

# 标注需求模板

## 标注内容

|            |                               |
|------------|-------------------------------|
| 标注目的       | 用于 Rag 相关任务训练                 |
| 标注类型       | 文本改写                          |
| 标注数据集描述    | 大语言模型用户输入的问题                  |
| 待标注目标      | 判断问题是否属于时效性问题，如果是则改写问题，添加时间约束 |
| 标注的数据规模    | 18097                         |
| 标注工具       | Excel                         |
| 交付数据格式     | Excel                         |
| 后续是否还有相同任务 | 有                             |
| 精度         | 99%                           |
| 重要性        | 重要                            |

# 标注要求

---

## 标注规则及示例

- 判断 `query` 是否包含时效性约束，例如：`xxx 最新的产品`、`今天 xxx 发生了什么`、`xxx 的价格`等，这种问题会随着时间的变更发生变化，用户更想要的是一个最新的信息
- 如果是时效性约束问题，就根据 `ref_time` 字段所指定的时间来改写问题，假定该时间是当前日期。
- 一般是添加 `xxxx` 年；一些更关心具体时效的就是添加 `xxxx` 年 `x` 月，涉及到周的日期也是改写到月就可以；只有具体 `query` 提到了今天、昨天、明天这种具体日期的，才需要添加 `xxxx` 年 `x` 月 `x` 日；如果 `query` 提到的是一个时间范围，比如近几年的情况，就不需要改写
- 注意标注时主要考虑以下类型需要改写：
  1. **天气**：今天天气怎么样、明天有雨吗
  2. **价格**：今日油价、今天的黄金价格、广州的房价、`xxx` 的股价等各种强时效性的价格，一般指定了今天昨天明天等具体日期的需要添加年月日，指定了模糊时间的，比如最近、近期、本周、上周、这几天等需要添加年月，什么时间相关信息都没有的只需要添加到年，比如房价、手机价格等
  3. **时事新闻类**：最近的热点事件、俄乌进展、`openai` 发布什么新产品、`google` 大会发布了什么等等没有指定时间，但其实是需要时效约束的问题，具体添加时间的规则 and 价格一致
  4. **时效性事实**：随着时间改变的事实类问题，比如 `xxx` 公司的 `ceo` 是谁、`xx` 市的市长是谁、今年的高考时间、`xxx` 证书考试的时间、中国的人口数、中国的 `GDP` 等
  5. **排名**：带有时效性的排名类问题，`NBA` 季后赛 `xx` 队和 `xx` 队比分、`xx` 羽毛球世界排名第几、`xxx` 公司是世界 `500` 强吗等
  6. **不完整时间描述**：对于不完整的时间描述要补全为标注 `xxxx` 年 `x` 月 `x` 日 的格式，比如 `23 年腾讯财报`，改写为 `2023 年腾讯财报`
- 注意以下例子不需要改写：
  1. **年龄**：某某多少岁、多大了，问年龄类的问题不需要改写
  2. **隐含时效的热点问题**：介绍一下胖猫事件、中信银行宁波分行为什么被罚款、雷军辟谣高考分数等，没有明显时间提示的近期热点事件，这种不需要标注时间
  3. **问题中有完整的时间描述**：比如 `2024 年 6 月` 发生了啥，`xx` 公司 `2023 年` 财报等等。这种已经明确指定了日期时间信息的不需要改写。
  4. **非查询类问题**：比如人设对话、闲聊、逻辑推理、数学、代码等 `query` 不需要改写，仅改写对于事实类的查询问题。
- **特殊**：标注文件中已经包含了预标注，需要标注同学判断是否标注正确，如果不正

确要手动修改，如果不需要改写的数据就留空就可以。

| B                          | C              | D  |
|----------------------------|----------------|--|
| query                      | ref time       | time_rewrite   |
| 今年济南市计划完成哪些城市道路建设任务？       | 1938年9月6日星期二   | 1938年济南市计划完成哪些城市道路建设任务？  |
| 克尔曼沙赫的人口在过去几年里有何变化？        | 1981年6月14日星期日  |  |
| 佛登、上海所债券托管数据的净增量是多少？       | 1989年7月13日星期四  | 截止1989年7月佛登、上海所债券托管数据的净增量是多少？                                  |
| 小皮奎在9月10日第二次发表声明给国际汽联后，国   | 1932年9月13日星期二  | 小皮奎在1932年9月10日第二次发表声明给国际汽联后，国际汽联主席马克斯·莫斯利如何回应？                 |
| 2年白酒暴利茅台的毛利率超过了多少？         | 2021年7月28日星期二  | 2021年白酒暴利茅台的毛利率超过了多少？  |
| 这次政治局会议提出了哪些重要的政策信号？       | 1948年11月22日星期一 | 1948年政治局会议提出了哪些重要的政策信号？  |
| 1年内英总产量增长了多少？              | 1968年12月10日星期二 | 1967年肉英总产量增长了多少？   |
| 金融街的董事会换届完成了吗？             | 1917年1月2日星期二   | 截止1917年1月2日金融街的董事会换届完成了吗？                                      |
| 习近平在澳门回归5周年大会上发表了什么样的讲话    | 2025年9月17日星期三  | 习近平在2014年澳门回归15周年大会上发表了什么样的讲话？                                 |
| 住建部今年的主要工作重点是什么？           | 1953年2月27日星期五  | 1953年住建部的主要工作重点是什么？  |
| BATJ 今年以来的股价表现如何？          | 1997年12月18日星期四 | 1997年BATJ的股价表现如何？  |
| 今年渔民捕鱼有什么变化？               | 2024年8月7日星期三   | 2024年渔民捕鱼有什么变化？  |
| 石嘴山市今年的主要任务是什么？            | 2000年2月18日星期五  | 2000年石嘴山市的主要任务是什么？   |
| 过去的三个月，左右废塑料市场的那只手到底是什么    | 1965年7月24日星期六  | 1965年4月至6月，左右废塑料市场的那只手到底是什么？                                   |
| 今年计划执行情况如何？                | 1946年7月4日星期四   | 1946年计划执行情况如何？   |
| 年初理财类理财产品赎回的原因是什么？         | 2009年1月13日星期二  | 2009年初理财类理财产品赎回的原因是什么？   |
| 6日的专家会诊结果是什么？              | 2002年5月23日星期四  | 2002年5月6日的专家会诊结果是什么？   |
| 公司三季度新开店数量是多少？             | 1993年10月10日星期日 | 1993年第三季度公司新开店数量是多少？   |
| 今年，淘大更成立了什么类型的工厂？          | 1915年8月14日星期六  | 1915年淘大更成立了什么类型的工厂？  |
| 请问昨天晚上发生了什么事情？             | 1948年2月27日星期五  | 1948年2月26日晚上发生了什么事情？   |
| 中信证券预计4月MLF和LPR利率将迎来等幅下降，后 | 2007年5月14日星期一  | 中信证券预计2007年4月MLF和LPR利率将迎来等幅下降，后续存款基准利率的调降仍然可期，是否意味着央行将进一步降准操作？ |
| 3年前澳门青州非法危险品仓库爆炸的原因是什么？    | 1916年7月4日星期二   | 1913年澳门青州非法危险品仓库爆炸的原因是什么？                                      |
| 五大连池赛特利牛排原料近期价格是多少？        | 1998年8月13日星期四  | 1998年8月五大连池赛特利牛排原料价格是多少？                                       |
| 油价周二走低的原因是什么？              | 1945年10月27日星期六 | 1945年10月24日油价走低的原因是什么？   |
| 本月事业方面的运势如何？               | 2028年1月22日星期六  | 2028年1月事业方面的运势如何？  |
| 胡蜂号在8月6日进行了什么样的飞行演练？       | 1992年9月27日星期日  | 胡蜂号在1992年8月6日进行了什么样的飞行演练？                                      |

## 注意事项

[ 所有上面没有提到的注意事项都可以列在这里 ]

## 文档更新记录

| 日期         | 版本号  | 更改人 | 更新内容 |
|------------|------|-----|------|
| 06/12/2024 | V1.0 | 伍文成 | 创建文档 |

## 标注须知：