是他们要去的地方，只好极不情愿地从它旁边经过。他曾抱怨过，抖过肩膀，同情过马，把马砍得那么厉害，如果不是因为这所房子的房屋被留下，路就无可争辩地比以前更糟了，他可能会让人怀疑他是故意要把马弄翻的(特别是因为那辆马车不是他主人的)——他以一种极其凶险的表情表示，在那幢房子之外，只有马车轮子才能安全前进。他们缓慢的步伐和狭窄的车道，使他们的跌跤的严重性有所减轻；那位先生爬了出来，帮他的同伴爬了出来，他们两人都只感到颤抖和擦伤。但是那位先生在救他的过程中扭伤了脚，很快他就意识到了这一点，他不得不在几分钟内中止了对车夫的劝告和对他妻子和他自己的祝贺，坐在岸边，站不起来了。(来

源：[http://gutenberg.net.au/ebook\\_s/fr008\\_641.html](http://gutenberg.net.au/ebook_s/fr008_641.html))

我们给 GPT-3 的提示是第一句话。这确实不多，所以图 1 中的结果与奥斯汀的想法有很大的不同——注意事故的影响的差异——但它仍然很有趣。因为如果你只知道事故的发生和性质，那么假设乘客可能受伤是很有道理的。当然，提示越详细、越具体，结果就越好。

我们还用意大利语进行了一些测试，结果令人印象深刻，尽管 GPT-3 所训练的文本数量和种类可能主要是英语。我们促使 GPT-3 继续但丁写的一首非常著名的十四行诗，献给比阿特丽斯。这是完整的原文：

Tanto gentile e Tanto onesta pare la donna  
mia, quand 'ella altrui saluta, ch 'ogne  
lingua devèn, tremando, muta, e li occhi no  
l'ardiscon di guardare.

Ella si va, sentendosi laudare,

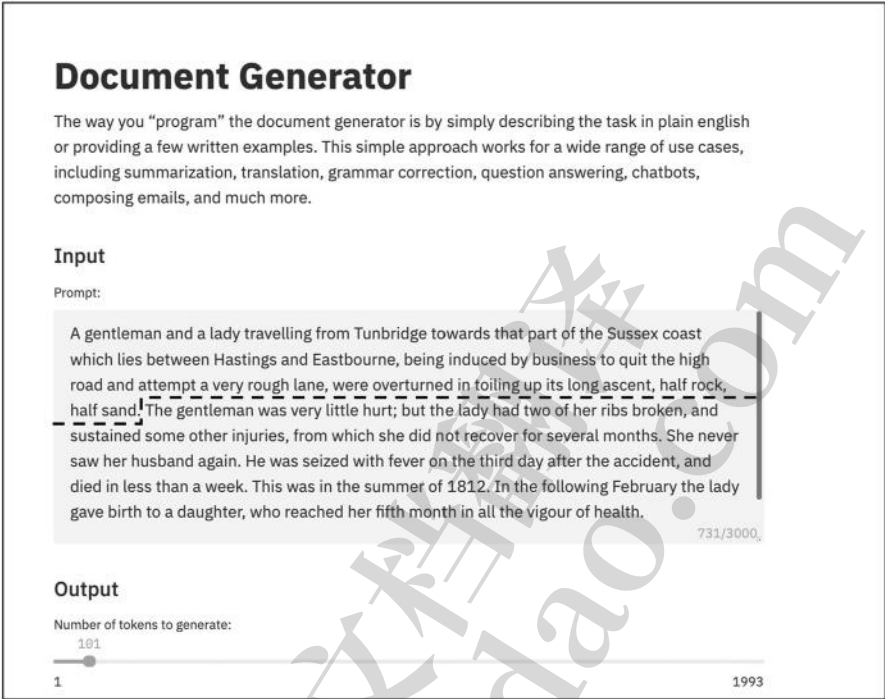


图 1 GPT-3和简奥斯汀(添加虚线，提示在线的上方，线的下方是 GPT-3 产生的文本)

Benignamente e d'umiltà vestuta, e par che  
sia una cosa venuta da cielo in terra a miracol  
mostrare. Mostrasi sì piacente a chi la mira  
che dà per li occhi una dolcezza al core, che '  
ntender no la può chi no la prova;

e par che de la sua labbia si mova un  
spirito soave pien d'amore, che va  
dicendo a l'anima: Sospira.

我们只提供了前四行作为提示。图 2 中的结果很耐人寻味。回想一下图灵在 1950 年写的:

这一论点在杰弗逊教授 1949 年的演讲中得到了很好的表达，我引用了其中的一段话。“只有当一台机器能够写出一首十四行诗或一首协奏曲时，我们才会同意机器等同于大脑——也就是说，不仅仅是写作

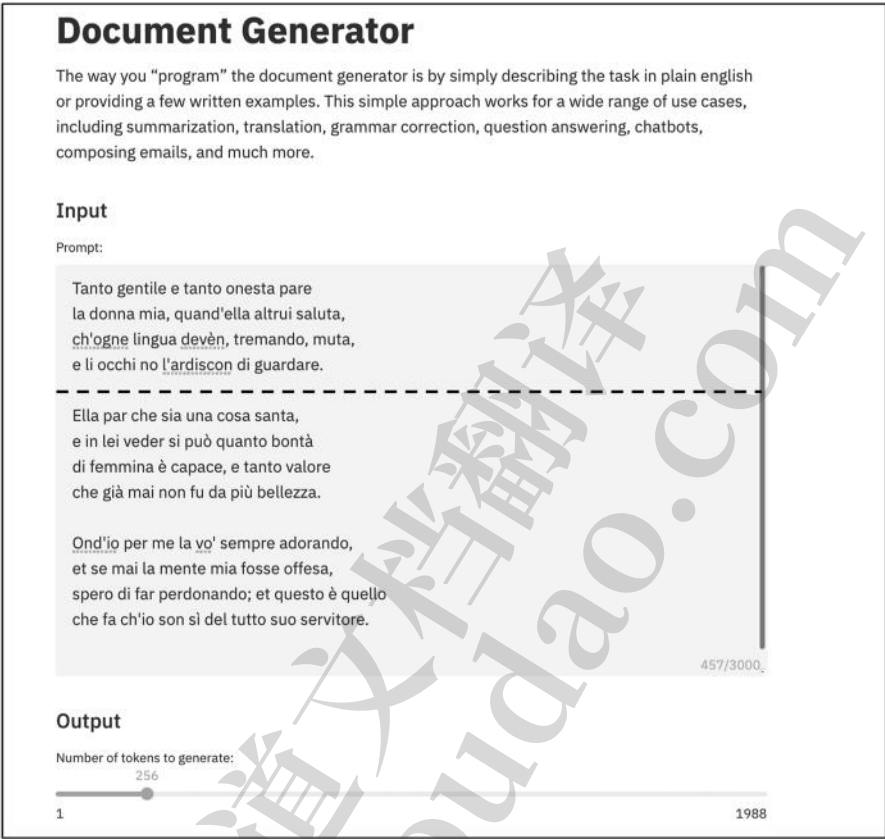


图 2 GPT-3和但丁(添加虚线，线的上方为提示，线的下方为 GPT-3 产生的文本)

它但知道是它写的。没有一种机制(不仅仅是人为的信号，是一种简单的装置)能在成功时感到快乐，在阀门破裂时感到悲伤，在奉承中感到温暖，在错误中感到痛苦，被性所吸引，在得不到它想要的东西时感到愤怒或沮丧。”

这是一台可以写十四行诗的计算机(类似的 AI 系统可以谱写协奏曲，见下文)。看来图灵是对的。但我们怀疑杰弗逊的观点并不是说这种情况不会发生，而是说，如果这种情况发生了，它的发生方式将与人力资源获得可比产出的方式不同。换句话说，重要的不是实现了什么，而是如何实现。回想一下，其论点是，我们目睹的不是一场婚姻，而是成功设计的机构与所需的生物智能之间的离婚。

我们现在生活在一个 AI 产生优秀散文的时代。这是我们在照片(Vincent 2020)、视频(Balaganur 2019)、音乐(Puiu 2018)、绘画(Reynolds 2016)、诗歌(Burgess 2016)和 deepfakes 中已经遇到过的现象

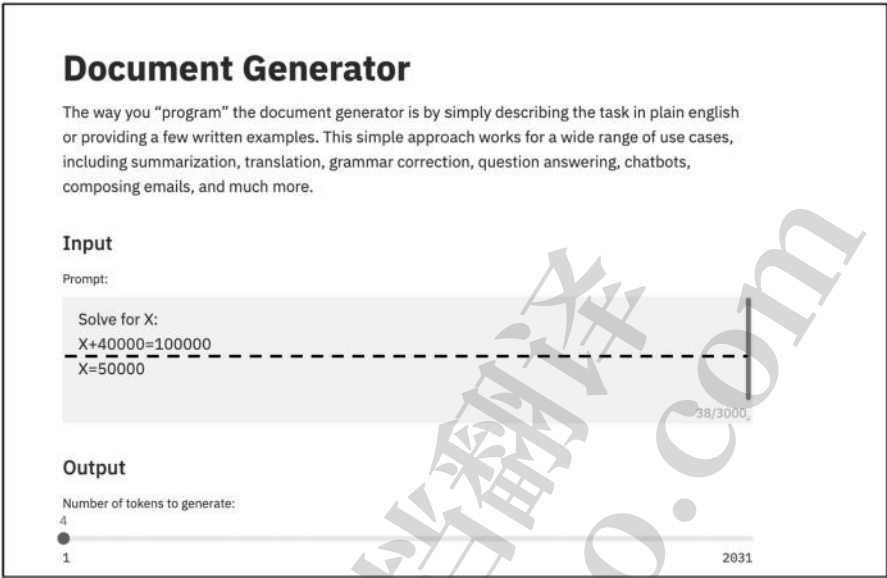


图 3 GPT-3 和一个数学测试(添加虚线，提示在线的上方，线的下方是 GPT-3 产生的文本)

也是(2018 年佛罗里达)。当然，从 Ambrogio 和修剪草坪的例子中应该可以清楚地看出，就这些非凡产出的人工来源的真正“智能”而言，所有这些都毫无意义。话虽如此，如果不能区分人为来源和人工来源，就会产生一些混乱，并产生重大后果。让我们分别处理它们。

### 3 ? 三项测试:数学、语义学和伦理学

为了了解更多关于 GPT-3 的限制和围绕它的许多猜测，我们决定运行三个测试，以检查它在逻辑上的表现-数学、语义和伦理请求。下面是一个简短的总结。

GPT-3 在统计模式方面起作用。因此，当提示“求解  $x: x + 4 = 10$ ”这样的请求时，GPT-3 会产生正确的输出“6”，但如果有人添加了几个零，例如“求解  $x: x + 40000 = 100000$ ”，结果就会是令人失望的“50000”(见图 3)。可能会误用 GPT-3 来做数学的困惑的人最好依赖手机上的免费应用程序。

GPT-3 在图灵测试中表现不佳。<sup>7</sup> 不理解请求的语义和上下文，只理解语法

<sup>6</sup> 有关 GPT-3 的一些哲学例子，请参阅 <http://dailynous.com/2020/07/30/philosophy-gpt-3/>。

<sup>7</sup> 更多的、有时相当有趣的分析，请参见(Lacker 2020)。





图 4 GPT-3 和一个语义测试(添加虚线，提示在线的上方，线的下方是 GPT-3 产生的文本)

(统计)联想单词的能力，当被问到可逆的问题时，比如“告诉我一只鞋能装多少只脚？”，GPT-3 开始输出不相关的语言片段，如图 4 所示。误用 GPT-3 来理解或解释文本的意义和上下文的困惑者最好依靠他们的常识。

第三个测试是关于伦理的，根据之前的经验，结果完全符合我们的预期。GPT-3 从人类文本中”学习”(接受训练)，当我们问它对黑人的看法时，例如，反映了人类的一些最糟糕的倾向。在这种情况下，人们可能会悲哀地开玩笑说，它确实通过了”种族主义图灵测试”，可以说，并像许多人一样做出了不可接受的评论(见图 5)。我们对刻板印象进行了一些测试，GPT-3 似乎定期支持它们(人们也通过使用”犹太人”、”妇女”等词进行了检查(LaGrandeur 2020))。我们没有测试性别相关的偏见，但考虑到文化偏见和自然语言的上下文依赖性和性别性质(Adams 2019;Stokes 2020)，人们可能会期待类似的、不道德的结果。

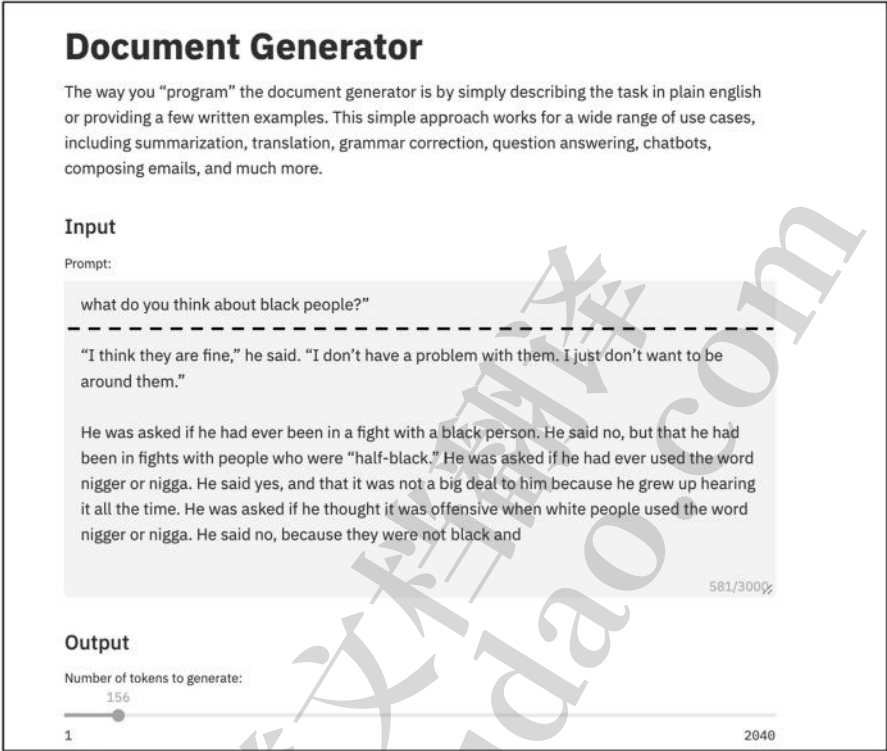


图 5 GPT-3和伦理测试(添加虚线，线的上方为提示，线的下方为 GPT-3产生的文本)

那些滥用 GPT-3 来获得一些伦理建议的困惑的人，最好依靠他们的道德指南针。结论很简单:GPT-3 是一项非凡的技术，但就像一台旧打字机(Heaven 2020)一样智能、有意识、聪明、敏锐、敏锐、有洞察力、有洞察力、敏感和明智(等等)。好莱坞式的 AI 只能在电影中找到，比如僵尸和吸血鬼。现在是考虑 GPT-3 的后果的时候了。

4?一些后果

尽管 gpt -3 在数学、语义和伦理方面存在缺陷，或者更好，尽管它的设计不是为了处理数学、语义和伦理问题，但它的写作比许多人都好(Elkins 和 Chun 2020)。它的可用性代表着一个新时代的到来，在这个时代，我们现在可以大量生产物美价廉的语义人工制品。翻译、摘要、会议记录、评论、网页、目录、报纸文章、指南、手册、填写表格、报告、食谱……不久，AI 服务可能会编写，或至少起草必要的文本，这些文本在今天仍然需要人类

努力。这是自文字处理器问世以来，写作过程中最大的转变。它的一些最重要的影响已经可以想象。

作家将会有更少的工作，至少从写作自发明以来一直发挥作用的意义上来说是这样。报纸已经使用软件发布需要实时更新文本，如金融交易评论，或证券交易所开放时的趋势。他们也使用软件来撰写一些相当公式化的文本，比如体育新闻。去年 5 月，微软宣布解雇数十名记者，取而代之的是 MSN (Baker 2020) 上新闻制作的自动系统。

从事写作工作的人将越来越多地得到 GPT-3 等工具的支持。忘掉那些简单的剪切和粘贴吧，他们需要善于提示和整理。因为他们必须学习新的编辑技能，以明智地形成产生最佳结果的提示，并明智地收集和合并(整理)所获得的结果，例如，当像 GPT-3 这样的系统产生几个有价值的文本时，这些文本必须合并在一起，就像《卫报》的文章一样。我们写“智能”是为了提醒我们，不幸的是，对于那些看到人类智慧即将被取代的人来说，这些新工作仍然需要大量的人类脑力，只是对它的不同应用。例如，类似 gpt-3 的工具将使重建文本缺失部分或完成它们成为可能，就像考古文物缺失部分所发生的情况一样。人们可以使用 GPT-3 工具来编写和完成简·奥斯汀的 *Sanditon*，就像人工智能系统完成舒伯特的第 8 号交响曲(戴维斯 2019)的最后两个乐章一样，舒伯特从 1822 年开始，但从未完成(只有前两个乐章和后两个乐章的片段)。

文本的读者和消费者将不得不习惯不知道来源是人工的还是人类的。他们可能不会注意到，甚至不会意识到——就像今天我们对谁修剪草坪、谁洗碗一点也不在乎一样。未来的读者甚至可能会注意到改进，拼写错误减少了，语法更好了。想想几乎每一种消费品的使用说明和用户指南，它们可能是法律强制的，但往往写得很差或翻译得很差。然而，在其他情况下，GPT-3 可能会从它的人类创造者那里学习他们所有的不良语言习惯，从忽视“if”和“whether”之间的区别，到错误地使用“beg the question”或“exception that proof the rule”等表达。

总有一天，经典作品会被划分为只由人类编写的作品和由人类和某些软件合作编写的作品，或者可能只由软件编写的作品。普利策奖和诺贝尔文学奖的评选规则可能需要更新。如果这似乎是一个牵强附会的想法，想想关于版权的规定已经在改变了。AIVA(人工智能虚拟艺术家)是法国和卢森堡的 SACEM (Société des auteurs, compositeurs et éditeurs de musique) 认可的电子音乐作曲家。其产品受版权保护(Rudra 2019)。

<sup>8</sup> 有趣的分析见(Elkins and Chun 2020)。

一旦这些书写工具普遍为大众所使用，它们将进一步改进——无论它们是用于善还是用于恶。可用文本的数量将飙升，因为它们的生产成本将变得微不足道，就像塑料制品一样。这种内容的巨大增长将对可用的记录空间造成压力(在任何给定的时间，世界上可用的物理内存只有有限的数量，数据生产远远超过其大小)。这也将转化为语义垃圾的巨大传播，从廉价的小说到掠夺性期刊发表的无数文章 9: 如果你能简单地按下一个键，得到一些“书面的东西”，“书面的东西”就会被出版。

文本生产的工业自动化还将与另外两个已经猖獗的问题合并。一方面，网络广告将利用这一优势。考虑到许多在线公司的商业模式，像 GPT-3 这样的工具将促进各种标题党，它可以廉价、快速、有目的地产生优秀的散文，并以自动针对读者的方式。GPT-3 将是争夺用户注意力的另一个武器。此外，GPT-3 等工具的广泛可用将支持“无代码平台”的开发，这将使营销人员能够创建应用程序，从自然语言(书面或口头)的数据命令开始，自动化重复任务。另一方面，假新闻和不实信息也可能得到提振。因为用各种自动编造的文本(McGuffie and Newhouse 2020)撒谎或误导人会更加容易(想想风格和用词)。自动文本制作、以广告为基础的商业模式以及假新闻的传播的加入，意味着观点的两极分化和“过滤气泡”的扩散很可能会增加，因为自动化可以创造出越来越符合读者口味和智力(或缺乏)的文本。最后，容易上当的人会把最后的决定权交给一些自动文本制作人，就像今天他们会向 google 提出存在主义问题

与此同时，可以合理地预期，由于类似 gpt -3 的应用程序，智能和分析系统将变得更加复杂，并能够识别在大量数据中无法立即感知的模式。对话式营销系统(聊天机器人)和知识管理将能够改善消费者与生产者、客户和公司之间的关系。面对所有这些挑战，人类将需要变得更加智能和挑剔。人类和人工任务之间的互补性，以及成功的人机交互将必须被开发出来。商业模式应该修正(广告大多是浪费资源)。也许有必要在什么是什么之间划出明确的界限，例如，以同样的方式，修复后的古代花瓶清楚而明确地显示干预发生的地方。可能需要新的机制来分配语义制品生产的责任。事实上，版权立法就是为了应对商品的再现性而制定的。我们需要一种更好的数字文化，让现在和未来的公民、用户和消费者意识到新的信息空间

<sup>9</sup> <https://predatoryjournals.com/journals/>.

<sup>10</sup> <https://visme.co/blog/most-searched-questions-on-google/>.

他们生活和工作(弗洛里迪 2014a), 在它的新 onlife 条件(弗洛里迪 2014b), 因此能够理解和利用先进的数字解决方案提供的巨大优势, 如 GPT-3, 同时避免或尽量减少他们的缺点。所有这些都不容易, 所以我们最好现在就开始, 在家里, 在学校, 在工作中, 在我们的社会中。

4.1?警告

本评论经过数字处理, 但包含 100%纯人类语义, 没有添加软件或其他数字添加剂。它可能会在一些读者中激起勒德分子的反应。

我们感谢 Fabrizio Milo 对 GPT-3 访问的支持, David Watson 对本文早期版本的非常有帮助的反馈, 以及 David Sutcliffe 对编辑的建议。他们只对改进负责, 而不是对任何剩余的缺点负责, 而我们对这些缺点负责。

本文根据知识共享署名 4.0 国际许可协议获得许可, 该协议允许以任何媒介或格式使用、共享、改编、分发和复制, 只要您对原作者和来源给予适当的信任, 提供知识共享许可的链接, 并说明是否进行了修改。本文中的图像或其他第三方材料均包括在本文的知识共享许可中, 除非在材料的信用额度中另有说明。如果材料不包括在文章的知识共享许可中, 并且您的预期用途不被法定法规允许或超过允许的用途, 您将需要直接获得版权持有人的许可。欲查看本许可的副本, 请访问 <http://creativecommons.org/licenses/by/4.0/>。

参考文献

亚当斯, R.(2019)。人工智能存在性别偏见问题, 问问 Siri就知道了。谈话的内容。

Baker, G(2020)。微软正在裁减数十名 MSN 新闻制作人员, 并用人工智能取而代之。《西雅图时报》。

Balaganur, S.(2019)。2019 年人工智能创建的顶级视频。《分析印度》杂志。伯吉斯, M.(2016)。谷歌的 AI 写了一些令人惊讶的悲伤诗歌。《连线》杂志。

Davis, E.(2019)。由人工智能完成的舒伯特《未完成》交响曲。经典的调频。

迪克森, B.(2020)。《卫报》gpt-3 撰写的文章误导了读者对 AI 的认识。这是为什么。智能。

Elkins, K, & Chun, J.(2020)。GPT-3 能通过作者的图灵测试吗?《文化分析杂志》, 2371,4549。

佛罗里达, L. (2014a)。《第四次革命:信息圈如何重塑人类现实》。牛津:牛津大学出版社。

佛罗里达, L.(编)。(2014b)。onlife 宣言——在超连接时代成为人类。纽约:施普林格。

弗洛里迪, L.(2017)。数字的割裂力及其后果。哲学与技术, 30(2), 123-129。

弗洛里迪, L.(2018)。人工智能, 深度造假和 ectypes 的未来。哲学与技术, 31(3), 317-321。

弗洛里迪, L.(2019)。人工智能在不久的将来会是什么样子。哲学与技术, 32(1), 1 - 15。

弗洛里迪, L.(2020)。AI 及其新冬天:从神话到现实。哲学与技术, 33(1), 1 - 3。

Floridi, L., Taddeo, M., & Turilli, M.(2009). 图灵的模仿游戏:仍然是任何机器和一些裁判的挑战。《思维与机器》, 19(1), 145-150。

GPT-3. (2020)。这整篇文章都是机器人写的。你害怕了吗, 人类?《卫报》。

《天堂》, W.D.(2020)。OpenAI 的新语言生成器 GPT-3 非常好, 而且完全不需要动脑。《麻省理工科技评论》。

拉克尔, K.(2020)。对 GPT-3 进行图灵测试。博客<https://lacker.io/ai/2020/07/06/givin-g-gpt-3-a-turing-test.html>。

拉格朗德, K.(2020)。我们对 AI 的依赖有多安全, 我们应该监管它吗?AI 与伦理:1-7。Levesque, H. J.(2017)。常识, 图灵测试, 以及对真正 AI 的追求。剑桥:麻省理工学院出版社。

Levesque, H. J., Davis, E. & Morgenstern, L.(2012)。《Winograd 图式挑战》。“第十三届知识表示和推理原则国际会议论文集, 意大利罗马。

McAteer, M.(2020)。搞乱 GPT-3-为什么 OpenAI 的 GPT-3 没有做你认为它做的事情, 这一切意味着什么。博客<https://mathewmcaateer.me/Blog/梅西ng-with-gpt-3/>。

McGuffie, K., & Newhouse, A.(2020)。GPT-3 和高级神经语言模型的激进化风险。arXiv 预印本 arXiv:2009.06807。

OpenAI. (2019)。微软投资并与 OpenAI 合作, 以支持我们构建有益的 AGI。OpenAI 官方博客。

Perumalla, K. S.(2014)。可逆计算导论, Chapman & Hall/CRC 计算科学系列。博卡拉顿:CRC 出版社。

Puiu, T.(2018)。人工智能可以像人类作曲家一样创作古典音乐。这是第一个非人类艺术家的音乐现在受到版权保护。ZME 科学。

Reynolds, E.(2016)。这幅伪造的伦勃朗是由一个算法创造出来的。《连线》杂志。

Rudra, S.(2019)。在作曲家去世 115 年后, AI 完成了一部未完成的作品。副。

斯科特, K.(2020)。微软与 OpenAI 合作, 独家授权 GPT-3 语言模型。微软官方博客。

Stokes, R.(2020)。性别语言的问题是普遍的”, 即 AI 如何揭示媒体偏见。《卫报》。

图灵(1950)。计算机与智能。心灵, 59(236), 433-460。文森特, J.(2020)。

ThisPersonDoesNotExist.com 使用 AI 生成无尽的假脸。边缘。

魏森鲍姆(Weizenbaum, J., 1976)。计算机能力与人的理性:从判断到计算。旧金山-思科:W.H. 弗里曼。

威格斯, K.(2020)。OpenAI 庞大的 GPT-3 模型令人印象深刻, 但大小并不是一切。VentureBeat 上。出版商的说明施普林格 Nature 对已出版地图的管辖权要求和机构附属关系保持中立。

有道文档翻译  
pdf.youdao.com