

## API接口描述

名称	类型	描述	请求参数	响应内容
/agent/invoke	POST	创建智能体并调用，直接返回结果或中断数据	request=AgentRequest( user_id: str query: str system_message: Optional[str] )	<p>完成状态:</p> <pre>response = AgentResponse(     session_id=session_id,     status="completed",     result=result )</pre> <p>中断状态:</p> <pre>response = AgentResponse(     session_id=session_id,     status="interrupted",     interrupt_data=interrupt_data )</pre> <p>失败状态:</p> <pre>error_response = AgentResponse(     session_id=session_id,     status="error",     message=f"处理请求时出错: {str(e)}" )</pre>
/agent/resume	POST	恢复被中断的智能体执行，等待执行完成或再次中断	request=InterruptResponse( user_id: str session_id: str response_type: str args: Optional[Dict[str, Any]] = None )	<p>完成状态:</p> <pre>response = AgentResponse(     session_id=session_id,     status="completed",     result=result )</pre> <p>中断状态:</p> <pre>response = AgentResponse(     session_id=session_id,     status="interrupted",     interrupt_data=interrupt_data )</pre> <p>失败状态:</p> <pre>error_response = AgentResponse(     session_id=session_id,     status="error",     message=f"处理请求时出错: {str(e)}" )</pre>
/agent/status/{user_id}	GET	获取当前用户的状态	user_id	<pre>response = SessionStatusResponse(     user_id=user_id,     session_id=session_id,     status=status,     last_query=last_query,     last_updated=last_updated,     last_response=last_response )</pre>
/system/info	GET	获取系统状态信息	-	<pre>response = SystemInfoResponse(     sessions_count=sessions_count,     active_users=active_users )</pre>
/agent/session/{user_id}	DELETE	删除用户会话	user_id	<pre>response = {     "status": "success",     "message": f"用户 {user_id} 的会话已删除" }</pre>

数据模型定义		
名称	描述	数据模型
AgentRequest	客户端发起的智能体请求	<pre>class AgentRequest(BaseModel):     # 用户唯一标识     user_id: str     # 用户的问题     query: str     # 系统提示词     system_message: Optional[str] = "你会使用工具来帮助用户。如果工具使用被拒绝, 请提示用户。"</pre>
AgentResponse	智能体给予的响应	<pre>class AgentResponse(BaseModel):     # 会话唯一标识     session_id: str     # 三个状态: interrupted, completed, error     status: str     # 时间戳     timestamp: float = Field(default_factory=lambda: time.time())     # error时的提示消息     message: Optional[str] = None     # completed时的结果消息     result: Optional[Dict[str, Any]] = None     # interrupted时的中断消息     interrupt_data: Optional[Dict[str, Any]] = None</pre>
InterruptResponse	客户端给予的反馈响应	<pre>class InterruptResponse(BaseModel):     # 用户唯一标识     user_id: str     # 会话唯一标识     session_id: str     # 中断响应类型     response_type: str     # 如果是edit, response类型, 可能需要额外的参数     args: Optional[Dict[str, Any]] = None</pre>
SystemInfoResponse	系统信息响应	<pre>class SystemInfoResponse(BaseModel):     # 当前系统内会话总数     sessions_count: int     # 当前活跃的用户     active_users: List[str]</pre>
SessionStatusResponse	会话状态信息响应	<pre>class SessionStatusResponse(BaseModel):     # 用户唯一标识符     user_id: str     # 会话唯一标识符     session_id: Optional[str] = None     # 状态: not_found, idle, running, interrupted, completed, error     status: str     # error时的提示消息     message: Optional[str] = None     # 上次查询     last_query: Optional[str] = None     # 上次更新时间     last_updated: Optional[float] = None     # 上次响应     last_response: Optional[AgentResponse] = None</pre>

中断响应类型	
名称	描述
accept	接受工具调用
reject	拒绝工具调用
edit	修改工具参数后调用工具
response	不调用工具直接反馈信息

会话状态	
名称	描述
not_found	未发现会话
idle	会话空闲
running	会话运行中
interrupted	会话已中断
completed	会话已完成
error	未知会话错误