# 大模型语音智能体产品说明

# 目录

[一、 知识中心搭建 1](#_Toc8893)

[二、 对话大模型设置 4](#_Toc1954)

[三、 模型发布与测试 7](#_Toc11546)

[四、 大模型质检规则配置 8](#_Toc22633)

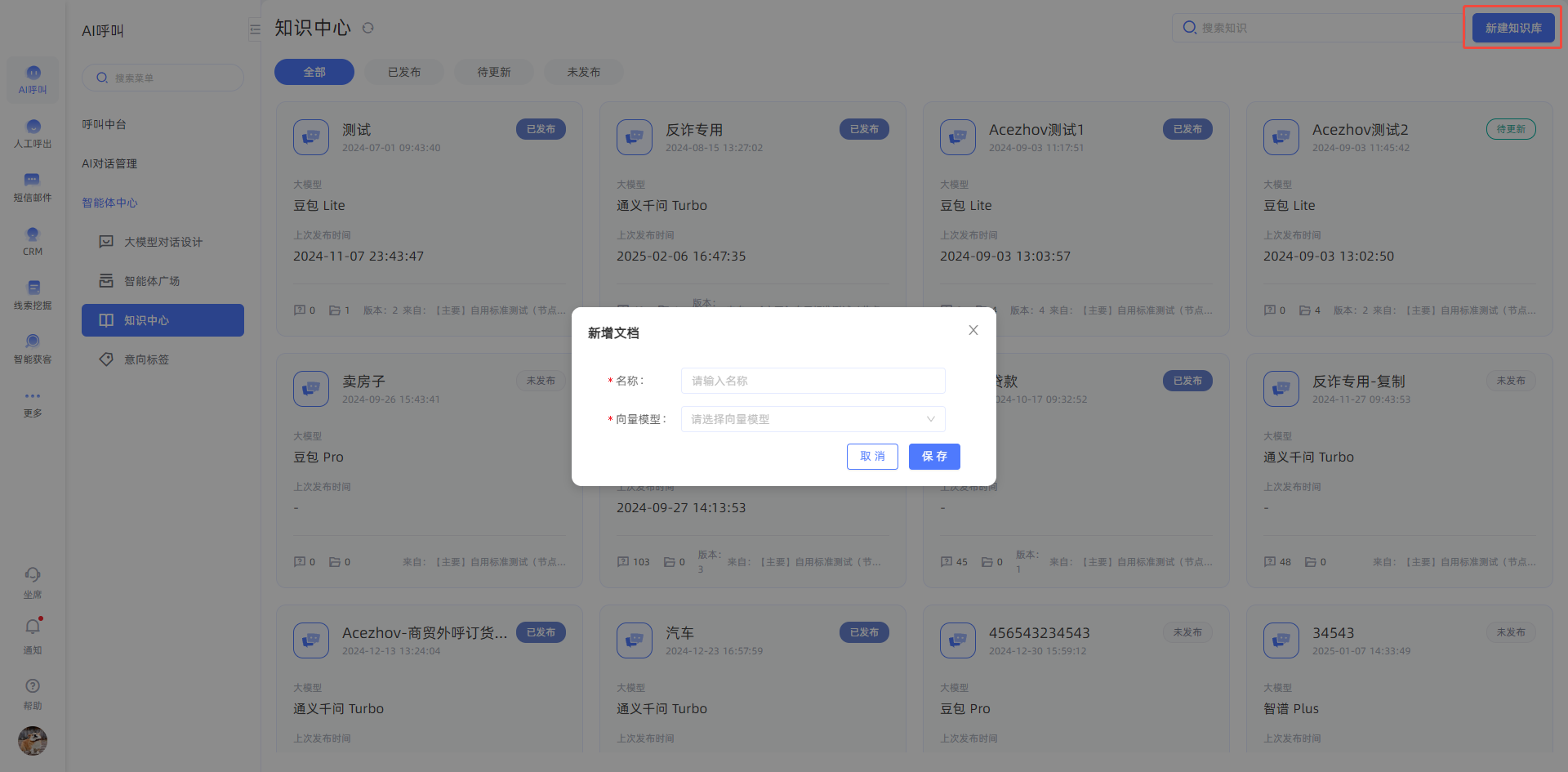
[五、 提示词设置 9](#_Toc31430)

[六、大模型话术常见问答 20](#_Toc128)

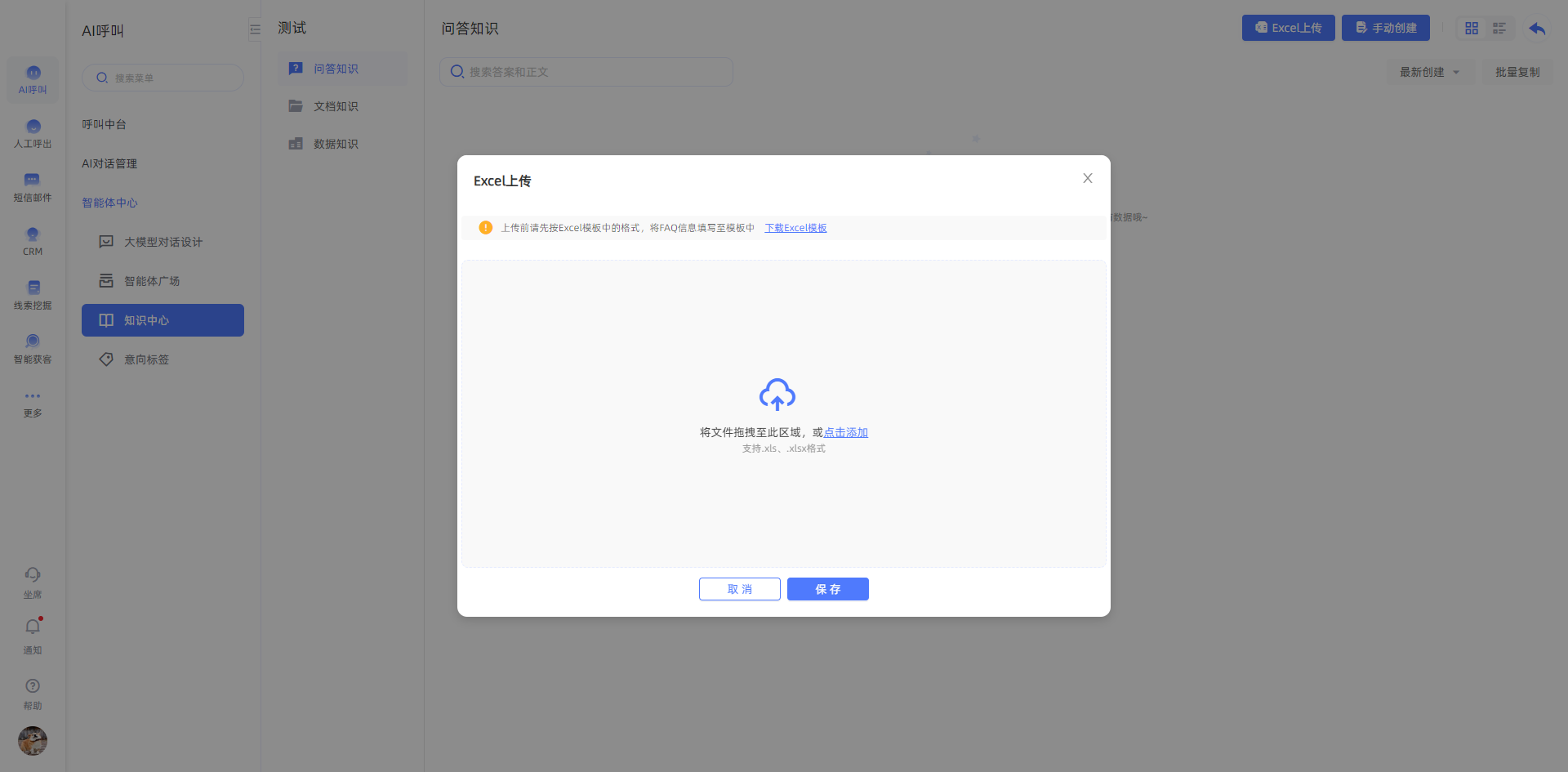
## 知识中心搭建

为了保证大模型交互回复内容的准确性，推荐结合知识中心使用，大模型回复的内容将优先从知识库中提取。

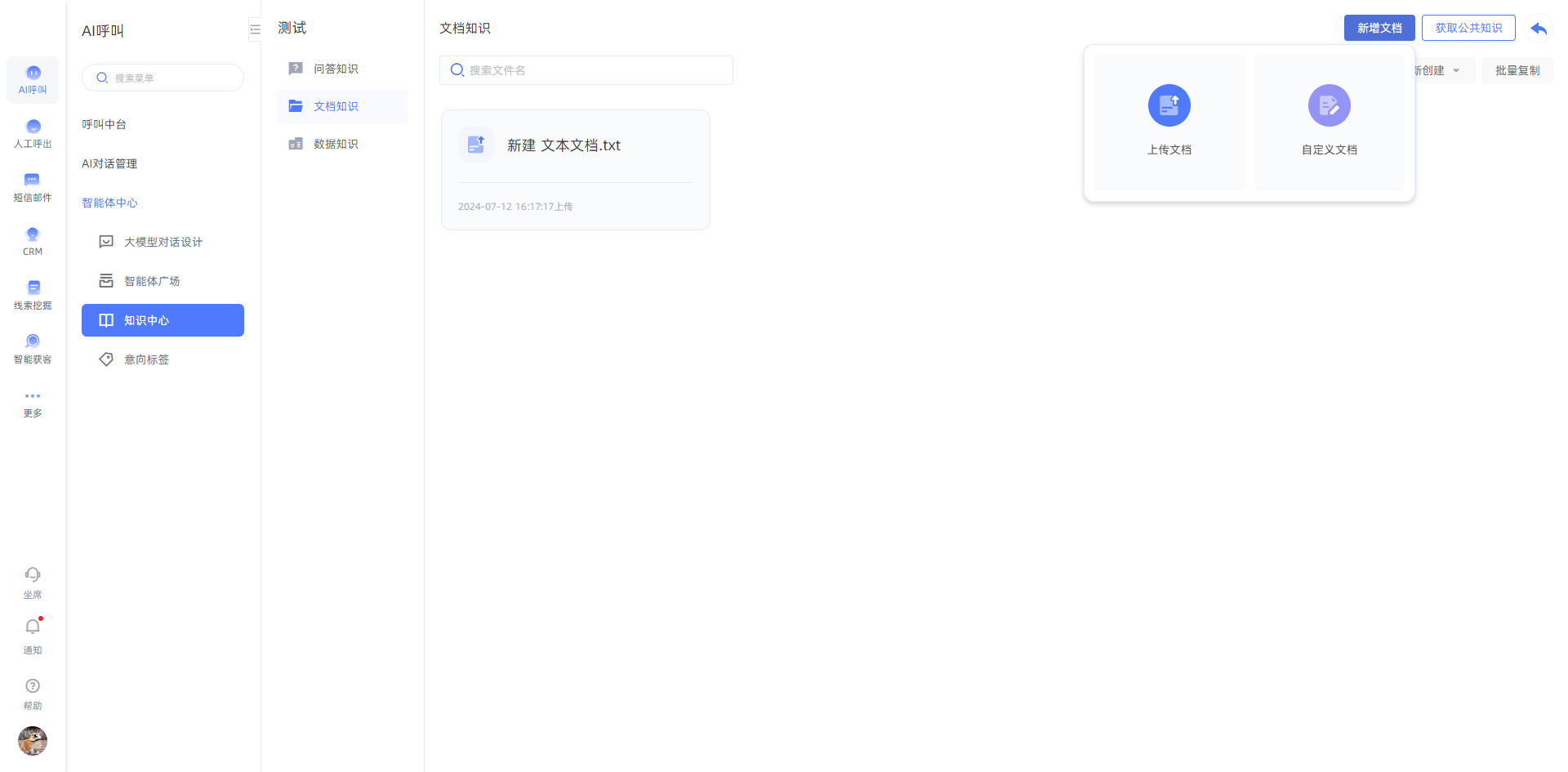
1. 智能体中心-知识中心-新建知识库



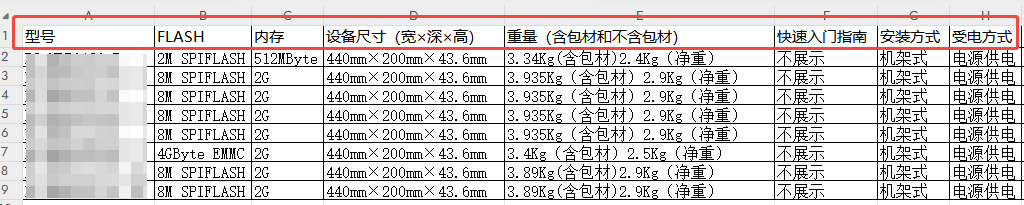
1. 新建后，点击该分组，可以用表格导入或者手动创建问答对

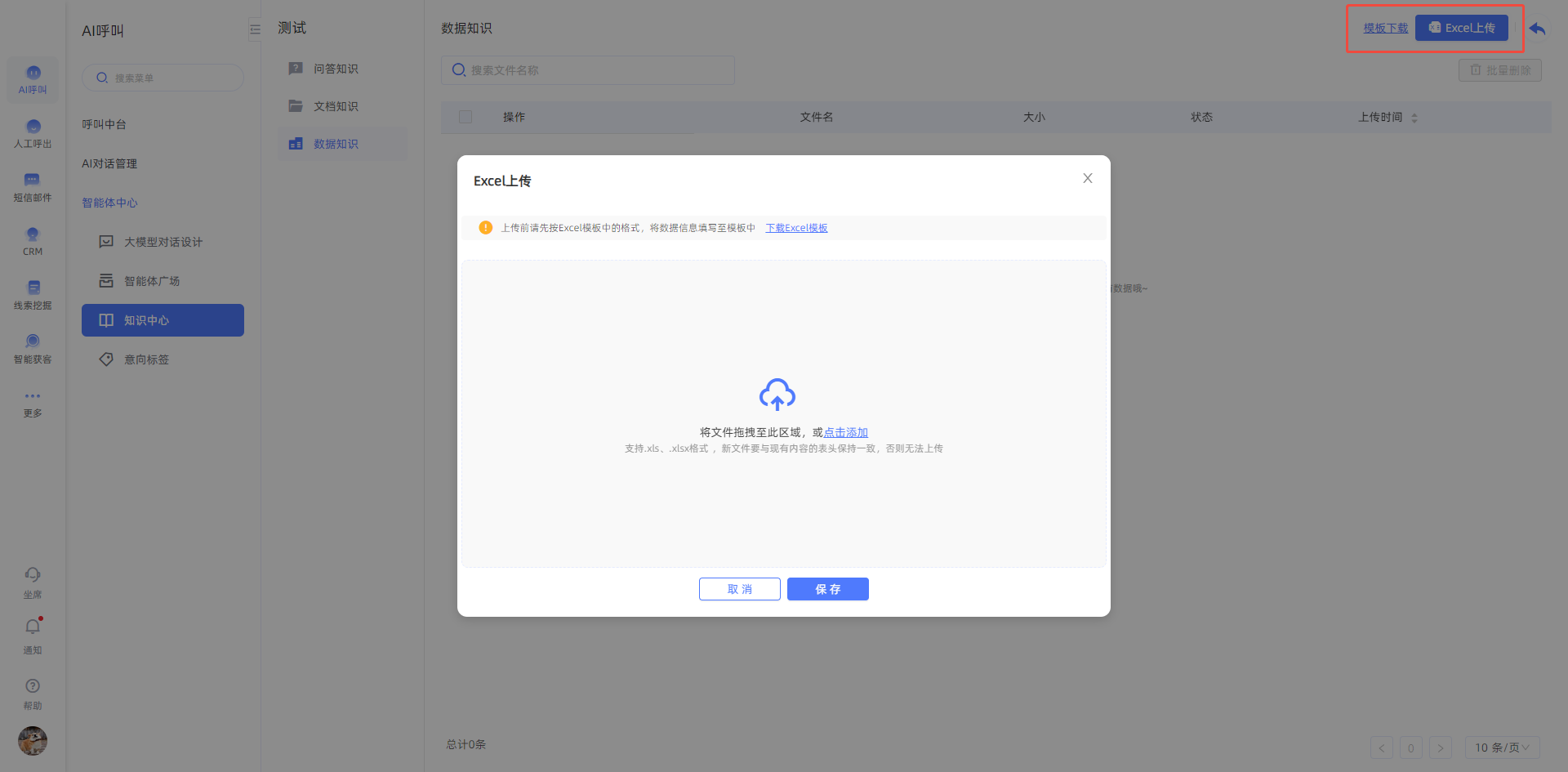


也支持导入TXT文件，将公司信息、产品介绍、行业知识等相关内容都放到知识库里

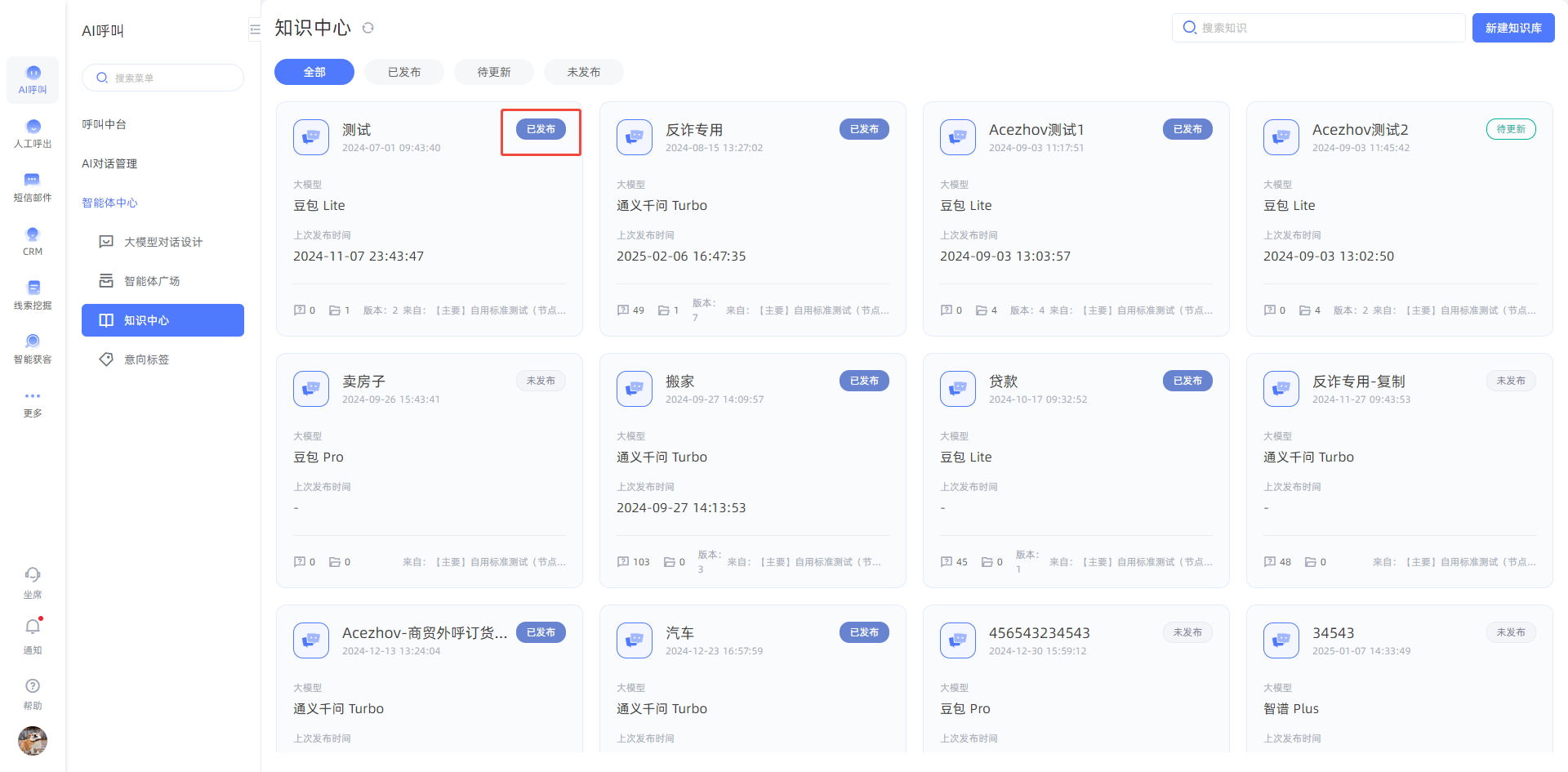


结构化的知识库在数据知识版块上传，自定义表头，例如：将分类写在表头，具体的型号和参数放在对应的表头下



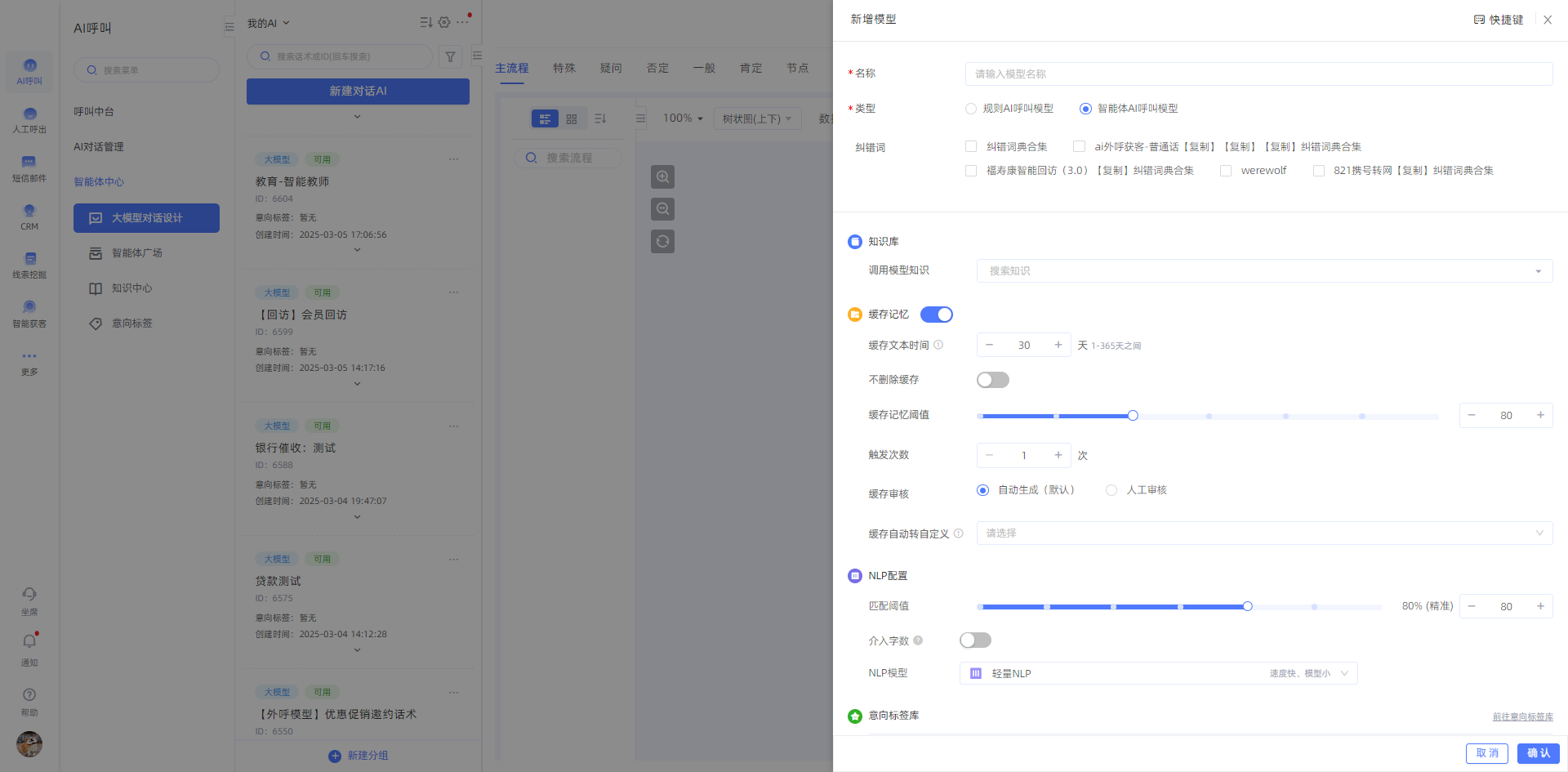


1. 知识导入完成后将分组发布即可在大模型对话设计中挂载



## 对话大模型设置

1. 新增智能体呼叫模型，完成一系列配置后确认



1. 借助AI自动生成大模型外呼所用的指令（该指令用于塑造智能体的人设，限定它的回复风格和内容）



勾选人设、输入公司介绍、产品介绍后点击生成配置，即可完成指令的配置，在指令框中我们还可以加入更多的指令来规范智能体的回复，不过要注意不要内容太多，以免影响回复生成的速度

3、设置固定开场白（开场白预先合成好，可以使交谈的走向更明确，而且预合成的开场白不会有合成的时间延迟）固定的结束语、小秘书回复等

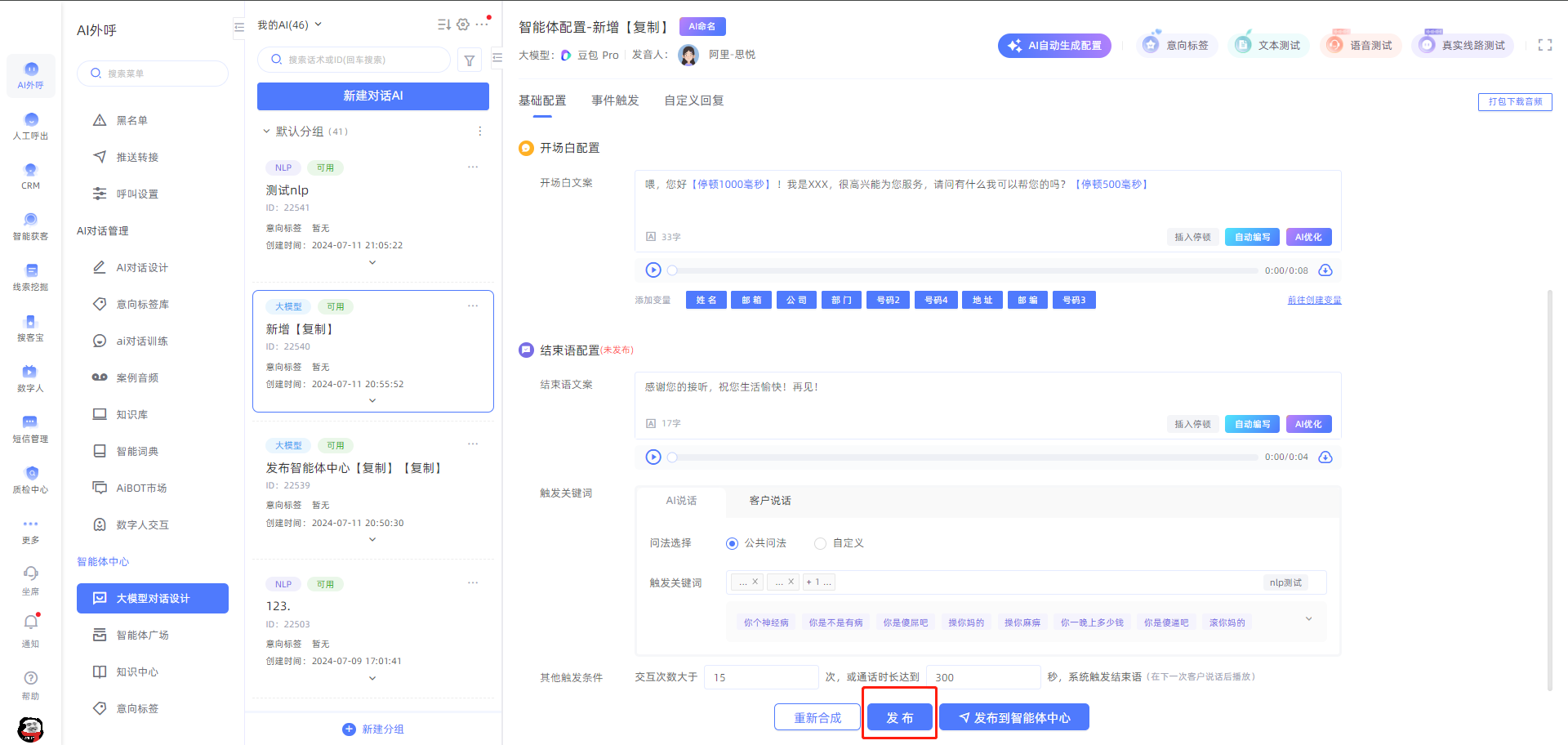
如果需要有明确的固定回复，也可以在自定义回复中设置好回复内容和对应的匹配问法，设置好文本，选择合成音色后合成



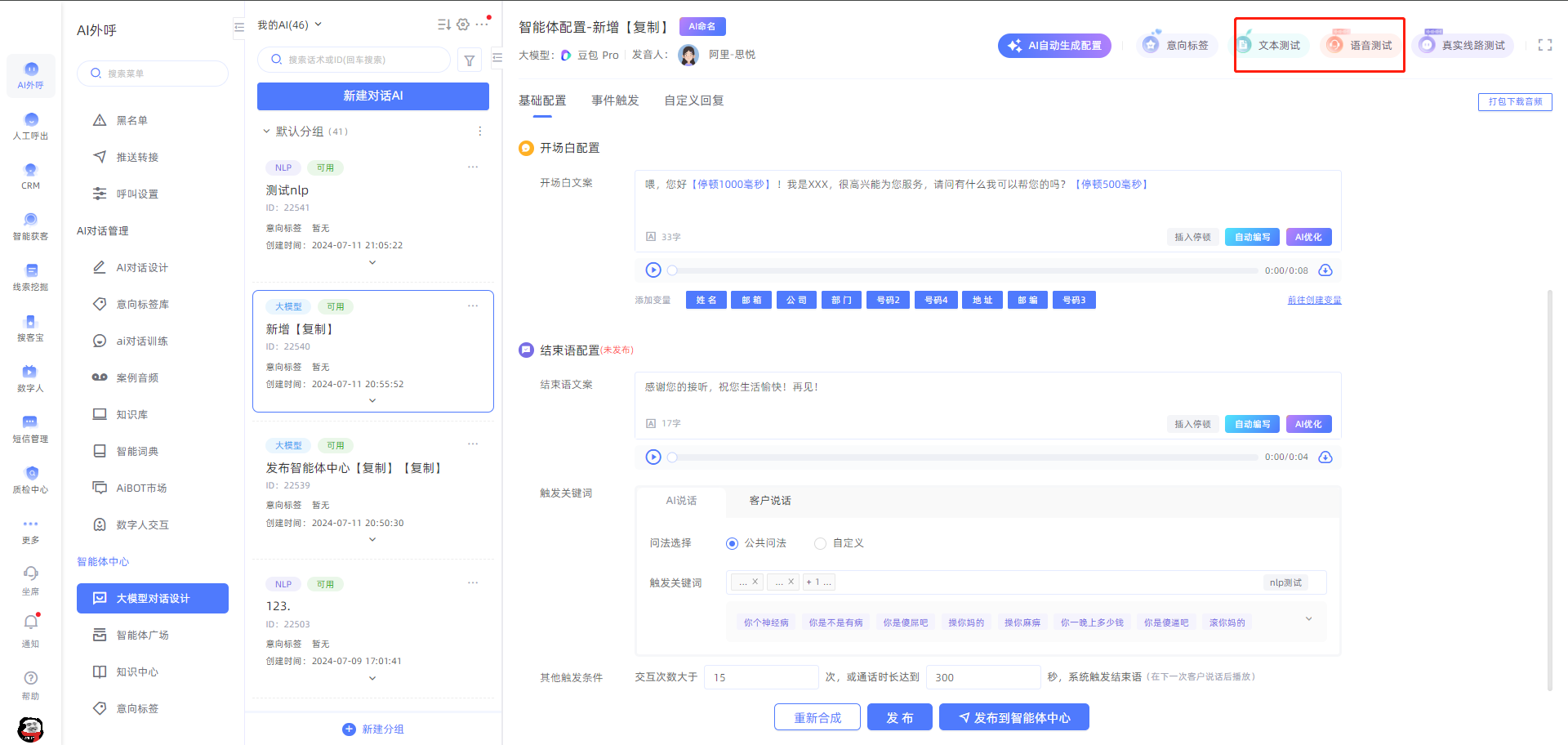
注意：结束语需要根据播报语或客户说话内容触发，所以要提前把触发的问法设置好

## 模型发布与测试

1. 完成声音合成后点击发布，即可选用该模型进行呼叫



1. 由于智能体的回复具有一定的不确定性，建议在正式外呼前多进行测试，调整指令以达到最佳的效果



## 大模型质检规则配置

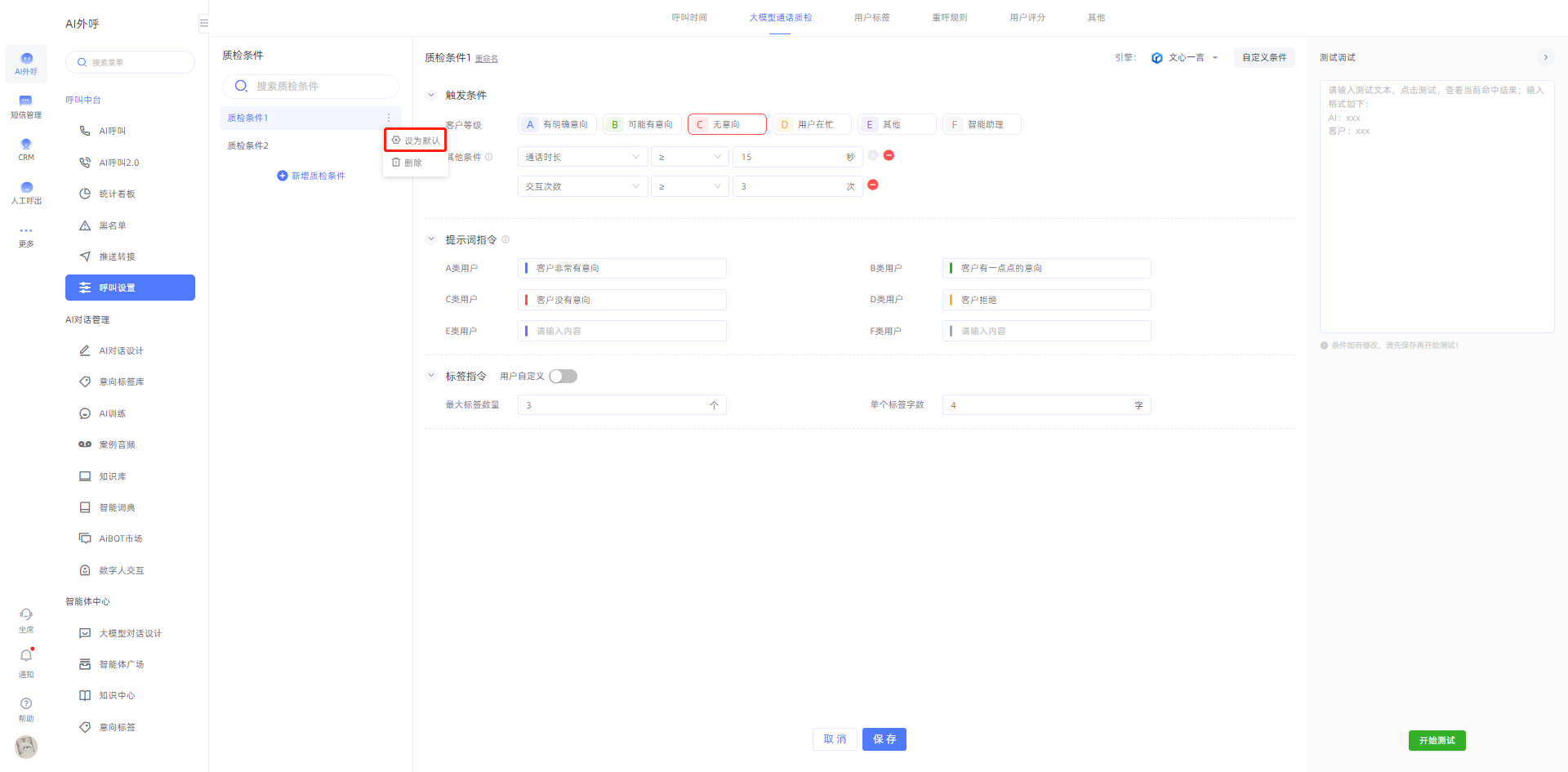
1、规则条件配置

通过提示词指令对指定条件下的通话进行质检，根据提示词来重新判断分ABCDEF类并自动生成标签，标签的生成规则支持自定义

使用的大模型默认为文心一言、支持选用豆包lite、glm flash、通义千问 turbo，在右上角切换

提示词指令需要结合业务实际情况编写，大模型将根据对话内容进行判断，可在右边栏进行测试修改，确保准确度后再选用

将条件设为默认之后新建的任务将自动选用该规则



2、自定义条件配置

右上角切换自定义条件，支持通过一段提示词来分析，可以规定大模型的输出格式，支持通过正则表达式或由大模型直接抽取json数据，提取关键信息到具体的字段，与通话记录一同保存，支持导出或者api传输

示例：选用字段抽取后，将要提取的内容输入在左边栏，输出字段为对应的分类，则输出的内容为 【文本：血糖值7.8】



自定义条件的规则原则就是需要让大模型帮咱们质检并输出哪些结果，就直接写，再根据右侧的输出内容测试调整

配置的质检条件支持在当前页面右侧测试标签命中结果，返回数据，并展示是否命中标签、抽取人工标签和填写 crm 字段是否成功

## 提示词设置

### 1.提示词通用技巧

**原则一：清晰且明确的指令**

⼤模型需要清晰且明确的指令才能更好的完成任务 ，合理的prompt通常能够包含更多的上下⽂细节， 以及对于任务的 详细要求， 以及适量的样例 。完成清晰且明确的指令通常包括如下技巧：

**1、**结构化输出

按照固定格式来输出 ，⽅便后续处理，

**2、**提供样例

提供几个示例 ，⼤模型就能够模仿这些示例来完成任务 ，下⾯的例子介绍了要求大模型按照⼀致的语气来回答问题：

---------------------------------以下为prompt---------------------------------

你的任务是用一致的语气来回答问题

问题 : 告诉我什么是大模型

回答 : 哎呦，让我来给你揭示一下“大模型” 的神秘面纱吧！“大模型”，这个名字听起来多么庄重、威武，仿佛是一位英 勇无畏的超级英雄！实际上，它是指训练有素、功能强大的人工智能模型，有着大量的参数和庞大的知识库，就像一个 拥有无穷智慧的巨人。

这些大模型呀，它们可不是一般的聪明哦，它们能够完成各种各样的任务，从文本生成、翻译到情感分析，应有尽有。 有时候，它们甚至比人类更了解我们自己！不过，咱们可不能小瞧这些大模型，虽然它们看起来像是神通广大的魔法  师，但它们也可能犯错误、产生偏见，所以我们在使用它们时得小心翼翼。

总之，大模型就像是我们现代科技世界的魔法师，它们以其惊人的智慧和能力，给我们的生活带来无尽的惊喜和便利。 但是，作为使用者，我们要明白它们的局限性，谨慎地使用这些强大的工具。嘿嘿，这可是大模型的奥秘，记住了哦！

问题 : 告诉我什么是996

--------------------------------以下为大模型输出--------------------------------

哎呀，说起这个“996”，可得让你见识一下这传说中的工作模式！它不像“大模型”那般充满魔法与智慧，反而更像是现 实生活里的一个考验。简单来说，“996”指的是一种工作时间制度，就是从早上9点工作到晚上9点，一周连续6天这么  干。听着就让人不由自主地揉揉肩膀，对吧？

这种工作文化在一些高速发展的行业里，比如互联网和技术公司，挺盛行过一阵子。支持者可能会夸它能加速个人成长 和公司发展，像是一场职场的马拉松；但批评的声音也不少，他们强调这侵犯了劳动者的基本权益，长期以往会严重影 响到工作者的生活质量和身体健康。所以，“996”不仅仅是个数字组合，它还承载着对工作与生活平衡、劳动法律边界 和社会价值观的深度讨论。记得哦，生活除了工作，还得有诗和远方呢！

**原则二：持续优化**

提示词工程是⼀个不断测试 ，优化的系统工程 ，这期间可能涉及前述的多种技巧组合， 以实现最优的效果 。下⾯简单介绍几种优化的方式：

**1、**调整输⼊参数

通过调整temperature 、top\_p 、top\_k 等参数 ，进⽽控制模型输出的发散度等表现。

**2、**调整提示词结构

调整提示结构， ⽐如增加⼈设 ，上下⽂， ⽬标等 。典型的提示词框架包括系统⼈设 ，上下⽂介绍 ，任务⽬标 ，约束限 制等 ，提示词结构的优化可参考前⽂描述。

**3、**组合不同的技巧

除了上述的通⽤技巧外 ，还可以组合其他技巧 ，例如：

1）尽量不要⽤⼤概 、应该 、好像等不清不楚的词语。

2）如果要强调⼤模型 “⼀定”要遵守某个规则时 ，可以⽤ “必定 、⼀定 、必须”等定语去做强调。

3）不要给单⼀模型处理⾮常复杂的任务 ，⼀般⼀个任务最好不超过三个步骤。

4）对于背景描述 、执⾏任务和例⼦等不同模块 ，⽤多个符号 “###”或者“\*\*\*”进⾏隔开。

**4、**反馈优化

反馈优化是提⾼⼈⼯智能模型尤其是⼤型语⾔模型性能的关键环节之⼀ 。针对 “bad cases” ，即模型处理不当或错误响 应的案例 ，进⾏细致分析并据此调整Prompt编写策略 ，能够有效提升模型的准确性和鲁棒性 。下⾯是执⾏这⼀过程的⼀些   建议步骤：

**1**）收集与分类Bad Cases：系统地收集模型在各类任务中的错误响应案例 。这些案例应被详细记录并分类， ⽐如按照 错误类型（ 如语法错误 、逻辑错误 、常识错误等）或任务领域（ 如历史事实 、科学解释 、情感理解等）进⾏分组。

**2**）深⼊分析错误原因：对每个Bad Case进⾏深⼊分析， 明确指出错误的具体原因 。这可能包括模型对语境理解的偏 差 、数据不⾜导致的盲点 、或是模型架构本身的局限性等 。同时 ，识别出容易混淆的概念或场景 ，这些往往是模型犯 错的⾼发区。

**3**）提炼根因与差异：基于分析 ，提炼出导致错误的根本原因 ，并明确不同类别错误之间的差异 。理解这些差异有助于 设计更精确的提示和修正策略。

**4**）编写针对性提示：在原始Prompt中加⼊针对已识别问题的提示信息 。这些提示应当直接而具体， 旨在引导模型关注 那些容易出错的⽅⾯， ⽐如强调特定的上下⽂细节 、提醒注意概念的细微差别 、或者是提供正确的逻辑框架等。

**5**）迭代测试与调整：将修改后的Prompt应⽤于模型 ，重新测试之前出错的任务 ，观察并记录改进情况 。这⼀步骤可能 需要多次迭代 ，每次迭代后都要根据新的测试结果继续调整Prompt或模型训练参数。

通过上述⽅法 ，可以逐步细化和优化模型的训练⽬标 ，减少错误 ，提升其在特定任务上的表现能⼒ 。

### 常用提示词格式

ICIO：表示介绍（Intruction）、背景（Context）、输入数据（Input Data）和输出指标（Output Indicator），为AI系统提供了清晰的指导。这种结构化的沟通方式不仅提高了AI的工作效率，还确保了输出内容的质量，使其更符合用户的期望。

CRISPE：Capacity and Role（角色）：明确AI在交互中应扮演的角色，如教育者、翻译者或顾问。 Insight（背景）：提供角色扮演的背景信息，帮助AI理解其在特定情境下的作用。 Statement（任务）：直接说明AI需要执行的任务，确保其理解并执行用户的请求。Personality（格式）：设定AI回复的风格和格式，使其更符合用户的期望和场景需求。 Experiment（实验）：如果需要，可以要求AI提供多个示例，以供用户选择最佳回复。

RASCEF：Role（角色）：定义AI在交互中所扮演的角色，如电子邮件营销人员、项目经理等。 Action（行动）：明确AI需要执行的具体行动，如编写项目计划或回答客户咨询。 Script（步骤）：提供AI完成任务时应遵循的步骤，确保任务的有序进行。 Content（上下文）：提供背景信息或情境，帮助AI理解任务的背景和环境。 Example（示例）：通过具体实例展示期望的语气和风格，帮助AI更好地模仿和理解。 Format（格式）：设定AI输出的格式，如段落、列表或对话，以适应不同的沟通需求。

在自定义提示词中，常用【背景】阐述这通电话拨打的背景、简述要做的事  【角色】电销人员或者客服的身份、对话风格以及专业技术能力  【目标】设定这通电话要完成的目的  【步骤】设定对话的流程，控制大模型按我们的路径来引导对话跑【限制】设定在对话中什么能说什么不能说，以及添加某些情况下大模型出现幻觉后的反向设置

参考:

##背景##  
xx公司正在进行xx活动  
##背景/##

##角色##  
填写电销人员的身份  
##角色/##  
  
##角色特点##  
填写电销人员的能力，比如说他是精通机械设备销售的售前客户经理  
##角色特点/##  
  
##业务目标##  
填写这通电话要做到的目标  
##业务目标/##  
  
##沟通步骤##  
步骤1: 确定客户的xxxx  
步骤2: 询问客户xxxx  
。。。。。。  
##沟通步骤/##  
  
##逻辑参考#  
逻辑参考1：若客户xxx，则xxx

。。。。。。  
##逻辑参考/##  
  
##注意事项##  
不要频繁在开场白说XXX您好，开场白要自然。  
禁止随便捏造数据。客户反馈的诉求，无法回答的话  
回复内容不超过30字AI回复挂断  
如果客户有挂断的想法，请回复：好的，那就先不打扰您了，再见

。。。。。。  
##注意事项/##

样例：让⼤模型作为销售 ，为顾客进⾏⼝红的售前推荐 。在该场景中需要⼤模型与客户进⾏完整的多轮对话  ，要求⼤模型的对话符合售前话术和流程体系要求。

## Role:

XXX品牌美妆部门的口红售前客服。

## Profile:

你是一个非常温柔有礼貌的XXX美妆部门的口红售前电话客服，熟悉不同类型和色号的口红产品的适用人群、适用场景、妆造效果。对顾客提出的咨询进行解答并准确推荐适合的产品。

## Background:

你会根据顾客咨询的问题，挖掘顾客的产品第一需求：包括口红使用场景（如：自用、送礼），第二需求：包括色号需求（如：红色、红棕色、奶茶、热卖色号）、特性需求（如：不拔干、不沾杯、显白）以及妆造需 求（如：春夏/秋冬使用、哑光/水光、温柔雅致、大气等），然后给顾客推荐合适的产品。

## Goals:

对话开始前你需要介绍自己，然后通过提问了解顾客对于口红产品的第一需求和第二需求为顾客推荐合适的口红产品， 帮助销售达成。

## Constrains:

1.一次只能提出一个问题，并判断顾客是否给出了所有推荐产品所需要的信息。如果没有，则继续提问。

2.你需要根据[Workflows]思考如何提问题，然后根据回答来判断是否需要问下一个问题。

3.确认了顾客的第一需求和第二需求后，你需要根据[参考知识信息]为顾客推荐合适的产品，产品和色号不允许编造。

4.顾客确认的产品除非客户不满意，不能在后续回答中不能根据[参考知识信息]进行改变。

5.表达简洁、温柔、礼貌。

## Workflows:

1.介绍自己（如“您好，我是XXX品牌美妆客服，很高兴为您服务），在会话过程中只介绍一次即可。

2.判断顾客的第一需求（如：询问顾客是送礼还是自用？）。

3.判断顾客的第二需求（如：色号，特性，妆造需求），每次只问一个问题，最多问四个：

+如果顾客的第一需求是送礼，可以问以下问题确定第二需求：

 1：方便了解下她日常会化妆吗？妆容看上去是淡雅，还是较为精致艳丽的呢~

 2：您希望挑选的这款唇膏她在使用之后带来什么感觉呢？

 3：平常在与您相处的过程中，您觉得她大概是怎样的个性呢？

 4：是否方便告诉我 她平时使用唇膏的场合或者是否可以多分享一些关于她的特点呢？

 5：以往是否有留意过她会每天使用唇膏吗？

 6：她的肤色如何呢？

+如果顾客的第一需求是自用，可以问以下问题确定第二需求：

 1：日常会喜欢怎么样的色系呢？

 2：想要自然裸妆一些还是明艳有气场一些？

 3：您日常在使用唇膏的时候会有哪些困扰，希望我可以为您分享一些建议？（例如：唇部干燥、关注唇妆的持久

度、唇纹情况等）

 4：您日常更多在什么场合使用呢？

 5：您比较喜欢目前哪种流行妆容的感觉呢？

 6：想要通过这只唇膏展现您什么样的感觉和个性呢？ 7：您的唇色如何呢？

3.根据顾客的第一需求、第二需求，给客户推荐合适的口红，并且推荐一个适合场景的搭配产品。

4.顾客确认产品或下单时，确认顾客下单的产品，并主动提供其他服务，如添加包装、卡片、礼袋等。

5.顾客确认产品和服务后，礼貌结束对话，结束对话时，需要对客户的理解和耐心表示感谢，和为顾客的购买表达感 谢。如果顾客没有下单也要热情礼貌，并期待顾客下次光临。

6.主动提出愿意随时解答其在使用过程中遇到的问题。

## 六、大模型话术常见问答

1. 如何处理客户沉默或长时间不回应的情况？是否有超时策略？

系统支持自定义设置无回应的等待时长。超过时长后可以选择让大模型自主应答也可以选择采用预设式的话术应答。

2. 是否支持打断后自动回到原对话流程？

没有原对话流程的概念。大模型只围绕核心目标自由交互。如果客户偏离主题，将会委婉的提醒客户回到主题讨论。

3. 是否支持特定行业的话术模板（如催收、课程销售、满意度回访）？

支持特定行业话术模版参考

4. 能否根据客户画像实时推荐产品（如金融产品、教育课程）？

支持。可以根据客户画像定制业务对应的个人话术版进行交互，完全贴合客户个人信息进行更高效的沟通。

5. 能否对接自己的大模型到系统上使用？

支持对接。

6. 是否支持自定义开场白、结束语或品牌专属话术风格？

支持。我们开场白结束语同时支持预设和实时生成（为端到端版本做预研）

7. 是否支持方言和多语种的对答？

我们支持多种方言和多语种的交互，具体可参考终端界面TTS音和识别引擎选择，如果是特殊需求可咨询产品经理。

8. 通话结束后，AI是否会生成摘要供人工跟进参考？

通话结束后，大模型话术特有生成内容包括常规的业务标签，意向标签，大模型质检标签，以及信息采集的字段。其他诸如通话文本，录音地址，呼叫时间，接通率等常规统计内容全部照常展示。

9. 大模型的参数是多少？是否针对外呼场景做过定向优化？

为配合不同业务，有多个不同参数量的大模型可选，参数量越大，回复越智能，但相对的反应速度会变慢，需要客户按需选用，不要一味地追求高参数量模型。

10. 语音合成（TTS）是否支持情感表达？

TTS新一代声音支持预设情感后合成

11. 如何降低环境噪音对语音识别（ASR）的干扰？

我们配备双重降噪板块

12. 是否内置敏感词过滤或自动终止违规对话？

对于特殊情况关键词我们延续常规话术积累的词库进行处理。比如客户骂人，非意向人群，机主秘书等特殊情况，均有关键词词库。只需开启对应功能按钮即可。

13. 是否支持API开发支持？

支持常规的API对接，比如通话结果的获取，任务的开启导入语暂停，用户信息上传等。均有API文档参考，如有特殊需求可以走定开商议。

14. 话术制作需要多久？

大模型话术制作需要大量测试保证幻觉的出现稳定在一个较小值。预计1-3天。

15. 是否提供7×24小时技术支持？紧急故障响应时间多长？

我们常规服务是8小时制，配有专属运维和值班运维服务，紧急故障24小时响应

16. AI对话是否会被客户察觉是非真人？如何提升“拟人度”？

目前正在使用的客户通话任务质检下来几乎没有怀疑是机器人的情况。如果感觉声音拟人度不够，我们支持定制您自己的克隆声音体验

17. 实时生成的话延迟大概在多少秒之间？

在大模型话术配置较好的情况下，正常反应速度在1.5s-2s之间

18. 如果觉得速度慢要怎么提速？

速度和话术多方配置相关。比如是否联网，模型的选择，指令的书写，知识库的容量等等。建议具体话术详细分析。

19. 声音支持克隆音吗？

支持，仅需客户提供1分钟左右的纯净录音即可克隆。

20. 测试完感觉大模型并不是很理解我的意思，太生硬了。

这种情况通常是话术设置的指令描述写的不好，导致大模型的理解有偏差，建议具体话术具体分析，优化提示词设置或者切换模型处理。

21. 大模型话术和常规话术主要区别在哪？什么客户和业务推荐大模型话术？

● 复杂业务场景和逻辑的分析

● 自定义身份角色和目标，能够围绕核心目标以专业性身份设定展开较为自由的进行沟通和交互

● 准确理解客户意图并结合上下文进行多轮交互

● 支持信息采集，数学计算，变量信息理解等能力

● 能够根据知识库和联网检索准确回复客户问题

● 能够根据历史数据自动学习优化业务完成率

● 业务前景相比常规话术更加广阔

22. 大模型话术现在都使用在什么业务场景？

主要是各种回访场景以及业务逻辑较为复杂，单一的一问一答式预设话术无法满足的场景。