

大语言模型对员工效率的影响

——基于通信数据的实证研究

陆迪 吴祁 张迎峰

目录

附录 I 大语言模型的发展与应用	1
附录 II 稳健性检验	3
附录 III 构造 LLM share 关键词	8
附录 IV 调查问卷及分析	9
参考文献	13

附录 I 大语言模型的发展与应用

（一）大语言模型的发展

语言是人类自幼形成并一生发展的关键交流能力，而机器要具备类人类语言理解与交流能力，需依赖强大的 AI 算法，这是一项长期的研究挑战。大语言模型（LLM）的诞生与发展就是应对这一挑战的具体表现。2022 年底，OpenAI 公司推出了 ChatGPT，开启了大语言模型的突破发展阶段，被普遍认为是生成式 AI 的里程碑。相比于之前的生成式对话产品，ChatGPT 在大范围连续对话能力、生成内容质量、语言理解能力和逻辑推理能力上都得到大幅提升（Zhao 等，2023）。在此之后，LLM 呈现爆发式增长，各大公司和研究机构都在发布各种不同类型的大语言模型。图 I 1 展示 GPT-3.5 发布以来至 2023 年底国内外主要大模型发布时间表，其中国内的大语言模型发布晚于国外，在 2023 年 6 月呈现出“井喷式”增长。2023 年 8 月 31 日，国内 11 家通过《生成式人工智能服务管理暂行办法》备案的 AI 大模型产品陆续上线，面向全社会开放。目前，国外大模型除 GPT 系列外，还有 Claude、Gemini、Llama 等；国内则有“文心一言”“盘古”“混元”“通义千问”等。

LLM 正在改变我们与信息交互的方式，甚至在转变我们的工作方式。世界经济论坛在 2023 年发布的《未来工作：大语言模型和工作》白皮书¹中指出：“由于所有工作任务的 62% 都涉及语言，大语言模型可能会对各种职业都产生重大影响。大语言模型本身功能就很丰富，加之其使用门槛很低，导致其极快被市场采用和接受。这些都表明，未来许多工作任务及与其高度相关的职业都可能受到大语言模型的影响。”目前，这些技术可以做内容的原创，从大量数据中产生见解，以接近人类的准确性翻译语言，甚至可能做出复杂的决策。因此，大语言模型的应用可能会导致劳动力生产率显著地提高，并创造新类型的就业机会。



图 I 1 国内外主要大语言模型时间表

¹ ‘Jobs of Tomorrow: Large Language Models and Jobs’, the World Economic Forum, 2023 年 9 月 18 日, 来源: <https://www.weforum.org/publications/jobs-of-tomorrow-large-language-models-and-jobs/>.

（二）国产大语言模型应用：文心一言

作为国产大语言模型的代表，百度文心一言于 2023 年 3 月 16 日举行新闻发布会并邀请第一批内测，并于 2023 年 8 月 31 日率先向全社会开放，其用户累计数量于 2023 年 12 月超过 1 亿。8 月 31 日 0 点，文心一言宣布率先向全社会全面开放。用户可以在应用商店下载“文心一言 APP”或登录“文心一言官网”体验。之后，其他几家通过《生成式人工智能服务管理暂行办法》备案的 AI 大模型产品也陆续上线。百度文心 API 大模型日均调用量达千万次，位列中国第一，每月服务企业超过 1 万家¹。文心一言当前主要的功能包括以下五项：文学创作、商业文案创作、数理逻辑推算、中文理解与多模态生成²。2023 年 9 月 22 日，文心一言发布了首份《文心一言用户使用报告》，其中显示超过 70% 的文心一言用户为上班族。从行业分布上来看，IT 互联网、教育、社会公共管理、医药卫生、能源采矿化工、金融保险、日化百货等七大行业占比超五成³。

除了文心一言，国内其他主要的大模型包括华为发布的“盘古”、腾讯发布的“混元”、阿里巴巴发布的“通义千问”等。这些模型的推出时间集中于 2023 年，功能侧重和具体表现不尽相同。考虑到样本期间其他国产大模型的推出或公开的影响，本文探究以“文心一言”为首的国产大模型“井喷”的冲击下，中国国产通用大模型对企业员工工作模式与运营成本的影响。

对于 ChatGPT 等海外大模型，尽管其技术先进且功能强大，但在中国市场，由于以下原因，其普及和使用受到了一定限制：首先，GPT 模型早期在中国的知名度相对较低，许多潜在用户并不了解其功能和应用场景。即使在美国，GPT 模型刚推出时，作为一种新技术，了解和使用的人也寥寥无几。The verge 2023 年 6 月的美国调查结果显示，高达 43% 的被调查者甚至没听说过 ChatGPT⁴。其次，由于网络环境的限制，中国用户访问 GPT 等国际模型可能需要借助 VPN 等工具，高级功能往往需要付费订阅和海外账户，这无疑增加了使用门槛。相较之下，“文心一言”等作为国内首批推出的大模型，提供了免费版本，用户可以在没有经济负担的情况下体验大模型带来的便利。

¹ 来源：http://www.199it.com/archives/1668678.html#google_vignette。

² 来源：<https://github.com/szwnba/affweb/issues/310>。

³ 来源：<https://www.163.com/dy/article/IF8UVC8P05199KH8.html>。

⁴ 来源：<https://www.theverge.com/c/23753704/ai-chatgpt-data-survey-research>。

附录 II 稳健性检验

1. 进一步限制样本

首先，参考姚加权等（2024），我们通过三种方式限制样本并进行回归分析：（1）剔除金融行业公司；（2）剔除信息传输、软件和信息技术服务业以及科学研究和技术服务行业；（3）剔除当年处于 ST 和 *ST 状态的上市公司。结果展示在表 II 1 第（1）—（4）列。此外，我们还将样本中日加班率高于分布 99% 分位数的极端值去除（99% 分位数为 0.5）后回归，结果汇报在表 II 1 第（5）—（8）列。最后，为了进一步减轻数据上可能存在的偏误¹，我们根据上市公司办公地址，排除办公地址仅为部分办公楼层的企业²，最终保留 636 家上市公司中的 357 家进行稳健性检验。回归结果如表 II 1 第（9）—（12）列所示。以上稳健性检验结果均与主回归结果保持一致。

2. 拓展到非上市公司

为探究研究结果的普适性，我们对广东省非上市公司进行了千分之一随机抽样，并对抽样的非上市公司样本进行了通信大数据的收集以及日度加班率的计算。基于补充数据，我们增加了对表 3 基准回归的稳健性检验，结果见表 II 1 第（13）—（16）列。结果显示，拓展样本到非上市公司，基准回归结果仍然是稳健的。

3. 控制行业-月固定效应

在样本时间区间内，行业之间的趋势差别可能受到众多因素影响，这可能会对模型估计结果产生干扰。例如，大模型冲击程度较高的行业（如互联网等），在这段时间还可能受到与大模型无关的其他行业或政策因素的影响，包括但不限于数字化转型和远程办公带来的变化。我们引入“行业-月”固定效应以控制每个行业在不同月份所面临的特定影响。我们将基准回归中的行业层面的 LLM 冲击指标替换为企业层面受 LLM 冲击指标（LLM share），并在回归中加入“行业-月”固定效应，结果汇报在表 II 2。我们发现，加入“行业-月”固定效应后，LLM 冲击的估计结果仍然是稳健的³。

3. 加班率替代指标

主结果中对加班的定义可能会把一些上班时间晚或者从事夜班工作的员工职业误判为加班，即使这些员工 24 小时内的总工作时长并没有超过 8 小时，或者样本中缺失在家工作的员工⁴。为了进一步验证研究结果的稳健性，我们计算并引入了加班率的替代指标，“日度工作时长超过 8 小时的员工占比”，作为日度加班率的补充测度⁵。我们将基准回归中的因变量替换为替代指标，稳健性检验结果（见表 II 3）与基准回归结果一致。

¹ 通信大数据的特性带来了一定客观存在的局限性。例如，在中央商务区等高楼密集区域，一个地址可能同时对应到多家企业。这种情况下，不同企业的员工可能被归为同一地址某个企业的员工，导致度量误差并可能造成估计偏误。

² 比如一品红药业股份有限公司（股票代码：300723），其办公地址“广东省广州市黄埔区广州国际生物岛寰宇一路 27 号云润大厦 15-19 层”为大楼的部分楼层。

³ 由于招聘数据覆盖上市企业数量有限，表 II 2 中样本量相对于正文表 3 有所变动。

⁴ 在我们的广义双重差分（DID）设定中，除非员工在大模型冲击前未在家工作，而在冲击后改变模式选择在家工作，这一问题不太可能对估计结果造成显著的估计偏误。此外，考虑到我们关注的是 LLM 冲击在短期内的效应，样本时期较短，且模型中已经控制了月份固定效应，这种偏差的影响应当是可控的。

⁵ 我们发现原始构造的加班工作员工数量和补充构造的超过 8 小时工作员工数量呈正相关（相关系数为 0.926），原始构造的日度加班率和替代指标也呈正相关的关系（相关系数为 0.016），这表明两种指标在度量上的一致性。

表 II 1 稳健性检验 1

Panel A 因变量：日度加班率（overtime_ratio）	剔除特定行业及 ST 和*ST 公司				去除日加班率最高 1% 极端值			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post×exposure_gpt4	-0.0274** (0.0114)				-0.0193** (0.0083)			
Post×exposure_hl		-0.0310** (0.0125)				-0.0233** (0.0096)		
Post×high_exposure_gpt4			-0.0060** (0.0025)				-0.0054** (0.0022)	
Post×high_exposure_hl				-0.0059** (0.0025)				-0.0056** (0.0022)
within_2_weeks	0.0050** (0.0022)	0.0050** (0.0022)	0.0050** (0.0022)	0.0050** (0.0022)	0.0076*** (0.0015)	0.0076*** (0.0015)	0.0076*** (0.0015)	0.0076*** (0.0015)
Constant	0.1290*** (0.0021)	0.1290*** (0.0019)	0.1260*** (0.0006)	0.1260*** (0.0006)	0.1200*** (0.0016)	0.1199*** (0.0015)	0.1176*** (0.0006)	0.1177*** (0.0006)
Observations	118,430	118,430	118,430	118,430	131,640	131,640	131,640	131,640
R-squared	0.081	0.081	0.081	0.081	0.159	0.159	0.159	0.159
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel B 因变量：日度加班率（overtime_ratio）	去除非独栋办公楼的公司				非上市公司			
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Post×exposure_gpt4	-0.0234* (0.0122)				-0.0947* (0.0515)			
Post×exposure_hl		-0.0260*				-0.0932*		

		(0.0140)			(0.0554)			
Post×high_exposure_gpt4			-0.0064**				-0.0177*	
			(0.0031)				(0.0097)	
Post×high_exposure_hl				-0.0065**				-0.0194**
				(0.0031)				(0.0097)
within_2_weeks	0.0072***	0.0072***	0.0072***	0.0072***				
	(0.0021)	(0.0021)	(0.0021)	(0.0021)				
Constant	0.116***	0.116***	0.113***	0.113***	0.3894***	0.3870***	0.3764***	0.3769***
	(0.0023)	(0.0022)	(0.0008)	(0.0008)	(0.0087)	(0.0082)	(0.0017)	(0.0018)
Observations	69,727	69,727	69,727	69,727	88,299	88,299	88,299	88,299
R-squared	0.223	0.223	0.223	0.223	0.163	0.163	0.163	0.163
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

注：本表使用正文表 3 模型，（1）—（4）列回归样本中剔除了金融行业公司，信息传输、软件和信息技术服务业以及科学研究和技术服务行业，以及当年处于 ST 和*ST 状态的上市公司。（5）—（8）列回归样本中剔除了日加班率高于分布 99%分位数（0.5）的极端值。（9）—（12）列回归样本为广东省上市公司去除部分办公地址为非独栋楼宇，保留了 636 家中的 357 家。（13）—（16）列回归样本为广东省内随机抽样的非上市公司，共包含 928 家。***、**、*分别表示在 1%、5%、10%水平下显著，括号内为公司层面的聚类标准误。

表 II 2 稳健性检验 2：控制行业-月固定效应

	(1)	(2)
因变量：日度加班率 (overtime_ratio)		
Post×LLM_share	-0.025* (0.015)	
LLM_share	0.014 (0.010)	
Post ×high LLM_share		-0.013** (0.006)
high LLM_share		0.003 (0.004)
within_2_weeks	0.014*** (0.004)	0.014*** (0.004)
Constant	0.122*** (0.0004)	0.123*** (0.001)
Observations	36,475	36,475
R-squared	0.105	0.106
Firm FE	Yes	Yes
Industry*Month FE	Yes	Yes

注：本表使用正文表 3 模型，“LLM_share”为基于企业招聘文本数据构建的企业层面的大模型暴露度指标。***、**、* 分别表示在 1%、5%、10%水平下显著，括号内为公司层面的聚类标准误。

表 II 3 稳健性检验 3：加班率替代指标

	(1)	(2)	(3)	(4)
因变量：日度工作时长超过 8 小时的员工占比				
Post×exposure_gpt4	-0.0542* (0.0287)			
Post×exposure_hl		-0.0568* (0.0323)		
Post×high_exposure_gpt4			-0.0138* (0.0073)	
Post×high_exposure_hl				-0.0147** (0.0073)
within_2_weeks	0.0323*** (0.0038)	0.0324*** (0.0038)	0.0323*** (0.0038)	0.0323*** (0.0038)
Constant	0.5428*** (0.0056)	0.5414*** (0.0052)	0.5359*** (0.0019)	0.5363*** (0.0020)
Observations	131,982	131,982	131,982	131,982
R-squared	0.295	0.295	0.295	0.295
Firm FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes

注：本表使用正文表 3 模型，将因变量替换为公司“日度工作时长超过 8 小时的员工占比”，即当天工作超过 8 小时员工数除以总员工数。***、**、* 分别表示在 1%、5%、10% 水平下显著，括号内为公司层面的聚类标准误。

附录 III 构造 LLM share 关键词

参考 Alekseeva 等（2021），并结合招聘文本数据，我们构造 AI 以及大模型相关技能词典。词典中的词汇为与 AI 和大模型知识或使用相关软件直接相关的词汇，例如“人工智能”“深度学习”“大数据”“语音识别”等。采用的关键词具体如下表。

表 III1 AI 及大模型相关技能关键词

人工智能（AI）	机器学习	深度学习	神经网络	强化学习
卷积神经网络	循环神经网络	长短期记忆（LSTM）	支持向量机（SVM）	梯度提升机（GBM）
自然语言处理	语音识别	语音合成	机器翻译	问答系统
人机交互	人机对话	语音交互	智能语音	聊天机器人（Chatbot）
对话系统	TensorFlow	Keras	PyTorch	算法设计
模型优化	决策树	随机森林	OpenCV	图像算法
颜色算法	计算机视觉	图像识别	特征识别	特征提取
模式识别	知识图谱	知识表示	深度神经网络	智能芯片
AI 芯片	智能传感器	边缘计算	分布式计算	物联网
IoT	增强现实（AR）	虚拟现实（VR）	智能硬件	嵌入式系统
可穿戴产品	智能家居	智能音箱	智能农业	智能教育
智能医疗	智能保险	智能投顾	智能养老	智能环保
智能零售	智能政务	智能监管	智能电网	智能运输
自动驾驶	无人驾驶	车辆控制	增强智能	智能计算
智算	智能搜索	智能客服	智能体	机器人流程自动化
商业智能	大数据营销	大数据风控	大数据分析	大数据处理
大数据管理	大数据平台	大数据运营	SQL	Hadoop
Spark	数据仓库	数据可视化	数据挖掘	文本分析
情感分析	推荐系统	内容推荐	用户行为分析	客户数据分析
客户关系管理	CRM	供应链优化	物流优化	库存管理
需求预测	SIEM	气象算法	雷达算法	算法创新
仿真	SLAM			

注：作者基于相关文献和招聘数据文本整理所得。

附录 IV 调查问卷及分析

(1) 调查背景与样本特征。2024 年 9 月 3 日至 2024 年 9 月 22 日，我们对三家位于深圳的上市公司总部（分部）的 61 位员工开展了问卷调查。这三家公司分别为：(a) 大晟时代文化投资股份有限公司总部（股票代码 600892，主营影视投资与制作、游戏开发等业务），(b) 中国工商银行深圳某支行（股票代码：601398），以及 (c) 深圳市瑞凌实业股份有限公司总部（股票代码：300154，主营逆变焊割设备供应）。按照前文构建的行业受 LLM 冲击指数，前两家公司所在行业受 LLM 冲击程度较高，而第三家公司所在行业受 LLM 冲击程度较低。

被调查员工的年龄主要集中在 23-35 岁（占 47.54%）和 36-45 岁（占 47.54%），性别分布中男性占 34.43%，女性占 65.57%。调查对象覆盖了企业不同岗位，包括管理类、技术类、营销与销售类、财务类、人力资源类、行政类与客户服务类岗位的员工，具体岗位分布参见图IV1 以及表IV3。本次调查旨在从多岗位视角了解大语言模型的实际应用情况及其对公司劳动生产率和运行成本的影响。

(2) 大语言模型使用情况。调查数据显示，超过半数的员工（50.82%）在工作中使用过大语言模型，且存在明显的结构性差异：(a) 性别差异。其中，男性员工的使用比例为 41.95%，低于总体水平；(b) 受 LLM 冲击程度不同的企业间存在差异。属于冲击程度较高行业的两家公司被调查员工的大模型使用比率为 54.76%，而属于冲击程度较低行业的一家公司被调查员工的大模型使用比率为 38.89%。这一差异交叉验证了我们构建的 LLM 冲击指标的区分度和可信度。值得注意的是，在未使用员工中，63.33%的受访者表示其工作岗位需求与大模型功能不匹配，36.67%则是因为对大模型认知不足，这一发现反映了大语言模型在部分岗位的适用性和认知普及程度的局限。

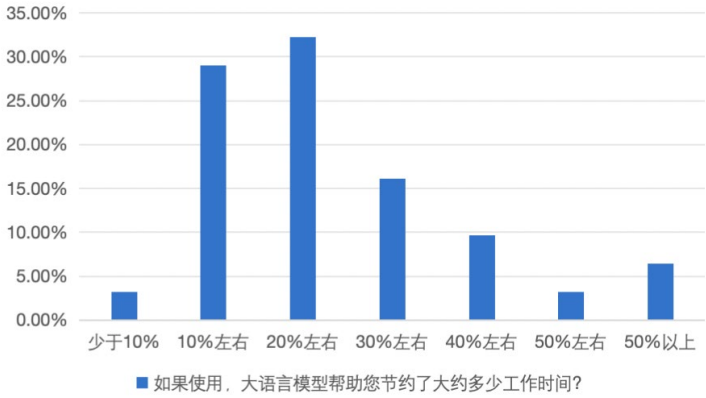
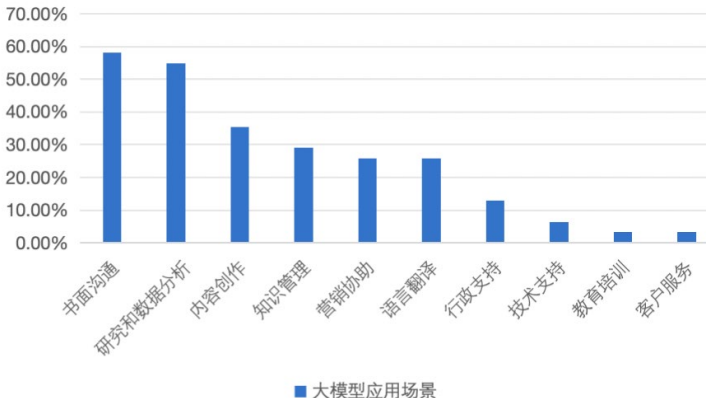
在使用大语言模型的员工中，百度推出的文心一言是最常使用的模型，其次是字节豆包、KIMI 和 ChatGPT。然而，调查中 100%的受访员工均仅使用免费版本的大语言模型。此外，截至调查时，这些企业尚未为员工购买任何付费大模型。这表明，大语言模型的企业级应用仍处于初期阶段，员工自发使用免费版本的行为在短期内走在了企业技术采用的前面，这与以往企业主导的 AI 技术应用模式形成鲜明对比。值得注意的是，大部分员工是在 2023 年 8 月后才开始使用大语言模型，这从侧面表明 ChatGPT 等国际产品在国内使用存在一定的门槛，得到国内用户的了解和认可需要时间。

(3) 大模型对工作效率的影响。在使用大模型的被调查员工中，61%的员工认为大模型帮助节约了 10%~20%的工作时间。这些模型在多个工作场景中协助员工完成任务，并提高了工作效率。具体而言，大语言模型的主要应用场景包括（选项可多选）：

- (a) 书面沟通：撰写和完善电子邮件、信件和报告（58.06%）；
- (b) 研究和数据分析：进行信息搜索、事实核查和分析数据（54.84%）；
- (c) 内容创作：起草各种内容，例如博客文章、社交媒体帖子、时事通讯和广告内容（35.48%）。
- (d) 知识管理：构建企业知识库，支持员工快速检索信息和学习新知识（29.03%）。
- (e) 营销协助：制定营销策略、编写广告模板、捕捉潜在客户（25.81%）。

表IV1 大模型使用调查问卷主要结果

调查时间	2024 年 9 月 3 日—9 月 22 日
调查对象	三家深圳上市公司总部/支行员工共 61 名
年龄分布	23-35 岁（47.54%）、36-45 岁（47.54%）
性别分布	男性 34.43%、女性 65.57%
岗位覆盖	管理类、技术类、营销与销售类、财务类、人力资源类、行政类、客户服务类
大模型使用比例	50.82%， - 其中男性使用比例 41.95%； - 属于冲击程度较高行业的两家公司使用比例 54.76%，属于冲击程度较低行业的一家公司使用比例 38.89%

未使用大模型的原因	工作中用不到（63.33%）、缺乏了解（36.67%）																						
第一次使用的大模型	国内大模型：文心一言（45.16%）、讯飞星火（9.68%）、KIMI（3.23%）、通义千问（3.23%） 国外大模型：ChatGPT（38.71%）																						
使用的其他大模型（多选）	文心一言（48.39）、字节豆包（32.26%）、KIMI（32.26%）、ChatGPT(19.35%)																						
使用大模型版本	100%受访者使用免费版																						
企业购买大模型情况	未购买付费版																						
首次使用大模型时间	2023 年 8 月后开始使用占 64.52%																						
如果使用，大语言模型帮助您节约了大约多少工作时间？	 <table border="1"><caption>如果使用，大语言模型帮助您节约了大约多少工作时间？</caption><thead><tr><th>节约时间范围</th><th>占比</th></tr></thead><tbody><tr><td>少于10%</td><td>3.00%</td></tr><tr><td>10%左右</td><td>29.00%</td></tr><tr><td>20%左右</td><td>32.00%</td></tr><tr><td>30%左右</td><td>16.00%</td></tr><tr><td>40%左右</td><td>10.00%</td></tr><tr><td>50%左右</td><td>3.00%</td></tr><tr><td>50%以上</td><td>6.00%</td></tr></tbody></table>	节约时间范围	占比	少于10%	3.00%	10%左右	29.00%	20%左右	32.00%	30%左右	16.00%	40%左右	10.00%	50%左右	3.00%	50%以上	6.00%						
节约时间范围	占比																						
少于10%	3.00%																						
10%左右	29.00%																						
20%左右	32.00%																						
30%左右	16.00%																						
40%左右	10.00%																						
50%左右	3.00%																						
50%以上	6.00%																						
大模型应用场景（多选）	 <table border="1"><caption>大模型应用场景</caption><thead><tr><th>应用场景</th><th>占比</th></tr></thead><tbody><tr><td>书面沟通</td><td>58.00%</td></tr><tr><td>研究和数据分析</td><td>55.00%</td></tr><tr><td>内容创作</td><td>35.00%</td></tr><tr><td>知识管理</td><td>30.00%</td></tr><tr><td>营销协助</td><td>25.00%</td></tr><tr><td>语言翻译</td><td>25.00%</td></tr><tr><td>行政支持</td><td>12.00%</td></tr><tr><td>技术支持</td><td>5.00%</td></tr><tr><td>教育培训</td><td>3.00%</td></tr><tr><td>客户服务</td><td>3.00%</td></tr></tbody></table>	应用场景	占比	书面沟通	58.00%	研究和数据分析	55.00%	内容创作	35.00%	知识管理	30.00%	营销协助	25.00%	语言翻译	25.00%	行政支持	12.00%	技术支持	5.00%	教育培训	3.00%	客户服务	3.00%
应用场景	占比																						
书面沟通	58.00%																						
研究和数据分析	55.00%																						
内容创作	35.00%																						
知识管理	30.00%																						
营销协助	25.00%																						
语言翻译	25.00%																						
行政支持	12.00%																						
技术支持	5.00%																						
教育培训	3.00%																						
客户服务	3.00%																						

注：本表为作者基于三家位于深圳的上市公司总部（分部）的 61 位员工开展问卷调查结果整理而成。

表IV2 调查问卷

- *1. 您的年龄（单选）：

☐16-22 岁

☐23-35 岁

☐36-45 岁

☐46-60 岁
- *2. 您的性别（单选）：

☐男

☐女
- *3. 您工作的主要内容（可多选）：

☐管理类：涉及企业战略规划、决策制定、团队领导和监督等。

☐技术类：包括研发、工程设计、系统开发、数据分析等需要专业技能的工作。

☐营销与销售类：涉及市场调研、产品推广、销售策略制定和客户关系管理等。

- ☐ 财务类：包括会计、审计、财务分析、预算控制和投资评估等。
- ☐ 人力资源类：涉及招聘、培训、员工关系和薪酬福利管理等。
- ☐ 行政类：包括日常办公管理、文档处理、会议组织和物资采购等。
- ☐ 客户服务类：提供客户咨询、售后服务和客户反馈处理等。
- ☐ 运营类：涉及生产管理、供应链协调、物流规划和库存控制等。

*4. 您在工作中是否使用大语言模型？（单选，回答“否”直接跳到 13 题）

- ☐ 是 ☐ 否

*5. 如果使用，您第一次使用的是哪个大语言模型？（单选）

- ☐ ChatGPT ☐ Claude ☐ 文心一言 ☐ 字节豆包
- ☐ KIMI（月之暗面） ☐ 通义千问 ☐ 百川
- ☐ 讯飞星火 ☐ 天工 ☐ 智谱清言
- ☐ 其他 _____ （请在空白处填写内容）

*6. 您是从什么时候开始第一次使用大语言模型的？（单选）

- ☐ 2023 年 8 月之前 ☐ 2023 年 9 月 ☐ 2023 年 10 月
- ☐ 2023 年 11 月 ☐ 2023 年 12 月 ☐ 2024 年 1 月
- ☐ 2024 年 2 月 ☐ 2024 年 3 月 ☐ 2024 年 4 月
- ☐ 2024 年 5 月 ☐ 2024 年 6 月 ☐ 2024 年 7 月
- ☐ 2024 年 8 月（本月）

*7. 您还使用什么其他的大语言模型吗？（可多选）

- ☐ 无 ☐ ChatGPT ☐ Claude ☐ 文心一言
- ☐ 字节豆包 ☐ KIMI（月之暗面） ☐ 通义千问 ☐ 百川
- ☐ 讯飞星火 ☐ 天工 ☐ 智谱清言
- ☐ 其他 _____ （请在空白处填写内容）

*8. 如果使用大语言模型，您用的是免费账户还是付费账户？（单选）

- ☐ 免费账户 ☐ 付费账户（只要有一个付费账户就请选该选项）

*9. 如果您使用付费账户，是个人付费账户还是公司付费账户？（单选）

- ☐ 个人付费账户 ☐ 公司付费账户

*10. 如果使用，您在工作中如何使用大语言模型？（可多选）

- ☐ 书面沟通：撰写和完善电子邮件、信件和报告。
- ☐ 研究和数据分析：进行信息搜索、事实核查和分析数据。
- ☐ 技术支持：协助编码、调试和软件配置。
- ☐ 营销协助：制定营销策略、编写广告模板、捕捉潜在客户。
- ☐ 知识管理：构建企业知识库，支持员工快速检索信息和学习新知识。
- ☐ 内容创作：起草各种内容，例如博客文章、社交媒体帖子、时事通讯和广告内容。
- ☐ 行政支持：创建项目大纲、起草会议记录、促进项目管理、时间管理和自动化任务。
- ☐ 教育和培训：协助创建教育内容和培训。
- ☐ 客户服务：自动化客户咨询响应，提供常见问题解答，提高服务效率。
- ☐ 语言翻译：实时翻译多语言内容，帮助企业拓展国际市场。

*11. 如果使用，大语言模型帮助您节约了大约多少工作时间？（单选）

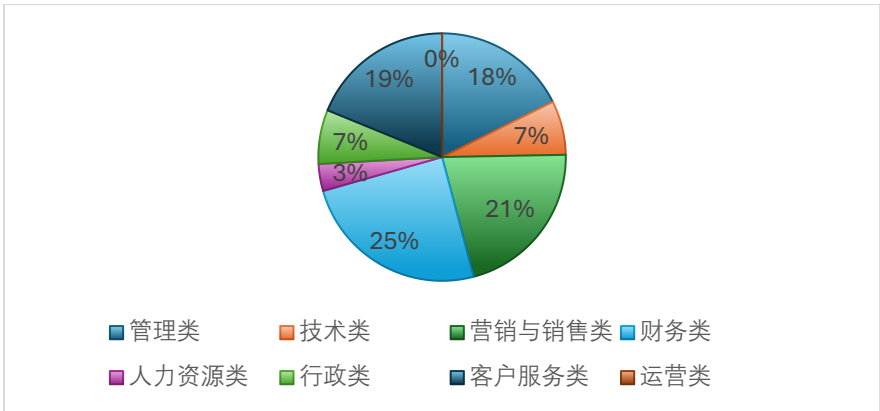
- ☐ 少于 10% ☐ 10%左右 ☐ 20%左右 ☐ 30%左右

☐40%左右 ☐50%左右 ☐50%以上

*12.从第一次使用开始，您大概用了多久时间才开始稳定使用大模型辅助工作？（可多选）
☐1~2 天 ☐1 个星期或更久 ☐用了几天觉得不好用，就停止继续使用了。

*13.您至今没有使用大语言模型的主要原因是？（可多选）
☐没有听说过
☐觉得工作中用不到
☐觉得很难使用，不想为此额外花时间学习
☐觉得不靠谱，怕使用会给工作带来纰漏
☐其他_____（请在空白处填写内容）

*14.如果没有使用，您后面计划使用哪一款大语言模型？（没有计划请填写“无”）
_____（请在空白处填写内容）



图IV1 被调查人员岗位分布

注：调查问卷中岗位可多选，即一人可能符合多个岗位分类。各岗位的主要工作内容如下：“管理类”涉及企业战略规划、决策制定、团队领导和监督等；“技术类”包括研发、工程设计、系统开发、数据分析等需要专业技能的工作；“营销与销售类”涉及市场调研、产品推广、销售策略制定和客户关系管理等；“财务类”包括会计、审计、财务分析、预算控制和投资评估等；“人力资源类”涉及招聘、培训、员工关系和薪酬福利管理等；“行政类”包括日常办公管理、文档处理、会议组织和物资采购等；“客户服务类”提供客户咨询、售后服务和客户反馈处理等；“运营类”涉及生产管理、供应链协调、物流规划和库存控制等。此问卷调查受访者中“运营类”岗位员工人数为0。

表IV3 被调查人员使用大模型情况：分企业—岗位类型统计

岗位	岗位人数（一人可能符合不止一个岗位分类）	使用大模型人数	使用大模型人数占比
高冲击行业企业			
管理	12	8	67%
技术	2	1	50%
营销	15	14	93%
财务	11	4	36%
人力	3	2	67%
行政	5	3	60%
客服	15	6	40%

高冲击行业企业总计	63	38	60%
低冲击行业企业			
管理	3	0	0%
技术	3	3	100%
营销	2	0	0%
财务	10	5	50%
人力	0	0	
行政	1	0	0%
客服	1	0	0%
低冲击行业企业总计	20	8	40%

注：按照文中构建的行业受 LLM 冲击指数，两家公司所在行业受 LLM 冲击程度较高，一家公司所在行业受 LLM 冲击程度较低。

参考文献

[1] Alekseeva, L., J. Azar, M. Giné, S. Samila, and B. Taska, “The Demand for AI Skills in the Labor Market” , *Labour Economics*, 2021, 71, 102002.

注：该附录是期刊所发表论文的组成部分，同样视为作者公开发表的内容。如研究中使用该附录中的内容，请务必在研究成果上注明附录下载出处。