

# 煤矿安全大模型——矿途智护者

1	项目简介 .....	1
2	项目架构图 .....	1
3	数据构造模块 .....	2
3.1	从煤矿安全领域法律法规等文件获取数据 .....	2
3.2	从煤矿从业人员入职考试题库获取数据 .....	4
3.3	煤矿历史事故案例、事故处理报告获取数据 .....	5
3.4	自我认知数据获取 .....	7
4	模型微调模块 .....	8
5	RAG 模块 .....	8
6	知识图谱构建模块 .....	9
7	部署模块 .....	11
8	测试评估 .....	11

## 1 项目简介

近年来，国家对煤矿安全生产的重视程度不断提升。为了确保煤矿作业的安全，提高从业人员的安全知识水平显得尤为重要。鉴于此，目前迫切需要一个高效、集成化的解决方案，该方案能够整合煤矿安全相关的各类知识，为煤矿企业负责人、安全管理人员、矿工提供一个精确、迅速的信息查询、学习与决策支持平台。为实现这一目标，我们利用包括煤矿历史事故案例、事故处理报告、煤矿安全领域法律法规、安全操作规程、规章制度、技术文档以及煤矿从业人员入职考试题库等在内的丰富数据资源，通过微调 **InternLM2** 模型，构建出一个专门针对煤矿事故和煤矿安全知识智能问答的煤矿安全大模型。

本项目的特点如下：

- 支持煤矿安全领域常规题型解答，如：单选题、多选题、判断题、填空题等（针对煤矿主要负责人及安管人员、煤矿各种作业人员）
- 支持针对煤矿安全领域法律法规、安全规程规章制度、技术等文档内容回答（如《中华人民共和国矿山安全法》、《煤矿建设安全规程》）
- 支持煤矿历史事故案例，事故处理报告查询，提供事故原因详细分析、事故预防措施以及应急响应知识

## 2 项目架构图



总体技术路线如下：

- 收集煤矿历史事故案例、事故处理报告、煤矿安全领域法律法规、安全操作规程、规章制度、技术文档以及煤矿从业人员入职考试题库等在内的丰富数据资源。
- 借助 **GLM-4**, **GPT-4o** 大模型，通过编写 **Prompt** 脚本从文本自动抽取问答对，包括单轮对话和多轮对话，同时借助人工收集题库，构造自我认知数据，最后人工审核，转变成 **Xtuner** 微调数据格式。
- 微调方面，使用 **Xtuner** 工具，采用 **QLora** 方法分别对 InternLM2-Chat-1.8B、InternLM2-Chat-7B、InternLM2-Math-7B、InternLM2-Chat-20B 模型进行微调，并分别命名为 CoalMineLLM\_InternLM2-Chat-1\_8B、CoalMineLLM\_InternLM2-Chat-7B、CoalMineLLM\_InternLM2-Math-7B、CoalMineLLM\_InternLM2-Chat-20B。同时采用 **Full Finetuning** 对 InternLM2-Chat-1.8B 模型全参微调，并命名为 CoalMineLLM\_InternLM2-Chat-1\_8B-full,使用 **LMDeploy** 工具，采用 **W4A16** 对 InternLM2-Chat-7B 模型量化，并命名为 CoalMineLLM\_InternLM2-Chat-7B-4bit。
- 为了模型高效准确回答来自用户的相关问题，我们采用 **RAG** 技术，针对用户的 query，检索来自向量数据库和知识图谱的相关知识信息，来增强回答效果。
- 部署模块，借助 **LMDeploy** 工具，进行高效推理部署到 **streamlit**、**gradio** 展示，以及部署到 **OpenXlab** 平台上。
- 在知识图谱构建模块，借助 **GLM-4**，编写信息抽取 **Prompt** 抽取三元组关系，最后存入到 **Neo4j** 中。（现在正在使用信息抽取指令数据集来微调 InternLM2 模型。）
- 在评测模块，针对煤矿事故处理报告、规章制度文件等数据源，采用人工和 **GPT-4o** 方式生成单轮和多轮测试数据集，以及收集未在训练集中出现的考试题库数据。采用主观、客观评估，针对选择题、判断题等题型，采用客观评估；其他类型问题，使用 **GPT-4o** 和专家来评估。

### 3 数据构造模块

#### 3.1 从煤矿安全领域法律法规等文件获取数据

使用 GLM-4 模型，通过编写指令，抽取数据。其 Prompt 模板如下：

```
def return_random_prompt():
```

```
    system_prompt = "根据下面提供有关煤矿安全领域文本，请你仔细通读全文，你需要依据该文本：\n\n#####\n{}\n#####\n\n尽可能给出多样化的问题和对应的回答。我们将用于人工评估 GLM-4 模型对问答对数据的完成情况。要求:\n".format(content)
```

```
    system_prompt += "1. 生成问题有价值且遵守该文本信息，回答准确专业。\\n"
```

```
    system_prompt += "2. 生成问答对不能重复。\\n"
```

```
    system_prompt += "3. 问题多样化，同个问题可以换成不同表述方式，但意思保持不变。\\n"
```

```
    system_prompt += "4. 为问题生成作为<input>，不应该只包含简单的占位符。<input>应提供实质性的内容问题，具有挑战性。字数不超过" + str(random.randint(80, 120)) + "字。\\n"
```

```
    system_prompt += "5. <output>应该是对问题的适当且真实的回答，不能只回复答应或拒绝请求。如果需要额外信息才能回复时，请努力预测用户意图并尝试回复，但不能胡编乱造。<output>的内容应少于" + str(random.randint(512,1024)) + "字。\\n\\n"
```

```
    system_prompt += "请给出满足条件的 20 条 JSON 格式数据，并存储在一个列表中，便于整理使用，不要输出无法的字符，只要列表形式存储 JSON 数据\\n"
```

```
    return system_prompt
```

最终变成数据格式形如：

```
[
    {
        "conversation": [
            {
                "system": "你是一个煤矿安全领域的知识达人，你对相关煤矿安全规章制度、技术等文档非常熟悉。请你专业正确地解答用户想问的煤矿安全相关问题。",
                "input": "请提供新建矿井立井井筒冻结法施工的具体要求。",
                "output": "新建矿井立井井筒冻结法施工需要遵守以下要求：冻结深度必须穿过风化带延深至稳定的基岩 10m 以上，第一个冻结孔必须全孔取芯，钻孔时必须测定钻孔的方向和偏斜度，偏斜度超过规定时必须及时纠正，冻结管必须采用无缝钢管并焊接或螺纹连接，开始冻结后必须经常观察水文观测孔的水位变化，并在确定冻结壁已交圈后才能进行试挖。"
            }
        ]
    },
    {
        "conversation": [
```

“system”: “你是一个煤矿安全领域的知识达人，你对相关煤矿安全规章制度、技术等文档非常熟悉。请你专业正确地解答用户想问的煤矿安全相关问题。”，

“input”: “煤矿安全规程中，对于新建矿井的立井井筒冻结法施工有哪些具体要求？”，

“output”: “新建矿井立井井筒冻结法施工需要遵守以下要求：冻结深度必须穿过风化带延深至稳定的基岩 10m 以上，第一个冻结孔必须全孔取芯，钻孔时必须测定钻孔的方向和偏斜度，偏斜度超过规定时必须及时纠正，冻结管必须采用无缝钢管并焊接或螺纹连接，开始冻结后必须经常观察水文观测孔的水位变化，并在确定冻结壁已交圈后才能进行试挖。”

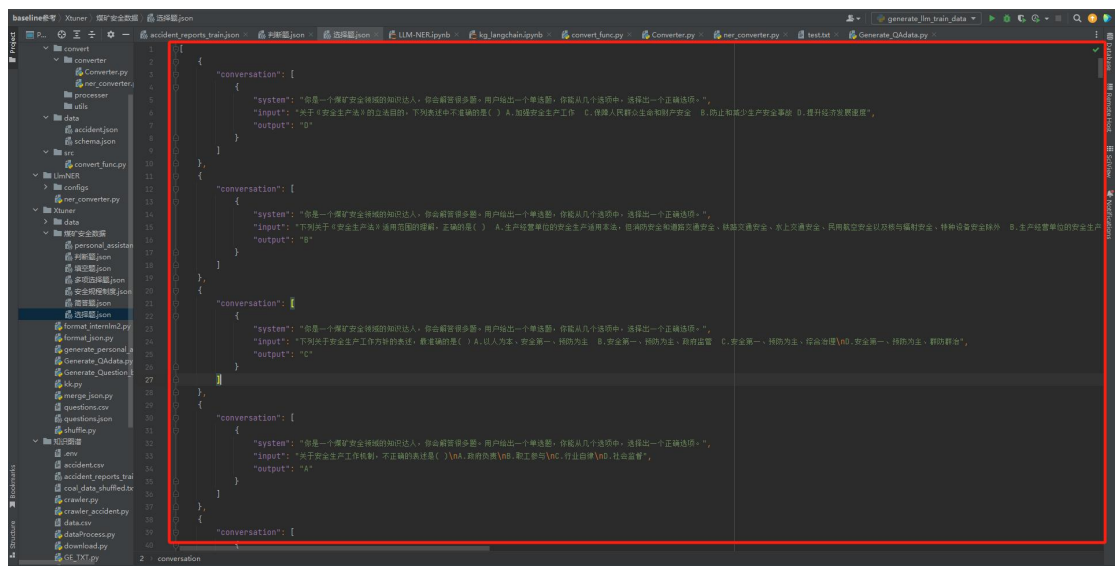
```
    }  
  ]  
  },  
  ...  
]
```

### 3.2 从煤矿从业人员入职考试题库获取数据

人工收集数据形如（以选择题为例）：

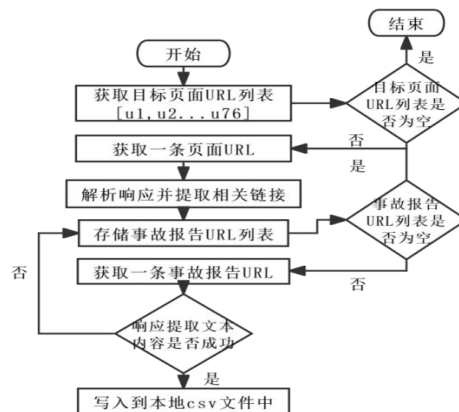
1	input	output
2	关于《安全生产法》的立法目的，下列表述中不准确的是( ) A.加强安全生产工作 C.保障人民群众生命和财产安全 B.防止和减少生产安全事故 D.提升经济发展速度	D
3	下列关于《安全生产法》适用范围的理解，正确的是( ) A.生产经营单位的安全生产适用本法，但消防安全和道路交通安全、铁路交通安全、水上交通安全、民用航空安全以及核与辐射安全、特种设备安全除外 B.生产经营单位的安全生产，适用本法；有关法律、行政法规对消防安全和道路交通安全、铁路交通安全、水上交通安全、民用航空安全以及核与辐射安全、特种设备安全另有规定的，适用其规定 C.生产经营单位的安全生产，适用本法；消防安全和道路交通安全、铁路交通安全、水上交通安全、民用航空安全以及核与辐射安全、特种设备安全参照适用本法有关规定 D.生产经营单位的安全生产，适用本法；消防安全和道路交通安全、铁路交通安全、水上交通安全、民用航空安全以及核与辐射安全、特种设备安全，适用其他有关法律、行政法规的规定	B
4	下列关于安全生产工作方针的表述，最准确的是( ) A.以人为本、安全第一、预防为主 B.安全第一、预防为主、政府监管 C.安全第一、预防为主、综合治理 D.安全第一、预防为主、群防群治	C
5	关于安全生产工作机制，不正确的表述是( ) A.政府负责 B.职工参与 C.行业自律 D.社会监督	A
6	某公司董事长由上一级单位总经理张某兼任，张某长期在外地，不负责该公司日常工作。该公司总经理安某在国外脱产学习，期间日常工作由常务副总经理徐某负责，分管安全生产工作的副总经理姚某协助其工作。根据《安全生产法》有关规定，此期间对该公司的安全生产工作全面负责的人是( ) A.安某 B.张某 C.徐某	C

最后通过脚本，变成：



### 3.3 煤矿历史事故案例、事故处理报告获取数据

先通过编写爬虫脚本，从煤矿安全网爬取事故报告数据，其爬取流程图，如下所示。



爬取后，然后整理成如下所示：



最后借助 GPT-4o，基于指令生成多轮对话数据，形如：

```
[
{
  "conversation": [
    {
      "system": "你是一名煤矿安全领域的知识达人，提供有关煤矿安全规程、事故预防措施和应急响应的知识。",
      "input": "你好！",
      "output": "您好，我是一名煤矿安全领域的知识达人，请问有什么可以帮助您？"
    },
    {
      "input": "2023 年 4 月 27 日陕西省府谷县国能业有限公司矿发生了什么事",
      "output": "2023 年 4 月 27 日，陕西省府谷县国能业有限公司矿 201-1 切",
      "input": "眼改造掘进工作面发生了一起顶板事故，导致 1 人死亡。",
      "output": "事故直接原因是 201-1 新切眼掘进工作面未按《作业规程》要求",
      "input": "进行临时支护，工人违章进入空顶区域作业，被冒落的顶板砸压。"
```

```

    },
    {
        "input": "那么，这起事故的间接原因有哪些？",
        "output": "事故间接原因包括现场安全管理混乱、技术管理存在漏洞、安全生产主体责任落实不到位、安全教育培训不到位和安全监管有差距。具体表现在作业规程执行不严格、安全技术措施审批把关不严、安全管理人员配备不足、安全管理人员岗位职责不明确等方面。"
    },
    {
        "input": "为了避免类似事故，煤矿企业应该采取哪些措施？",
        "output": "为了避免类似事故，煤矿企业应严格执行作业规程，加强现场安全管理，确保临时支护措施到位，强化技术管理，建立完善的技术管理体系，落实安全生产主体责任，加强安全教育培训，提高职工安全意识和风险辨识能力，同时加强安全监管，确保监管工作细致认真。"
    },
    {
        "input": "感谢您的详细解答！",
        "output": "不客气，很高兴能帮助您。如果您还有其他问题，欢迎继续提问。"
    }
]
},...
]

```

### 3.4 自我认知数据获取

通过如下脚本，生成 20 条自我认知数据

```

#coding=utf-8
import json
# 设置用户的名字
name = 'A100 换你 AD 钙奶团队'
# 设置需要重复添加的数据次数
n = 200
# 初始化 OpenAI 格式的数据结构
data = [
    {
        "conversation": [
            {
                "system": "",
                "input": "你好,你是谁？",

```



"output": "您好！我是{}研发的煤矿安全领域大语言模型。我们的团队成员来自安徽理工大学、中国矿业大学、北京化工大学、中国人民解放军总医院等多所高校/企业的博士、硕士、本科生以及个人开发者，共同参与了我的研发工作。我具备强大的语言处理能力和丰富的知识储备，旨在为煤矿企业负责人、安全管理人员、矿工等用户提供关于煤矿事故、煤矿安全规程规章制度及相关安全知识的智能问答服务。通过与我互动，您可以快速获取准确、全面的信息查询、学习和决策支持。我致力于成为您在煤矿安全管理领域的重要助手，帮助您提高安全意识、掌握安全规程、预防和应对煤矿事故，保障矿工的生命安全和企业的可持续发展。如果您有任何关于煤矿安全方面的问题或需求，欢迎随时向我提问，我将竭诚为您提供帮助。".format(name)

```
    }
  ]
}
```

# 通过循环，将初始化的对话数据重复添加到 data 列表中

```
for i in range(n):
```

```
    data.append(data[0])
```

# 将 data 列表中的数据写入到一个名为'personal\_assistant.json'的文件中

```
with open('煤矿安全数据/personal_assistant.json', 'w', encoding='utf-8') as f:
```

```
    # 使用 json.dump 方法将数据以 JSON 格式写入文件
```

```
    # ensure_ascii=False 确保中文字符正常显示
```

```
    # indent=4 使得文件内容格式化，便于阅读
```

```
    json.dump(data, f, ensure_ascii=False, indent=4)
```

## 4 模型微调模块

环境安装、数据准备、模型准备、下载配置文件、修改配置参数、模型训练、转换成 huggingface 格式、模型合并、测试模型。详见：<https://github.com/yaosenJ/CoalQA/blob/main/finetune/README.md>

## 5 RAG 模块

模块目的：

根据用户的问题，检索对应信息以增强回答的专业性，使 CoalQA 的回答更加专业可靠。检索内容包括但不限于以下几点：

- 煤矿安全相关的问答对
- 煤矿安全相关的案例

技术路线：

这个 RAG 系统各部分的技术路线分别是：

- Embedding 模型：BAAI/bge-small-en-v1.5
- LLM 基座：CoalMineLLM\_InternLM2-Chat-1\_8B CoalMineLLM\_InternLM2-Chat-7B CoalMineLLM\_InternLM2-Chat-20B CoalMineLLM\_InternLM2-Math-7B
- 使用 BAAI/bge-reranker-large 做检索后精排



- 向量数据库:
- FAISS: 是 Facebook 开源的一个高效的向量相似性搜索库。它为大规模向量检索提供了多种索引结构和搜索算法,能够快速查找与给定向量最相似的向量。FAISS 擅长处理超高维度的向量数据,在图像检索、推荐系统等场景有广泛应用。

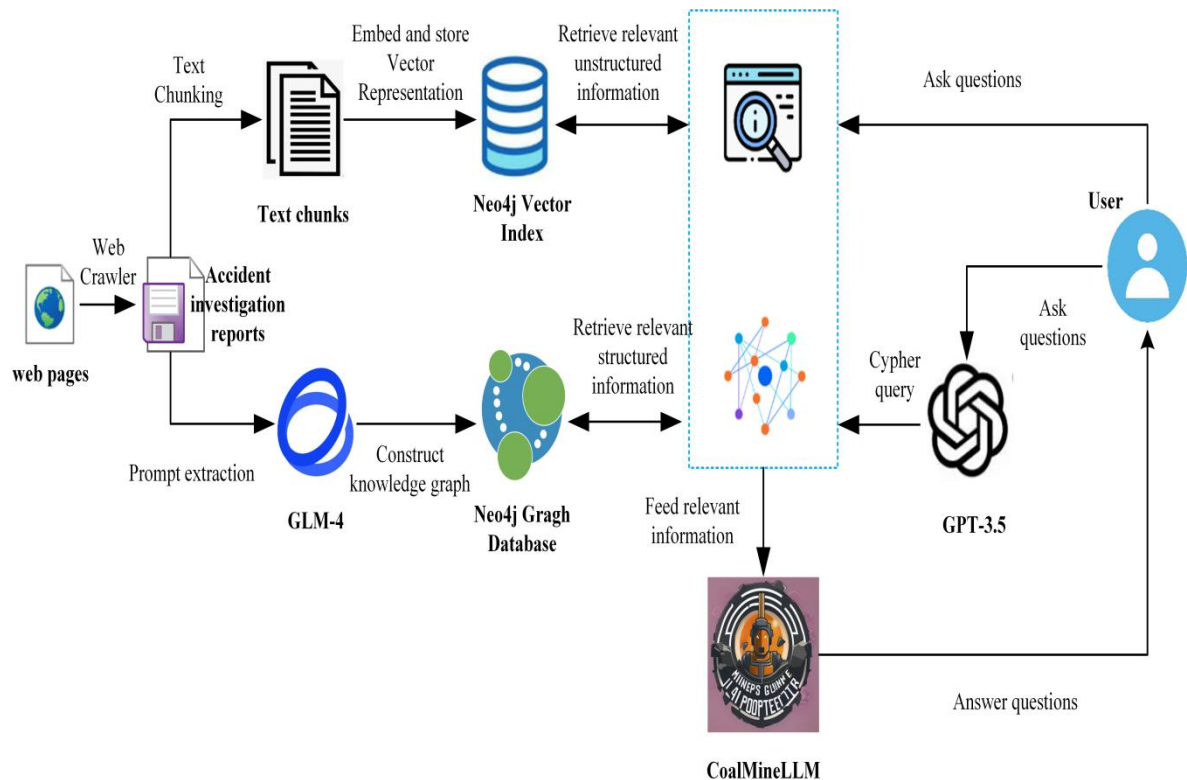
RAG 详细指南见: <https://github.com/yaosenJ/CoalQA/blob/main/rag/README.md>

后续计划

- 利用评测框架 RAGAS 对系统进行评估
- 构建二分类数据对向量模型以及精排模型进行训练以提升性能

使用 Neo4j 和 Langchain 集成非结构化和图知识增强煤矿事故 QA, 见:

[https://github.com/yaosenJ/CoalQA/blob/a1b7af0e04f708e60eb00c8f12fe08092947d133/demo/CoalMineLLM-InternLM2\\_Chat-1\\_8B-integrated-qa-neo4j-langchain.ipynb](https://github.com/yaosenJ/CoalQA/blob/a1b7af0e04f708e60eb00c8f12fe08092947d133/demo/CoalMineLLM-InternLM2_Chat-1_8B-integrated-qa-neo4j-langchain.ipynb)



## 6 知识图谱构建模块

使用 GLM-4 对煤矿事故信息抽取, 如:

**instruct:** "你是专门进行实体抽取的专家。请从 input 中抽取出符合 schema 定义的实体, 不存在的实体类型返回空列表。请按照(实体, 实体类别)的格式回答。"

**"schema":** [时间,事故地点,煤矿名称,煤矿事故,死亡人数,受伤人数,经济损失]

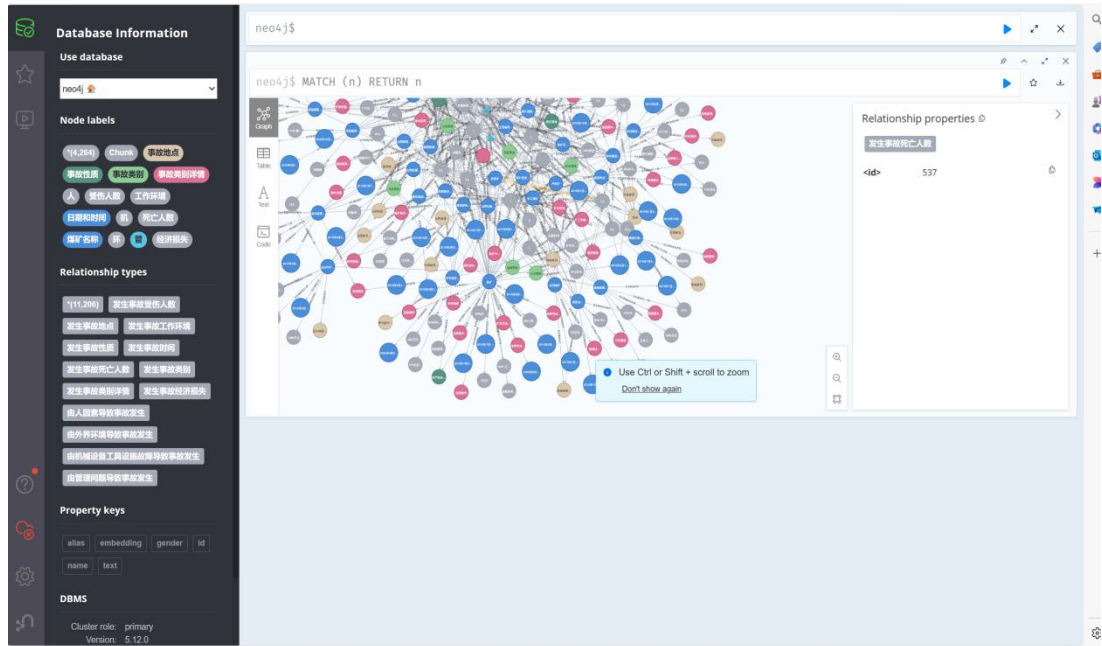
**"input":** "2023 年 8 月 21 日, 陕西省延安市延川县新泰煤矿(以下简称新泰煤矿)发生一起重大瓦斯爆炸事故, 造成 11 人死亡、11 人受伤, 直接经济损失 1919.2 万元"

**"output":** (2023 年 8 月 21 日,时间),(陕西省延安市延川县,事故地点),(新泰煤矿,煤矿名称),(重大瓦斯爆炸事故,煤矿事故),(11 人, 死亡人数),(11 人, 受伤人数),(1919.2 万元, 经济损失)

最终抽取信息内容, 如下所示:

#	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	序号	日期和时间	事故地点	煤矿名称	工作环境	事故类别	事故类别详细性质	人	机	管	环	死亡人数	受伤人数	经济损失			
2	790	2000年8月4日	某地	南翼皮带输送机事故		违规操作	违规操作	✓				0人	1人	7万元			
3	516	2005年4月30日	贵州省毕节市地区	31#主斜井	瓦斯事故	瓦斯爆炸	生产安全责任事故	✓		✓	✓	12人	0人	280万元			
4	1007	2010年1月2日	湖南省邵阳县	水口煤矿	矿井主斜井	其他事故	矿工失稳滑责任事故	✓		✓	✓	1人	0人	28.9万元			
5	2	2010年1月5日	湖南省湘潭市	湖南省湘潭市立井	三火事故	电缆短路着火		✓	✓	✓		34人	0人	未提供经济损失信息			
6	507	2010年1月5日	湖南省湘潭市	湖南省湘潭市立井	工作面火灾事故	短路着火	生产安全责任事故	✓	✓	✓	✓	34人	0人	2962万元			
7	1008	2010年1月5日	湖南省娄底市	良田煤矿	老井+259m	机电事故	接线柱触电责任事故	✓	✓	✓	✓	2人	0人	62万元			
8	1006	2010年1月25日	湖南省娄底市	蛇形山煤矿	1122工作面	顶板事故	矸石冒落	责任事故	✓	✓	✓	1人	0人	34万元			
9	759	2010年1月27日	安徽省宿州市	芦岭矿	II 826-2工	顶板事故	冒顶事故	✓	✓	✓		1人	0人	未提供经济损失信息			
10	995	2010年1月29日	某地	某矿	瓦抽队7803回巷	机电事故	违规操作麻花钻杆	✓	✓	✓		0人	1人	未提供经济损失信息			
11	730	2010年1月30日	某地	某矿	13051工作	其他事故	违章操作号责任事故	✓		✓		0人	1人	未提供经济损失信息			
12	729	2010年2月7日	某地	某矿	13051工作	其他事故	违章作业号责任事故	✓		✓		0人	1人	未提供经济损失信息			
13	758	2010年2月18日	安徽省淮北市	海孜煤电公司	1033外段	顶板事故	顶板掉矸	✓	✓	✓		1人	0人	未提供经济损失信息			
14	898	2010年2月21日	山西省长治市	潞安集团潘	1312出煤巷	水害事故	老空区水量积聚	✓		✓	✓	1人	0人	3972万元			
15	996	2010年2月23日	某地	某矿	综采二15102回采	机电事故	违章摘取液压枪	✓		✓	✓	0人	1人	未提供经济损失信息			
16	3	2010年3月1日	内蒙古自治	神华集团乌	16号煤层	水害事故	底板突水	✓		✓		32人	7人	未提供经济损失信息			
17	765	2010年3月1日	内蒙古自治	神华集团乌	16号煤层	回风	水害事故	承压水突破隔水带形成导水通道	✓		✓	32人	7人	4853万元			
18	880	2010年3月1日	某地	神华集团乌	16号煤层	水害事故	透水	✓	✓	✓		32人	7人	4853万元			
19	899	2010年3月1日	内蒙古自治	神华集团乌	16号煤层	水害事故	灰夹水大量责任事故	✓	✓	✓	✓	32人	7人	4853万元			
20	4	2010年3月15日	河南省新乡市	东兴煤矿	西大巷第一	火灾事故	电缆着火	责任事故	✓	✓		25人	0人	1145万元			
21	614	2010年3月15日	河南省郑州市	东兴煤矿	西大巷第一	火灾事故	违规使用才责任事故	✓		✓		25人	0人	1145万元			
22	771	2010年3月15日	某地	某矿	高档普采工	顶板事故	顶板掉矸	✓	✓			1人	0人	未提供经济损失信息			
23	5	2010年3月21日	湖南省常德市	湖南省常德市	老空区	水害事故	老空区积水涌出	✓		✓		13人	0人	未提供经济损失信息			
24	6	2010年3月22日	河南省驻马店市	泌阳县顺达矿	业有限公司	水害事故	透水	✓		✓		0人	11人	未提供经济损失信息			
25	7	2010年3月28日	山西省临汾市	中煤集团-20101回风	水害事故	老空区积水涌出	✓		✓			38人	0人	未提供经济损失信息			
26	764	2010年3月28日	山西省临汾市	王家岭煤矿	20101回风	水害事故	透水事故	✓		✓	✓	38人	0人	未提供经济损失信息			
27	859	2010年3月28日	山西省临汾市	王家岭煤矿	北翼盘区11	水害事故	废弃小窑老责任事故	✓		✓	✓	38人	0人	未提供经济损失信息			
28	861	2010年3月28日	山西省临汾市	王家岭煤矿	北翼盘区11	水害事故	小窑老空水责任事故	✓		✓	✓	38人	115人	未提供经济损失信息			
29	8	2010年3月31日	河南省伊川县	河南伊川矿	二1煤11021	瓦斯事故	煤矿瓦斯突出	✓		✓		48人	0人	2728.4万元			
30	505	2010年4月20日	江西省高安市	江西省高安丰井	瓦斯事故	煤与瓦斯突出	✓		✓	✓		12人	0人	未提供经济损失信息			
31	506	2010年4月23日	江苏省徐州市	张双楼煤矿	综采工作面	顶板事故	顶板再次来压导致石	✓		✓		1人	0人	未提供经济损失信息			
32	9	2010年5月4日	某地	神华集团某综采	工作面	顶板事故	冒顶事故	✓	✓			1人	0人	未提供经济损失信息			
33	876	2010年5月7日	内蒙古自治	内蒙古自治	灵东煤矿	综采工作面	顶板事故	顶板再次来压导致石	✓		✓	2人	0人	未提供经济损失信息			
34	903	2010年5月7日	内蒙古自治	内蒙古自治	灵东煤矿	综采工作面	顶板事故	冒落大量石块	✓		✓	2人	0人	未提供经济损失信息			
35	10	2010年5月8日	湖北省恩施州	恩施州恩施	水井湾煤矿	瓦斯事故	瓦斯爆炸	✓	✓	✓		10人	6人	580万元			
36	11	2010年6月6日	某地	神华集团某综采	工作面	运输事故	违章单人作业	✓				1人	0人	未提供经济损失信息			
37	452	2010年6月10日	某地	某矿	运安工东副斜井	运输事故	矿车挤压	✓		✓		0	1	未提供经济损失信息			
38	755	2010年6月15日	甘肃省金昌市	金昌市金昌	二矿17分	段其他事故	矿运机械在巷道帮拱	✓				1人	0人	未提供经济损失信息			
39	689	2010年6月16日	安徽省淮南市	淮南矿业(集团)有限	公司31223工	作面顶板事故	事故是由来压责任事故	✓		✓		3人	未提及	61.89万元			
40	997	2010年6月17日	某地	某矿	综采二13采区	机电事故	高压管爆裂	✓		✓		0人	1人	未提供经济损失信息			
41	12	2010年6月21日	河南省平顶山市	平顶山矿业	井上井下	放炮事故	井下炸药爆炸	✓		✓		49人	24人	未提供经济损失信息			

然后通过编写脚本文件，存储到 neo4j 图数据库中。



现已经构造好信息抽取数据，来微调 InternLM2 模型。

```
1 {
2   {
3     "text": "2008年9月13日13时50分, 洛阳市新安鑫泰煤业有限公司(以下简称鑫泰煤业公司)发生一起重大透水事故, 造成10人死亡, 1人受伤, 直接经济损失282.53万元。",
4     "entity": [{"entity": "2008年9月13日13时50分", "entity_type": "事故时间"}, {"entity": "洛阳市新安县", "entity_type": "事故地点"}, {"entity": "鑫泰煤业有限公司", "entity_type": "组织机构"}, {"entity":
5   },
6   {
7     "text": "2010年3月15日20:30, 新密市东兴煤业有限公司主井西大巷第一绕巷发生电缆着火事故, 当班入井31人, 截止3月16日凌晨2:00, 抢救工作结束, 6人安全升井, 25人遇难。",
8     "entity": [{"entity": "2010年3月15日20:30", "entity_type": "事故时间"}, {"entity": "新密市东兴煤业有限公司", "entity_type": "组织机构"}, {"entity": "主井西大巷第一绕巷", "entity_type": "工作面"}, {"entity":
9   },
10  {
11    "text": "2010年3月15日20:30, 新密市东兴煤业有限公司主井西大巷第一绕巷发生电缆着火事故, 当班入井31人, 截止3月16日凌晨2:00, 抢救工作结束, 6人安全升井, 25人遇难。",
12    "entity": [{"entity": "2010年3月15日20:30", "entity_type": "事故时间"}, {"entity": "新密市东兴煤业有限公司", "entity_type": "组织机构"}, {"entity": "主井西大巷第一绕巷", "entity_type": "工作面"}, {"entity":
13  },
14  {
15    "text": "2008年3月26日18时15分, 湖南省郴州市永兴县香樟乡刘家洲煤矿发生一起重大火灾与瓦斯突出事故, 造成13人死亡, 3人轻伤, 直接经济损失347万元。",
16    "entity": [{"entity": "2008年3月26日18时15分", "entity_type": "事故时间"}, {"entity": "湖南省郴州市永兴县香樟乡", "entity_type": "事故地点"}, {"entity": "刘家洲煤矿", "entity_type": "煤矿名称"}, {"entity":
17  },
18  {
19    "text": "2014年11月26日21时55分, 辽宁阜新矿业(集团)有限责任公司僧帽大井主井采区综采工作面发生一起重大煤尘爆炸事故, 截至12月1日已造成28人死亡、50人受伤。",
20    "entity": [{"entity": "2014年11月26日21时55分", "entity_type": "事故时间"}, {"entity": "辽宁省", "entity_type": "事故地点"}, {"entity": "阜新矿业(集团)有限责任公司僧帽大井主井采区", "entity_type": "组织机构"}, {"entity":
21  },
22  {
23    "text": "2016年4月25日, 铜川照金煤矿透水事故, 造成11名被困井下的矿工全部遇难。",
24    "entity": [{"entity": "2016年4月25日", "entity_type": "事故时间"}, {"entity": "铜川", "entity_type": "事故地点"}, {"entity": "照金煤矿", "entity_type": "煤矿名称"}, {"entity": "透水事故", "entity":
25  },
26  {
27    "text": "2023年6月21日10时20分左右, 徐州矿务(集团)新隆天山矿业有限责任公司(以下简称天山公司)南主斜井发生一起运输事故, 造成1人死亡, 直接经济损失231.16万元(截止6月26日)。",
28    "entity": [{"entity": "2023年6月21日10时20分左右", "entity_type": "事故时间"}, {"entity": "徐州矿务(集团)新隆天山矿业有限责任公司", "entity_type": "组织机构"}, {"entity": "天山公司", "entity_type": "工作面"}, {"entity":
29  },
30  {
31    "text": "2006年8月16日20时40分, 平朔公司安家岭井工2号井综掘二队在进行8903回风顺槽掘进工作面的可伸缩胶带运输机检修时发生一起运输死亡事故, 造成一人死亡。",
32    "entity": [{"entity": "2006年8月16日20时40分", "entity_type": "事故时间"}, {"entity": "平朔公司安家岭井工2号井综掘二队", "entity_type": "组织机构"}, {"entity": "8903回风顺槽掘进工作面", "entity_type":
33  },
34  {
35    "text": "2011年7月6日18时45分, 山东省平庄防务煤业有限公司井下-225米水平运输下山底溜车场一台空气压缩机着火, 引发事故造成28人被困。",
36    "entity": [{"entity": "2011年7月6日18时45分", "entity_type": "事故时间"}, {"entity": "山东省", "entity_type": "事故地点"}, {"entity": "平庄防务煤业有限公司", "entity_type": "组织机构"}, {"entity":
37  },
38  {
39    "text": "2008年9月13日13时50分, 洛阳市新安鑫泰煤业有限公司(以下简称鑫泰煤业公司)发生一起重大透水事故, 造成10人死亡, 1人受伤, 直接经济损失282.53万元。",
40  },
41 }
```

然后通过脚本把上面数据, 转成如下指令数据。

```
{
  "instruction": '{"instruction": "你是专门进行实体抽取的专家。请从 input 中抽取出符合 schema 定义的实体, 不存在的实体类型返回空列表。请按照 JSON 字符串的格式回答。", "schema": ["事故时间", "事故地点", "组织机构", "煤矿名称", "事故类别", "工作面", "死亡人数", "受伤人数", "经济损失"], "input": "2016 年 4 月 25 日, 铜川照金煤矿透水事故, 造成 11 名被困井下的矿工全部遇难。"}'
  "output": '{"事故时间": ["2016 年 4 月 25 日"], "事故地点": ["铜川"], "煤矿名称": ["照金煤矿"], "事故类别": ["透水事故"], "工作面": ["井下"], "死亡人数": ["11 名"]}'
}
```

## 7 部署模块

openxlab 部署

- 登陆 OpenXLab, 创建 Streamlit 应用
- 选择配置, 创建应用, 如果需要更多的硬件资源, 在这里进行申请
- 等待构建、启动
- 项目体验


本地部署、基于 LMDeploy 量化部署, 详见 <https://github.com/yaosenJ/CoalQA> 中部署模块。

## 8 测试评估

Whether use RAG for generate

Yes

No



Max Length

2848

8

32768

Top P

0.12

0.00

1.00

Temperature

0.10

0.00


1.00

Clear Chat History

Whether use RAG for generate

Yes

No



Max Length

2848

8

32768

Top P

0.10

0.00

1.00

Temperature

0.13

0.00

1.00

Clear Chat History

## 煤矿安全大模型--矿途智护者

你好，你是谁？

您好！我是A100换你AD钙奶团队研发的煤矿安全领域大语言模型。我们的团队成员来自安徽理工大学、中国矿业大学、北京化工大学、中国人民解放军总医院等多所高校/企业的博士、硕士、本科生以及个人开发者，共同参与了我的研发工作。我具备强大的语言处理能力和丰富的知识储备，旨在为煤矿企业负责人、安全管理人员、矿工等用户提供关于煤矿事故、煤矿安全规程与规范、煤矿工法等煤矿安全管理方面的知识和答案查询、学习和决策支持。通过与我互动，您可以快速获取准确、全面的信息查询、学习和决策支持。我致力于成为您在煤矿安全管理领域的重要助手，帮助您提高安全意识、掌握安全规程、预防和应对煤矿事故，保障矿工的生命安全和企业的可持续发展。如果您有任何关于煤矿安全方面的问题或需求，欢迎随时向我提问，我将竭诚为您提供帮助。

What is up?

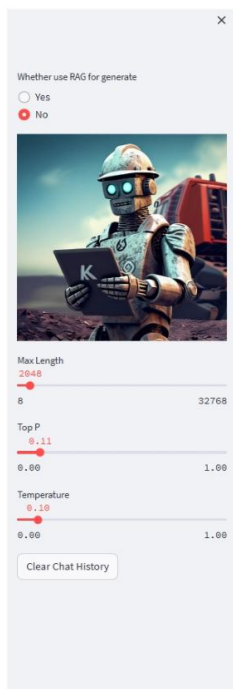
## 煤矿安全大模型--矿途智护者

简述修改后的《安全生产法》十大亮点是什么？

- 修改后的《安全生产法》有以下十大亮点：
- 坚持以人为本，强化安全生产工作中的法治原则。
  - 明确了安全生产工作的领导责任，规定了安全生产法定的安全生产监督管理部门和从业人员的安全生产主体责任。
  - 增加了有关从业人员的安全生产教育和培训的内容，要求企业、事业单位对从业人员进行安全教育和培训。
  - 强化了安全生产主体责任的落实，规定了未依法履行安全生产职责的违法行为。
  - 对一般安全生产事故和较大事故的定性定量标准进行了修订，增加了较大事故的定性定量标准。
  - 规定了事故隐患的排查和整改要求，增加了鼓励和规定的规定。
  - 增加了有关鼓励和规定的规定，鼓励企业实施安全生产主体责任，并对社会力量参与安全生产监督管理的规定。
  - 明确了事故调查和追究责任的程序和要求。
  - 增加了有关安全生产领域的数据统计和信息公开的要求。
  - 提出了建立安全生产领域行政执法与刑事责任追究相结合的机制的要求。

What is up?





请教你一个问题：矿山救护队在处理矿井火灾事故时应遵循哪些原则？

矿山救护队在处理矿井火灾事故时应遵循以下原则：控制烟雾蔓延，防止爆炸，防止风流逆转，保护人员安全，创造有利灭火条件，确保人员撤退和预防瓦斯爆炸。

帮我解答一下单选题：矿山、金属冶炼建设项目和用于生产、储存、装卸危险物品的建设项目，应当按照国家规定() A.进行安全条件论证 B.进行安全评价 C.分别进行安全条件论证和安全评价 D.进行安全条件论证或者进行安全评价

B

下面是一道填空题，请在括号中填上合适的文本内容。填空题：按照《安全生产法》的规定，()对全国安全生产工作实施综合监督管理。

国务院

你好，你能够做判断题吗？

是的，我能够做判断题。请问有什么可以帮助您？

这句话：“生产经营单位必须制定本单位生产安全事故应急救援预案。”正确还是错误？

正确


What is up?

请问您能提供有关黑河市兴边矿业有限公司事故的信息吗？


R 2023年6月8日6时，黑河市兴边矿业有限公司东部一区二段底分层采煤工作面发生顶板事故，造成1人死亡，直接经济损失182.33万元。事故直接原因是工作面放炮崩歪单体液压支柱，工人在空顶情况下违章打设支柱，冒落的岩石砸倒支柱，支柱砸伤其头部致死。事故间接原因包括：1. 技术管理存在不足，如对复合顶板管理技术认知不足，作业规程贯彻不认真等；2. 安全管理混乱，如违章实施爆破作业，未执行“一炮三检”和“三人连锁爆破”制度，未按规定给井下作业人员配发人员位置监测识别卡等；3. 现场安全检查责任不落实，现场安检员和带班安全矿长对于存在的违章行为未及时发现并制止；4. 职工安全培训教育不到位，现场作业人员安全风险辨识能力不足，安全意识不强。


What is up?

 2023年3月22日子长县中达焦家沟煤业有限公司发生了什么事故？

 2023年3月22日2时56分左右，子长县中达焦家沟煤业有限公司5309轨道顺槽发生其他事故，造成1人死亡，直接经济损失175万元。

 事故的直接原因是什么？


 事故的直接原因是矿井将不符合公称压力强度要求的管路作为瓦斯抽放管使用，在对瓦斯抽放管进行气密性试验的过程中瓦斯抽放管连接法兰处漏气，现场作业人员违反安全技术措施在瓦斯抽放管未泄压完毕的情况下用梅花扳手紧固漏气的法兰螺丝，维修过程中管路受力失衡爆裂，作业人员被喷出的冲击波击中，导致事故发生。

 事故的间接原因有哪些？

 事故的间接原因包括：

1. 安全生产责任落实不力，管理制度执行不严格。
2. 现场安全管理差，违章冒险作业。
3. 技术管理有漏洞，技术人员业务素质低。
4. 安全教育培训效果差，职工安全意识淡薄。
5. 安全监管存在不足。

 如何改进企业的安全生产责任落实？

 改进企业的安全生产责任落实可以采取以下措施：

1. 明确各级管理人员和员工的安全生产职责，确保每个岗位都有相应的安全管理制度和操作规程。
2. 加强对安全生产责任的培训和落实，使员工充分了解和掌握安全生产知识，严格遵守安全生产规定。
3. 定期进行安全生产检查，发现问题及时整改，确保安全生产责任落实到位。



如何加强现场安全管理，避免违章冒险作业？



加强现场安全管理，避免违章冒险作业可以采取以下措施：

1. 安排专职安全管理人员跟班，实时监督现场作业。
2. 定期进行安全检查，及时发现并制止违章冒险作业行为。
3. 制定并严格执行安全操作规程，确保每一步作业都符合规程要求。
4. 加强安全培训，提高员工安全意识和风险辨识能力。
5. 建立健全安全管理制度，确保安全生产责任落实到位。



如何提高技术管理水平？



提高技术管理水平可以采取以下措施：

1. 加强安全技术规范的学习和掌握，确保技术人员了解相关标准和要求。
2. 定期进行安全技术改进和创新，提出新的安全技术措施。
3. 加强现场作业经验的总结和反思，发现并改进技术管理中的不足。
4. 组织专家评审，确保技术措施的可行性和风险评估。
5. 确保技术措施的落实和监督，及时发现并制止违章操作行为。



感谢您的详细解答！



不客气，很高兴能帮助到您。如果您还有其他问题，欢迎继续提问。