

中文引用格式:唐海洋,刘振翼,陈东平,等. ChatSOS:基于大语言模型的安全工程知识问答系统[J]. 中国安全科学学报,2024,34(8):178-185.

英文引用格式:TANG Haiyang, LIU Zhenyi, CHEN Dongping, et al. ChatSOS: large language model-based knowledge Q&A system for safety engineering[J]. China Safety Science Journal, 2024, 34(8): 178-185.

ChatSOS:基于大语言模型的安全工程 知识问答系统*

唐海洋,刘振翼 副教授,陈东平 副教授,初庆钊** 副教授
(北京理工大学 机电学院,北京 100081)

中图分类号:X913.2

文献标志码:A

DOI: 10.16265/j.cnki.issn1003-3033.2024.08.1901

资助项目:北京理工大学爆炸科学与技术国家重点实验室自主课题(ZDKT21-01);北京理工大学科技创新项目(2022CX01028)。

【摘要】 为解决大语言模型在安全工程领域应用时面临的语料库规模、输入处理能力和隐私性限制等问题,以2013—2023年间共117篇爆炸事故调查报告为基础构建向量数据库,利用大语言模型的生成式能力,进行提示工程,提出一个基于大语言模型的安全工程知识问答(Q&A)系统——ChatSOS;与ChatGPT大语言模型相比,ChatSOS能够通过整合外部知识库,使大语言模型根据用户的输入信息,从数据库中检索相关语料,并深入分析。结果表明:ChatSOS具备深入分析问题、自主分配任务的能力,能够详尽总结事故报告并提出建议;通过结合外部知识库解决基础大模型在安全工程领域语料不足和语料实时性不高的问题,避免了使用新数据集微调模型可能导致的模型性能下降等问题,提升了大语言模型在安全工程领域的应用能力。

【关键词】 ChatSOS; 大语言模型; 安全工程; 知识问答(Q&A)系统; 事故调查; 向量数据库

ChatSOS: large language model-based knowledge Q&A system for safety engineering

TANG Haiyang, LIU Zhenyi, CHEN Dongping, CHU Qingzhao

(School of Mechatronic Engineering, Beijing Institute of Technology, Beijing 100081, China)

Abstract: To address the limitations of large language models in safety engineering, such as the corpus size, input processing capabilities and privacy concerns, ChatSOS, a Q&A system based on large language models, was developed. Based on 117 explosion incident reports from 2013 to 2023, a vector database to enhance the system's capability was constructed. ChatSOS integrated prompt engineering and external knowledge base to retrieve and analyze relevant data from the database. Compared to ChatGPT, ChatSOS integrated the external knowledge base, so that the big language model could retrieve the relevant corpus from the database according to the user's input information and make in-depth analysis. The results show that ChatSOS excels in in-depth professional problem analysis, autonomous task allocation, and providing detailed summaries and recommendations based on incident reports. By combining with the external

* 文章编号:1003-3033(2024)08-0178-08; 收稿日期:2024-02-20; 修稿日期:2024-05-25

** 通信作者:初庆钊(1992—),男,山东烟台人,博士,副教授,主要从事危险物质安全分析等方面的研究。E-mail:chuqz@bit.edu.cn。

knowledge database, the limitations of the large language model's professional corpus in safety engineering are overcome, which prevents performance degradation associated with fine-tuning on new datasets, broadens the application of large language models in this field, and paves the way for future advancements in automation and intelligent systems.

Keywords: chat safety oracles (ChatSOS); large language model; safety engineering; knowledge question answering (Q&A) system; accident investigation; vector database

0 引言

在如今信息技术快速发展的背景下,知识获取和管理受到前所未有的冲击与挑战。信息渠道逐渐增多,相较于传统的图书或基础互联网检索,人工智能的发展为检索提供了更为便捷的方式,如利用文本挖掘技术,从大量非结构化数据中提取有价值的信息^[1],但信息的真实性和完整性难以保证。对于事故调查报告,受到保密性和信息复杂程度限制,目前相关信息的收集和管理仍极度依赖人工,且关系型数据库管理系统(Relation Database Management System, RDBMS)的维护与管理也需要专业人员^[2-3],导致这种传统的信息检索方式无法满足迅速准确获取知识的需求,成为安全事故调查领域突出的痛点。

在这一背景下,GPT^[4-5],LLaMA,PaLM^[6]等大语言模型在自然语言处理任务方面的快速发展引人注目,其理解和响应人类指令技术突出。例如:VAGHEFI等^[7]在模型中开通了联合国政府间气候变化专门委员会第6次评估报告的访问权限,显著提升了模型回答的准确性,展示了大语言模型作为未来气候科学研究工具的潜力;BRAN等^[8]设计了基于大语言模型的代理系统,整合多种专家设计的化学计算工具,成功规划了一种驱虫剂和3种有机催化剂的合成过程,提高了大语言模型在化学领域的性能。然而,大语言模型在安全工程领域的应用,如事故调查报告的收集和管理,还存在诸如信息滞后、语料缺乏和模型自身“幻觉”等问题,在安全性评估等高度专业化的应用中并不够可靠。

鉴于此,笔者拟建立事故数据库,提出一套结合大语言模型和向量数据库的安全事故数据管理分析方法,构建一个基于大语言模型的安全工程知识问答(Question Answering, Q&A)系统——ChatSOS(Chat Safety Oracles),以期为安全事故报告的自动化处理提供新的方向,并促进安全工程领域的技术创新。

1 ChatSOS构建的基本理论

1.1 大语言模型

大语言模型是目前人工智能领域内的研究热点,属于自然语言处理领域内的分支之一,大语言模型通过在大量文本上训练,以捕捉语言的统计学规律,并自动调整模型参数。语言模型本质上是计算一系列的概率,分析并量化给定文本序列,其中,N元语法模型是概率语言模型的基础,预测过程计算公式为^[9]:

$$p(w_1, w_2, \dots, w_n) = \prod_{i=1}^n p(w_i | w_1, w_2, \dots, w_{i-1}) \quad (1)$$

$$p(w_i | s) = \frac{C(s, w_i)}{C(s)} \quad (2)$$

式中: $p(w_1, w_2, \dots, w_n)$ 为词序列出现的概率; w_i 为词序列 $[w_1, w_2, \dots, w_n]$ 中的一个单词; $p(w_i | s)$ 为在给定上下文 s 的条件下,某一单词 w_i 出现的估计概率; $C(s, w_i)$ 为上下文 s 与单词 w_i 共同出现的频率; $C(s)$ 为上下文 s 出现的频率。

模型通过特定训练和调优过程后,本身可以具备一定的思维链能力,能将复杂任务拆分为多个可解决的中间步骤。在ChatSOS中,相应的大语言模型主要作为生成式能力的核心,使用提示词激发模型的推理能力,引导生成结果的大致方向。

1.2 数据库系统

传统的RDBMS(如结构化查询语言数据库)以表格的形式存储数据,可保证数据的一致性和完整性,但对于数据的形式要求比较严苛(结构化数据),难以提取非结构化数据(如自然语言)中的信息进行储存和分析,这会增加对有效信息提取与保存的难度。同时RDBMS在处理大规模、高维度数据的表现不够理想。向量数据库在处理数据时,通常会使用“文本-向量”嵌入模型,将文本转化为高维度的向量,在向量嵌入过程中,嵌入模型会捕捉文本中的语义信息,语义相似的文本片段被编码成在

高维度向量空间中相近的位置,使得向量数据库在检索时能通过向量间的距离来快速检索出语义相似的条目,保证了向量数据库的检索泛化能力。

对于知识问答系统而言,语言模型作为内核,数据多为报告与总结形式的非结构化自然语言文本,且数据量通常比较大。因此,采用向量数据库能有效规避 RDBMS 对数据形式的严格要求和其在处理大规模、高维度数据性能不足的缺陷,提高数据处理和分析的效率。

1.3 提示工程

对于不同研究团队推出的基座大模型,根据模型训练语料和对齐标准的不同,模型的各项能力指标(如泛化能力)也会有区别。如 GPT-3.5-Turbo 的训练语料库截止于 2021 年 9 月,其对于此时间点后发生的事件无法作出回答,或是出现大模型幻觉。这本质上是源于训练数据的不足,除了数据时效性的问题,专业领域数据的敏感性也是导致这一问题的另一大原因。

为解决大语言模型训练语料不足的问题,最大程度保证回答的真实性与准确性,进行提示工程,注入相关知识片段,并将预测目标与具有对应语义关系的词句相联系,整合到用户提问中辅助回答。这种方式并不会对模型本身的参数造成影响,而是间接利用模型的逻辑推理能力和一定的泛化能力。相比直接对模型微调,应用提示工程的成本相对较低,效果显著且结果基本可控。文中所提到的提示模板主要采用将预设场景和相关知识注入问答模板的方式进行设计。

2 ChatSOS

以大语言模型作为生成式内容核心,基于提示工程,使用向量数据库与 MySQL 数据库,分别以“向量+标志符”和“标志符+文本”的形式储存数据,构建一套包含知识存储、提取与注入的 ChatSOS。

2.1 系统框架

通过分析系统不同部分的功能需求,基于提示工程和数据库建设技术构建知识问答系统架构,如图 1 所示。该系统包括 3 个基本层级,包括输入层、运算层和输出层。

运算层部分可细分为 2 个模块:知识库和问答处理。在安全工程知识库的构建中,使用 2013—2023 年间共 117 篇爆炸事故调查报告作为细分应用场景下的语料储备,系统对原始文档执行格式标准化处理,确保文档格式的一致性;通过文本预处理和分割,将文档拆分成可管理的片段;“文本到向量”的转换过程中,通过使用预训练的嵌入模型将文本数据映射到高维向量空间,以便捕捉语义信息。每个生成的向量及其对应的文本片段都会被分配一个唯一标志符。向量存储在专门的向量索引系统 Milvus 中,以支持高效的相似性检索,而文本片段及其元数据则存储在关系型数据库 MySQL 中,以便进行结构化查询与数据处理。在问答处理过程中,构建代理来处理信息检索、需求分析、问题拆解、模型调用、用户交互等方面的问题。

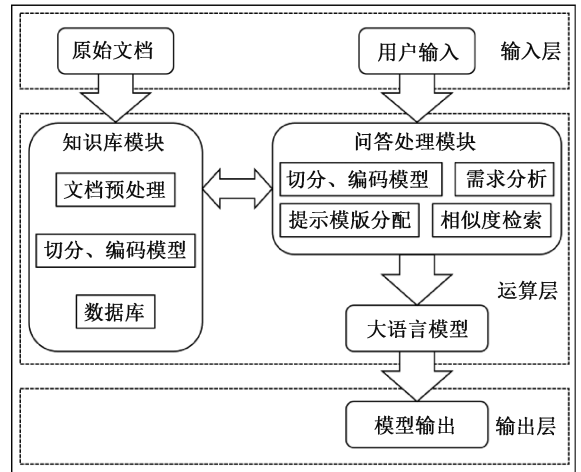


图 1 基于提示工程和数据库操作的知识问答系统架构

Fig. 1 Framework of knowledge Q&A system based on prompt engineering and database operation

大语言模型容易产生误导性和“幻觉”现象,主要是因为缺乏真实可靠的语料或使用了误导性的训练语料。因此,通过分析不同用户群体的需求(表 1),针对性地构建提示模板,通过使用提示模板,引导任务需求并整合相关专业信息,以提高模型推理时的目标领域知识权重,从而优化模型输出。初步试验表明:这种方法能有效减少模型在处理不熟悉问题时生成不准确或与偏离事实信息的频率。

表 1 行业需求分析

Table 1 Industry demand analysis

场景	需求	具体形式
政府部门	总结、报告、通知	政府公文
科研工作者	数据分析、研究成果分享	学术论文、研究报告、学术演讲

续表 1

场景	需求	具体形式
(本专业)企业从业者	安全风险评估、事故预案	安全风险评估报告、安规、安全培训资料
(跨专业)企业从业者	跨部门协作、了解紧急应对措施	跨部门协作指南、应急相应指南
公众	公共安全教育、灾害应对	科普文章、新闻稿件

2.2 知识问答处理流程

ChatSOS 在处理知识问答任务时的工作流程如图 2 所示。在数据处理过程中,用户查询首先通过北京智源人工智能研究院通用嵌入模型编码为向量表示,此后这一向量表示被用于向量数据库中的相似度检索。根据预设的算法,数据库返回相似度得分最高的前 k 个结果(k 为数据库检索结果中从排序结果中从前往后取出的结果数量)。得益于向量保留的丰富语义信息和高效的向量检索算法,ChatSOS 在处理多样化和复杂的用户查询时表现出卓越的泛化能力,利用 t 分布-随机近邻嵌入技术^[10],从高维向量空间降维至二维平面,其结果如图 3 所示,通过数据点之间的相对位置和集群趋势反映出数据点在语义上的关联性。相同的颜色代表其有类似的语义,位置相近的点可能会有更高的语义相似程度。但实际数据所包含的高维空间结构在映射后可能会被扭曲和压缩,可能不会显示出如此规则的分布。与此同时,模型代理接管问答任务并解析用户查询,根据不同场景需求选择合适的提示模板。在接收到数据库的查询结果后,模型代理会将这些结果作为补充知识融入提示模板中。最终,这些整合后的提示词被输入到大型模型中,系统随后将模型的输出结果反馈给用户。

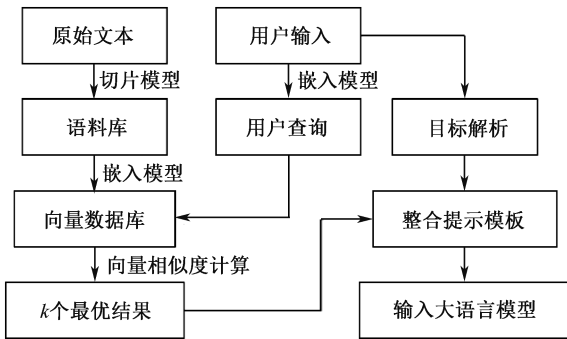


图 2 ChatSOS 问答任务处理流程

Fig. 2 ChatSOS Q&A processing workflow

3 大语言模型对比测试

3.1 大语言模型输出评价标准设计

调查现有评价大语言模型的常用标准^[11],

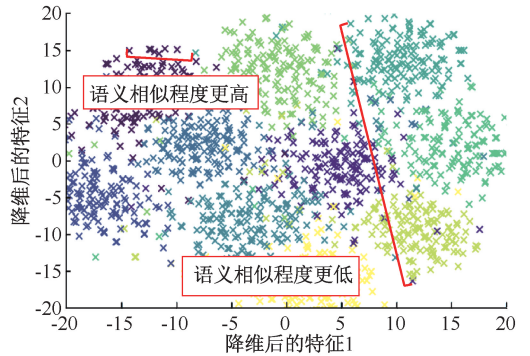


图 3 数据点的集群分布

Fig. 3 Cluster distribution of data points

BOMMASANI 等^[12]在问答测试中将大语言模型回答的准确性、鲁棒性和效率纳入评价体系。考虑到专业评价指标与实际用户使用需求,增加了适应性和简洁性 2 个维度,为基于知识库的问答系统设计了一套评分标准,见表 2。

表 2 系统问答响应评分

Table 2 Rating chart for system Q&A response

评分标准	分数(0~5)
准确性	—
可靠性	—
适应性	—
简洁性	—
快速性	—
加权总分	—

该评分表从 5 个维度评价 ChatSOS 输出:准确性评判回答问题的准确程度;可靠性评判系统回答的一致性;适应性评判系统对问题要求作出的调整;简洁性评判回答是否冗长且直接相关;快速性则评价系统的响应速度。

3.2 提示模板构建

分析不同人群需求,通过针对性地构建提示模板,采用“ n -shot learning”的方式,实现了适应性强的多样化提示方法,其中, n 为问答模板的数量。针对不同场景需求设计的提示模板基本流程举例说明见表 3。

3.3 大语言模型问答效果对比

依照 3.1 节提出的标准,设计一系列针对

表3 提示模板设计
Table 3 Prompt template designing

场景预设 & 知识注入	你是专业的人工智能助手,根据给定信息回答用户提问。回答时逐步思考,若依给定知识无法回答,只需说明给定信息不足,绝对禁止编造。已知信息,用户提问
事故分析	执行分析事故报告任务,从“事故概况、事故原因、责任分析、事故影响及预防、总结与建议”等多方面合理分析
问答模板	<p>Q1:请分析 2007 年湘西凤凰塌桥事故</p> <p>A1:1. 事故概括:时间—2007 年 8 月 13 日;地点—湖南省湘西凤凰县;环境条件—堤溪沅江大桥建设中;事故等级—严重责任事故;伤亡—64 人死亡,22 人受伤;直接经济损失—3974.7 万元</p> <p>2. 责任分析:调查认定施工、建设单位严重违规,监理单位、质量监督部门失职,湘西自治州和凤凰县两级政府及湖南省交通厅、公路局监管不力</p> <p>3. 事故原因:大桥主拱圈砌筑材料未满足规范和设计要求,施工工序不合理,主拱圈砌筑质量差,降低了结构整体性和强度,随着施工荷载增加导致坍塌</p> <p>4. 事故影响:人员伤亡及直接经济损失</p> <p>5. 事故预防:提升工程管理和现场监管,完善设计审查流程,强化责任意识等</p>

ChatSOS 的测试项目,包括通识性问答、格式规范任务和具体事故分析。测试中,将 ChatSOS 与 GPT 3.5-Turbo 模型的输出结果对比分析(表 4—表 6),

其中 ChatSOS 通过应用程序接口调用 GPT3.5-Turbo 作为其内核的大型语言模型,以验证所提出的改进结构对问答任务提升的有效性。

表4 通识性问答测试
Table 4 Performance evaluation of general knowledge Q&A

用户查询	ChatSOS	GPT 3.5-Turbo
为什么会发生燃气泄漏,有什么预防办法吗?	<p>根据给定的信息,可以得知以下情况可能导致燃气泄漏:</p> <ol style="list-style-type: none"> 燃气管道老化腐蚀:老旧燃气管道设施腐蚀造成的事故比例增加,这类事故隐蔽性强、不易及时发现,易酿成严重后果 管道破损:管道的老化破损、脱落或动物咬噬等情况可能导致燃气泄漏 第三方施工破坏:燃气管网中,第三方施工破坏事故数量占比较高,可能施工不慎损坏燃气管道 车辆撞击:地上燃气管道防护措施不到位,车辆撞击导致燃气管道泄漏 <p>针对燃气泄漏的预防措施包括:</p> <ol style="list-style-type: none"> 管道维护与更新:加强老旧燃气管道的更新改造,及时检修更换老化腐蚀的管道,减少泄漏风险 定期检查:定期对燃气管道进行检查和维护,确保管道的完整性和安全性 连接软管和燃气灶具的检查:用户端连接软管问题仍然是引发户内事故的主要原因,应定期检查连接软管和燃气灶具的安全性,及时更换老化损坏的软管和灶具 燃气阀门的正确使用:用户在使用燃气时,应正确关闭燃气阀门,避免瓶阀未关闭引发事故 燃气灶具的正确使用:用户在使用燃气灶具时,应避免无人看管和干烧,加强对用户安全使用燃气知识的宣传教育,推广使用有防干烧功能的灶具 加强安全宣传教育:对用户和相关从业人员进行燃气安全宣传教育,提高他们的安全意识和使用燃气的技能 <p>综上所述,通过管道维护与更新、定期检查、正确使用燃气阀门和灶具、加强安全宣传教育等措施,可以预防燃气泄漏事故的发生</p>	<p>燃气泄漏可能由多种原因引起,包括设备故障、人为错误、管道损坏、不当使用或安装等以下是一些常见的燃气泄漏原因和预防措施:</p> <p>燃气泄漏原因:</p> <ol style="list-style-type: none"> 设备故障或老化:燃气设备、管道或阀门因长时间使用或制造缺陷可能出现故障,导致泄漏 人为错误:错误的操作、不当维修、不规范的安装和施工可能导致管道破裂或连接处松动,引发泄漏 管道损坏:管道遭受外部压力、物理损坏、地震或自然灾害可能导致管道破裂 不当使用:例如,在不通风的空间使用燃气器具可能导致积聚的燃气泄漏 <p>燃气泄漏的预防和安全措施:</p> <ol style="list-style-type: none"> 定期检查和维修:定期检查家庭用具、管道和燃气设备,确保其状态良好,没有泄漏 合格安装:由专业人员合格安装和维护燃气设备,确保安装符合相关安全标准 使用燃气报警器:在可能泄漏的区域安装燃气报警器,能够及时检测到燃气泄漏并发出警报 通风:确保室内通风良好,特别是在使用燃气设备的封闭空间内 安全使用:使用燃气器具时遵循使用说明,避免长时间不间断使用,确保炉具在使用后关闭 防止损坏:避免在地下或易受损害的地方铺设燃气管道,以减少外部压力引起的破裂风险 应急处理:如果怀疑有燃气泄漏,立即打开门窗通风,离开泄漏现场,不要使用手机或电器设备,以免引发火源,然后拨打紧急救援电话 <p>请注意,如果你怀疑家中有燃气泄漏,绝对不能使用明火、电器开关,或者手机等可能引发火源的设备,应立即采取安全措施并呼叫紧急救援电话</p>

表 5 格式规范类任务测试

Table 5 Performance evaluation of format standardization task

用户查询	ChatSOS	GPT 3. 5-Turbo
<p>按照政府公文的写作规范,发布一份关于太阳岛洗浴中心事故的政府事故调查结果通知。</p>	<p>【政府事故调查结果通知】 尊敬的各相关方: 根据《生产安全事故报告和调查处理条例》,经过对驻马店市开发区太阳岛洗浴休闲会馆事故进行全面调查,现将调查结果通知如下: 一、事故概况:2021年11月18日16时57分,太阳岛洗浴休闲会馆男浴区桑拿房内的桑拿炉发生爆炸,造成3人死亡,3人受伤,直接经济损失约500万元。 二、事故调查经过:1.事故发生前,太阳岛会馆正处于翻新装修施工的尾声,计划于11月19日开业。2.事发当日,工作人员××*调试设施设备时,打开了男浴区桑拿炉控制开关,并调试温度在65~68℃之间自动控温。3.然而,由于未及时发现和消除桑拿炉存在的安全隐患,导致爆炸事故的发生。 三、事故原因分析:1.太阳岛洗浴休闲会馆在装修施工期间未能严格按照安全生产规定进行管理,未落实安全生产主体责任,未督促指导沐浴场所落实安全管理规定。2.太阳岛洗浴休闲会馆未能加强沐浴行业安全日常检查,未及时消除事故隐患。3.太阳岛洗浴休闲会馆对桑拿炉存在的安全隐患未能及时发现和处理,导致爆炸事故的发生。 四、责任追究:根据调查结果,对驻马店市开发区太阳岛洗浴休闲会馆的管理部门、相关责任人员存在管理不到位、监管不力等问题,将依法进行相应责任追究。 五、安全隐患整改:驻马店市开发区将加强对沐浴行业的管理工作,督促沐浴场所落实安全生产主体责任和相关安全管理规定,加强日常检查,及时消除事故隐患,确保公众的安全和利益。 六、救援和善后工作:事故发生后,市委市政府成立了事故抢险救援指挥部,组织了现场救援、调查和善后处置工作,全力排查安全隐患,防范发生次生事故,同时组织医疗救治工作,确保受伤人员得到及时救治。 七、倡议:广大市民和相关单位加强安全意识,严格按照安全生产规定进行管理,共同维护社会的安全稳定。 特此通知。 驻马店市政府 日期:××××年××月××日</p>	<p>太阳岛洗浴中心事故的政府事故调查结果通知 [日期] 尊敬的市民: 根据国家法律法规和《政府信息公开条例》,为向社会公众及时、公正、公开地通报太阳岛洗浴中心事故的调查结果,将事故详细调查情况通告如下: 一、事故基本情况 2022年月日,太阳岛洗浴中心发生火灾事故。事故发生后,市政府立即启动应急响应,组织消防、安监、公安等相关部门全力展开救援和调查工作。 二、事故原因 经过调查,事故的直接原因是洗浴中心内部设施设备存在安全隐患,加之管理不善,未能采取有效措施加以防范。具体原因包括(但不限于):设备老化、维护不善、员工培训不足、应急预案不健全等。 三、事故处理措施 1.责任追究:对相关责任人进行严肃问责,追究其领导责任和监管责任。 2.安全整改:对全市范围内的洗浴、娱乐等场所进行全面排查,督促整改存在的安全隐患。 3.加强监管:加大对公共场所的日常监管力度,确保市民生活、娱乐场所的安全。 4.加强宣传教育:加强火灾、安全知识的宣传,提高市民的自防能力。 五、结语 政府将一如既往地高度重视人民群众的生命财产安全,对于类似事故将继续加大监管力度,切实保障市民的生命和财产安全。感谢广大市民对政府工作的理解和支持,也请大家加强自身安全意识,共同维护社会安定与和谐。 特此通知。 [政府机关名称][日期]</p>

注:*具体事故涉及人员名称可被 ChatSOS 有效抓取,但是受个人隐私保护影响,测试结果中相关数据已被人为省略。

在处理符合特定格式规范的任务中,GPT-3. 5-Turbo 表现出较好的适应性,能准确理解政府通知的格式要求,将内容细分成多个条目总结。然而,由于缺少具体数据,GPT-3. 5-Turbo 的回答出现了所谓的“模型幻觉”,在总结太阳岛洗浴中心事故的基本情况时给出错误的事故时间,在分析事故原因时未能明确指出具体的致因因素,而是给出了一些较为通用的模糊表述,且在之后处理措施部分的总结中也出现了类似的模糊回答。另一方面,ChatSOS 在输出结果中准确地描述了事故的基础情况,并在事故原因分析中细化到了桑拿炉的安全问题,其余关键点如责任追究、隐患排查、救援及后续处理工作等

也展现了较为全面的考量。此外,在未提供具体负责机构名称的情况下,ChatSOS 凭借其深度的信息理解和推理能力,通过现有数据推导出了具体的负责机构为驻马店市政府,也再次强化了其输出结果的可信度。

在具体事故分析任务中,GPT-3. 5-Turbo 的表现受限于其训练语料,未能提供详细的事故分析。在分析山东四方安装工程公司较大爆燃事故时,ChatSOS 通过提示工程注入了相关知识条目和问答模板,全面解析了事故。ChatSOS 全面概述了该事故,并依照上下文详尽划分了事故涉及各方的责任,包括四方公司、施工单位、东营港经济开发区管委会

表 6 事故分析任务测试

Table 6 Performance evaluation of accident analysis task

用户查询	ChatSOS	GPT 3. 5-Turbo
请简要分析一下山东四方安装工程公司较大爆燃事故	<p>事故概括:山东四方安装工程公司在东营港经济开发区瑞林公司的碳钢系统污油罐改造施工过程中发生较大爆燃事故。事故发生于2022年7月26日上午9时30分,造成3人死亡,2人轻伤,直接经济损失约512.8万元。</p> <p>责任分析:事故中,山东四方安装工程公司未落实施工单位安全生产主体责任,对施工队管理不到位。东营港经济开发区管委会及其经济发展局、应急管理局也未认真履行安全生产监管职责。</p> <p>事故原因:直接原因是四方公司施工人员在污油罐存有油污、罐顶中心透气孔敞开的情况下,超出《特殊动火安全作业证》许可作业范围,违规在罐顶中心透气口附近进行焊接作业。间接原因包括四方公司安全管理混乱、安全生产责任制严重不落实、隐患排查治理不到位以及对外来施工队伍管理不到位等。</p> <p>事故影响:事故造成3人死亡,2人轻伤,直接经济损失约512.8万元。此外,事故还对周边环境造成了一定的影响。</p> <p>事故预防:为预防类似事故的发生,可采取以下措施:加强施工单位的安全生产责任制落实,完善安全管理制度和操作规程;加强隐患排查治理工作,确保施工现场的安全条件;加强对外来施工队伍的管理,确保其安全生产条件符合要求;加强安全风险分析,确保动火作业的风险辨识和管控措施到位。</p> <p>爆炸威力分析:根据事故描述,污油罐发生了爆燃,罐体撕裂。具体的爆炸威力需要进一步的详细分析和调查</p>	<p>很抱歉,根据本模型截至2021年9月的知识,无法提供有关山东四方安装工程公司较大爆燃事故的详细信息。请注意,由于模型无法获取实时信息,也无法提供截至目前的最新事件。如果这是最近发生的事件,模型建议查阅可靠的新闻源或官方声明以获取最新的信息和准确的分析</p>

及相关监管部门。此外,它能够区分事故的直接原因与间接原因,评估事故的影响并提出预防策略。ChatSOS在处理分析任务时具备深度和广度,充分证明了其在处理复杂任务上的优势。

依照评分标准对 ChatSOS 与 GPT-3. 5-Turbo 打分,各项评分对比如图 4 所示,轴上的点代表每项指标的得分情况。在处理通识性问题的问答任务中,由于训练语料充足,GPT-3. 5-Turbo 的回答覆盖面较

广,确保了较高的准确度和可靠性。在讨论燃气泄漏的具体原因及预防措施时,GPT-3. 5-Turbo 进行了详尽的分类,如将泄漏原因细分为设备故障和人为疏忽等多个方面。另一方面,ChatSOS 凭借其知识库的深入整合,能够在提供回答时援引具体数据和模板,提供了实际案例,如在回答中提到由于动物啃咬或车辆碰撞引起的燃气管道损伤,从而在广泛性与细致性之间实现了有效的平衡。

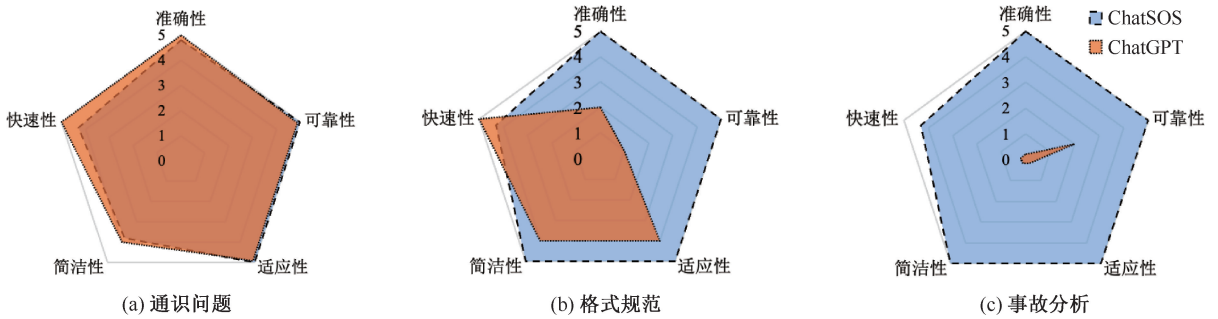


图 4 各项评分对比

Fig. 4 Comparison of scores across categories

4 结论

1) 和传统数据库直接检索方法相比,结合大语言模型和向量数据库,通过文本与向量的转化,在安全事故调查知识库检索方面,显著提高了知识库检索的效率和泛化能力,可为大语言模型提供详尽准确的语料。

2) ChatSOS 提升了大语言模型在安全工程领域的应用效果,并快速响应用户对安全事故的

提问,同时作出深入分析,通过模型优化和定制化应用,更好地满足安全工程领域对事故分析的需求。ChatSOS 在事故调查问答中回答的准确性、可靠性和适应性方面远好于 ChatGPT 大语言模型。

3) 安全工程领域的智能化分析与决策需求较多,未来将进一步探究更多实际应用场景,丰富和完善 ChatSOS 的工具组件,以验证和拓展模型的实用性和有效性。

参 考 文 献

- [1] 王亮, 赖佳燕, 张自欣, 等. 基于文本挖掘和改进 DEMATEL 法的化工事故关键因素识别[J]. 中国安全科学学报, 2024, 34(3): 20-28.
WANG Liang, LAI Jiayan, ZHANG Zixin, et al. Identification of critical factors in chemical accidents based on text mining and improved DEMATEL method[J]. China Safety Science Journal, 2024, 34(3): 20-28.
- [2] 李光耀, 孙振文, 朱军, 等. 危险化学品爆炸火灾事故数据库的构建及应用[J]. 消防科学与技术, 2023, 42(4): 575-579.
LI Guangyao, SUN Zhenwen, ZHU Jun, et al. Database construction and application of hazardous chemical explosion and fire accidents[J]. Fire Science and Technology, 2023, 42(4): 575-579.
- [3] 刘全龙, 张晓霖, 张悦倩, 等. 基于事故致因的标准化煤矿建筑规范库构建[J]. 中国安全科学学报, 2023, 33(5): 128-133.
LIU Quanlong, ZHANG Xiaolin, ZHANG Yueqian, et al. Construction of a database of standardized coal mine building regulation based on accident causation[J]. China Safety Science Journal, 2023, 33(5): 128-133.
- [4] RADFORD A, WU J, CHILD R, et al. Language models are unsupervised multitask learners[J]. OpenAI Blog, 2019, 1(8): 1-24.
- [5] BROWN T, MANN B, RYDER N, et al. Language models are few-shot learners[J]. Advances in Neural Information Processing Systems, 2020, 33: 1 877-1 901.
- [6] CHOWDHERY A, NARANG S, DEVLIN J, et al. PaLM: scaling language modeling with pathways[J]. Journal of Machine Learning Research, 2022: DOI: 10.48550/arXiv.2204.02311.
- [7] VAGHEFI S A, STAMMBACH D, MUCCIONE V, et al. ChatClimate: grounding conversational AI in climate science[J]. Communications Earth & Environment, 2023, 4: DOI: 10.1038/s43247-023-01084-x.
- [8] BRAN A M, COX S, SCHILTER O, et al. Augmenting large-language models with chemistry tools[J]. Nature Machine Intelligence, 2024, 6: 525-535.
- [9] 王乃钰, 叶育鑫, 刘露, 等. 基于深度学习的语言模型研究进展[J]. 软件学报, 2021, 32(4): 1 082-1 115.
WANG Naiyu, YE Yuxin, LIU Lu, et al. Language models based on deep learning: a review[J]. Journal of Software, 2021, 32(4): 1 082-1 115.
- [10] VAN DER MAATEN L, HINTON G. Visualizing data using t-SNE[J]. Journal of Machine Learning Research, 2008, 9: 2 579-2 605.
- [11] CHANG Yupeng, WANG Xu, WANG Jindong, et al. A survey on evaluation of large language models[J]. ACM Transactions on Intelligent Systems and Technology, 2024, 15(3): 1-45.
- [12] BOMMASANI R, LIANG P, LEE T, et al. Holistic evaluation of language models[J]. Annals of the New York Academy of Sciences, 2023, 1525(1): 140-146.

作者简介: 唐海洋 (2000—),男,湖南益阳人,硕士研究生,研究方向为大语言模型及相关应用场景等。E-mail: 3120230263@bit.edu.cn。

