```
前端参数
                         "query": str, # 用户输入
                         "user_id": str, #用户ID
                         "conversation_id": str, # 对话框ID
                         "conversation_name": str, # 对话框名称
                         "history_len": int, # 从数据库中取历史消息的数量
                         "history": Union[int, List[History]],# 历史对话,设为一个整数可以从数据库中读取历史消息
                         "stream": bool, # 流式输出
                         "model_name": str, # LLM 模型名称
                         "temperature": float, # LLM 采样温度
                         "max_tokens": Optional[int], #限制LLM生成Token数量, 默认None代表模型最大值
                         "prompt_name": str, # 使用的prompt模板名称
                         "db": AsyncSession # 数据库会话
ConversationModel
                       "ConversationModel": {
                         "__tablename__": "conversation", # 表名称
"id": "String(32)", # 会话ID
                         "user_id": "String(32)", # 用户ID
                         "name": "String(50)", # 对话框名称
                         "chat_type": "String(50)", # 聊天类型
                         "create_time": "DateTime", # 创建时间
                         "relationships": {
                           "user": "UserModel", # 会话与用户的多对一关系
                           "messages": "MessageModel" # 会话与消息的一对多关系
     UserModel
                      "UserModel": {
                      "__tablename__": "user", # 表名称
"id": "String(32)", # 用户ID
                      "username": "String(255)", # 用户名
                      "password_hash": "String(255)", # 密码的哈希值
                      "relationships": {
                      "conversations": "ConversationModel" # 用户与会话的一对多关系
   MessageModel
                       "MessageModel": {
                       "_tablename_": "message", # 表名称
                       "id": "String(32)", # 聊天记录ID
                        "conversation_id": "String(32)", # 会话ID
                       "conversation": "ConversationModel", # 聊天记录与会话的关系
                       "chat type": "String(50)", # 聊天类型
                       "query": "String(4096)", # 用户问题
                       "response": "String(4096)", # 模型回答
                       "meta_data": "JSON", # 记录知识库id等,以便后续扩展
                       "feedback_score": "Integer", # 用户评分
                       "feedback reason": "String(255)", # 用户评分理由
                       "create time": "DateTime" # 创建时间
```

