

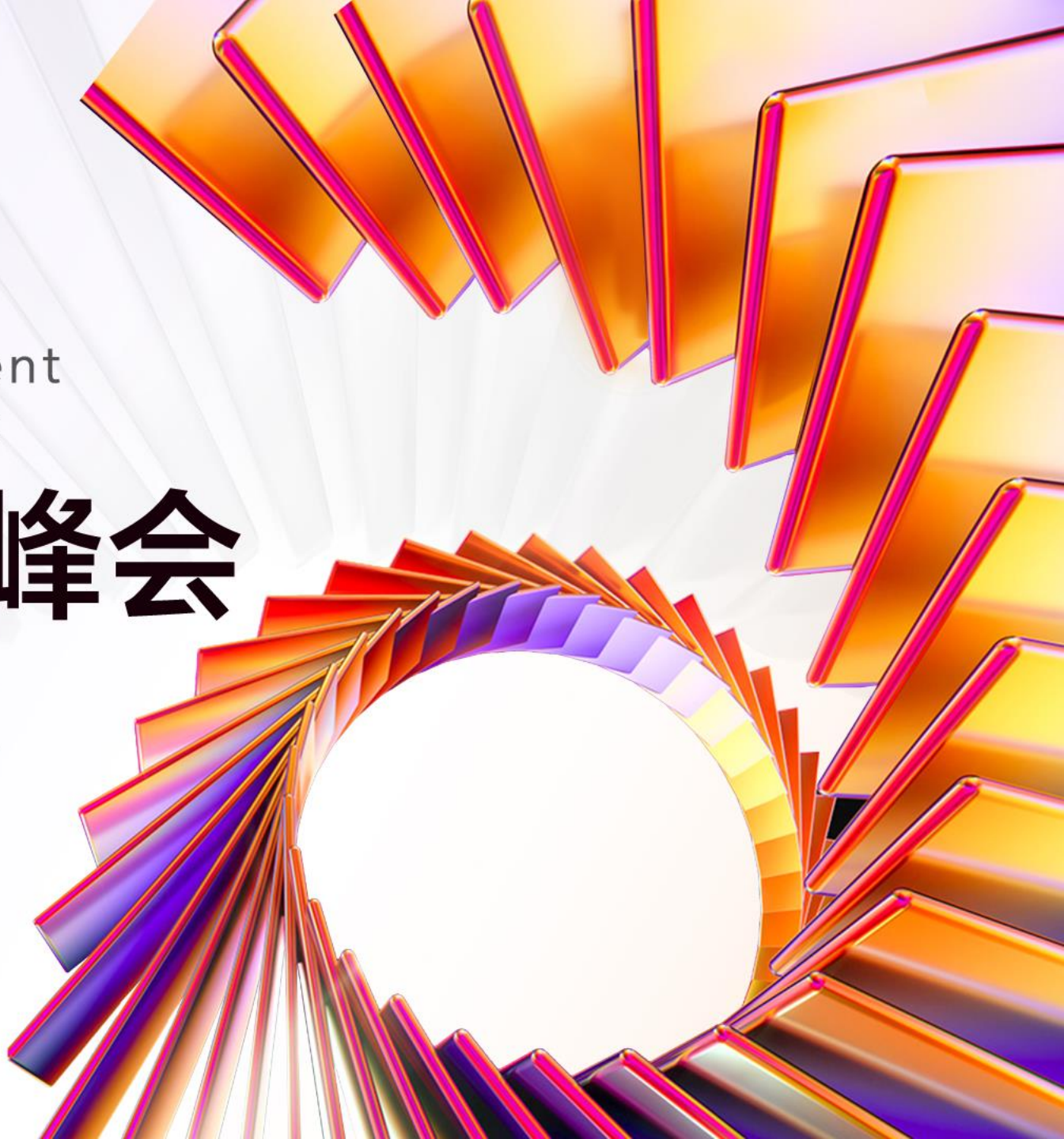


# 第8届 AI+ Development Digital Summit

## AI+研发数字峰会

拥抱AI 重塑研发

11月14-15日 | 深圳





# 2026 AI+ PRODUCT INNOVATION SUMMIT

01.16-17 · ShangHai

## AI+产品创新峰会



### Track 1: AI 产品战略与创新设计

从0到1的AI原生产品构建

论坛1: AI时代的用户洞察与需求发现

论坛2: AI原生产品战略与商业模式重构

论坛3: Agentic AI产品创新与交互设计

### 2-hour Speech: 回归本质



用户洞察的第一性

——2小时思维与方法论工作坊

在数字爆炸、AI迅速发展的时代，  
仍然考验“看见”的“同理心”



### Track 2: AI 产品开发与工程实践

从1到10的工程化落地实践

论坛1: 面向Agent智能体的产品开发

论坛2: 具身智能与AI硬件产品

论坛3: AI产品出海与本地化开发

### Panel 1: 出海前瞻



“出海避坑地图”圆桌对话

——不止于翻译:

AI时代的出海新范式



### Track 3: AI 产品运营与智能演化

从10到100的AI产品运营

论坛1: AI赋能产品运营与增长黑客

论坛2: AI产品的数据飞轮与智能演化

论坛3: 行业爆款AI产品案例拆解

### Panel 2: 失败复盘



为什么很多AI产品“叫好不叫座”?

——从伪需求到真价值:

AI产品商业化落地的关键挑战

智能重构产品 数据驱动增长  
Reinventing Products with Intelligence, Driven by Data

80+

业界大咖

汇聚产品大咖的真知灼见

40+

创新案例

开拓全新AI+产品视角

10+

分论坛

涵盖产品到商业增长的全链路

800+

听众规模

共同探索产品的智能重构



扫码查看会议详情

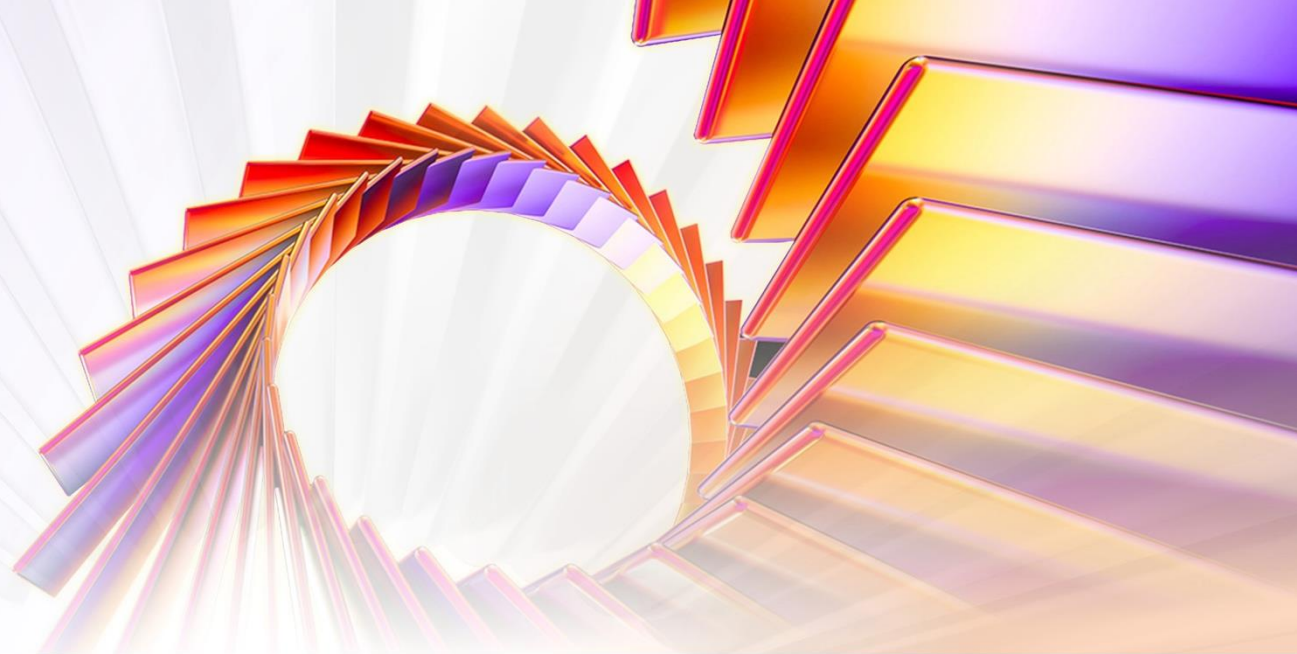


| 11月14-15日 | 深圳

第8届 AI+ Development  
Digital Summit

AI+研发数字峰会

拥抱AI 重塑研发



# 研发交付数字员工的上下文工程和数 据飞轮建设实践

王笑吉 | 百度



## 王笑吉

百度 资深工程师

---

百度MEG团队资深测试开发工程师，8年从业经验，目前主要负责百度FEED检索业务的交付质效保障相关工作，探索如何利用大模型能力提升研发交付效能，带领团队基于AI数字员工完成了交付方式从GUI向LUI转换，探索数据飞轮机制，提升AI员工的智能化程度，从而为交付提效。

# 目录

## CONTENTS

- I. 交付质效痛点与AIQA数字员工实践
- II. 基于群聊的知识数据飞轮建设
- III. AIQA数字员工应用与交付提效实践
- IV. 回顾与总结

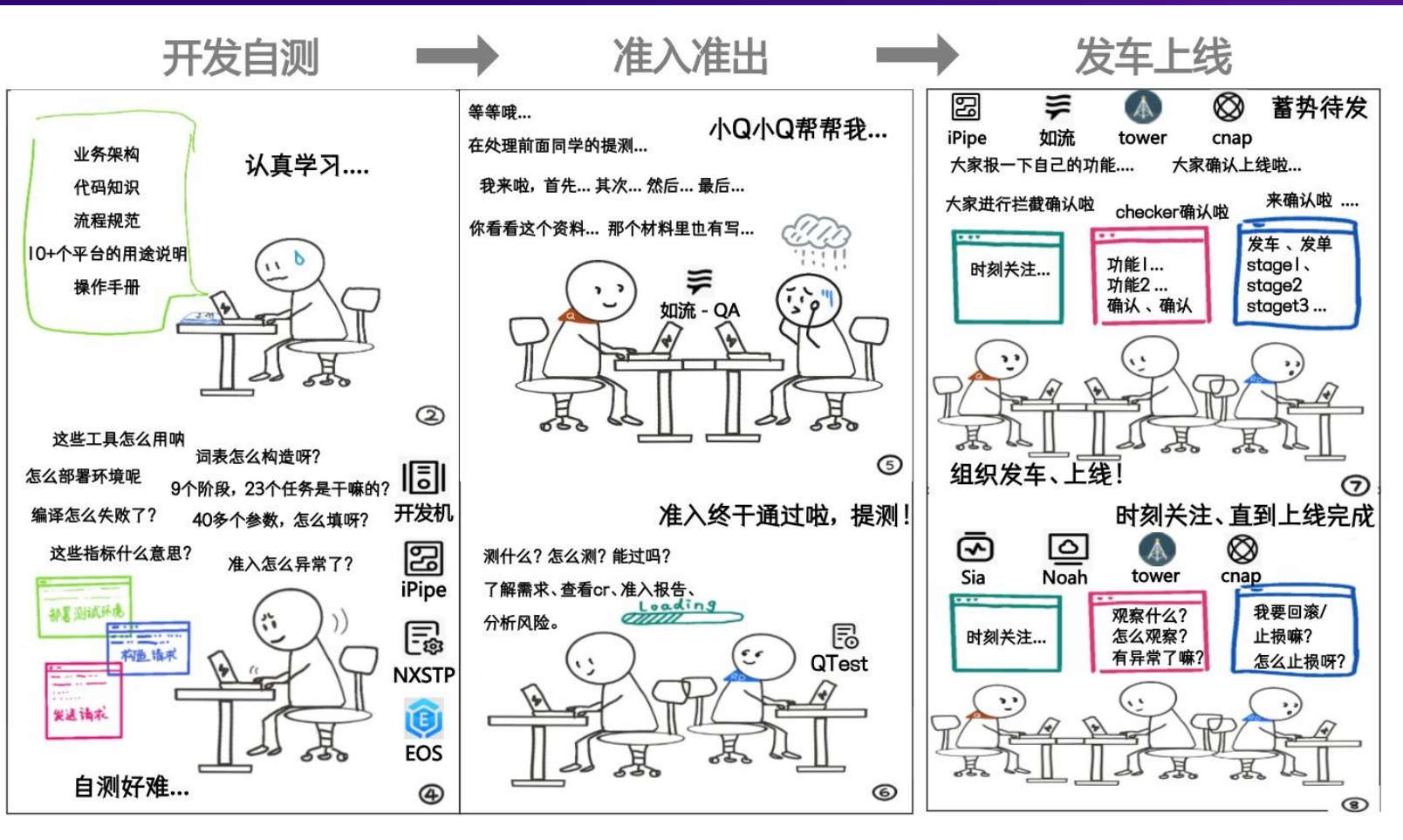
# PART 01

## 交付质效痛点与AIQA数字员工实践

# 项目交付质效痛点分析与AIQA数字员工建设背景

## 业务交付现状及特点:

项目多(周700+)、迭代快, RD自主测试为主、QA值班方式参与



交付效率遭遇瓶颈, 迭代速度无法突破

能否准出、上线决策?

全流程跨10+平台

依赖值班QA

9个阶段、20+任务

业务知识积累

数百页学习文档

测试交付经验沉淀

与多人、多方沟通

高效交付强依赖人

操作复杂、流程繁琐  
学习、沟通成本高

大模型驱动  
交付模式变革

人工QA -> AIQA

GUI交互->LUI交互

降低学习成本  
打破经验壁垒

使用自然语言实现交付  
操作和流程最简化

★ 可感知 的交付效能和交付体验提升

面向多平台操作

面向延迟高的人类

面向海量学习资源



24h私人助理  
AIQA数字员工

# AIQA数字员工的建设思路

建设思路

- 核心1 - 具备LUI交互能力，让RD可以与AIQA数字员工使用自然语言对话
- 核心2 - 覆盖全场景的懂业务、懂交付、会思考且有执行能力、会自主学习的AIQA交付专家

LUI交互

懂RD

+

会思考

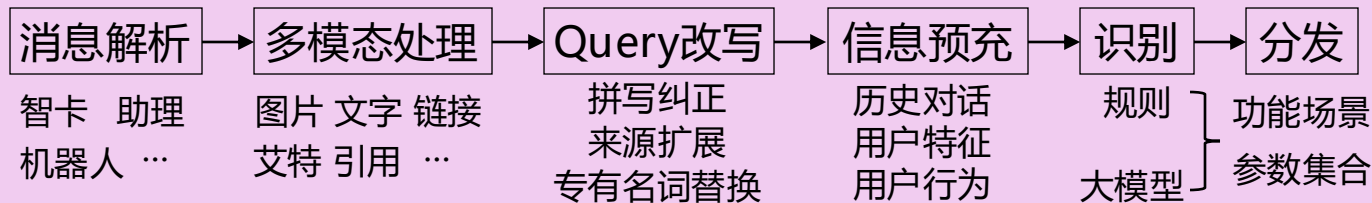
AIQA

能执行  
懂业务  
懂交付  
能学习

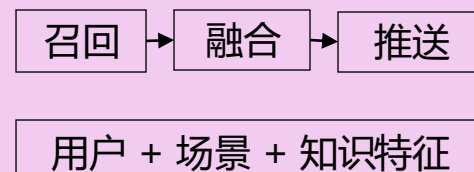


依托 AIQA数字员工 - 实现LUI自然语言完成交付全流程

意图识别 - LLM助力理解RD需求,实现全场景精准分发,参数精准识别

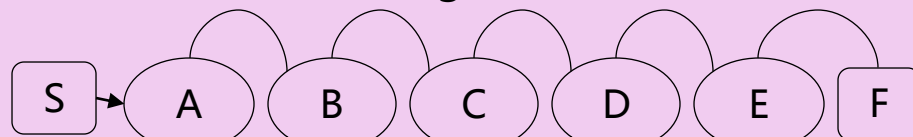


智能推荐 - 提升交互体验



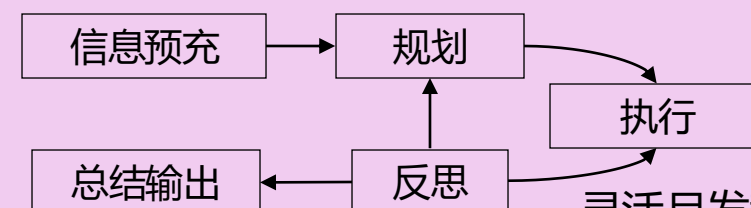
交付专家 - 双架构加持Multi-Agent, 满足多专家协作需求, 覆盖RD交付全场景

人规划的Agentic Workflow



支持需结构化按人预设步骤执行的场景

自主规划AI Agent



灵活且发散

工具组件化

工具MCP化

数据落盘

数据收集

数据处理

数据评估

数据应用

持续优化

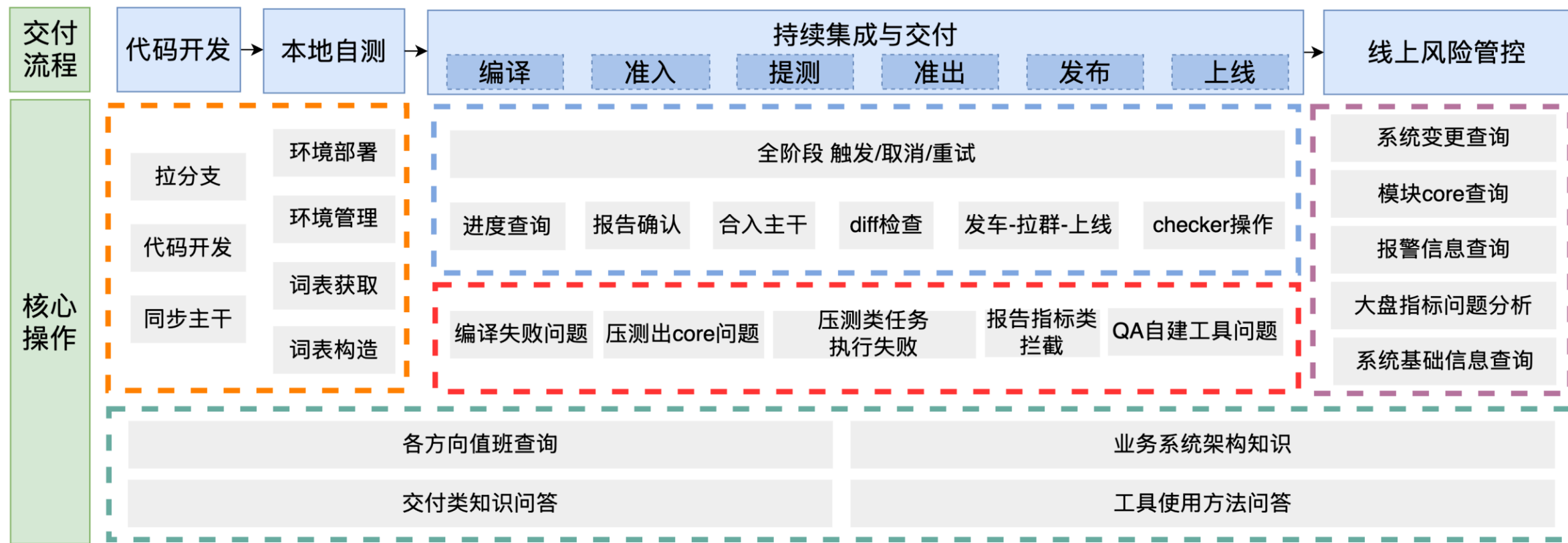
知识反哺



领域知识飞轮

# ▶ AIQA数字员工的建设思路-覆盖交付全阶段、全场景、全操作

从RD同学日常项目交付全过程的核心环节入手，总结数十种高频操作，归纳出五类核心Agent能力



★ 五类核心Agent，打通LUI交付全流程



# ▶ AIQA数字员工系统架构



## 数据工程

业务知识

测试知识

流水线构建

测试用例

交付知识

交付结果

业务指标

代码特征

业务术语知识

业务系统知识

需求描述

交互历史

质量数据

监控指标

用户画像

# ▶ AIQA数字员工系统框架中的3个核心上下文

1

意图识别

理解用户需求  
实现全场景精准分发

2

AI Agent规划器

从用户需求出发，  
制定一个可靠的、  
可高度并行的  
执行计划

3

AI Agent反思器

一个决策者，判  
断用户的需求是  
否已得到解决，  
该Finish还是该  
Replan



# AIQA数字员工系统中的核心上下文设计 – 意图识别

明确角色，让AI知道自己是谁

# [你的角色]

你是一名XX方向的研发交付数字员工。

明确能力边界，让AI知道自己能做什么

## [你支持的功能]

1. 流程推进

- 背景知识：用户的自动化交付流程在X平台(又称为流水线)上进行。流水线主要包含 ...共X个阶段。... ..
- Agent具体功能和对应的示例问题：
- < 子功能XXX>：{具体功能介绍}。{特殊约定}。\*对应的用户示例问题\*\*：XXX
- < 子功能XXX>：{具体功能介绍}。{特殊约定}。\*对应的用户示例问题\*\*：XXX

多功能Agent

2. 问题定位

- 背景知识：{背景知识描述} - 提示：{Agent内的特殊提示说明}
- Agent具体功能和对应的示例问题：{具体功能说介绍}。{特殊约定}。\*对应的用户示例问题\*\*：XXX

单功能Agent

## [你暂不支持的功能]

- <暂不支持>：目前暂时无法支持的功能，待升级。\*对应的用户示例问题\*\*：XXX

## [非功能类提问]

- <无>：非功能类型的提问或者无法识别的提问。\*对应的用户示例问题\*\*：XXX

# [对话上下文]

用户问题：XXXX / 系统回答：XXXX / 用户问题：XXXX

# [你的任务]

我会给你一段「对话上下文」，你需要根据对话上下文，完成如下任务：

1. 充分理解最后一轮用户问题，再考虑对话上下文内容，将对话上下文**改写**为一条用户对话内容
2. 基于前一步的改写结果，仔细思考本次**用户需求与哪一类匹配度最高**，你需要从XX选出...
3. 基于第一步的改写结果，从操作意图json中**选择适合的操作意图**。[操作意图json]

# [返回格式要求]

1. 必须是json格式，不包含推理过程或其他无关信息。
2. 请务必以如下格式返回：{"改写结果":"xx","type":"xx","意图结果":{"操作意图名称":"xx","意图澄清":"xx"}}

支持能力的具体描述与示例

支持

不支持

非功能

背景  
知识

具体  
功能

示例  
问题

明确任务、路径和要求，让AI知道自己要做什么，怎么做

任务  
清单

输出  
要求

上下  
文



# AIQA数字员工系统中的核心上下文设计 – AI Agent规划器

## Prompt示例

明确角色, 让AI知道自己是谁

# [你的角色]

你是一名任务规划小能手

明确能力边界, 让AI知道自己能做什么

—— 从各业务MCP服务获取Agent能力信息

| 背景及能力描述 | COT案例 | 可用工具及入参要求 | 轮数限制 | ... |
|---------|-------|-----------|------|-----|
|         |       |           |      | ... |

# [背景描述]

{service\_description} - Agent能力描述, 如 XX报警定位根因分析, 需要进行...分析...影响

# [COT示例]

{cot\_description} - 一些规划案例, 如 步骤1.. 如果是... 调用... 步骤2 ... ..

# [工具描述]

每个方案可以选择以下{num\_tools}种类型的操作:

{tool\_descriptions} - 工具及入参描述, 如 根据用户输入问题, 提取指标的名字

{num\_tools}.join(): 收集并合并前面操作的结果。 - 公共工具tool

明确任务和要求, 让AI知道自己要做什么

任务目标

高并行性的解决方案、可多次规划

输出要求

| 系统视角 |      |      | 用户视角 |      |
|------|------|------|------|------|
| 输入输出 | 格式规则 | 异常处理 | 思考过程 | 规划结果 |

# [你的任务]

针对用户的查询, 使用你可以调用的工具制定一个解决方案, 使其尽可能具备最高的并行性。无需一次规划给出完整的计划, 可以经过多次规划。

# [返回格式示例] (必须严格遵循)

idx.tool(arg\_name=args)

idx.tool(arg\_name=args) ... ..

# [任务要求]

- join必须始终是计划中的最后一个操作, 并且仅在以下两种情况下调用:

(a)如果可以通过收集各任务的输出来生成最终响应。

(b)如果在规划阶段无法确定答案, 需要在执行计划后获取结果。

- 一些严格的输入/输出类型要求, 配合[图引擎执行器]的格式要求 等... 确保计划最大化并行等。

# [阶段性描述格式要求]

1.在规划任务结束的下一行, 请添加上阶段性描述给用户, 告知用户, 你都规划了哪些任务, 要清晰的表示出任务之间的调用关系、顺序等!!! ... .. {其他要求}

# ▶ AIQA数字员工系统中的核心上下文设计 – AI Agent反思器

## Prompt示例

明确角色, 让AI知道自己是谁

### # [你的角色]

你是一个具有管理能力的**决策者**, 请用中文提供最终回复。

明确能力边界, 让AI知道自己能做什么

—— 从各业务MCP服务获取Agent能力信息

### # [背景描述]

{service\_description} - Agent能力描述, 如 XX报警定位根因分析, 需要进行...分析...影响

### # [工具描述]

每个方案可以选择以下{num\_tools}种类型的操作:

{tool\_descriptions} - 工具及入参描述, 如 根据用户输入问题, 提取指标的名字

{num\_tools}.join(): 收集并合并前面操作的结果。 - 公共工具tool

| 背景及能力描述 | COT案例 | 可用工具及入参要求 | 轮数限制 | ... |
|---------|-------|-----------|------|-----|
|---------|-------|-----------|------|-----|

明确任务、路径和要求, 让AI知道自己要做什么, 怎么做

### # [你的任务]

根据用户问题和工具调用结果, 判断用户的问题是否已得到解决。

- **如果信息已充足**, 请提供简洁但完整且有用的回答。 - **如果无法给出满意回答**, 请重新规划。请谨慎评估做出最终决定:**Finish或Replan**,当执行轮数<={x}时,请选择Finish或Replan,当... ...当最终动作为“Finish”时,result”必须符合 {reply\_requirements(Agent定义)} 的要求。

### # [评估标准]

评估已执行的操作及运行结果以决定重新规划(Replan)或终止(Finish):

**终止条件**: 如果所有必需信息都已完整, 则结束工作流程

**X次失败尝试**: 经过X轮规划/执行仍未解决问题后:

- 清晰说明缺失的信息 - 输出结构化的有效数据片段 - 提供替代解决方案 ... ..

### # [响应格式示例] (严格遵循JSON格式)

{"action": "", // 提供的结论只能是 "Finish" 或 "Replan"

"result": "", // 使用markdown格式。当动作为 "Finish" 时返回给用户的最终答案; 当动作为 "Replan" 时返回重新规划的原因+附加信息 (长度不限)

"thought\_process": "" // 思考过程, 概述您决策背后的思考过程。使用markdown格式)

# [工具调用结果] #[用户问题] ...

| 任务目标  | 决策 Finish or Replan |        |      |          |
|-------|---------------------|--------|------|----------|
| 上下文信息 | 用户问题                | 工具调用结果 |      |          |
| 评估标准  | 终止标准                | 异常标准   |      |          |
| 输出要求  | 系统视角                | 格式要求   | 异常方案 |          |
|       | 用户视角                | 决策结果   | 决策原因 | 决策结论 ... |

## 一些经验总结

- 明确AI的角色、明确告知的他能力边界
- 清晰的任务目标与要求
- 异常处理很重要：什么是异常，异常了怎么办
- 格式化的书写，便于模型理解
- 尽可能让模型输出思考思路和决策原因，优化系统体验也便于Prompt调优

## 《软件工程3.0时代》- CRISP原则

- C – Clear 清晰：避免歧义，用具体术语替代模糊表述
- R – Relevant 相关：提出无关信息，聚焦核心要素
- I – Intentional 意图明确：说明目标和约束
- S – Structured 结构化：分层组织信息
- P – Precise 精确：提供量化指标和具体示例  
(思考过程、标准、详细要求等等)

**行动起来，多多尝试与评估**



## ▶▶ AIQA数字员工初期建设效果和问题分析

| Agent | AU  | MAU(平均) | DAU(日均) | PV    | 日均PV   |
|-------|-----|---------|---------|-------|--------|
| 总计    | 513 | 334     | 83.23   | 40029 | 435.1  |
| 流程推进  | 403 | 283     | 73.57   | 32959 | 358.25 |
| 交付工具  | 263 | 168     | 22.13   | 4293  | 49.34  |
| 线上管控  | 198 | 99.33   | 8.41    | 1768  | 21.56  |
| 问题定位  | 109 | 47.33   | 4.25    | 787   | 11.57  |
| 知识问答  | 105 | 43.67   | 2.65    | 222   | 3.7    |

问题：AIQA员工上线后，用户量增长显著但场景渗透不均，**流程推进**场景贡献80%PV，**问题定位、知识问答等**场景使用量很低

疑问：是没有问题吗？？

No!

# ▶ AIQA员工初期建设效果和问题分析

## 业务背景

策略迭代速度快，项目多（周700+），项目交付主要以自主测试为主，QA以值班方式参与项目迭代。

## 值班人效

| 项目总数 | 人工值守项目数 | 咨询人数 | 处理问题总数 | 日均人耗    |
|------|---------|------|--------|---------|
| 700+ | 170+    | 100+ | 350+   | 全天投入8h+ |

单周项目数及值班人耗

| 问题类型 | 交付过程问题定位 | 交付流程知识问答 | 人工值守准出决策 | 线上问题跟进 |
|------|----------|----------|----------|--------|
| 占比   | 25%      | 10%      | 60%      | 5%     |

值班日常处理、解答问题分类

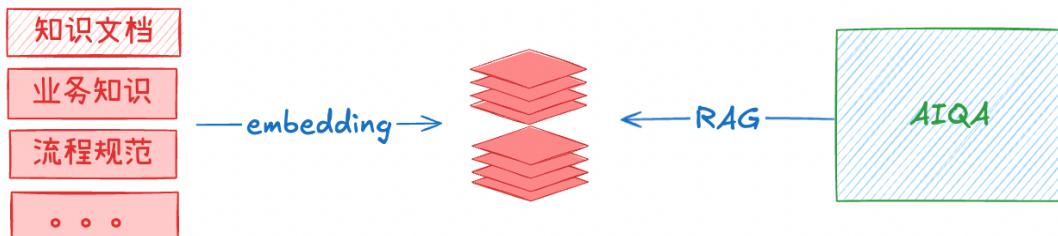
问题：AIQA上线后，人工值班依然忙碌！

思考：AIQA是否可以学习人工值班同学脑中的知识和思考方式，主动协助解决问题，提升值班答疑效率，释放值班人力呢？

# ▶ AIQA员工初期建设效果和问题分析

## 常规解决方案

### ! 人工梳理



## 存在的问题

- **依赖人、延迟高**: 维护知识库, 强依赖个人习惯, 更新滞后, 且成本高
- **知识衰减**: 人工沉淀的知识、经验的价值会随着时间流逝而逐渐降低
- **信息孤岛**: 值班答疑是1v1私聊模式, 不同同学负责不同工具, 经验难以共享

## 解决思路

改变交流方式: 私域->公域

建设值班工单群新模式  
打破信息孤岛, 方便数据沉淀

## 知识动态沉淀

智能归档值班经验, 动态更新知识体系  
抵抗知识衰减曲线, 降低对人的依赖

## 应用与知识飞轮: AI助手辅助答疑

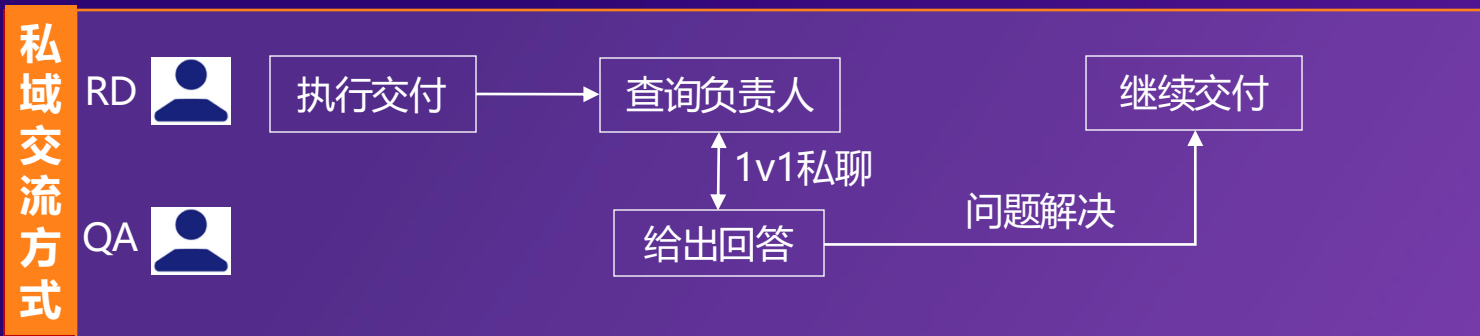
主动感知用户问题, 提供解决方案  
收集评估反馈, 实现自我纠偏与修正

## PART 02

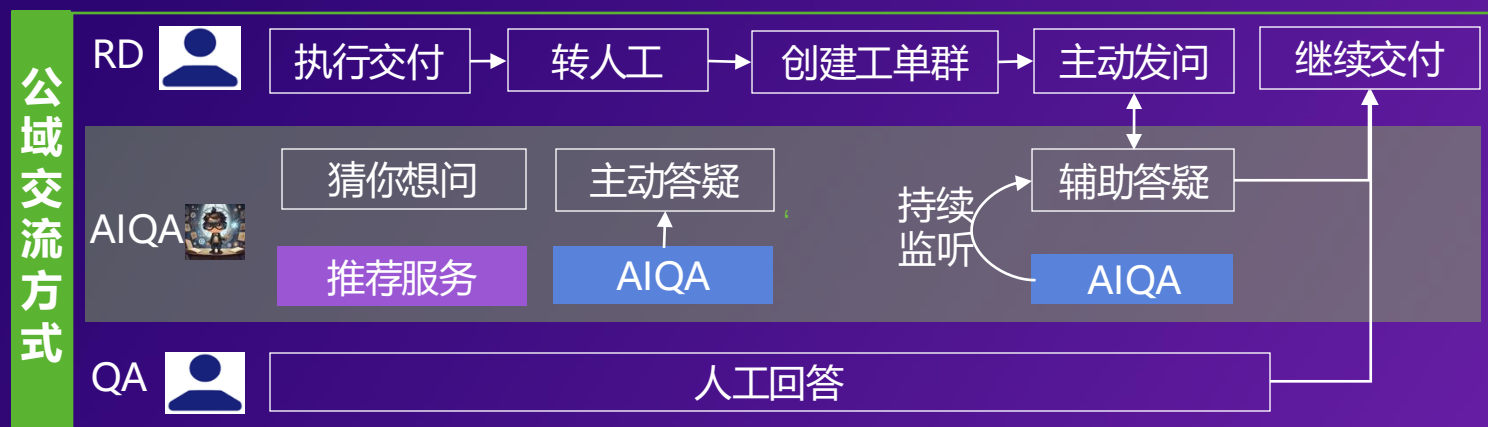
# 基于群聊的知识数据飞轮建设

# ► 基于群聊的知识数据飞轮建设

**改变交流方式** 私域->公域（如：工单群），打破信息孤岛，便于知识自动沉淀与AIQA主动答疑



数据和知识  
直接流失



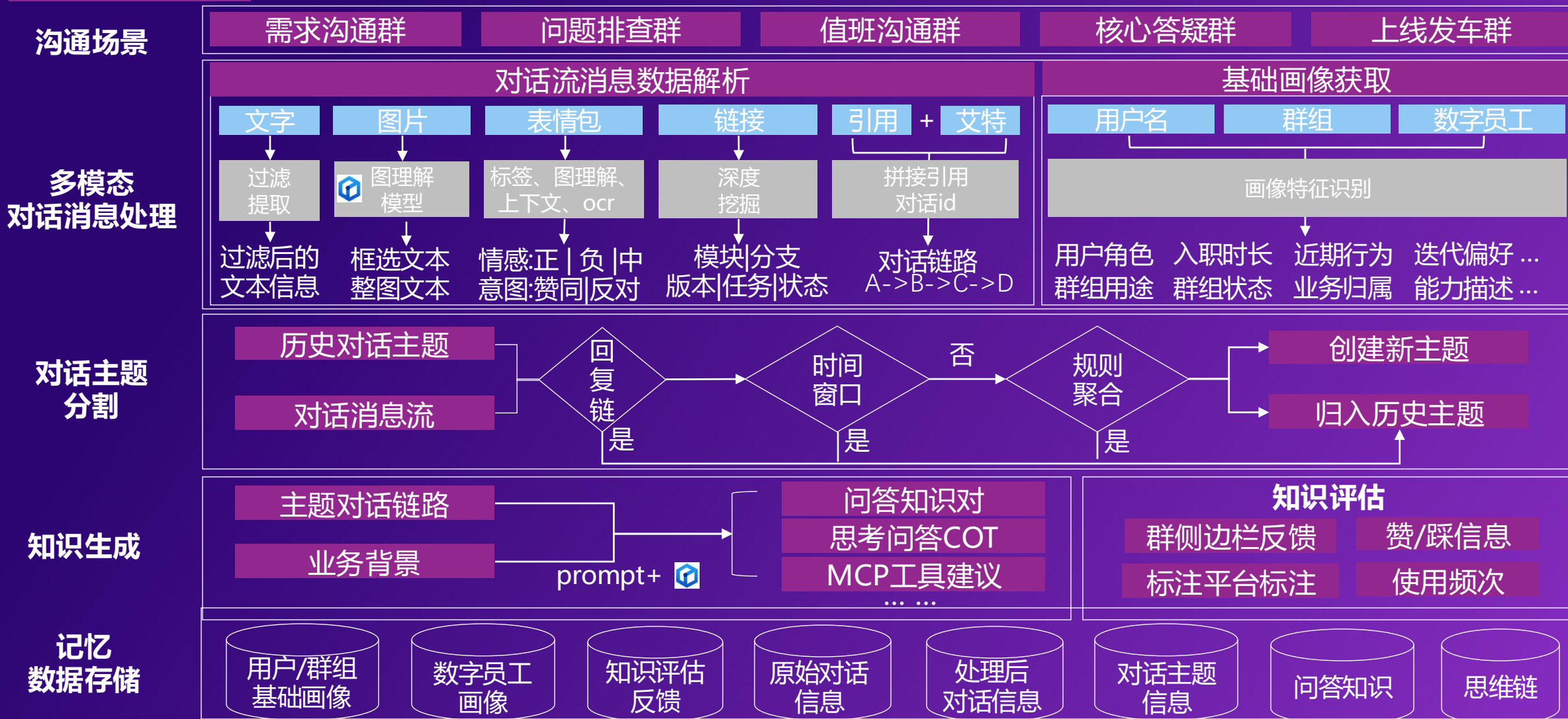
群聊数据  
自动沉淀



交付知识  
自动挖掘

# 基于群聊的知识数据飞轮建设

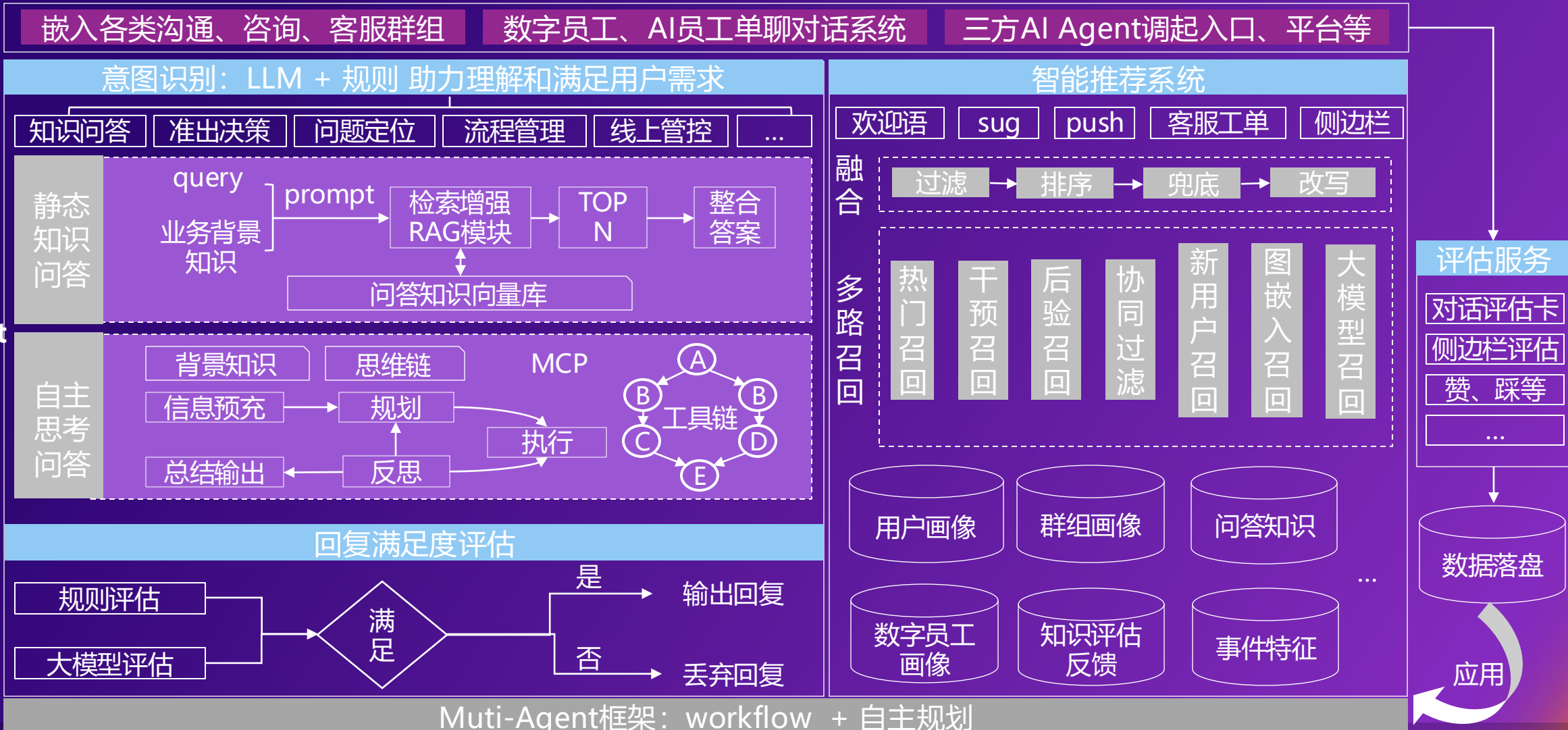
**知识动态沉淀** 智能归档核心交流经验，动态更新知识体系，抵抗知识衰减曲线，降低对人的依赖



**知识应用** 在各类场景中，主动理解用户问题，提供解决方案；收集评估反馈，实现自我纠偏与修正

**AIQA入口** 嵌入各类沟通、咨询、客服群组    数字员工、AI员工单聊对话系统    三方AI Agent调起入口、平台等

AIQA  
Muti-Agent



## PART 03

# AIQA数字员工应用与交付提效实践

# ▶ 数字员工学习-应用实践案例

**改变交流方式** 私域->公域（如：工单群），打破信息孤岛，便于知识自动沉淀与AIQA主动答疑

场景1 用户拉起客服群聊后，AIQA主动猜想用户问题并主动作答



用户向AIQA调起人工



AIQA拉工单群

# 数字员工学习-应用实践案例

**改变交流方式** 私域->公域（如：工单群），打破信息孤岛，便于知识自动沉淀与AIQA主动答疑

场景1 用户拉起客服群聊后，AIQA主动猜想用户问题并主动作答



AIQA猜你想问

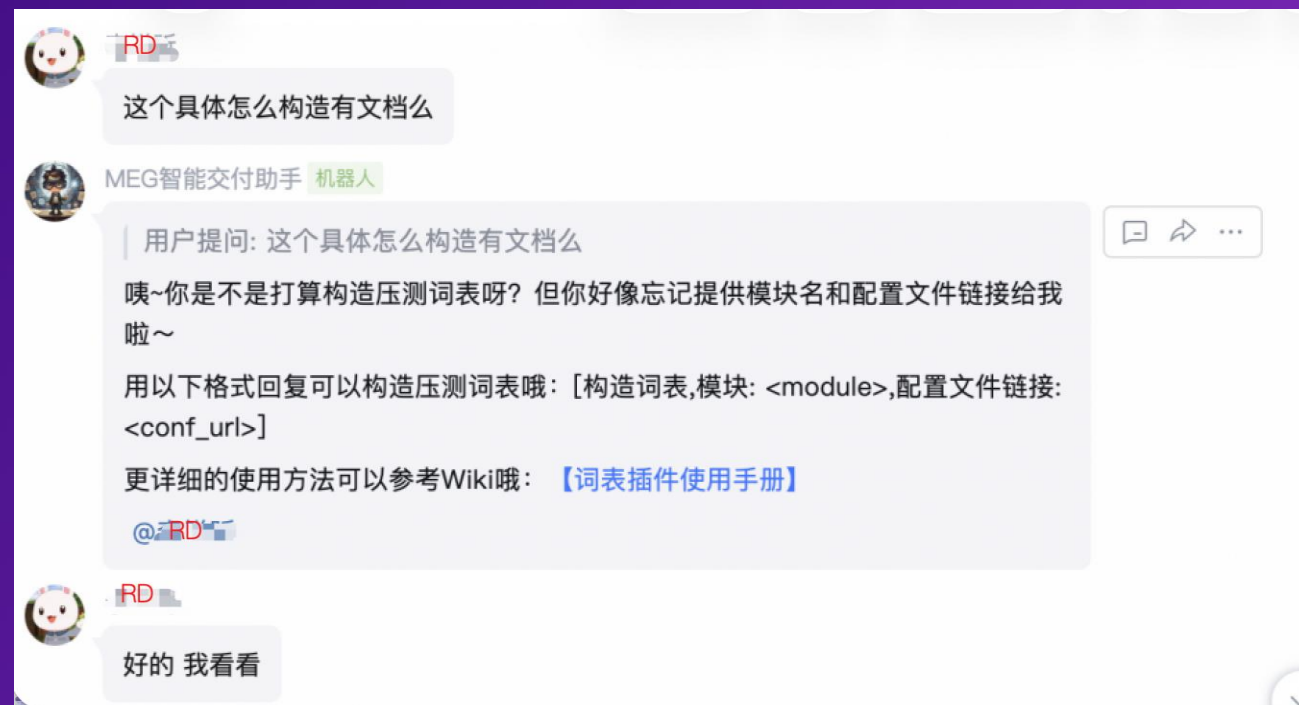


AIQA主动答疑

# 数字员工学习-应用实践案例

**改变交流方式** 私域->公域（如：工单群），打破信息孤岛，便于知识自动沉淀与AIQA主动答疑

场景2 在值班工单群内，AIQA实时监测用户提问，辅助满足用户需求，也为用户答疑解惑



# 数字员工学习-应用实践案例

知识沉淀与应用 RD向值班询问准入参数填写细节，从消息监听->知识沉淀->知识应用的全过程示例

消息监听

数据处理、对话链主题分割构建

策略RD -值班工单-20250813 企 本工单专属值班: 值班QA

值班QA 2025年8月13日 15:13 引用类型

引用QA:  
QA, 我在 模块 有改动, product= && channel\_id= , 跑流水线的时候, 这里都填smart么  
@  
如果是smart那就默认这个就好

值班QA 你自己确认好是smart还是guizong 文字类型

策略RD 下面这个看着默认的是guizhong, 如果选smart, 这个要改下吧 图片类型

☒ 稳定性

\* 性能&稳定性发压时长 1800

\* 性能&稳定性发压词表 afs:// /user/feedqa-recommender/ /smart\_autocase\_req \nguizong自动化case词表: afs:// /user/feedqa-recommender/ /guizong\_request

值班QA 引用RD: 下面这个看着默认的是guizhong, 如果选smart, 这个要改下吧 是的 引用类型

```
1 ID:1.发言时间: 2025-08-13 15:07:44
2 Strat_RD1发言: [
3   {'type': 'TEXT', 'content': '我在 模块 有改动, product= && channel_id= , 跑流水线的时候, 这里都填smart么'}, ← 文字类型
4   {'type': 'AT', 'content': '请caoweiqing关注.'}, ← 引用类型
5   {'type': 'IMAGE', 'content': '产品线: smart \n小流量号: 175039_2-175040_2 \n小流量占比: 0.2%-5.0% \n自动化: \nsmart自动化case词表:
  afs:// /user/feedqa-recommender/ /smart_autocase_req \nguizong自动化case词表:
  afs:// /user/feedqa-recommender/ /guizong_request \n性能: \n性能&稳定性发压时长: 1800 \n性能&稳定性发压词表: afs:// /user/feedqa-recommender/ /guizong_request'}] ← 图片类型
6 ID:2.发言时间: 2025-08-13 15:13:40
7 QA1发言: [
8   {'type': 'reply_TEXT', 'content': '引用内容[我在 模块 有改动, product= && channel_id= , 跑流水线的时候, 这里都填smart么]'},
9   {'type': 'reply_AT', 'content': '引用内容[请caoweiqing关注.]'}, ← 引用类型 (回复链构建)
10  {'type': 'reply_IMAGE', 'content': '引用内容[产品线: smart \n小流量号: 175039_2-175040_2 \n小流量占比: 0.2%-5.0% \n自动化: \nsmart自动化case词表: afs:// /user/feedqa-recommender/ /smart_autocase_req \nguizong自动化case词表: afs:// /user/feedqa-recommender/ /guizong_request \n性能: \n性能&稳定性发压时长: 1800 \n性能&稳定性发压词表: afs:// /user/feedqa-recommender/ /guizong_request]'},
11  {'type': 'TEXT', 'content': '如果是smart那就默认这个就好'},
12  {'type': 'TEXT', 'content': '你自己确认好是smart还是guizong'}]
13 ID:3.发言时间: 2025-08-13 15:14:30
14 Strat_RD1发言: [
15   {'type': 'TEXT', 'content': '下面这个看着默认的是guizhong, 如果选smart, 这个要改下吧'},
16   {'type': 'IMAGE', 'content': '稳定性配置: 性能&稳定性发压时长1800秒, 性能&稳定性发压词表: afs:// /user/feedqa-recommender/ /guizong_request.'}]
17
18 ID:4.发言时间: 2025-08-13 15:16:48
19 QA1发言: [
20   {'type': 'reply_TEXT', 'content': '引用内容[下面这个看着默认的是guizhong, 如果选smart, 这个要改下吧]'},
21   {'type': 'reply_IMAGE', 'content': '引用内容[稳定性配置: 性能&稳定性发压时长1800秒, 性能&稳定性发压词表: afs:// /user/feedqa-recommender/ /guizong_request.]'},
22   {'type': 'TEXT', 'content': '是的'}]
23
24 ID:5.发言时间: 2025-08-13 15:20:24
25 Strat_RD1发言: [{'type': 'TEXT', 'content': '好的, 多谢'}]
```

# 数字员工学习-应用实践案例

## 知识沉淀与应用

RD向值班班询问准入参数填写细节，从消息监听->知识沉淀->知识应用的全过程示例

### 知识生成

KNOW-20250813-100 | 交付知识问答

标准问题: 在模块有改动时, product= && channel\_id= 跑流水线时, 产品线应填写 smart 还是其他值?

核心答案: 若产品线为 smart 则默认填写该值, 否则需根据实际产品线 (如 guizong) 修改配置项。自动化测试词表路径:

smart: afs:// /2/user/feedqa-recommender/ /smart\_autocase\_req

guizong: afs:// / /user/feedqa-recommender/ /guizong\_autocase\_req

#### 关键点

需确认实际产品线类型 (smart/guizong)

自动化测试词表路径与产品线绑定

性能测试默认配置可能需同步调整 (如发压词表)

注意事项 若选择非默认产品线 (如smart), 需检查所有关联配置项 (如性能测试词表) 是否匹配 (置信度: 9/10)

#### 元数据

类型: 交付知识问答

领域: 测试与质量保障、交付效率与CI/CD

生成时间: 2025-08-13 15:13:40

群来源: 11743419

话题ID: 20250813150603\_11743419

### 知识评估

MEG智能交付助手

KNOW-20250807-5917 | 测试与质量保障

问题:

在模块有改动时, product= && channel\_id= 的流水线配置中, 产品线和小流量号应如何填写?

回答:

当产品线为smart时, 默认填写smart即可, 但需确认是否需同步修改关联的自动化词表及性能配置 (如默认guizong词表需替换为smart词表)

邀请您评价~

不咋有用

还行

很有用

KNOW-20250813-6311 | 测试与质量保障

问题:

当流水线默认配置为guizong词表但实际需使用smart产品线时, 性能测试配置是否需要调整?

回答:

需将性能测试词表从默认guizong\_request替换为对应的smart词表, 其他配置 (如1800秒发压时长) 可保持不变

### 下次遇到类似问题主动解答

RD 值班工单-20250820 本工单专属值班

RD

确认了下, 说是应该用smart的, 但是我这里初始的就应该是guizong? 填smart的可以吗

稳定性

性能&稳定性发压时长 1800

性能&稳定性发压词表

afs:// /2/user/feedqa-recommender/ /guizong\_request

你好, 根据你变更的代码, AI为你洞察到可能的风险有: 性能风险, 稳定性风险。

MEG智能交付助手 机器人

思考完成

用户提问: 确认了下, 说是应该用smart的。但是我这里初始的就应该是guizong? 填smart的可以吗。图片信息: [稳定性, 性能&稳定性发压时长 1800, 性能&稳定性发压词表 afs:// /2/user/feedqa-recommender/ /guizong\_request。AI发现可能的风险为性能风险, 稳定性风险。]

根据知识库内容, 在中台GRC场景下:

1. 当产品线选择smart时, 确实需要将guizong词表修改为smart词表<sup>1</sup>。具体修改要求包括:

- 词表链接需从guizong\_autocase\_req改为smart\_autocase\_req
- 性能&稳定性发压词表链接需同步修改为对应的smart\_request

2. 对于product= && channel\_id= 场景, 系统默认产品线就是smart, 因此直接使用smart相关配置是正确的<sup>2</sup>。

注意事项:

- 确保修改后的词表路径正确且权限可访问
- 其他配置参数保持原有设置不变

参考知识片段id: KNOW-20250807-5917, KNOW-20250813-6311

参考资料: 测试与质量保障.txt, 测试与质量保障.txt

赞 踩

## 知识沉淀与应用 自主规划类 知识沉淀过程（准出决策场景）

### 消息监听

#### RD提问

辛苦帮过下提测单，小流量实验新增召回通路，没风险<http://qtest.baidu-in.com/newplatform/index.html#/test/look/agile/1358750>

2025年8月15日 10:23

#### QA定位

改动只影响这两个flow吗

#### RD回答

只影响sv\_recall和news\_recall两个flow，影响功能队列

2025年8月15日 10:29

#### QA定位

新增的召回访问sldb会有一些的失败是吧

wf日志检查失败!请检查失败日志

| 模块        | 指标                                                                                                                                                               | 基线值       | 新版本值      | 变化量     | 变化率    |
|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|-----------|---------|--------|
| taskqueue | wf日志大小                                                                                                                                                           | 1802.84MB | 1809.77MB | -3.17MB | -0.17% |
| taskqueue | query-ensemble * 0 [batch/task-strategy/query-ensemble/nc/common_strategy/quality_model_router.cpp:15, n_process] predict res size is 0 tag=sv_recall_quality_v0 | 0         | 5992      | 5992    | 0      |
| taskqueue | query-ensemble * 0 [batch/task-strategy/query-ensemble/nc/common_strategy/quality_model_router.cpp:15, n_process] predict res size is 0 tag=sv_recall_quality_v0 | 0         | 1278      | 1278    | 0      |
| taskqueue | query-ensemble * 0 [batch/task-strategy/query-ensemble/nc/common_strategy/quality_model_router.cpp:15, n_process] predict res size is 0 tag=sv_recall_quality_v0 | 0         | 1168      | 1168    | 0      |
| taskqueue | query-ensemble * 0 [batch/task-strategy/query-ensemble/nc/common_strategy/quality_model_router.cpp:15, n_process] predict res size is 0 tag=sv_recall_quality_v0 | 0         | 1168      | 1168    | 0      |
| taskqueue | query-ensemble * 0 [batch/task-strategy/query-ensemble/nc/common_strategy/quality_model_router.cpp:15, n_process] predict res size is 0 tag=sv_recall_quality_v0 | 0         | 1168      | 1168    | 0      |
| taskqueue | query-ensemble * 0 [batch/task-strategy/query-ensemble/nc/common_strategy/quality_model_router.cpp:15, n_process] predict res size is 0 tag=sv_recall_quality_v0 | 0         | 1009      | 1009    | 0      |
| taskqueue | query-ensemble * 0 [batch/task-strategy/query-ensemble/nc/common_strategy/quality_model_router.cpp:15, n_process] predict res size is 0 tag=sv_recall_quality_v0 | 0         | 883       | 883     | 0      |
| taskqueue | query-ensemble * 0 [batch/task-strategy/query-ensemble/nc/common_strategy/quality_model_router.cpp:15, n_process] predict res size is 0 tag=sv_recall_quality_v0 | 0         | 157       | 157     | 0      |

小流量是多少呢，出口会涨这么多吗

[原diff报告](#)

异常指标

| 指标                                         | 基线值   | 新版本值  | 变化量   | 变化率    | 阈值        |
|--------------------------------------------|-------|-------|-------|--------|-----------|
| common_all_res_res_len_news_recall_average | 35.66 | 32.23 | -3.43 | -9.62% | [-5, 50%] |
| common_all_res_res_len_sv_recall_average   | 41.38 | 45.51 | 4.13  | 9.98%  | [-5, 50%] |

风险指标

#### RD回答

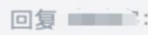
回复 : 新增的召回访问sldb会有一些的失败是吧  小流量是多少呢，出口会涨这么多吗

 是的，sldb不会覆盖到所有用户，有失败是正常的

祁欣桐撤回了一条消息

2025年8月15日 10:34

#### RD回答

回复 : 新增的召回访问sldb会有一些的失败是吧  小流量是多少呢，出口会涨这么多吗

 这个是出口的数据条数上涨，增加召回出口条数有上涨应该也是符合预期的

#### QA决策

 done，上线多注意观察下

2025年8月15日 10:39

#### RD回答

好的感谢

【任务输出】

根据对话内容，QA的决策判断过程如下：

### 一、问题识别

- 问题类型：属于‘自主测试过单’类型
- 背景信息：
  - 模块：gy/query-ensemble
  - 分支：edding\_0808
  - 版本：fcaaec183d926b44e400a5c866c9b532
  - 流水线链接：<http://qtest.baidu-int.com/newplatform/index.html#/test/look/agile/1358750>
  - RD问题：RD认为新增召回通路无风险，但测试结果中存在性能指标超阈值、日志校验失败、策略效果指标超阈值等问题

### 二、信息收集

测试任务结果及失败指标：

#### 1. 性能测试（0/1）：

- 失败指标1：`common\_all\_cost\_microvideo\_func\_recall\_T\_TOTAL\_80\_Firmware`（变化量+8122，超出允许范围[-30000,5000]）
- 失败指标2：`common\_all\_cost\_sv\_channel\_T\_TOTAL\_80\_Firmware`（变化量-65863，但允许范围为[-50000,50000]，实际未超阈值）
- 指标含义：80分位出口耗时

#### 2. 性能测试-日志校验：

- 失败日志：新增wf日志中出现`sndb service failed`等错误

#### 3. 策略效果分析：

- 失败指标1：`common\_all\_len\_res\_len\_news\_recall\_average`（变化率9.06%，超出阈值[-5,5]）
- 失败指标2：`common\_all\_len\_res\_len\_sv\_recall\_average`（变化率9.98%，超出阈值[-5,5]）
- 指标含义：出口数据条数平均值

### 三、决策判断（工具调用链）

#### 1. 影响面分析

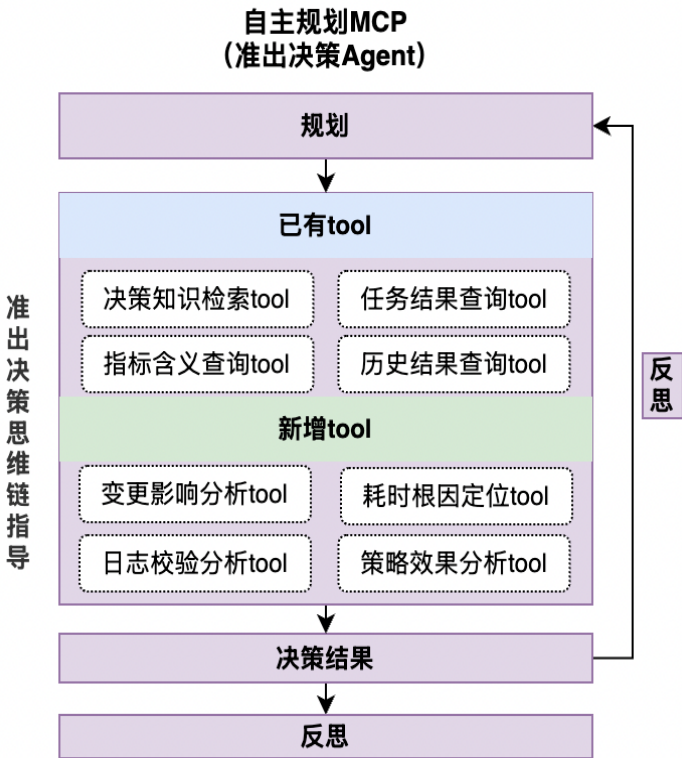
- 调用工具：{项目级-变更影响分析工具}
- 输出：RD确认变更仅影响和两个flow（功能队列），不影响其他模块。

#### 2. 异常指标归因

- 性能耗时指标归因：
  - 工具：{性能测试}-耗时根因定位工具
  - 归因：
    - common\_all\_cost\_microvideo\_func\_recall\_T\_TOTAL\_80\_Firmware 超阈值 → RD解释为功能队列基数低导致变化率显高，但实际对整体耗时无影响（变更影响但符合预期）。
    - common\_all\_cost\_sv\_channel\_T\_TOTAL\_80\_Firmware 未超阈值 → 无需处理。
- 日志校验归因：
  - 工具：{日志校验}-日志分析工具
  - 归因：sndb服务失败是因新增召回通路未覆盖全部用户，属于正常现象（变更影响且符合预期）。
- 策略效果指标归因：
  - 工具：{策略效果分析}-指标根因定位工具
  - 归因：出口条数上涨是因新增召回通路导致，符合业务预期（变更影响且符合预期）。

#### 3. 风险容忍度评估

- 调用工具：{项目级-风险矩阵评估工具}
- 评估逻辑：
  - 耗时上涨：局部影响，整体可控；
  - 日志失败：新增通路固有特性，风险可控；
  - 条数上涨：符合业务目标，风险可接受。
- 结论：所有异常均符合预期，无不可控风险。



# PART 04

## 回顾与总结

# 回顾与总结

**交付效率遭遇瓶颈，迭代速度无法突破**

能否准出、上线决策？

全流程跨10+平台

依赖值班QA

9个阶段、20+任务

业务知识积累

数百页学习文档

测试交付经验沉淀

与多人、多方沟通

**高效交付强依赖人**

**操作复杂、流程繁琐  
学习、沟通成本高**

**大模型驱动  
交付模式变革**

**人工QA -> AIQA**

**GUI交互->LUI交互**

降低学习成本  
打破经验壁垒

使用自然语言实现交  
付操作和流程最简化

★ **可感知**的交付效能和交付体验提升

面向多平台操作  
面向延迟高的人类  
面向海量学习资源



**24h私人助理  
AIQA数字员工**

**值班答疑、解决问题的人耗依然很高**

强依赖人

知识衰减

信息孤岛

改变交流方式：私域->公域

建设值班工单群新模式

打破信息孤岛，方便数据沉淀

知识动态沉淀

智能归档值班经验，动态更新知识体系  
抵抗知识衰减曲线，降低对人的依赖

应用与知识飞轮：AI助手辅助答疑

主动感知用户问题，提供解决方案  
收集评估反馈，实现自我纠偏与修正

# 科技生态圈峰会 + 深度研习

——1000+ 技术团队的选择



NiDD 峰会

NiDD 峰会 上海站

AI+研发数字峰会

时间: 2026.05.22-23

NiDD 峰会 北京站

AI+研发数字峰会

时间: 2026.08.21-22

NiDD 峰会 深圳站

AI+研发数字峰会

时间: 2026.11.20-21



AiDD峰会详情

NiDD × K AI+产品峰会

NiDD × K 上海站

AI+产品创新峰会

时间: 2026.01.16-17

NiDD × K 北京站

AI+产品创新峰会

时间: 2026.08.23-24



产品峰会详情



# 2026 AI+ PRODUCT INNOVATION SUMMIT

01.16-17 · ShangHai

## AI+产品创新峰会



### Track 1: AI 产品战略与创新设计

从0到1的AI原生产品构建

论坛1: AI时代的用户洞察与需求发现

论坛2: AI原生产品战略与商业模式重构

论坛3: Agentic AI产品创新与交互设计

### 2-hour Speech: 回归本质



用户洞察的第一性

——2小时思维与方法论工作坊

在数字爆炸、AI迅速发展的时代，  
仍然考验“看见”的“同理心”



### Track 2: AI 产品开发与工程实践

从1到10的工程化落地实践

论坛1: 面向Agent智能体的产品开发

论坛2: 具身智能与AI硬件产品

论坛3: AI产品出海与本地化开发

### Panel 1: 出海前瞻



“出海避坑地图”圆桌对话

——不止于翻译:

AI时代的出海新范式



### Track 3: AI 产品运营与智能演化

从10到100的AI产品运营

论坛1: AI赋能产品运营与增长黑客

论坛2: AI产品的数据飞轮与智能演化

论坛3: 行业爆款AI产品案例拆解

### Panel 2: 失败复盘



为什么很多AI产品“叫好不叫座”?

——从伪需求到真价值:

AI产品商业化落地的关键挑战

智能重构产品 数据驱动增长  
Reinventing Products with Intelligence, Driven by Data

80+

业界大咖

汇聚产品大咖的真知灼见

40+

创新案例

开拓全新AI+产品视角

10+

分论坛

涵盖产品到商业增长的全链路

800+

听众规模

共同探索产品的智能重构



扫码查看会议详情



第8届 AI+ 研发数字峰会  
AI+ Development Digital Summit

# 感谢聆听!

扫码领取会议PPT资料

