

# UNIT 产品使用手册

<http://unit.baidu.com>

v3.1.3

百度在线网络技术（北京）有限公司

(版权所有,翻版必究)

---

## 目录

1	产品概述	5
1.1	UNIT 能做什么	5
1.2	基本概念	6
1.2.1	机器人	6
1.2.2	技能	6
1.2.3	对话技能	6
1.2.4	问答技能	6
1.2.5	对话意图	6
1.2.6	词槽	6
1.2.7	词典	7
1.2.8	对话模板	7
1.2.9	特征词	7
1.2.10	特征词词典	7
1.2.11	对话样本	7
1.2.12	对话样本集	8
1.2.13	问答对	8
1.2.14	标签	8
1.2.15	模型	8
1.2.16	沙盒环境	8
1.2.17	生产环境	8
1.2.18	图谱知识库	9
2	登录 UNIT	10
3	我的技能	11
3.1	技能管理	11
3.1.1	新建技能	11
3.1.2	删除技能	13
3.1.3	重命名技能	14

---

3.1.4	分享技能 .....	15
3.2	自定义-对话技能 .....	17
3.2.1	意图管理 .....	17
3.2.2	训练数据 .....	26
3.2.3	技能训练 .....	40
3.2.4	测试 .....	43
3.2.5	日志分析 .....	44
3.2.6	技能发布 .....	47
3.2.7	技能设置 .....	52
3.2.8	技能测试如何优化 .....	55
3.3	自定义-问答技能 .....	55
3.3.1	问答管理 .....	56
3.3.2	技能训练 .....	57
3.3.3	测试 .....	58
3.3.4	技能设置 .....	59
3.4	预置技能 .....	错误!未定义书签。
3.5	技能对话 API 调用 .....	75
4	我的机器人 .....	77
4.1	新建机器人 .....	77
4.2	添加技能 .....	78
4.3	测试 .....	79
4.4	机器人对话 API 调用 .....	80
5	我的知识 .....	81
5.1	功能概述 .....	81
5.2	图谱知识库 .....	81
5.2.1	图谱知识库的作用 .....	81

---

5.2.2	图谱知识库的使用流程 .....	81
5.2.3	定义知识类目及属性 .....	82
5.2.4	上传原始数据 .....	87
5.2.5	图谱构建 .....	87
5.2.6	手动调优 .....	93
5.2.7	发布图谱 .....	93
5.3	问答知识库 .....	95
5.3.1	问答知识库的作用 .....	95
5.3.2	问答知识库的使用流程 .....	95
5.3.3	上传文档 .....	95
5.3.4	手动调优 .....	96
5.3.5	导出数据 .....	97
6	创新技术 .....	98
6.1	语义解析 离线使用 .....	98
6.1.1	业务判断 .....	98
6.1.2	获取文件 .....	98
6.1.3	优化文件 .....	99
6.1.4	集成并调用 SDK .....	99
6.2	对话式文档问答技能 .....	99
6.2.1	创建技能 .....	99
6.2.2	删除技能 .....	100
6.2.3	重命名技能 .....	101
6.2.4	文档上传 .....	101
6.2.5	技能训练 .....	104
6.2.6	测试 .....	105
6.2.7	日志分析 .....	110
6.2.8	技能 API 调用 .....	111

# 1 产品概述

随着 AI 技术和理念的兴起，很多产品都希望增加对话式的人机交互，给用户带来更高效更自然的用户体验。然而对话系统（对话技能、对话机器人）的研发对于大多数开发者而言却是一个很困难的工作，对技术和数据的要求都很高。

为此，百度将积累多年的自然语言理解与交互技术对外开放，推出了智能对话系统定制平台 UNIT（Understanding and Interaction Technology），将业界领先的技术能力输出给广大的您，以便降低对话系统的研发门槛

## 1.1 UNIT 能做什么

帮您打造“面向任务的理解与交互能力”，一项任务可以是解答用户的某个问题（如：今天天气怎么样、我的快递到哪儿了）、执行用户指令（如：去给我把咖啡端过来），甚至通过一系列交互引导用户达成某项需求（如：通过注册-选座-下单完成订票）。

通过 UNIT，可以快速、灵活的按自身业务需求定制对话技能，满足智能助手、智能家居、智能汽车、智能客服、机器人等各类产品的应用场景。

在实际应用的场景中，对话技能需要完成的任务可能是上述任务的组合：例如，在购物场景下，对话技能主要的任务是解答用户的问题(例如：我的快递到哪儿了)，但也会存在一些辅助任务，例如，执行指令类（例如：帮我下个单吧）或对话引导类(例如：引导用户通过一系列操作去完成了退货的操作)

您可以脑洞大开，尽情发挥想象力，让自己的用户感受到更丰富的业务帮助和互动体验。比如：

- 在银行营业厅，与机器人对话来申请一张信用卡
- 在线客服更聪明的回答用户的问题
- 在餐厅根据您的喜好得到菜单推荐
- 让家庭机器人帮您打开空调、把咖啡端过来
- 肚子饿了，帮您下单买个外卖
- 和朋友一起玩猜谜游戏
- .....

## 1.2 基本概念

### 1.2.1 机器人

机器人是指在一个完整对话场景下的对话能力。通常按行业垂类划分（例如：小度智能音箱中的对话场景就是一个机器人）。

### 1.2.2 技能

某一个对话场景下的对话能力（例如：小度智能音箱的点歌功能就是一个技能）。分为开发者自定义技能和 UNIT 平台预置技能。自定义技能又分为「对话技能」和「问答技能」，完全由用户配置。预置技能为 UNIT 平台为了使开发者降低成本而预置的通用场景的对话能力，部分预置技能目前支持添加词槽词典值来干预对话理解效果。

### 1.2.3 对话技能

对话技能在多轮对话中可以准确理解用户需求（意图）和实现需求的关键信息（词槽），适用于需要根据不同对话意图和词槽信息给出不同对话答复的对话场景。能够在一些特定的对话场景下高精度地理解并满足用户需求，提升用户体验和满意度。比如问天气技能、订火车票技能、听音乐技能等等。

### 1.2.4 问答技能

问答技能适用于问题问法多样，但答案相对固定的对话场景，不需要根据用户对话话语中的关键信息来设定不同的答复内容。比如针对各种规则制度、政策法规等信息的问答技能。

### 1.2.5 对话意图

对话意图（intent），是对话技能要理解的用户需求（例如用户说”换到中央台”，那对应电视控制技能技能下的对话意图就是”换台”；”北京天气”，对话意图是查天气）。定义对话技能下的对话意图时还需要设置对话意图关联的词槽（满足对话意图的必须条件），以及技能理解对话意图后给用户的回应。

### 1.2.6 词槽

是满足用户对话意图时的关键信息或限定条件，可以理解为用户需要提供的筛选条件。例如在查询天气时，词槽是地点和时间。

例如：“换到中央台”中的”中央台”就是一个”电视台词槽”，它会一定程度上影响系统对”换台”这个对话意图的执行。

## 1.2.7 词典

每个词槽都需要对应一套词典，技能的专名词识别服务会结合词槽绑定的词典来识别用户问话中的词槽。您可以自行导入词典，对话技能进行词槽识别时，会优先考虑自定义词典中的词汇。

## 1.2.8 对话模板

对话模板是您给对话系统按具体语法、句式做出的示范，教它在某一个特定语法、句式中，该如何理解对话意图，哪个词是重要信息，对应的词槽、特征词是什么。

例如：“[D:sys\_loc][D:sys\_time]天气如何”，上述标注表示可以将所有满足“[城市]+[时间]+天气如何”这一规则的 query 解析为 WEATHERINFO 对话意图。如“北京今天天气如何？”，“天津明天天气如何”等。其中“[D:sys\_loc]”表示所有城市词组成的集合，“[D:sys\_time]”表示所有时间描述组成的集合。

UNIT 中一个对话模板也可以使用多条模板片段组成，实现按片段去匹配用户 query，实现更强的对话意图泛化匹配能力，提高模板对用户 query 的召回率。

## 1.2.9 特征词

特征词表示具有某一类特征的词，通常被用于约束某条对话模板的匹配范围(例如：天气、天儿等类别关键词)或提供一定限度的泛化能力(例如订机票对话中『从北京到上海』的“到”就可以与“去、飞“等助词一起作为特征词)；活用特征词机制可以事半功倍的提高对话模板的精度和覆盖度。

## 1.2.10 特征词词典

您需自行导入词典，用于系统识别特征词。

## 1.2.11 对话样本

对话样本就是真实对话场景下用户表达目的的句子，需要您给系统做示范，教它在用户的具体句子里该如何理解对话意图，哪个词是重要信息，对应的词槽是什么。

例如：通过对话样本标注告诉机器人“三亚明天会不会下雨”与“三亚明日会下雨吗”都是询问天气的语句，其中“三亚”是对应城市 city 这个词槽，“明日”和“明天”都是 time 词槽。

这样的训练越多，机器人的理解能力便越强，这与在学习语言中的人类孩童的学习方式也是十分相似的。

## 1.2.12 对话样本集

是管理对话样本的集合，类似文件夹的作用，可以把不同来源的对话样本放在不同的对话样集中，也可以把不同时间段获取的对话样本分在不同的对话样本集中。

## 1.2.13 问答对

问题与答案的组合，称之为问答对。问答对支持一对一、一对多、多对一和多对多；当某一个问答对中包含多个答案时，答案随机呈现。

## 1.2.14 标签

对问答对进行分类标签化的管理，可以打在问答对上，对问答对进行标签化的管理，可以按业务分类、问题所属业务部门等自己需要的维度来管理问答对，提高问答对的管理效率。也可以在机器人调用接口和问答技能的调用接口中传入标签信息，在特定标签范围内的问答对中匹配相似问题。

## 1.2.15 模型

是您配置对话逻辑、导入训练数据后通过 UNIT 平台内置的各种学习引擎训练出的技能核心文件，模型需生效至沙盒环境后才可发挥作用。

## 1.2.16 沙盒环境

沙盒环境是 UNIT 平台提供给您创建技能，训练技能模型，对话测试，调试优化技能模型的环境。每个技能都配有一个沙盒环境，将训练好的技能模型生效到沙盒环境后，就可以与技能进行对话了，同时可接入到您自己的业务系统中使用。一个技能可以训练出多个版本的模型，但只能选择一个放到沙盒环境中。

## 1.2.17 生产环境

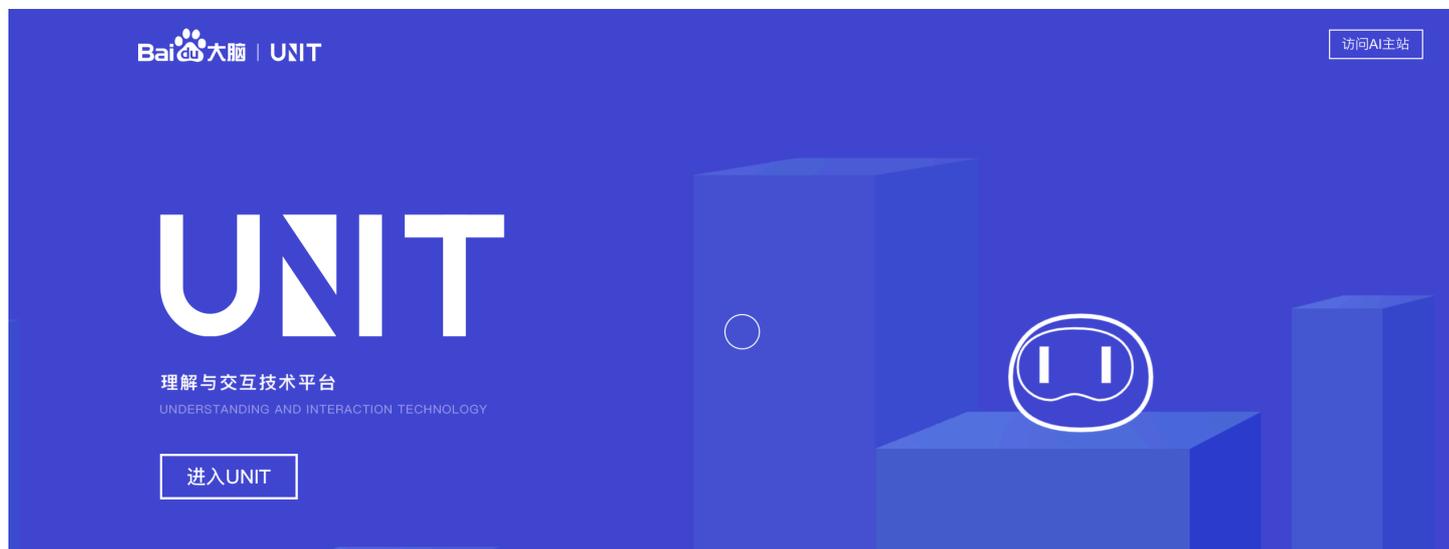
生产环境是 UNIT 平台在百度云上为您提供的可定制的稳定的对话服务环境。

## 1.2.18 图谱知识库

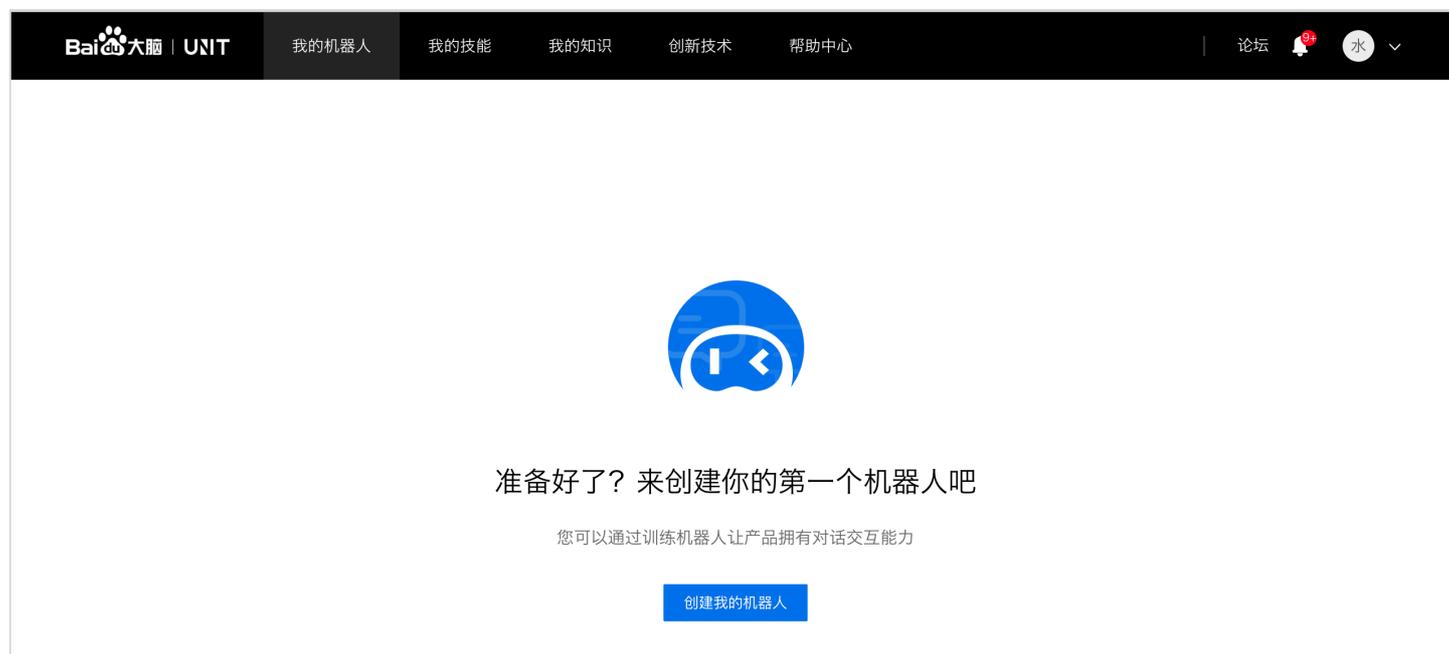
知识图谱 (Knowledge Graph) 本质上是语义网络, 是一种基于图的数据结构, 由节点(Point)和边(Edge)组成。在知识图谱里, 每个节点表示现实世界中存在的“实体”, 每条边为实体与实体之间的“关系”。知识图谱是关系的最有效的表示方式。更多知识图谱介绍见论坛: <http://ai.baidu.com/forum/topic/show/943731>

## 2 登录 UNIT

- 1) 用浏览器访问<http://unit.baidu.com> 进入UNIT系统官网，点击“进入UNIT”进入UNIT管理后台。

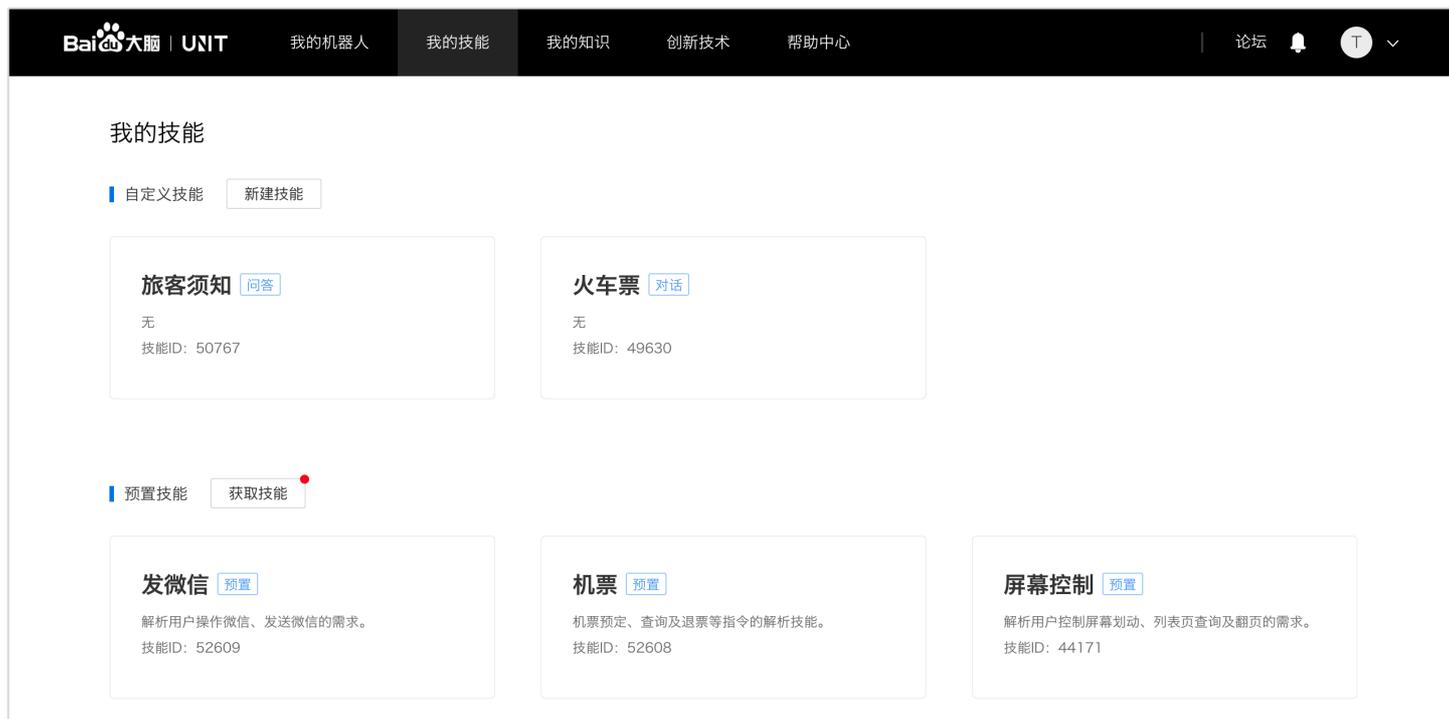


- 2) 如果尚未登录，请登录百度云。成功登录后，可以看到UNIT包含了“我的机器人”、“我的技能”、“我的知识”、“创新技术”4个模块。



## 3 我的技能

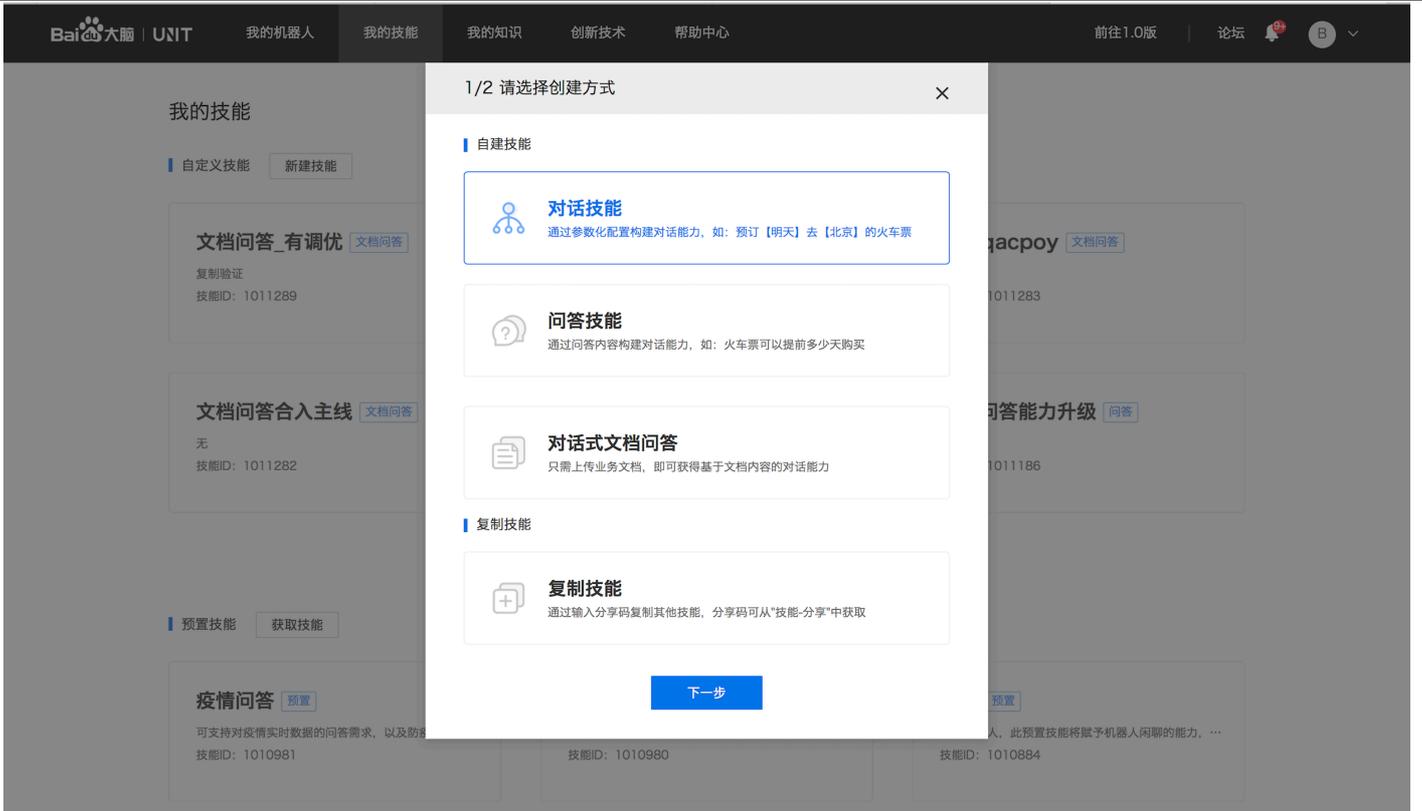
在某一个对话场景下的对话能力（例如：小度智能音箱的点歌功能就是一个技能），分为开发者自定义技能和 UNIT 平台预置技能。自定义技能又分为「对话技能」和「问答技能」，完全由用户配置。预置技能为 UNIT 平台为了使开发者降低成本而预置的通用场景的对话能力，部分预置技能目前支持添加词槽词典值来干预对话理解效果。在 UNIT 中我们区分“对话技能”、“问答技能”、“对话式文档问答技能”、“预置技能”4 种技能。



### 3.1 技能管理

#### 3.1.1 新建技能

- 1) 在“我的技能”页面可以新建“自定义技能”和“预置技能”，这里我们演示如何创建“自定义技能”。
- 2) 创建自定义技能有3种方式，可以选择自建对话技能或问答技能，也可以通过分享码复制自己或他人账号下已经创建好的技能及其数据。这里我们先选择“对话技能”，下一步填写技能信息。



### 技巧及注意点——我们在一个场景下应该创建对话技能还是问答技能？

- **对话技能**适合对话流程灵活的对话任务，如火车票的订、退、改签。需要识别用户意图并提取必需的关键信息，接着判定对话流程给出不同的答复。
- **问答技能**适合答复内容固定的对话任务，如旅客须知。只需匹配到问题就能返回对应的答复。
- **我的机器人**可以同时调用对话技能和问答技能，来满足同一场景多种类型的对话任务。

3) 技能信息包含“技能名称”与“技能描述”，完成填写后，点击创建技能。

2/2 请填写技能信息 ×

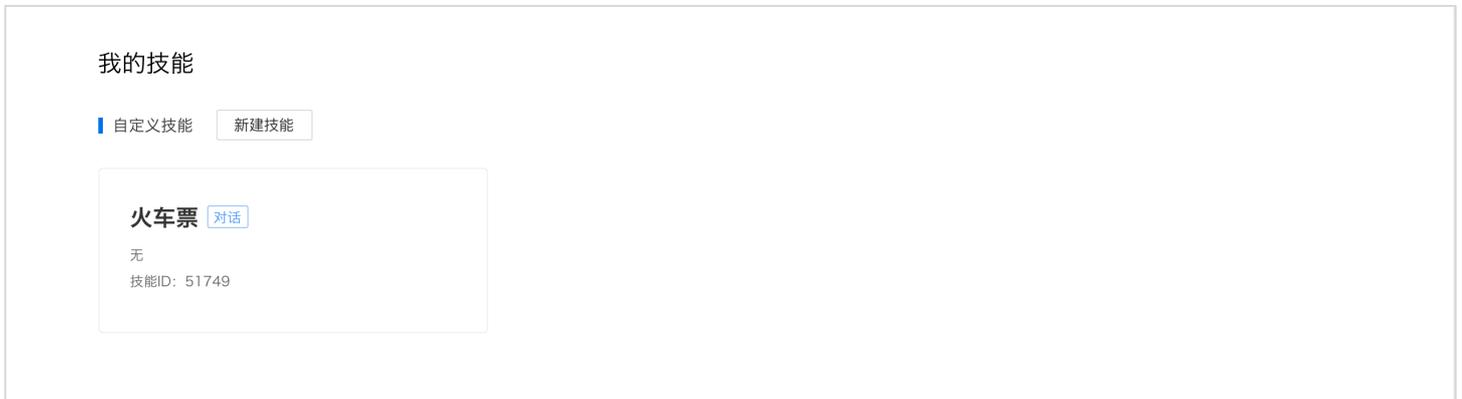
技能类型: 对话技能

\* 技能名称:

技能描述:

- 技能名称: 支持中文、英文、数字与”\_”, 最大长度30, 为必填项, 支持修改。
- 技能描述: 支持中文、英文、数字与常用符号, 最大长度50, 为非必填项, 支持修改。

4) 成功创建技能后, 会在技能列表中显示新创建的技能卡片。技能卡片除了展示技能名称与技能描述之外, 还会展示技能类型与技能ID。



- 技能类型: 您创建技能时选择的技能类型, 目前平台包含对话、问答、创新、预置四种技能类型。
- 技能ID: 您创建技能时, 系统会为其分配的唯一标识。

### 3.1.2 删除技能

点击技能卡片右上角的“操作”按钮, 显示操作列表。点击“删除该技能”后会要求输入技能名称进行二次确认。

**火车票** 对话

无  
技能ID: 51749

分享技能  
重命名技能  
删除该技能

删除技能 ×

技能删除后不可恢复，所有与其相关的信息都将删除，您确定要删除技能 "火车票" 吗？

删除

### 3.1.3 重命名技能

点击技能卡片右上角的“操作”按钮，显示操作列表。点击“重命名技能”后会可修改技能名称和技能描述。

**旅游方式** 对话

无  
技能ID : 72095

分享技能  
重命名技能  
删除该技能

### 重命名技能 ×

技能类型: 对话技能

\*技能名称:  ×

技能描述:

### 3.1.4 分享技能

- 1) 分享技能是指把已经创建好的技能及其数据在任意账号下复制多个备份。
- 2) 点击技能卡片右上角的“操作”按钮，显示操作列表。点击“分享技能”后可以复制分享码。

### 分享技能 ×

技能分享码: **2905a4**

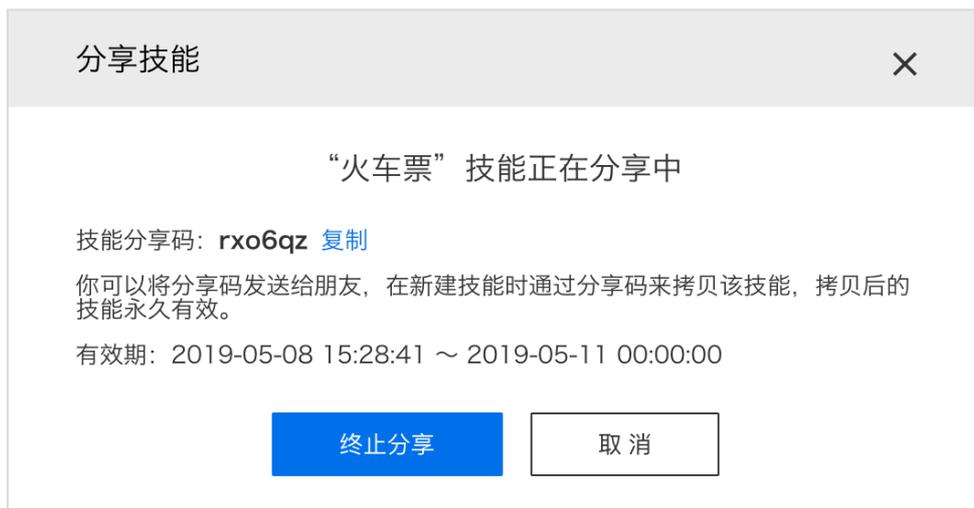
你可以将分享码发送给朋友，在新建技能时通过分享码来拷贝该技能，拷贝后的技能永久有效。

请选择分享码有效时长:  ▼

- 3) 在分享有效期内（3/7/15/30 天），利用这个分享码可以在任何账号下复制这个技能。然后在新建自定义技能时，选择创建方式为“复制技能”，并在下一步输入“分享码”即可复制被分享的技能。



终止分享：当完成分享后，可以提前结束分享有效期。点击技能卡片右上角的“操作”按钮，显示操作列表，点击“查看分享状态”，接着点击弹窗中的“终止分享”来结束当前或后续的复制任务。



## 3.2 自定义-对话技能

对话技能在多轮对话中可以准确理解用户需求（意图）和实现需求的关键信息（词槽），适用于需要根据不同用户意图和词槽信息给出不同对话答复的对话场景。能够在一些特定的对话场景下高精度地理解并满足用户需求，提升用户体验和满意度。比如问天气技能、订火车票技能、听音乐技能等等。

### 3.2.1 意图管理

对话意图 (intent)是对话技能要理解的用户需求（例如用户说”换到中央台”，那对应电视控制技能技能下的对话意图就是”换台”；”北京天气”，对话意图是查天气）。定义对话技能下的对话意图时还需要设置对话意图关联的词槽（满足对话意图的必须条件），以及技能理解对话意图后给用户的回应。

意图管理包含设置意图基本信息、设置关联词槽、设置对话回应方式及其内容。



自定义技能 > 编辑对话意图

**设置意图基本信息**

- 意图名称: BOOK\_TICKET
- 意图别名:
- [+ 添加别名 ?](#)
- 描述:

**设置关联词槽 ?**

[添加词槽](#)

词槽名称	词槽别名	词典来源	词槽必填	澄清话术	澄清...
user_time	出发时间	自定义词典 / 系统词典	非必填	什么时间出发	1 上移 下移 ...

**对话回应 ?**

[<> 新推出DM Kit离线开源对话流程管理模块，可实现复杂对话流程的管理](#)

#### 3.2.1.1 设置意图基本信息

意图基本信息包含意图名称、意图别名、描述。意图名称是意图的唯一标识。意图别名是意图名称的口语化表述，建议为中文。

### 设置意图基本信息

意图名称: BOOK\_TICKET

意图别名:

订票

+ 添加别名 ?

描述:

实现订票的对话任务

- 意图名称: 支持大写英文、数字与下划线“\_”，最大长度30，为必填项，不支持修改。
- 意图别名: 只支持中文、英文、数字，最大长度20，数量为1-10个，支持修改。
- 意图描述: 支持中文、英文、数字与常见符号，最大长度50，非必填项，支持修改；

### 3.2.1.2 设置关联词槽

词槽是满足用户对话意图时的关键信息或限定条件，可以理解为用户需要提供的筛选条件。例如在查询天气时，词槽是地点和时间。

例如：“换到中央台”中的“中央台”就是一个“电视台词槽”，它会一定程度上影响系统对“换台”这个对话意图的执行。

### 设置关联词槽 ?

添加词槽

词槽名称	词槽别名	词典来源	词槽必填	澄清话术	澄清...	
user_time	出发时间	自定义词典 / 系统词典	非必填	什么时间出发	1	上移 下移 ...
user_from	出发站点	自定义词典 / 系统词典	非必填	您想从哪出发	2	上移 下移 ...
user_to	到达站点	自定义词典 / 系统词典	非必填	您要去哪	3	上移 下移 ...
user_trainnumber	车次	自定义词典 / 系统词典	非必填	请选择您的车次	4	上移 下移 ...

点击“添加词槽”打开新建词槽弹窗，添加词槽有两种方式，分别为新建自定义词槽和复用自定义词槽。这里演示如何新建自定义词槽：

1) 第一步 设置词槽的添加方式。选择添加方式为“新建自定义词槽”，填写词槽名称和词槽别名。词槽名称是词槽的唯一标识。词槽别名是词槽名称的口语化表述，建议为中文。

- 词槽名称：支持小写英文、数字与下划线“\_”，最大长度15，为必填项，不支持修改，默认以“user\_”开头。
- 意图别名：只支持中文、英文、数字，最大长度20，数量为1-10个，支持修改。

2) 第二步 为词槽选择词典。新建自定义词槽提供两种词典配置方式——自定义词典和复用系统词槽词典。



- 自定义词典由“自定义词典值”和“词典值黑名单”两部分组成。
  - 自定义词典值：开发者枚举的词槽可以匹配的值；
  - 词典值黑名单：开发者枚举的词槽匹配时需要屏蔽的值；
- 系统词槽词典：UNIT预置的词典，具有词量丰富、覆盖全面的特点，能有效提高开发效率。

**技巧及注意点——选择词典是，如果系统词槽词典能满足您的业务，尽量选它！**

我们建议在能选择词典时，尽量选择使用系统词槽词典，因为这能极大提升您的开发效率和技能理解能力！当系统词槽词典无法满足您的业务需求时可以添加自定义词典。UNIT 系统词槽说明：

<https://ai.baidu.com/forum/topic/show/869949>

3) 第三步 设置词槽与意图关联属性。

新建词槽 ×

1.设置添加方式 - 2.选择词典 - 3.设置词槽与意图关联属性

\*词槽必填: ?

必填     非必填

\*澄清话术: ?

普通澄清话术     枚举澄清话术 (词典值数量=0或>5时不可选)

请告诉我您的出发时间

+ 添加话术

澄清 3 轮后放弃要求澄清

上一步
确定

- **词槽必填**：选择“必填”，在用户提问中如果缺少该词槽值，技能会强制用户进行澄清，在澄清成功之前，不会触发其他任何动作，直至成功或超出最大澄清次数(理解失败)；且优先级高于『场景技能回应』中的触发规则
- **澄清话术**：普通澄清话术：每种澄清话术最多可以设置5条，在技能需要用户澄清词槽的情况下可以随机使用此处设置的不同话术来询问用户，让人机交互不那么单调。
- **枚举澄清话术**：当词槽词典值数量为1-5个时（选择任一系统词典时词典值数量肯定大于5），技能可自动在对话中枚举词典值（<enum\_value\_list>）让用户选择来完成词槽值的确认，比如词典值是您上传的一个商品套餐，共有3中套餐，可由技能自动枚举出来让用户选择澄清。
- **澄清轮数**：设置一个必须澄清的词槽最多澄清多少次，默认是3次。超过最大澄清次数技能将放弃要求澄清。

4) 第四步 点击确认成功添加一个词槽。按此步骤依次创建当前意图需要的词槽。创建的词槽会在列表中进行展示。

**设置关联词槽** ?

添加词槽

词槽名称	词槽别名	词典来源	词槽必填	澄清话术	澄清...	
user_time	出发时间	自定义词典 / 系统词典	非必填	什么时间出发	1	上移 下移 ...
user_from	出发站点	自定义词典 / 系统词典	非必填	从哪出发	2	上移 下移 ...
user_to	到达站点	自定义词典 / 系统词典	非必填	要到哪	3	上移
user_trainnumber	车次	自定义词典 / 系统词典	必填	请选择您的车次	4	上移

编辑词典值  
 编辑别名  
 解除关联

- 词槽列表中可设置词槽上下顺序，当多个词槽需要终端用户填充时，技能会按照列表从上往下的顺序进行对话澄清，来获取槽值。您可通过点击操作项中的**上移**或**下移**调整词槽澄清顺序。
- 在“...”展开菜单中，可以编辑词典值、编辑别名、把词槽从当前意图中解除关联。解除关联后，对应的词槽会从当前词槽列表中移除，但不会从技能中删除；

### 3.2.1.3 对话回应

技能对话回应支持三种方案：1.在 UNIT 平台上配置，2.使用 Taskflow 配置（云端对话流程管理模块，可实现复杂对话流程的管理），3.使用 DM Kit 配置（离线开源对话流程管理模块，可实现复杂对话流程的管理），您可在“技能设置-高级设置-对话回应设置”中选择使用哪种回应方案。



#### 3.2.1.3.1 在 UNIT 平台上配置

在UNIT平台上配置对话回应，支持两种动作：**答复**、**引导至对话意图**。**答复**是指让当前意图直接回复用户。**引导至对话意图**是指通过对话互动，将用户引导至别的对话意图上，。

两种动作可以任选其一，也可以组合使用。如果组合使用，可配置动作优先级，即同时满足触发规则时，技能优先以哪种类型的动作回应；默认答复优先。

回应方式设置
✕

请选择回应动作的类型：

答复

引导至对话意图

请选择回应动作的优先级：

1. 答复                      上移 下移

2. 引导至对话意图        上移 下移

确定
取消

### 3.2.1.3.1.1 答复

答复是指让当前意图直接回复用户，可以是固定的文本内容，也可以是一个标识（需要执行的函数名称）。

**答复** ?

文本内容  执行函数

为您找到以下车次和座位信息，确认下单吗

触发规则：

范围	词槽	关系/状态	内容	
且 会话过程中	user_time	已填充		X
且 会话过程中	user_from	已填充		X
且 会话过程中	user_to	已填充		X

+ 添加规则

+ 添加规则组

- 文本内容：支持中文、英文、数字与常见符号，最大长度50，为必填项，支持修改。
- 执行函数名称：支持英文、数字与下划线“\_”，最大长度30，为必填项，支持修改。

### 3.2.1.3.1.2 引导至对话意图

引导至对话意图是指通过对话互动，将用户引导至别的对话意图上。

**引导至对话意图** ?

1. 引导话术:

目标对话意图:

对话意图	展现文本内容	
WEATHER	哪里的天气怎么样	X
BOOK_TICKET	我还想再定一张火车票	X
BOOK_TICKET	再定一张票	X

+ 添加目标对话意图

触发规则:

范围	词槽	关系/状态	内容	
您还未设置触发规则，默认根据意图直接触发本条回应! <a href="#">马上创建</a> 或 <a href="#">智能生成</a>				

- 1) 首先配置引导话术，这个话术会反馈给用户，用于引导用户进一步答复。
- 2) 接着配置目标对话意图及其展示文本内容，在引导话术播报给终端用户后，终端用户的答复匹配到展示文本内容时，会命中对应的意图。
- 3) 单个引导话术支持最多配置3个目标对话意图。但是您可以添加多个引导话术及其触发规则。

### 3.2.1.3.1.3 触发规则

针对答复、引导至对话意图这两类动作，都可以配置触发规则。创建了规则后，在条件满足时，才会触发对应的动作。未创建触发规则，则在任意条件下都可以触发这个动作。

触发规则:

范围	词槽	关系/状态	内容	
且 会话过程中	user_time	已填充		X
且 会话过程中	user_from	已填充		X
且 会话过程中	user_time	已填充		X
且 会话过程中	user_trainnumber	已填充		X

+ 添加规则

+ 添加规则组

每条规则由3~4部分组成，包含如下：

- 范围：当前规则应用范围，必选项；

- 会话过程中：终端用户和系统的整个对话过程；若您设置过重置对话状态条件，会话过程范围则截止至最近一次重置对话状态时。
- 当前用户输入：本次对话用户输入的语句。
- 上轮对话意图：和用户对话过程中用户，用户的上一个意图；您设置的对话状态重置条件会影响当前条件的判断。
- 词槽：具体对应哪一个词槽，必选项；当“范围”选择的是“上轮对话意图”时，由于规则与“意图”相关，此时“词槽”项无需编辑。
- 关系/状态：词槽的关系或状态，必选项；
  - 已填充/未填充：用户输入的语句中，是否含有相关词槽的信息；
  - 包含/未包含：用户输入的语句中，是否含有相关词槽的信息且信息中是否包含某个值；
  - 等于/不等于：用户输入的语句中，是否含有相关词槽的信息且信息是否等于某个值；
- 内容：需要您填写相关内容，支持中文、英文、数字与常用符号，最大长度30。

点击“添加规则”可以增加规则项；点击“添加规则组”可以增加规则组，答复动作的触发规则中可以创建多个规则组，每个规则组间是“或”的关系，满足任意一组规则组即触发满足动作。

### 3.2.1.3.2 使用 Taskflow 配置

如果平台上默认的对话回应设置无法满足您更加个性化的对话流程控制，您可以选择线上开源的Task flow，对于复杂对话逻辑需求也可以借助Taskflow的云端编程能力在较低成本下实现。同时Taskflow中还提供资源调用能力，一键接入领域资源，让对话链路更加完整。需要注意的是，Taskflow只能在「对话技能」中使用，不能在问答技能中使用。在对话技能中的「技能设置--高级设置」的「对话回应设置」里选择「使用Taskflow」配置。



### 3.2.1.3.3 使用 DM Kit 配置

如果平台上默认的对话回应设置无法满足您更加个性化的对话流程控制，您可以选择下载离线开源的 DM Kit，在本地部署搭建 DM Kit 服务。从云端获取 UNIT 解析结果，在本地的 DM Kit 中对技能的回复及处理流程进行管理控制。如果在“技能设置-高级设置-对话回应设置”改成**使用 DM Kit 配置**，则不在需要在 UNIT 平台上配置技能回应。

对话回应 



当前正在使用离线开源对话流程管理 DM Kit

- 您已选择了“使用DM Kit配置”， 如果想在UNIT平台上配置（可实现单个意图的技能回应设置）， 请到[高级设置](#)中选择
- 使用DM Kit可以与平台的云端解析能力无缝对接， 获取技能模型解析的意图和词槽结果并完成复杂对话流程的流转控制

## 3.2.2 训练数据

技能的训练数据包含对话模板、对话样本、词典。

### 3.2.2.1 对话模板

对话模板是您给对话系统按具体语法、句式做出的示范，教它在某一个特定语法、句式中，该如何理解对话意图，哪个词是重要信息，对应的词槽、特征词是什么。

例如：“[D:sys\_loc][D:sys\_time]天气如何”，上述标注表示可以将所有满足“[城市]+[时间]+天气如何”这一规则的 query 解析为 WEATHERINFO 对话意图。如“北京今天天气如何？”，“天津明天天气如何”等。其中“[D:sys\_loc]”表示所有城市词组成的集合，“[D:sys\_time]”表示所有时间描述组成的集合。

UNIT 中一个对话模板也可以使用多个模板片段组成，实现按片段去匹配用户 query，实现更强的对话意图泛化匹配能力，提高模板对用户 query 的召回率。

< 返回我的技能

火车票

ID 51749

测试

意图管理

训练数据

对话模板

对话样本集

词槽管理

技能训练

日志分析

技能发布

技能设置

▲ 上移 ▼ 下移 删除模板 模板片段内容 搜索对话模板 口语化词管理

<input type="checkbox"/>	意图	模板内容	模板片段顺序	必须匹配		
		预订	0	是		
		[D:user_time]	0	否		
<input type="checkbox"/>	BOOK_TICKET	从	1	否		
		[D:user_from]	2	否	☑	...
		到	3	否		
		[D:user_to]	4	否		
		火车票	5	是		
		[D:user_time]	0	否		
<input type="checkbox"/>	BOOK_TICKET	[D:kw_booking]	1	是		
		[D:user_from]	2	否	☑	...
		[D:kw_to][D:user_to]	3	否		
		[D:kw_ticket]	4	是		
		[D:user_time]	0	否		
<input type="checkbox"/>	BOOK_TICKET	[D:kw_booking]	0	是		
		[D:user_trainnumber]	0	否	☑	...
		[D:kw_ticket]	0	是		
		[D:kw_seattype]	0	否		
<input type="checkbox"/>	BOOK_TICKET	[D:kw_booking][D:user_trainnumber][W:0-1][D:kw_ticket]	0	是	☑	...
<input type="checkbox"/>	BOOK_TICKET	[D:kw_booking][D:user_from][D:kw_to][D:user_to][W:0-1]	0	是	☑	...
<input type="checkbox"/>	WEATHER	我要	0	否	☑	...
		[D:kw_weather]	0	是		

### 3.2.2.1.1 新建对话模板

选择意图:

必须匹配	模板片段 <small>?</small>	顺序 <small>?</small>	
是	预订	0	+
			-
			×
否	[D:user_time]	0	+
			-
			×
否	从	1	+
			-
			×
否	[D:user_from]	2	+
			-
			×
否	到	3	+
			-
			×
否	[D:user_to]	4	+
			-
			×
是	火车票	5	+
			-
			×

+ 新增一行

插入词槽:

1) 选择意图: 当对话模板匹配到用户的语句时, 就会命中为模板选择的意图。

选择意图:

搜索意图名或意图别名

模板片段	意图名称	意图别名	顺序 <small>?</small>	
<input type="text" value="BOOK_TICKET"/>	订票		0	+
				-
				×
<input type="text" value="WEATHER"/>	查天气		0	+
				-
				×
<input type="text" value="SYS_OTHER"/>	SYS_OTHER		0	+
				-
				×
<input type="text" value="CMD_UPDATE_USER_FR..."/>	澄清出发站点		1	+
				-
				×
<input type="text" value="CMD_UPDATE_USER_TIME"/>	澄清出发时间		1	+
				-
				×
<input type="text" value="CMD_UPDATE_USER_TO"/>	澄清到达站点		2	+
				-
				×
否	到		3	+
				-
				×
否	[D:user_to]		4	+
				-
				×
是	火车票		5	+
				-
				×

置空意图

- 自定义的意图：开发者自己创建的意图。
- SYS\_OTHER 意图：因为技能模型具备泛化能力，如果您希望部分语句不要被识别，那么可以把对应的语句标注到这个意图上。
- 词槽澄清意图 (CMD\_UPDATE\_USER\_\$\_slot.name)：开发者每新建一个词槽，平台就会自动生成一个词槽澄清意图，提供填充或更新词槽值的识别能力。

2) 添加模板片段：符合一定规则的语义片段的抽象表达，如"到"，"到[目的地]"等。

- 模板片段的内容：
  - 文本
  - 词槽——[D:user\_xxx]：一类词的抽象，满足意图的必须条件或关键信息；用户可以新建自定义词槽，复用系统词典或者自己上传自定义词典，自定义词槽必须以 user\_开头。格式符合
  - 特征词——[D:kw\_xxx]：由口语化词、辅助表述词、标点符号等对用户输入的文本内容真实理解无影响的词组成，支持多个，用户可自定义且解析结果中不展示。特征词词典需要用户自己上传或添加。
  - 通配符——[W:x-x]：可以匹配指定长度的任意字符串；例如："[W:0-10]"，表示可匹配总字节数在 0 到10 之间的任意字符串。注意：一个汉字占两个字节，通配符不要过长，尽量不要在模板片段中只写单个通配符，否则有可能会影响模板的识别效果。
- 是否必须匹配：query在匹配当前模板时，是否必须存在当前模板片段所描述的局部语义片段
- 模型片段顺序：query 所需满足的一组模版片段之间的相对顺序，在匹配 query 时，位置小的模版片段在位置大的模版片段的左侧，其中位置为 0 表示可匹配 query 中的任意位置。

3) 设置阈值：query中可识别部分占总query长度的比例达到多少时，该query可被识别为所标注的意图。其中可识别成分包括词槽、特征词、口语化词；识别但不计入可识别部分的为通配符匹配的内容。

4) 解析是否返回所有词槽设置：默认选“是”，解析返回用户query中可以识别的所有词槽。如果选“否”，则只解析返回用户query中当前模板里插入的词槽。对话模板标注说明见：<http://ai.baidu.com/forum/topic/show/940955>

### 3.2.2.1.2 编辑对话模板

点击编辑图标在弹框中打开当前模板，操作项与新建相同。

意图		模板内容	模板片段顺序	必须匹配		
<input type="checkbox"/>	BOOK_TICKET	预订	0	是		...
		[D:user_time]	0	否		
		从	1	否		
		[D:user_from]	2	否		
		到	3	否		
[D:user_to]	4	否				
		火车票	5	是		

### 3.2.2.1.3 调整模板优先级

当我们有多个对话模板时，它们之间是有优先级的，在对话模板列表的上面的优先级要高于下面的，可以选中一条对话模板，然后执行上移、下移、置顶、置底的操作来调整优先级。

意图		模板内容	模板片段顺序	必须匹配		
<input type="checkbox"/>	BOOK_TICKET	预订	0	是		...
		[D:user_time]	0	否		
		从	1	否		
		[D:user_from]	2	否		
		到	3	否		
[D:user_to]	4	否				
		火车票	5	是		
<input checked="" type="checkbox"/>	BOOK_TICKET	[D:user_time]	0	否		...
		[D:kw_booking]	1	是		
		[D:user_from]	2	否		
		[D:kw_to][D:user_to]	3	否		
		[D:kw_ticket]	4	是		
<input type="checkbox"/>	BOOK_TICKET	[D:user_time]	0	否		...
		[D:kw_booking]	0	是		
		[D:user_trainnumber]	0	否		
		[D:kw_ticket]	0	是		
		[D:kw_seatype]	0	否		
<input type="checkbox"/>	BOOK_TICKET	[D:kw_booking][D:user_trainnumber][W:0-1][D:kw_ticket]	0	是		...
<input type="checkbox"/>	BOOK_TICKET	[D:kw_booking][D:user_from][D:kw_to][D:user_to][W:0-1]	0	是		...
<input type="checkbox"/>	WEATHER	我要	0	否		...
		[D:kw_weather]	0	是		

### 3.2.2.1.4 导入/导出对话模板

您可以导入本地文件，文件中的模板将覆盖线上当前技能中全部对话模板；也可以导出当前技能中全部对话模板；

对话模板列表
配置说明

+ 添加对话模板

导入
导出

▲ 上移 ▼ 下移
🗑️ 删除模板
模板片段内容 ▼
搜索对话模板
口语化词管理

☐	意图 ▼	模板内容	模板片段顺序	必须匹配	
		预订	0	是	
		[D:user_time]	0	否	
		从	1	否	
☐	BOOK_TICKET	[D:user_from]	2	否	🗑️ ...
		到	3	否	
		[D:user_to]	4	否	
		火车票	5	是	

#### 【离线文件导入格式】

```
ROUTE 0.8 [D:user_start]###1#@#0#@## [D:kw_to]###2#@#1#@## [D:user_end]###3#@#1 1
WEATHERINFO 0.8 [D:sys_loc]###0#@#1#@## [D:sys_time]###0#@#1#@##天气###0#@#1 0
```

#### 【示例第一行格式解读】

其中上例中ROUTE表示当前模板意图，0.8表示阈值，使用#@##分割模板中的多个模板片段，#@##分割每个模板及其属性，最后的“1/0”表示“是/否”在解析结果中返回所有词槽

则上例中，第一行模板有三个模板片段：

[D:user\_start]为第一个模板片段，1表示顺序，0表示非必填

[D:kw\_to]为第二个模板片段，2表示顺序，1表示必填

[D:user\_end]为第三个模板片段，3表示顺序，1表示必填

#### 【格式说明】

- 1、每行编辑一个对话模板，上传模板时需上传模板对应意图及阈值约束，置于行首，模板中模板片段间用#@##分割，每个模板片段有顺序、是否必填属性，这两种属性与模板片段间用#@##分割。
- 2、tab键用于分割意图、阈值、模板片段组和是否返回所有词槽，切勿使用空格分割。
- 3、模板间排列顺序会影响优先级，首行模板优先级最高，编写时请注意模板排序。
- 4、是否返回所有词槽，如果不写，导入时会默认填写

### 3.2.2.1.5 新建/编辑特征词

1) 新建特征词，需要填写特征词名称、描述、词典值。



- 特征词名称：支持限小写英文、数字与下划线“\_”，最大长度20，为必填项，不可修改，须以“kw\_”开头；
- 特征词描述：支持中文、英文、数字与常用符号，最大长度50，为非必填项，可修改；
- 词典值：一行一个，第一次新建时在弹框中最多支持创建200个词典值，可以在创建后继续追加词典值。

2) 在创建特征词后从特征词列表中点击某个特征词进入词典值管理页面，支持导入/导出自定义词典值、导入/导出词典值黑名单，黑名单为自定义词典屏蔽词词典，用于屏蔽某个词典；优先级高于自定义词典值；

对话模板 > 特征词 kw\_seattype 📄 导入 📄 导出

搜索全部词典 🔍 搜索

**词典值 (4词)** 🗑️ 清空词典值

点击Enter添加自定义词典值

词典值	
无座	🗑️
硬座	🗑️
硬卧	🗑️
软卧	🗑️

< 1 > 共1页 4条

**词典黑名单 (0词)** 🗑️ 清空词典黑名单

点击Enter添加词典值黑名单

### 3.2.2.1.6 口语化词管理

口语化词将在语义解析时会被自动忽略掉，以此来提高解析的精准率。比如通用场景下：嗯我看一下明天的天气吧 --> 明天天气，这里把"嗯、我、看一下、的、吧"等口语化词都去掉了。又比如特定场景里：预定如家酒店 --> 预订如家，在定酒店的场景里省略了“酒店”，可以将“酒店”配置为口语化的词。

**对话模板列表** ⓘ 配置说明

⊕ 添加对话模板 📄 导入 📄 导出

▲ 上移 ▼ 下移 🗑️ 删除模板 模板片段内容 🔍 搜索对话模板 口语化词管理

<input type="checkbox"/>	意图 ▼	模板内容	模板片段顺序	必须匹配	
		预订	0	是	
		[D:user_time]	0	否	
		从	1	否	
<input type="checkbox"/>	BOOK_TICKET	[D:user_from]	2	否	📄 ...
		到	3	否	
		[D:user_to]	4	否	
		火车票	5	是	

平台预置了200多个口语化词，开发者可以进行增删改。调整后也可以“恢复系统默认值”

搜索全部词典 Q 搜索

**词典值 (217词)**

清空词典值 恢复系统默认值

点击Enter添加词典值，单个词典值不能超过6个字符

词典值	
啦	
!	
?	
。	
,	
哈	
暖	

### 3.2.2.2 对话样本

对话样本就是真实对话场景下用户表达目的的句子，需要您给对话系统做示范，教它在用户说的具体句子里，该如何理解对话意图，哪个词是重要信息，对应的词槽是什么。

例如：通过对话样本标注告诉机器人“三亚明天会不会下雨”与“三亚明日会下雨吗”都是询问天气的语句，其中“三亚”是对应城市 city 这个词槽，“明日”和“明天”都是 time 词槽。

这样的训练越多，机器人的理解能力便越强，这与在学习语言中的人类孩童的学习方式也是十分相似的。

[< 返回我的技能](#)

**火车票**  
ID 51749

测试

意图管理

训练数据 ▾

对话模板

对话样本集

新建对话样本集
合并样本集
推荐对话样本 NEW

<input type="checkbox"/>	ID	名称	样本集来源	标注进度	操作
<input type="checkbox"/>	1049	<a href="#">查天气</a>	自建	0 / 0	导入 导出 ...
<input type="checkbox"/>	1048	<a href="#">订票</a>	自建	0 / 0	导入 导出 ...

< 1 > 共1页 2条

### 3.2.2.2.1 新建对话样本集

- 1) 对话样本集：点击“新建对话样本集”在弹出的窗口中填写对话样本集名称。支持中文、英文、数字与下划线，最大长度30



### 3.2.2.2.2 标注对话样本

- 1) 进入一个对话样本集后，对话样本列表中展示意图、对话样本及其标注的词槽，标注状态。



- 2) 添加一条样本

对话样本集 > 订票 导入 导出

帮我订明早的火车票

筛选:    共2条, 已标注2条, 未标注0条, 待确认0条

已选择0条样本

<input type="checkbox"/>	意图	对话样本	标注状态
<input type="checkbox"/>	BOOK_TICKET	订 [D:user_from:北京] 到 [D:user_to:上海] 的高铁票, [D:user_time:明天上午] 的	已标注 ...
<input type="checkbox"/>	BOOK_TICKET	[D:user_from:北京] 到 [D:user_to:上海] [D:user_time:周五晚上] 的票	已标注 ...

### 3) 为其标注意图

已选择0条样本

<input type="checkbox"/>	意图	对话样本	标注状态
<input type="checkbox"/>	请标注意图	帮我订明早的火车票	未标注 ...
<input type="checkbox"/>	BOOK_TICKET	订票	已标注 ...
<input type="checkbox"/>	WEATHER	查天气	已标注 ...
<input type="checkbox"/>	SYS_OTHER	SYS_OTHER	已标注 ...
<input type="checkbox"/>	CMD_UPDATE_USER_FR...	澄清出发站点	已标注 ...
<input type="checkbox"/>	CMD_UPDATE_USER_TIME	澄清出发时间	已标注 ...
<input type="checkbox"/>	CMD_UPDATE_USER_TO	澄清到达站点	已标注 ...

15 < 1 > 共1页 3条

### 4) 为其标注词槽

已选择1条样本

<input type="checkbox"/>	意图	对话样本	标注状态
<input checked="" type="checkbox"/>	BOOK_TICKET	帮我订 [明早] 的火车票	待确认 <input type="button" value="确认"/> ...
<input type="checkbox"/>	BOOK_TICKET	订 [D:user_from:北京] 到 [D:user_to:上海] 的高铁票, [D:user_time:明天上午] 的	已标注 ...
<input type="checkbox"/>	BOOK_TICKET	[D:user_from:北京] 到 [D:user_to:上海] [D:user_time:周五晚上] 的票	已标注 ...

词槽名      词槽别名

user\_from      出发站点

user\_to      到达站点

user\_trainnumber      车次

**user\_time**      出发时间

15 < 1 > 共1页 3条

### 5) 确认标注状态

已选择1条样本 修改意图 删除样本 批量确认

<input type="checkbox"/> 意图	对话样本	标注状态
<input checked="" type="checkbox"/> BOOK_TICKET	帮我订 [D:user_time:明早] 的火车票	待确认 <span>确认</span> ...
<input type="checkbox"/> BOOK_TICKET	订 [D:user_from:北京] 到 [D:user_to:上海] 的高铁票， [D:user_time:明天上午] 的	已标注 ...
<input type="checkbox"/> BOOK_TICKET	[D:user_from:北京] 到 [D:user_to:上海] [D:user_time:周五晚上] 的票	已标注 ...

15 < 1 > 共1页 3条

### 3.2.2.2.3 合并样本集

合并样本集操作将吧指定的样本及其数据合并为同一个样本。

新建对话样本集 合并样本集 推荐对话样本 NEW

<input type="checkbox"/>	ID	名称	样本集来源	标注进度	操作
<input checked="" type="checkbox"/>	1049	查天气	自建	0 / 0	导入 导出 ...
<input checked="" type="checkbox"/>	1048	订票	自建	3 / 3	导入 导出 ...

### 3.2.2.2.4 推荐对话样本

基于已有的对话样本，利用百度大数据挖掘对话样本，标注后快速提升技能模型的理解能力。

新建对话样本集 合并样本集 推荐对话样本 NEW

<input checked="" type="checkbox"/>	ID	名称	样本集来源	标注进度	操作
<input checked="" type="checkbox"/>	1049	查天气	自建	0 / 51	导入 导出 ...
<input checked="" type="checkbox"/>	1048	订票	自建	3 / 3	导入 导出 ...

< 1 > 共1页 2条

为了保证挖掘出质量更高的数据，需要参与勾选的对话样本集至少包含 50 条对话样本。

你可以调节推荐策略及参数：

- 策略 A-基于百度大数据挖掘：从百度搜索大数据中挖掘相似的用户 query。
- 策略 B-基于同义的推荐：基于同义词，生成同义 query。

推荐对话样本
✕

推荐策略及参数:

策略A-基于百度大数据挖掘 ?
 策略B-基于同义的推荐 ?

推荐样本数量上限:  ?

策略A与策略B样本比例  :

### 3.2.2.2.5 导入导出样本集

支持为单个对话样本集导入对话样本。也支持从单个对话样本集中导出所有的对话样本。

☐	ID	名称	样本集来源	标注进度	操作
☐	1049	<a href="#">查天气</a>	自建	0 / 51	导入 导出 ...
☐	1048	<a href="#">订票</a>	自建	3 / 3	导入 导出 ...

#### 【离线文件导入导出格式】

北京明天天气如何?

明天上海天气怎样?   SYS\_WEATHER

广州明天是什么天气?    SYS\_WEATHER

杭州明天是什么天气?    SYS\_WEATHER sys\_loc:杭州 sys\_time:明天

成都后天是什么天气?    SYS\_WEATHER sys\_loc:成都 sys\_time:后天

南京后天是什么天气?    SYS\_WEATHER sys\_loc:成都 sys\_time:后天    cfmd **【格式说明】**

- 1、每行编辑一条对话样本，最大长度 125。
- 2、上传对话样本时可以仅上传对话样本，也可以在上传对话样本的同时上传相应的标注结果和标注状态；标注结果包含意图和词槽，没有词槽时只标意图即可（未标意图只标注词槽或标注的词槽不属于意图的视为无效标注结果，无效标注结果上传时会移除，但不影响对话样本内容上传）。
- 3、标注没有词槽的对话样本只需对话样本后直接添加意图即可。
- 4、标注对话样本词槽请按 XXX:YYY 格式上传，XXX 代表词槽名称，YYY 代表取词范围。
- 5、tab 键用于分隔对话样本、对话样本意图与对话样本词槽。

6、如果在导入之前对包含意图、词槽的样本进行了人工确认，确认无误的可以在样本后面加上 cfmd 标识，导入 UNIT 后就直接是"已标注"状态，没有带 cfmd 标识的导入后有意图、词槽的样本信息是"待确认"状态。只有"已标注"状态的样本才能用于深度学习训练。

### 3.2.2.3 词槽管理

词槽名称	词槽别名	词典来源	关联意图	操作
user_from	出发站点	自定义词典 / 系统...	CMD_UPDATE_USER_FROM ; BOOK_TICKET	编辑别名 删除
user_to	到达站点	自定义词典 / 系统...	CMD_UPDATE_USER_TO ; BOOK_TICKET	编辑别名 删除
user_time	出发时间	自定义词典 / 系统...	CMD_UPDATE_USER_TIME ; BOOK_TICKET	编辑别名 删除
user_trainnumber	车次	自定义词典 / 系统...	CMD_UPDATE_USER_TRAINNUMBER ; BO...	编辑别名 删除

词槽管理列表中包含了当前技能所有的词槽，您可以在这里重新配置词槽的别名和词典。点击单行词槽，进入词槽的词典配置页面。

词槽列表 > 词槽 user\_from (出发站点)

搜索全部词典 [搜索]

**系统词典**  on

sys\_loc\_country  
  sys\_loc\_province  
  sys\_loc\_city  
  sys\_loc\_county  
  sys\_loc\_train\_station  
 修改

**自定义词典**  on ?

点击Enter添加自定义词典值

清空自定义词典值

自定义词典值 (0词)

- 自定义词典由“自定义词典值”和“词典值黑名单”两部分组成。
  - 自定义词典值：开发者枚举的词槽可以匹配的值；
  - 词典值黑名单：开发者枚举的词槽匹配时需要屏蔽的值；
  - 系统词槽词典：UNIT预置的词典，具有词量丰富、覆盖全面的特点，能有效提高开发效率。

## 3.2.3 技能训练

### 3.2.3.1 训练并生效新模型

当您的技能已经定义了自定义技能里的意图，并在训练数据中加入了对话模板或对话样本（并进行了标注）等，就可以训练技能了。



- 训练数据包含对话模板、对话样本集：
  - 对话模板：默认生效全部模板，训练时引用技能中最新模板信息；
  - 对话样本集：包含自建样本和系统负例样本：
- 训练策略包含快速生效、深度训练。
  - 快速生效：对上面所选数据使用更快速的训练方式，能够保证效果的快速迭代
  - 深度训练：使用深度学习模型，对上面所选数据进行深度学习，能够保证学习到更精细的区分模型，但效率较低。
- 描述：记录本次模型训练做了哪些调整和设置，支持中文、英文、数字与常用符号，最大长度50，选填项；

**技巧及注意点**——我们升级了训练策略，把机器学习与对话模板的规则技术进行了融合，在调用 api 时选择设置不同的参数将可以获取不同的解析效果和结果，

详见：<https://ai.baidu.com/forum/topic/show/868426>

### 3.2.3.2 模型管理

< 返回我的技能

**火车票**

ID 51749

测试

**沙盒环境** ?

模型版本: v3

更新时间: 2019-05-07 14:01:12

状态: 运行中 ?

训练并生效新模型

**模型列表** ?

版本	描述	训练时间	训练进度 ?	操作
v3		2019-05-07 14:00:53	● 训练完成	运行中 删除
v2		2019-05-07 14:00:26	● 训练完成	生效到沙盒 删除
v1		2019-05-07 14:00:02	● 训练完成	生效到沙盒 删除

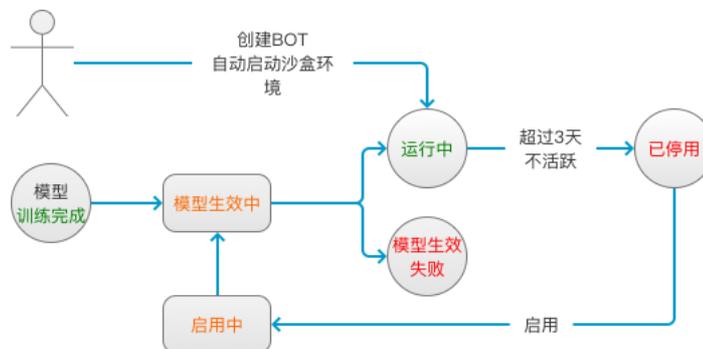
模型管理页面显示了当前技能的沙盒环境状态（包含沙盒中当前技能模型的版本、更新时间、状态）和模型列表。

#### 1) 沙盒环境信息：

- 模型版本：沙盒中当前技能模型的版本；
- 更新时间：沙盒环境最近一次更新的时间；
- 状态：该技能沙盒环境的状态，分为“启用中”、“运行中”、“模型生效中”、“模型生效失败”和“已停用”：
  - 启用中：沙盒环境处在正在启动过程中，在这个状态时不能训练新模型。
  - 运行中：沙盒环境已经启动完成，处于运行状态时可以训练模型或把已有模型生效到沙盒，当运行中沙盒里有技能模型时还可以与技能对话（可以在unit平台上的对话窗口对话，也可以使用对话api接口）；
  - 模型生效中：目前生效一个模型大概需要30秒左右。如果训练一个新模型，训练完成后会自动把最新的模型生效到沙盒，沙盒中原有的模型会被替换。也可以单独把不在沙盒中的模型生效到沙盒；
  - 模型生效失败：模型生效异常，此时可以重试，但多次重试都还失败时可加入 UNIT 百度 Hi 群：1617262 或 QQ 群：805312106 联系值班客服协助解决；
  - 已停用：沙盒环境因长时间未使用处于停止状态，需重启使用；

沙盒环境的状态在技能列表页也会展示，点击技能列表中沙盒状态为“已停用”旁边的“启动”，也可启动沙盒环境。

沙盒状态变成流程：



当技能超过5天没有任何操作（平台网站没有任何操作，对话API也没有调用），UNIT系统会自动停用当前不活跃技能的沙盒环境，收回相关资源供需要的技能使用，如需继续使用被停用的技能沙盒环境，可手动启用。模型列表展示有版本、描述、训练时间、进度信息，可操作生效至沙盒和删除。

- 版本：每次训练都会生成一个新版本的模型。
- 描述：模型在训练时您填写的模型信息，可以记录本次训练做了哪些调整，方便后续回顾对比；
- 训练时间：模型被训练的开始时间；
- 训练进度：模型当前进度，包含初始化、训练中、训练完成、训练失败4种状态：
  - 初始化：此时不可进行生效到沙盒或删除操作；
  - 训练中：此时不可进行生效到沙盒或删除操作；
  - 训练完成：执行『训练并生效新模型』并训练完成后会立即执行模型生效的操作，此时也不可进行任何操作；
  - 训练失败：模型训练失败，只可删除，不可生效到沙盒；
- 操作：
  - 生效至沙盒：模型列表中不在沙盒环境中的模型当训练状态为训练完成时，可以单独『生效到沙盒』替换沙盒中已有的模型。
- 删除：删除该模型；

### 3.2.4 测试



- 对话记录区：展示您和技能对话内容，左侧为技能“说”的内容，右侧为您输入的内容；技能解析结果中会含有答复、动作、意图、词槽信息和查看JSON按键：
  - 答复：技能答复给您或终端用户的内容；
  - 动作：技能当前答复对应的动作名称；
  - 意图：技能对当前对话样本解析出的意图；
  - 词槽：技能对当前对话样本解析出的词槽；
  - 查看JSON：点击后，可以查看技能解析的详细信息，包括当前意图和词槽是用哪个策略识别的，如果是对话模板识别的，还会包含具体是哪条模板识别的（如果当前意图和词槽识别错了，可以据此找到对应的对话模板进行修改调整）；
- 重置对话状态：会清空技能记录的所有历史对话信息（例如之前技能识别的意图、词槽信息）；

- 操作区下方式一个文本框，您可直接再次输入对话样本信息去验证技能的实例，按enter发送消息。在消息输入框中按键盘的向上键可以掉出历史消息快速发送。

**技巧及注意点说明**——UNIT平台上的测试对话窗口只是供你「测试体验」对话理解效果的，实际你把技能在集成到自己的产品中，可以根据UNIT平台解析的意图词槽进行一系列的处理给出用户多种形态的反馈，比如给用户播放一首歌、一段视频、一个候选文字链list，一个图文消息等等。

## 3.2.5 日志分析

### 3.2.5.1交互学习日志

UNIT 新增了在对对话过程中 用户可以人工干预技能解析结果，让对话更加灵活。

返回我的技能

火车票  
ID 51749

测试

意图管理

训练数据

技能训练

日志分析

交互学习日志

时间范围: 2019-05-03 23:24:16 至 2019-05-09 23:24:16

已选择0条记录 批量清除 加入训练数据 从训练数据移出

搜索被反馈对话

反馈时间	被反馈对话	反馈人	反馈来源	模型识别意图	反馈后意图	加入训练数据	操作
05-09 23:...	我要北京西站西广场	UNIT_DEV...	web测试	BOOK_TICKET user_to: 北京西站	TAXI user_to: 北京西站	未加入   加入	查看JSON   清除

15 < 1 > 共1页 1条

下面举例说明：

- 1) 这是模型对“我要北京西站西广场”解析的结果，认为是“订火车票”。

我要北京西站西广场

请选择您的车次

动作: book\_ticket\_user\_trainnumber\_clarify

意图: BOOK\_TICKET

词槽: user\_to 北京西站

[添加词槽](#)

[查看JSON](#)

2) 用户开始通过对话表达自己的真实想法，是想要“打车”。

我不要订票，我说的是打车

好的正在为你打车

动作: taxi\_satisfy

意图: TAXI

词槽: 无

[添加词槽](#)

[查看JSON](#)

3) 交互学习后，对应用户再次说出同样的语句，模型会解析为“打车”而不是“订火车票”。

我要北京西站西广场

好的正在为你打车

动作: taxi\_satisfy

意图: TAXI

词槽: 无

[添加词槽](#)

[查看JSON](#)

上图中展示了通过对话人工干预模型解析结果的过程。并会产生一条交互学习日志

时间范围: 2019-05-03 23:24:16 至 2019-05-09 23:24:16

已选择0条记录 批量清除 加入训练数据 从训练数据移出 搜索被反馈对话

反馈时间   
 被反馈对话   
 反馈人   
 反馈来源   
 模型识别意图   
 反馈后意图   
 加入训练数据   
 操作

05-09 23:... 我要北京西站西广场 UNIT\_DEV... web测试 BOOK\_TICKET TAXI ● 未加入 | 加入 查看JSON | 清除  
 user\_to: 北京西站 user\_to: 北京西站

如果把这条交互学习记录「加入」训练数据，那在下次训练新的模型时，对应的反馈记录会参加训练。这样的反馈记录多了以后，训练出来的模型会越来越聪明。

### 3.2.5.2 日志分析

< 返回我的技能

火车票 ID 51749 测试

意图管理 训练数据 技能训练 日志分析 交互学习日志 对话日志 技能发布 技能设置

来源: 沙盒环境 注: 最新只能查看2小时前的日志 转成对话样本 导出日志明细 日志分析说明

2019-05-03 19:50:03 至 2019-05-09 17:50:03 更多筛选 对话内容 搜索对话内容

对话时间	用户问	模型答	模型动作	识别意图	置信度	识别槽位	user/log id
05-09 16:12:00	我要订机票	您还想查天气吗?	guide	BOOK_TICKET	100	未识别词槽	...
05-09 16:12:03	是	正在为您查询天气	satisfy	WEATHER	100	未识别词槽	...
05-09 16:12:05	我要订机票	您还想查天气吗?	guide	BOOK_TICKET	100	未识别词槽	...
05-09 16:12:07	不	您还想查天气吗?	guide	BOOK_TICKET	100	未识别词槽	...
05-09 16:12:10	不	您还想查天气吗?	guide	BOOK_TICKET	100	未识别词槽	...
05-09 16:10:42	我要订机票	请选择您的车次	clarify	BOOK_TICKET	100	未识别词槽	...
05-09 16:10:43	1	请选择您的车次	clarify	BOOK_TICKET	100	未识别词槽	...
05-09 16:10:50	我要订机票	请选择您的车次	clarify	BOOK_TICKET	100	未识别词槽	...
05-09 16:10:56	213	对不起,我没有理解澄清...	failure	BOOK_TICKET	100	未识别词槽	...
05-09 16:11:01	23	我不知道该怎么答复您。	failure	未识别意图	0	未识别词槽	...
05-09 16:09:58	我要订机票	请选择您的车次	clarify	BOOK_TICKET	100	未识别词槽	...

日志分析 在 UNIT 中是你优化技能对话理解效果 必不可少的重要工作。

它不仅可以帮助你发现技能识别错误的用户需求(意图)，还可帮助你发现当前技能没有覆盖的一些核心用户需求，以及技能已定义意图的使用频率。

定期分析用户对话日志，从中找出未识别意图与识别错误的意图，然后做进一步分析，并做出相应的优化，可以让你的技能对话理解效果越来越好。总详细介绍：<https://ai.baidu.com/forum/topic/show/865317>

## 3.2.6 技能发布

### 3.2.6.1 沙盒/生产环境

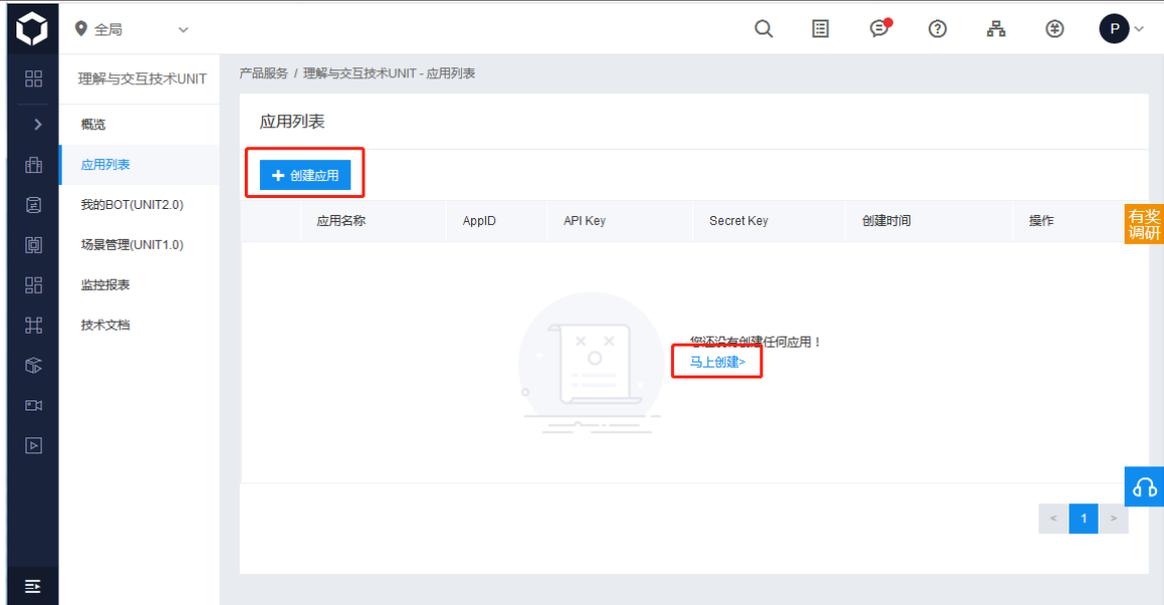
The screenshot shows the 'Skill Release' interface. On the left is a sidebar with navigation options: '返回我的技能', '火车票 ID 51749' (with a '测试' button), '意图管理', '训练数据', '技能训练', '日志分析', '技能发布', '沙盒/生产环境' (highlighted), '发布到DuerOS', and '接入微信'. The main content area is titled '沙盒/生产环境' and contains three sections: 1. '发布至沙盒环境': Explains that the sandbox is for training and testing, with a '测试窗口' and '对话API接口'. It notes that training automatically publishes the latest model to the sandbox, which has a default limit of 3 calls per second and unlimited daily calls. 2. '发布至生产环境': States that when the model in the sandbox reaches a certain effect, it can be published to the production environment. It notes that production environment QPS and call limits are the same as the sandbox but more stable, and can only be called via '对话API接口'. 3. '接口API Key / Secret Key': Provides instructions to check 'UNIT对话API文档' for interface addresses and parameters. It states that API Key / Secret Key should be obtained from the '百度云' application list, and that the same keys are used in both environments. A blue button '获取API Key / Secret Key' is located at the bottom of this section.

如果您需要在自己的 app、pc、web 端调用你的技能对话服务，需要先获取 API key/Secret Key，随后点击“获取 API key/Secret Key”按钮，进入百度云控制台：

#### 3.2.6.1.1 创建应用并关联技能

点击页面左侧导航中的“应用列表”，进入应用管理页。

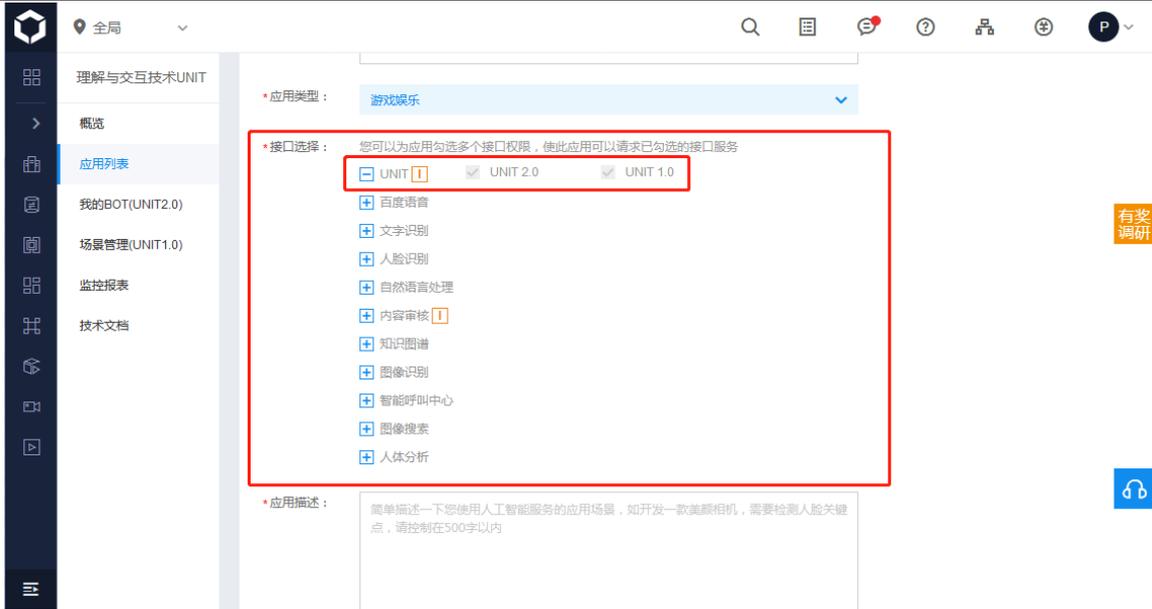
- 1) 点击页面中的“创建应用”进入创建应用页新建应用；未创建过应用时，也可以点击“马上创建”完成操作。



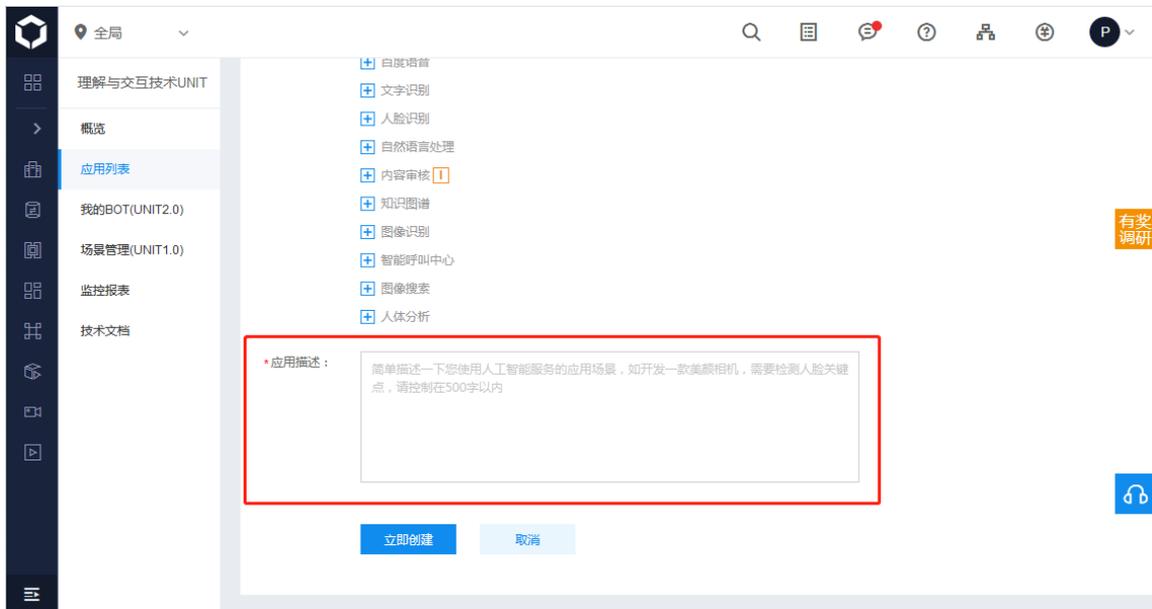
2) 在创建应用页，先填写应用名称、并选择应用类型。



3) 接着选择需要授权的接口，默认选中“UNIT”分类并勾选UNIT 2.0和UNIT 1.0的接口；除了“UNIT 技能”类以外，还可以选择其他方向接口，便于应用调用。



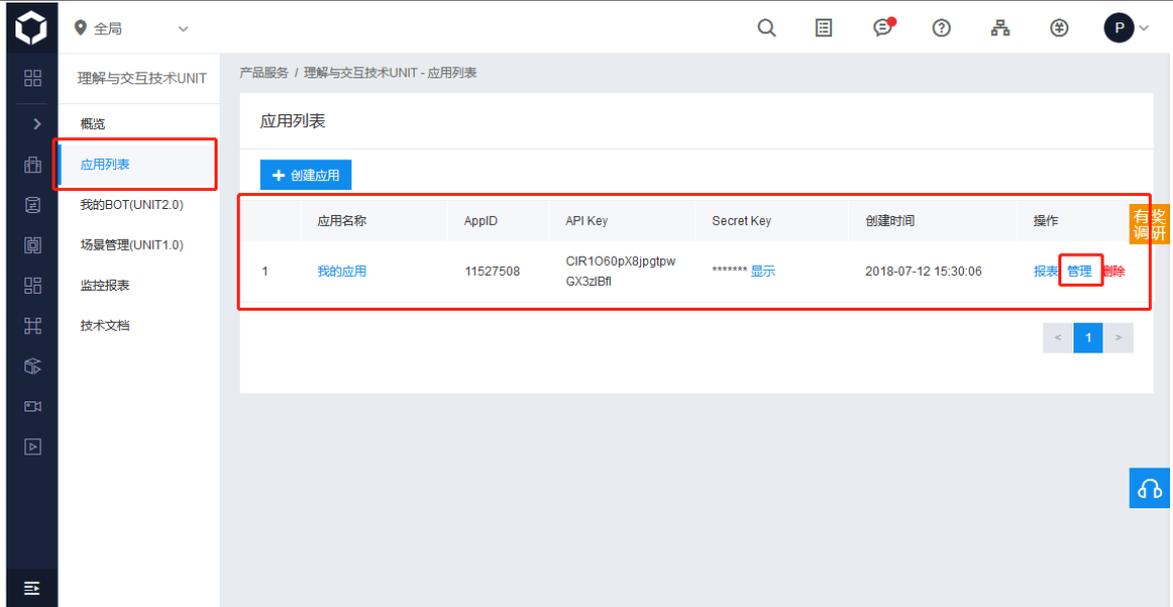
4) 然后按照实际情况填写应用描述信息。



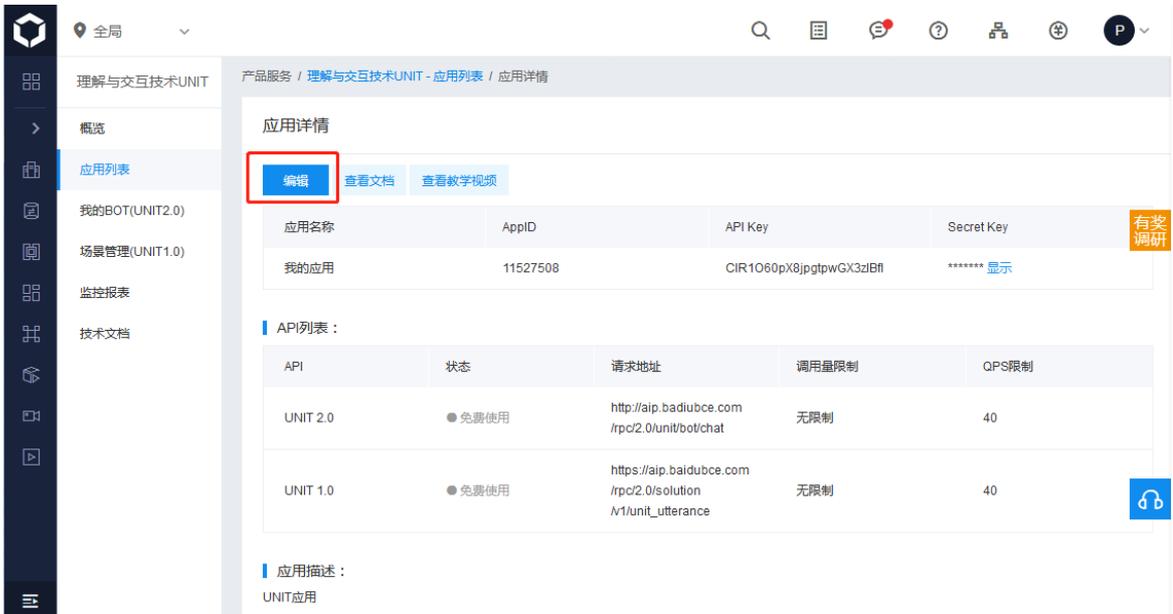
5) 编辑完成后，点击“立即创建”保存，完成应用创建。

### 3.2.6.1.2 获取 API Key 和 Secret Key

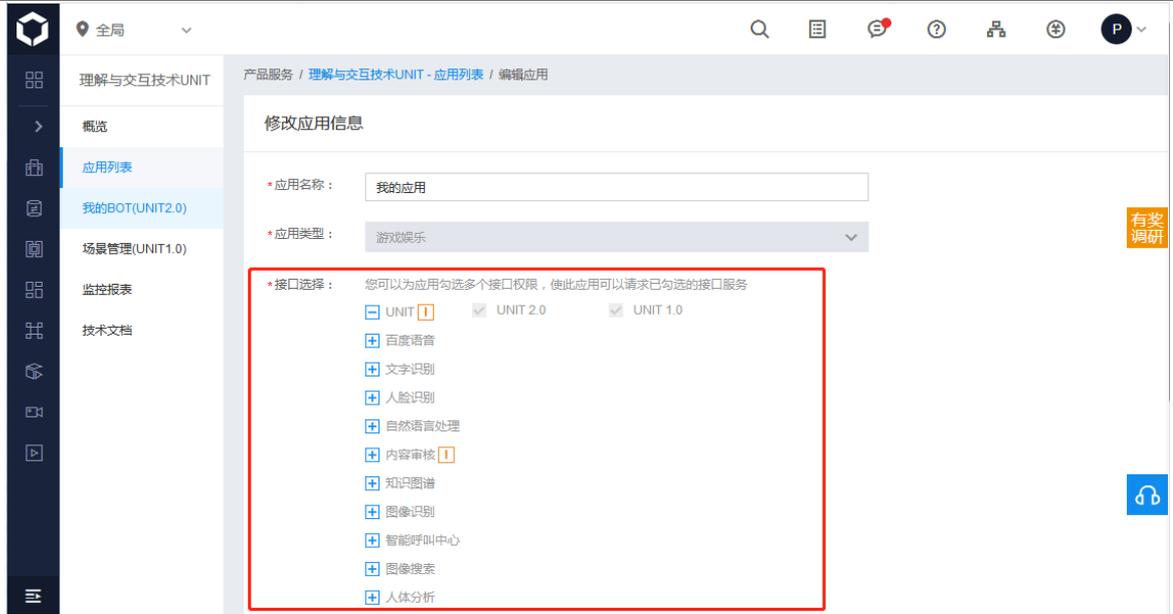
- 1) 点击页面左侧导航中的“应用列表”，进入应用列表页，即可拷贝API Key和Secret Key。
- 2) 在右侧列表中找到要修改的应用，点击“管理”进入应用详情页。



3) 在应用详情页面点击“编辑”，打开修改应用信息页。



4) 在修改应用信息页，修改关联接口和应用信息。



### 3.2.6.2 发布到 DuerOS

详细操作按指引流程完成即可。



### 3.2.6.3 接入微信

不仅仅是聪明的24h免费客服，会订酒店、机票的语音助理、风趣幽默的点餐小助手、旅行途中的解说专家、新手妈妈的百科全书、.....都可以在UNIT平台上快速搭建后与你的微信公众对接，让你的微信公众号智能有趣起来。

< 返回我的技能

**火车票**  
ID 51749

测试

意图管理

训练数据 >

技能训练

---

日志分析 >

技能发布 >

沙盒/生产环境

发布到DuerOS

**接入微信**

技能设置 >

#### 接入微信后能做什么

接入微信后你可在公众号上使用UNIT接收与回复用户咨询



#### 接入前准备

接入前您需要认证微信号并迁移自动回复

- 1 你的公众号必须是**认证过的微信订阅号或服务号**，否则无法正常回复顾客对话
- 2 授权后，公众号之前配置的自动回复可能和通过UNIT配置的回复冲突，建议您**将自动回复迁移至百度UNIT**
- 3 如果您使用外部资源，您需填写资源API地址，地址可以修改
- 4 一个微信公众号最多只能绑定一个技能模型，绑定第2个技能模型时之前绑定的技能模型会自动解除绑定

[授权微信公众号](#)

## 3.2.7 技能设置

### 3.2.7.1 高级设置

技能高级设置，设置项包括技能识别异常答复话术设置、清空对话session的条件设置等。

< 返回我的技能

火车票

ID 51749

测试

---

意图管理

训练数据 >

技能训练

---

日志分析 >

技能发布 >

技能设置 >

高级设置

DM Kit

如何优化

### 技能模型识别异常答复话术设置

语音识别异常

意图识别异常

词槽识别异常

动作名	动作描述	答复话术 <span style="font-size: 12px;">?</span>
asr_fail_action	完全未听清	我没听清你说的话，您能再说一遍吗?
asr_ambiguous_action	识别多个但不确定	您说的是<asr_value_list>[1]还是<asr_value_list>[2]?
asr_unsure_action	识别但不确定	您说的是<asr_value>吗?

重置答复话术

### 清空对话session的条件设置 ?

用户切换意图时: 不清空

用户意图达到满足时: 不清空

固定对话次数: 不清空

- 1) 技能识别异常答复话术设置，包括语音识别异常、意图识别异常、词槽识别异常 三种异常的答复话术设置，系统默认的话术可以修改，但话术中的”<abc.....>”结构的内容为变量，建议不要去掉，它可以动态获取对话中的意图、词槽的别名或值。修改系统默认话术后可以重置恢复默认话术。

### 技能模型识别异常答复话术设置

语音识别异常

意图识别异常

词槽识别异常

动作名	动作描述	答复话术 <span style="font-size: 12px;">?</span>
asr_fail_action	完全未听清	我没听清你说的话，您能再说一遍吗?
asr_ambiguous_action	识别多个但不确定	您说的是<asr_value_list>[1]还是<asr_value_list>[2]?
asr_unsure_action	识别但不确定	您说的是<asr_value>吗?

- 2) Unit平台支持切换意图清空session（历史词槽解析结果将被清空）、用户意图达到满足时清空、固定轮数清空3种方式。

- 用户切换意图清空：技能对终端用户输入内容解析出的意图与之前不一致时，清空重置对话session；
- 用户意图达到满足时清空：命中意图并完成词槽必填澄清后，就认为满足意图；
- 固定轮数清空：设置固定数值，清空N轮前的SESSION信息；

重置回复话术

**清空对话session的条件设置** ?

用户切换意图时: 不清空 ▾

用户意图达到满足时: 不清空 ▾

固定对话次数: 不清空 ▾

### 3) 对话回应设置:

在UNIT平台上配置对话回应，支持两种动作：**答复**、**引导到对话意图**。**答复**是指让当前意图直接回复用户。**引导到对话意图**是指通过对话互动，将用户引导至别的对话意图上，。

**对话回应设置** ?

在UNIT平台上配置

使用DM Kit配置（离线开源对话流程管理模块，可实现复杂对话流程的管理） [查看详情](#)

### 4) 技能训练-快速生效设置高级设置

**快速生效高级设置**

基于样本实例策略的召回阈值: 0.70   ?

- 阈值越高，模型基于已标注样本可泛化识别出的新说法更准确，但可泛化识别的说法种类会偏少；
- 阈值越低，模型基于已标注样本可泛化识别的新说法更多，但识别出来的样本的准确率会偏低

### 5) 技能训练-深度生效设置高级设置

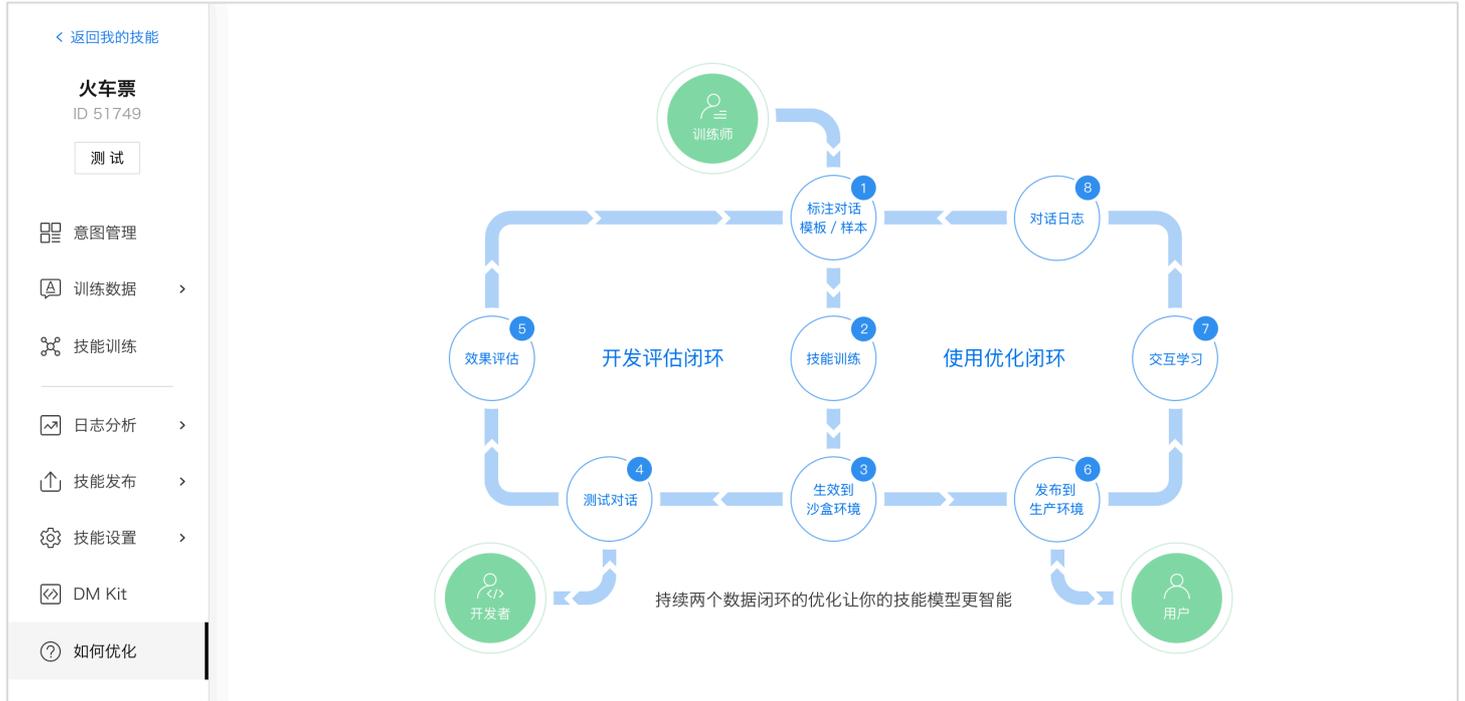
**深度训练高级设置**

意图槽位一体化深度训练:  启用  停用 ?

该策略训练耗时较长，5万样本训练预计耗时2.5小时，10万样本预计5小时。当您拥有大于5万的高质量对话样本标注数据时才建议启用该策略，启用后可在原有的深度训练模型上进一步提升整体对话理解效果。

### 3.2.8 技能测试如何优化

通过图示告诉开发者如何闭环优化技能模型。



### 3.3 自定义-问答技能

问答技能适用于问题问法多样，但答案相对固定的对话场景，不需要根据用户对话话语中的关键信息来设定不同的答复内容。比如针对各种规则制度、政策法规等信息的问答技能。

< 返回我的技能

**旅客须知**  
ID 50767

测试

问答管理

技能训练

技能设置

全部问答 / 3条

添加问答对

您可对第一个问题进行搜索

问题	问答对ID	标签
退票要加收多少费用	146285	退票
军人可以享受多大的优惠	3551	优惠 军人
多高的孩子乘坐火车可以半价票	2780	订票 优惠 儿童

15 < 1 > 共1页 3条

## 3.3.1 问答管理

### 3.3.1.1 问答对

问答集 > 编辑问答对

**标签**

订票 × 优惠 × 儿童 × +

**问题**

普通问题 问题模板

\* 标准问题:

相似问题:  最多可添加999个相似问题

**答案**

\* 答案1:

新增答案

- 问答对包含问题和答案，当终端用户的问题匹配到某一个问答对的问题时，就会随机返回对应的一个答案。
- 其中问题又划分为了普通问题和问题模板，用户语句与问题匹配时优先级为普通问题原句匹配>问题模板匹配>语义相似度匹配。
- 一个问答技能支持最多设定 100 万条问答对。

### 3.3.1.2 标签



- 标签可用于高效管理大量的问答对，每个问答对可打上 0-10 个标签，可以按标签快速灵活筛选问答对。
- 可以在调用 API 时，在请求参数中传入标签，将识别范围圈定为含有传入标签的问答对集合。您可以在测试对话框中体验这个功能。
- 一个问答技能支持最多设定 1000 个标签。

### 3.3.2 技能训练



模型列表展示有版本、描述、训练时间、进度信息，可操作生效至沙盒和删除。

- 版本：每次训练都会生成一个新版本的模型。
- 描述：模型在训练时您填写的模型信息，可以记录本次训练做了哪些调整，方便后续回顾对比；
- 训练时间：模型被训练的开始时间；
- 训练进度：模型当前进度，包含初始化、训练中、训练完成、训练失败4种状态：
  - 初始化：此时不可进行生效到沙盒或删除操作；
  - 训练中：此时不可进行生效到沙盒或删除操作；

- 训练完成：执行『训练并生效新模型』并训练完成后会立即执行模型生效的操作，此时也不可进行任何操作；
- 训练失败：模型训练失败，只可删除，不可生效到沙盒；
- 操作：生效至沙盒：模型列表中不在沙盒环境中的模型当训练状态为训练完成时，可以单独『生效到沙盒』替换沙盒中已有的模型；
- 删除：删除该模型。

### 3.3.3 测试

你可以在平台对技能进行测试



- 对话记录区：展示您和技能对话内容，左侧为技能“说”的内容，右侧为您输入的内容；技能解析结果中会含有答复、查看JSON按键：
  - 答复：技能答复给您或终端用户的内容；
  - 查看JSON：点击后，可以查看技能解析的详细信息，如果当前意图和词槽识别错了，可以据此找到对应的对话模板进行修改调整

- 操作区下方式一个文本框，您可直接再次输入对话样本信息去验证技能的实例，按enter发送消息。

**技巧及注意点说明**——UNIT 平台上的测试对话窗口只是供你「测试体验」对话理解效果的，实际你把技能在集成到自己的产品中，可以根据 UNIT 平台解析的意图词槽进行一系列的处理给出用户多种形态的反馈，比如给用户播放一首歌、一段视频、一个候选文字链 list，一个图文消息等等。

### 3.3.4 技能设置

可以在技能设置中调节问答阈值，修改技能设置后，需要重新训练方可生效。



## 3.4 自定义-对话式文档问答技能

### 3.4.1 文档上传

目前该技术仅支持 txt、word 格式的文档。

- 1) 在技能列表里点击新建技能的卡片，可进入技能训练的主页面：



2) 点击“上传文档”，会弹出以下界面，上传符合格式的文档；也可以直接下载示例文档进行体验。



完成文档类型选择，并点击“上传文档”后，进入以下界面，如果您需要调整要上传的文档，可以点击文档名右侧的“删除”icon；

上传文档 ×

- 当前版本可支持.txt、.docx格式的文档上传
- 建议给文档以恰当的名字命名，以能简明代表文档内容，这对于模型训练很重要，比如《海淀公园景区介绍》，而不是《文档1》
- 每篇文档的大小上限是10M
- 每个技能下文档数量的上限是99篇，文档的整体大小上限是10M
- [文档示例下载](#)

航空公司行李信息说明.txt 🗑️

确定 取消

3) 点击“确定”后，完成上传文档；

4) 文档上传完成后，进入以下界面，可点击“设置”中的“删除”删除当前文档，删除时会弹出二次确认的窗口；点击“下载”下载当前文档；

< 返回对话式文档问答

**行李信息说明**  
ID 51622

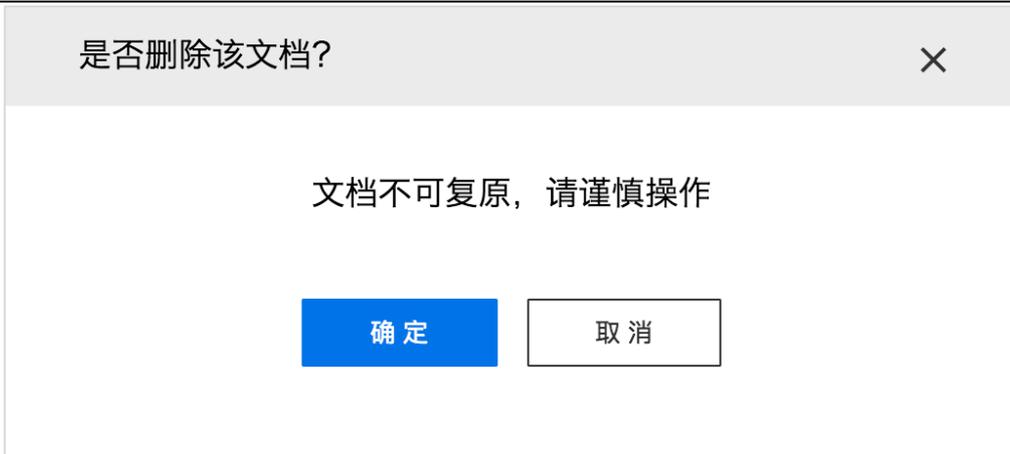
测试

文档训练 | 日志分析 > | 技能设置

重新训练 ● 训练已完成 ?

**上传文档**  
格式: txt 大小<10M 编码: utf-8

**航空公司行李信息说明.txt**  
随身携带物品一般规定: 1. 随身携带物品  
每位旅客以10  
大小 下载 删除



### 3.4.2 技能训练

- 1) 文档上传完成后进入以下界面，点击“训练”按钮，开始模型训练，训练过程中不可对文档进行任何编辑；

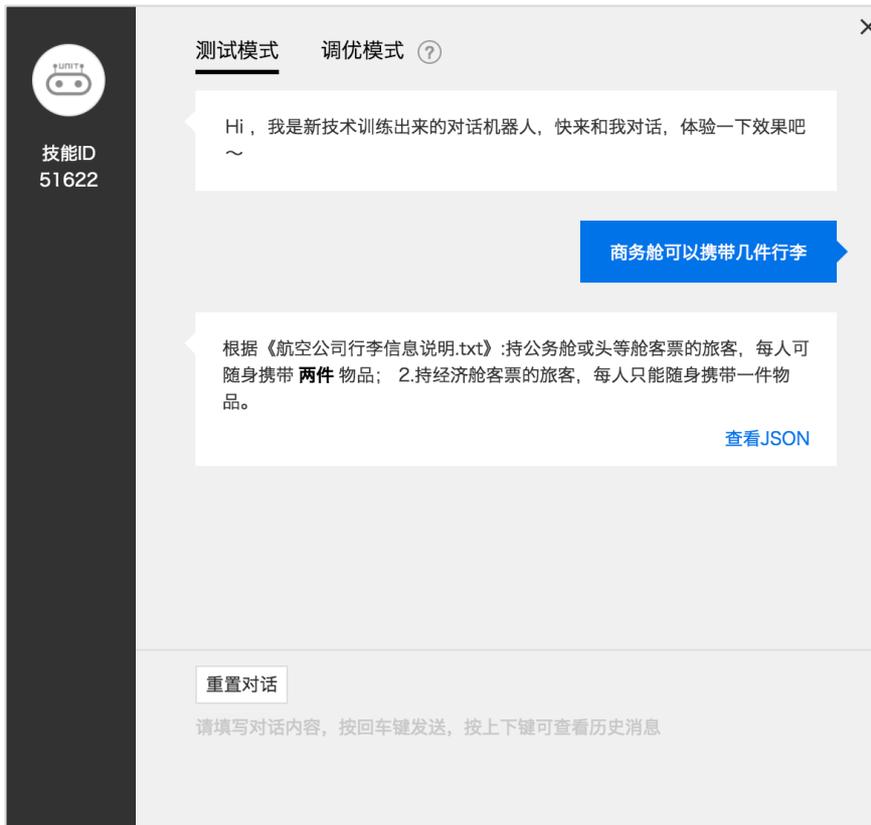


- 2) 训练完成后，界面如下，“训练”按钮变成“重新训练”，点击“测试”进入测试界面。



### 3.4.3 测试与调优

1) 点击“测试”进入以下界面，可以体验对话；对话有两种形态：一种是直接对话，如下图：



2) 一种是回复中提供选项，用户可以进行选择，或者直接输入更多信息



3) 如果对对话体验不满意，可以点击“调优模式”进入调优；



4) 调优模式下，可以展开每个候选项查看完整的选项内容：



5) 查看并确定正确选项后, , 点击单选按钮进行选择, 并点击确认, 提交并保存答案



6) 点击“确定”后, 会呈现针对这个问题完整的答案, 如下图:



7) 如果对之前选择的答案不满意，可以点击“删除并重新选择”按钮，进行重新选择



8) 点击“删除并重新选择”按钮后，会呈现选择前的选项状态；

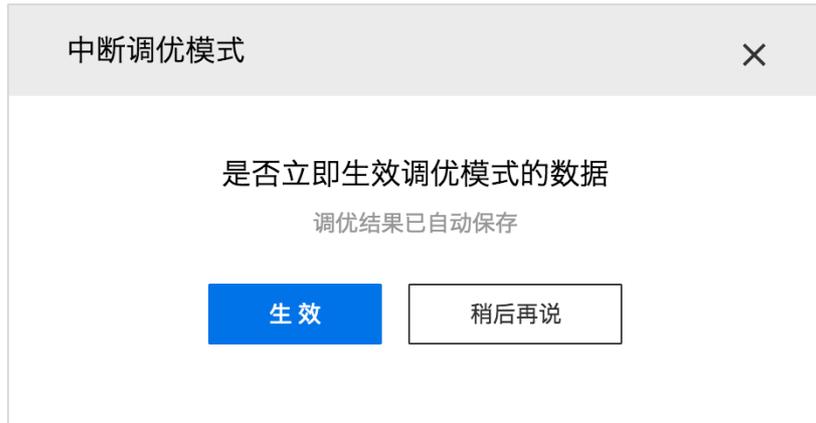


9) 调优完成后, 点击“生效”按钮, 调优结果会在模型中生效; 之后即可在“测试模式”下体验调优后的对话效果了;



10) 如果调优后没有点击“生效”就要离开调优界面, 会弹出二次确认, 点击“生效”后, 调优结果就生效

到模型中了；



### 3.4.4 日志分析

#### 3.4.4.1 交互学习日志

调优过程中产生的对话日志可在“交互学习日志”中实时查看；

< 返回对话式文档问答

**行李信息说明**  
ID 51622

测试

文档训练

日志分析

交互学习日志

对话日志

技能设置

时间范围: 2019-05-03 19:11:22 至 2019-05-09 19:11:22

已选择0条记录 批量删除 搜索被反馈对话

	反馈时间	被反馈对话	反馈前答案	反馈后答案	操作
<input type="checkbox"/>	05-09 19:00:06	重新选择	根据《航空公司行李信息说明.txt》:持公务舱或头等舱客票的旅客,每人可随身携带两件物品; 2.持经济舱客票的旅客,每人只能随身携带一件物品。	根据《航空公司行李信息说明.txt》:..... 2.持经济舱客票的旅客,每人只能随身携带一件物品。√ 根据《航空公司行李信息说明.txt》:3.旅客应将携带的液态物品(液体、.....气溶胶)盛放在容积不超过100毫升(ml)的容器内携带。√	<a href="#">查看JSON</a>   <a href="#">删除</a>
			根据《航空公司行李信息说明.txt》:每件随身携带物品的体积均不得超过20x55x40厘米(国内运输)或三边之和不超过115CM(国际运输)。√	根据《航空公司行李信息说明.txt》:..... 1. 随身携带物品的重量,..... √	
				根据《航空公司行李信息说	

### 3.4.4.2 对话日志

测试过程中产生的日志和正常对话过程中产生的日志可在“对话日志”中查看；对话日志只能查看两小时之前的日志。

< 返回对话式文档问答

**行李信息说明**  
ID 51622

测试

文档训练

日志分析

交互学习日志

对话日志

技能设置

来源: 沙盒环境    注: 最新只能查看2小时前的日志    [导出日志](#)

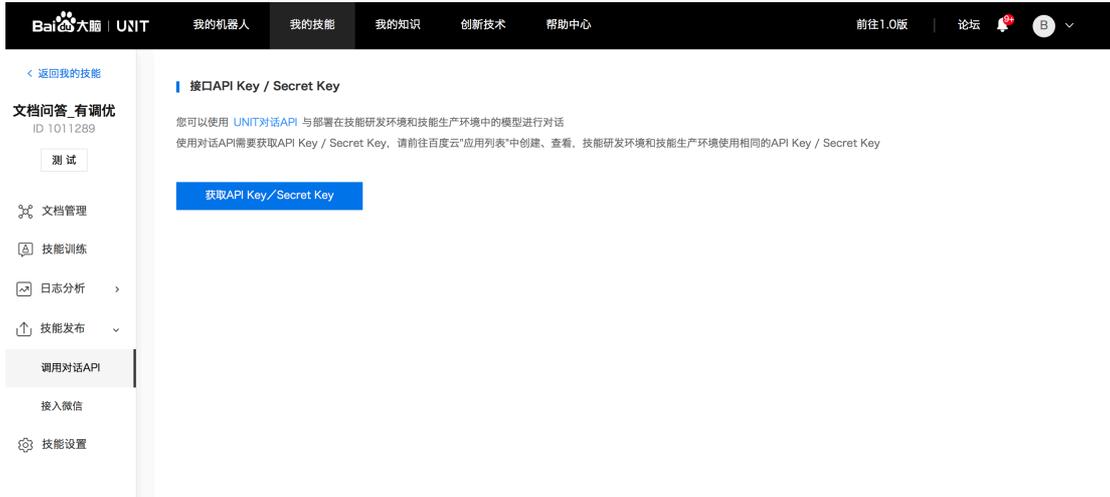
🕒 2019-05-03 19:16:20 至 2019-05-09 17:16:20
更多筛选
对话内容

对话时间	用户问	模型答	模型动作	user id	log id
05-09 16:5...	经济仓可以携带几件...	根据《航空公司行李...	satisfy	1149547170_1	cb48b0c9-cb9a-...
05-09 16:5...	商务仓可以携带几件...	根据《航空公司行李...	satisfy	1149547170_1	23e90502-047d-...
05-09 16:5...	经济舱可以携带几件...	根据《航空公司行李...	satisfy	1149547170_1	c67372d9-794e-...
05-09 16:5...	那些行李不可托运	如果引导项中没有您...	satisfy	1149547170_1	009bd69d-ca9c-...
05-09 16:5...	哪些行李不可托运	根据《航空公司行李...	satisfy	1149547170_1	0f3cdc08-8d4d-...
05-09 14:5...	经济舱可以带几件行李	根据《航空公司行李...	satisfy	1149547170_1	c4d9b845-779c-...
05-08 20:2...	哪些行李不可以托运	根据《航空公司行李...	satisfy	1149547170_1	7be73ea8-cc05-...
05-08 20:2...	刀具可以带上飞机吗?	如果引导项中没有您...	satisfy	1149547170_1	973ef861-58c9-...
05-06 17:3...	服务犬能上飞机吗	如果引导项中没有您...	satisfy	1149547170_1	99f3279f-9569-...
05-06 17:3...	活体动物 符合规定的 ...	根据《航空公司行李...	satisfy	1149547170_1	aa7b8b5b-515d-...

## 3.4.5 技能发布

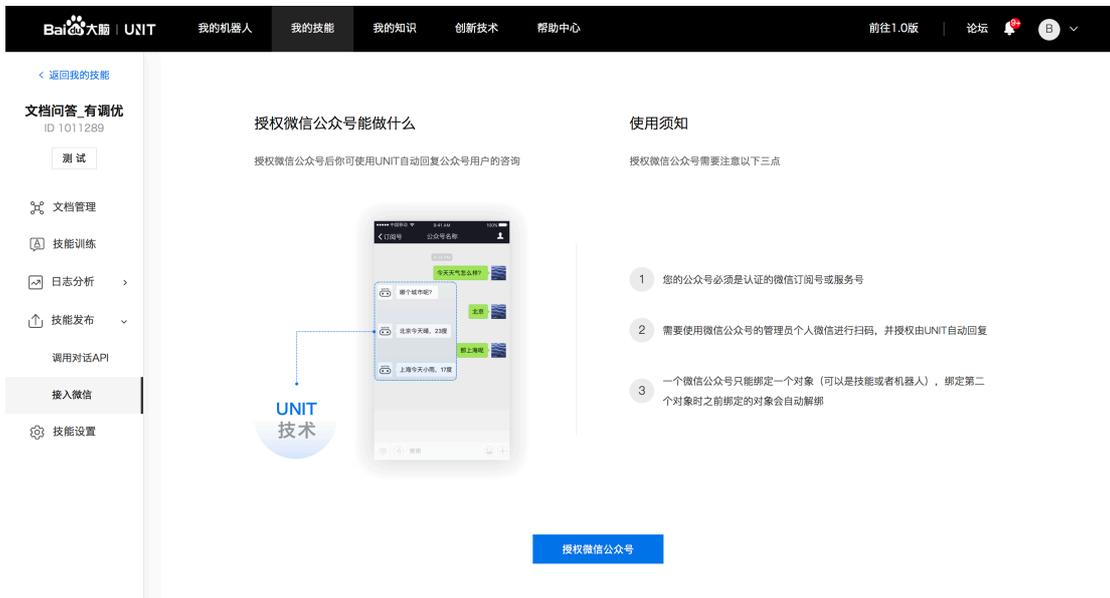
### 3.4.5.1 调用对话 API

如果您需要在自己的 app、pc、web 端调用你的技能对话服务，需要先获取 API key/Secret Key，点击下面的按钮后会打开百度云页面，进入百度云控制台。



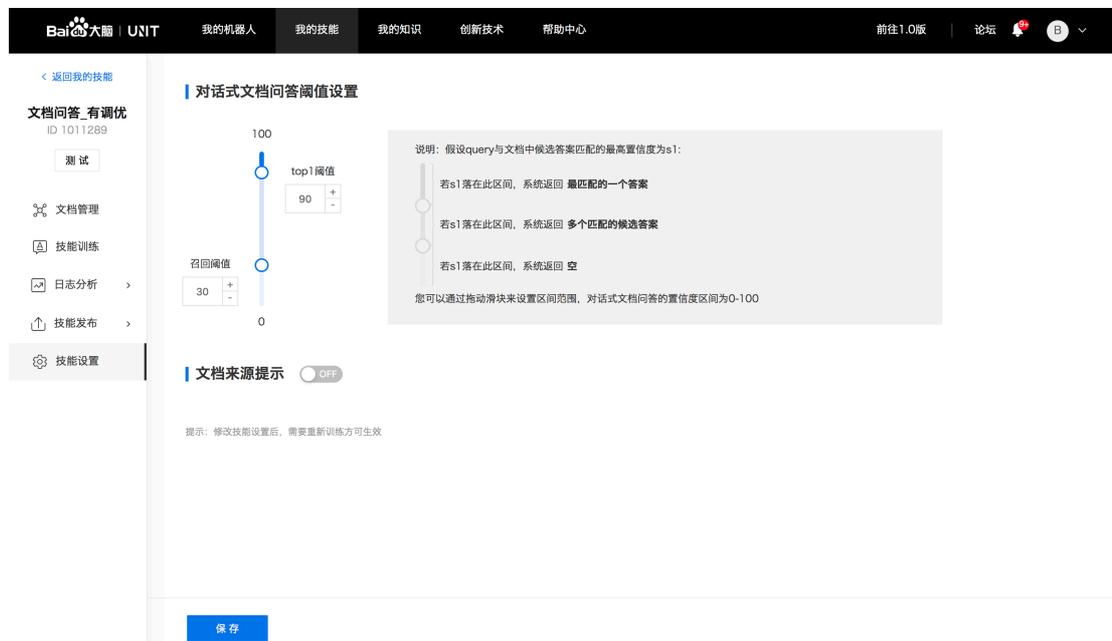
### 3.4.5.2 接入微信

如果您想基于业务文档快速搭建智能助手，可以在 UNIT 平台上快速搭建后与你的微信公众对接，让你的微信公众号智能有趣起来。



## 3.4.6 技能设置

技能设置，设置项包括文档问答的阈值以及文档来源提示等。



## 3.5 预置技能

目前 UNIT 平台预置了 20 多个技能，大家可以根据需要直接调用，不必进行重复的技能定义、训练，减少工作量：



闲聊、问候、成语问答、智能对联、智能写诗、名词解释、智能问答-----这 7 个预置技能同时预置了对话答复的资源，开发者朋友 测试后效果满意的话可直接集成到自己的产品中使用。

屏幕控制、发短信、发微信、提醒、通讯录、闹钟、机器人控制、打电话-----这 8 个预置技能只有用户意图、词槽理解的能力，可以集成在手机 app、智能音箱、机器人等智能硬件中，欢迎试用。

故事、音乐、电视剧、订咖啡、电影、天气、火车票、机票、外卖-----这 9 个预置技能 也只有用户意图、词槽的理解能力，在这些技能的对话场景中还需要开发者根据这些技能的理解的意图、词槽信息查询相应的资源来满足用户。

您可以在“我的技能”模块直接添加已有的预置技能：

## 我的技能

自定义技能

新建技能

### 旅客须知 问答

无  
技能ID: 50767

### 火车票 对话

无  
技能ID: 49630

预置技能

获取技能

### 发微信 预置

解析用户操作微信、发送微信的需求。  
技能ID: 52609

### 机票 预置

机票预定、查询及退票等指令的解析技能。  
技能ID: 52608

### 屏幕控制 预置

解析用户控制屏幕划动、列表页查询及翻页的需求。  
技能ID: 44171

其中『机票，打电话，订咖啡，音乐，故事，闹钟，通讯录，提醒，发微信，发短信，屏幕控制』11个预置技能，平台还开放了添加词槽自定义词典值的功能，来满足开发者自己场景中个性化词槽的识别问题，详细介绍见：

<https://ai.baidu.com/forum/topic/show/942709>。

## 机票

开发者: UNIT官方技能    版本号: 1.0.1    使用方式: 技能调用

技能测试

调用说明

技能需要单独调用或添加到机器人中使用, 点击[调用说明]获取具体调用方式

### 技能简介

机票预定、查询及退票等指令的解析技能。

### 意图及词槽

意图名	问法	词槽 (可自定义添加词典) <span>?</span>
DING_PIAO/订票	帮我订张机票 购买北京去香港的机票 帮我订张国航的机票	user_dep出发地点 <input type="checkbox"/> user_arr到达地点 <input type="checkbox"/> user_time出发时间 <input type="checkbox"/> user_company航空公司 <input type="checkbox"/> user_carbin机舱等级 <input type="checkbox"/> user_num航班号 <input type="checkbox"/> user_airtype飞机型号 <input type="checkbox"/>

请选择要添加的预置技能:

- 打电话 可定制  
解析用户通过语音控制电话呼入...
- 电影  
解析理解用户点播电影、搜索电...
- 天气  
满足用户查询天气、天气指数、...
- 机票 可定制  
机票预定、查询及退票等指令的...
- 问候**  
常用问候用语和日常对话中打招...
- 闲聊  
闲聊机器人，此预置技能将赋予...

更多技能正在路上，敬请期待~  
没有找到你想要的？[反馈我们](#)，我们会尽快建设

你好  
小笨蛋  
你好聪明呀

你好

你好啊~  
动作: built\_faq\_hello\_satisfy  
意图: BUILT\_FAQ\_HELLO  
词槽: faq\_question 你好

重置对话

在此开始体验预置技能，按Enter可直接发送

技能详情 获取该技能

### 3.6 技能对话 API 调用

主要针对 API 开发者，描述百度理解与交互技术技能对话接口服务的相关内容。

UNIT 平台技能对话 API: <http://ai.baidu.com/docs#/UNIT-v2-API/top>

文档目录



百度项目

视觉技术

自然语言

语言处理基础技术

语言处理应用技术

理解与交互技术UNIT

快速开始

概念介绍

实用范例

相关下载

机器人对话API文档

技能对话API文档

资源API说明

常见问题

文档中心 &gt; 自然语言 &gt; 理解与交互技术UNIT &gt; 技能对话API文档

请输入您想了解的关键词



## UNIT对话API文档

### 简介

Hi, 您好, 欢迎使用百度理解与交互技术 (UNIT) 对话API服务。

本文档主要针对API开发者, 描述百度理解与交互技术对话接口服务的相关内容。如果您对文档内容有任何疑问, 可以通过以下几种方式联系我们:

1. 在百度云控制台内提交工单, 工单类型请选择人工智能-理解与交互技术UNIT服务;
2. 进入[UNIT开发者论坛](#)发帖交流;
3. 当前文档仅用于对话API, 管理API请移步[相关下载](#)查阅《UNIT技能管理API说明文档》;

### 文档约定

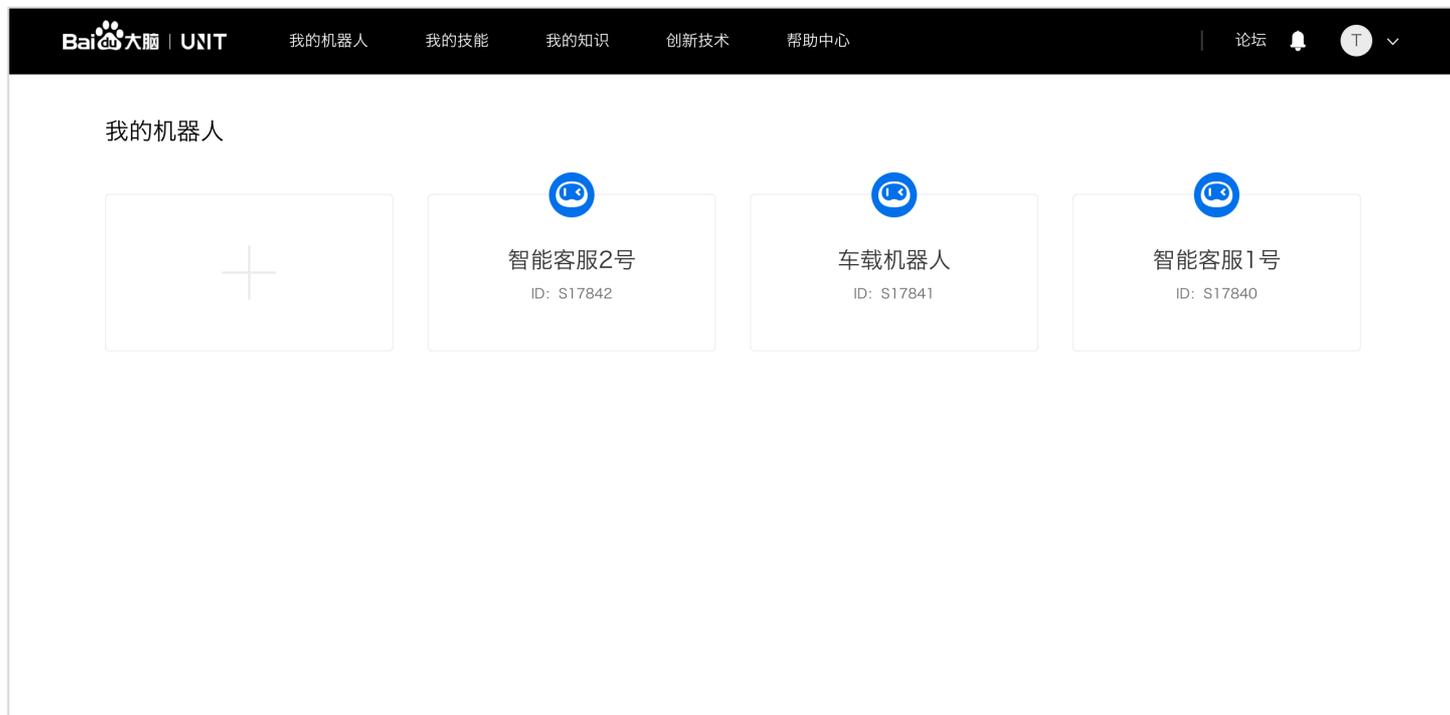
#### 参数结构约定

#### 本文目录

[UNIT对话API文档](#)[简介](#)[文档约定](#)[对话接口描述](#)[请求说明](#)[请求参数说明](#)[请求参数详细说明](#)[响应参数说明](#)[响应参数详细说明](#)[请求示例代码](#)[错误信息](#)

## 4 我的机器人

机器人是指在一个完整对话场景下的对话能力。通常按行业垂类划分（例如：小度智能音箱中的对话场景就是一个机器人）。



### 4.1 新建机器人

新建机器人需要填写机器人名称和机器人描述。一个账号可以创建 50 个机器人。

新建机器人
✕

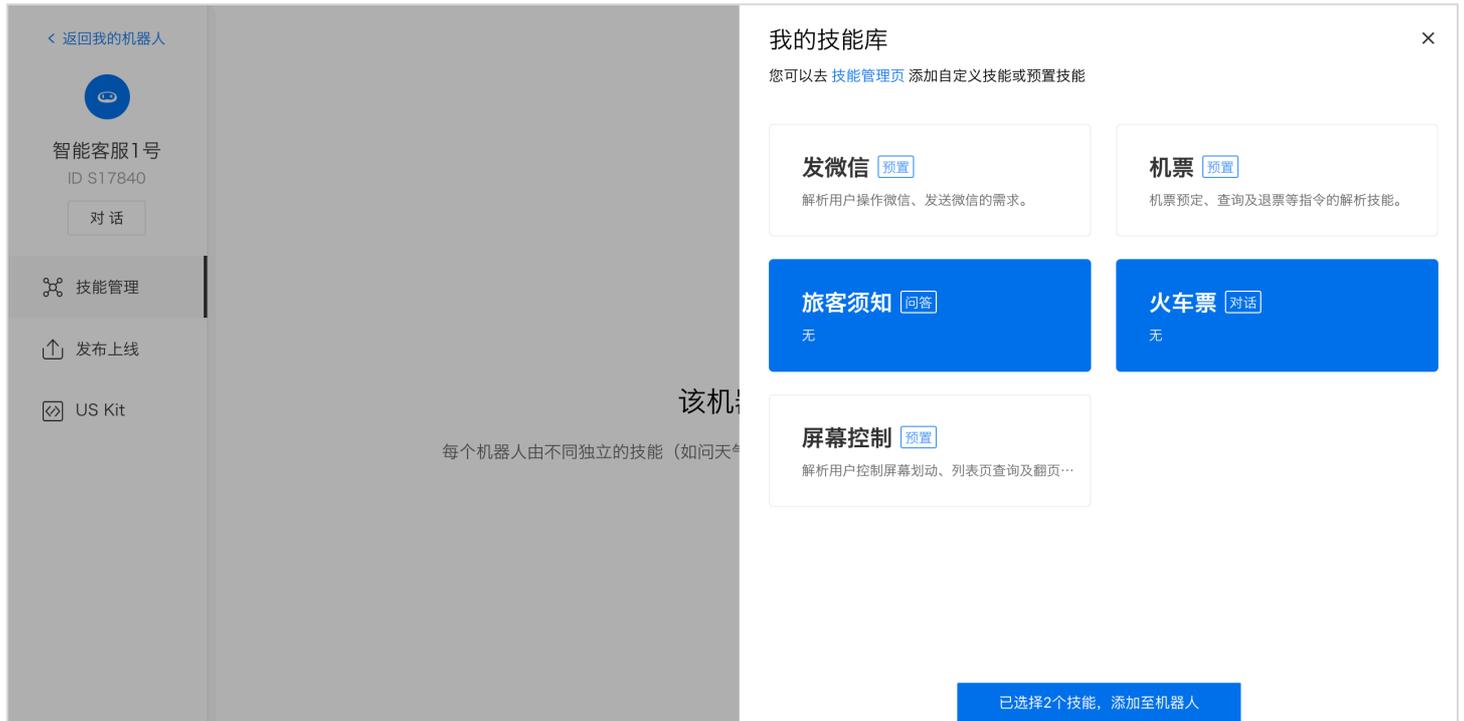
\* 机器人名称:

机器人描述:

- 机器人名称：支持中文、英文、数字和下划线，最大长度30，为必填项，可修改。
- 机器人描述：支持中文、英文、数字、常用符号，最大长度50，为非必填项，可修改。

## 4.2 添加技能

创建好机器人后，可以为其添加 5 个技能。“对话技能”、“问答技能”、“预置技能”、“对话式文档问答技能”都可以被添加到机器人中。



与此同时，还提供了技能间拖动排序的能力，告诉机器人哪个技能的优先级最高。



## 4.3 测试

您可以在平台对机器人进行测试



- 对话记录区：展示您和技能对话内容，左侧为技能“说”的内容，右侧为您输入的内容；技能解析结果中会含有答复、动作、意图、词槽信息和查看JSON按键：
  - 答复：技能答复给您或终端用户的内容；
  - 动作：技能当前答复对应的动作名称；
  - 技能：包含技能ID和技能名称；
  - 意图：技能对当前对话样本解析出的意图；
  - 词槽：技能对当前对话样本解析出的词槽；
  - 查看JSON：点击后，可以查看技能解析的详细信息，包括当前意图和词槽是用哪个策略识别的，如果是对话模板识别的，还会包含具体是哪条模板识别的（如果当前意图和词槽识别错了，可以据此找到对应的对话模板进行修改调整）；
- 重置对话状态：会清空技能记录的所有历史对话信息（例如之前技能识别的技能、意图、词槽信息）；

- 操作区下方式一个文本框，您可直接再次输入对话样本信息去验证技能的实例，按enter发送消息。在消息输入框中按键盘的向上键可以掉出历史消息快速发送。

**技巧及注意点说明**——UNIT平台上的测试对话窗口只是供你「测试体验」对话理解效果的，实际你把机器人在集成到自己的产品中，可以根据UNIT平台解析的意图词槽进行一系列的处理给出用户多种形态的反馈，比如给用户播放一首歌、一段视频、一个候选文字链list，一个图文消息等等。

## 4.4 机器人对话 API 调用

主要针对 API 开发者，描述百度理解与交互技术机器人对话接口服务的相关内容。

UNIT 平台机器人对话 API: <https://ai.baidu.com/docs#/UNIT-v2-service-API/top>

The screenshot shows the Baidu AI Open Platform documentation page for the UNIT Robot Dialogue API. The page is in Chinese and features a dark navigation bar at the top with the Baidu logo and various menu items like '产品服务', '解决方案', '合作伙伴', '开发资源', 'AI加速器', 'AI市场', '资讯', '社区', and '控制台'. A search bar is located on the right side of the navigation bar. The main content area is titled 'UNIT机器人对话API文档' and includes a '简介' (Introduction) section. The introduction text reads: 'Hi, 您好, 欢迎使用百度理解与交互技术 (UNIT) 机器人对话API服务。' followed by a paragraph stating that the document is primarily for API developers and describes the related content of the Baidu Understanding and Interaction Technology Robot Dialogue Interface Service. It also provides two ways to contact support: 1. Submitting a ticket in the Baidu Cloud console, and 2. Posting questions in the UNIT developer forum. A table of contents is visible on the right side of the page, listing sections such as '简介', '文档约定', '对话接口描述', '请求说明', '请求参数说明', '请求参数详细说明', '响应参数说明', '响应参数详细说明', and '请求示例代码'.

## 5 我的知识

### 5.1 功能概述

UNIT 对话系统的核心能力包含语言理解、对话流程管理、知识建设。“我的知识-图谱/问答知识库”是知识建设能力的重要组成部分，为开发者提供了知识挖掘与管理工具。这些积累的知识可用于提高模型理解能力、完善对话流程的控制与答复。

我的知识由图谱知识库和问答知识库共同组成。

- 您可以通过图谱知识库完成：
  - 构建自己的知识图谱，从数据源文档中挖掘实体、属性、属性值。
  - 导出图谱知识库里的各种属性值作为对话技能的词槽词典值；
  - 在对话技能的对话流程管理中，进行数据查询，利用查询结果来判断对话流程或作为答复内容。
- 您可以通过问答知识库完成：
  - 上传数据源文档，高效、低成本的挖掘问题及其答案。
  - 导出问答知识库的问答对作为问答技能的训练数据。

### 5.2 图谱知识库

#### 5.2.1 图谱知识库的作用

- 构建自己的知识图谱，从数据源文档中挖掘实体、属性、属性值。
- 导出图谱知识库里的各种属性值作为对话技能的词槽词典值；
- 在对话技能的对话流程管理中，进行数据查询，利用查询结果来判断对话流程或作为答复内容。

#### 5.2.2 图谱知识库的使用流程



## 5.2.3 定义知识类目及属性

### 5.2.3.1 为什么要做“知识定义”？

“知识定义”，也称数据描述规范（Schema），是知识图谱数据生产所必需的一套规范。“知识定义”用于描述规范化数据的结构。拥有一套完善的数据描述规范，也是计算机理解知识的必要条件。

建立“知识定义”是创建知识图谱的首要步骤。知识图谱的所有构建策略全部依赖于知识定义进行。知识定义的是否准确与全面将直接影响知识图谱的构建效果，为了保证图谱知识的丰富与全面，定义的知识类目及属性（Schema）应尽量全面的覆盖到知识文档中数据的类别和属性。

### 5.2.3.2 如何定义知识类目及属性？

定义知识类目及属性有三种方式，分别为：手动新建、导入预置类目、导入类目文件。这里我们选择手动新建的方式来构建一个有关“城市”的简单 Schema。

【源文件示例】

杭州城市介绍↵

**地理位置**：位于中国东南沿海，浙江省北部，钱塘江下游北岸，京杭大运河南端。↵

**面积**：16596 平方公里↵

**行政区划**：辖上城、下城、拱墅、江干、西湖、滨江、萧山、余杭 8 个区，富阳、建德、临安 3 个县级市，桐庐、淳安 2 个县。↵

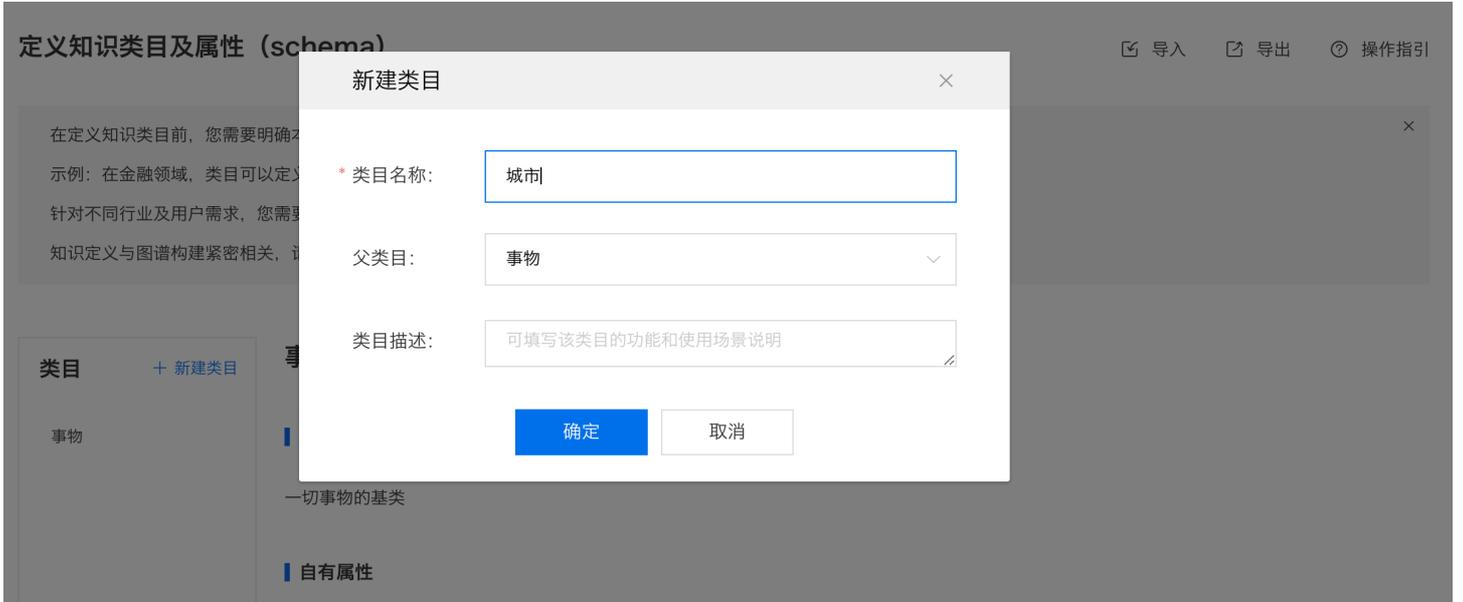
基于以上数据，我们根据需要，定义知识类目：

类目	属性	属性说明	属性别名（  隔开）	属性值类型	单多值
城市	地理位置	城市地理位置	位置	文本	单
城市	面积	城市面积	地域面积	文本	单
城市	行政区划	城市行政区划	行政区划分 区划	文本	单

您需注意：知识定义与图谱构建紧密相关，请不要频繁修改，一旦修改则需要重新进行图谱构建。

【具体操作】

- 1) 新建一个类目，类目名称输入“城市”，父类目默认为“事物”（“事物”类目为一切其他概念的基类，无法更改）。



2) 接着为这个类目添加属性：在属性名称中输入“地理位置”、数据类型选择“文本”、单多值选择单值、属性别名输入“位置”、属性描述为非必填项，可按需填写。

定义知识类目及属性 (schema) 导入 导出 操作指引

在定义知识类目前，您需要明确本次构建图谱知识库需要产出的知识的分类及对应属性。  
示例：在金融领域，类目可以定义为金融、信用卡、储蓄卡等。用户比较关注的属性有年费、卡种、发行区域等。  
针对不同行业及用户需求，您需要定义对应的知识类目及属性。  
知识定义与图谱构建紧密相关，请不要频繁修改，一旦修改则需要重新进行图谱构建。

类目 + 新建类目 城市

事物

城市 + 0

类目描述 暂无

自有属性

\* 属性名称:  属性别名:

\* 数据类型:   属性描述:

添加至属性列表

定义知识类目及属性 (schema) 导入 导出 操作指引

在定义知识类目前，您需要明确本次构建图谱知识库需要产出的知识的分类及对应属性。  
示例：在金融领域，类目可以定义为金融、信用卡、储蓄卡等。用户比较关注的属性有年费、卡种、发行区域等。  
针对不同行业及用户需求，您需要定义对应的知识类目及属性。  
知识定义与图谱构建紧密相关，请不要频繁修改，一旦修改则需要重新进行图谱构建。

类目 + 新建类目 城市

事物

城市

类目描述 暂无

自有属性

\* 属性名称:  属性别名:

\* 数据类型:   属性描述:

添加至属性列表

3) 然后，分别创建另外两个属性：面积、行政区划。

属性名称	属性别名	数据类型	单/多值	描述	操作
行政区划	行政区划分,区划	文本	单值	城市行政区划	<a href="#">编辑</a> <a href="#">删除</a>
面积	地域面积	文本	单值	城市面积	<a href="#">编辑</a> <a href="#">删除</a>
地理位置	位置	文本	单值	城市地理位置	<a href="#">编辑</a> <a href="#">删除</a>

在添加属性的过程中您需要注意：

- 属性类型除了可以填写常用的文本、日期、数值，还可选择已创建的其他类目来建立边关系，详情请查看建边流程及说明。
- 属性别名是指当前属性名称的相似表达，填写多个属性别名有利于在图谱构建时更好的抽取数据。
- 单多值是指属性值是单个答案还是多个答案，多个答案即多值。
- 一个类目所具有的属性，既包括该类目的自有属性，也包括继承自其父类的所有属性。

### 5.2.3.3 建边规则说明

建立边关系能实现更灵活的知识查询。下面举例说明：

有类目“学生”及其一个属性“就读学校”（下图表格标黄部分）；有类目“学校”及其一个属性“地址”。

当把“就读学校”的属性值类型选为另外一个类目“学校”后，可以完成类似“小明的就读学习的地址是什么”这种更灵活的知识查询，这里小明代表类目“学生”下的一个实体。

类目	属性	属性说明	属性别名	属性值类型	单多值
学校	所属地区	学校所属地区		文本	单
学校	创办时间	学校创办时间		文本	单
学校	地址	学校地址		文本	多
学校	类别	学校类别		文本	单
学生	年龄	学生年龄		文本	单
学生	性别	学生性别		文本	单
学生	就读学校	学生就读学校		文本	单

“学生”类目与“学校”类目之间要建立“就读学校”的边关系，平台操作如下：“就读学校”属性的属性类型选择“学校”即可建立边关系。

学校

学生

**自有属性**

\* 属性名:  属性别名:

称:

\* 数据类型: 学校 单值

型: 数值

事物

城市

名胜古迹

十大风景名胜区

学校

学生

添加至属性列表

属性名称	数据类型	单/多值	描述	操作
性别	文本	单值	学生性别	<a href="#">编辑</a> <a href="#">删除</a>
年龄	文本	单值	学生年龄	<a href="#">编辑</a> <a href="#">删除</a>

**继承属性**

### 5.2.3.4 导入预置类目

在这个方式下，您可以根据需要选择导入系统自动定义的 schema。

【具体操作】

我的机器人
我的技能
我的知识
创新技术
帮助中心
论坛
百

#### 定义知识类目及属性 (schema)

导入
导出
操作指引

在定义知识类目前，您需要明确本次构建图谱知识库需要产出的知识的分类及对应属性。

示例：在金融领域，类目可以定义为金融、信用卡、储蓄卡等。用户比较关注的属性有年费、卡种、发行区域等。

针对不同行业及用户需求，您需要定义对应的知识类目及属性。

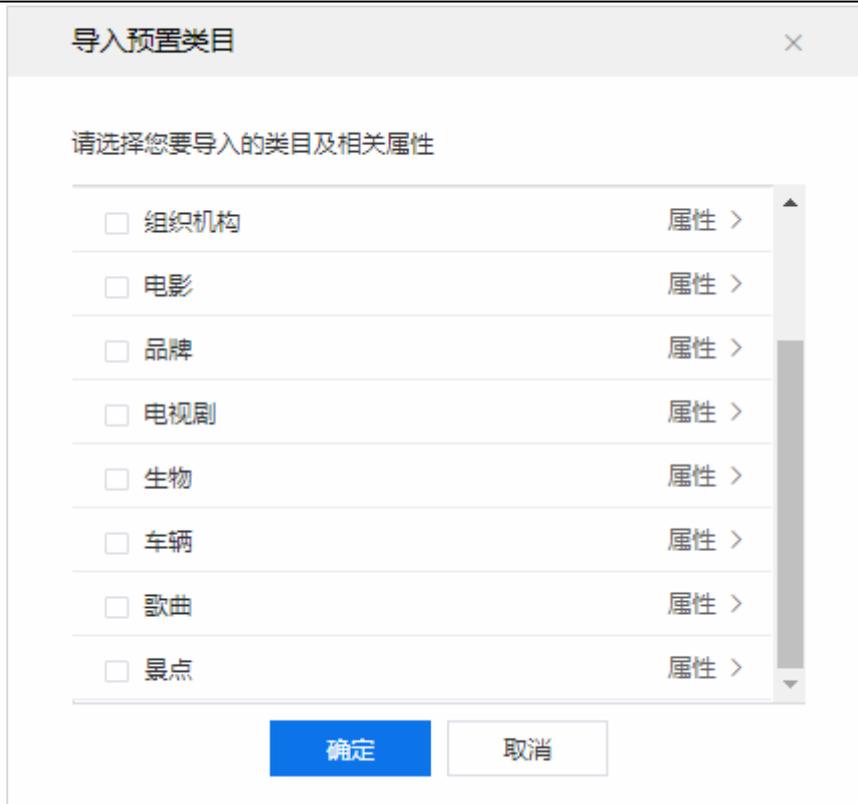
知识定义与图谱构建紧密相关，请不要频繁修改，一旦修改则需要重新进行图谱构建。

导入预置类目

导入类目文件

×

根据您需要的行业选择预置类目，导入即可。



## 5.2.4 上传原始数据

定义知识类目及属性 (Schema) 后，您需要上传构建图谱知识库的原始数据文件，按照定义好的 Schema 抽取实体、属性、属性值。建议您按照模板中推荐的格式整理原始数据文件，并进行上传。



## 5.2.5 图谱构建

完成知识类目及属性定义 (Schema)，并成功上传数据源文件后即可开始图谱构建。知识图谱的生产过程分为5步：



- 知识抽取：从上传的所有文档中抽取图谱知识数据。
- 属性映射：将文档中抽取出的属性映射到“知识定义”阶段所定义的属性上。
- 数据清洗：针对日期、数值、文本类属性值制定清洗规则，对齐数据表示结构（符合数据类型约束）。
- 消歧归一：系统会根据策略判断两个或多个实体描述是否为同一实体。如是，会将其融合为一个实体。

构建完成当您更新知识类目及属性（Schema）更新、删减原始数据文件或更新策略规则时，需二次启动图谱构建。



- 当知识类目及属性（Schema）与策略规则皆未更新，仅原始数据文件有新增时，建议选择“增量更新”。
- 当知识类目及属性（Schema）或策略规则有更新，或原始数据文件有删减，则默认全量启动。

根据数据源文件的数量与配置的策略，知识图谱生产的过程将耗时几分钟至几个小时不等。生产完成后，您可以查看详细的知识抽取结果、属性映射结果、数据清洗结果、消歧归一结果。

### 5.2.5.1 知识抽取

知识抽取为上传的所有文档中抽取图谱知识（即：spo 三元组）数据。您能够查看本次图谱生产中所有的数据源文件，列表中会显示文件名称、抽取实体数量；

< 返回我的知识

旅游城市

ID: K1998

知识定义

我的数据

**图谱构建**

我的知识

版本发布

[构建图谱](#)
历史构建记录
操作说明

### 图谱构建日志

知识抽取结果
属性映射结果
数据清洗结果
归一消歧结果

对2个文件完成知识抽取，抽取实体数2个，图谱知识6条，未抽取知识的文件数量为0

序号	文件名	抽取实体个数	操作
1	厦门城市介绍.docx	1	<a href="#">查看抽取详情</a>
2	杭州城市介绍.docx	1	<a href="#">查看抽取详情</a>

在“查看抽取详情”中您可以看到该文件中抽取出的全部实体、属性、属性值，以及该实体的属性值所在的位置信息。

图谱构建 > 抽取详情

操作说明

**抽取的图谱知识**

杭州城市

**地理位置:**

中国东南沿海

位于中国东南沿海，浙江省北部，...

**面积:**

16596平方公里

**行政区划:**

辖上城、下城、拱墅、江干、西湖...

文档内容 (杭州城市介绍.docx)

**杭州城市介绍**

**地理位置:** 位于中国东南沿海，浙江省北部，钱塘江下游北岸，京杭大运河南端。

**面积:** 16596平方公里

**行政区划:** 辖上城、下城、拱墅、江干、西湖、滨江、萧山、余杭8个区，富阳、建德、临安3个县级市，桐庐、淳安2个县。

**旅游特色:** 杭州，是我国著名的历史文化名城，七大古都之一。是长江三角洲重要中心城市和中国东南部交通枢纽。

杭州历史悠久，源远流长。杭州建县于2200多年前的秦代，时称“钱唐”，至隋开皇年间改称杭州，其名自此始。到唐代中期，杭州发展成“珍异所聚、商贾并臻”的商业都市。五代吴越和南宋又两代建都，历时200多年，是杭州发展史上的鼎盛时期，号称“东南第一州”。元时，杭州为东南重镇，经济繁荣，风景优美，被意大利旅行家马可·波罗赞叹为“世界上最美丽华贵的天城”。到明清，杭州经济社会文化发展仍处全国前茅。

杭州文化璀璨，积淀深厚。从新石器时期后期开始，先后出现过极具特色的良渚文化、吴越文化、南宋文化和明清文化，形成了一个完整的文化发展系列，为中华文明史增添了亮丽的一笔。杭州文化孕育和荟萃了众多为祖国、为家乡的政治、经济、文化和社会发展作出卓越贡献的政治家、科学家、艺术家、民族英雄、忠臣义士、贤明州官和能工巧匠，他们为杭州留下了一篇篇光耀千古的华章，构成了“文化之邦”的形象和品位。西湖的一草一木、一屋一椽，无不带着历史的印记。

**最佳旅游时间:** 四季皆宜。

## 5.2.5.2 属性映射

属性映射是将文档中抽取出的属性映射到“知识定义”阶段所定义的属性上。属性映射结果中会按类目显示每个类目中各属性的映射情况。当状态为“未映射”或“已映射”但您认为映射结果不正确时，您可以重新配置属性映射规则来持续优化属性映射策略。更新属性映射规则后，需重新进行图谱构建，方可生效。

类目: 城市

注意: 所有策略规则更新, 需重新进行“图谱构建”, 方可生效

序号	文档中抽取的属性	实体覆盖率	属性值	映射属性
1	旅游特色	100%	杭州, 是我国著名的历史文化名城, 七大古都之一。...	请选择映射属性
2	最佳旅游时间	66.7%	四季皆宜	请选择映射属性
3	地理位置	100%	位于中国东南沿海, 浙江省北部, 钱塘江下游北岸, ...	不映射
4	交通指南	100%	2000年12月29日全新投入使用的杭州萧山国际机场...	交通指南
5	面积	100%	16596平方公里	行政区划
6	美食小吃	100%	杭州菜以它的质优价廉风靡江南。杭州菜属中国八大...	面积
7	行政区划	66.7%	辖上城、下城、拱墅、江干、西湖、滨江、萧山、...	地理位置
8	区划	33.3%	下辖翔安区、集美区、海沧	最佳旅游时间

### 5.2.5.3 数据清洗

数据清洗是针对日期、数值、文本类属性值制定清洗规则, 对齐数据表示结构。数据清洗结果中会显示具体哪个类目下的哪个属性的数据类型进行了清洗, 并显示出清洗后的数据格式。

类目: 化妆品

注意: 所有策略规则更新, 需重新进行“图谱构建”, 方可生效

序号	属性	属性值	属性值类型	清洗规则
1	官网	http://www.laneige.com.cn	文本	<input type="text" value="[a-zA-z]+://(\w+(-\w+)*)"/> ⓘ
2	评价时间	3/2/2019	日期	<input type="text" value="yyyy年mm月dd日"/> ▾
3	用户评价	是我自己比较满意的, 这个色号很...	文本	<input type="text" value="请填写正则表达式"/> ⓘ
4	电子邮箱	邮箱: lanzhi_demo@lanzhi.com	文本	<input type="text" value="[A-Za-z\d]+([-_][A-Za-z"/> ⓘ
5	联系方式	400-820-8016	文本	<input type="text" value="\d{3}-\d{4}-?\d{8} \c"/> ⓘ
6	适合人群	女性**哈哈	文本	<input type="text" value="[男 女]性"/> ⓘ
7	价格	259.01	数值	<input type="text" value="保留小数点后3位"/> ▾

您需注意清洗规则为3类:

- 日期: 下拉框选择清洗类型 (yyyy-mm-dd、yyyy年 mm 月 dd 日、mm/dd/yyyy)
- 数值: 下拉框选择清洗类型 (整数、保留小数点后1位、保留小数点后2位、保留小数点后3位、保留小数点后4位、保留小数点后5位)
- 文本: 您需要输入正则表达式

【正则表达式示例】用户问法: XS最高配最新售价10888元, 则需要填写的正则表达式为: \d+元

更新清洗规则后, 需重新进行图谱构建, 方可生效。

### 5.2.5.4 消歧归一

消歧归一是系统会根据策略判断两个或多个实体描述是否为同一实体。如是, 会将其融合为一个实体。您能够:

- 查看系统对相同实体的归一融合情况: 对于从文档中抽取出的实体, 系统将根据属性、属性值对其进行归一融合。
- 手动配置归一规则来优化知识数据。

更新归一规则后, 需重新进行图谱构建, 方可生效。

配置规则
ⓧ

当满足以下条件时，会将实体视为 相同 实体

[+ 添加规则](#)

属性	值类型	规则	
<span style="border: 1px solid #ccc; padding: 2px;">血型</span> <span style="float: right;">ⓧ</span>	<span style="border: 1px solid #ccc; padding: 2px;">文本</span>	<span style="border: 1px solid #ccc; padding: 2px;">完全相同</span>	<span style="font-size: 1em;">🗑</span>
<span style="border: 1px solid #ccc; padding: 2px;">身高</span> <span style="float: right;">ⓧ</span>	<span style="border: 1px solid #ccc; padding: 2px;">文本</span>	<span style="border: 1px solid #ccc; padding: 2px;">完全相同</span>	<span style="font-size: 1em;">🗑</span>
<span style="border: 1px solid #ccc; padding: 2px;">出生日期</span> <span style="float: right;">ⓧ</span>	<span style="border: 1px solid #ccc; padding: 2px;">文本</span>	<span style="border: 1px solid #ccc; padding: 2px;">完全相同</span>	<span style="font-size: 1em;">🗑</span>

确定
取消

最后您可以查看历史构建记录。

Baidu 大脑 | UNIT
我的机器人
我的技能
我的知识
创新技术
帮助中心
论坛 百

返回我的知识
图谱构建 > 历史构建记录
ⓧ 操作说明

旅游城市

ID: K1998

知识定义

我的数据

图谱构建

我的知识

版本发布

时间	图谱知识总数	生产耗时(分钟)	其他指标
05-08 12:53	6	1	已清洗数据
05-08 13:03	9	1	已映射属性
05-08 13:07	6	1	归一实体

- 导入文件
- 图谱知识总数
- 生产耗时(分钟)
- 抽取实体数
- 已映射属性
- 已清洗数据
- 归一实体

## 5.2.6 手动调优

您可以在“我的知识”中查看所有抽取出的知识，包含实体、属性、属性值，也可以手动编辑优化，同时也支持新增实体，手动调优的数据实时生效，无需重新进行图谱构建。

The screenshot shows the '我的知识' (My Knowledge) section for the entity '杭州城市' (Hangzhou City). The right-hand panel displays the following entity data:

- 实体名称: 杭州城市
- 所属类目: 城市
- 地理位置: 中国东南沿海, 位于中国东南沿海, 浙江省北部, 钱塘江下游北岸, 京杭大运河南端
- 行政区划: 辖上城、下城、拱墅、江干、西湖、滨江、萧山、余杭8个区, 富阳、建德、临安3个县级市, 桐庐、淳安2个县
- 面积: 16596平方公里

A red box highlights the '编辑数据' (Edit Data) button in the top right corner of the entity data panel.

## 5.2.7 发布图谱

您可以对生产出的知识进行发布，不限制发布次数，但只保留 5 个版本，发布后的数据仅供查看，无法更改。

The screenshot shows the '版本发布' (Version Release) section for the entity '杭州城市'. The table below lists the versions:

版本	版本描述	发布时间	图谱知识数	操作
v3	--	2019-05-08 13:12:36	5	● 当前版本 <a href="#">查看数据并导出</a>
v2	--	2019-05-08 13:12:19	5	<a href="#">应用该版本</a> <a href="#">查看数据并导出</a> <a href="#">删除</a>
v1	--	2019-05-08 13:11:30	5	<a href="#">应用该版本</a> <a href="#">查看数据并导出</a> <a href="#">删除</a>

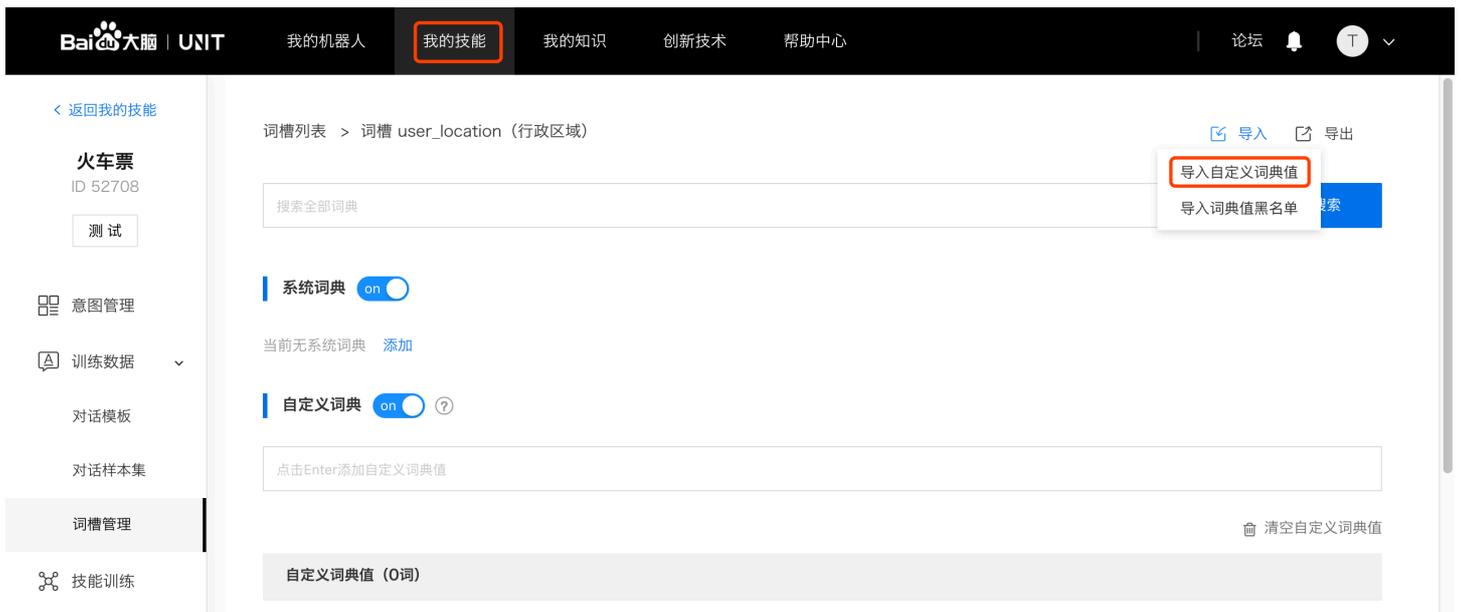
A red box highlights the '查看数据并导出' (View Data and Export) button for the current version (v3).

您可以根据需要导出知识数据，可选择导出“属性值”或“属性别名”。



接下来，应用图谱知识库来提升对话技能的理解和交互能力，目前包含以下两种方式：

- 导出图谱知识库里的各种属性值/属性别名作为对话技能的词槽词典值；



- 您可以应用任一版本，在对话技能的 taskflow（对话流程云端管理工具）中可以查询对应版本的图谱知识库数据，利用查询结果来判断对话流程或作为答复内容。
- taskflow 可以在 对话技能-技能设置-高级设置-对话回应设置 中启用并了解详细使用流程

## 5.3 问答知识库

### 5.3.1 问答知识库的作用

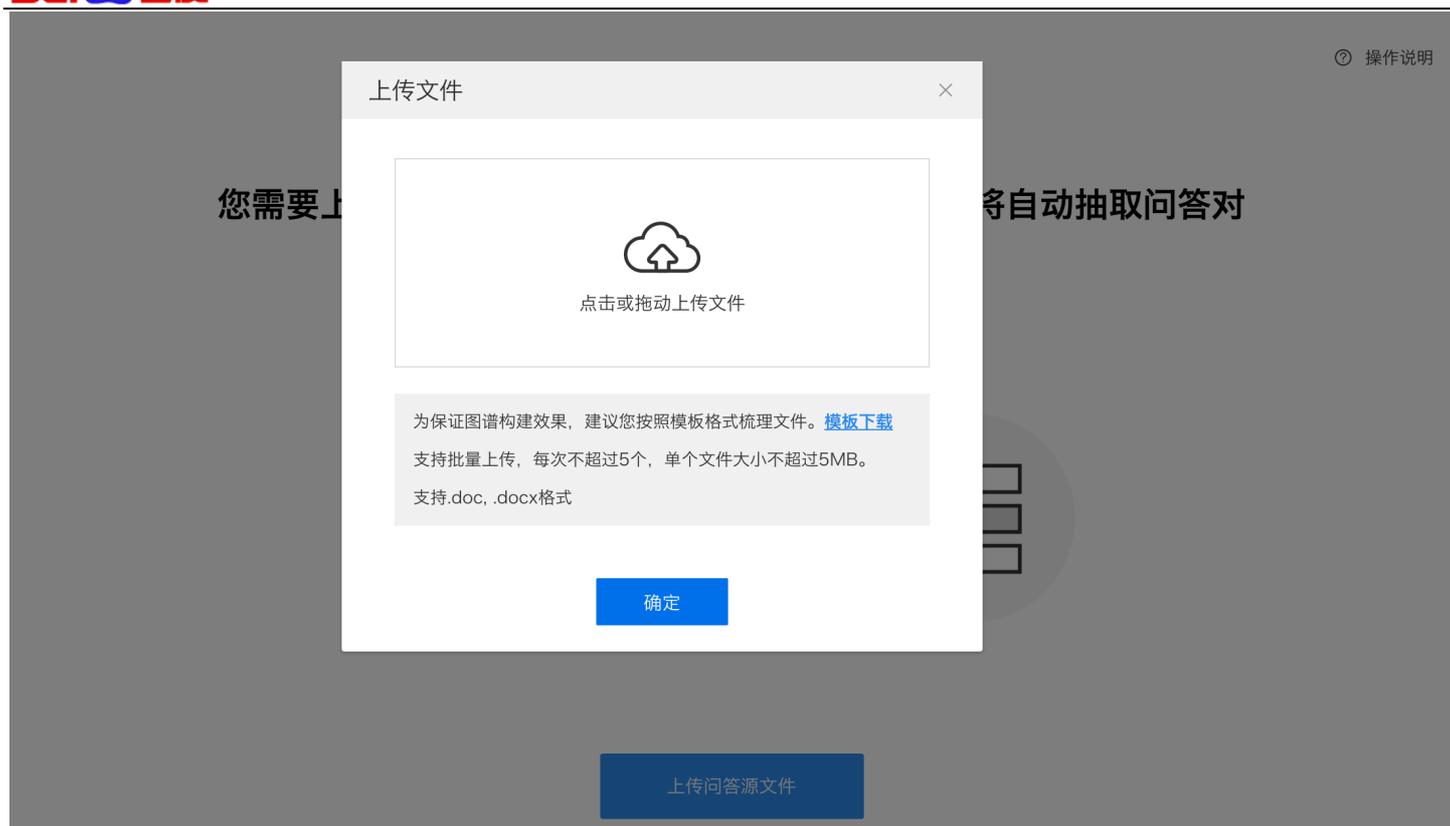
- 上传数据源文档，高效、低成本的挖掘问题及其答案。
- 导出问答知识库的问答对作为问答技能的训练数据。

### 5.3.2 问答知识库的使用流程



### 5.3.3 上传文档

建议您按照模板中推荐的格式整理原始数据文件，并进行上传。



### 5.3.4 手动调优

此项功能您可以添加问答对，也可对已有问答对进行编辑。

编辑问答对
✕

**问题**

\* 问题

相似问题

**答案**

\* 答案

确定
取消

### 5.3.5 导出数据

导出问答对后的文件格式，可以直接在“我的技能-问答技能-问答管理”中，导入为问答对。

我的机器人
我的技能
我的知识
创新技术
帮助中心
论坛 百

导出问答对

搜索问题内容

ⓘ 操作说明

Baidu 大脑 | UNIT
我的机器人
我的技能
我的知识
创新技术
帮助中心
论坛 T

返回我的技能

**旅客须知**  
ID 50767

测试

问答管理

技能训练

技能设置

全部问答 / 3条

导入

导出

添加问答对

<input type="checkbox"/> 问题	问答对ID	标签
<input type="checkbox"/> 退票要加收多少费用	146285	退票
<input type="checkbox"/> 军人可以享受多大的优惠	3551	优惠 军人
<input type="checkbox"/> 多高的孩子乘坐火车可以半价票	2780	订票 优惠 儿童

15
<
1
>
共1页 3条

## 6 创新技术

创新技术目前包括了语义解析离线使用和对话式文档问答，未来会持续开放更多创新技术。

### 6.1 语义解析 离线使用

语义解析离线使用的定位是服务于终端设备，对于语义解析能力的 Edge Computing 可实现在设备无网络时进行语义解析能力。可应用于智能车载、智能家居、智能穿戴等场景。

#### 6.1.1 业务判断

首先要判断自己的业务是否需要断网使用语义解析能力。

#### 6.1.2 获取文件

目前平台已在智能车载、智能家居等多个领域中实现了离线解析的能力。您无需自行配置，我们将为您提供现成的场景配置文件。目前支持智能车载场景下的车控指令文件下载，如需其他场景，可直接联系我们 ([unit-innovation@baidu.com](mailto:unit-innovation@baidu.com)) 。

# 02

## 获取场景文件

- **获取离线场景文件，直接使用或按标准修改**

目前平台已在智能车载、智能家居等多个领域中形成了离线解析的能力。您无需自行配置，我们将为您提供现成的场景配置文件。目前支持智能车载场景下的车控指令文件下载，如需其他场景，可直接联系我们 ([unit-innovation@baidu.com](mailto:unit-innovation@baidu.com)) 。



智能车载



智能家居



智能穿戴

## 6.1.3 优化文件

已支持的离线场景效果均有保障，您可以使用训练工具对场景文件进行训练。如需优化离线场景的效果，您可以根据《数据优化配置说明》的修改标准进行优化。最后将训练后得到的模型文件按照 Step4 集成入 SDK 即可。

### 03

修改并训练模型

- **使用训练工具将场景训练完成**

已支持的离线场景效果均有保障，您可以使用训练工具对场景文件进行训练。如需优化离线场景的效果，您可以根据《数据优化配置说明》的修改标准进行优化。最后将训练后得到的模型文件按照Step4集成入SDK即可。

[下载训练工具](#) [数据优化配置说明](#)

## 6.1.4 集成并调用 SDK

目前我们仅支持 Android 4.1 及以上，请下载 Android SDK，并按照《集成及调用说明》，将 Step3 中训练好的模型集成入 SDK 即可使用。

### 04

集成SDK并调用

- **将训练好的模型集成入SDK后，即可直接调用**

目前我们仅支持Android 4.1及以上，请下载Android SDK，并按照《集成及调用说明》，将Step3中训练好的模型集成入SDK即可使用。

[下载安卓SDK](#) [集成及调用说明](#)

## 6.2 对话式文档问答技能

### 6.2.1 创建技能

- 1) 登录后点击页面左侧导航中“创新技术”，点击“对话式文档问答”，进入对话式文档问答的主页面，新用户点击“新建技能”，填写技能信息。

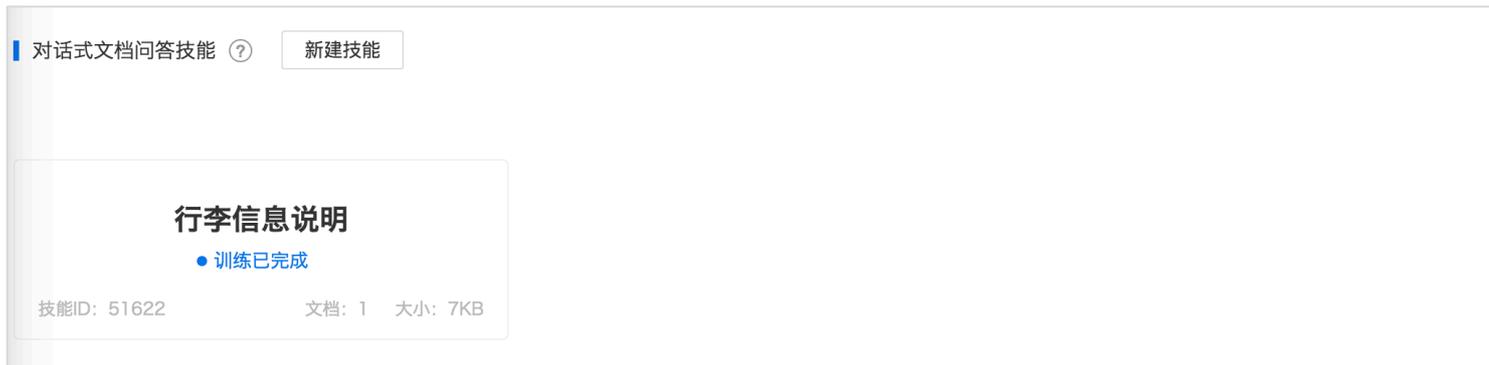
### 新建技能 ×

\* 技能名称:

技能描述:

2) 填写技能名称与技能描述。技能名称支持中文、英文、数字与”\_”，最大长度30，必填项，不可重复；技能描述支持中文、英文、数字与常用符号，最大长度50，非必填项。填写完成后，点击确定。

3) 技能创建完成后会出现在“对话式文档问答技能”列表和“我的技能”列表中。



4) 技能卡片展示技能名称、技能状态、技能文档信息、技能ID。

- 技能名称：您创建技能时填写的技能名称信息。
- 技能状态：当前该技能的训练状态。
- 技能文档信息：包括文档的数量和总大小。
- 技能ID：系统为该技能自动生成的ID。

## 6.2.2 删除技能

点击技能卡片的“操作”按钮，显示操作列表。点击“删除技能”后会要求输入技能名称二次确认：



### 6.2.3 重命名技能

点击技能卡片的“操作”按钮，显示操作列表。点击“重命名技能”后可对技能名称和技能描述进行更改



### 6.2.4 文档上传

目前该技术仅支持纯文本文档的训练，文档内容支持两种形式：普通非结构话文档和 FAQ 文档。

5) 在技能列表里点击新建技能的卡片，可进入技能训练的主页面：



6) 点击“上传文档”，会弹出以下界面，请选择需要上传的文档类型并按照相关说明上传符合格式的文档；您也可以直接下载示例文档进行体验。





完成文档类型选择，并点击“上传文档”后，进入以下界面，如果您需要调整要上传的文档，可以点击文档名右侧的“删除”icon；



7) 点击“确定”后，完成上传文档；

8) 文档上传完成后，进入以下界面，可点击“设置”中的“删除”删除当前文档，删除时会弹出二次确认的窗口；点击“下载”下载当前文档；

< 返回对话式文档问答

行李信息说明  
ID 51622

测试

文档训练

日志分析 >

技能设置

重新训练 ● 训练已完成 ?

上传文档  
格式: txt 大小<10M 编码: utf-8

航空公司行李信息说明.txt  
随身携带物品一般规定: 1. 随身携带物品  
每位旅客以10  
大小: 7KB  
下载  
删除

是否删除该文档? X

文档不可复原, 请谨慎操作

确定 取消

## 6.2.5 技能训练

3) 文档上传完成后进入以下界面, 点击“训练”按钮, 开始模型训练, 训练过程中不可对文档进行任何编辑;

< 返回对话式文档问答

行李信息说明  
ID 51622

测试

文档训练

日志分析 >

技能设置

训练 ● 未训练 ?

上传文档  
格式: txt 大小<10M 编码: utf-8

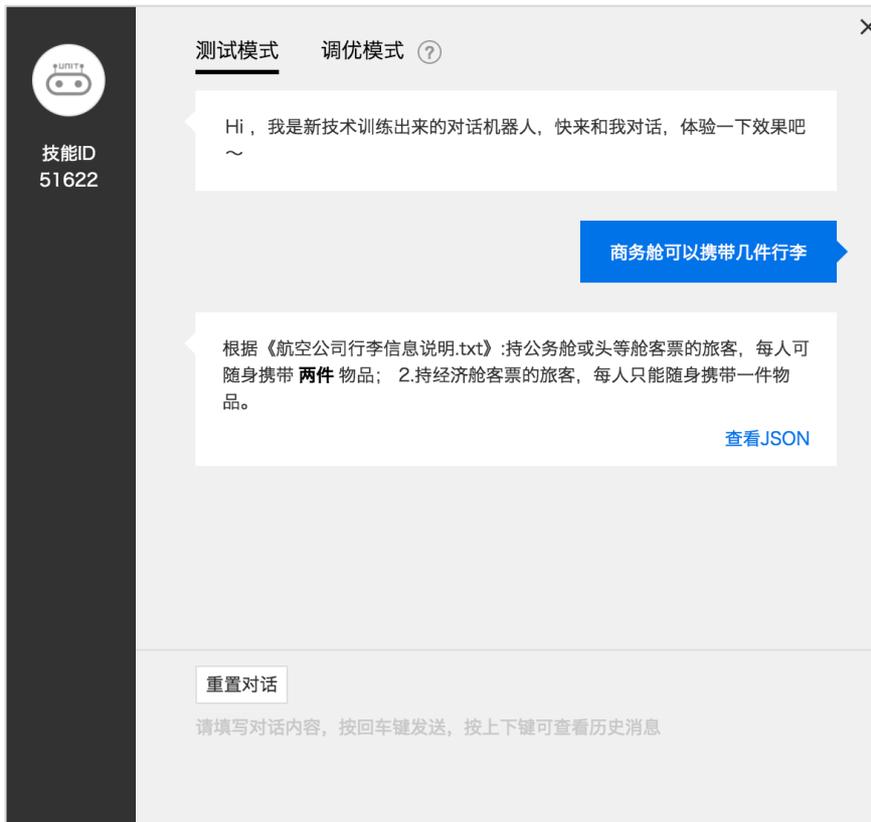
航空公司行李信息说明.txt  
随身携带物品一般规定: 1. 随身携带物品的重量,  
每位旅客以10  
大小: 7KB

4) 训练完成后，界面如下，“训练”按钮变成“重新训练”，点击“测试”进入测试界面。



## 6.2.6 测试

11) 点击“测试”进入以下界面，可以体验对话；对话有两种形态：一种是直接对话，如下图，



12) 一种是回复中提供选项，用户可以进行选择，或者直接输入更多信息



13) 如果对对话体验不满意，可以点击“调优模式”进入调优；



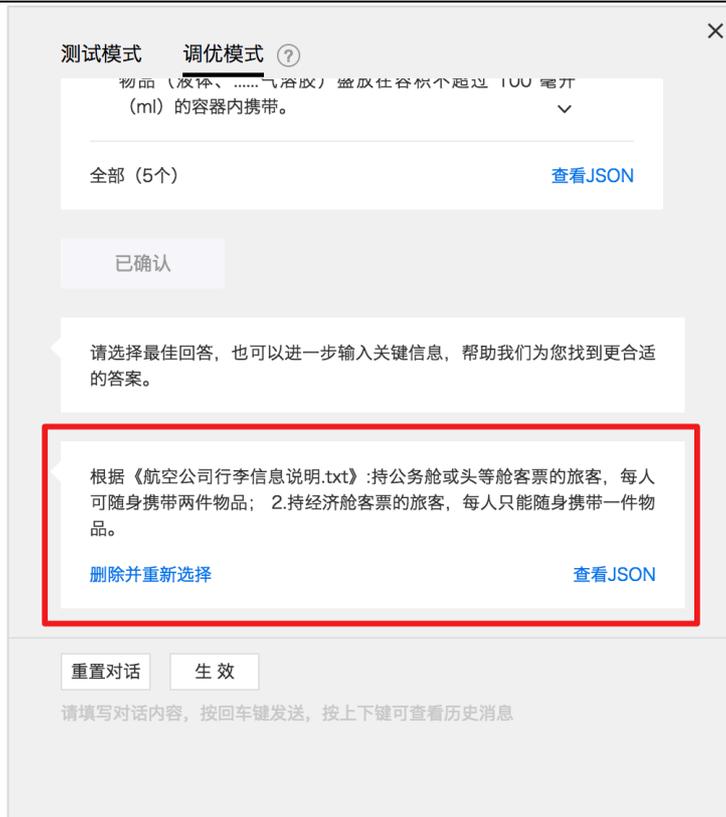
14) 调优模式下，可以展开每个候选项查看完整的选项内容；



15) 查看并确定正确选项后, 点击单选按钮进行选择, 并点击确认, 提交并保存答案



16) 点击“确定”后, 会呈现针对这个问题完整的答案, 如下图:



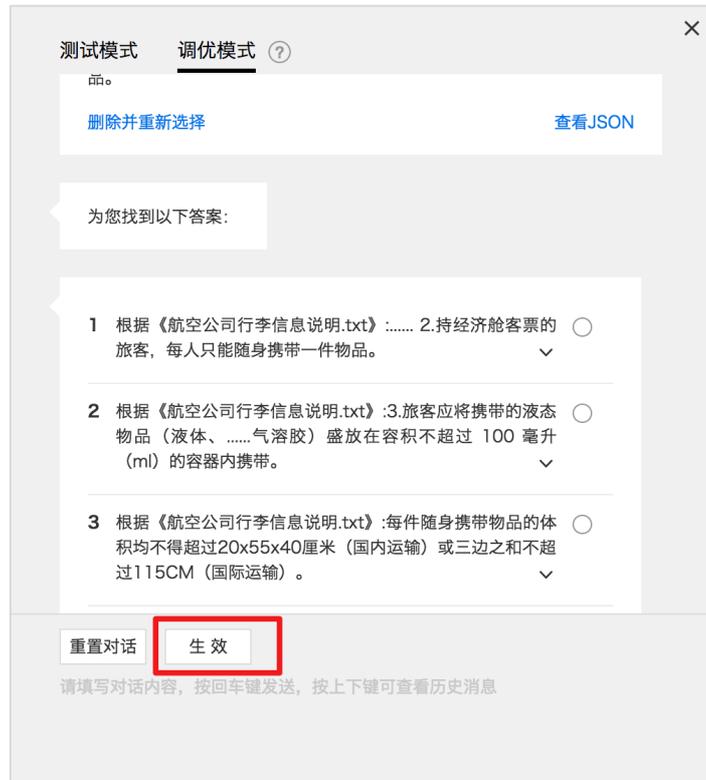
17) 如果对之前选择的答案不满意, 可以点击“删除并重新选择”按钮, 进行重新选择



18) 点击“删除并重新选择”按钮后, 会呈现选择前的选项状态;



19) 调优完成后, 点击“生效”按钮, 调优结果会在模型中生效; 之后即可在“测试模式”下体验调优后的对话效果了;



20) 如果调优后没有点击“生效”就要离开调优界面, 会弹出二次确认, 点击“生效”后, 调优结果就生效到模型中了, 也可以点击;



## 6.2.7 日志分析

### 6.2.7.1 交互学习日志

调优过程中产生的对话日志可在“交互学习日志”中实时查看；

< 返回对话式文档问答

**行李信息说明**  
ID 51622

测试

文档训练

日志分析

交互学习日志

对话日志

技能设置

时间范围: 2019-05-03 19:11:22 至 2019-05-09 19:11:22

已选择0条记录 批量删除 搜索被反馈对话

<input type="checkbox"/>	反馈时间	被反馈对话	反馈前答案	反馈后答案	操作
<input type="checkbox"/>	05-09 19:00:06	重新选择	根据《航空公司行李信息说明.txt》:持公务舱或头等舱客票的旅客,每人可随身携带两件物品;2.持经济舱客票的旅客,每人只能随身携带一件物品。	根据《航空公司行李信息说明.txt》:..... 2.持经济舱客票的旅客,每人只能随身携带一件物品。 根据《航空公司行李信息说明.txt》:3.旅客应将携带的液态物品(液体、.....气溶胶)盛放在容积不超过100毫升(ml)的容器内携带。 根据《航空公司行李信息说明.txt》:每件随身携带物品的体积均不得超过20x55x40厘米(国内运输)或三边之和不超过115CM(国际运输)。 根据《航空公司行李信息说明.txt》:..... 1. 随身携带物品的重量,.....	<a href="#">查看JSON</a>   <a href="#">删除</a>

### 6.2.7.2 对话日志

测试过程中产生的日志和正常对话过程中产生的日志可在“对话日志”中查看；对话日志只能查看两小时之前的日志。

### 行李信息说明

ID 51622

测试

文档训练

日志分析

交互学习日志

对话日志

技能设置

来源: 沙盒环境 注: 最新只能查看2小时前的日志

导出日志明

2019-05-03 19:16:20 至 2019-05-09 17:16:20

更多筛选

对话内容

搜索对话内容

对话时间	用户问	模型答	模型动作	user id	log id
05-09 16:5...	经济仓可以携带几件...	根据《航空公司行李...	satisfy	1149547170_1	cb48b0c9-cb9a-...
05-09 16:5...	商务仓可以携带几件...	根据《航空公司行李...	satisfy	1149547170_1	23e90502-047d-...
05-09 16:5...	经济舱可以携带几件...	根据《航空公司行李...	satisfy	1149547170_1	c67372d9-794e-...
05-09 16:5...	那些行李不可托运	如果引导项中没有您...	satisfy	1149547170_1	009bd69d-ca9c-...
05-09 16:5...	哪些行李不可托运	根据《航空公司行李...	satisfy	1149547170_1	0f3cdc08-8d4d-...
05-09 14:5...	经济舱可以带几件行李	根据《航空公司行李...	satisfy	1149547170_1	c4d9b845-779c-...
05-08 20:2...	哪些行李不可以托运	根据《航空公司行李...	satisfy	1149547170_1	7be73ea8-cc05-...
05-08 20:2...	刀具可以带上飞机吗?	如果引导项中没有您...	satisfy	1149547170_1	973ef861-58c9-...
05-06 17:3...	服务犬能上飞机吗	如果引导项中没有您...	satisfy	1149547170_1	99f3279f-9569-...
05-06 17:3...	活体动物 符合规定的 ...	根据《航空公司行李...	satisfy	1149547170_1	aa7b8b5b-515d-...

## 6.2.8 技能 API 调用

主要针对 API 开发者, 描述百度理解与交互技术技能对话接口服务的相关内容。

UNIT 平台技能对话 API: <http://ai.baidu.com/docs#/UNIT-v2-API/top>

文档目录



自然语言

视觉技术

自然语言

语言处理基础技术

语言处理应用技术

理解与交互技术UNIT

快速开始

概念介绍

实用范例

相关下载

机器人对话API文档

技能对话API文档

资源API说明

常见问题

文档中心 &gt; 自然语言 &gt; 理解与交互技术UNIT &gt; 技能对话API文档

请输入您想了解的关键词



## UNIT对话API文档

### 简介

Hi, 您好, 欢迎使用百度理解与交互技术 (UNIT) 对话API服务。

本文档主要针对API开发者, 描述百度理解与交互技术对话接口服务的相关内容。如果您对文档内容有任何疑问, 可以通过以下几种方式联系我们:

1. 在百度云控制台内提交工单, 工单类型请选择人工智能-理解与交互技术UNIT服务;
2. 进入[UNIT开发者论坛](#)发帖交流;
3. 当前文档仅用于对话API, 管理API请移步[相关下载](#)查阅《UNIT技能管理API说明文档》;

### 文档约定

#### 参数结构约定

#### 本文目录

[UNIT对话API文档](#)[简介](#)[文档约定](#)[对话接口描述](#)[请求说明](#)[请求参数说明](#)[请求参数详细说明](#)[响应参数说明](#)[响应参数详细说明](#)[请求示例代码](#)[错误信息](#)