

阿里云智能对话机器人

产品白皮书

产品版本：4.4.0

文档版本：2021.04

法律声明

阿里云提醒您阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。

2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。

3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。

4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。

5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。

6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

目录

法律声明.....	1
目录.....	2
一、 产品概述	5
1.1 产品介绍	5
1.2 产品优势	7
1.3 客户价值	8
二、 应用场景	8
2.1 在线人机交互.....	8
2.2 语音人机交互.....	11
三、 功能简介	13
3.1 系统管理与账户配置.....	13
3.1.1 租户及用户的管理功能	13
3.1.2 系统关键信息概览功能	13
3.1.3 高级模型配置功能	13
3.1.4 标签管理	13
3.1.5 全局名词	14
3.1.6 自定义登录页配置功能	14
3.2 系统集成服务.....	14
3.2.1 会话接口 API 调用功能	14

3.2.2 Chatui 一键部署功能	14
3.3 智能问答.....	14
3.3.1 问答机器人管理	14
3.3.2 知识挂载功能	14
3.3.3 闲聊功能	15
3.3.4 自定义问答策略配置	15
3.3.5 知识引用与上传下发	15
3.3.6 自定义多视角管理	15
3.3.7 多轮问答能力-对话工厂.....	15
3.3.8 单轮问答能力-FAQ 问答及管理.....	15
3.3.9 文档问答能力-MRC	16
3.3.10 图谱问答能力-KBQA 知识图谱问答	16
3.3.11 表格问答能力-Table QA.....	16
3.3.12 英文问答机器人	16
3.3.13 卡片工坊	17
3.4 高级运营工具模块.....	17
3.4.1 问答训练	17
3.4.2 模型训练平台	17
3.4.3 问法聚类工具	17
3.4.4 全局问答配置	17
3.4.5 问答能力测评	17
3.4.6 FAQ 挖掘	18

3.4.7 智能荐句	18
3.4.8 测试窗功能	18
3.5 高级运营工具模块.....	18
3.5.1 API 接口管控支持	18
3.5.2 健康度检查	18
3.5.3 审核功能/审核中心.....	18
3.5.4 实时会话分析	19
3.5.5 数据看板功能	19
3.5.6 权限管理功能	19
3.5.7 业务空间管控功能	19
3.5.8 系统日志功能	19
3.5.9 错误码查询文档	20
3.5.10 登录登出功能	20
四、 技术特性	20
4.1 技术架构.....	20
4.2 技术优势	21
4.2.1 完备的技术体系	21
4.2.2 双十一大规模运用和技术沉淀	22
4.2.3 基于本体结构的 FAQ 库.....	23
4.2.5 多轮对话框架	25
4.2.6 问答训练.....	26

五、 应用部署	28
5.1 部署架构	28
5.2 应用部署图	29
5.3 架构特点	29
5.4 基于飞天敏捷部署，自带运维与监控能力	30
附录一 名词解释	31
附录二 阿里云介绍	32
概要	32
整体能力	33
市场地位	36
平台特点	37
开放性	39

一、 产品概述

1.1 产品介绍

阿里云智能对话机器人包括公有云和私有化部署两个版本。

阿里云智能对话机器人（原云小蜜）是一款基于自然语言处理

（NLP）和人工智能（AI）技术提供智能会话能力的云服务，开发者无需亲自掌握 NLP、AI 等技术，只要基于阿里云智能对话机器人的会话 API 即可在多种终端（网站、移动 APP、智能硬件）中构建自己的智能会话机器人界面，如网页、APP、智能硬件及实体机器人。 阿里

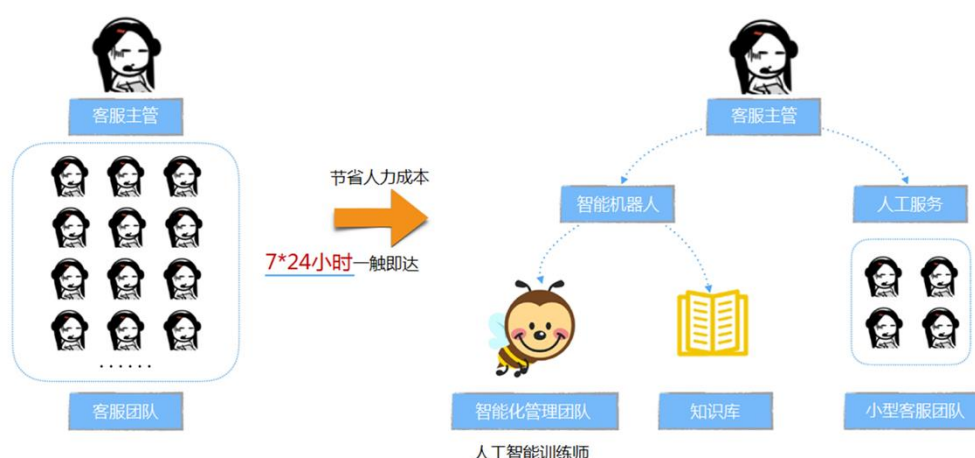
云智能对话机器人提供基于本体 FAQ 库、机器阅读等技术的问答服务，是构建智能客服系统的利器；同时基于对话工厂的多轮对话配置工具还可以将现有业务系统集成到机器人中，实现诸如订单查询，物流跟踪，自助退货等类型的自助客户服务。

- 阿里首款智能助理，以AI和人工结合方式提供智能购物、极速服务的私人助理体验。
- 与手机淘宝深度整合
- 正式上线：2016.4
- 当前智能解决率:89%（2016双11: 93.5%）
- 覆盖阿里巴巴生态圈二十余个业务BU
- 每天服务500万名消费者



阿里云智能对话机器人是人工智能在客服领域的典型应用，使得传统大团队客服模式转变为“人工智能+小客服团队”模式，极大的降低了客服领域的人员成本和学习成本，同时也提高了客户响应速度和用户体验。

“单一人工客服”向“智能机器人+人工服务”模式转变



1.2 产品优势

易于使用：无需成为 NLP、AI 专家，即可使用智能会话的能力

稳定的自动伸缩的服务：完全部署在云端的云服务，无需关心基础设施的扩容和缩容

基于本体 FAQ 库的问答：通过 FAQ 库功能录入问答形式的业务知识，会话机器人即可通过 NLP 和 AI 能力智能地理解机器人访问者的询问，寻找匹配度较高的答案

集成现有系统实现业务自助处理：通过对话工厂的功能可以将会话机器人与现有的业务系统集成起来。会话机器人可以理解访问者意图、通过现有业务系统对访问者的意图进行处理

便于使用的 API：提供标准会话 API，简单易上手。

1.3 客户价值

➤ 提升客户体验：

秒极响应；24 小时；高峰响应；基于实时需求和障碍预测主动服务

➤ 降低运营成本：

智能对话机器人解决简单/一般问题；智能辅助提升效率和弹性能力

➤ 服务数据化：

客户原声沉淀与分析；问题与发现、在线服务数据闭环

➤ 提升客服小二幸福感和工作价值：

智能辅助，提效，问题与发现，快速定位客户问题

➤ 提升管理效率：

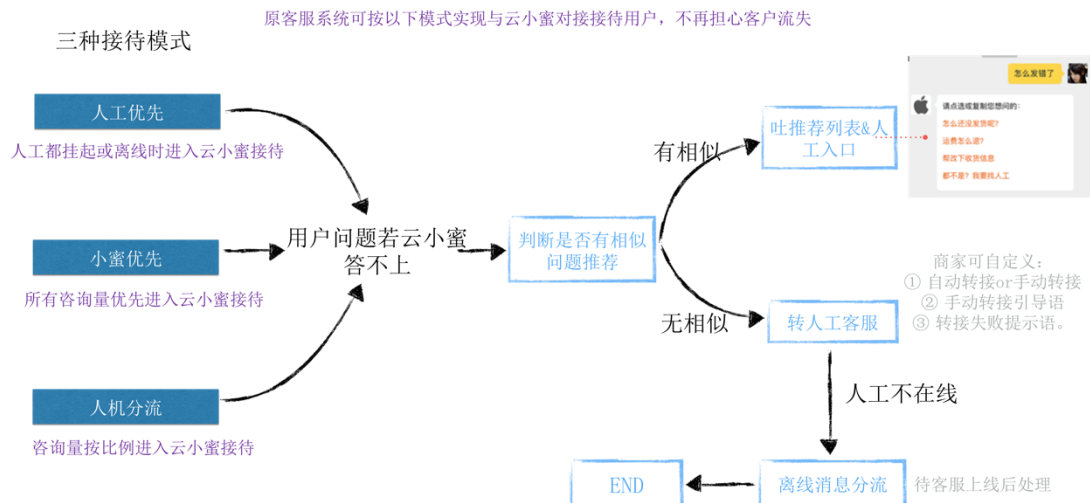
提升服务团队/呼叫中心管理效率、降低数据安全和资金风险

二、 应用场景

2.1 在线人机交互

用户在线提交的问题通过阿里云智能对话机器人机器人的会话接口提交给机器人，由机器人理解后响应作答，再返回给用户。

机器人作答可以和人工客服无缝对接，以便更好的为用户服务，以下是三种可行的接待模式：



典型应用场景：

➤ FAQ 式知识问答

可通过网站、APP 等渠道接入阿里云智能对话机器人，当用户进入阿里云智能对话机器人，可首先呈现欢迎语，一些常见热点或者个性化的相关问题，如果无法满足当前服务诉求时，无缝转接人工服务





➤ 商品导购



➤ 生活助理。如查天气

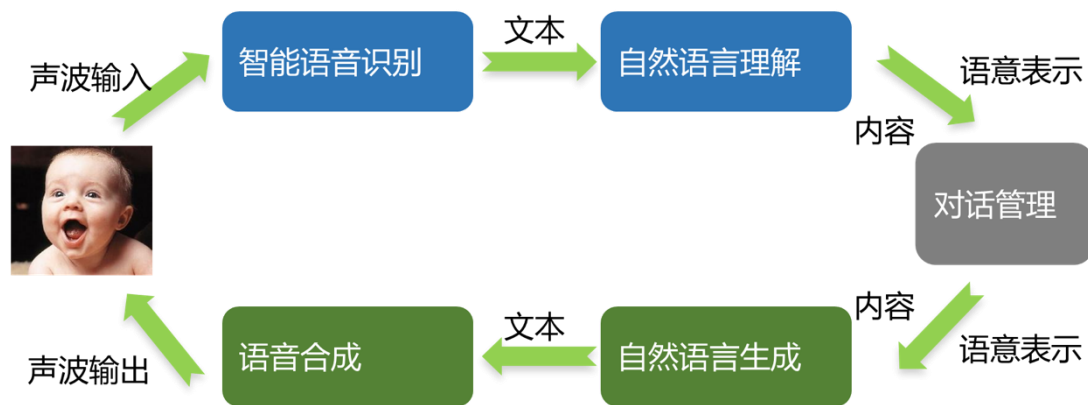


➤ 业务办理。如订机票



2.2 语音人机交互

阿里云智能对话机器人+语音识别可实现基于语音的人机交互应用。



用户的呼入语音先通过 ASR 语音转写服务转换为文本，再调用阿里云智能对话机器人会话接口进行会话理解和返回文本结果，然后再调用 TTS 服务合成为自然语音，返回给用户。整个过程无缝衔接，用户感觉到的完全是自然语音对话交互。

典型应用场景：

- 与 IVR 结合实现智能语音导航
- 智能知识问答，响应用户业务咨询
- 客服话席智能助手（金牌话术）

机器人可根据用户的问题快速给出知识问答答复或者推荐话术，引导客服人员给出更优答复，也降低了客服人员的学习成本和服务差异性。

智能语音导航/交互场景



智能外呼/回访场景



三、 功能简介

3.1 系统管理与账户配置

3.1.1 租户及用户的管理功能

针对客户对租户和用户管理的需求，提供系统下租户及用户的增删改查，密码重置

3.1.2 系统关键信息概览功能

账号总览支持系统级相关信息的浏览，便于客户了解自身的系统的性能限制级授权情况，可支持查看系统到期时间、已分配/已购租户数、已分配/已购机器人数量、产品 AccessKey，NLP 并发总数等信息。

3.1.3 高级模型配置功能

高级设置是针对客户侧调整相似问法和同义词的上限，选择开关 LDAP mode 并配置参数信息。

3.1.4 标签管理

标签管理是对各项内容进行标签标记的统一管理功能，能够支持标签的构建和引用，从根本上解决原先标签系统管理复杂、无法统一管理、缺乏引用机制、标签重合等问题，新的标签管理功能支持标签的增删改查，需要引用标签的功能模块可以直接搜索或者选择标签进行

引用。

3.1.5 全局名词

全局名词功能作为提升问答效果的干预手段之一，配合全局对话中控，为客户提供简单快捷的同义词配置方式，支持同义词的增删改查、批量操作，支持双环境切换，全局名词的配置将直接影响算法分词效果以及对同义表述的识别。

3.1.6 自定义登录页配置功能

支持自定义登录页、自定义 Title 功能。

3.2 系统集成服务

3.2.1 会话接口 API 调用功能

支持会话接口的集成调用，提供问答会话接口 API。

3.2.2 Chatui 一键部署功能

支持集成 Chatui,可在 chatui 框架下自定义对话前端界面。

3.3 智能问答

3.3.1 问答机器人管理

支持多个问答机器人的构建，支持依据业务域构建机器人，功能上支持机器人实例的新建，删除，编辑，包括名称、备注等。

3.3.2 知识挂载功能

支持多轮对话、知识图谱、FAQ 问答、闲聊、MRC、TableQA 的知识挂载功能，系统支持

不同机器人挂载不同知识,服务不同场景的能力,挂载后的机器人具备对应的知识问答能力。

3.3.3 闲聊功能

支持机器人闲聊,可以通过配置进行开关,支持内置的闲聊语料包。

3.3.4 自定义问答策略配置

支持依据问答需求自定义问答策略,在问答策略配置中机器人会对用户问题与 FAQ 库里内容匹配,并根据匹配相似度做出回答,用户可配置直出优先或推荐优先的策略;还可以对推荐的最大问题数量进行配置,当超出推荐阈值时,按配置数量返回相似度最高知识条数。

3.3.5 知识引用与上传下发

支持知识引用与上传下发功能,支持集团式知识分发。

3.3.6 自定义多视角管理

支持知识引用与上传下发功能,支持集团式知识分发。

3.3.7 多轮问答能力-对话工厂

支持画布式构建多轮对话,系统支持不同节点的组合和对话流的构建,支持意图、服务、变量、实体的构建,对话工厂能够适用于各项多轮对话场景,支持问答业务流的构建,节点上支持判断、意图、回复、服务、函数等节点,其他功能上支持全局变量、全局实体、全局服务等全局引用功能和转人工配置、内置意图的情绪识别功能,高效构建对话流。

3.3.8 单轮问答能力-FAQ 问答及管理

支持 FAQ 问答能力,支持 FAQ 的增删改查及导入功能,支持 FAQ 分类目管理、查询检索

功能，FAQ 问答同时可支持富文本问答及卡片式问答功能，支持丰富的答案表达形式。

3.3.9 文档问答能力-MRC

机器阅读理解问答，支持从非结构化文档中问答的机器阅读理解能力，支持上传文档解析，问答文档片段内容。

3.3.10 图谱问答能力-KBQA 知识图谱问答

基于知识图谱的问答，支持图谱 shcema 构建、管理，实体的增删改查，支持逻辑查询、多体查询、问答效果干预（包括图谱引擎模块拒识、属性高级语义配置，统一实体同义词，在重构之后，可以降低 badcase 错误率，提升问答效果，解决问答缠绕问题）等 KBQA 高级问答功能。

3.3.11 表格问答能力-Table QA

支持表格化问答能力，针对客户表格内容的问答功能，支持客户上传表格信息，自动化解析表格信息，在线编辑表格，表格管理及发布等功能功能。包括表格管理和表格配置两个模块，能够针对复杂表格数据进行快速检索和问答，且支持多轮问答、条件问答、比较问答等多种问答能力，同时支持选表干预、同义词干预、答案生成干预、SQL 干预等多种干预能力，TableQA 的上线，对于基于 Excel 表格数据的问答场景，例如 114 查号、公安查询车牌、基金理财产品查询等数据场景等具备非常好的支撑。

3.3.12 英文问答机器人

支持配置英文问答机器人，提供英文问答能力。

3.3.13 卡片工坊

支持多模态知识样式的构建（配合 chatui 使用），支持不同卡片类型的调用和引用。

3.4 高级运营工具模块

3.4.1 问答训练

拉取机器人问答不佳的问题，人工可通过增加扩展问、或者标注问题到对应意图的方式进行标注，训练机器人能够在下次遇到此类问题时正确作答。

3.4.2 模型训练平台

在用户发出问题时，模型能够解析问题的语意来将问题匹配到对应的知识，可以进行模型训练及测评可使模型变的更合理，使线上机器人回答问题时定位更准确，系统支持不同模型的训练功能。

3.4.3 问法聚类工具

为客户提供高效冷启动工具，针对用户上传的话术测试集进行无监督聚类，帮助用户进行FAQ 的快速冷启动。

3.4.4 全局问答配置

支持配置包括全局服务、全局同义词、全局视角等配置，增强问答效果。

3.4.5 问答能力测评

可使用现有的话术测试机器人的回答准确率，需要先下载话术模板，按模板填写问题，在话术测试集进行测试。

3.4.6 FAQ 挖掘

支持从非结构化文档中红自动化挖掘 FAQ 的功能，帮助客户快速冷启动，支持通过 QA 结构的文档、内容篇章结构的文档、人人对话日志、非结构化文档内容、指定 URL 内容中挖掘 FAQ。

3.4.7 智能荐句

基于 FAQ 问法及用户输入的问法自动化推荐相似问或意图话术，支持用户快速选用话术，大幅提升 FAQ 的配置效率，降低人工维护成本，提升系统的智能化程度。

3.4.8 测试窗功能

通过全局测试窗测试各个问答模块的 NLU 解析结果、问答链路结果、对话中控结构等复杂问答信息，帮助客户精准定位问题，快速 debug，提升问答效果。

3.5 高级运营工具模块

3.5.1 API 接口管控支持

支持通过 API 管控知识、机器人等关键数据及功能配置。

3.5.2 健康度检查

支持健康度检查功能，通过健康度检查检测重复、冗余 FAQ、相似问和对话流的用户问法。

3.5.3 审核功能/审核中心

审核功能从各个问答模块拆出，为全局提供统一审核中心机制，支持通过配置审核级别的方式，实现一到三级的分级审核，新增数据审核、审核开关、审核数据提交界面，提供全局一

致性的审核体验。

3.5.4 实时会话分析

按时间段对热门问题进行分析查询，并以报表的形式展示 top10 问题占比和问题出现的频次，展示各个问题的名称和出现频次，对聚类不达预期的问题，输入问题规则，可更好优化算法，可看到每次对对话的详情信息，对话的 seession_id、渠道信息、对话时间、对话内容等。

3.5.5 数据看板功能

支持包括对话轮次、热门问题、冷门问题、对话流问题、接待人次、对话轮次、会话平均轮次、点评率、点赞率、点踩率、无答案率、有效知识占比、有效对话流占比进行数据分析。

3.5.6 权限管理功能

支持客户权限管控功能，支持权限管理、自定义权限、自定义角色等功能。

3.5.7 业务空间管控功能

业务空间管控功能能够较好的帮助用户区分不同业务场景的对话机器人，并且实现业务与业务之间的数据隔离、模型隔离、语言区分，未来业务将作为精细化对话模型的基础设施，为不同行业提供个性化行业模型。

3.5.8 系统日志功能

查询用户变更、权限、登录等系统操作日志。

3.5.9 错误码查询文档

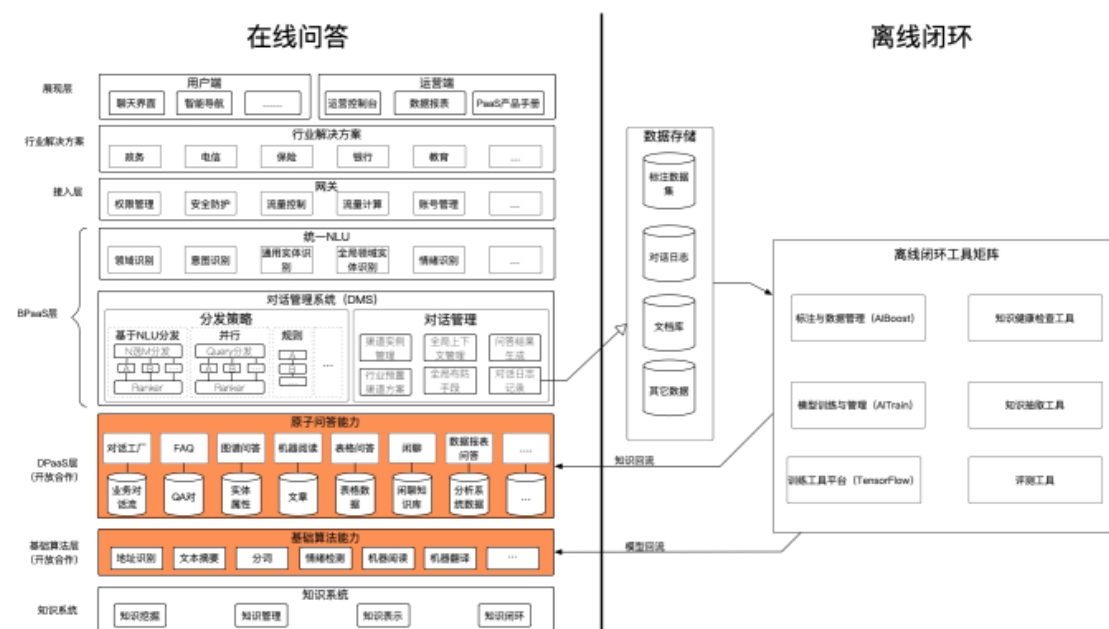
为了方便用户更编辑的查询错误码，我们在系统的右上角入口新增错误码说明的查询文档，通过点击链接可进入的错误查询文档，搜索对应的错误码可以查询到对应的错误和解决方法。

3.5.10 登录登出功能

系统支持用户依据分配的账号密码完成登录登出、账号修改等基础功能。

四、 技术特性

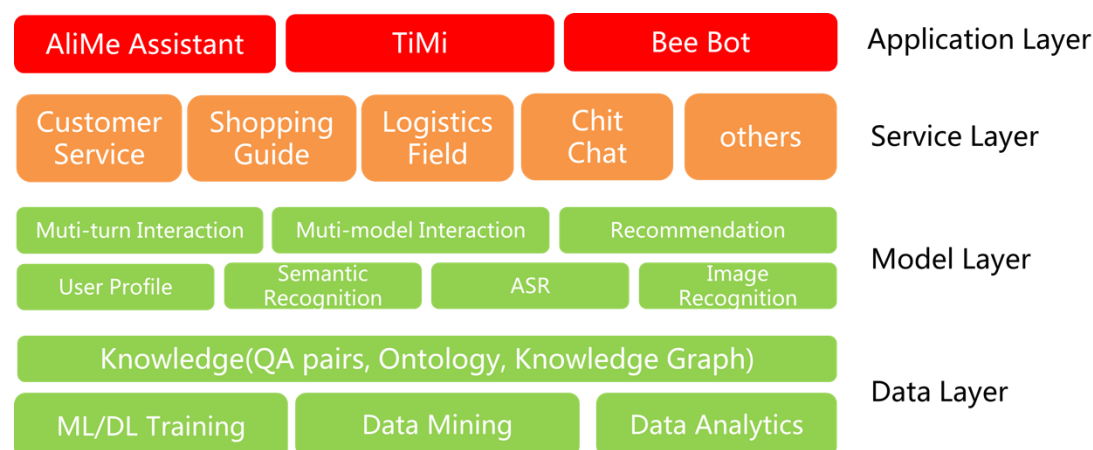
4.1 技术架构



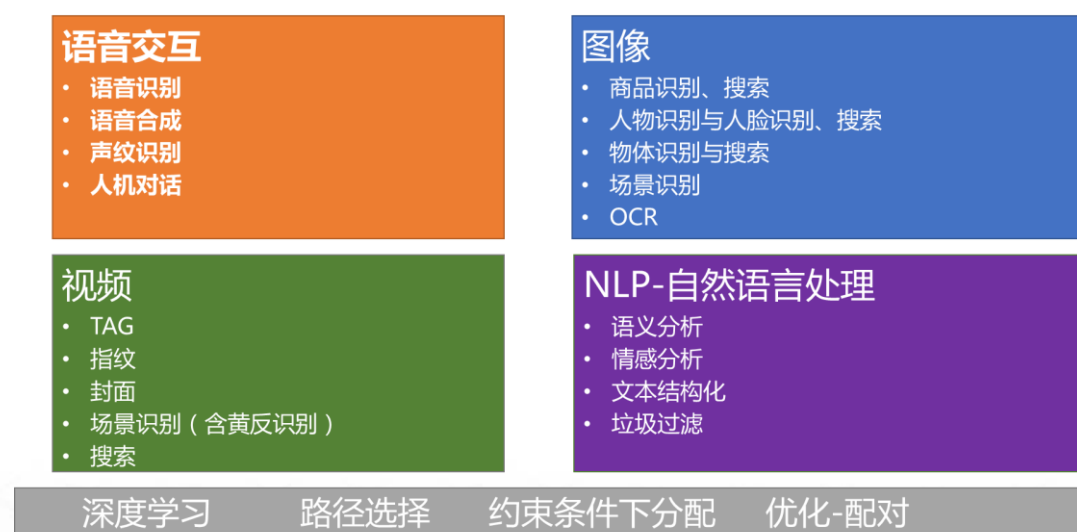
4.2 技术优势

4.2.1 完备的技术体系

阿里云智能对话机器人是阿里巴巴在自然语言处理、人工智能、智能客服领域多年来积累的智能会话能力的输出。



阿里的研究领域包括了机器学习、大数据挖掘、自然语言处理、多媒体识别、语音、图片识别等各个领域，并汇集了这些领域的前沿科学家。这些是阿里云智能对话机器人智能机器人能够在阿里集团、双十一战场，以及对外输出获得成功的一个重要优势。



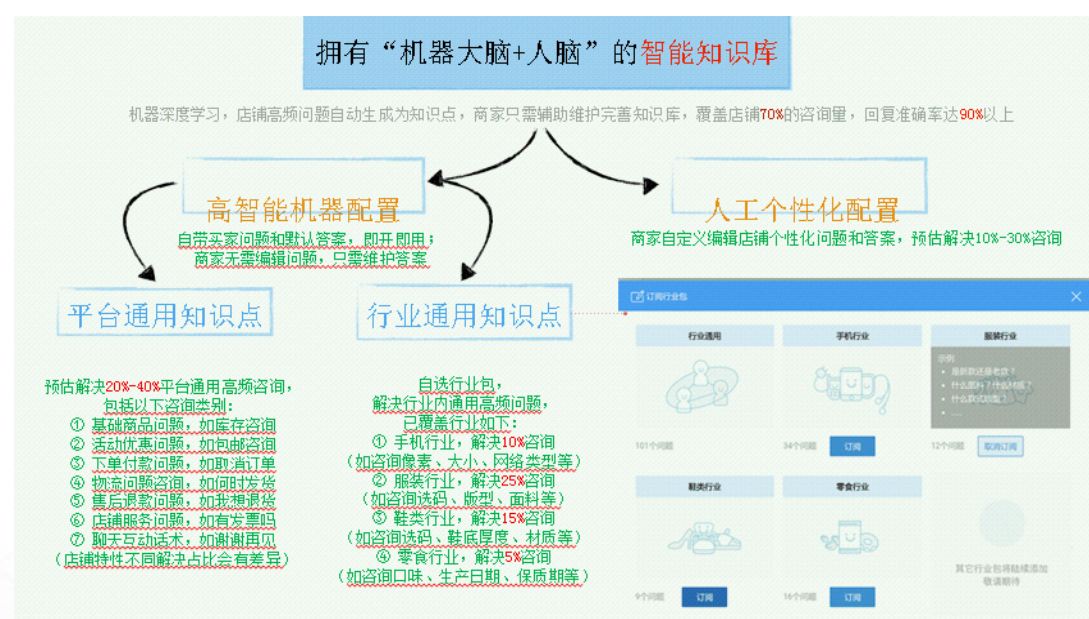
4.2.2 双十一大规模运用和技术沉淀

阿里云智能对话机器人是阿里小蜜系列产品首次向阿里集团之外的开放和赋能，集成了阿里小蜜在自然语言处理、人工智能、智能客服领域多年来的积累，以及历经了多次双十一考验沉淀下来稳定性、可靠性和准确性。

- 阿里首款智能助理，以AI和人工结合方式提供智能购物、极速服务的私人助理体验。
- 与手机淘宝深度整合
- 正式上线：2016.4
- 当前智能解决率:89%（2016双11: 93.5%）
- 覆盖阿里巴巴生态圈二十余个业务BU
- 每天服务500万名消费者



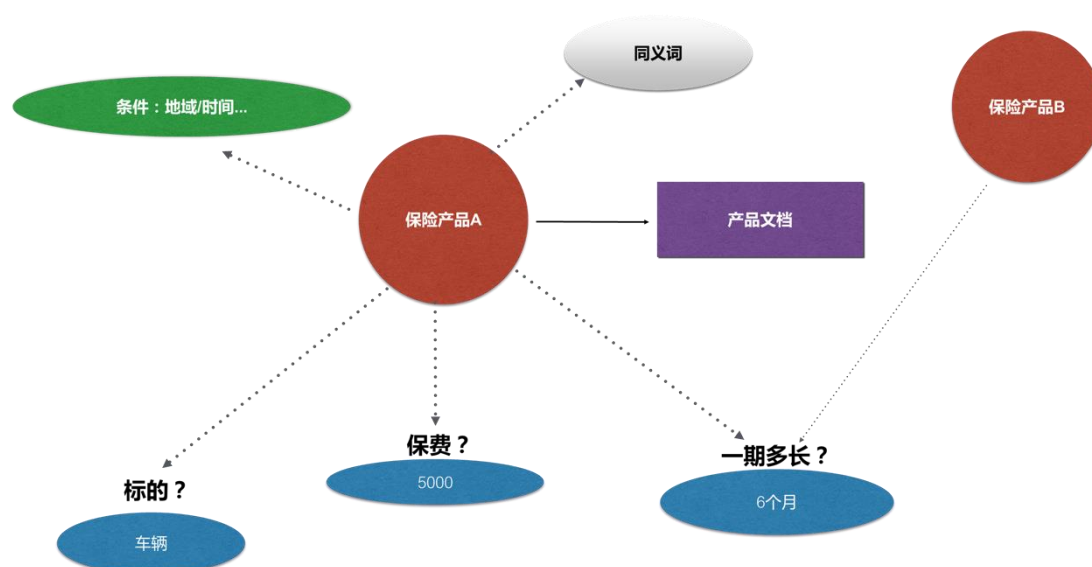
同时，阿里云智能对话机器人也利用阿里的大数据挖掘平台，收集挖掘各通用行业领域范围内知识语料，以更好的应对不同的应用场景。



4.2.3 基于本体结构的 FAQ 库

通过 FAQ 库功能录入问答形式的业务知识，会话机器人即可通过 NLP 和 AI 能力智能地理解机器人访问者的询问，寻找匹配度较高的答案。

阿里的本体知识库设计荣获国家专利，在国际大赛中机器阅读模型排名世界第 8、在业界首次实现中文机器阅读、在业界首次实现商用化的机器阅读模型。



阿里的知识库学习能力，支持从产品文档中进行阅读，并通篇理解整个文档，从中提取本体，比如保险产品 A，并在本体的基础上提取各种属性，比如保险的标的、保费、期限、条件等。可以说阿里的机器学习能力能够模仿人进行阅读理解并回答特定的问题。在此基础上：

- 实现上下问继承
- 实现问题反问

- 实现机器阅读的基础
- 大幅度减少相似问法

下图是阿里为某移动公司实施的本体理解和机器人问答示例。

飞享系列套餐 [点击编辑](#)

操作人：吉仁

更新时间：2017/08/22 13:56:44

同义词 [+ 添加同义词](#)

[飞享充值 ×](#)

[飞享优惠 ×](#)

[飞享活动 ×](#)

[飞享套餐 ×](#)

[飞享套 ×](#)

[飞享 ×](#)

简述 [编辑简述](#)

飞享系列套餐活动规则 一、活动时间：2017年8月4日-2017年8月31日 二、参与条件：广东移动139邮箱客户 三、参与方式：活动期间，广东移动用户通过邮箱短信验证码的方式进入活动页面，按提示登录查询话费余额等信息，点击【戳我抢1G】即可参与抢礼活动。四、活动细则：1. 活动期间，客户通过活动页面按提示登录查询话费余额等信息，点击【戳我抢1G】即可参与活动获得1G流量。2、同一号码每天可参与抢礼一次，同一号码活动期间只可中奖一次。3、中奖客户需用中奖手机号码于本活动开展期间，下载并登录和彩云手机客户端，按照领奖路径领取1G流量奖品。五、其他注意事项 1. 【重要】中奖客户需用中奖手机号码于本活动开展期间，下载并登录和彩云手机客户端，按照领奖路径领取1G流量奖品。2. 活动奖品说明：100万份1G省内7天2G/3G/4G通用流量，注：奖品为广东省内通用2G/3G/4G流量，每日限时（北京时间上午10:00:00至当日23:59:59）限量发放，抢完即止（本活动中奖概率约为90%，仅供参考，具体以实际参与为准）。活动奖品将于客户下载并登录和彩云APP后，自行在“活动专区的领奖专区”领取。获赠流量在领奖后的7个工作日内进行赠送，流量到账生效时间以短信为准；获赠流量不能用于转移、赠送、共享。3. 活动期间，手机号出现停机、换号、注销等情况，导致无法成功领取流量奖品，则视为客户主动放弃，不再进行补赠。4、以下情况不能获得流量奖品赠送：非广东移动客户、限定客户、无限量套餐客户、GPRS无限量套餐客户、4G随心王套餐客户、多号通的副号、一卡双号副号、家庭服务计划副号、万能副卡等虚拟号码客户、内部员工号码、集团代付/对公托收（如政企动力100行业卡）客户、集团全额统付号码、万花筒客户、手机状态异常（停机、销户、欠费、转品牌未生效等状态）号码；另有部分非实名制客户、M-Free卡客户、无线通、随E行、移动总机、固话、商话、公话、快捷通和金卡快捷通及CMWAP20元封顶套餐用户无法享受流量赠送。5、此活动为回馈粉丝的福利活动，如发现作弊或恶意信息将剥夺中奖资格，中国移动将保留追索全部损失的权利。最终核验的中奖结果以系统数据为准。

关联问题 [+ 添加关联问题](#)

问题：手机欠费是否还能参加？

[编辑](#) [删除](#)

回答：手机欠费的话是不能参加的

问题：分几个月返还啊？

[编辑](#) [删除](#)

回答：预存话费自办理当月起分10个月平均返还

问题：返还的费用及时到账么？

[编辑](#) [删除](#)

回答：赠送话费是即时到账

阿里云智能对话机器人的本体技术在阿里内部、通信行业、税务行业等多个行业都已经形成了成熟的案例。

4.2.5 多轮对话框架

客服机器人的多轮对话能力是业界考验机器人成熟度的重要指标，只有具备多轮对话能力的机器人才能支撑比较复杂的业务场景。阿里云智能对话机器人通过意图配置的方式，可以轻松实现多轮对话的场景，如下图所示：



所以阿里的客服体系先从场景进行划分，比如催收、回访、贷前咨询等多个场景，再在场景中配置各种意图，比如贷前咨询场景下的，是否能够提供折扣，贷款额度咨询等意图，而在每个意图之下，又能进一步支撑多轮对话。

4.2.6 问答训练

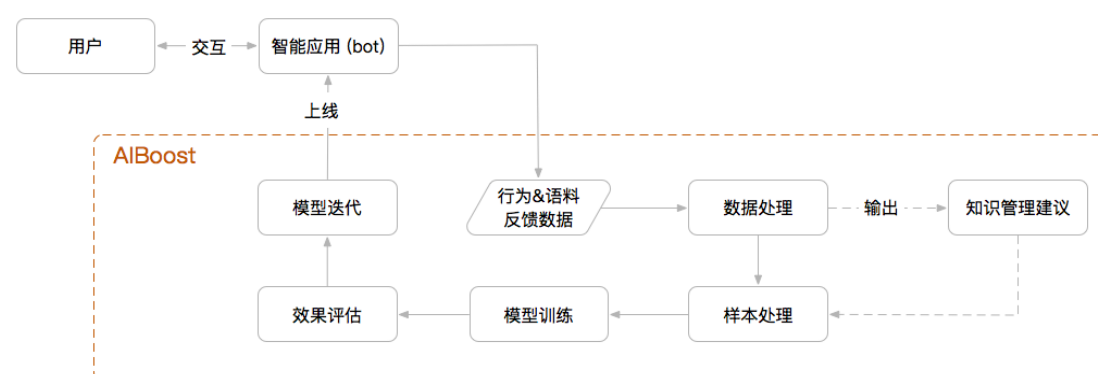
阿里的 AIBoost 是一个辅助智能产品优化的智能训练平台。在智能产品上线后，AIBoost 通过对数据进行智能化处理，充分协同智能训练师与算法工程师的工作，快速迭代优化线上效果。

知识回流的目标是辅助提升智能产品的解决率、覆盖率等关键指标，充分利用数据智能，提升机器智能训练的自动化程度，降低人工成本及时间成本。

知识回流的用戶包括智能训练师（泛指銀行內參與問答訓練的管理者及員工）和算法工程師。

如下圖所示，用戶和機器人的交互過程會完整的記錄和保留。

AIBoost 會將這些數據進行處理，對於需要更新 FAQ 庫的操作，提出知識管理的建議，對於機器人模型有待提升的點進行新的模型訓練和效果評估，並將升級後的客服機器人推廣上線。



整個知識回流的过程原理简单但处理过程实际很复杂，也是阿里的核心竞争力之一。首先要根据历史记录进行质量的判断，对于高质量完成的交互不需要回流，判断的过程用到一些简单的规则比如客户是否

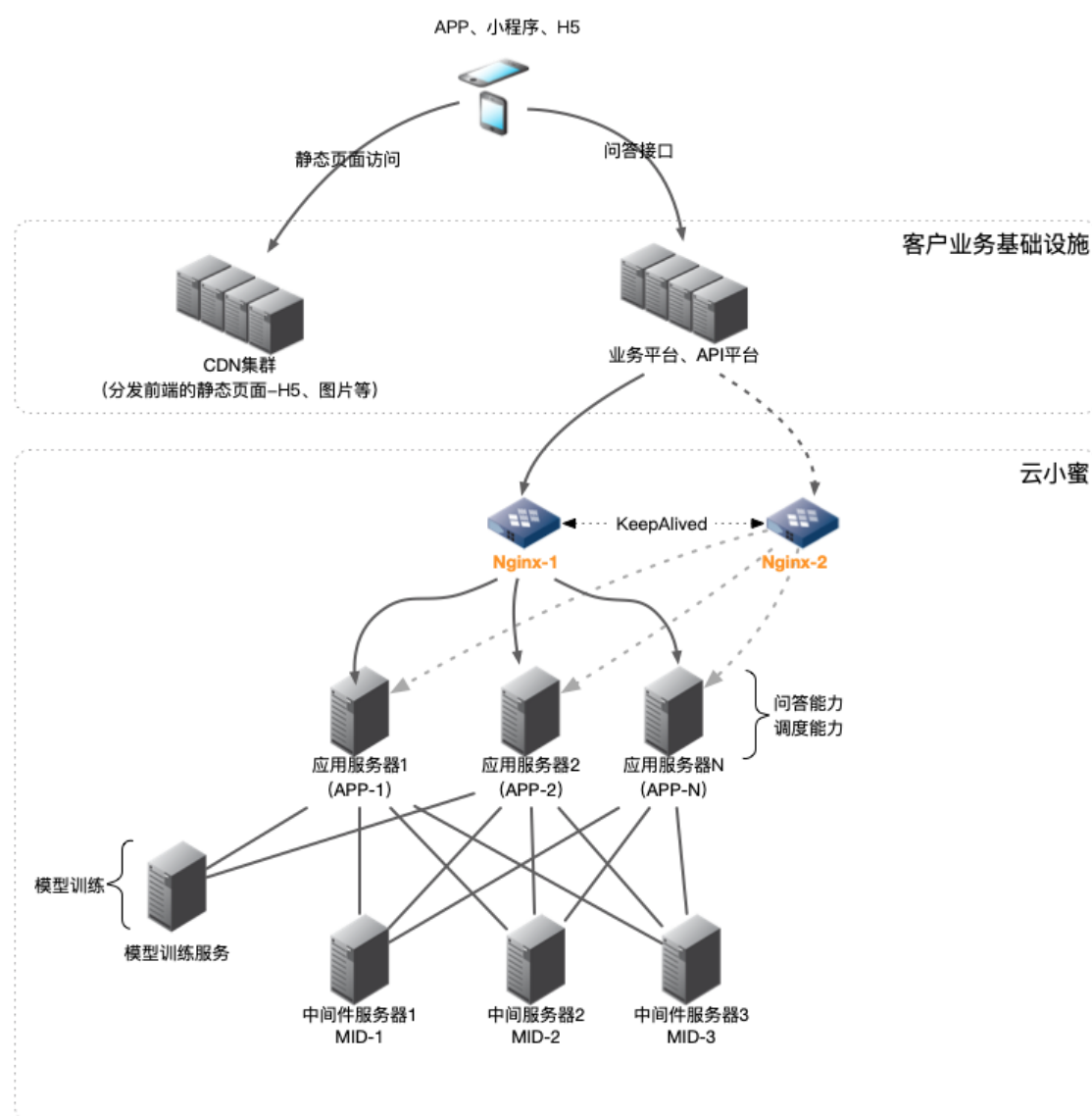
满意，也会用到复杂的算法比如聚类分析；对于问题对话回流到 AIBoost 中，再进一步细分为没回答的、回答错的和客户不满意的三类，针对这三类问题进行语料聚类 and 实体抽取，发现能够改进的模型或者知识点，在反馈到训练过程中，在 FAQ 库增加知识、或者对 FAQ 库进行纠错或者标注，最终完成模型优化。

阿里云智能对话机器人的知识回流特点在于：

- 维护成本低：不在需要维护相似问法，只要更少的 Q&A，通过机器阅读的方法实现问题的推理归类 and 匹配。
- 训练简单：问题数据可以实时回流，自动聚类，人工训练手段丰富，可以简单快速的完成训练过程。
- 上下文问答：能够支持一定层次的推理，支持逐步反问确定问题，大大提升了机器人的泛化理解能力。
- 自定义灵活：可以人工自定义场景、自定义接口调用、使得模型上线更加快速。

五、应用部署

5.1 部署架构

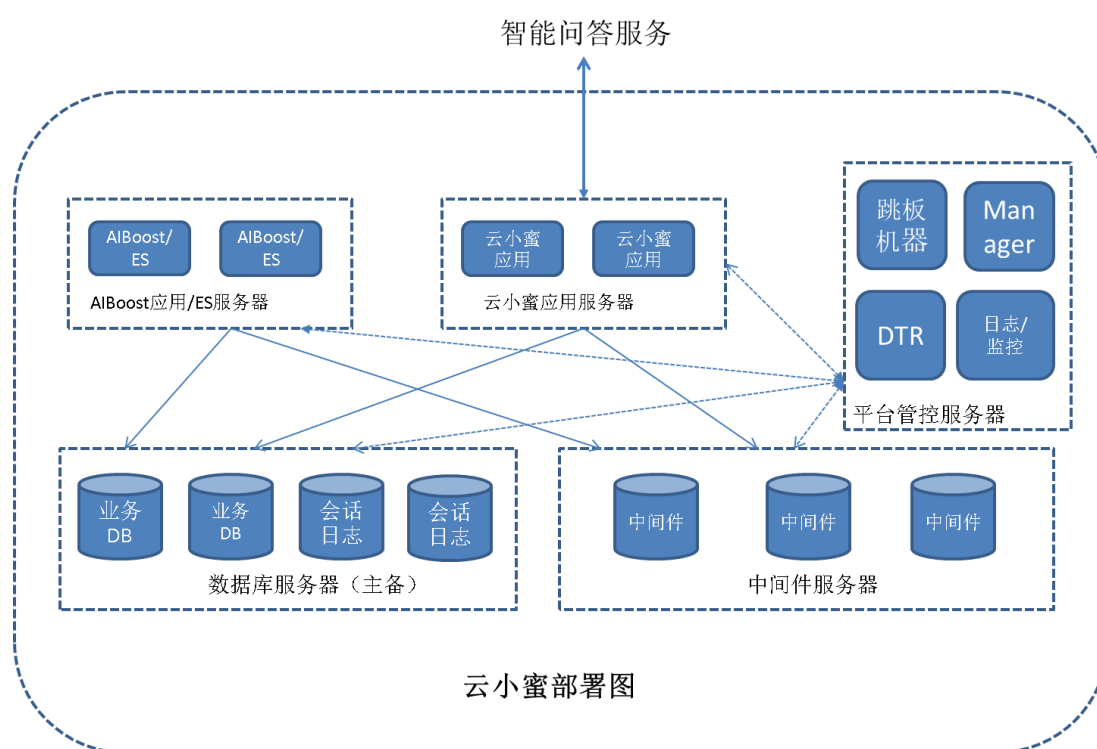


整个智能客服系统均采用内网部署方式实施，需要外网出口的只有 2 部分内容：前端 H5 页面、1 个问答接口。前端 H5 页面因涉及到客户加速问题，需要使用现有的 CDN 系统进行加速。文档接口直接走现有的 API 平台对外开放即可。

APP-1、APP-2、APP-N 是部署应用的机器，根据实际流量动态调整机器数量，4 台支持 500 QPS（业务接口延迟为 50ms 内）。

MID-1、MID-2、MID-3 是三台部署中间件的机器，根据实际流量可以调整机器数量（主要需要调整 Redis 集群数量），3 台中间件支持 800 路并发。

5.2 应用部署图



5.3 架构特点

本方案不只是应用本身，还提供运维监控能力；不只是提供智能会话，还提供了智能训练能力。

- 采用阿里云的飞天敏捷(容器化) 部署方式，应用实例、数据库、均可动态扩展。
- 采用分层架构，每层承载不同的功能逻辑，并且可以灵活扩展
- 数据库分为智能问答（业务）数据库、Alboost 数据库、会话数据库，其中会话日志数据库可横向扩展，用于快速数据存储及检索。各数据库均采用主备模式
- 其他缓存类中间件也采取多节点模式。

5.4 基于飞天敏捷部署，自带运维与监控能力

- 服务的 crash 监控和自动重启
- 程序日志查询和监控
- 灰度升级



附录一 名词解释

术语	其它叫法（英文）	含义
阿里云智能对话机器人	Intelligent Chatbot	阿里云的一款基于自然语言处理（NLP）和人工智能（AI）技术，面向开发者提供智能会话能力的云服务，包括公有云和私有化部署两个版本
AI	Artificial Intelligence	人工智能，是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。
NLP	Natural Language Processing	自然语言处理，实现人机间自然语言通信，是人工智能的一个子领域
ASR	Automatic Speech Recognition	自动语音识别，是一种将人的语音转换为文本的技术
TTS	Text To Speech	从文本到语音，是人机对话的一部分，让机器能够说话
AIBoost		人工智能算法迭代工具
API	Application Programming Interface	应用程序编程接口，是一些预先定义的函数，目的是提供应用程序与开发人员基于某软件或硬件得以访问一组例程的能力，而又无需访问源码，或理解内部工作机制的细节。
APP	Application	应用程序的缩写
私有化部署		私有化部署将软件直接部署在内网的本地服务器中，数据把握在企业内部。
SaaS 部署		企业通过购买 SaaS 服务提供商的服务，从而获得相应的所需功能，但企业通过该服务所处理的数据将无一例外的保存在 SaaS 服务提供商的服务器中。
IVR	Interactive Voice Response	互动式语音应答，用电话即可进入服务中心，可以根据操作提示收听手机娱乐产品，也可以根据用户输入的内容播放有关的信息。
对话工厂	Dialog Studio	对话工厂允许您按照既定的话术套路设计用户与机器人之间的对话，通过“Slot Filling(填槽)”技术，机器人可以与用户在对话中完成业务流程处理。当用户与机器人的会话匹配到意图时，机器人会按照预先编排的多轮对话来收集所需参数。当意图的全部参数收集完毕，机器人会调用意图处理逻辑中配置的服务来进行业务处理。以此可以实现：机器人订单查询、退货、订票等。

对话流	Dialog	Dialog Studio 管理业务的容器，比如订票，查物流，交话费为一个对话流（Dialog）。
意图	Intent	意图，可以以 NLU 和规则处理等形式被触发。
正则名词	LGF	正则名词的格式就是正则表达式，正则表达式是一种可以用于模式匹配和替换的规范，一个正则表达式就是由普通的字符（例如字符 a 到 z）以及特殊字符（元字符）组成的文字模式，它用以描述在查找文字主体时待匹配的一个或多个字符串。正则表达式作为一个模板，将某个字符模式与所搜索的字符串进行匹配。
同义词名词		在生活和业务当中，有很多不同的词语代表着同一种意思，例如：北京、帝都、中国首都。这几个词语都是指同一个城市，那么我们可以通过配置的方式将这些词语指定到一个固定的专有名词下面。

附录二 阿里云介绍

概要

阿里云是阿里巴巴集团旗下的公司，目前是中国最大最先进和全球第三大云计算技术和服务提供商。阿里云不仅对阿里巴巴集团内部的电商、金融、物流等业务提供云计算服务，同时 2010 年阿里云开始正式对外提供公有云服务，2013 年阿里云单集群规模 5000 台服务器的飞天集群正式开始投入生产，这标志着阿里云从此成为世界上独立拥有相关技术能力的屈指可数的公司之一，2014 年起阿里云陆续为公安部、海关总署、国家税务总局、新华社、中石油、中石化、中国联通、中国邮政等大型政府和企业提供云计算和大数据服务。

云计算是国家未来经济发展重要的信息基础设施，阿里云希望和国内云计算生态企业一起，致力于为中国政府和企业提供世界领先的云计算和大数据服务。

阿里云的使命是“提供无法计算的价值”，为大型政企用户服务的目标是将阿里巴巴集团基于自身海量业务实践沉淀的经验、技术和产品输出和赋能，进而帮助客户实现业务价值。

整体能力



刚过去的天猫双 11 全球狂欢节，有 225 个国家和地区消费者参与，当天支付峰值达到 25.6 万笔/秒。科技已经变成生产性支撑。



双 11 的背后，我们整体构架了全球最大混合云的平台，让整体 IT 成本下降 50%。同时，有超过 1200 个点 CDN 节点，让全球消费者能够快捷的接入到淘宝天猫的网站。当天，整体防御了 15 亿次以上的攻击。



现在，阿里云已经成为国家的基础设施，参与了国家电网、12306、人保金融、海关中国等一大批重点客户的建设。



今天我们跟国家密码管理局、贵州省密码管理局、数据通信技术研究所以一起，共同就云上贵州整体安全体系进行深度合作，这个合作我们已经开展了一年，我们今年可以正式向市场公布，国密局跟阿里云一起携手提供国家级的加密保护。



广东的制造业在全世界排名第五位，非常了不起。整个广东有 36 万家制造业企业，制造业和生产服务业整体占比 65%。去年，广东在阿里巴巴平台上整体的交易额达到 8756 亿。



目前，我们正在为超过 10 万家广东企业提供付费服务，其中有上万家规模以上企业。这是广东对阿里云技术和服务的认可。阿里云在广东一定会把工业作为我们最重要的服务方向，广东也是阿里云新制造战略的大本营。



珠江啤酒、OPPO、南航、珠江钢琴、金立、美的等源于广东的企业已经使用阿里云的技术和服务。



阿里云服务于广东，必须扎根于广东。我们在广东更多地投入研发人员。我们今天宣布阿里云正式成立广东研发中心，同时我们招募1000名云计算和人工智能工程师与广东制造业深度合作。

市场地位

阿里云在国内市场占据绝对领先的地位，其他第三方的市场调研及评价见：<https://www.aliyun.com/analyst-reports/>



平台特点

- 完整性：阿里云平台提供了体系完备、功能完整、IT 架构全覆盖的服务能力，包括互联网架构、大数据平台、整体安全的服务能力。阿里云平台既能支持三层架构的 IT 系统，也能提供 DT 服务，是支撑政府信息化跨越式发展的基础。基于阿里云，可以快速搭建 OLTP，OLAP，大数据，内容管理和搜索等多种、全面的企业应用。

- 扩展性：阿里云平台具备云基础资源的快速、弹性、持续和自动化供给服务能力，提供大规模、分布式集群的管控能力，通过增加物理设备，实现总体网络资源、计算/存储资源、内存资源和数据库资源的自动扩展；提供横向扩展能力，可根据业务、管理要求，跨机房、跨地区以 1 万台服务器规模增加云计算节点。
- 开放性：阿里云平台以 API 的方式提供开放接口，可以和第三方软件产品集成、兼容，实现阿里云平台与异构 IT 环境应用的互联互通。
- 成熟性：阿里云平台采用先进的设计思想和方法，符合技术发展趋势，既可以适应地理环境、应用场景、运维能力以及投资规模等客观因素，又可灵活地配置云服务种类和硬件设备，满足业务管理的需求。同时，阿里云平台经过高并发、高可靠的实践验证，确保平台的稳定和可靠。
- 安全性：阿里云平台支持冗余、自恢复、高可扩展编程模型，允许应用系统从不可避免的硬件、软件错误中恢复，确保应用系统的正常运行和数据存储的安全可靠。同时阿里云平台提供远程、跨节点的容灾机制，保证业务连续性。可按照云服务的使用范围以及层次，提供相应的云服务安全体系，并与安全防护体系、安全运维体系相结合，形成完整的阿里云平台防护体系。

- 自主可控：阿里云平台所有产品均实现了完全国产化，全部由阿里云开发并拥有全部代码，具有自主知识产权，符合国家网络与信息安全规范。

开放性

阿里云计算平台系统架构中由阿里云自主开发的云操作系统——飞天大规模分布式计算平台。飞天系统可以将上万台服务器的超大规模集群进行整合，使得集群以一台超级计算机的形态展示在其他服务以及应用面前。并在这台超级计算机之上提供 OSS、OTS、MaxCompute、ECS 以及 RDS 等服务，实现了在一个通用云平台上提供不同的云服务，为实现数据共享，打通信息孤岛提供了架构上的支持。阿里云“飞天”云操作系统完美的满足了先进云计算平台所必需的底层基础技术支撑。飞天系统负责管理数据中心 Linux 集群的物理资源，控制分布式程序运行，隐藏下层故障恢复和数据冗余等细节，有效地提供弹性计算和负载均衡的服务。飞天系统所提供的主要服务包括：资源管理、安全、远程过程调用等构建分布式系统常用的底层服务；分布式文件系统；任务调度；集群部署和监控等。

开放不等于开源，今天开源实现很多还处在使用技术的阶段，远远没有经过企业级平台产品必须的层层测试，包括单元测试、功能测试、压力测试、极端破坏性攻击测试、集成测试。成为开放标准往往是核心的技术或者平台需要经过比开源严格的多的企业级综合测试和验证。

开放主要指得是平台级的产品有对外开放的 API 或者服务能否进行数据层；数据操作层级的兼容和互通。开放的另外一个条件是有一定数量的生态合作伙伴基于该平台进行二次开发和应用开发。开放的深一层含义是指平台是具备核心竞争力的平台，无论是强大的计算能力，还是多维和实时的分析处理能力，这些能力在同类平台产品中处于领导者的地位。阿里云计算平台和大数据的核心是开放的技术体系和架构。

开放的另外一个层面是数据平台必须已经传统 500 强企业和互联网企业都使用过；传统 500 强企业代表了复杂的业务应用流程，互联网企业代表着用户承压和对新技术的率先验证使用。今天阿里的大数据和云平台已经为世界 500 强传统行业企业应用；同时被世界最大型的互联网公司实践应用过。

