中国“人工智能-大语言模型技术”暴露指数

构建方法与数据说明

北京大学国家发展研究院 AI 与经济学实验室、智联招聘

发布日期：2025.09.19

本技术文档系统介绍了研究团队在构建“人工智能-大语言模型技术”（**AI-LLM）暴露指数**过程中所采用的方法与数据基础。随着大语言模型技术的快速发展和相关数据的持续更新，与《管理世界》已刊论文[1]（张丹丹等，2025）相比，本次公开发布的指数基于一套更加透明、严谨且适配性强的新测算方案。本方案在延续张丹丹等（2025）核心思路的基础上，对任务提取与匹配等关键环节进行了细化和优化，显著增强了指数构建过程的可追踪性与可复现性。同时，当前版本在处理招聘平台数据中职业分类标准差异和跨期不一致性问题方面也做出了改进，进一步提升了指数在时间序列上的稳健性。通过这些系统性优化，研究团队为后续基于本指数的学术研究与政策分析提供了更为可靠、通用的工具。该新构造的方法在于航、张丹丹、张润博和李泓宇（2025）的论文中首次被详细阐述，并已应用于该文的实证分析部分。

一、指数构建方法

借鉴已有研究的成熟做法，特别是 Eloundou 等（2024）的指数构建方法，本指数基于 O*NET [2] 数据库对职业工作内容的分解与聚类信息，结合中国招聘平台大数据，构建了 AI-LLM 暴露指数。具体测算过程中，采用了 O*NET 中两个互补的细化指标——详细工作活动（Detailed Work Activity, DWA）和任务（Task）。图 1 展示了 O*NET 数据库中不同层级指标的含义与结构。以“电工”职业为例，其工作内容可拆分为若干具体“任务”，例如：“准备草图或遵循蓝图，确定电线或设备的位置，以确保符合建筑和安全规范”，以及“提供材料或服务的初步草图或成本估算”等。这些“任务”由 O*NET 通过实地调查、问卷与访谈方式收集整理，总 19,265 项。

O*NET 会将内容相似的任务聚类为 2,087 项详细工作活动（DWA）。例如，前述两项任务均可归类到同一个 DWA——“绘制建筑或安装图纸”。在更高层级，O*NET 还设有约 300 项中间工作活动（Intermediate Work Activity, IWA）和约 41 项广义工作活动（General Work Activity, GWA）等指标。相比之下，DWA 和 Task 层级的指标粒度更细、信息量更丰富，能够更有效地区分不同职业的任务结构差异。因此，AI-LLM 暴露指数构建将以 DWA 和 Task 层级为基础。

[1] 作者包括：张丹丹、于航、李力行、胡佳胤、莫怡青、李泓宇。

[2] O*NET （职业信息网络， Occupational Information Network）由美国劳工部开发和维护，它涵盖了大约 1,000 种职业对应的 2,087 种工作任务和 19,265 种详细任务要求。

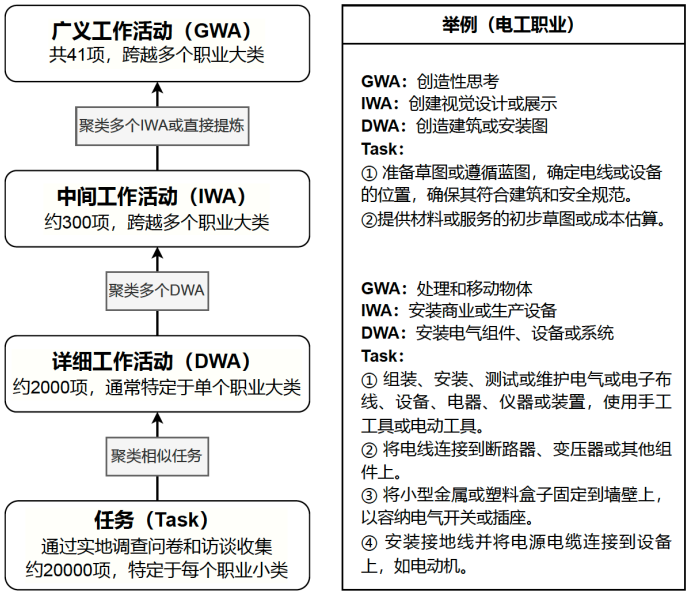

图 1：O*NET 不同层级指标定义和示例

注：该图引用了张丹丹等（2025）文中 P63 的图 1。

图 2 展示了 AI-LLM 暴露指数测算的完整技术路线，目标是基于招聘广告中的职位描述，评估每个职位受到大语言模型人工智能技术影响的程度，并在职业大类层面进行汇总，从而构建一个科学、标准化且具有通用性的指数体系。

在路线 1 中，O*NET 中的 2,087 个详细工作活动（DWA）作为定义工作内容的基础作为指数测算的基础；在路线 2 中，则使用 O*NET 中最为细致的 19,265 个任务（Task）作为分析单位，构建基于任务层级的暴露指数。在这一过程中，每个招聘岗位的文本描述可与最多 20 个 O*NET“任务”建立关联，从而形成岗位与任务之间的一对多映射关系。随后，我们利用大模型对这些“任务”受“大语言模型技术”影响的程度进行评分，并在职业或职业组别层面对评分进行加权平均，最终得到基于任务的大语言模型人工智能技术暴露指数。具体步骤如下：

第一步，在大模型的辅助下，我们对每一个详细工作活动（d_i）或者任务（t_i）进行暴露度评分（具体提示词见附录 A），分别记为s_i^d或者s_i^t。值得注意的是，Eloundou 等（2024）的另一贡献在于，他们比较了人工评分与大模型评分的差异，并验证了大模型在人工智能暴露度评估中的准确性与稳定性。

第二步，将每个招聘广告j中的职位描述文本与 O*NET 全部详细工作活动（DWA）和任务（Task）进行匹配，筛选出广告j中涉及的所有 DWA（$d_x, \ldots, d_y$）和 Task（$t_x, \ldots, t_y$）并确定它们在该岗位中的重要性权重（$w_x, \ldots, w_y, \sum_{x,\ldots,y} w = 1$）。

第三步，将第一步获取的相关 DWA 暴露度评分（$s_x^d, \ldots, s_y^d$）和 Task 暴露度评分（$s_x^t, \ldots, s_y^t$）按重要性权重加权求和，得到每个岗位j在 DWA 层面和 Task 层面的暴露度指数。

第四步，对同一职业或职业组别中所有岗位j的指数取平均值，即可得到该职业（或职业组别）的暴露度，进而构造出基于 DWA 和基于 Task 的大语言模型人工智能技术暴露指数。

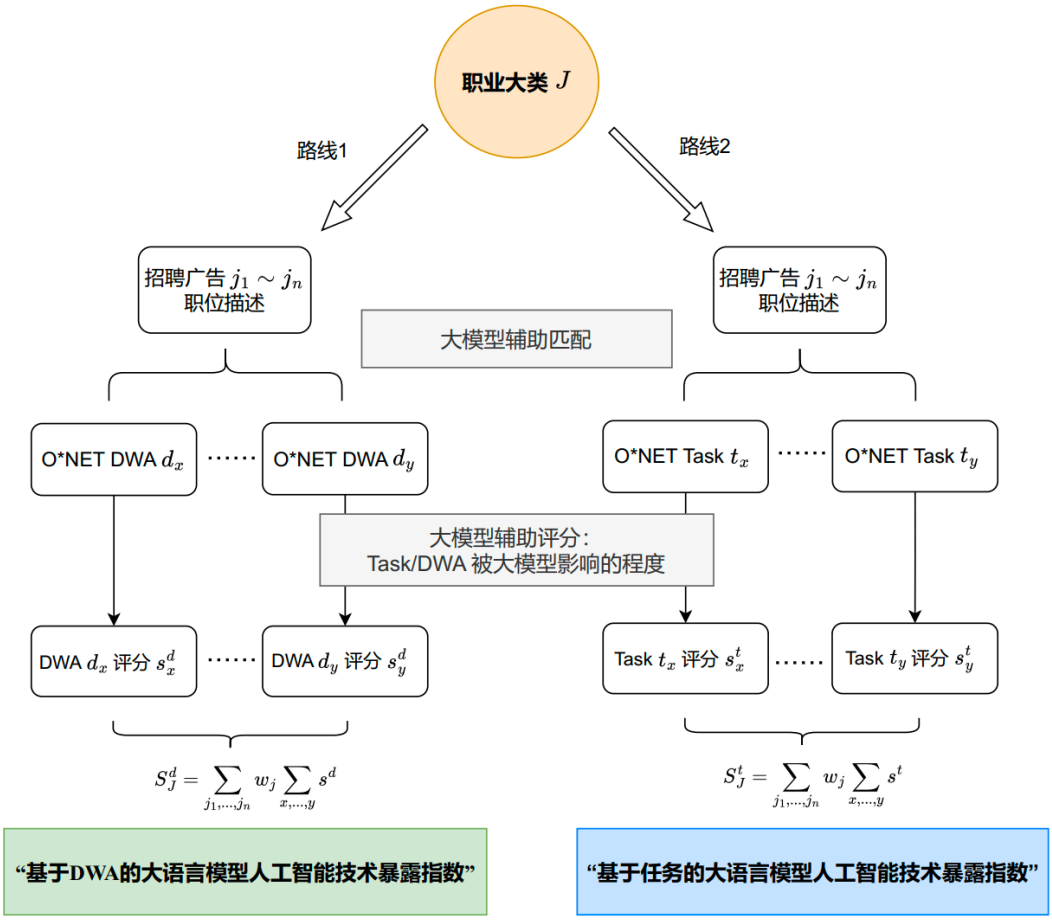

图 2：AI-LLM 暴露指数构建策略

注：该图引用了张丹丹等（2025）文中 P64 的图 2。

对步骤二（广告文本和 O*NET 中的 Task 匹配方法）的详述

为了确保整个任务匹配过程更加透明和可追踪，结合 Hampole 等（2025）的思路与招聘广告数据的特征，将任务提取流程细化为四个步骤（见图 3），能够在处理流程中保留清晰的中间输出。此外，该方法还支持对提取结果进行针对性优化，例如按照最高相似度阈值进行筛选和取舍，从而在保持精度的同时提高结果的稳定性与可复现性。

第一步，提取含工作任务的信息。从招聘广告（JD）中过滤出与工作任务相关的有效信息。招聘广告文本通常包含工作任务、任职要求、福利待遇以及礼貌性表述等，其中仅工作任务与任职要求与指数构建直接相关，其他冗余信息会干扰后续处理，因此不予考虑。本步骤采用大语言模型结合提示词工程（具体提示词见附录 B 任务一）实现抽取。为确保输出的逻辑性与稳定性，提示词结构遵循“五段式设计”：①任务描述；②详细说明；③逻辑指引；④输出模板；⑤参考示例。

第二步，提炼具体工作任务。在第一步过滤结果的基础上，根据 O*NET 中任务（Task）的语义表述，从中文 JD 中进一步提炼出具体工作任务，并翻译为英文（具体提示词见附录 B 任务二）。此步骤在翻译与重述时尽可能保留中文 JD 的原始语义与细节，以保证与岗位真实需求的匹配度。

第三步，清理、审查工作任务。为剔除无效或低质量的任务表述，对第二步的输出进行进一步筛查。例如，去除过于笼统或无实际信息量的描述（如“按要求完成其他任务”），以及与岗位职责无关的内容。本步骤通过更严格的筛选与提取规则（具体提示词见附录 B 任务三），最终过滤掉约 10%的无效任务表述。

第四步，匹配标准任务。将前三步提取并筛选得到的英文工作任务，与 O*NET 标准任务进行文本相似度匹配。具体方法为：利用 Sentence-BERT 框架，将招聘广告中的任务和 O*NET 的 Task 分别编码为 768 维向量（使用在英语任务匹配中表现最佳的开源模型 paraphrase-MPNet-base-v2），计算两者的余弦相似度。对每条招聘任务，选取相似度最高的 O*NET Task 作为匹配结果。本研究的匹配相似度均值约为 0.724，表明整体匹配质量较高。

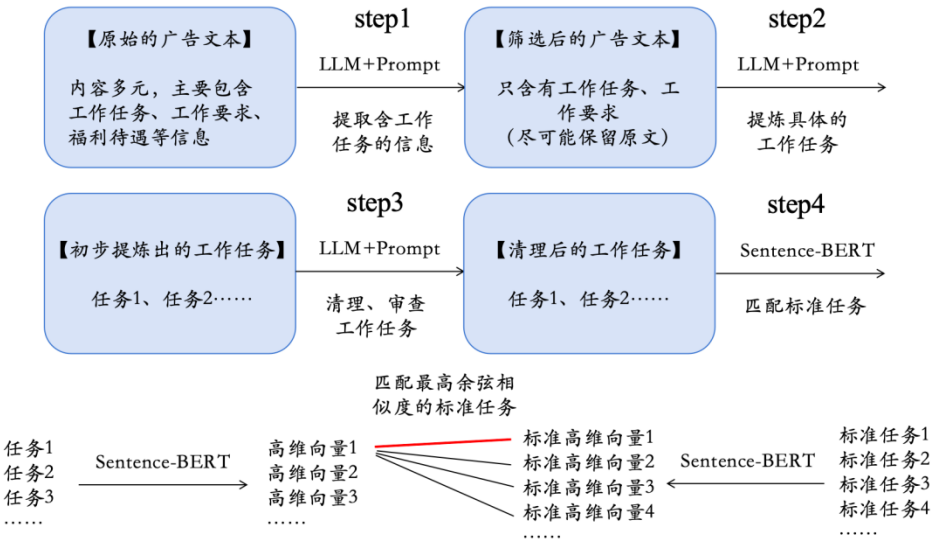

图 3：从招聘广告中提取出标准化的任务的步骤

与《管理世界》论文方法的差异

在《管理世界》论文（张丹丹等，2025）中，AI-LLM 暴露指数的核心环节是通过大语言模型直接从招聘广告文本中提取标准化任务。这一做法虽然可行，但由于依赖提示词反复调整，中间步骤难以验证，因而在可解释性和可控性上存在局限。相比之下，本次发布的新指数在方法上进行了优化，特别是在任务提取环节引入了更透明、可追踪的流程，从而显著提高了结果的严谨性与可复现性。这一点也是新指数与论文方法之间最主要的差异。新的构造方法详见于航、张丹丹、张润博、李泓孛（2025）。

招聘广告数据来源

测算数据来自在线招聘平台——智联招聘的招聘广告。智联招聘成立于 1994 年，是中国最大的在线招聘平台之一，拥有约 3.74 亿活跃个人用户和 1436 万企业用户，对城镇劳动力市场具有一定代表性。需要注意的是，在线招聘广告通常更集中于城市地区、相对年轻且高学历人群，以及中高薪酬岗位，因此在将结果推广至整体劳动力市场时需谨慎解释。

JD 样本覆盖 **2018 年 1 月 1 日至 2025 年 7 月 31 日**，通过全国范围内的分层随机抽样获得。具体而言，2018–2024 年每年均随机抽取约 20 万条招聘广告，涵盖发布时间、职

位要求、薪酬待遇及完整职位描述等信息；2025 年虽仅覆盖前 7 个月，但抽取量同样约为 20 万条。

整个市场 AI-LLM 暴露指数的年度变化趋势

依据新测算方法，研究团队计算了每个年度基于 Task 层级的 AI-LLM 暴露指数（见图 4）。具体而言，通过大模型辅助打分，结合专家评估，先计算 Task 的暴露度，再根据 JD 和 Task 的对应关系，计算出 JD 的暴露度，并按年份 JD 的构成汇总到整个市场，得到分年度的 AI-LLM 暴露指数。

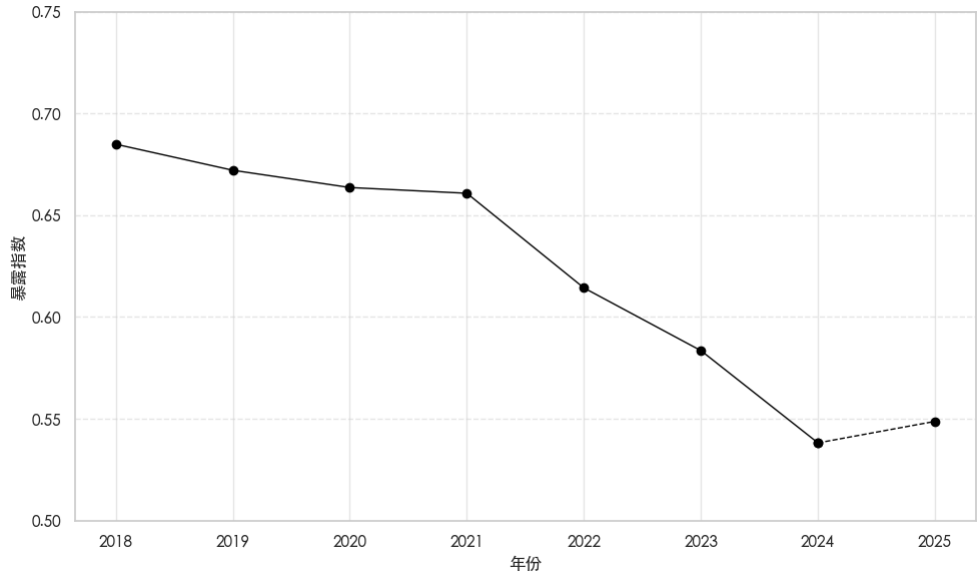

图 4：AI-LLM 暴露指数的年度变动趋势

注：2025 年的 AI-LLM 暴露指数基于 1 月 1 日-7 月 31 日期间的 JD 数据计算，非整年度测算。

二、指数发布说明

常用术语

- SOC 体系：全称 Standard Occupational Classification system，是美国劳动统计局对职业的划分体系，被广泛应用于劳动经济学研究，最新版于 2018 年发布；
- SOC 详细职业：SOC 体系中的最小职业单位，最新版共 1,016 个，编码格式 XX-XXXX.XX;
- SOC6 位职业：部分 SOC 详细职业汇总而成，最新版共 867 个，编码格式 XX-XXXX，或 XX-XXXX.00；
- SOC6 位职业第 24 版（旧版），2018 年之前的 SOC 体系版本，用于与 Felten 指数（基于第 24 版）的 SOC 分类体系 6 位职业进行匹配，同样有 867 个；
- SOC minor group（张丹丹等（2025）中翻译为“职业组别”），为 SOC 详细职业汇总得到，最新版共 98 个职位，编码格式 XX-XX；
- 智联职业分类：智联招聘平台根据其海量岗位数据构建了标准化的职业分类体系，该分类体系按照职业类型的细分程度分为三个层级，研究团队采用了第二级分类。

说明：为了统一使用 SOC 系统的职业分类，我们构造了智联内部的职业分类体系与 SOC 体系的对照表。智联招聘从 2023 年开始正式启用了新的内部职业分类体系；我们尽可能地保持了智联职业分类与 SOC 体系对应的合理性和一致性，但体系变更仍可能导致 SOC 职业层面暴露指数在 2023 年出现跳跃。这一跳跃在样本量较小的 SOC 职业组别和样本量较小的城市（间接法）指数中可能更明显。

发布指数清单如下：

1、基期指数：基于基期年份（2018-2021 年）的分职业暴露指数

职业分类级别：

（1）SOC 职业组别（SOC minor group）；为保证有效性，保留基期样本占比最多的前 60 个 SOC 职业组别。

（2）SOC 详细职业（SOC detailed occupation）；为保证有效性，保留基期样本占比最多的前 100 个 SOC 详细职业。

（3）智联二级职业；为保证有效性，保留基期样本占比最多的前 100 个智联二级职业。

2、动态指数：2018-2024 年年度分职业暴露指数

职业分类级别：

（1）职业组别（SOC minor group）；为保证有效性，保留基期样本占比最多的前 60 个 SOC 职业组别。

（2）SOC 详细职业（SOC detailed occupation）；为保证有效性，保留基期样本占比最多的前 100 个 SOC 详细职业。

（3）智联二级职业；为保证有效性，保留基期样本占比最多的前 100 个智联二级职业。

更新频率：年度。

参考文献

张丹丹、于航、李力行、胡佳胤、莫怡青、李泓宇，“中国人工智能技术暴露度的测算及其对劳动需求的影响：基于大语言模型的新证据”，*《管理世界》*，2025. 41(7):59-72。

于航、张丹丹、张润博、李泓宇，“人工智能技术进步与劳动力市场错配：来自中国岗位和求职者匹配数据的证据”，*工作论文*，2025.

Tyna Eloundou, Sam Manning, Pamela Mishkin, and Daniel Rock (2024), “GPTs are GPTs: Labor Market Impact Potential of LLMs”, *Science*, 384(6702):1306-08

Menaka Hampole, Dimitris Papanikolaou, Lawrence D.W. Schmidt, and Bryan Seegmiller (2025), “Artifical Intelligence and The Labor Market”, *NATIONAL BUREAU OF ECONOMIC RESEARCH Working Paper* 33509, http://www.nber.org/papers/w33509

附录 A：

Eloundou 等（2024）对 Task 打分的提示词（Prompt）

（见 Eloundou 等（2024）Supplementary Materials 中的 P10-11）

E Exposure Taxonomy

Consider the most powerful OpenAI large language model (LLM) This model can complete many tasks that can be formulated as having text input and text output where the context for the input can be captured in 2000 words. The model also cannot draw up-to-date facts (those from <1 year ago) unless they are captured in the input.

Assume you are a worker with an average level of expertise in your role trying to complete the given task. You have access to the LLM as well as any other existing software or computer hardware tools mentioned in the task. You also have access to any commonly available technical tools accessible via a laptop (e.g. a microphone, speakers, etc.). You do not have access to any other physical tools or materials.

Please label the given task according to the taxonomy below.

E0 — No exposure

Label tasks E0 if direct access to the LLM through an interface like ChatGPT or the OpenAI playground cannot reduce the time it takes to complete this task with equivalent quality by half or more.

If a task requires a high degree of human interaction (for example, in person demonstrations) then it should be classified as E0.

E1 — Direct exposure

Label tasks E1 if direct access to the LLM through an interface like ChatGPT or the OpenAI playground alone can reduce the time it takes to complete the task with equivalent quality by at least half. This includes tasks that can be reduced to:

- Writing and transforming text and code according to complex instructions,
- Providing edits to existing text or code following specifications,
- Writing code that can help perform a task that used to be done by hand,
- Translating text between languages,
- Summarizing medium-length documents,
- Providing feedback on documents,
- Answering questions about a document, or
- Generating questions a user might want to ask about a document.

E2 — Exposure by LLM-powered applications

Label tasks E2 if having access to the LLM alone may not reduce the time it takes to complete the task by at least half, but it is easy to imagine additional software that could be developed on top of the LLM that would reduce the time it takes to complete the task by half. This software may include capabilities such as:

- Summarizing documents longer than 2000 words and answering questions about those documents
- Retrieving up-to-date facts from the Internet and using those facts in combination with the LLM capabilities
- Searching over an organization’s existing knowledge, data, or documents and retrieving information

Examples of software built on top of the LLM that may help complete worker activities include:
- Software built for a home goods company that quickly processes and summarizes their up-to-date internal data in customized ways to inform product or marketing decisions
- Software that is able to suggest live responses for customer service agents speaking to customers in their company's customer service interface
- Software built for legal purposes that can quickly aggregate and summarize all previous cases in a particular legal area and write legal research memos tailored to the law firm's needs
- Software specifically designed for teachers that allows them to input a grading rubric and upload the text files of all student essays and have the software output a letter grade for each essay
- Software that retrieves up-to-date facts from the internet and uses the capabilities of the LLM to output news summaries in different languages

E3 – Exposure given image capabilities

Suppose you had access to both the LLM and a system that could view, caption, and create images. This system cannot take video media as inputs. This system cannot accurately retrieve very detailed information from image inputs, such as measurements of dimensions within an image. Label tasks as E3 if there is a significant reduction in the time it takes to complete the task given access to a LLM and these image capabilities:
- Reading text from PDFs,
- Scanning images, or
- Creating or editing digital images according to instructions.

Annotation examples:

Occupation: Inspectors, Testers, Sorters, Samplers, and Weighers
Task: Adjust, clean, or repair products or processing equipment to correct defects found during inspections.
Label (E0/E1/E2/E3): E0
Explanation: The model does not have access to any kind of physicality, and more than half of the task (adjusting, cleaning and repairing equipment) described requires hands or other embodiment.
Occupation: Computer and Information Research Scientists
Task: Apply theoretical expertise and innovation to create or apply new technology, such as adapting principles for applying computers to new uses.
Label (E0/E1/E2/E3): E1
Explanation: The model can learn theoretical expertise during training as part of its general knowledge base, and the principles to adapt can be captured in the text input to the model.
Activity: Schedule dining reservations.
Label (E0/E1/E2/E3): E2
Explanation: Automation technology already exists for this (e.g. Resy) and it's unclear what an LLM offers on top of using that technology (no-diff). That said, you could build something that allows you to ask the LLM to make a reservation on Resy

for you.
Activity: Negotiate purchases or contracts.
Label (E0/E1/E2/E3): E2
Explanation: You could have each party transcribe their point of view and then feed this to an LLM to resolve any disputes (E3). That said, many people would need to buy into using new technological tools to accomplish this (system).
Occupation: Allergists and Immunologists
Task: Prescribe medication such as antihistamines, antibiotics, and nasal, oral, topical, or inhaled glucocorticosteroids.
Label (E0/E1/E2/E3): E2
Explanation: The model can provide guesses for different diagnoses and write prescriptions and case notes. However, it still requires a human in the loop using their judgment and knowledge to make the final decision.

附录 B:

任务一：提取含工作任务的信息

【第一部分】

System Prompt:

Your task is to review a job posting from an online recruitment platform and extract the job tasks and requirements directly from the job posting while ensuring accuracy and structure. Please use the following instructions, recommended steps, samples and output format to respond to user inputs.

User Prompt:

【第二部分】**Instructions:**

(1)The job posting may describe multiple aspects of the position, but you should focus only on job tasks (the activities and responsibilities the employee will be expected to perform) and requirements (the education, age, experience, qualifications, skills, or other personal characteristics the employer expects the candidate to have).

(2)Extract only the raw text corresponding to "job tasks" and "requirements." Ignore all other content, such as salary, benefits, company descriptions, and work arrangements.

(3)Classify correctly. Do not rely solely on section titles. Some sections labeled "Job Responsibilities" may actually describe required skills. Sometimes, job tasks or requirements may be scattered in separated places of the posting; make sure to capture all of the relevant information. Use the following rules: If a sentence describes actions performed on the job, it belongs in job tasks. If a sentence describes skills, qualifications, experience, age, gender, or other personal attributes, it belongs in the requirements.

(4)Pay attention to the negative semantics. Sometimes, in a recruitment advertisement, it states "There is no need to do a certain task." In such cases, do not categorize this statement under "job tasks" or "requirements".

(5)Ignore promotional or motivational content. Do not include statements about career paths, income potential, or generalized encouragement (e.g., "This could change your life!").

(6)Maintain original phrasing. Copy the relevant text verbatim. Do not summarize or paraphrase. If the information is scattered, group all relevant parts together under the appropriate section without modifying their structure.

(7)Handle ambiguous cases: At most times, if a sentence could belong to both sections, place it in the most relevant category based on primary intent. But some job tasks can be derived from a requirement statement. For example, from the requirement statement "Be proficient in using relevant architectural design software such as AUTO CAD, PHOTOSHOP, and 3D MAX", we know that the job includes the task "Use architectural design software such as AUTO CAD, PHOTOSHOP, and 3D MAX". In this case, classify this kind of statement as both a job task and a requirement.

(8)Avoid duplication. If the same information appears multiple times in different sections, include it only once in the most appropriate category.

【第三部分】**A three-step process is suggested:** First, filter out the original job

posting and only keep information about job tasks and job requirements. Second, reread the original job posting to ensure that no tasks or requirements are neglected in your answer. Third, check the job tasks and requirements in your answer are correctly classified based on the instructions.

【第四部分】**Output format** (Do not include anything else in the answer. If no relevant information is found, return "No relevant information found" under the respective section like "Job Tasks: No relevant information found | Requirements: No relevant information found"): Job Tasks:{1.Task1;2.Task2;...}| Requirements:{1.Requirements1;2.Requirements2;...}.

【第五部分】For reference, here is an example of correctly applied filters: Original job posting:"销售行政/商务;销售行政专员/助理;办公室助理;岗位职责： 1、根据领导安排执行业务计划，并且有独立的销售能力 2、协助领导做好业务细节； 3、日常客户接待； 4、协助上级处理其他日常工作。 任职要求： 1.踏实认真，严谨细心，抗压力强； 2.具备较强的组织能力、沟通协调能力和责任心； 3.较好的亲和力、判断能力，应变力较好，执行力强； 4.具有较强的文字组织能力和语言表达能力，熟练日常办公操作； 5.工作态度积极乐观，自信、正直，良好的职业精神和团队协作力； 6.有同岗位经验的优先录取 公司福利： 五险、绩效奖金、全勤奖、带薪年假、员工旅游、节日福利、领导 nice" A good answer: "Job Tasks:1.根据领导安排执行业务计划，并且有独立的销售能力。2.协助领导做好业务细节。3.日常客户接待。4.协助上级处理其他日常工作。| Requirements:1.踏实认真，严谨细心，抗压力强。2.具备较强的组织能力、沟通协调能力和责任心。3.较好的亲和力、判断能力，应变力较好，执行力强。4.具有较强的文字组织能力和语言表达能力，熟练日常办公操作。 5.工作态度积极乐观，自信、正直，良好的职业精神和团队协作力。 6.有同岗位经验的优先录取。"

Here is the job posting you need to filter: < job post content >

任务二：提炼具体工作任务

System Prompt:

Your task is to extract specific tasks from the provided "Job Tasks" text in a recruitment poster. Please use the following instructions, recommended steps, samples and output format to respond to user inputs.

User Prompt:

Instructions:

(1)For structure and content, format each task to be similar to O*NET tasks, such as "Direct or coordinate an organization's financial or budget activities to fund operations, maximize investments, or increase efficiency", "Attend receptions, dinners, and conferences to meet people, exchange views and information, and develop working relationships".

(2)For length and language, each task that you extract is expected to be between 10 - 30 words in English.

(3)Maintain consistency with the original text description as much as possible.

(4)Some tasks in "Job Tasks" are long and complex such as "Operation, assembly, testing, inspection, packaging", split it into small tasks which are consistent with the O*NET tasks.

(5)Some tasks are simple and general such as "Product packaging", integrate and summarize moderately in combination with the context.

(6)Specific information such as place names, company names, and product names should not appear in the tasks you extract. Instead, generalized nouns and language should be used. For example, "Be responsible for the good credit of customers of Huabei and Jiebei of Ant Group" should be cleaned up as "Maintain good credit history of borrowers of the company’s micro loan product". "1. Short-distance freight driver; 2. Fixed short-distance routes in Shenzhen, 2 to 3 trips per day, with each trip taking 4 to 5 hours" should be comprehensively summarized as "Drive short-distance freight along fixed routes".

(7)The aim is to calculate the text similarity between these extracted tasks and O*NET tasks.

A two-step process is suggested: First, extract specific tasks from the provided "Job Tasks" text and organize the language and structure to make it comparable with O*NET tasks. Second, reread the given "Job Tasks" text and check the extracted job tasks to ensure that there are no duplications or omissions, and that the granularity and description of the tasks are consistent with those in O*NET.

Output format (Do not include anything else in the answer. There should be no quotation marks at the very front and the very end of the output. If no relevant information is found, return "No relevant information found"): {Task1 | Task2 | ...}.

For reference, here is an example of correctly applied filters: Original "Job Tasks" text: "Job Tasks:1.根据公司提供的客户名单，电话邀约上门咨询。2.为客户提供专业，热情的课程咨询，以达成签约。3.做好客户服务，提成客户满意度，增加转介绍。4.完成每月的招生指标。" A good answer: "Make phone calls to invite clients to come for on-site consultations according to the client list provided by the company. | Provide clients with professional and enthusiastic course consultations to achieve contract signing. |

Provide good customer service, enhance customer satisfaction, and increase referrals. | Achieve the monthly enrollment target."

Here is the text from the recruitment poster: < job tasks and requirements >

任务三：清理、审查工作任务

System Prompt:

Your task is to filter a "Job Tasks" list. Please use the following instructions, samples and output format to respond to user inputs.

User Prompt:

(1)Some task descriptions contain non-English characters, please return "This task is multilingual."

(2)Some task descriptions are too broad, please return "This task is too broad." For example, the description "Complete other tasks assigned by the leader" is too general—we cannot determine the specific nature of the tasks from such a vague statement.

(3)Some task descriptions are actually requirements, please return "This task is a requirement." For example, the description "Master foundational JAVA and J2EE technologies" outlines a requirement rather than a specific, actionable task.

(4)Some task descriptions are overly detailed, including specific information like place names, company names, and product names. You should generalize such specifics and return the revised task. For example, "Be responsible for the good credit of customers of Huabei and Jiebei of Ant Group" should be generalized to "Maintain good credit history for borrowers of the company's microloan products." Similarly, "Travel to Japan for business assignments for one year" becomes "Travel to foreign countries for business assignments for one year."

(5)For qualified task descriptions, simply return the original text.

The input is a set of tasks like "Task1 | Task2 | ...". The output format should be (Do not include the quotation mark): "Processed Task1 | Processed Task2 | ...".

For reference, here is an example of correctly applied filters: Original text: "与客户对接，洽谈合作. | Follow the leader's arrangements and work conscientiously. | Have less than one year of experience. | Fixed short-distance routes in Shenzhen, 2 to 3 trips per day, with each trip taking 4 to 5 hours | Make no more than 20 phone calls daily to targeted clients." A good answer: "This task is multilingual. | This task is too broad. | This task is a requirement. | Drive short-distance freight along fixed routes. | Make no more than 20 phone calls daily to targeted clients."

Here is the text you need to filter: < extracted tasks >