



聊个工作流



想一想

我需要什么

- 确定好需要的结果
- 需要 json 格式
- 需要题目和判分规则
- 需要难一点的题目
- 需要50题
- 需要题目差异较大

大模型的弱点有哪些

- 长文本输出
- 逻辑不清晰
- 忽略细节
- 处理前后不一

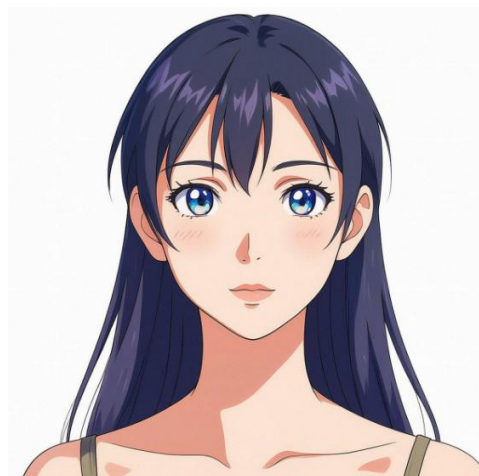
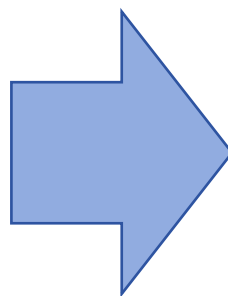


可以抽象成什么 workflows



具体化任务

1. workflows 每一步需要的结果是什么，需要如何做(CoT)，有没有好的示例
2. workflows 每一步是否耦合太多，每一步需要大模型什么能力，如何拆分不同的步骤使其更可控
3. workflows 每一个步骤之间需要哪些信息交互，信息的传递如何才能更高效



聊一聊

与大模型交流

生成第一个可用结果

目标：快速将想法转化为一个符合格式要求、可被评测的初始题目。

观察

1. 熟悉模型能力边界：基于一个核心点子，不考虑完美性，先构建一个包含多轮对话和多个约束条件的JSON结构。

2. 针对缺陷添加提示：运行API测试，观察模型输出。针对模型暴露的弱点（如忽略某个约束、格式错误），在Condition中强化约束描述或修改指令表述。

3. 同大模型对齐任务理解：观察自己的理解和大模型理解的是否相同。

主题约束：限定输出内容必须围绕给定主题进行

📌 结构流程约束

结构约束：规定输出的组织结构（分几部分、有无标题等）

流程约束：控制输出内容的先后顺序和执行业务

边界约束：定义系统执行任务时的能力边界和行为规范

其他约束：不属于上述分类的复杂或非典型约束要求

请生成一个测试集 必须包含 流程约束 和 排版约束 并使用复杂输出结构, 仅仅输出结果

json

复制

```
{
  "id": 3,
  "messages": [
    {
      "role": "user",
      "content": "你是一位专业的学术论文撰写助手，请按照以下要求生成一篇关于人工智能伦理的学术论文：\n\n## 论文结构要求\n1. 必须包含：
    }
  ],
  "condition": [
    {
      "msg_index": 1,
      "constraint_type": "流程约束",
      "constraint_detail": "必须严格按五步流程执行：大纲规划-摘要引言-三个章节-结论参考文献-最终整理，每个步骤后标注【步骤完成】，不得
    },
    {
      "msg_index": 2,
      "constraint_type": "排版约束",
      "constraint_detail": "必须使用LaTeX数学格式（如 $E=mc^2$ ），二级三级Markdown标题，关键词语加粗，项目符号列表，完整的学术论文排
    },
    {
      "msg_index": 3,
      "constraint_type": "结构约束",
      "constraint_detail": "必须包含摘要、引言、三个章节（每章2-3小节）、结论、参考文献的完整论文结构，层次分明"
    }
  ]
}
```

🔍 搜索

AI搜索

AI阅读

AI写作

AI画图

有问题，尽管问，shift+enter换行

Hunyuan

T1·深度思考

联网搜索

🔍

与大模型交流

模块测试

目标：将复杂题目生成任务拆解为独立模块，逐一验证每个约束条件的有效性。

观察

1. 观察每个模块的输入输出是否符合自身认知，正确，如判分模块
2. 观察多轮对话大模型对于提示词的细节理解

TIPS

1. 较好的结果用来做few-shot
2. 使用混元模型的过程中可以积累评分维度（哪里不满意）

```
"condition": [
  {
    "msg_index": 1,
    "constraint_type": "流程约束",
    "constraint_detail": "必须严格按五步流程执行：大纲规划→摘要引言→三个章节→结论参考文献→最终整理，每个步骤后标注【步骤完成】，不得跳步或逆序"
  },
  {
    "msg_index": 2,
    "constraint_type": "排版约束",
    "constraint_detail": "必须使用LaTeX数学格式（如 $E=mc^2$ ），二级三级Markdown标题，关键词语加粗，项目符号列表，完整的学术论文排版结构"
  },
  {
    "msg_index": 3,
    "constraint_type": "结构约束",
    "constraint_detail": "必须包含摘要、引言、三个章节（每章2-3小节）、结论、参考文献的完整论文结构，层次分明"
  }
]

使用以上规则对之前生成的内容打分，按照
"""
json
[[{"id":int,"score":0/1},...]
"""
的json格式输出
```

```
json
[
  {
    {"id": 1, "score": 1}, // 严格按五步流程执行（大纲+摘要引言+三章+结论参考文献+整理），每步标注【步骤完成】
    {"id": 2, "score": 1}, // 规范使用LaTeX公式（如 $\frac{d}{dx}f(x)$ ），二级/三级标题，关键词语加粗，项目符号列表
    {"id": 3, "score": 1} // 完整包含六大部分（摘要/引言/三章/结论/参考文献），层次清晰且符合学术规范
  }
]
```

AI搜索 AI阅读 AI写作 AI画图

有问题，尽管问，shift+enter换行

Hunyuan T1·深度思考 联网搜索

分享 收藏 发送

由智谱AI生成 内容安全

做一做

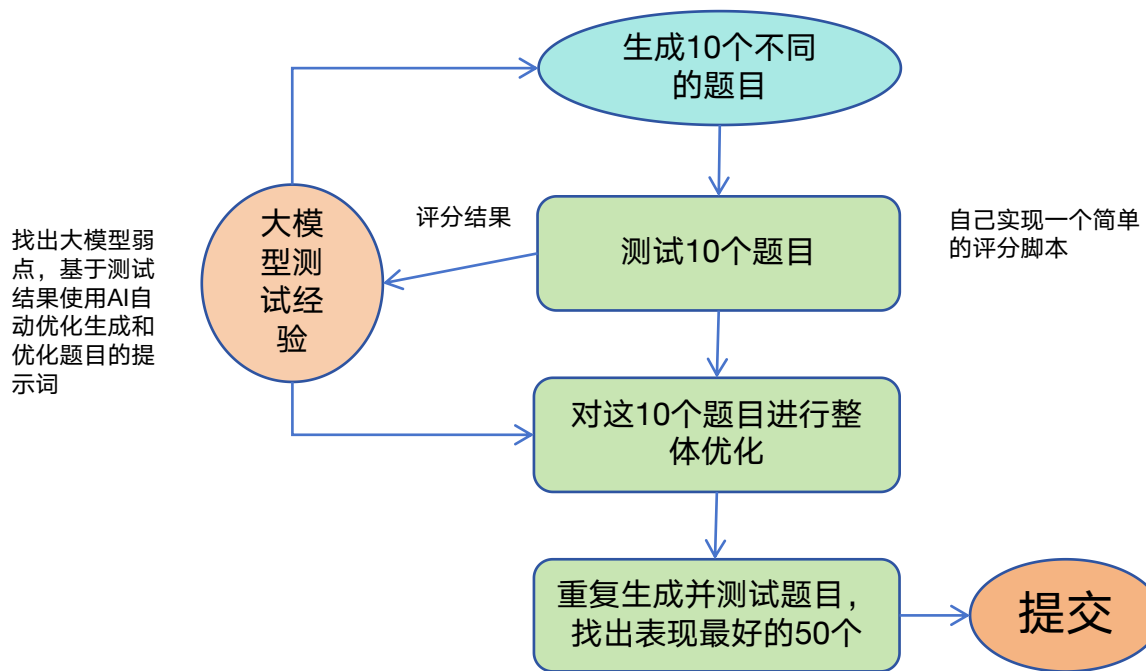
做成 workflow

完成最小MVP

使用代码或者其他 workflow 工具，将聊天串成一个工作流。

**雕花

更细致的控制提示词，增加few-shot，堆砌更多流程。



优化上面的测试集，其中 messages -> content 字段必须严格约束大模型输出，使用更多的约束条件，content 应该详细而具体。

并且上面的测试集修改的更加复杂，修改方法为：

1. 输出更多更长的json内容，如输出3个->10个，2. 更复杂的json结构
3. 增加更多的约束条件，例如不能出现xx字，4. 更严格的判分条件，想办法使得每个判分条件更加严格

messages 必要时改为多轮对话，

```
messages -> [{"role": "user", "content": "XXX"}, {"role": "assistant", "content": "XXX"}, {"role": "user", "content": "XXX"}]
```

但是最后一个元素的role必须是user 仅仅输出结果判分。

```
json
{
  "id": 4,
  "messages": [
    {
      "role": "user",
      "content": "你是一位顶尖的学术研究专家，请按照最严格的标准生成一篇关于\"人工智能伦理框架构建\"的学术论文。以下是详细要求：

## 核心结构要求（必须完全遵守）
1. 必须包含10个固定组成部分：
- 摘要（200-250字严格限制）
- 引言（300-350字）
- 第一章：理论基础（必须包含3个子章节：1.1 定义与范畴/1.2 历史发展/1.3 核心挑战）
- 第二章：技术分析（必须包含4个子章节：2.1 算法偏见/2.2 数据隐私/2.3 自主决策/2.4 人机关系）
- 第三章：实践框架（必须包含3个子章节：3.1 伦理准则/3.2 治理模型/3.3 评估指标）
- 案例研究（2个真实案例，每个案例必须包含：背景/伦理冲突/解决方案/启示）
- 挑战与展望（必须讨论5个未来挑战）
- 结论（150-200字）
- 参考文献（至少15篇权威文献，MLA格式）

## 严格流程约束（必须逐步执行）
1. 【必须显示】首先输出完整论文大纲（包含所有章节字数分配）
2. 【必须显示】然后分步撰写：
- 先完成摘要（严格200-250字，多一字少一字都扣分）
- 接着完成引言（严格300-350字）
- 然后按顺序完成：第一章->第二章->第三章
- 接着完成案例研究（两个案例必须严格区分）
- 最后完成挑战展望-结论-参考文献
3. 每个部分完成后必须标注【步骤x完成：部分名称】
4. 不得跳过任何步骤或改变顺序

## 极端排版约束
1. 必须使用标准学术论文格式：
- 二级标题（##）用于章节
- 三级标题（###）用于子章节
- 四级标题（####）用于案例小节
2. 关键词必须同时满足：
- 加粗显示
- 首次出现时附加英文原词（如：算法偏见(algorithmic bias)）
3. 所有数学公式必须使用$$包裹的LaTeX格式
```

谢谢观看