

阿里小蜜机器阅读理解技术揭秘

张佶 算法专家

Powered by 阿里巴巴智能服务事业部 & IDST



AiCon

全球人工智能与机器学习技术大会

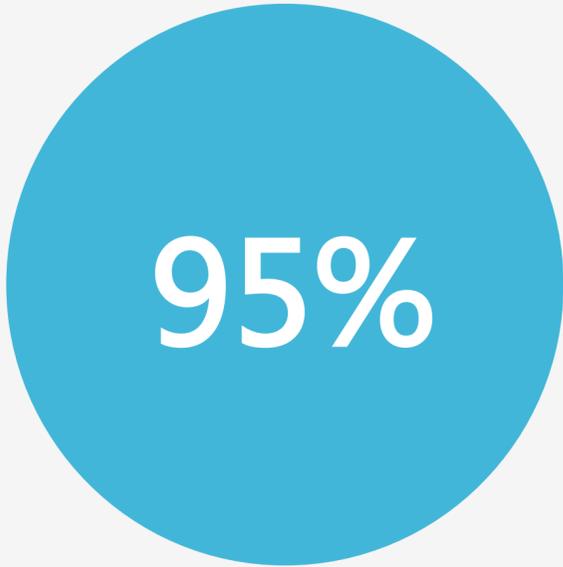
助力人工智能落地

2018.1.13 - 1.14 北京国际会议中心



扫描关注大会官网

阿里小蜜-双十一服务主力



阿里小蜜
智能服务占比



阿里小蜜
智能解决率



店小蜜
对话轮次

TABLE OF CONTENTS 大纲

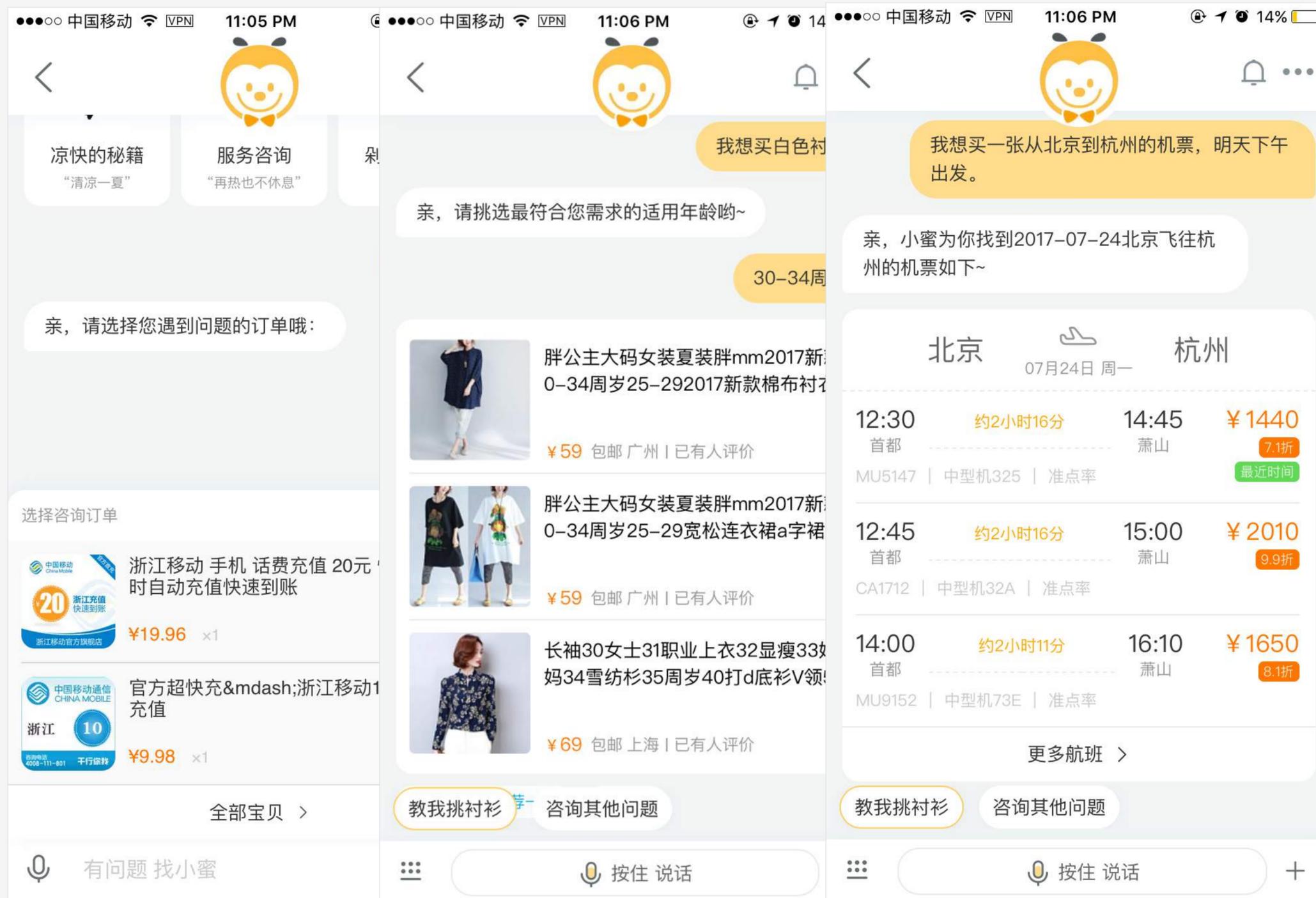
- 阿里小蜜平台介绍
- 机器阅读理解技术概览
- 业务场景及技术实践
- 挑战与展望

各领域生态圈的赋能



产品展示

- 阿里小蜜
- 客户服务
- 导购
- 通用助理
- 闲聊
- 运营活动
-



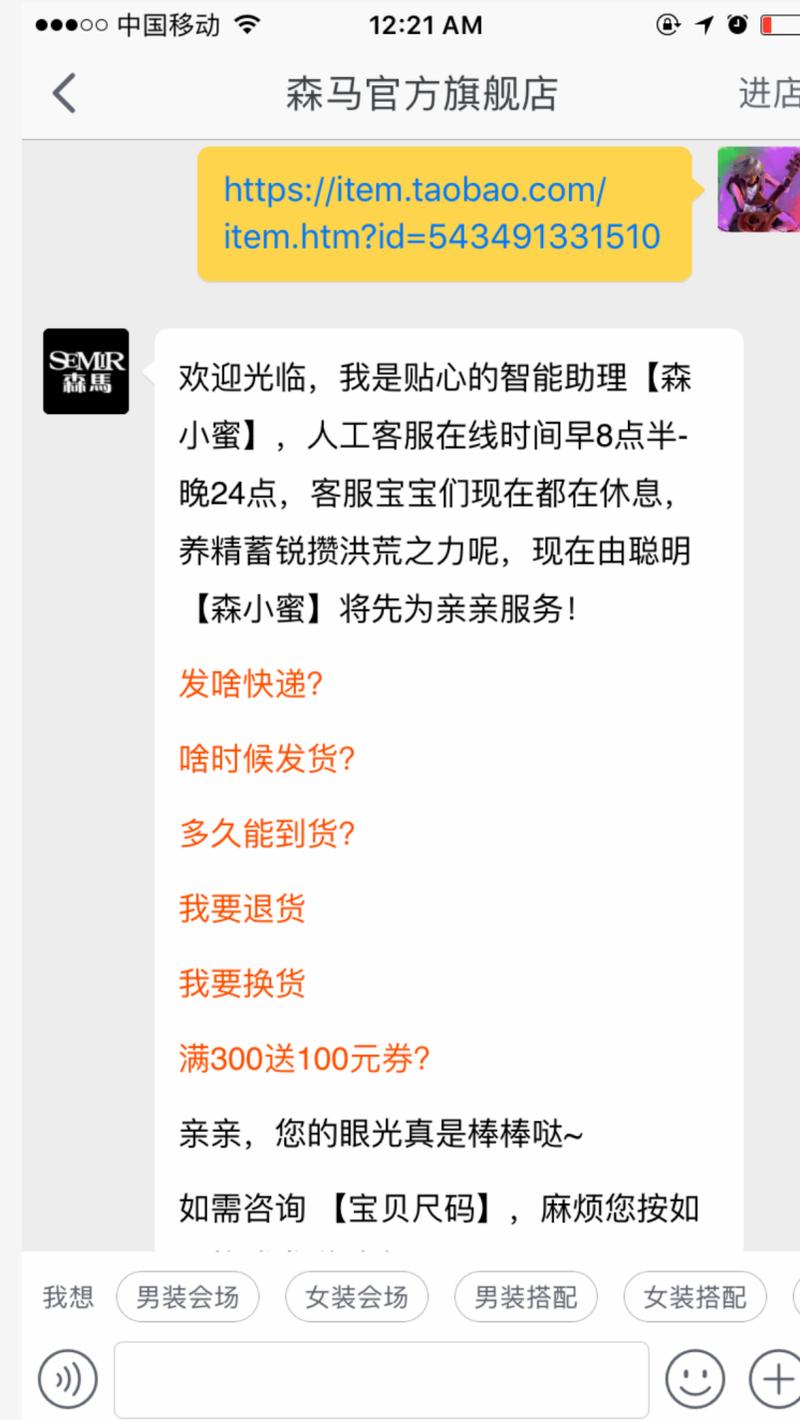
产品展示

- 店小蜜

- 通用模型：物流

- 行业模型：手机、服饰

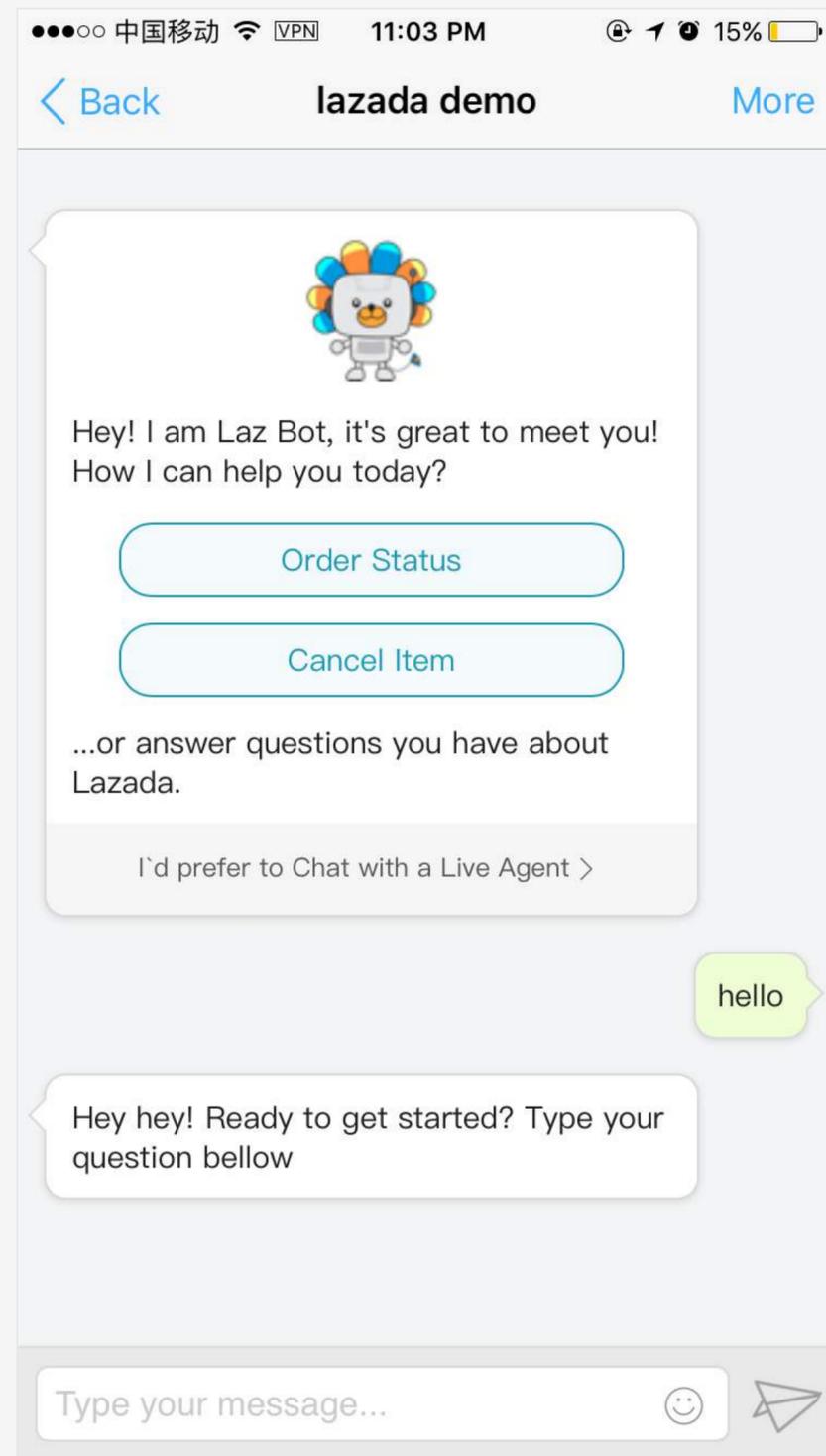
- 自定义QA模型



产品展示

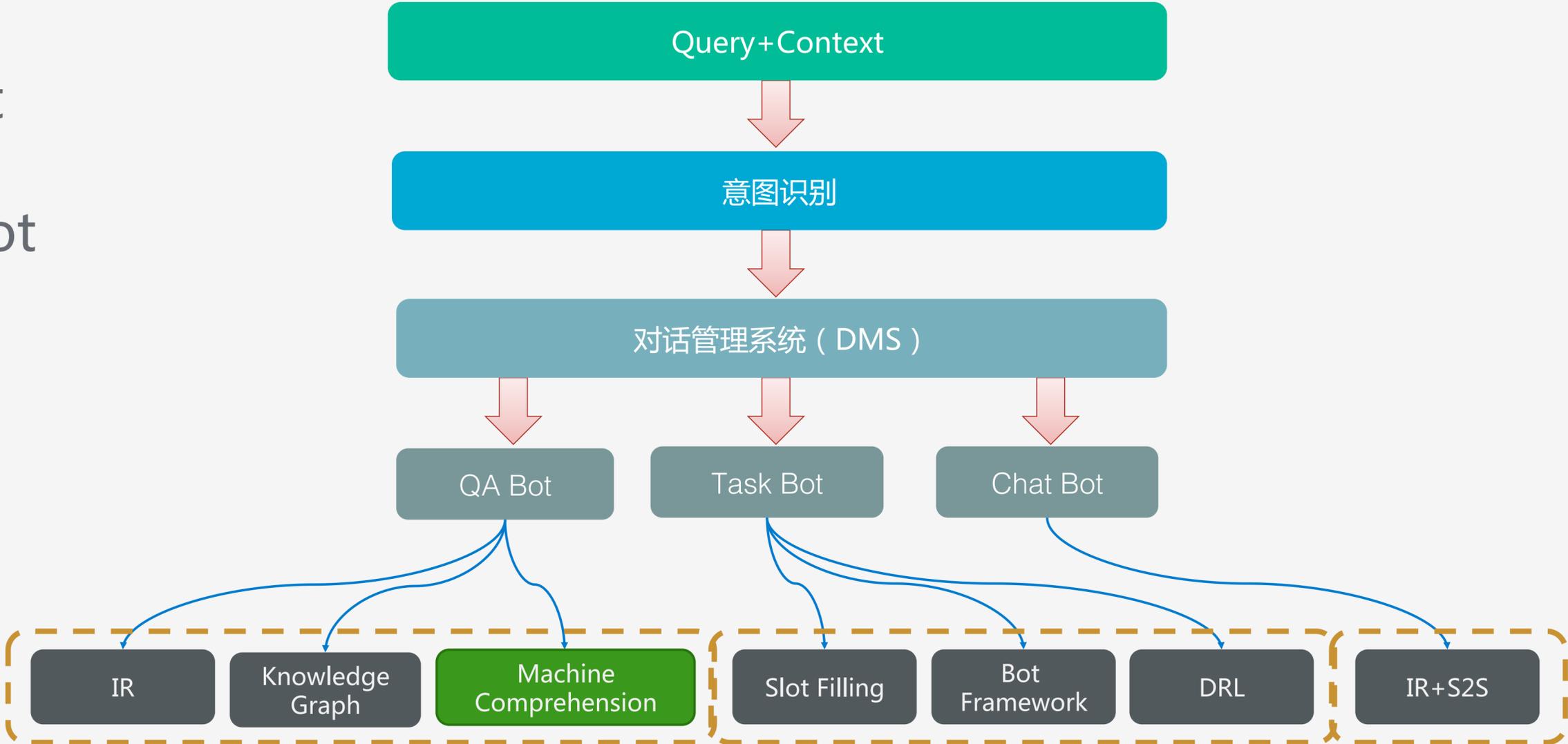
- 企业小蜜

- 完整的企业智能服务方案
- 本体知识库：企业百科式的知识库
- 预置领域包
- 多语言支持



分领域的模型支持

- 知识问答：QA Bot
- 任务解决：Task Bot
- 闲聊：Chat Bot



一种常见的QA Bot场景

- 问答的线索来自一篇文章（新闻、百科、邮件等）
- 其中包含很多的数字、地点、名称以及一些客观事实描述
- IR方式：人工逐个添加知识点“问答对”非常耗时，且难以穷尽所有问题，知识覆盖率低
- 知识图谱方式：图谱构建代价大、领域特点强

西溪园区位于文一西路969号，北靠文一西路，东临常二路，西靠高教路，南依创新路，隶属余杭区五常街道。一期项目总占地面积约16.4万平方，总建筑面积约29.3万平方，主体建筑有6幢办公楼，2幢停车楼，总工位数有14000余个，机动车位2279个，非机动车位2000多个，其中电动车充电车位820个。二期项目总占地面积约9.6万平方，总建筑面积约为15万平方，主体建筑有2幢办公楼，约5000个工位，1幢展示厅，近1号楼有中式园林的董事局办公室，将按建筑进度分期投入使用。二期预计可提供机动车位1200个。外围物业为上海明华物业管理有限公司，为园区提供安全秩序维护、设备管理维护、园区保洁等专项服务。



Q: 中式园林在哪里？

A: 近1号楼

Q: 西溪园区地址？

A: 文一西路969号

Q: 西溪园区二期有多少机动车位

A: 1200

Q: 西溪园区的物业公司叫什么名字？

A: 上海明华物业管理有限公司

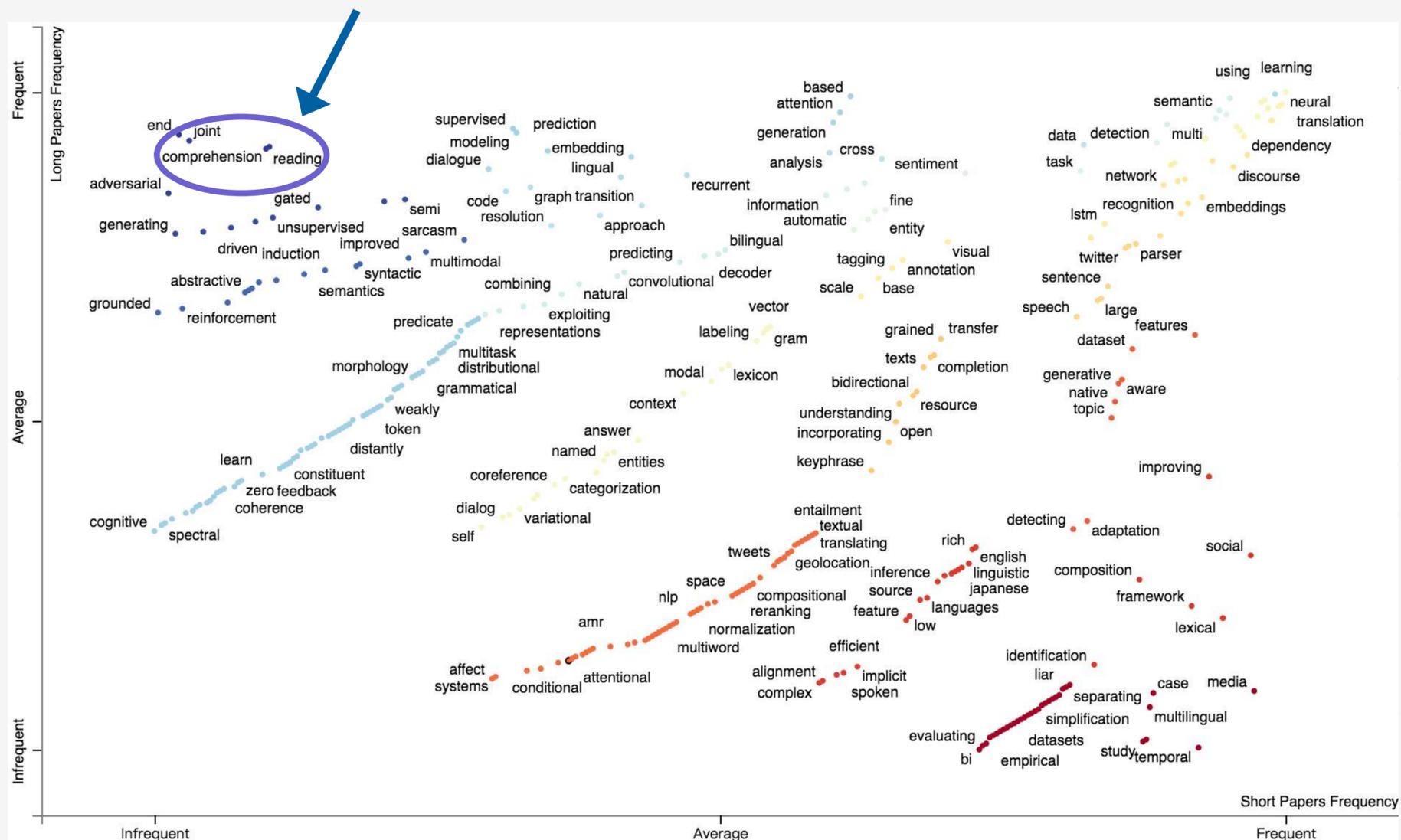
TABLE OF CONTENTS 大纲

- 阿里小蜜平台介绍
- 机器阅读理解技术概览
- 业务场景及技术实践
- 挑战与展望

机器阅读理解

- 基于深度学习技术, 将非结构化文章和问题作为输入, 自动寻找答案
 - ✓ 不依赖外部知识库和FAQ
 - ✓ 端到端进行模型训练
 - ✓ 学术会议上最热门的关键词和任务
 - ✓ 公开数据集推动了该领域的快速进步

Reading Comprehension



Visualizing the Corpus of ACL 2017 Paper Titles

公开数据集-推理型

- 推理型问答：Facebook bAbI
 - ✓ 由人工构造的由若干简单事实形成的英文文章
 - ✓ 回答基于推理的问题
 - ✓ 20个评测任务：事实、计数、是否等
 - ✓ 数据量小：1000个训练数据、1000个测试数据

Task 1: Single Supporting Fact

Mary went to the bathroom.
John moved to the hallway.
Mary travelled to the office.
Where is Mary? **A:office**

Task 2: Two Supporting Facts

John is in the playground.
John picked up the football.
Bob went to the kitchen.
Where is the football? **A:playground**

Task 3: Three Supporting Facts

John picked up the apple.
John went to the office.
John went to the kitchen.
John dropped the apple.
Where was the apple before the kitchen? **A:office**

Task 4: Two Argument Relations

The office is north of the bedroom.
The bedroom is north of the bathroom.
The kitchen is west of the garden.
What is north of the bedroom? **A: office**
What is the bedroom north of? **A: bathroom**

Task 5: Three Argument Relations

Mary gave the cake to Fred.
Fred gave the cake to Bill.
Jeff was given the milk by Bill.
Who gave the cake to Fred? **A: Mary**
Who did Fred give the cake to? **A: Bill**

Task 6: Yes/No Questions

John moved to the playground.
Daniel went to the bathroom.
John went back to the hallway.
Is John in the playground? **A:no**
Is Daniel in the bathroom? **A:yes**

Task 7: Counting

Daniel picked up the football.
Daniel dropped the football.
Daniel got the milk.
Daniel took the apple.
How many objects is Daniel holding? **A: two**

Task 8: Lists/Sets

Daniel picks up the football.
Daniel drops the newspaper.
Daniel picks up the milk.
John took the apple.
What is Daniel holding? **milk, football**

Task 9: Simple Negation

Sandra travelled to the office.
Fred is no longer in the office.
Is Fred in the office? **A:no**
Is Sandra in the office? **A:yes**

Task 10: Indefinite Knowledge

John is either in the classroom or the playground.
Sandra is in the garden.
Is John in the classroom? **A:maybe**
Is John in the office? **A:no**