

基于大数据的智能问答

浙江大学计算机学院

赵洲

2018年10月

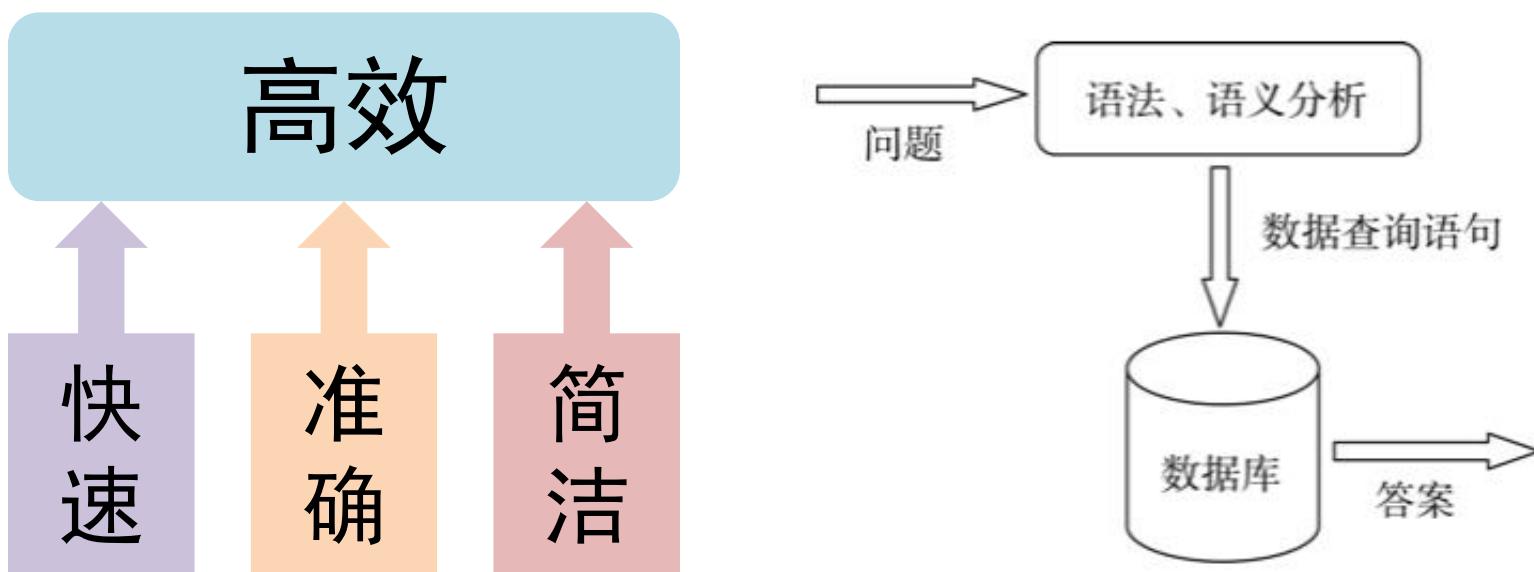
目 录

- 一. 智能问答介绍与历史发展
- 二. 基于大数据的问答应用
- 三. 核心问题和关键技术
- 四. 总结

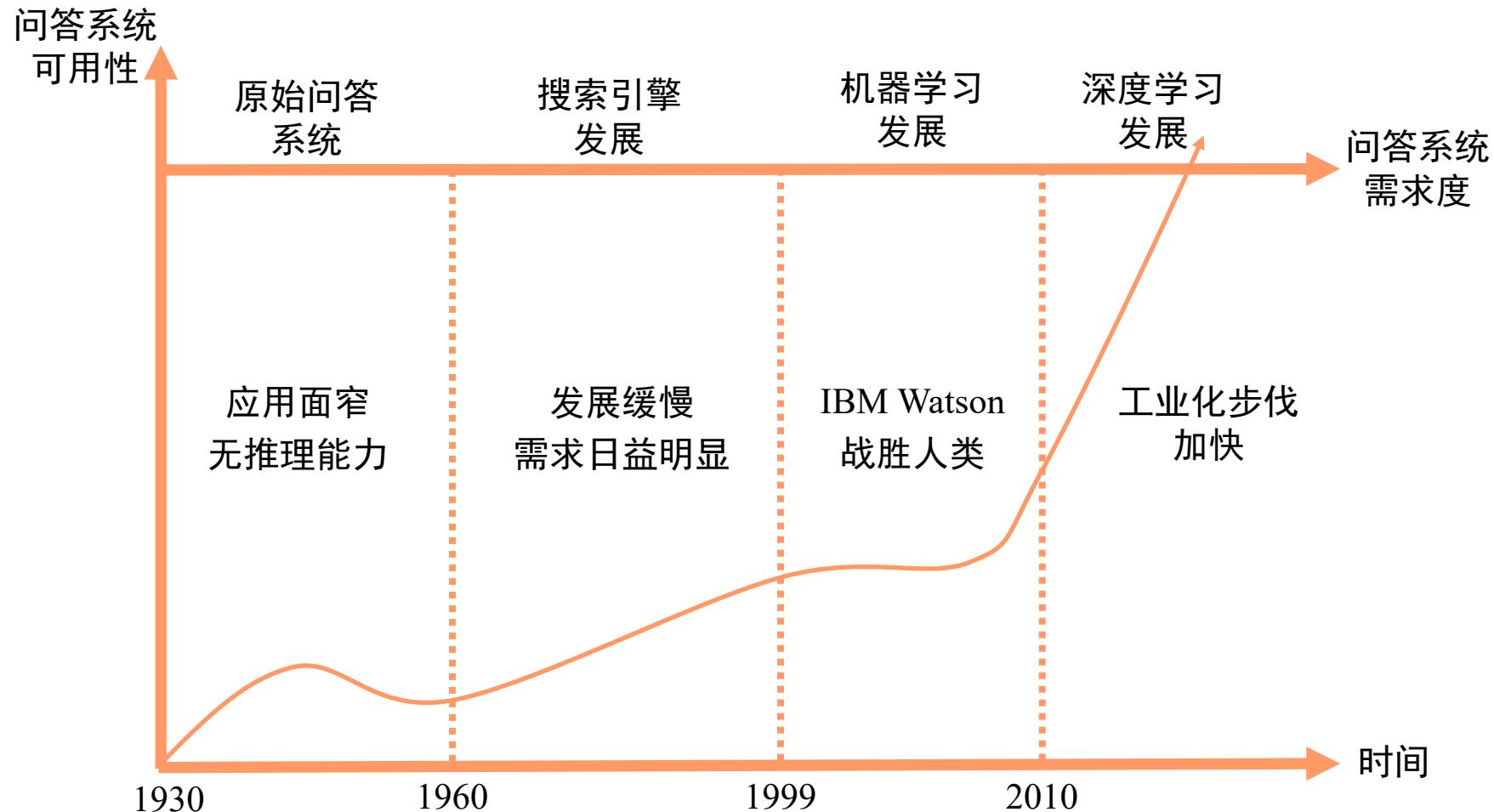
一、智能问答介绍历史发展

介绍

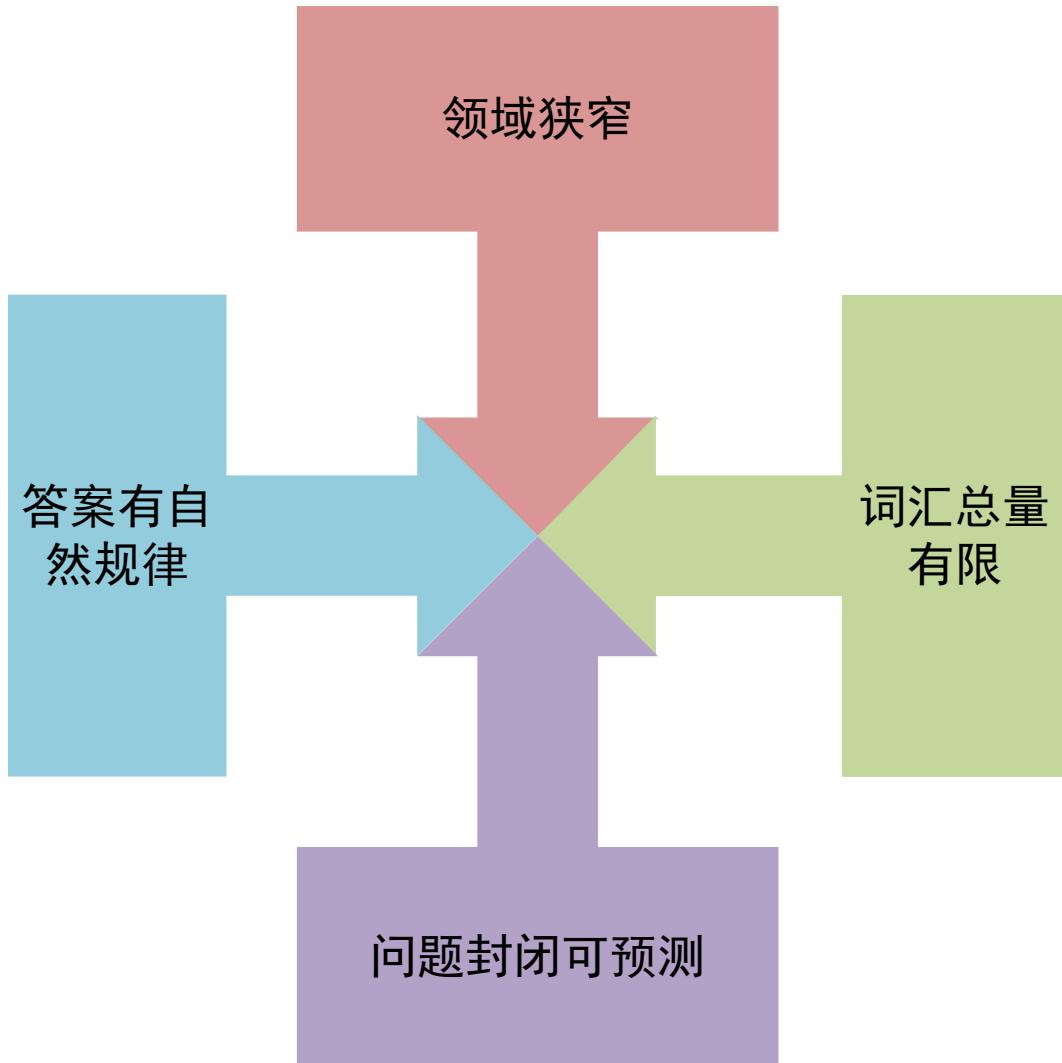
问答系统(QA), 是指接受用户以自然语言形式描述的提问(例如:世界上最高的山是哪座?),并从大量的异构数据中查找能回答该提问的准确、简洁答案(例如:喜马拉雅山)的信息检索系统。



历史发展



出现

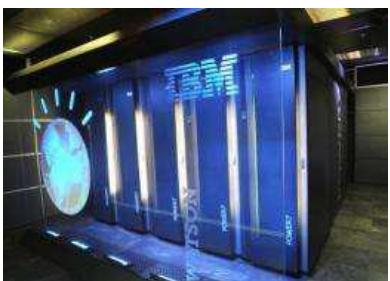
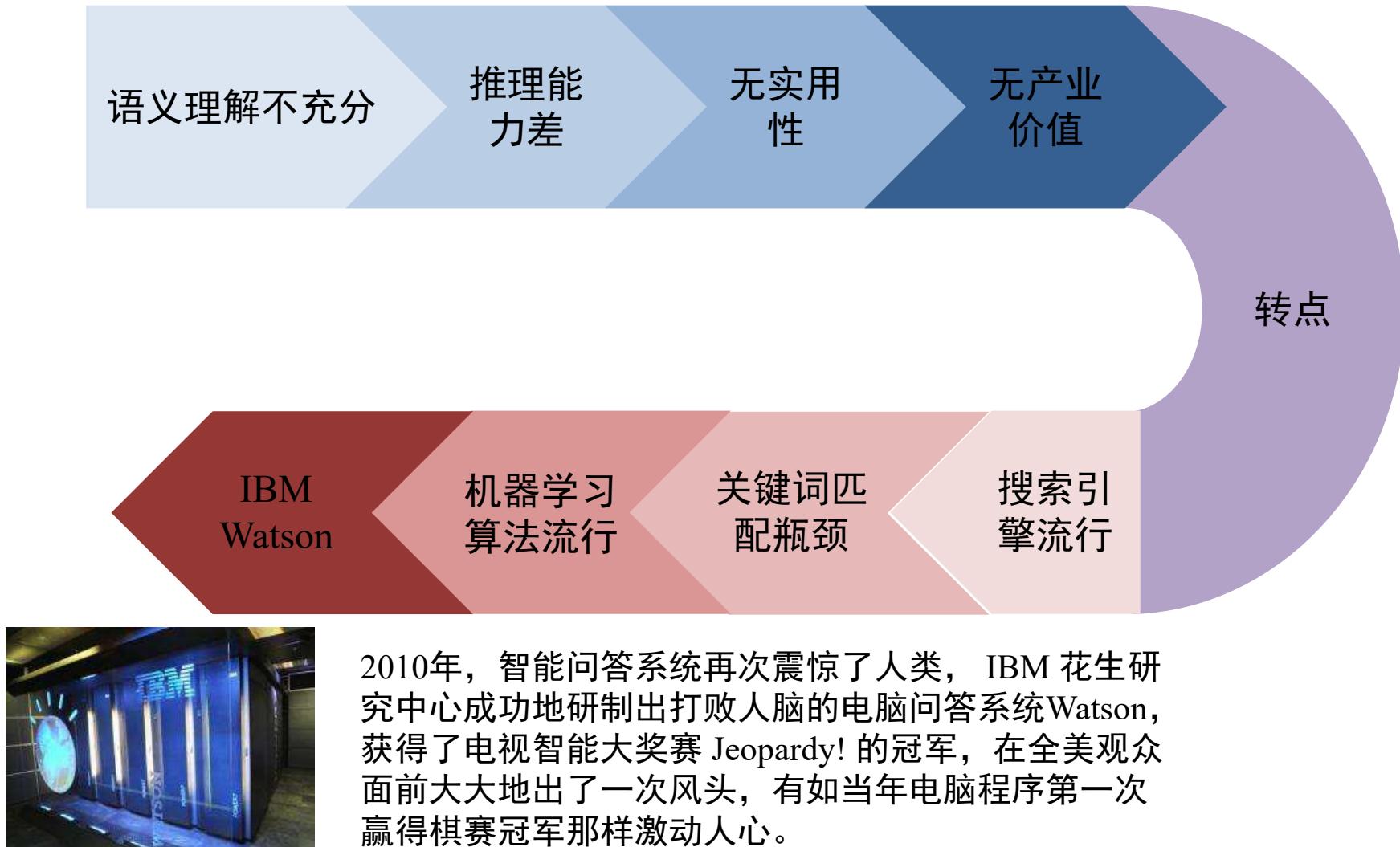


LUNAR系统



SHRDLE积木问答系统

陨落与爆发



现状

1

问答系统发展的多样化

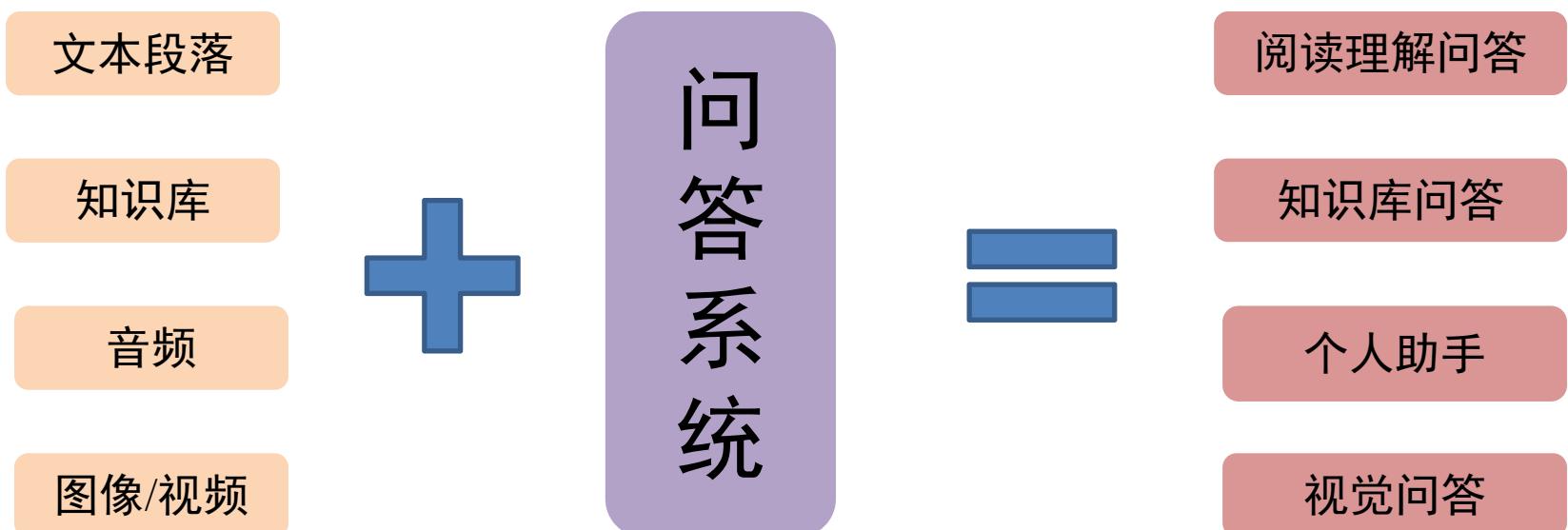
2

问答系统工业化脚步加快

二、基于大数据的智能问答应用

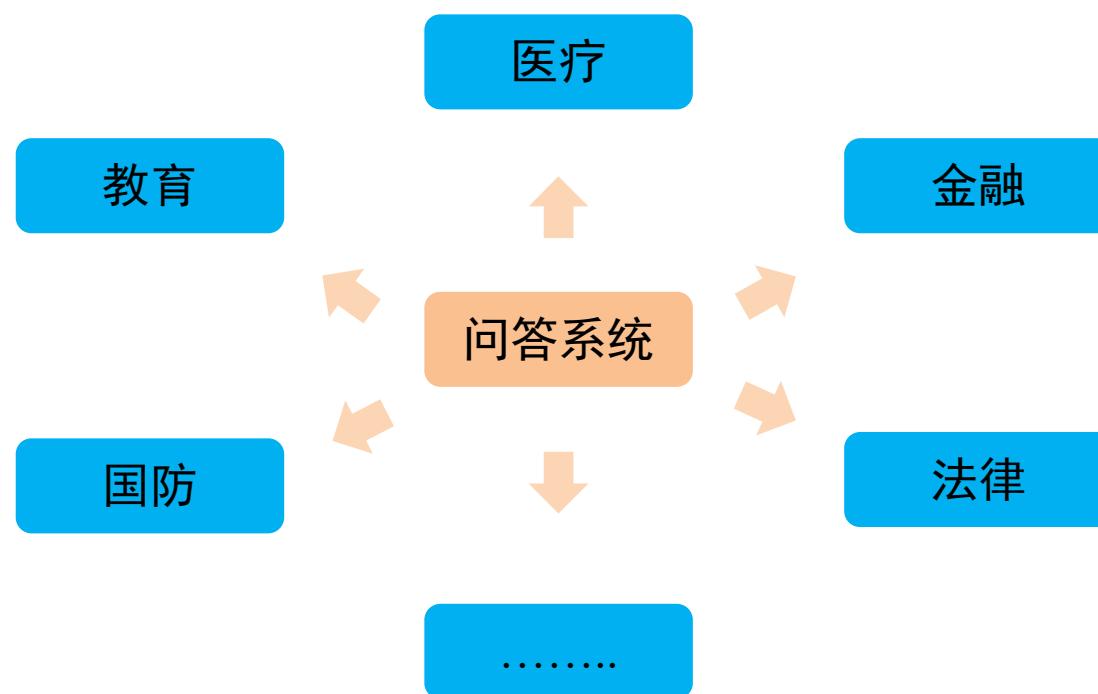
基于多元异构数据的问答系统

如今的问答系统不再拘泥于传统的文本检索形式，而是在此基础上，进一步融合知识库、音频、图像、视频等信息，形成不同模态的问答系统。



应用广泛

当前，问答系统已经被深入应用于诸多应用领域中，并悄然改变着这些行业的服务形态。

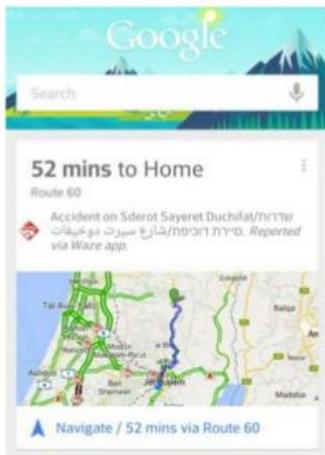


总体概况

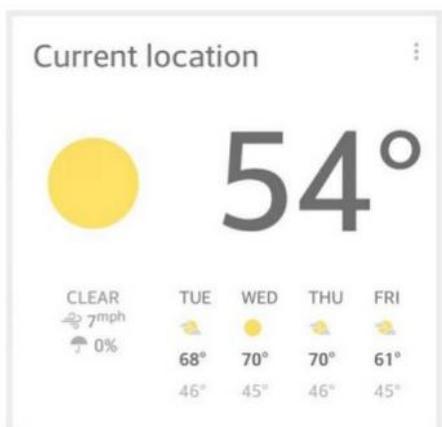


Google Assistant

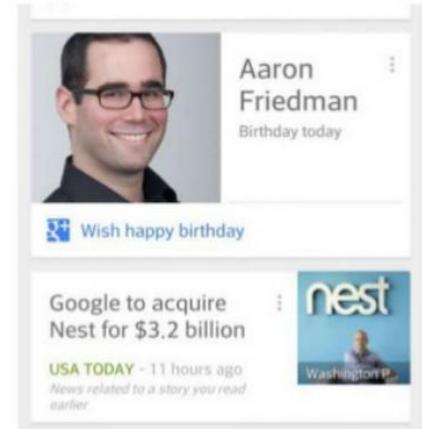
Traffic



Weather



Birthdays, articles



Sports previews & results

Heat	92
Knicks	102
Final, January 10	
TEAM	W L PCT STRK
HEAT	27 9 .750 L1
KNICKS	13 22 .371 W3
Box score	

Google Assistant

book me a flight from Paris to Rome this Saturday

Here are some flight options for you

Flights from Paris to Rome

Sat, Oct 8 – Wed, Oct 12 Sponsored

Non-stop flights 2h+

- Ryanair from €112
- Vueling from €170
- Alitalia from €175
- Air France from €178
- Iberia from €204
- British Airways from €204
- easyJet from €224

Connecting flights 3h 50m+

- Swiss from €193
- Lufthansa from €205
- Other airlines from €294

More Google flight results

+ Say something... 

show me the capital of Australia in the map

Here you go



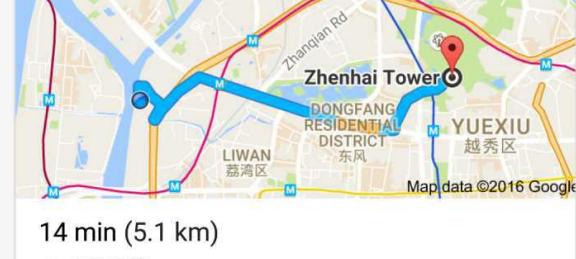
Canberra, Australian Capital Territory 2601, Australia

Search results What ocean is Australia in?

navigate to Guangzhou Museum

Ready to go

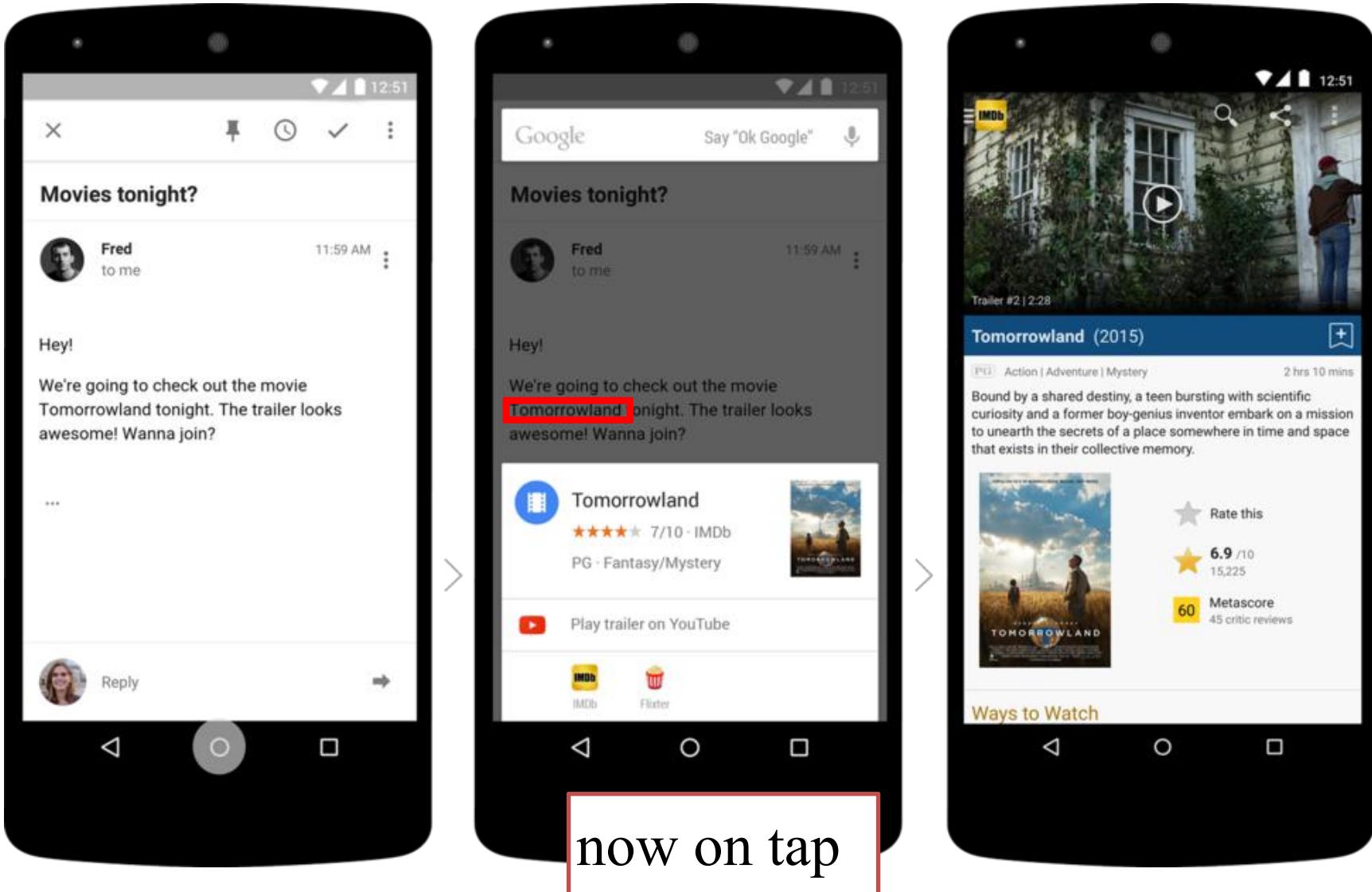
From Your location To Zhenhai Tower, Zhenhai Road, Yuexiu, Guangzhou,...



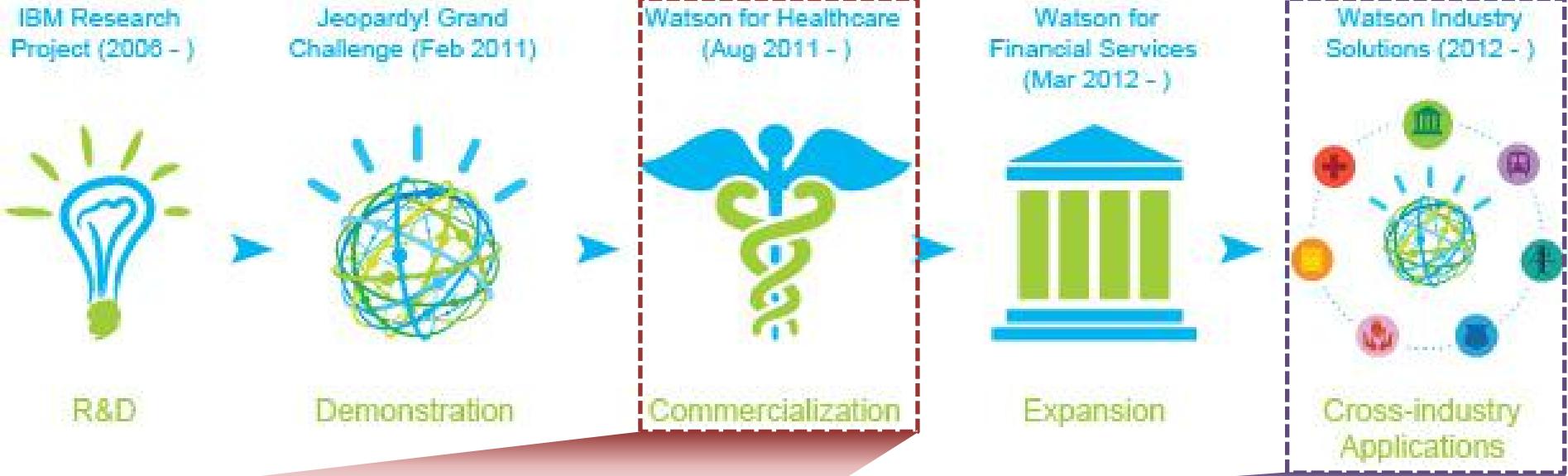
14 min (5.1 km)
via 东风西路

Search results Opening hours Walking dir 

Google Assistant



IBM Watson



A screenshot of the IBM Bluemix Watson catalog page. At the top, there's a search bar with the placeholder "Type to search". Below it, a section titled "Watson" describes building cognitive apps. A grid of 16 hexagonal icons represents different Watson services, each with a name and status (e.g., IBM, IBM BETA, Third Party). The services include:

- AlchemyAPI
- Concept Expansion
- Concept Insights
- Dialog
- Document Conversion
- Language Translation
- Natural Language Classifier
- Personality Insights
- Relationship Extraction
- Retrieve and Rank
- Speech To Text
- Text to Speech
- Tone Analyzer
- Tradeoff Analytics
- Visual Recognition
- Cognitive Commerce™
- Cognitive Graph
- Cognitive Insights™

Microsoft Cortana



User talks
to Cortana



Cortana
invokes skills

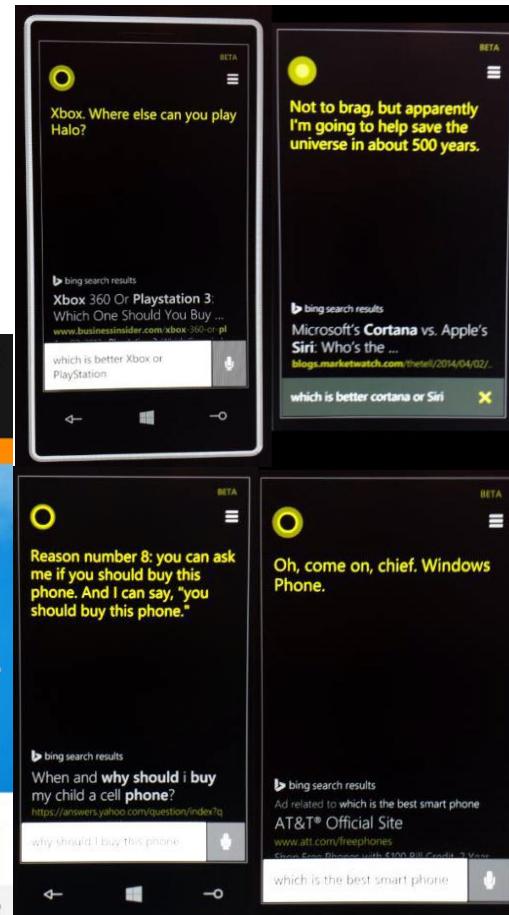
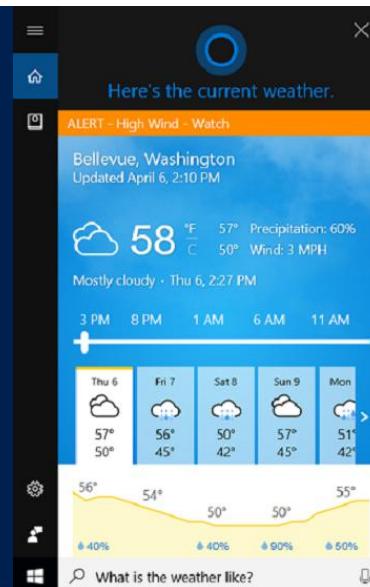
CORTANA SKILL
Cortana user profile data
Microsoft Bot Framework
Cognitive Services

Skills connect
to services



CORTANA DASHBOARD
Deployment tools
Publishing pipeline

CORTANA SKILLS KIT

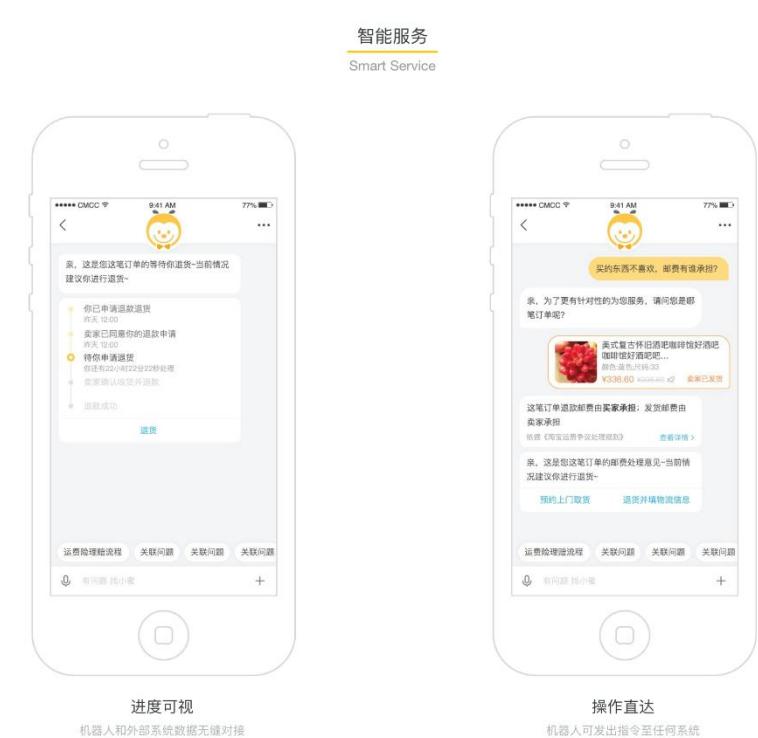


Cortana Analytics Suite

Transform data into intelligent action



阿里小蜜



电商领域的私人助理实践-阿里小蜜



| 基于阿里海量消费数据，结合线上、线下
的生活场景需求，以智能+人工的模式提供
智能导购、服务、助理的业务体验

- ## | 电商领域的平台化开放

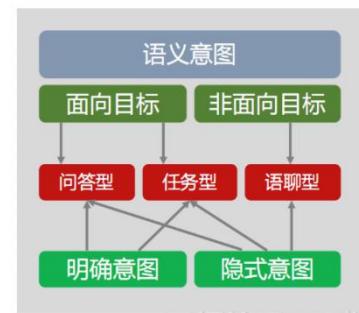


| 语义意图识别 :

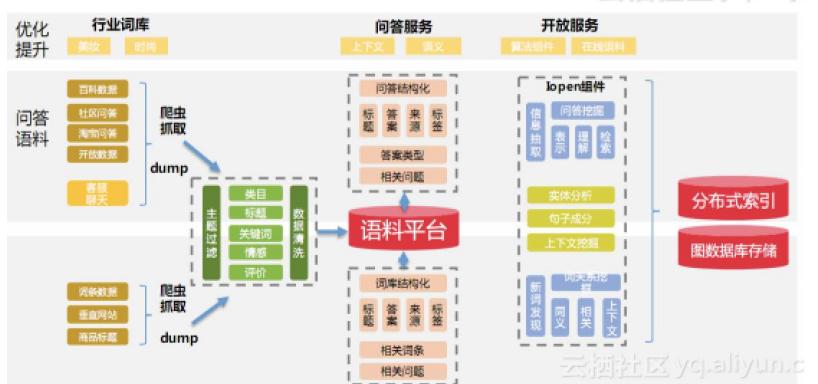
- 目标和非目标任务的细分
 - 意图的明确与推理匹配

| 面向3类业务划分并建立技术体系：

- 问答型：“密码忘记怎么办？”
 - 任务型：“我想订一张明天从北京到杭州的机票”
 - 语聊型：“我心情不好”



知识图谱构建



百度DuerOS

DUER OS

DuerOS是百度度秘事业部研发的对话式人工智能系统

是百度全球领先人工智能技术的重要应用之一，借助百度的信息与服务构成的巨大生态。DuerOS拥有海量数据，能通过自然语言完成对硬件的操作与对话交流，为用户提供完整的服务链条。作为一款开放式的操作系统，DuerOS通过云端大脑时刻进行自动学习让机器具备人类的语言能力。



- 娱乐影音
 - 信息查询
 - 知识教育
 - 生活服务
 - 出行路况
 - 实用工具
 - 个人助手
 - 指令控制
 - 智能家居
 - 聊天休闲
- 音乐、电影、电视剧、综艺、短视频、有声书、电台等功能
- 咨询、天气、股票、航班、体育（NBA）、问答、菜谱、图片等功能
- 百科、故事、儿歌、成语、育儿、诗词、文库等功能
- 美食、电影、外卖、酒店、购物、打车、保洁、旅游、充值多种生活O2O服务
- 地图、路线、路况、限行、违章、周边查询等功能
- 翻译、时间、计算、汇率、单位换算等功能
- 日程管理、闹钟、提醒、备忘、记事等功能
- 打电话、发短信、设置闹钟、通讯录查询、应用调起等功能
- 指令控制、状态查询、使用手册、客服电话等功能
- 聊天、笑话、藏头诗、成语接龙、游戏等功能

APPLE SIRI

Explore some of the things you can ask Siri



The Basics



What's New



Staying in Touch



Getting Organized



Sports



Entertainment



Out and About



HomeKit



Tips and Tricks



Getting Answers



"What does the fox say?"



"How do you spell supercalifragilisticexpialidocious?"



"How high is El Capitan?"



"What's an 18% tip on \$85?"



"Did the groundhog see its shadow?"

"When is sunset in Paris?"



■ intelligent service



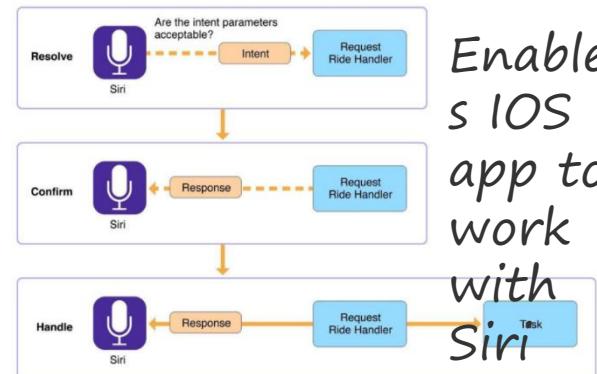
Web Services and APIs



Domain & Task Models

Guided Dialog

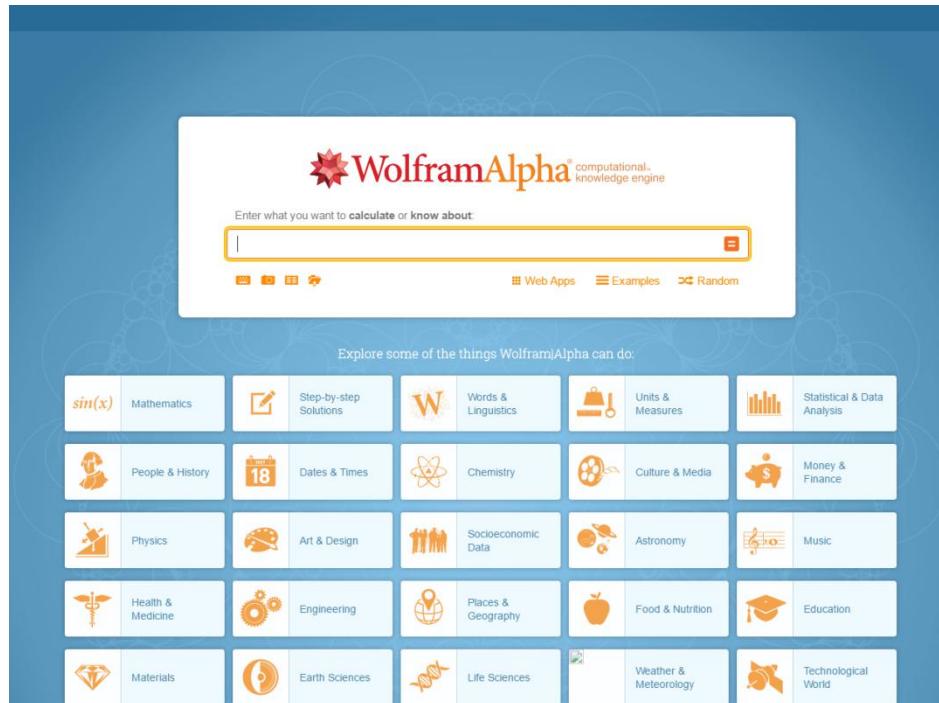
■ system architecture



■ Sirikit

WolframAlpha

WolframAlpha是开发计算数学应用软件的沃尔夫勒姆研究公司开发出的新一代的搜索引擎，能根据问题直接给出答案的网站，用户在搜索框键入需要查询的问题后，该搜索引擎将直接向用户返回答案，而不是返回一大堆网页链接。

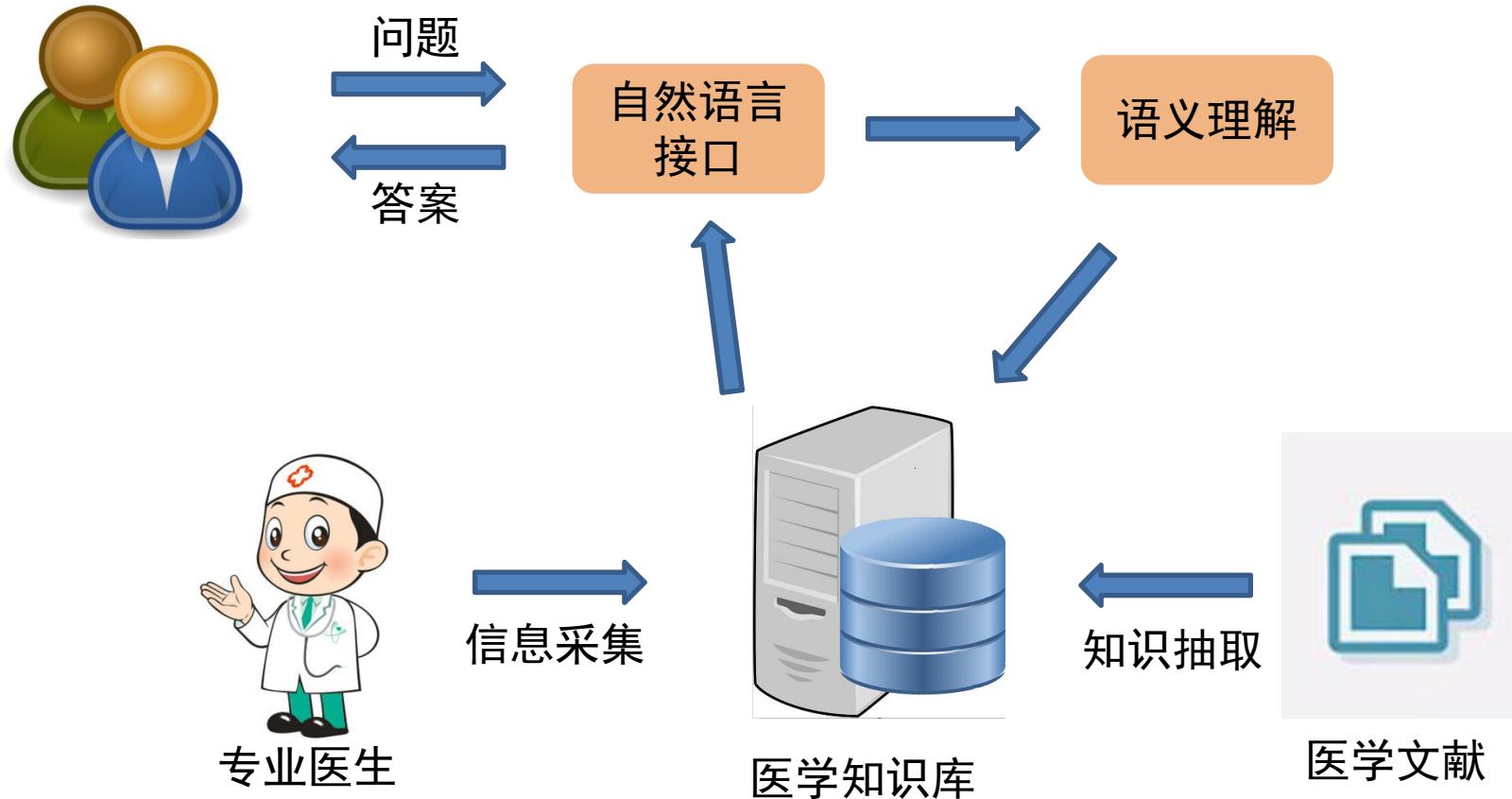


医疗：医学知识查询

- 无需像操作数据库一样去机械化地查找数据
- 使用自然的语言和查询系统去交流
- 查询系统在接收到用户问题后
- 进行自我分析与数据检索，并返回给用户最贴切的答案。

The image shows a medical consultation interface. It features two interactions between a doctor (represented by a blue icon with a white stethoscope) and a patient (represented by a black silhouette with a white cross).
1. Top Interaction:
Patient: 白癜风能根治吗?
Doctor: 白癜风可以治疗，但不是每个方案对每个患者都有效，有治愈的案例，但是具体到每个患者，则不能预期是否能根治。
2. Bottom Interaction:
Patient: 得了肝炎一定会慢化吗
Doctor: 也不是的。肝炎是否慢性化和感染的病毒类型、感染年龄、个人体质等密切相关。一般来说，甲肝和戊肝感染来的快去的也快，不易慢性化。常说的慢性肝炎主要是指病毒感染事件超过半年的乙肝。大约每20个成年乙肝患者.....

医疗：医学知识查询



相关企业



以搜索引擎的检索方式，为用户提供医疗信息查询服务，**用户不仅可以查到疾病、症状与药品的相关信息，还可以检索到相似病例与解决方法，供用户参考。**



专注于提供医疗数据服务，并建立云端的随访系统，为患者提供疾病管理服务



在业务数据、结构化数据、非结构化数据、业务数据的基础上，建立大数据池。在海量数据的基础上为用户提供信息查询服务。

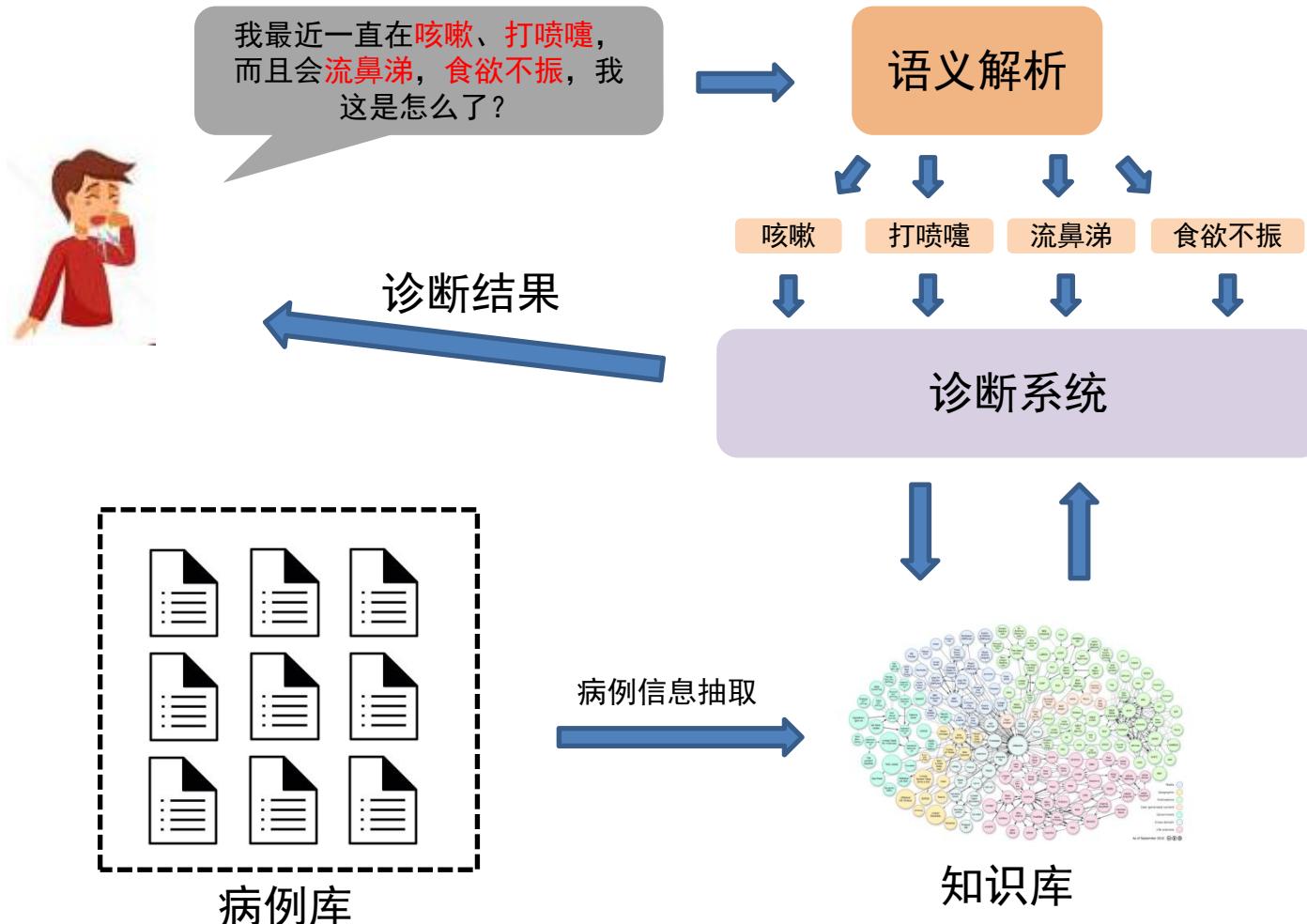


医疗：智能问诊

用户可以向机器人说明自己的症状，机器人会根据用户的描述，判断用户所患疾病，并根据判断，对用户之后需要如何诊断与治疗做出相应的推荐。



医疗：智能问诊



Airdoc



相关企业



以人为核心的智慧病历矩阵；以疾病为核心的临床科研数据矩阵；**多源异构医疗数据处理技术**；大数据科研辅助分析引擎开发。



利用深度学习技术，**在千万级医学图像数据及专家标注下进行训练**，建立模型，实现特定病变组织的检测、组织分割、影像识别分类等多种医学计算机视觉技术。

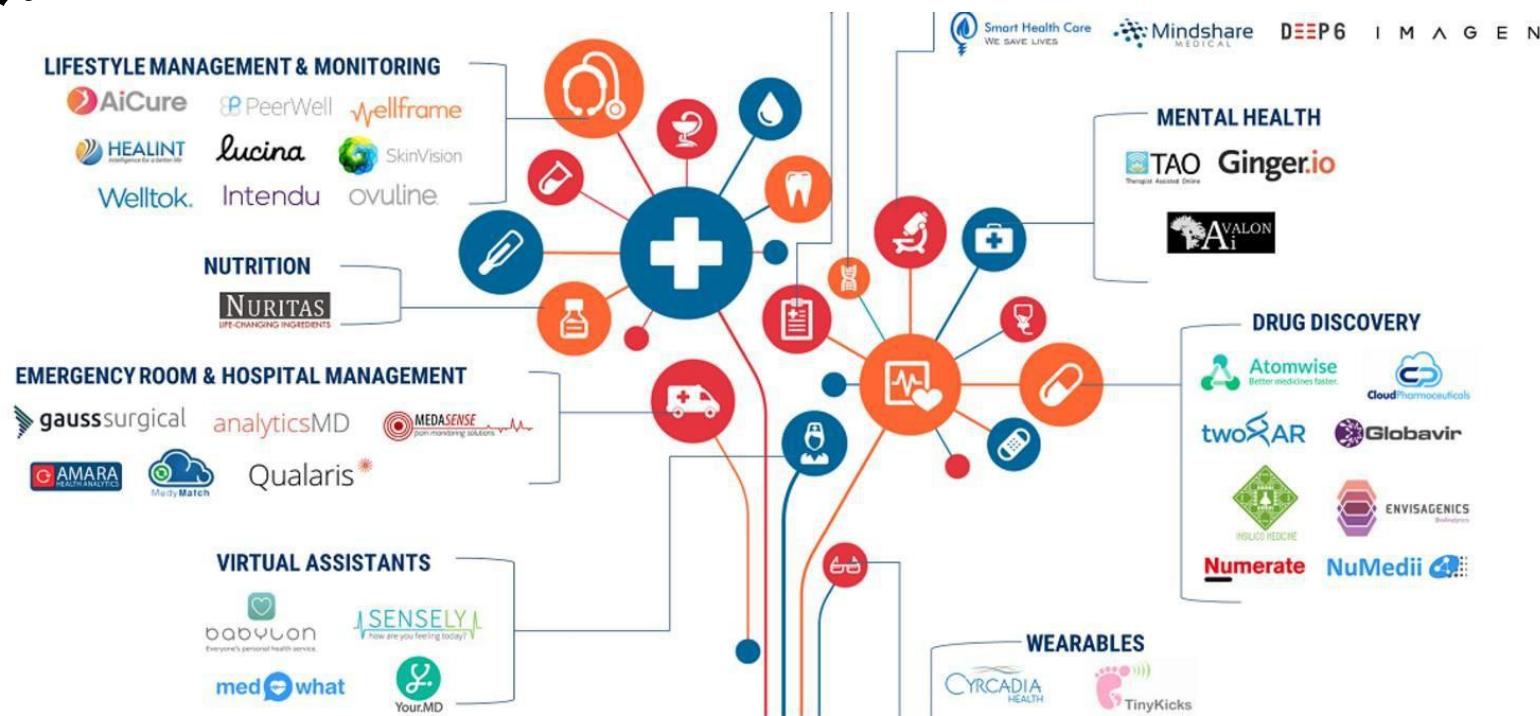


推出了智能问诊机器人，它是个健康管家。向患者提供疾病自诊、导诊、挂号、问药、健康咨询。

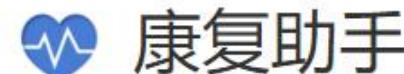
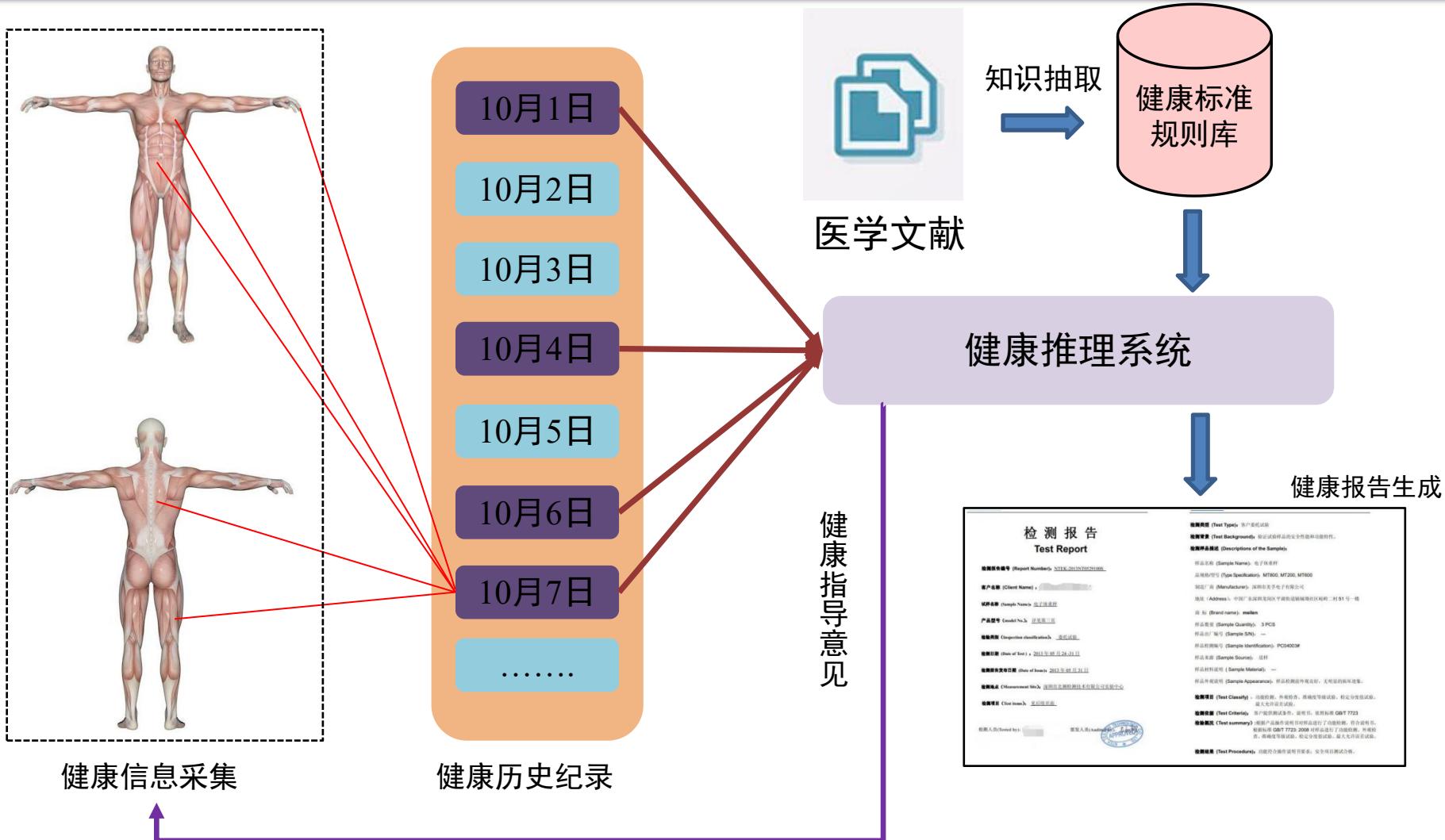


医疗：健康管理

以个人助手的形式，随时记录用户的健康状态，并根据与用户的交流，判断用户最近健康走势。综合已有健康数据，对用户的健康状态做出一个判断，并给出相应建议。



医疗：健康管理



相关企业



深耕“健康管理”市场，进行行业的融合创新，建立“精准健康管理”平台，**集成健康配餐服务，计划定制，评测与报告。**



其**健康风险基因检测系统**可以预测疾病风险，通过检测各种癌症、心脑血管等高发疾病相关的易感基因，预测自身健康风险，做到早发现、早预防、早干预，防患未然，成就健康人生。



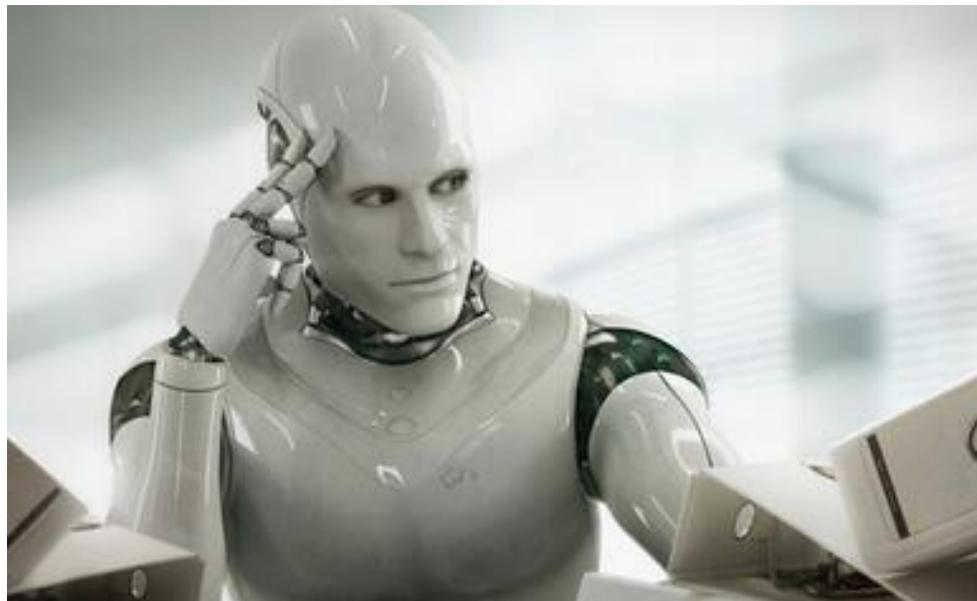
康复助手

提供全面的治疗指南，**针对性的康复提醒，人性化的康复关怀**。根据用户健康状况，与相仿的疾病案例，供用户自我学习，早日康复。

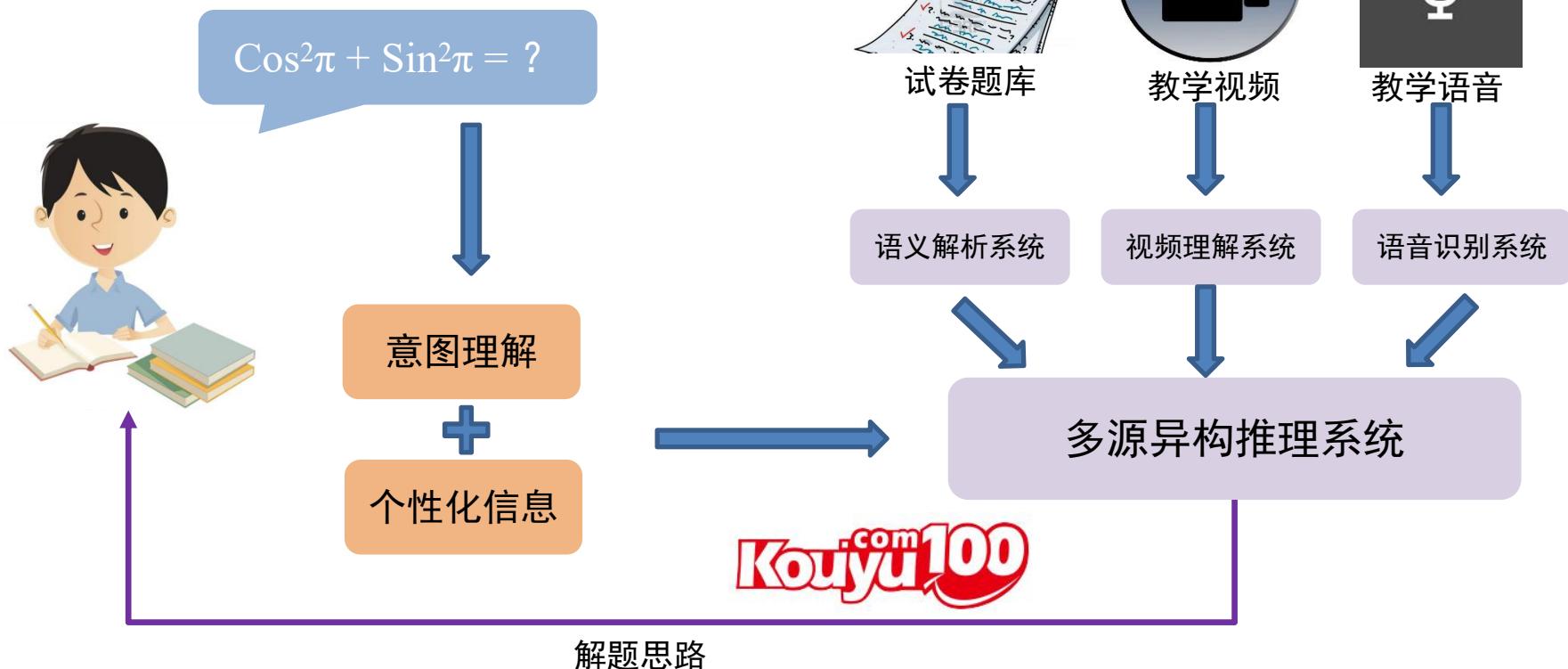


教育：个性化作业辅导

问答系统在理解学生所要学习内容后（文本、图片、音频、视频），学生可以向其提出各种问题。系统可以根据学生过往的学习习惯，制定相关策略，一步一步帮助学生解决问题，辅助学生找到问题的答案。



教育：个性化作业辅导



数理化学习用辅导团



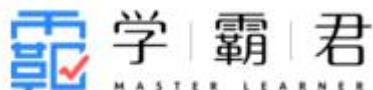
学 | 霸 | 君
MASTER LEARNER

作业帮 | 一课
见证进步每一课

相关企业



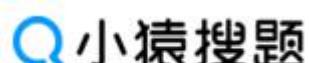
从“家庭作业”切入，解决“家庭作业”这一高频且刚需的任务。**家长可以将小孩子作业拍照上传**，辅导团实时响应，在15分钟内完成对每道题的检查批改，并给出每道错题的详细解析。



基于数据积累和对人工智能技术，**利用自然语言理解和深度学习为核心技术**，实现了将人工智能技术引入到教育事业中，升级传统的教育模式，开启教育事业创新与变革的大门，或将产生强劲助力。

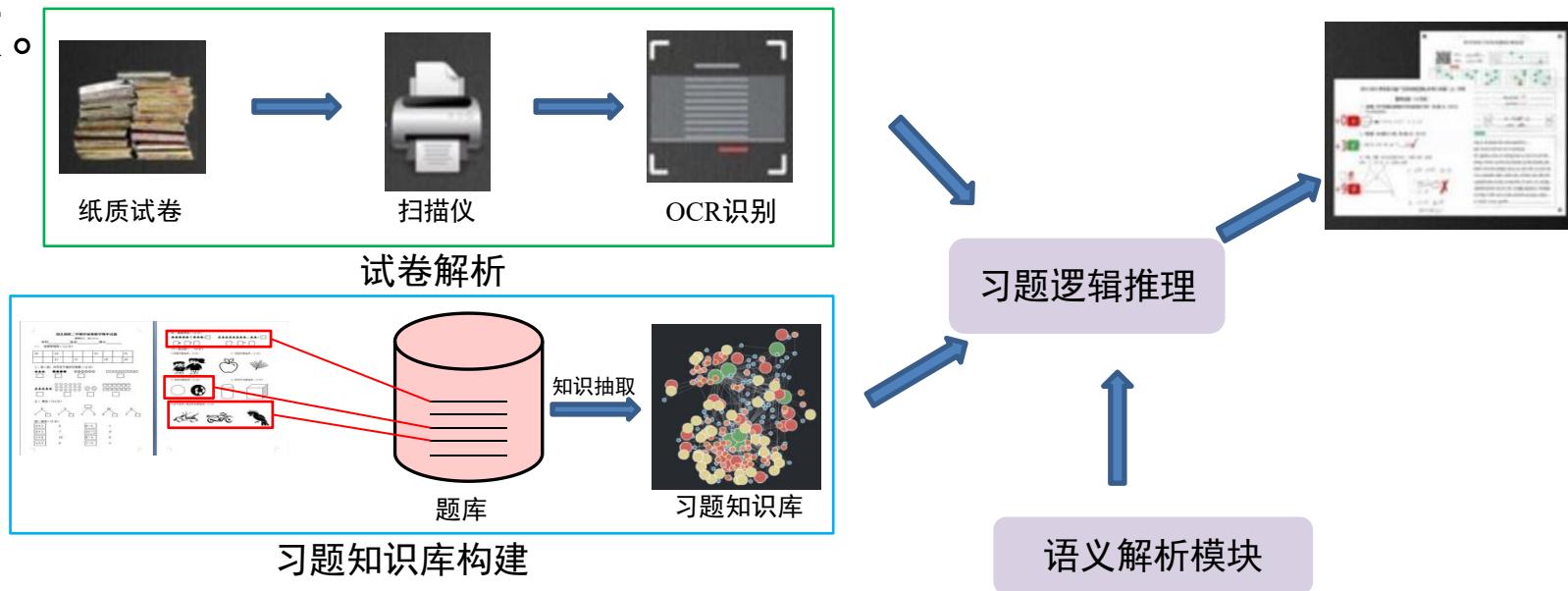


作业帮是百度知道团队内部打造的面向K12的问答学习平台，除了拍题、语音搜题等功能，还**集合日产生百万条数据的UGC**，并在答疑后端接入包括学大、好未来等视频内容2万余条。



教育：试卷批改

不同于现在已有的试卷批改系统，基于问答的试卷批改系统不仅可以批改客观题，更善于批改主观题。批改试卷的过程中，系统会理解学生的主观表达，从中凝练答案主旨，和标准答案进行相似度匹配，并给出相应的分數。



相关企业



主要由数据扫描系统、OMR自动评分系统、网上评卷系统、数据监控系统构成，以计算机网络技术和图像处理技术为依托，采用专业扫描阅读设备，对各类考试答卷和文档进行扫描和处理，实现了客观题机器自动评卷和主观题老师网上评卷相结合的高效评卷模式。

研究范围包括语音，图像，自然语言，掌握了语音识别，图像识别，自然语言理解，作文评价，人机对话等多项人工智能技术。

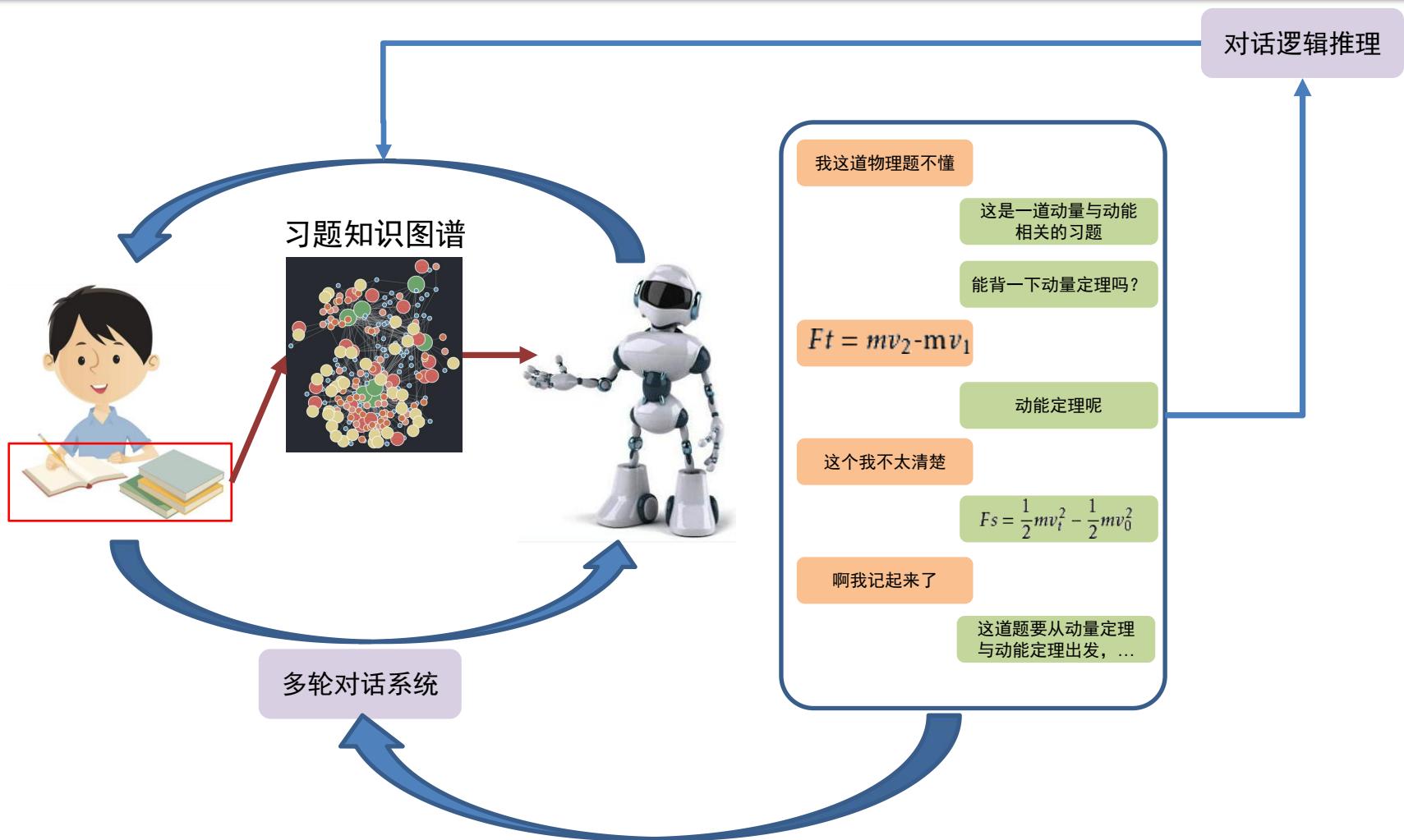
使用图像识别技术，能够自动批改学生作业

教育：教育机器人

教育机器人可以主动与学生接触，通过个性化方式传授给学生所需知识，并在此期间会记录学生的行为反馈与意识反馈，根据反馈结果做出下一步判断，比如是继续传授学生新的知识还是让学生巩固已学过的知识。



教育：教育机器人



相关企业



开发的**智能摄像头机器人**，支持**一对一英语对话**，**百科知识讲解**，**讲故事**等等育儿功能



依托小i机器人的智能硬件云平台及垂直领域专业知识库，具有**精准语义分析能力**，结合**语音识别技术**、**语音合成技术**、**人脸识别技术**及**声源定位技术**等多种先进技术



陪伴孩子讲故事、说笑话、聊心声...沟通交流情感，做到能够理解复杂的语义并且配合生动的表情。丰富孩子的知识学习，同步小学英语，纠正孩子发音标准，学练考让孩子英语成绩飞速提高

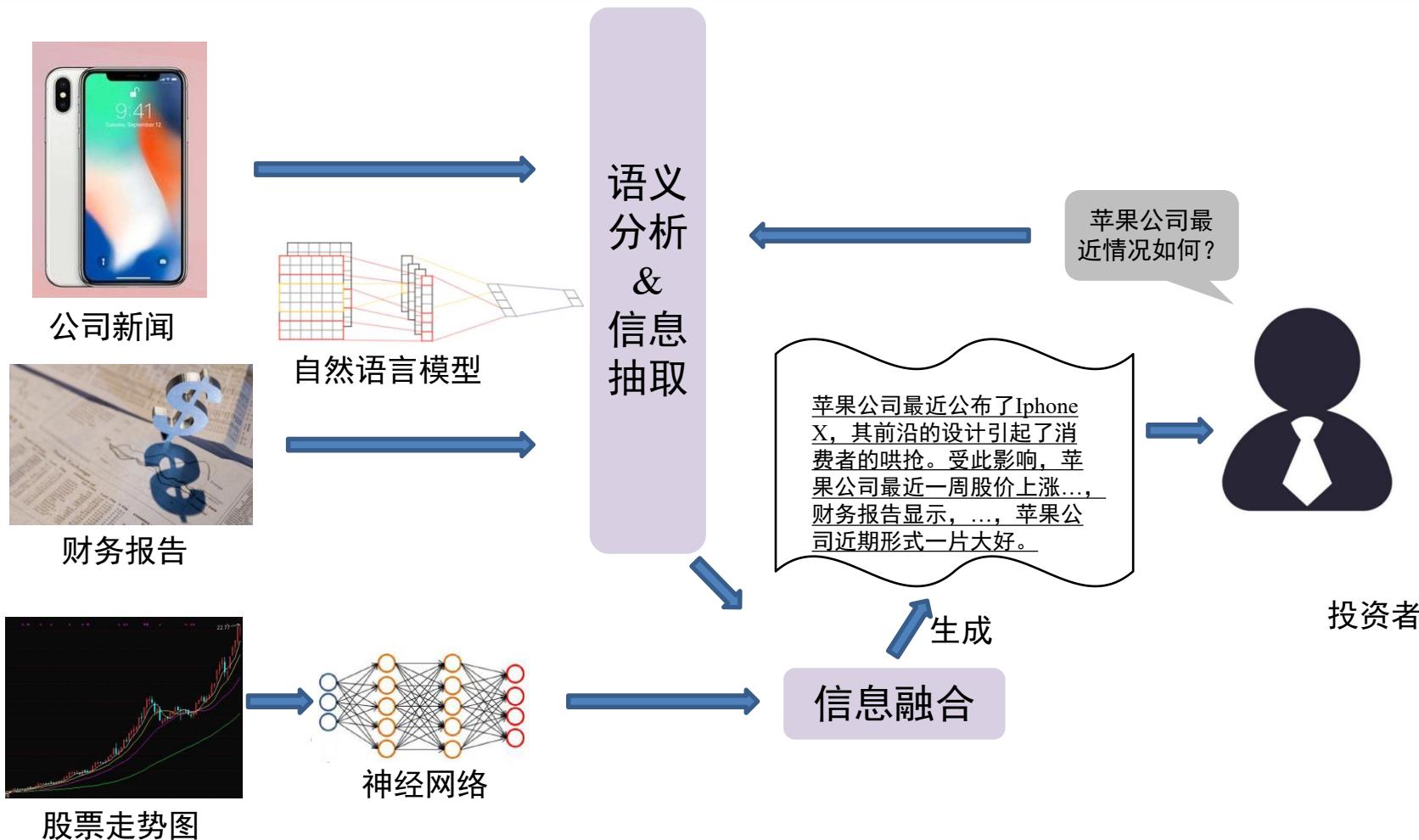


金融：金融信息咨询

用户可以通过日常交流的方式，向金融问答系统提问，比如可以询问某只股票最近的走势、某家公司公布的财务情况等。系统会通过语义分析，理解用户问题，生成准确答案，反馈给用户。



金融：金融信息咨询



相关企业



使用命名实体、关系检测等自然语言处理技术，通过对**企业研报与大盘的分析**，得到**企业分析报告**。



文因互联是一家智能金融技术与服务提供商，主要提供**自动化公告摘要、自动化研报摘要、自动化报告写作、金融查询机器人、金融搜索**等智能金融核心工具。



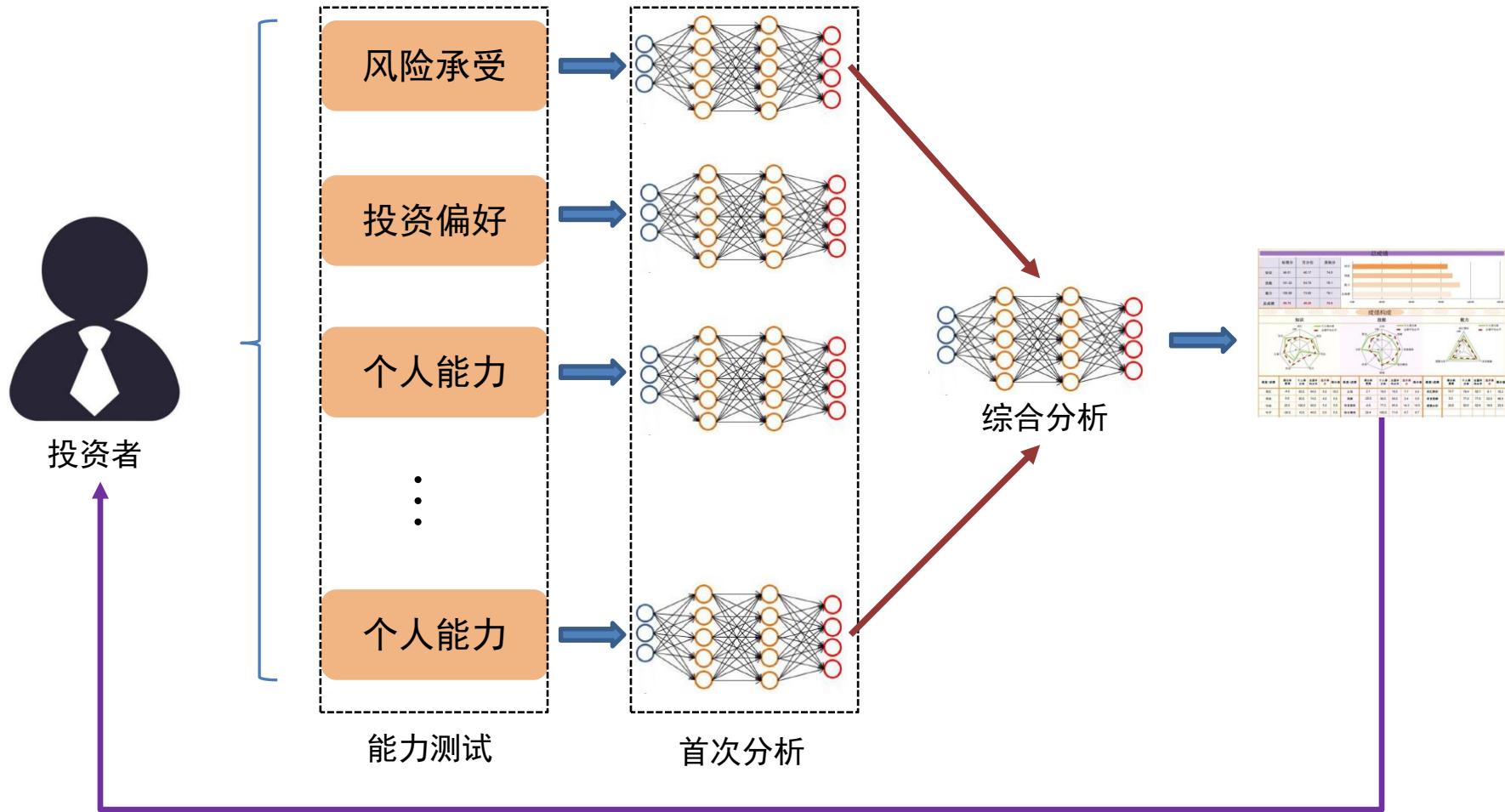
建立关系图谱，直观呈现产品维度的竞品分析、公司参控股关系、投资机构参控股关系及投资业绩行业大数据；挖掘公司的投资价值、投资机构的业绩回报，洞察行业潜力；拥有覆盖7亿用户的行为数据，精准描绘用户画像以及3000万企业画像

金融：智能投资顾问

用户可以向该机器人说明自己的需求，机器人在对用户进行风险承受能力的评估后，会根据用户的需求，向用户推荐理想的股票、基金等金融投资产品。



金融：智能投资顾问



相关企业



通过投资者在页面中回答几个简单的问题，智能化地为投资者制定出组合投资方案，并从多个维度对京东金融上的基金产品进行排名，为用户挑选基金产品提供参考。



全国首家运用人工智能进行投资管理的公司，其开发的第一代证券投资人工智能系统(SIAI)，每天能够自主地进行投资分析，并生产出数量不等的投资策略，进而进场交易，在投资期内，SIAI 系统能够实时地根据市场的信号进行调仓和买卖。



在蚂蚁金服的业务场景下运用人口智能技术进行资产配置、财经信息分析等一系列的创新和应用。



法律：法律咨询机器人

不同于传统的关键词匹配技术，法律咨询机器人会根据用户描述，将用户所希望得到的法律法规返回给用户，同时，机器人还会帮助用户解读法律条文，并会根据用户的描述与这些法律法规给出用户需求合理性的判断。

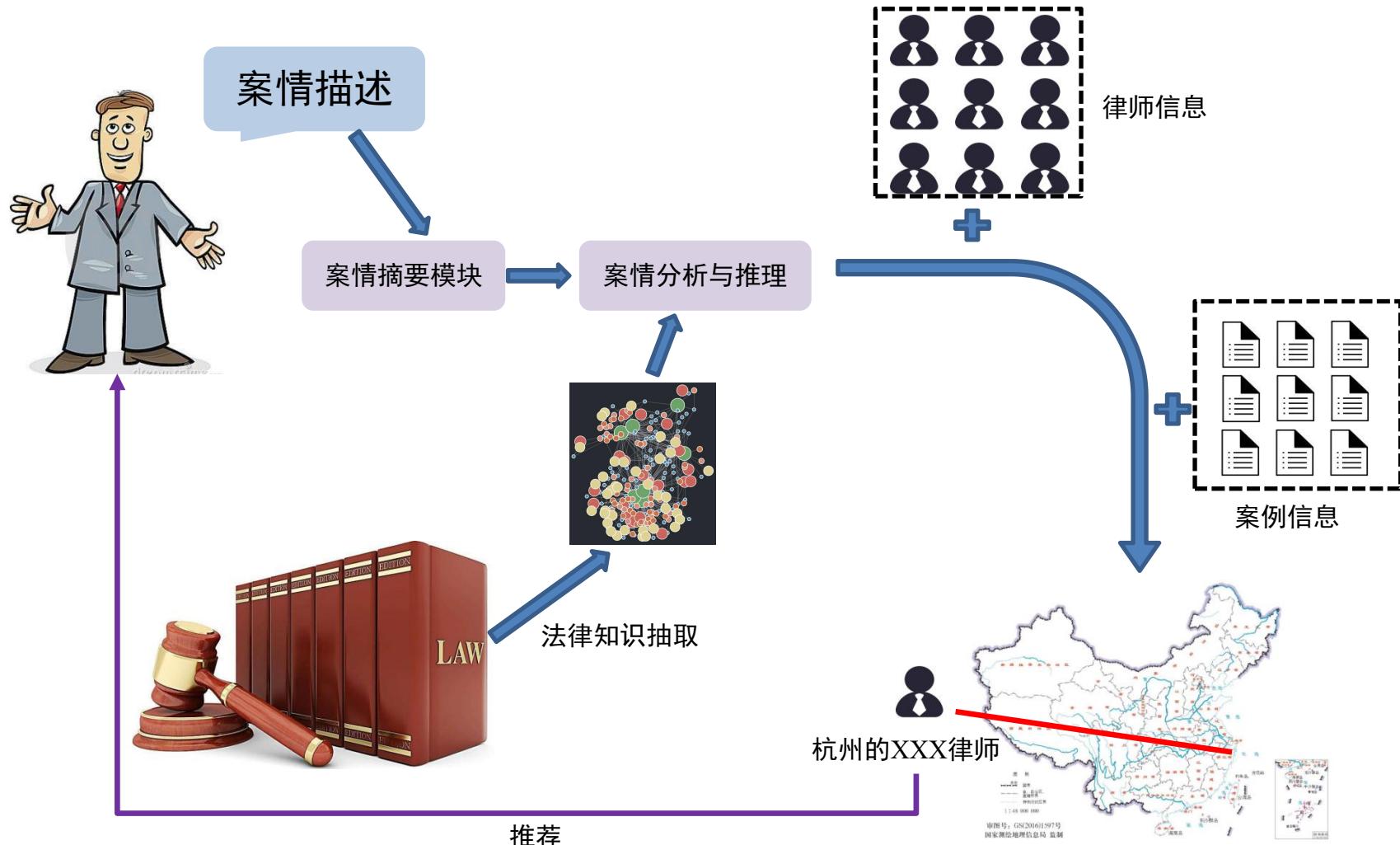


法律：律师推荐

问答系统通过对上千万份裁判文书的理解与匹配，结合用户的自身情况，推荐给用户合适的律师与有优势的诉讼地点。



法律：律师推荐



相关企业



法律谷是iPIN团队孵化的智能法律服务产品，在此平台上只需输入案情描述，便能搜索、查看到过往法院受审、判定的相关案例与案例的代理律师。



数据驱动智能法律服务平台，目前专注解决女性离婚用户的法律问题。法里通过大数据和人工智能高效地连接法律服务，同时用人工智能解决基础的法律服务，从而高效地为律师筛选案源，和律师是良好的合作伙伴而非竞争对手。



将律师与案例数据关联匹配，提高律师互联网影响力。无讼案例是目前中文世界最精确、易用的案例检索工具。其产品“法小淘”通过与众不同的算法以及独一无二的关键词系统，将最精准的检索结果前置，一举突破了传统案例检索工的繁琐。

智能家居

用户可以通过语音和家中的智能设备进行交互，智能设备在理解用户语言后，识别用户意图，并做出相应的反应。



LifeSmart 云起智能



相关企业



从衣食住行，围绕生活方方面面，并以手机与互联网为中心，构造智能家居生态。小米智能家居可以通过手机APP或者互联网进行控制。

LifeSmart 云起智能

通过物联网和人工智能的结合为用户提供安全、舒适、节能以及多姿多彩的家居生活体验，以时尚的智能电子产品改变人们的生活方式。



专注并拥有行业领先语音识别及语言处理技术。所有智能家具产品均可以通过语音交互与控制。



ORVIBO

技术难点



语义理解



知识提取凝练



人机对话



多元异构数据推理

当前问答系统面临的四大问题

仅可以识别结构简单的问题，在理解复杂的长文本问题方面，效果不理想

如何向法院申请限制令？

我借了XXX钱给XXX，但他
一直不还，现在欠条也丢
了…



当前问答系统面临的四大问题

在推理方面的能力还是比较孱弱，无法深层的发掘问题中不同对象的关系，并在这种关系上做出合理的推理与解释。

感冒了应该吃什么药？

我头疼，还有些发烧，而且一直在流鼻涕，精神不振，我得了什么病？



当前问答系统面临的四大问题

无法和用户之间形成自然的多次交流。大多数情况下，只能根据用户的上一次问题，做出相应的回答

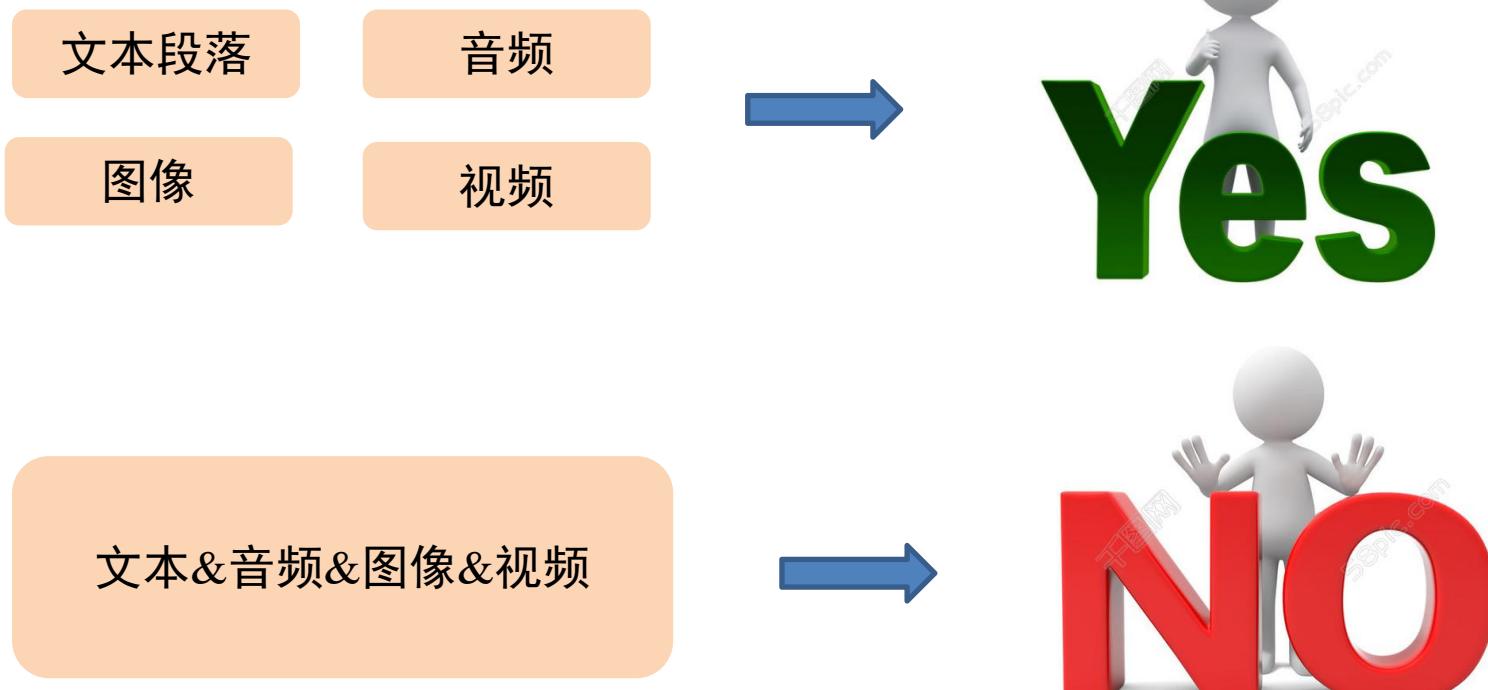
问 2.如果提供合格的样品，可以进行免费的检测与定做吗？
答 可以的；由客户提供合格的样品，我们可以为样品做一个材料体测然后根据客户实际要求定做合格的原料产品。

我想知道约好的Party在几点
你有三个约会，分别是和Alex、你姐姐和Jeff，你指的是哪一个？
我想知道Alex邀请我去的那个
约定的时间在下午4点
谢啦
不客气



当前问答系统面临的四大问题

只能在单一的媒介（文本、音频、图像、视频）中完成问答，无法在不同媒介之间建立关系并完成问答任务。



四大难题



语义理解



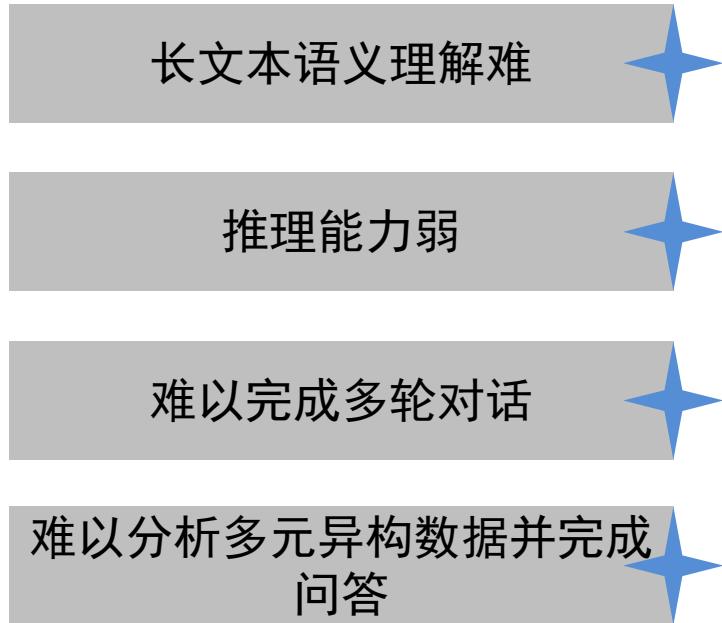
知识提取凝练



人机对话



多元异构数据推理



三、核心问题和关键技术

核心问题

- 机器阅读理解
- 知识库构建与问答
- 视觉理解与问答
- 场景理解与多轮对话

机器阅读理解

- 机器阅读理解与人阅读理解面临问题类似。
- 给定一篇文章，针对文章提出一些问题，机器回答这些问题来证明理解文章所传达的主旨内容。

Towards the Machine Comprehension of Text: An Essay

Christopher J.C. Burges
Microsoft Research
One Microsoft Way
Redmond, WA 98052, USA

December 23, 2013



Facebook的推理理解任务

■ 给定系列事实：

- ◆ John picked up the apple.
- ◆ John went to the office
- ◆ John went to the kitchen.
- ◆ John dropped the apple.

■ 提出问题：

- ◆ Where was the apple before the kitchen?

■ 答案推理：

- ◆ office

DeepMind的完型填空任务

- 给出一篇新闻和一个问题，把问题中某个实体抽掉，预测被抽掉的实体。
- 该数据来自真实新闻数据，9万篇CNN和22万篇DailyMail新闻。

Original Version	Anonymised Version
Context <p>The BBC producer allegedly struck by Jeremy Clarkson will not press charges against the "Top Gear" host, his lawyer said Friday. Clarkson, who hosted one of the most-watched television shows in the world, was dropped by the BBC Wednesday after an internal investigation by the British broadcaster found he had subjected producer Oisin Tymon "to an unprovoked physical and verbal attack." ...</p>	<p>the <i>ent381</i> producer allegedly struck by <i>ent212</i> will not press charges against the "<i>ent153</i>" host , his lawyer said friday . <i>ent212</i> , who hosted one of the most - watched television shows in the world , was dropped by the <i>ent381</i> wednesday after an internal investigation by the <i>ent180</i> broadcaster found he had subjected producer <i>ent193</i> " to an unprovoked physical and verbal attack . " ...</p>
Query <p>Producer X will not press charges against Jeremy Clarkson, his lawyer says.</p>	<p>producer X will not press charges against <i>ent212</i> , his lawyer says .</p>
Answer <p>Oisin Tymon</p>	<p><i>ent193</i></p>

Stanford的SQuAD挑战任务

- SQuAD是斯坦福发布的英文可变长答案数据集。
- 内容主要以维基为主，包含2万个小段落和10万个问题。

给定文档

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under **gravity**. The main forms of precipitation include drizzle, rain, sleet, snow, **graupel** and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals **within a cloud**. Short, intense periods of rain in scattered locations are called "showers".

答案推理

What causes precipitation to fall?
gravity

What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?
graupel

Where do water droplets collide with ice crystals to form precipitation?
within a cloud

SQuAD: 100,000+ Questions for Machine Comprehension of Text

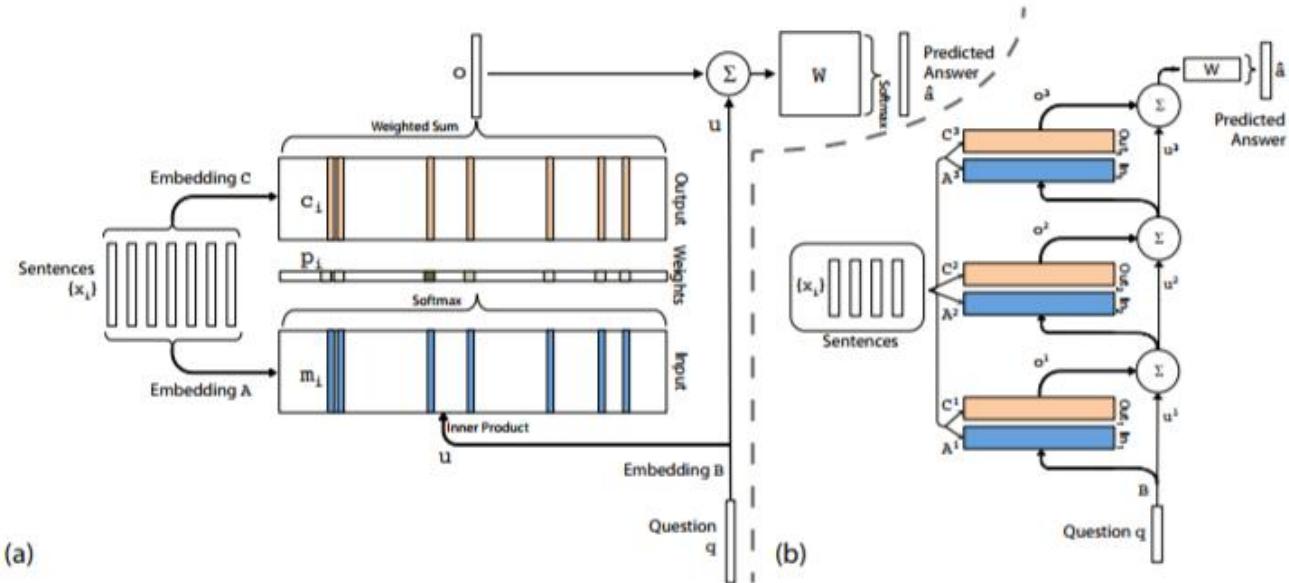
机器阅读技术

■ 阅读理解技术通常可以分为四部份：

1. 阅读整篇文章，初步理解再审题
2. 问题和文本中部份段落和内容做关联
3. 候选答案与问题进行对应后，进行证据融合辅证答案正确性
4. 针对挑选出的候选答案进行精晒，得到正确答案

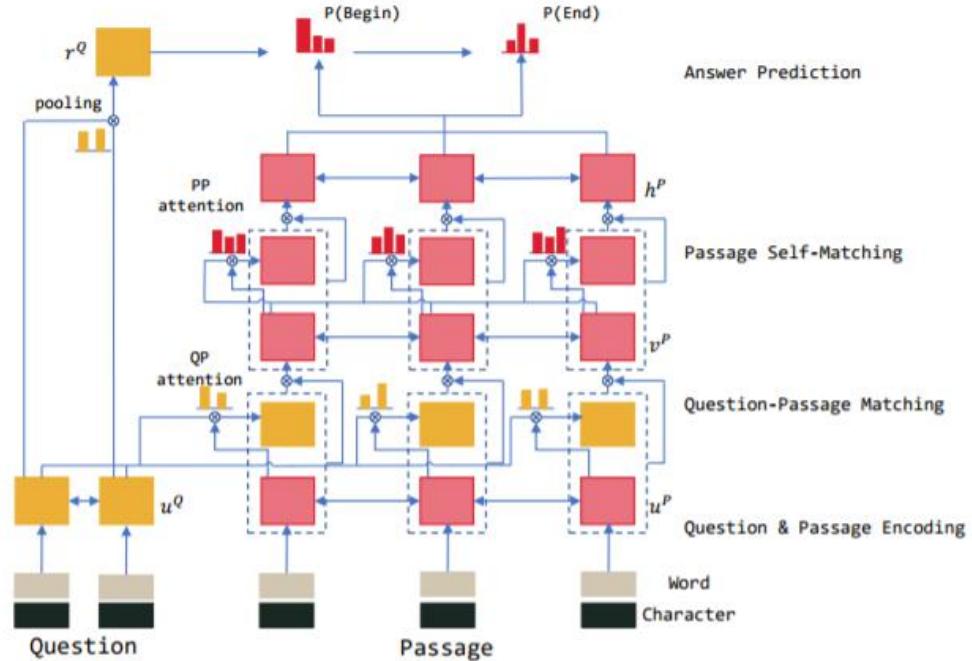
Facebook的记忆网络

- 解决循环神经网络记忆力随时间增长而消退的问题。
- 根据Memory中的知识进行分析推理。



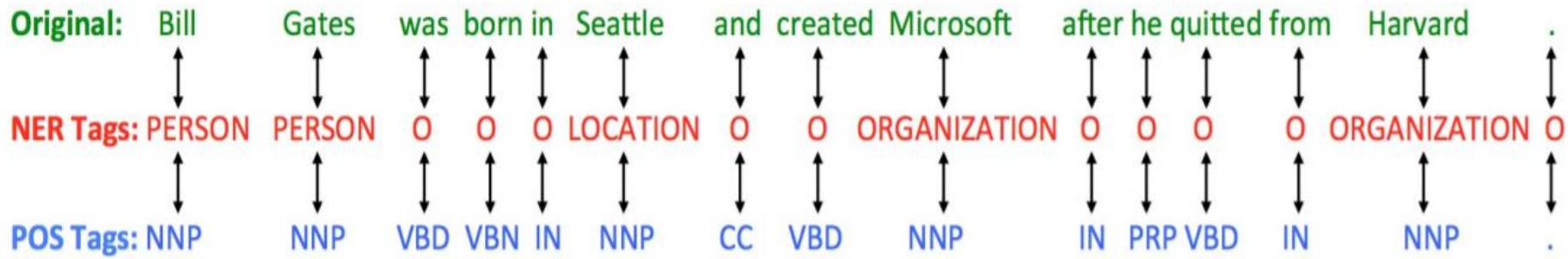
Microsoft的R-Net

- 双通道词向量特征
- Self-Attention机制
- 在斯坦福阅读问答比赛中摘得**榜首**



R-Net: Machine Reading Comprehension with Self-Matching Networks

我们的MEMEN网络

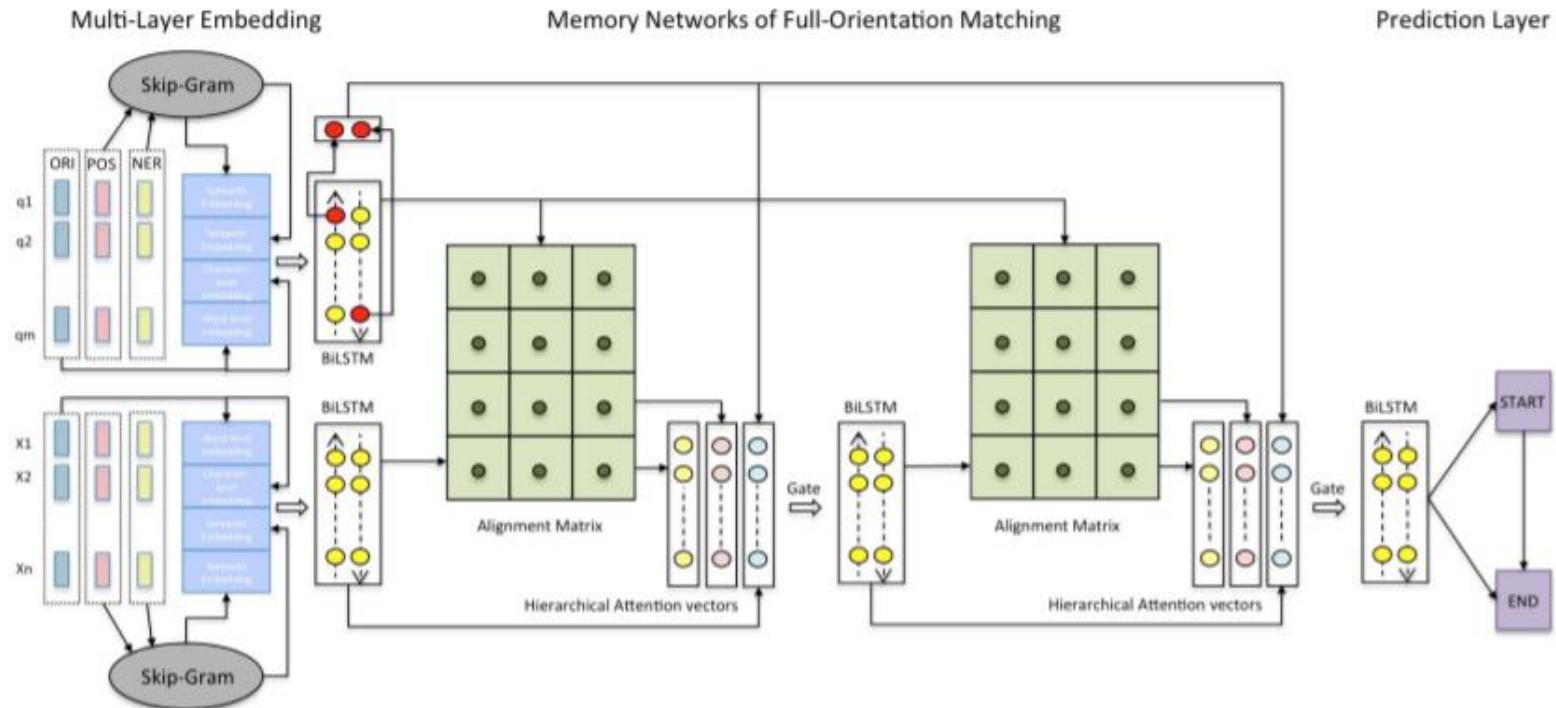


■ 对阅读文本进行语义分析：

- ◆ 识别文本中出现的实体
- ◆ 词性分析

■ 根据以上内容，学习一个语义空间，将原始文本映射在该空间中，并使用特征向量表达文本内容。

我们的MEMEN网络



- 使用BiLSTM为问题与阅读文本编码
- 使用记忆网络和注意力机制匹配问题与阅读文本内容
- 使用Pointer Network定位问题答案在原文中的位置

SQuAD挑战评测结果

Rank	Model	EM	F1
1 Mar 2017	r-net (ensemble) Microsoft Research Asia http://aka.ms/rnet	76.922	84.006
2 May 2017	MEMEN (ensemble) Eigen Technology & Zhejiang University	75.37	82.658
3 Mar 2017	ReasoNet (ensemble) MSR Redmond	75.034	82.552
4 May 2017	r-net (single model) Microsoft Research Asia http://aka.ms/rnet	74.614	82.458
5 May 2017	Mnemonic Reader (ensemble) NUDT & Fudan University https://arxiv.org/abs/1705.02798	73.754	81.863
6 Apr 2017	SEDT+BiDAF (ensemble) CMU https://arxiv.org/abs/1703.00572	73.723	81.53

- 微软
- Google
- Salesforce
- IBM
- Facebook
- 科大讯飞
- 卡内基梅隆大学
- 新加坡国立大学
- 清华大学
- 北京大学
- 复旦大学

机器阅读意义

■ 机器阅读是语义理解的经典问题

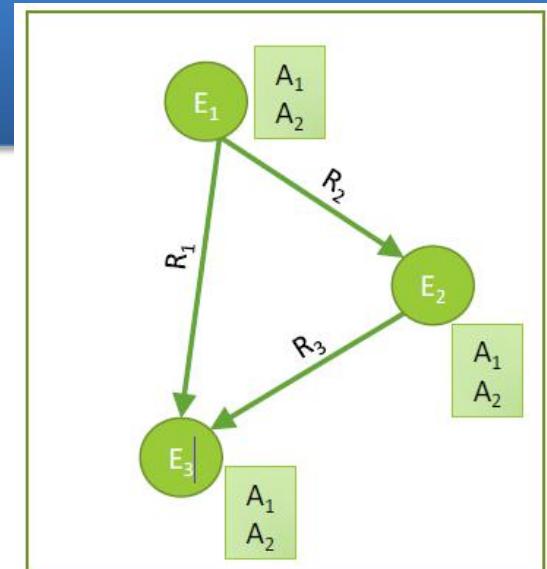
- ◆ 精准问答
- ◆ 短句分析

■ 落地应用

- ◆ 司法领域的电子卷宗
- ◆ 医疗领域的个人病例
- ◆ 电子产品说明书

知识库定义

- 知识表示为图的形式
 - ◆ 节点：实体
 - ◆ 边：实体之间关系



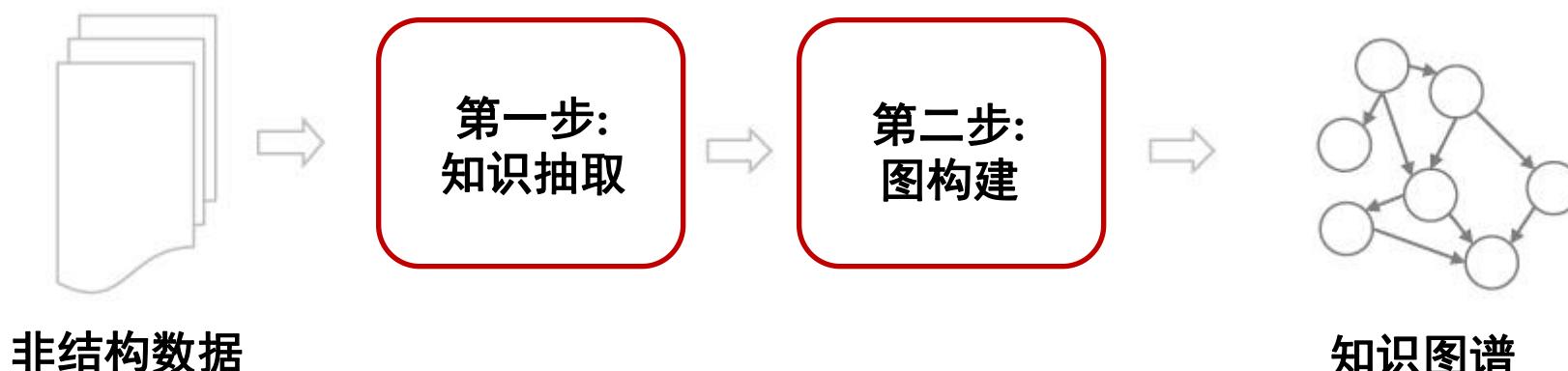
知识库构建

■ 知识抽取

- ◆ 非结构化文本
- ◆ 结构化文本

■ 图构建

- ◆ 概率模型
- ◆ 网络嵌入



知识图谱表达

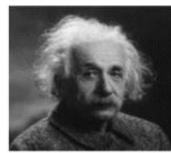
RDF 表达

```
<rdf:Description  
rdf:about="http://www.recshop.fake/cd/Empire Burlesque">  
  <cd:artist>Bob Dylan</cd:artist>  
  <cd:country>USA</cd:country>  
  <cd:company>Columbia</cd:company>  
  <cd:price>10.90</cd:price>  
  <cd:year>1985</cd:year>  
</rdf:Description>  
  
<rdf:Description  
rdf:about="http://www.recshop.fake/cd/Hide your heart">  
  <cd:artist>Bonnie Tyler</cd:artist>  
  <cd:country>UK</cd:country>  
  <cd:company>CBS Records</cd:company>  
  <cd:price>9.90</cd:price>  
  <cd:year>1988</cd:year>  
</rdf:Description>
```

三元组表达

Empire_Burlesque	artist	Bob_Dylan
Empire_Burlesque	country	USA
Empire_Burlesque	company	Clumbia
Empire_Burlesque	price	10.90
Empire_Burlesque	year	1985
Hide_your_hert	artist	Bonnie_Tyler
Hide_your_hert	country	UK
Hide_your_hert	company	CBS_Records
Hide_your_hert	price	9.90
Hide_your_hert	year	1988

</m/0jcx, /m/04m8, /m/019xz9>



/en/albert_einstein

Albert Einstein



/en/ulm

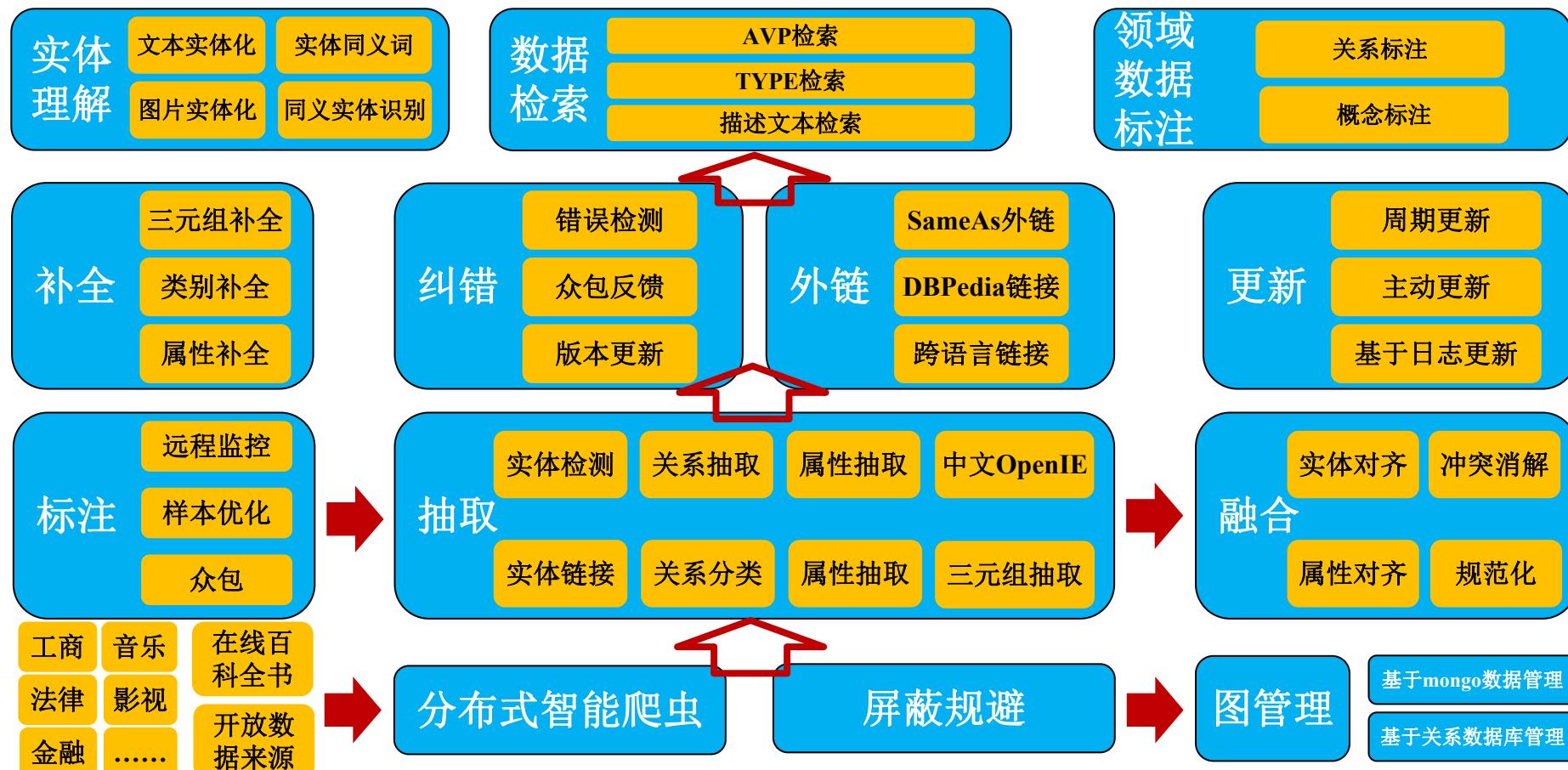
Ulm

/people/person/place_of_birth

Place of birth

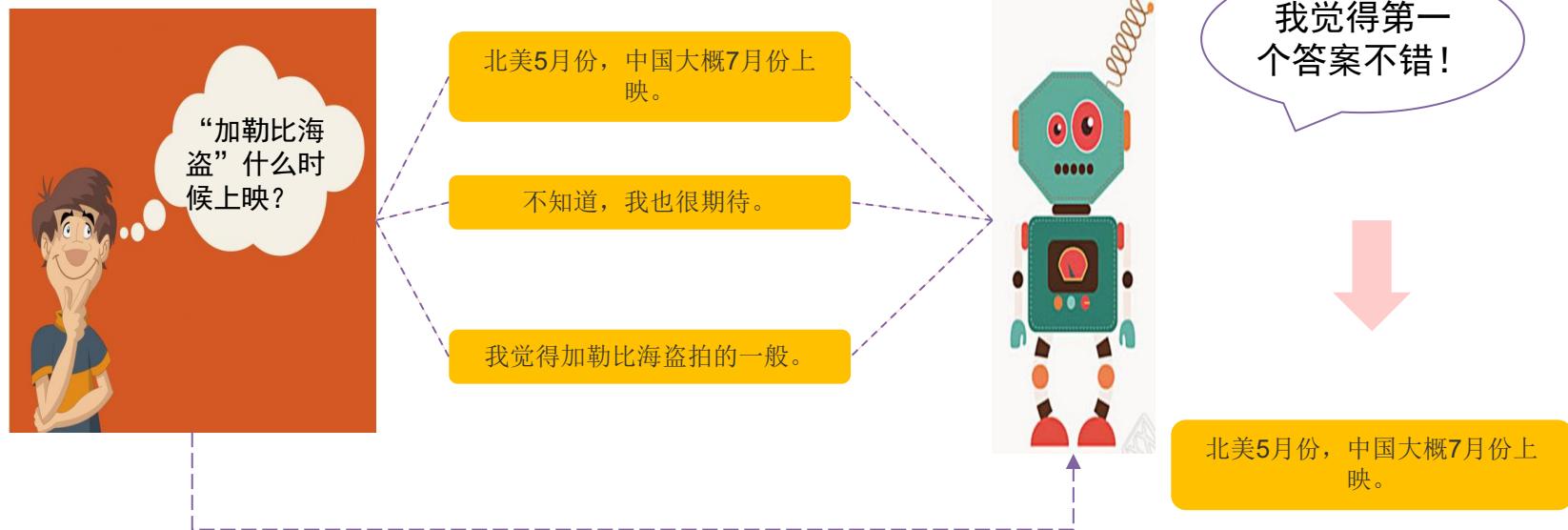
YAGO2 uses
SPOTL tuples
(SPO + Time
and Location)

知识库构建框架



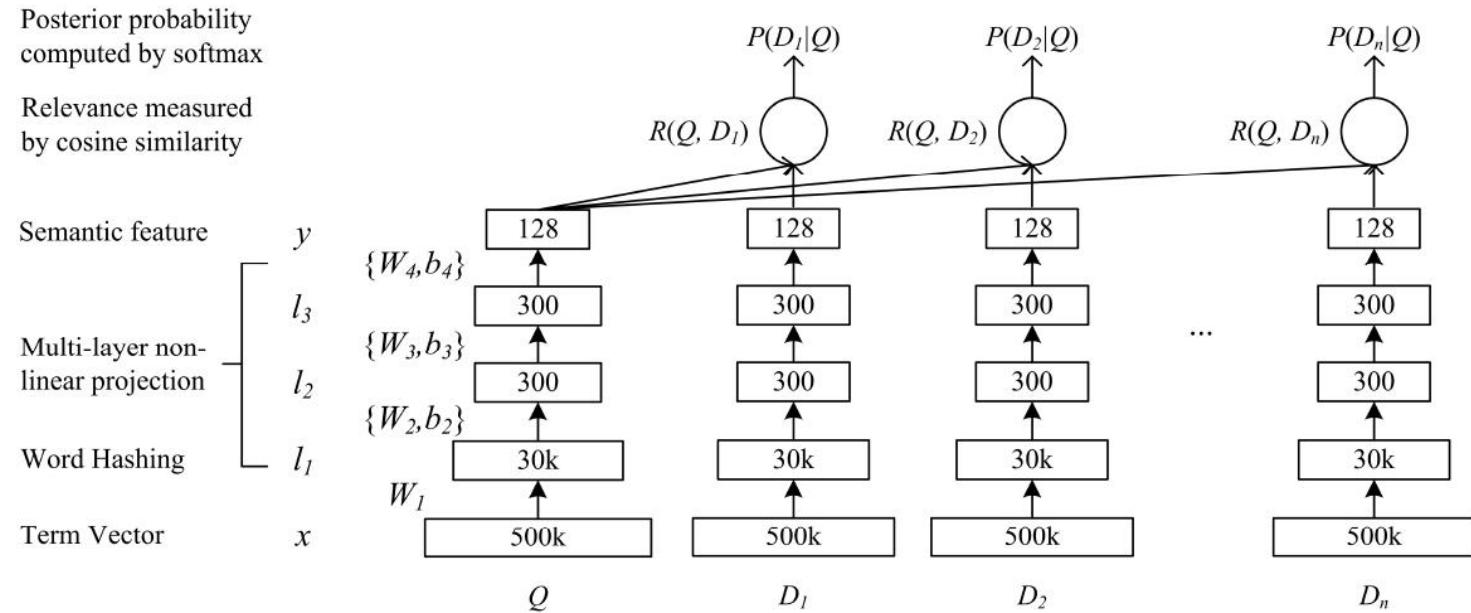
传统方法

- 传统的深度学习方法仅仅是通过语义，建立问题与问题、问题与答案之间的语义联系



微软的DSSM模型

■ 基于多层感知机构建了整个模型：



Huang, Po-Sen, et al. "Learning deep structured semantic models for web search using click-through data." Proceedings of the 22nd ACM international conference on Conference on information & knowledge management. ACM, 2013.

Baidu的CFO框架

■ 一阶知识推理：给定查询问句，对问句理解并在知识库中推理

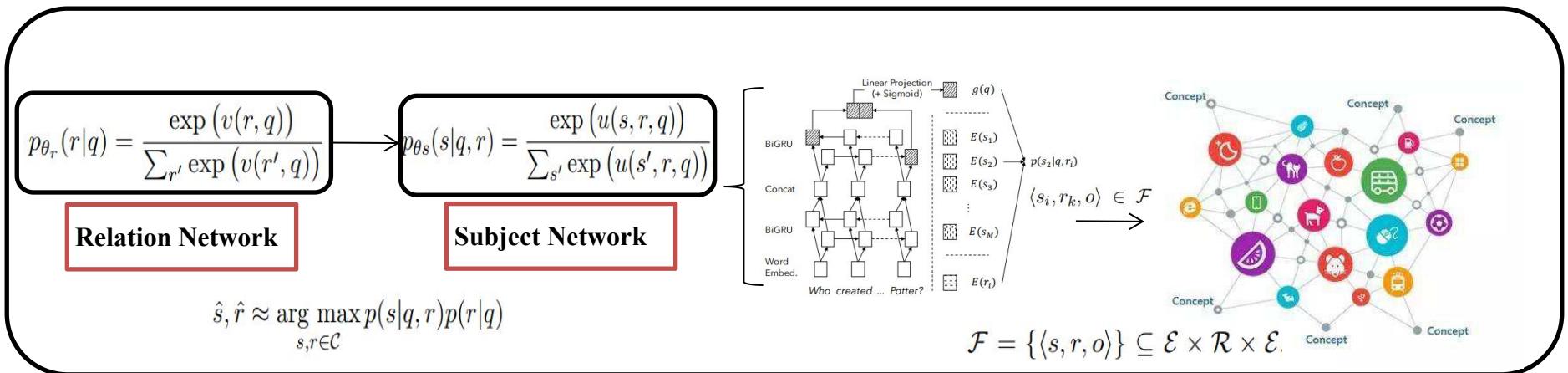
◆ **关系检测：** $p_{\theta_r}(r|q) = \frac{\exp(v(r, q))}{\sum_{r'} \exp(v(r', q))}$

◆ **对象推理：** $p_{\theta_s}(s|q, r) = \frac{\exp(u(s, r, q))}{\sum_{s'} \exp(u(s', r, q))}$

◆ **答案推理：** $\langle s_i, r_k, o \rangle \in \mathcal{F}$

■ 例子：“Who created the character Harry Potter?”

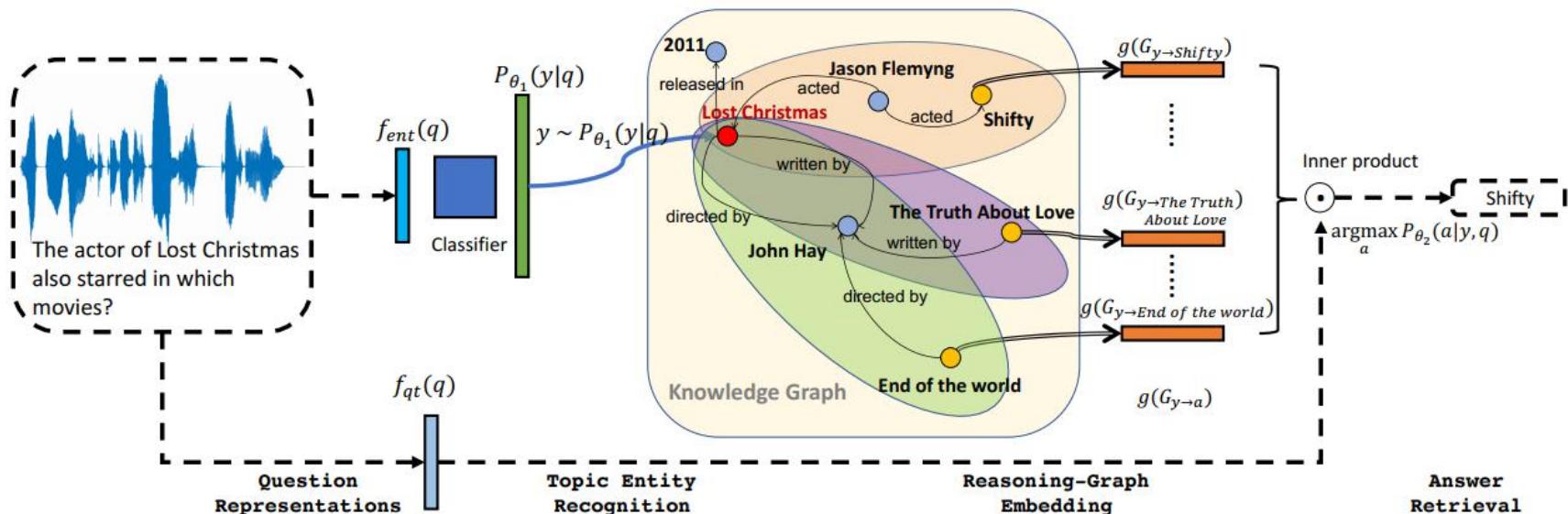
■ <?, CharacterCreatedBy, ?> => <HarryPotter, CharacterCreatedBy, ?> => <HarryPotter, CharacterCreatedBy, J.K.Rowling>



Amazon的Reasoning4KG

■ 主题驱动推理：给定问句与知识库，

- ◆ 主题检测： $P_{\theta_1}(y|q) = \text{softmax} \left(W_y^\top f_{\text{ent}}(q) \right)$
- ◆ 候选答案生成：2-hop scope of topic entity
- ◆ 答案排序： $P_{\theta_2}(a|y, q) = \text{softmax} \left(f_{\text{qt}}(q)^\top g(\mathcal{G}_{y \rightarrow a}) \right)$



Yuyu Zhang, Hanjun Dai, Zornitsa Kozareva, Alexander Smola and Le Song, "Variational Reasoning for Question Answering with Knowledge Graph", AAAI 2018

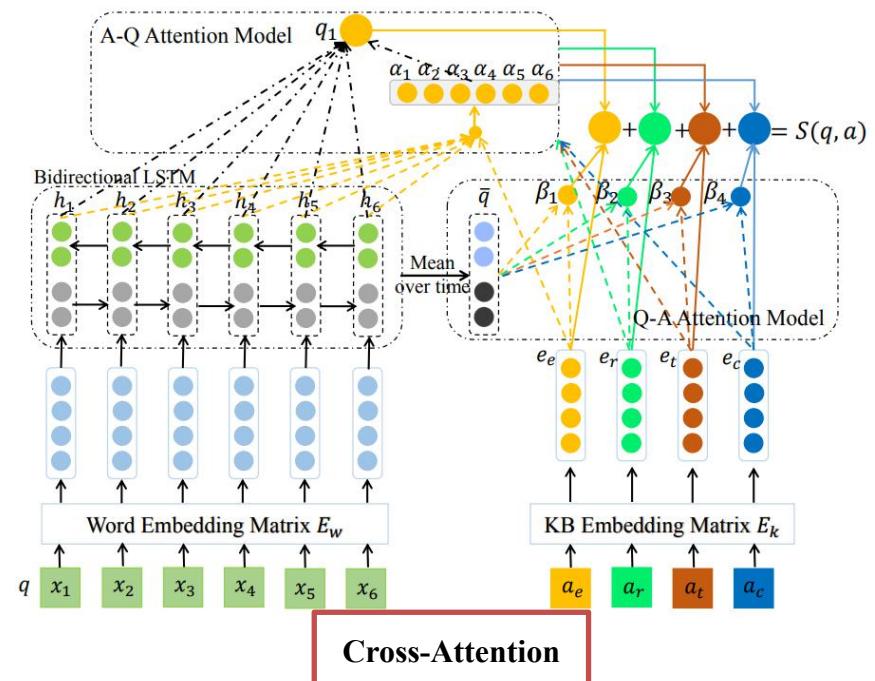
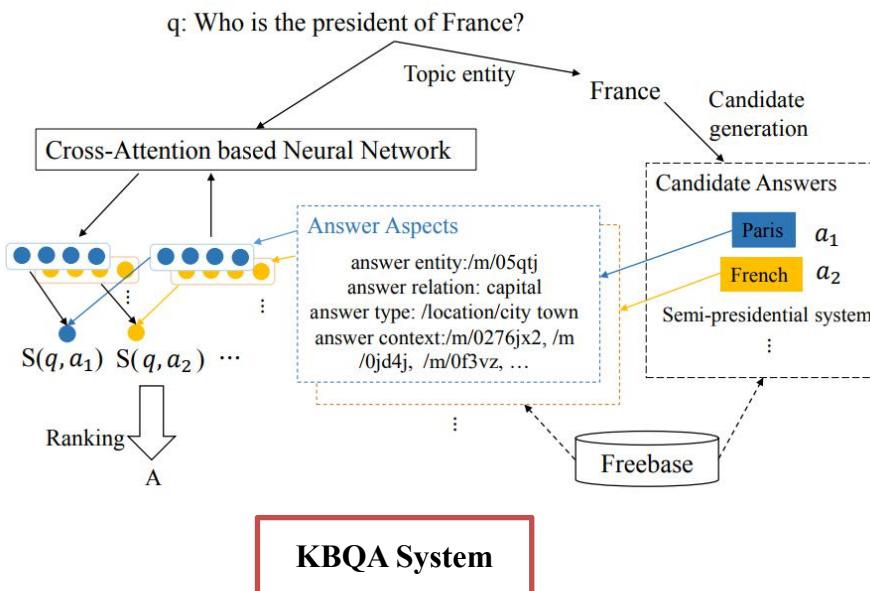
中科院&百度的KBQA

■ 互相聚焦知识推理：给定问句与知识库，

- ◆ 问句主题实体检测：Freebase API

- ◆ 候选答案生成：2-hop scope of topic entity

- ◆ 互相聚焦排序： $L_{q,a,a'} = [\gamma + S(q, a') - S(q, a)]_+$



Yanchao Hao, Yuanzhe Zhang, Kang Liu, Shizhu He, Zhanyi Liu, Hua Wu, Jun Zhao: An End-to-End Model for Question Answering over Knowledge Base with Cross-Attention Combining Global Knowledge. ACL (1) 2017: 221-231

视觉理解与问答

- 机器视觉理解与人阅读图片任务相类似。
- 给定一张图片，针对图片提出一些问题，机器回答这些问题来证明理解图片的内容。



What color are her eyes?
What is the mustache made of?

视频问答

- 给定一段视频，针对视频提出一些问题，机器通过理解视频内容和提出的问题，输出正确答案。



Q: What is the woman boiling in a pot of water?

A: Eggs

视频问答难点

■ 语义鸿沟：

- ◆ 如何打通文本问题与视觉视频之间的语义联系

■ 多视角：

- ◆ 如何理解视频中视觉语义不一致的情况

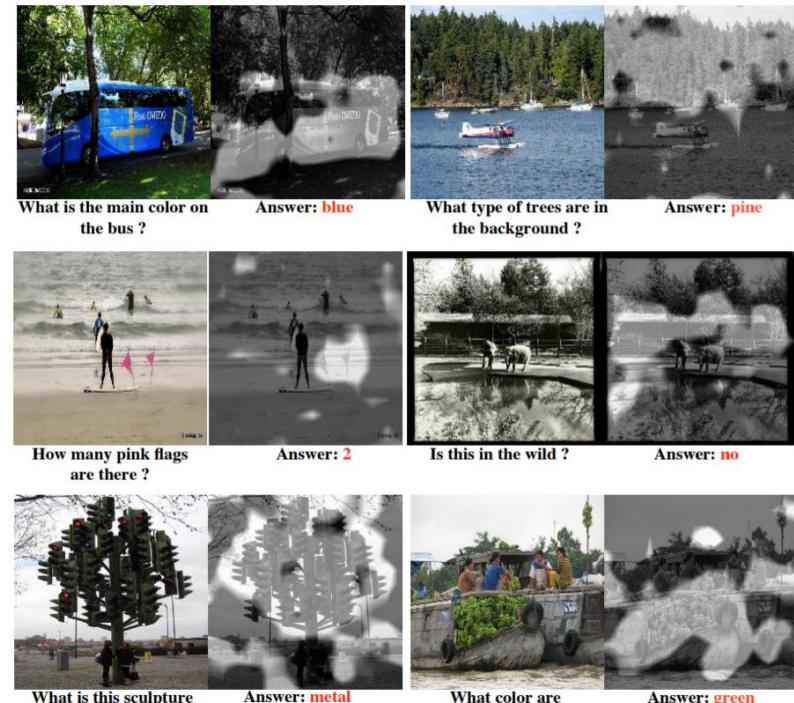
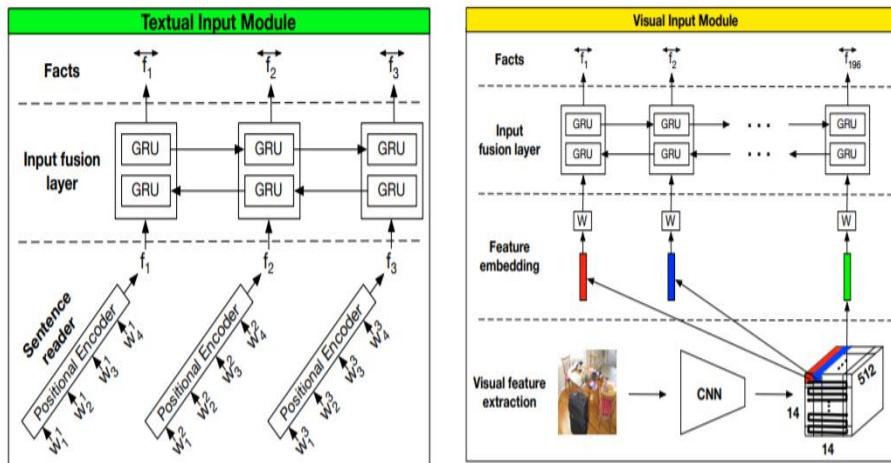


Salesforce公司的图像问答

■ 深度神经网络理解语义信息

- ◆ 卷积神经网络(CNN)理解图片语义信息
- ◆ 循环神经网络(RNN)理解问题语义信息

■ 多源语义融合与推理，答案推理

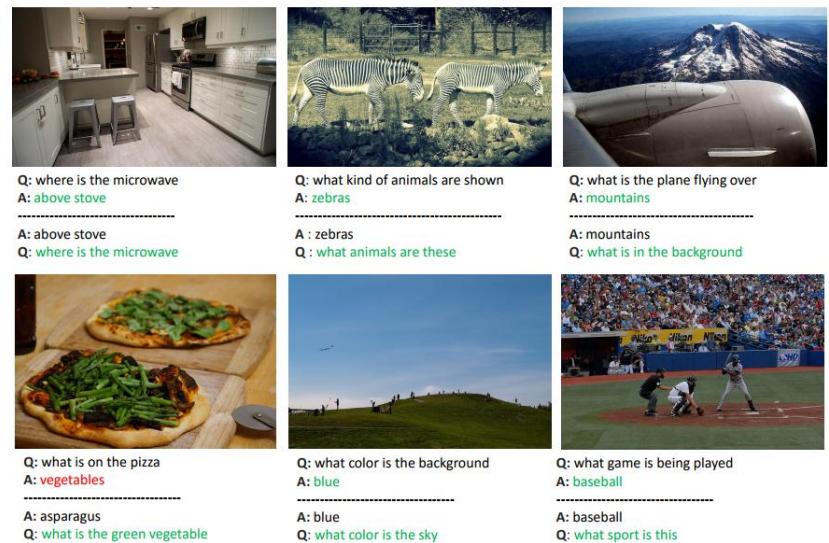
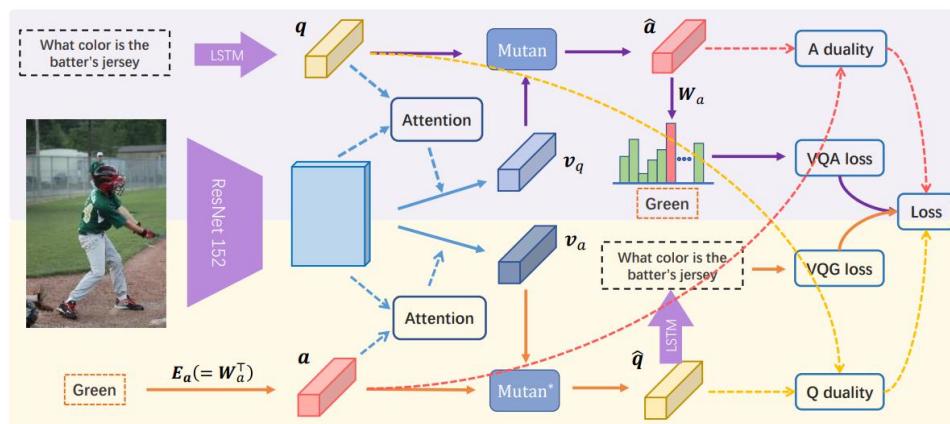


Xiong, Caiming, Stephen Merity, and Richard Socher. "Dynamic memory networks for visual and textual question answering." International Conference on Machine Learning. 2016.

Microsoft的iVQA图像网络

■ 可逆图像问答神经网络

- ◆ 问题 + 图片 → 答案
- ◆ 图片 + 答案 → 问题



Yikang Li, Nan Duan, Bolei Zhou, Xiao Chu, Wanli Ouyang, Xiaogang Wang: Visual Question Generation as Dual Task of Visual Question Answering

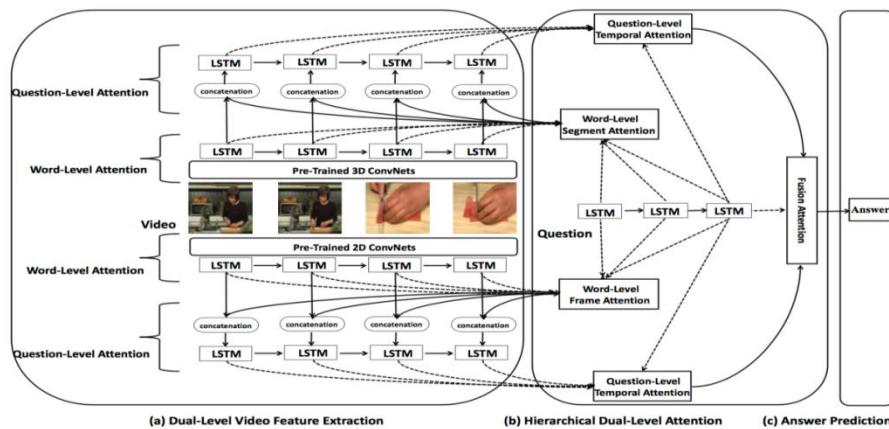
我们的视频问答网络

■ 双通道聚焦神经网络推理

- ◆ 动态特征
- ◆ 静态特征

■ 多层次聚焦神经网络推理

- ◆ 词语层次
- ◆ 问题层次



vegetables
fish
Pasta
air

Question: What is the woman slicing? **Answer:**

Zhou Zhao, Jinghao Lin, Xinghua Jiang, Deng Cai, Xiaofei He, Yueling Zhuang, Video Question Answering via Hierarchical Dual-Level Attention Network Learning, ACM Multimedia 2017

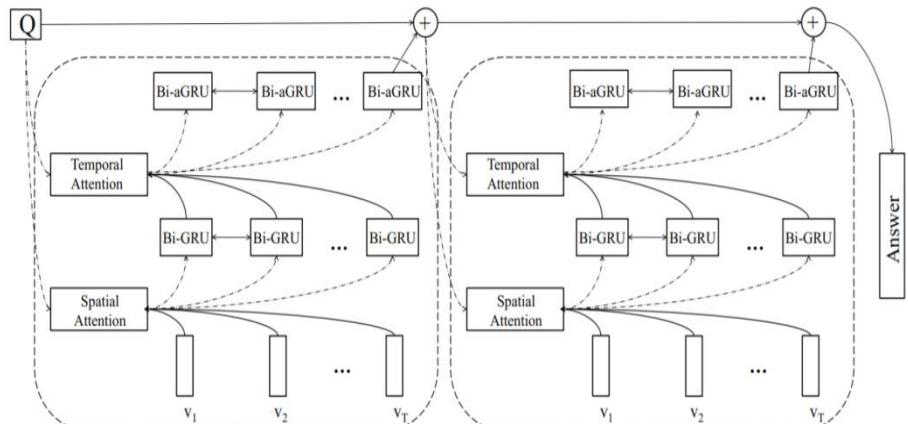
开放式视频问答网络

■ 空间时间联合聚焦推理

- ◆ 空间聚焦推理答案目标
- ◆ 时间聚焦推理重要视频帧

■ 基于多源融合的答案生成

- ◆ 基于循环神经网络的语言模型

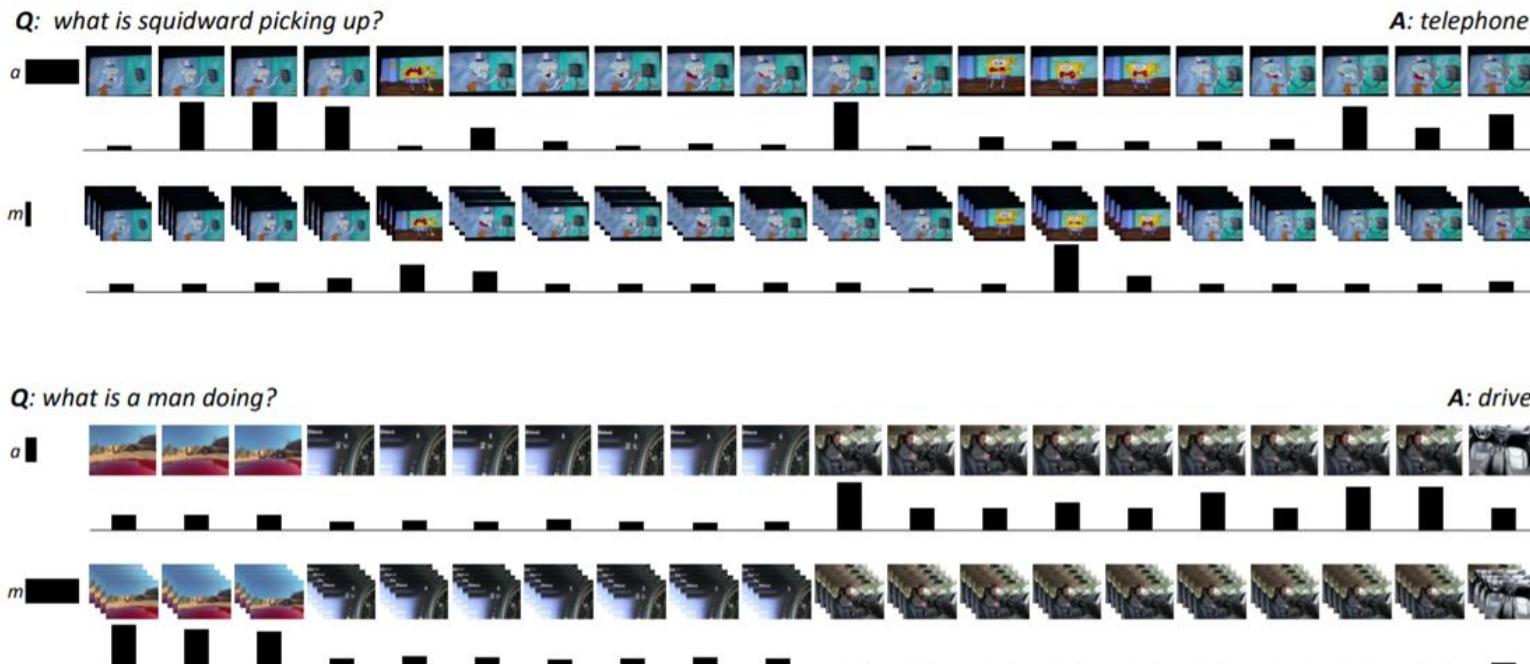


Question: What is the cat doing? Answer: playing with a tablet

更多例子

■ 双通道聚焦网络推理：

- ◆ 通道 a：静态特征推理
- ◆ 通道 m：动态特征推理



Visualization of the attention for two examples. *a* stands for appearance and *m* stands for motion.

Dejing Xu, Zhou Zhao, Jun Xiao, Fei Wu, Hanwang Zhang, Xiangnan He, Yueting Zhuang, Video Question Answering via Gradually Refined Attention over Appearance and Motion, ACM Multimedia 2017

场景理解与多轮对话

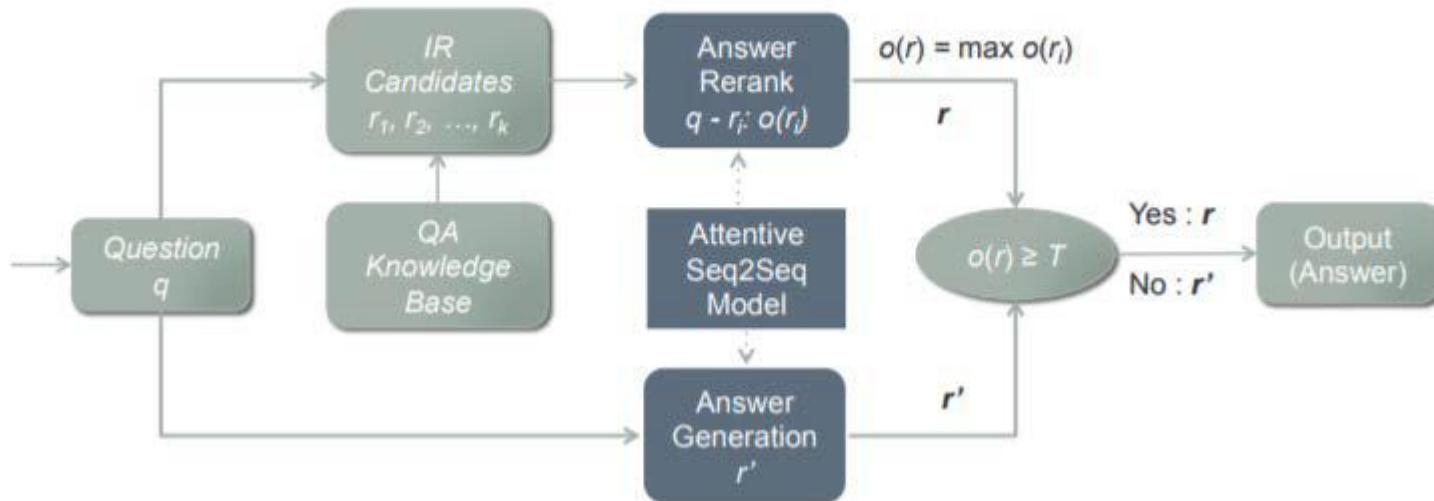
■ “合格”的聊天

- ◆ 表达内容明确，聊天逻辑清晰
- ◆ 每一句话都会针对之前的内容做出反应
(提问、回答、补充等)
- ◆ 主题变化循序渐进



Alibaba的AliMe Chat

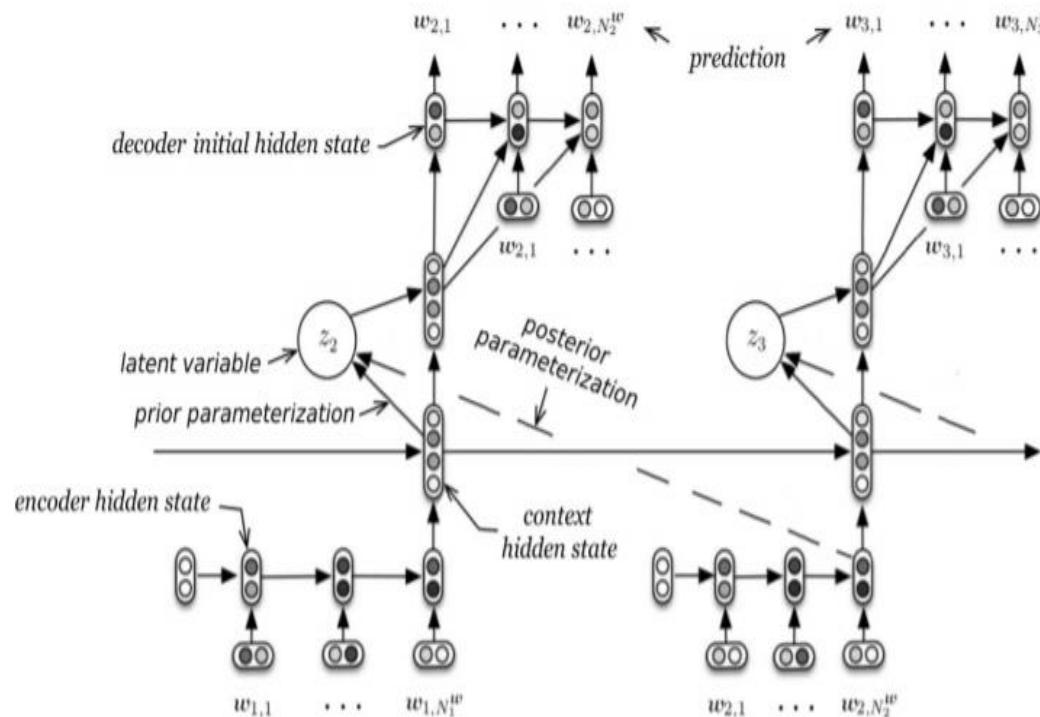
- 利用知识库、注意力机制等技术
- 利用Seq2Seq生成答案



Qiu, Minghui, et al. "AliMeChat: A Sequence to Sequence and Rerankbased Chatbot Engine." Annual Meeting of the Association for Computational Linguistics (ACL-55 Short Paper). 2017.

对话生成

- 使用深度神经网络提取对话中潜在的重要信息，通过使注意力机制，综合利用这些信息，引导对话系统生成更符合上下文的对话语句。



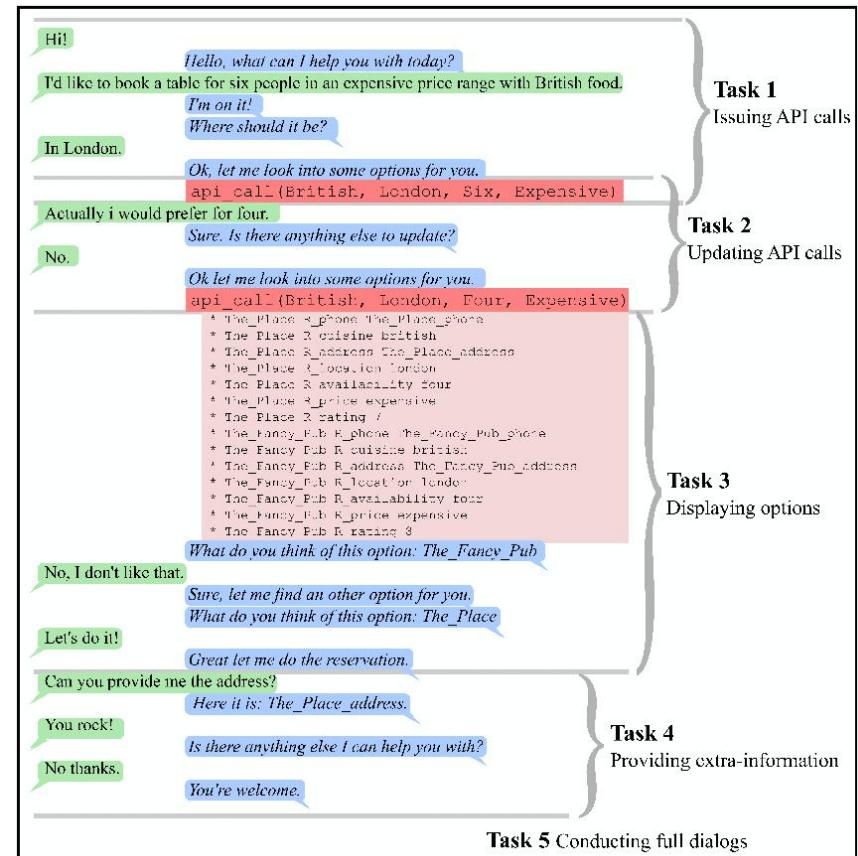
Serban, Iulian Vlad, et al. "A Hierarchical Latent Variable Encoder-Decoder Model for Generating Dialogues." AAAI. 2017.

Facebook的Goal-Oriented Dialog

■ 提出餐馆预约任务场景对话，技术基于

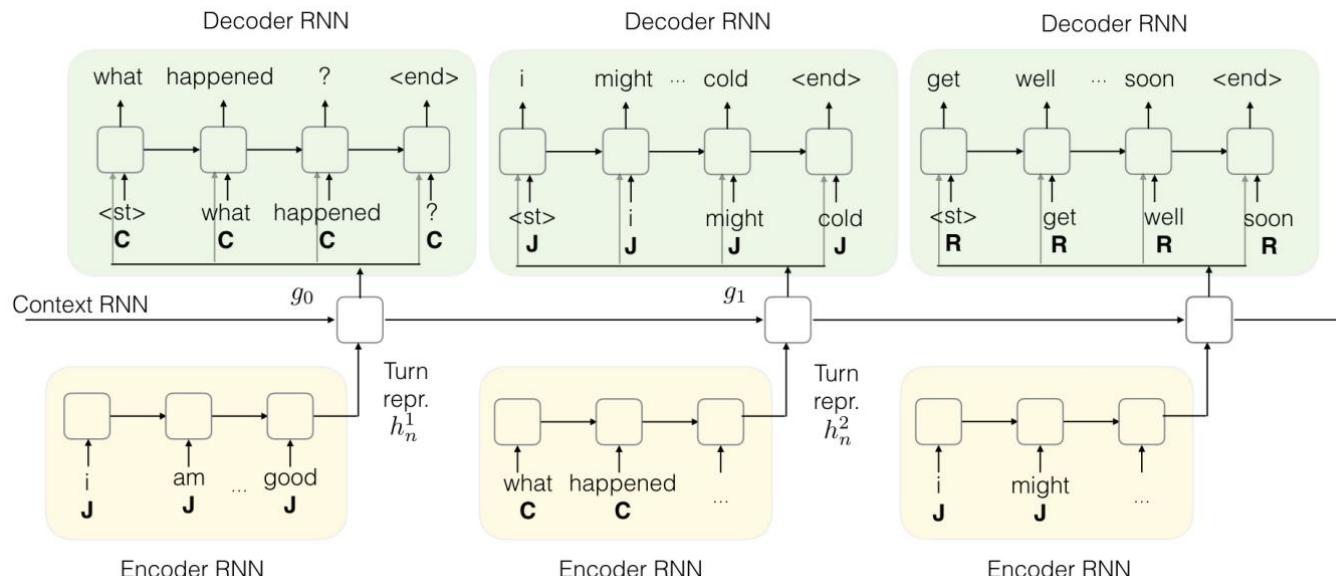
- ◆ 知识库
- ◆ Seq2Seq生成模型

■ 对于OOV词，用type分类检测



Snap的PNCM Chat

- 采用场景循环网络（Context RNN）理解当前对话场景
 - ◆ 记入每轮对话的状态
- 利用Seq2Seq模型生成对话
 - ◆ 融合当前场景和上轮对话来联合生成

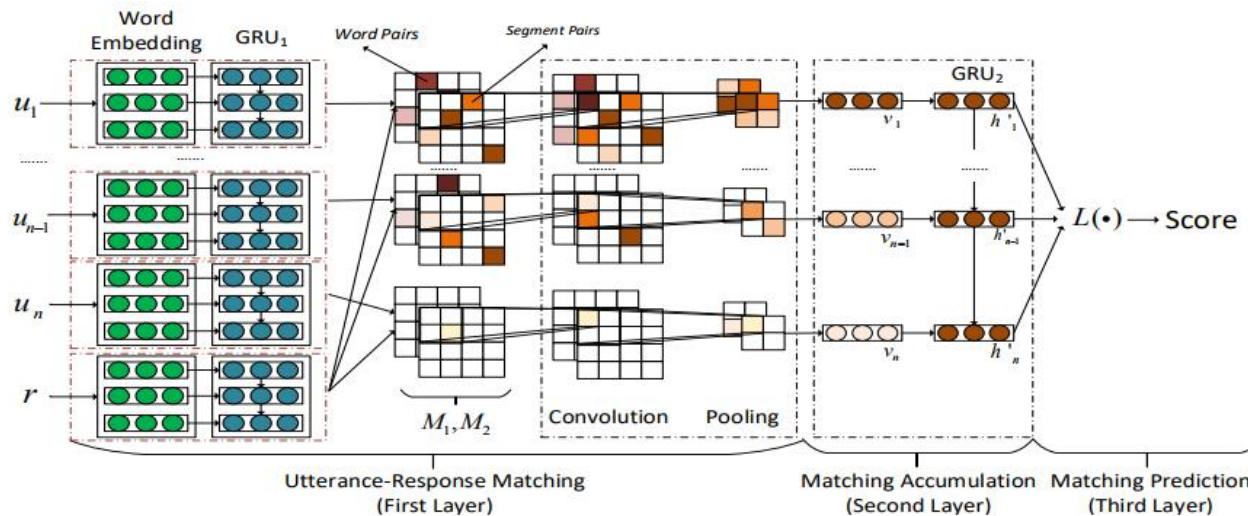


Satwik Kottur, Xiaoyu Wang, Vítor Carvalho: Exploring Personalized Neural Conversational Models. IJCAI 2017: 3728-3734

Microsoft的SMN Chat

■ 在端对端框架下的场景序列匹配模型

- ◆ 聚焦重要场景
- ◆ 序列化理解当前场景
- ◆ 多粒度匹配回复



Yu Wu, Wei Wu, Chen Xing, Ming Zhou, Zhoujun Li: Sequential Matching Network: A New Architecture for Multi-turn Response Selection in Retrieval-Based Chatbots. ACL (1) 2017: 496-505

CMU的RubyStar

■ 意图分析

- ◆ 32个预定义意图

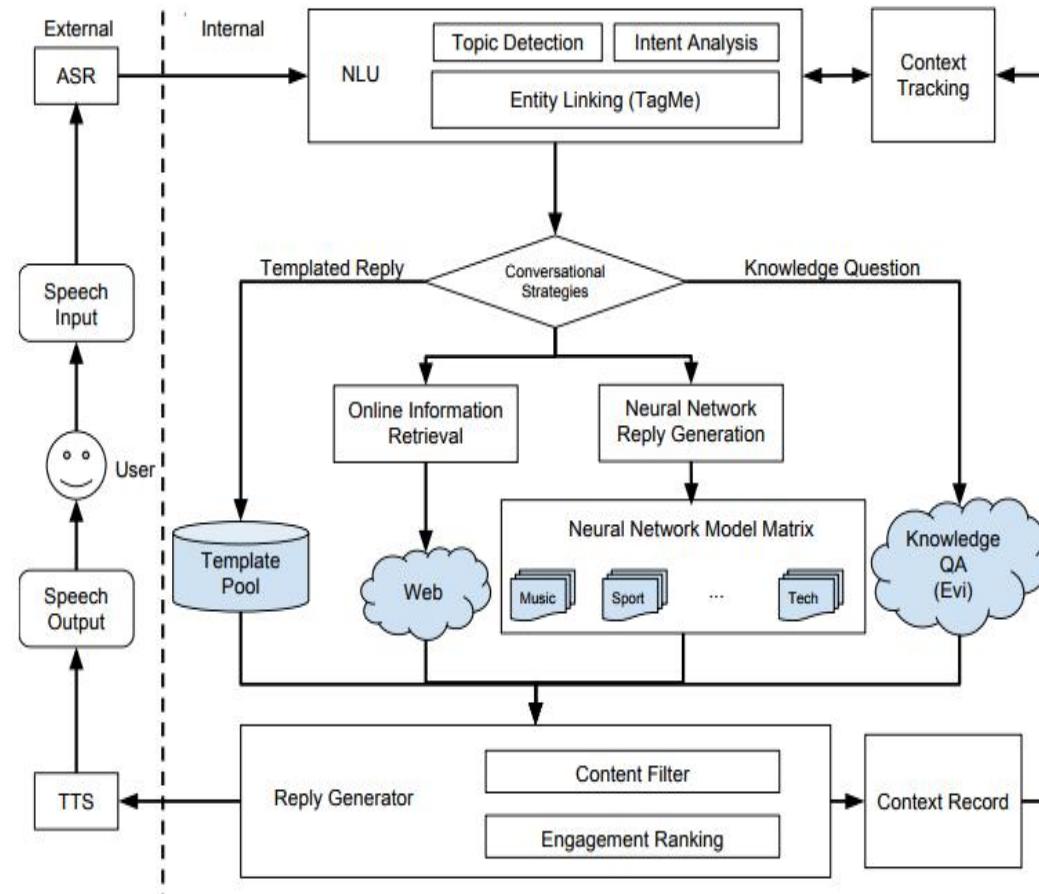
■ 主题检测

■ 信息检索

■ 对话生成

- ◆ 基于char-RNN

■ 组合排序模型



我们和腾讯的PRG Chat

- 使用Seq2Seq模型，首先对问题进行编码，在解码过程中，加入年龄、性别等个性化信息
- 使用了Dual Learning学习机制，在增强学习框架的基础上，加入个性化定制模型，产生聊天的回复

3.2.2 *Dual learning with policy gradient*. In this section, we propose a policy gradient [8] reinforcement learning algorithm to optimize long-term rewards of the *post agent* and the *response agent*. We first introduce the state, action, policy and reward of the reinforcement learning architecture below.

3.2.3 *State*. In time t , a state s is denoted by the input sequence of the *post agent* or the *response agent*.

3.2.4 *Policy*. We employ the stochastic policy gradient method to approximate a stochastic policy directly using an independent function approximator with its own parameters. In this paper, the policies for the *post agent* and the *response agent* are $P(s_{r_{mid}}|p; \theta_A)$ and $P(p|s_{r_{mid}}; \theta_B)$, whose inputs are the representations of the states, whose outputs are action selection probabilities.

3.2.5 *Action*. An action a in timestep k is the sequence to generate ($s_{r_{mid}}$ or p) by the policy ($P(s_{r_{mid}}|p; \theta_A)$ or $P(p|s_{r_{mid}}; \theta_B)$).

3.2.6 *Reward*. The policy gradient algorithm is a type of reinforcement learning method, which relies upon optimizing parametrized policy with respect to the expected return (long-term cumulative reward) by gradient descent. We assume that our model receives a reward r at each iteration. We discuss the major factors that contribute to the reward for post agent A below, and denote the total reward as r_A . Similarly, we can get the reward r_B for response agent B in the dual learning procedure.

四、总结

大数据下智能问答：从检索到理解

- 大数据时代，在医疗、教育、法律等垂直领域涌现多处智能问答应用场景
- 语义理解是提高智能问答性能的重要途径
 - ◆ 机器阅读理解
 - ◆ 知识库构建与问答
 - ◆ 视觉理解与问答
 - ◆ 场景理解与多轮问答

QA

谢谢！