

大富翁全链路排查机器人学习与技术沉淀

测试开发视角下的 Agent 排障能力、TTS 案例复盘与 3 天执行计划

一、学习目标

- 用 3 天建立大富翁排障使用方法，而不是停留在 Agent 概念理解。
- 能够复盘 TTS 朗读失败案例，并独立判断报告证据链是否完整。
- 掌握大富翁提问、追问、复核、沉淀的标准动作。

二、基础知识掌握边界

优先级	知识点	为什么要学	掌握标准	可跳过内容
1	LLM	理解大富翁如何理解问题、总结日志、生成报告。	知道模型输出可能不等于事实；结论必须回到工具证据。	模型训练、微调、推理加速。
2	Function Calling	理解自然语言如何变成结构化查询参数。	能看懂工具名、入参、时间范围、关键字段、返回结果。	模型内部协议适配。
3	Tool Calling	理解大富翁如何查询日志、链路、监控、工单、知识库。	能判断工具是否查对系统、参数是否准确、结果是否足够支撑结论。	工具框架源码和插件生命周期。
4	Agent	理解大富翁不是一次问答，而是多轮计划、查询、观察、归纳。	能画出 Agent Loop，并套用到 TTS 案例。	复杂 Agent 调度实现。
5	ReAct	排障过程本质是判断、查询、观察、修正、归因。	能把大富翁输出拆成 Thought/Action/Observation/Conclusion。	论文实验细节。
6	MCP	辅助理解工具标准化接入方式。	了解 Tools、Resources、Prompts 的概念即可。	MCP Server 开发和协议细节。
7	Multi-Agent	了解复杂排障可能拆成不同分析角色。	知道日志分析、监控分析、报告生成可以拆角色。	多 Agent 编排算法。

三、3 天详细执行计划

Day 1 上午：建立问题域与链路认知

目标：把 TTS 朗读失败拆成可排查的问题域，明确端侧、代理层、合成链路、模型侧的职责。

执行动作：

- 阅读大富翁 TTS 案例输出，标出所有关键实体：utdid、uid、chat_req_id、TTS sid、端类型、版本、设备、错误码。
- 画出 TTS 主链路：客户端点击朗读 -> tts-proxy 接收文本 -> audio-hub 启动合成 -> quark-voice 路由模型 -> 模型返回音频或错误 -> 客户端播放。
- 把案例中的系统名和日志名归类：客户端埋点、tts_ws_event、audio_hub、quark_voice_prod、tts-proxy。
- 整理 01:50 精确窗口和 13:47-13:50 命中窗口的差异，记录为什么两个窗口都要写进报告。

产出物：TTS 链路图一张。；字段词典一份：字段名、来源、含义、排障价值。；问题域拆解笔记：端侧问题、服务端问题、下游模型问题分别看什么证据。

验收标准：能不依赖原文，讲清 TTS 朗读从点击到失败的完整链路。

Day 1 下午：掌握 Agent 与工具调用的排障模型

目标：理解大富翁如何把自然语言排障请求转成工具查询，并认识模型输出与工具证据的边界。

执行动作：

- 把原始问题拆成结构化入参：utdid、日期、时间点、时间窗口、功能域、失败现象、排查范围。

- 推演大富翁可能调用的工具：查询端侧 trace、查询 tts_ws_event、查询 audio_hub、查询 quark_voice_prod、查询权限覆盖范围。

- 用 ReAct 表格复盘：判断是什么、调用什么、观察到什么、下一步为什么改变。

- 记录大富翁报告中的事实、推断和边界：事实来自日志，推断来自错误链路，边界来自 0 命中和权限限制。

产出物：一张 ReAct 复盘表。；一份工具调用推演表：工具、入参、输出、用于证明什么。；一份事实/推断/边界三列表。

验收标准：能解释大富翁为什么先查 01:40-02:00，再转向 13:47-13:50，并能指出这不是随意扩大范围，而是基于 0 命中后的复核动作。

Day 2 上午：复盘真实证据链

目标：把客户端、tts-proxy、audio-hub、quark-voice 的证据串成同一条失败链路。

执行动作：

- 从客户端证据开始：记录

chat_tts_kmp_synthesize、err_code=5000005、stream_terminal_reason=synth_error、received_audio_chunks=0。

- 从 tts-proxy 证据确认文本进入服务端：tts.text.commit、client.connection.error code=5000005、on_close reason=header_code_error。

- 从 audio-hub 证据确认合成开始后失败：SynthesisStarted 成功、SentenceBegin 成功、SentenceEnd NE003、SynthesisCompleted NE003。

- 从 quark-voice 证据确认下游模型侧根因：tts service overloaded:100032、retry exhausted attempts:3、finalSuccess:false、sessionSuccess:false。

- 校验这些证据是否能通过 chat_req_id 或 TTS sid 对齐到同一会话。

产出物：端侧到服务端证据链表。；关键日志摘录清单。；会话一致性校验记录：哪些字段用于串联，哪些字段不能单独作为根因依据。

验收标准：能说明为什么根因不是客户端播放失败，也不是用户主动停止，而是服务端下游 TTS 模型负载过高。

Day 2 下午：建立质量复核标准

目标：从测试开发视角定义大富翁报告是否可信的验收标准。

执行动作：

- 检查时间：是否同时说明 01:40-02:00 0 命中和 13:47-13:50 命中链路。

- 检查覆盖：是否列出千问 Android、夸克 Android、千问 iOS、夸克 iOS，以及 qwen-harmony 权限边界。

- 检查链路：客户端、tts-proxy、audio-hub、quark-voice 是否都出现，且顺序合理。

- 检查错误码：100032、NE003、header_code_error、5000005 是否形成传播链。

- 检查排除项：是否明确排除客户端播放失败、用户主动停止、单端误判。

- 检查报告用语：是否区分“已确认”“高度匹配”“无法覆盖”“需要补充”。

产出物：大富翁报告质量检查表。；错误码传播链图。；一份可复用的排障报告评分标准。

验收标准：能对任意一份大富翁排障报告做质量评审，指出证据不足、范围不清或结论越界的问题。

Day 3 上午：形成可复用提问与验证方法

目标：沉淀如何向大富翁提问，才能得到可验证、可复核、可交付的排障报告。

执行动作：

- 设计标准提问模板：设备/用户标识、时间窗口、端类型、功能域、失败现象、期望查询范围。

- 设计追问模板：要求补查某个服务、扩大时间窗口、补充某个 req_id/sid、说明权限边界。

- 设计验证模板：要求大富翁按时间线、证据链、错误码映射、排除项重写结论。

- 把 TTS 案例改写成 3 个不同质量的问题，对比大富翁可能输出差异。

产出物：标准提问模板 4 条。；追问模板 4 条。；报告重写/校验模板 3 条。；低质量问题与高质量问题对比表。

验收标准：能通过更准确的问题输入提升大富翁输出质量，并能通过追问补齐报告缺口。

Day 3 下午：完成沉淀文档与汇报材料

目标：把学习结果转化为可交付材料，能够用于团队内复盘和后续排障复用。

执行动作：

- 整理 1 页架构说明：大富翁如何通过 Hermes Agent 编排模型和工具完成全链路排障。

- 整理 1 页案例复盘：现象、时间线、链路、证据、根因、排除项、边界。
- 整理 1 页质量标准：查询范围、证据链、错误码、覆盖边界、报告措辞。
- 准备 5-10 分钟汇报提纲：学习内容、TTS 案例、方法沉淀、后续优化建议。

产出物：最终技术沉淀文档。； Mermaid
全链路流程图。； 汇报提纲。； 后续问题清单：哪些工具权限、字段说明、错误码文档还需要补齐。
验收标准：能完整讲清大富翁的使用方法、TTS 案例证据链、报告质量校验标准和后续改进点。

四、TTS 案例 ReAct 复盘

原始问题：@大富翁 全链路排查分析 utdid=ag2z+YxIB5kDAHd8mikGfdN1 在今天 01:50 左右TTS朗读失败的问题

阶段	本案例动作/观察	排障价值
Thought	01:50 左右 TTS 失败，需要先按设备和时间精确查端侧与服务端链路。	避免在没有时间范围的情况下泛查日志。
Action	按北京时间 01:40-02:00 查询干问 Android、夸克 Android、干问 iOS、夸克 iOS 和服务端 TTS 链路。	确认用户给定时间窗口是否存在可归因链路。
Observation	精确窗口端侧和服务端均 0 命中，qwen-harmony 无应用权限。	需要保留 0 命中和权限边界，不能直接给根因。
Thought	同设备可能存在时间语义差异或相邻窗口失败记录，需要复核。	线上排障要处理用户口述时间与日志时间不一致问题。
Action	查询 13:47-13:50 同设备 TTS 相关记录。	寻找高度匹配真实链路。
Observation	命中客户端 5000005、audio-hub NE003、quark-voice overloaded:100032。	形成错误传播链。
Conclusion	服务端下游 TTS 模型负载过高，导致客户端未收到音频分片并感知朗读失败。	结论来自多源证据闭环。

五、数据源与证据链

数据源	查询对象	关键证据	证明内容
客户端	chat_tts_kmp_synthesize / span detail	err_code=5000005; stream_terminal_reason=synth_error; received_audio_chunks=0	用户感知失败且未收到音频分片
tts-proxy	tts_ws_event / connection error / on_close	收到 tts.text.commit; client.connection.error code=5000005; on_close reason=header_code_error	文本进入服务端链路，错误返回客户端
audio-hub	SynthesisStarted / SentenceBegin / SentenceEnd / SynthesisCompleted	SentenceEnd NE003 / tts response failed, 服务负载过高; SynthesisCompleted NE003 / tts infer failed	合成启动成功，失败发生在下游推理返回阶段
quark-voice	quark_voice_prod / retry log	tts service overloaded:100032; retry exhausted attempts:3; finalSuccess:false; sessionSuccess:false	根因指向下游 TTS 模型负载过高

六、报告质量验收标准

检查项	标准	本案例样例	质量价值
时间窗口	必须同时说明精确窗口和命中窗口	01:40-02:00 0 命中; 13:47-13:50 命中	避免错误归因
端覆盖	列明已覆盖端和未覆盖端	四端已查; qwen-harmony 无权限	避免过度结论
会话一致性	使用 chat_req_id / sid 串联	8757f... / 4f5a...	避免跨会话拼接

检查项	标准	本案例样例	质量价值
错误传播	根因码到端侧码闭环	100032 -> NE003 -> header_code_error -> 5000005	证明因果链
排除项	说明为什么不是其他原因	received_audio_chunks=0; 主链路为 synth_error	提升结论可信度
报告边界	说明无法覆盖或需补充信息	qwen-harmony 权限; 如确认凌晨需补充端类型/req_id	保证可追溯

七、可复用提问模板

场景	提问模板
标准排查	@大富翁 全链路排查分析 utdid=【设备ID】在【北京时间时间范围】出现【功能失败】的问题，请覆盖端侧、tts-proxy、audio-hub、quark-voice，并输出证据链。
时间复核	@大富翁 请先确认【精确时间窗口】是否命中；若 0 命中，请复核同设备前后【N】分钟及可能的时间语义差异，并明确写出窗口差异。
证据链校验	@大富翁 请按 chat_req_id / sid 串联客户端、tts-proxy、audio-hub、quark-voice，输出每一层关键日志和错误码映射。
排除项追问	@大富翁 请说明本次是否可排除用户主动停止、客户端播放失败、单端异常，并给出对应证据。
报告重写	@大富翁 请按基本信息、时间线、失败链路、错误码传播、根因、排除项、覆盖边界重写报告。

八、技术沉淀结论

- 学习计划的核心不是泛学 Agent 概念，而是围绕大富翁真实排障输出建立可执行的验证方法。
- TTS 案例的关键质量点是时间窗口复核、端覆盖范围、会话一致性、错误码传播和排除项证明。
- 测试开发应把大富翁输出作为证据聚合入口，而不是无条件接受结论；报告必须能被日志、链路和错误码复核。
- 3 天结束后应产出字段词典、TTS 链路图、ReAct 复盘表、质量检查表、提问模板和最终沉淀文档。