


当知识图谱遇上聊天机器人

本文整理自狗尾草科技 CTO 王昊奋老师在广州知识图谱与问答系统论坛的演讲。

大家好，今天我演讲的题目是 When KG meets Chatbots，即当知识图谱遇上聊天机器人。具体来说，我们主要探讨一下，在聊天机器人（Chatbots）中，知识图谱（Knowledge Graph, KG）将如何融入，起到重要的支撑作用。



- 1 简介
- 2 为什么需要KG
- 3 需要什么样的KG
- 4 KG应用

 开放知识图谱

我的报告包括四块内容，首先是简介，全面介绍 Chatbots 的生态。其次，我会着重介绍一下为什么需要 KG，尤其是在 Chatbots 应用中引入知识图谱的价值。接着，我将根据自身在带领开发公子小白和 Holoera 等旨在创造虚拟生命的聊天机器人产品的经验以及结合 Chatbots 应用落地的思考，与大家一起分享一下 Chatbots 需要怎样的 KG，并列出具将 KG 应用于 Chatbots 所面临的核心挑战；最后，我将总结 KG 用于 Chatbots 的几大典型应用。

01

简介

开放知识图谱

聊天机器人历史



开放知识图谱

首先，我们来回顾一下聊天机器人的历史发展。对话问答是极具挑战的，伴随着人工智能的发展，Chatbos 往往被作为杀手级应用来证明人工智能技术已经达到了相当的高度。纵观其时间轴，我们不难发现从 2010 年起，Chatbots 产品或平台推出的频率较之前明显提升。与此同时，推出方也逐步从学术界主导渐渐转变为互联网巨头公司。这也意味着，随着大数据的积累、计算能力的提升、以及各项人工智能技术慢慢成熟，以及用

户市场的培育，Chatbots 慢慢从验证技术向应用落地转变。很多专家也戏称 2016 年为 Chatbots 元年。

我将从 SIRI 开始对目前知名的 chatbots 产品和平台做一个详细的介绍。在此之前，我不得不提到两个重要的早期产品。作为第一个吃螃蟹的，ELIZA 诞生于 1966 年，开发者为 MIT 的 Joseph Weizenbaum。它是一个模拟罗杰斯心理治疗的 BASIC 程序，是自然语言处理方面的先驱者。另一个不得不提及的便是 A.L.I.C.E，起源于美国国防部 DARPA 的一个项目，它诞生于符号 AI 鼎盛时期，基于规则和模板来处理问句理解和回复逻辑。他对 Chatbots 的后续发展起到了非常深远的影响。它的一个贡献是定义了 AIML，一种类 XML 的声明式语言，可以用来编写各种问答对，甚至多轮对话的对话逻辑。这种做法对于数据规模小，且领域知识相对充足的情况下，非常利于解决冷启动并提供表现不错的对话体验，同时也是对当前基于机器学习尤其是深度学习方法的一种补充。不仅可用来积累数据，也可以提供可解释性更强，且能针对个别用例做针对性地在线干预和修正。所以，到目前为止，大部分商用聊天机器人尤其是特定领域的对话应用沿用了这一思路。

聊天机器人简单分类



娱乐化

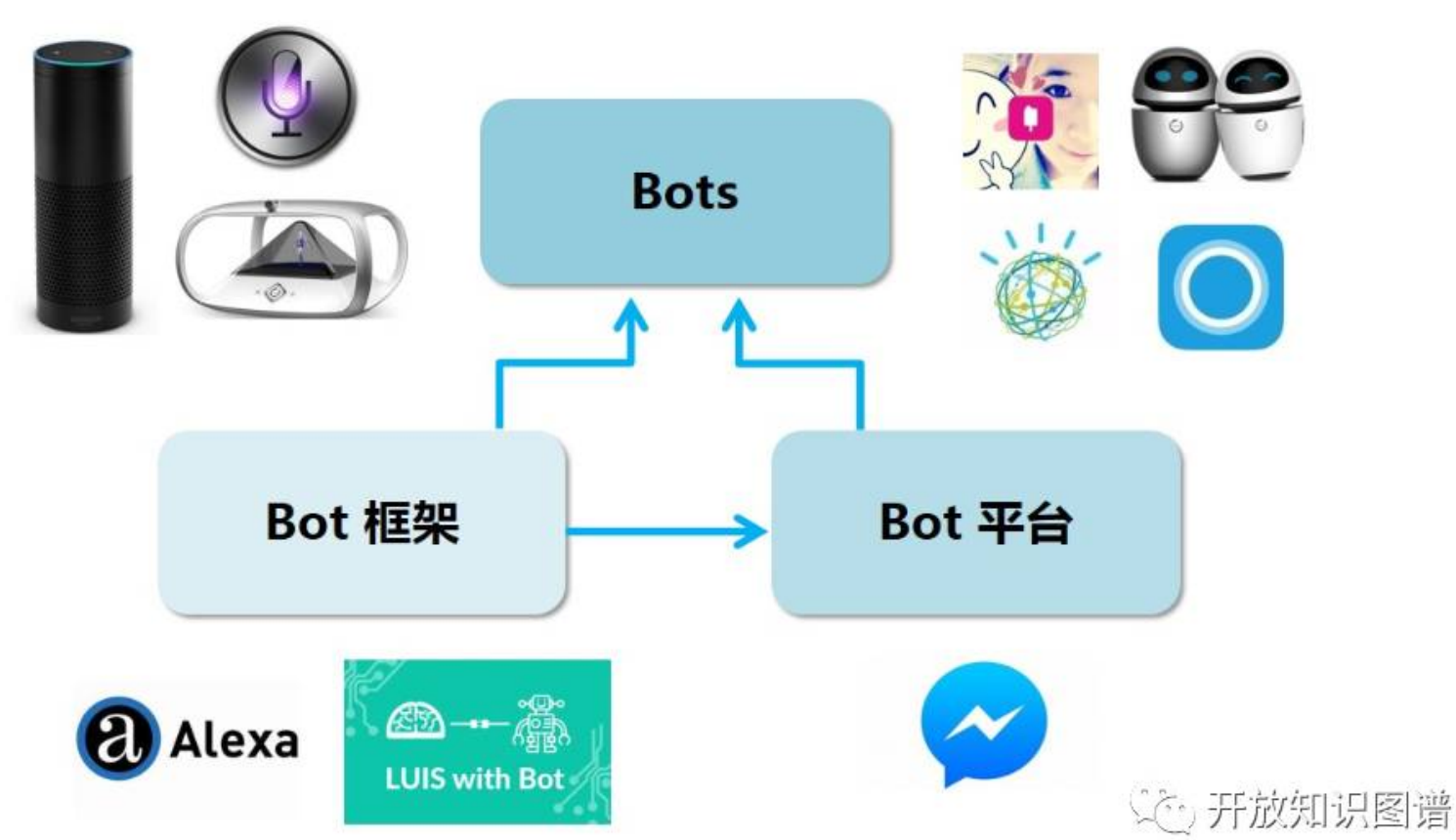


工具化

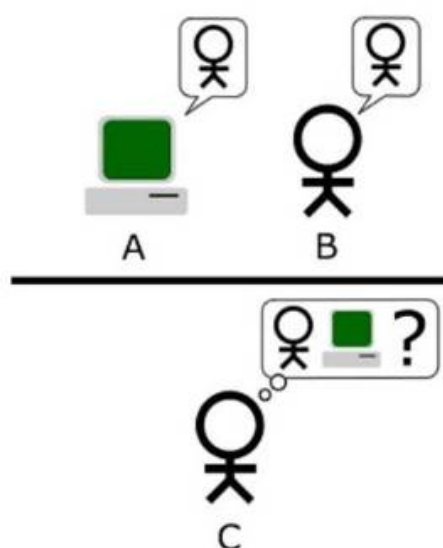
开放知识图谱

在作详细介绍之前，我们先对聊天机器人从分类角度进行一定的梳理。我们从最简单的二分类着手，即 Chatbots 根据其用途或使用场景，可以分为偏娱乐化或偏工具化。图中给出了一个鲜明的对比，白富美的 Eva 更擅长的是情感陪伴，而蓝领工人 Wall-E 则更擅长完成特定的工作。这两者

由于目标不同，所以往往在 Chatbots 设计和技术选型上会存在一定的差异。微软公司很早就洞察了这一点，针对偏娱乐化场景和偏工具化的应用分别推出小冰和小娜。



围绕 Bots 生态圈，有些在做实际的 Chatbots，既有硬件形态的 Amazon Echo 及狗尾草公司推出的公子小白和 Holoera 等，也有纯软件的如苹果的 Siri 和微软的小冰等。除此之外，为了加速实际 Bots 的研发，不少巨头或创业企业开始对外提供 Bot 框架（Bot Framework），以 SDK 或 Saas 服务的形态供第三方公司使用来构建特定应用和领域的聊天机器人，这里的典型代表包括支持 Echo 的 Amazon Alexa 和微软推出的包含在 Cognitive Services 大框架下的 LUIS with Bot 等。更进一步，除了提供开发 Bot 的 API，Bot 平台（Bot Platform）进一步考虑开发得到的 Bot 如何部署到一些常用平台中如微信或 Facebook 等。

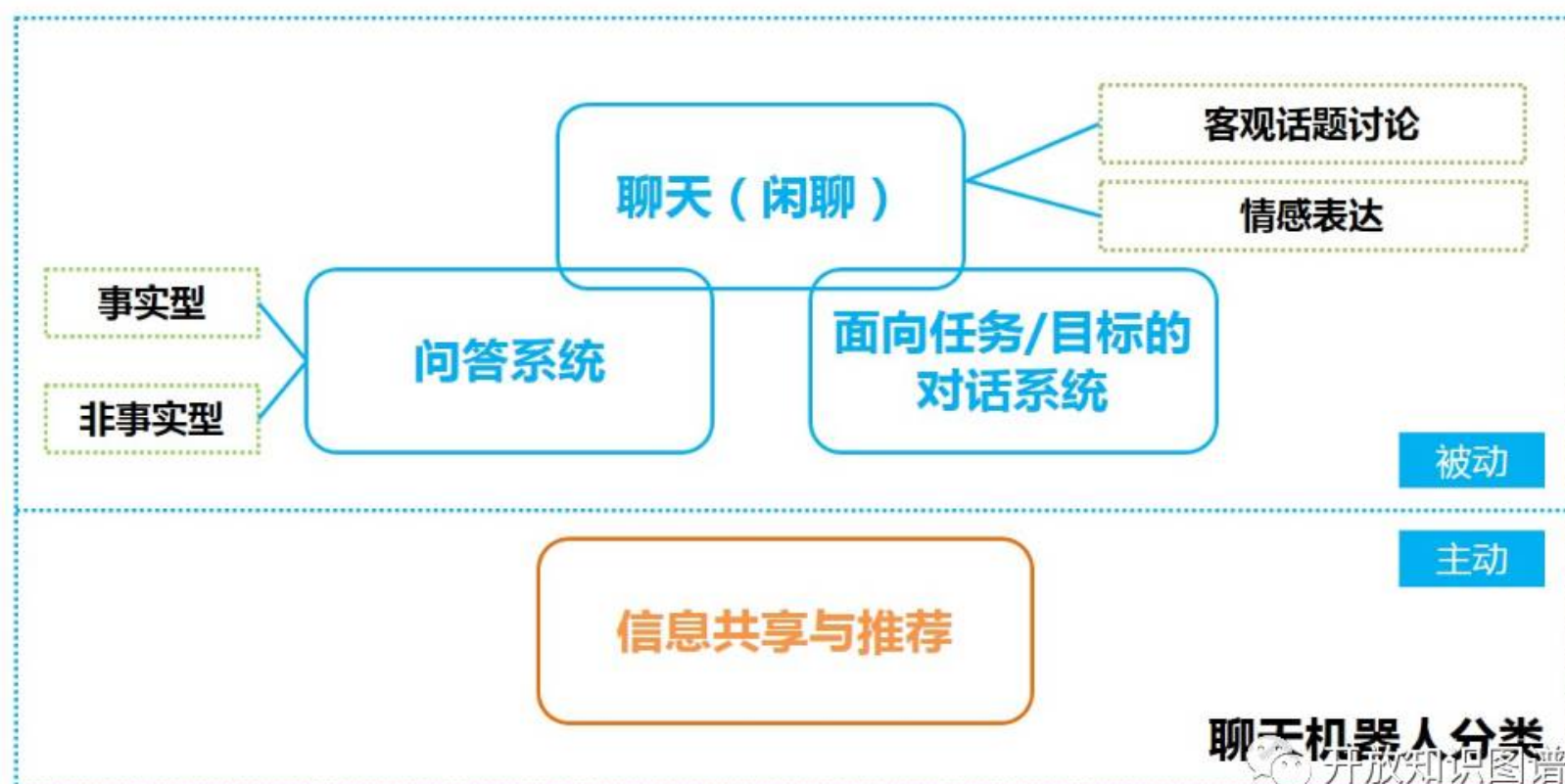


5分钟，30%蒙蔽



开放知识图谱

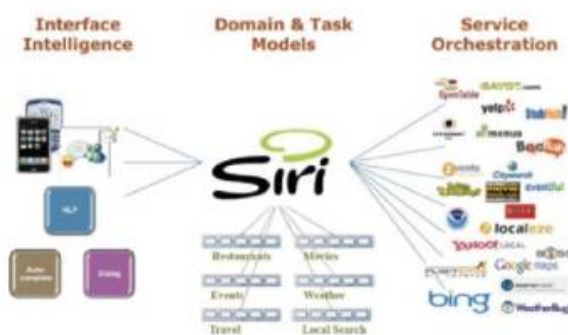
聊天机器人的成功离不开智能。而图灵测试常常被用来测试机器是否具有真的智能。图灵测试是指测试者在与被测试者（一个人和一台机器）隔开的情况下，通过一些装置（如键盘）向被测试者随意提问。进行一系列时长为 5 分钟的测试后，如果有超过 30% 的测试者不能确定出被测试者是人还是机器，那么这台机器就通过了测试，并被认为具有人类智能。图灵测试一词来源于计算机科学和密码学的先驱阿兰·麦席森·图灵写于 1950 年的一篇论文《计算机与智能》，其中 30% 是图灵对 2000 年时的机器思考能力的一个预测，目前我们已远远落后于这个预测。2014 年 6 月 7 号，正值图灵逝世 60 周年纪念日，一款名为尤金·古斯特曼（Eugene Goostmanz）的聊天机器人，它伪装成了一个用第二语言沟通的 13 岁乌克兰男孩儿，成功“骗过”了测试者，通过了图灵测试。不过事后有很多质疑。首先，这个聊天机器人号称只有 13 岁，并使用第二语言来回答问题，以此作为重大缺陷的借口。另外，测试者只有 5 分钟与之展开互动，大大增加了他们在短期内被“欺骗”的概率。尤金无法一直保持对话的顺畅性，他会不断重复自己的话，还经常使用聊天机器人典型的无推断型回应方式。



我们对聊天机器人按照功能和交互方式等二个维度可以进一步细分为四类。交互方式可以分为主动交互或被动交互。目前大家接触到的大部分聊天机器人都属于被动交互的范畴，即由用户发起对话，机器理解对话并作出相应的响应。而主动交互能更好体现机器人和用户之间的对等关系，即通过共享或推荐用户感兴趣或热点事件等由机器人首先发起。目前主动交互更多作为传统交互方式的一种补充，作为辅助手段，并未大规模得到广泛使用。从功能角度来看，聊天机器人可以细分为以闲聊为主的聊天，问答和面向任务或目标的对话。其中，若根据场景切分，闲聊可以进一步分为情感交流和针对客观话题的讨论。问答系统需要不仅考虑如 What、Who、Which、Where 和 When 等事实型问答（Factoid QA），也需要考虑如 How 和 Why 等非事实型问答。



2010



- 实体为自然语言理解提供消歧
- 对于解决不了的问题，Siri直接调用搜索引擎

VIV : Siri的核心人物Dag Kittlaus和Adam Cheyer创立，被认为是Siri的升级版

INTELLIGENCE BECOMES A UTILITY



VIV

语音个人助理

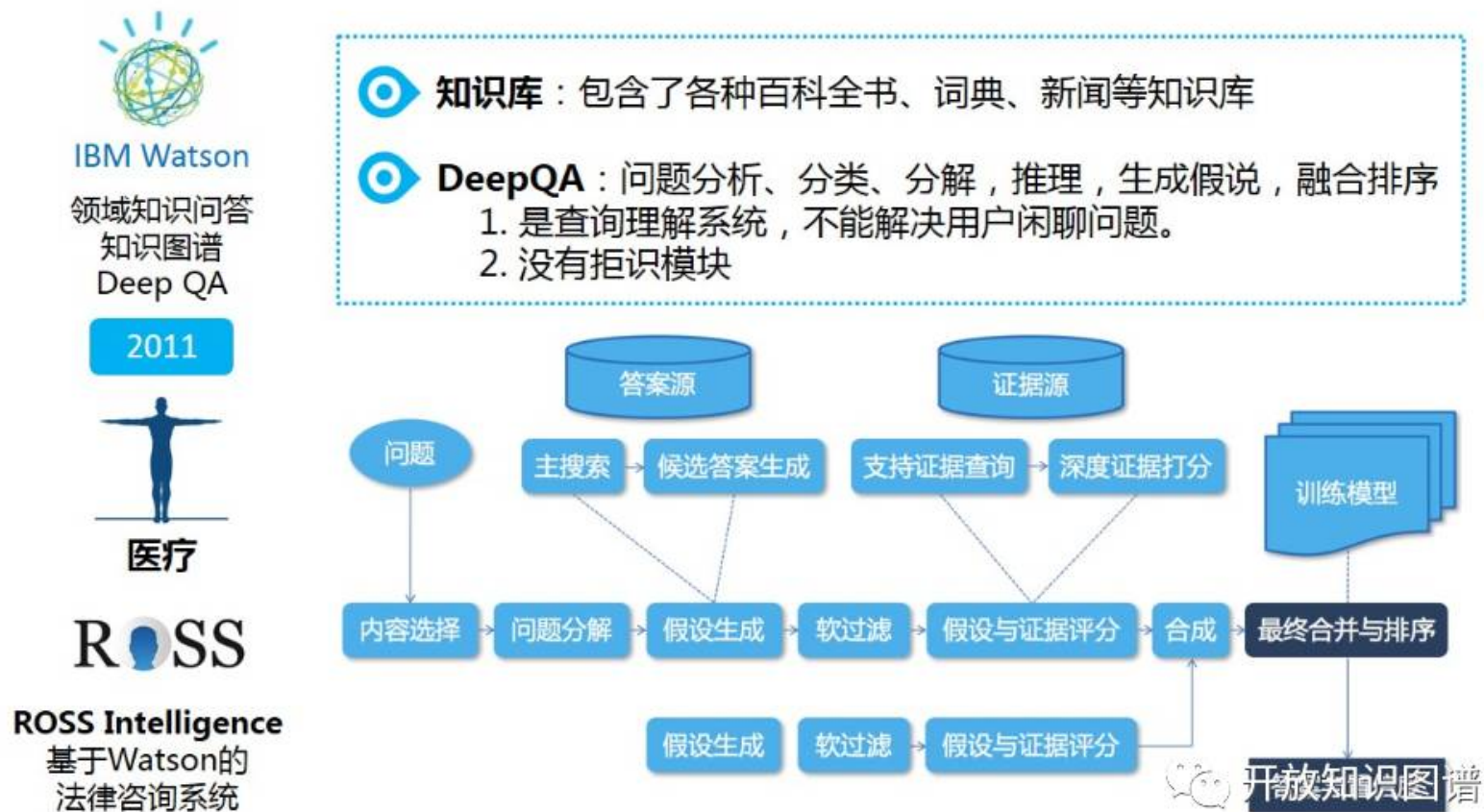
开放知识图谱

Siri 作为 iPhone 4S 推出时的一个亮点特征，定位是语音个人助理。在推出之时，引起了极大的轰动。虽然这么多年 Siri 的技术也不断提升，但我觉得他更大的价值是作为聊天机器人之门开启的开门砖，教育了用户和市场。回到刚刚的分类，Siri 其实是一个面向特定任务的对话系统。他对接了很多本地服务如通讯录、音乐播放等以及 Web 服务如订餐、订票和导航等功能服务。针对这些服务意图，他通过实体驱动的自然语言理解

(Natural Language Understanding, NLU) 来识别问句中涉及到的对象和相关服务，从而实现特定任务下的多轮功能交互。对于解决不了的问题，即服务意图范畴外的需求，则直接调用搜索引擎返回相关答案来返回。随后，Siri 的核心人员 Dag Kittlaus 和 Adam Cheyer 于 2016 年推出了 Viv。Viv 被认为是 Siri 的升级版，虽然其在多服务组合，服务编排等方面做了不少亮点工作，但背后的基本原理和定位和 Siri 无差异。



正如之前介绍的那样，微软针对娱乐化和工具化这两个截然不同的定位，分别推出了小冰和小娜（Cortana）。小娜，作为嵌入在 Windows 或 Windows Mobile 等微软操作系统内核的语音个人助理，承载着类似 Siri 或 Viv 的角色，它的目的是提升用户的工作效率，据说 Cortana 有 1.5 亿多用户，这也使得微软吸引到 Bengio 这样的大师作为顾问加入。另一方面，小冰是微软中国团队推出的娱乐聊天机器人。她的人设是一位 16 岁的少女。小冰是一个基于搜索的回复检索系统。通过各种基于深度学习的语义匹配算法，从海量的问答对语料中返回最佳的回复（Message response 而非 Answer）。小冰也会不定期推出新的技能供大家使用，这些技能往往包含了微软团队在图像理解、语音和自然语言理解方面的各种小应用尝试。更值得一提的是：微软针对日本、北美和欧洲等市场陆续推出了具有不同人设的少女如 Rinna、Tay 和 Zo，她们往往可以方便的通过微信、微博或 Twitter 等平台进行交流。



Watson 系统是典型的问答系统，其由 IBM 研究院在 2011 年推出，参加美国知识竞赛 Jeopardy!（危险边缘）并挑落人类冠军而名声大噪。相比 AlphaGo 或早年 IBM 研制的战胜卡斯帕罗夫的国际象棋人工智能程序深蓝，Watson 具有更清晰的商业路径。IBM 斥巨资成立医疗事业部，并与 MD Anderson 等知名医疗机构合作推出面对特定病种（尤其是癌症）的辅助诊断 AI 医生。与此同时，Ross Intelligence 依托 Watson 认知计算平台推出了法律咨询系统。回到技术层面，Watson 所用到的知识库是一个广义的知识库，不仅包含各种结构化知识、也包含各种文本语料和语言学知识。整个流程称为 Deep QA，包含问题分解、假设生成、基于证据的融合排序等关键步骤。这里的 Deep QA 并非指通过深度学习（Deep Learning）技术来提供问答。事实上，Watson 诞生于深度学习大热之前，这里的 Deep 是指通过深度解析（Deep Parsing）来实现对问句的真正理解。

众所周知，自然语言处理往往包含诸多步骤，传统的做法会将这些步骤通过串接的形式形成 pipeline 来完成从词法到句法最后到语义的过程。这样做法的最大问题在于，由于每一环节的组件都不是完美的，通过串接方式形成的系统易产生错误传递和放大，导致整个系统无法在生产环境中使用。针对传统 NLP 的局限，Watson 通过如下手段来进行改善：1) 对于每个组件，采用多种实现算法并集成来避免单点失败从而增加每一环节的准确性和鲁棒性；2) 在候选答案生成阶段以覆盖率为主要目的，避免过早的将正确答案过滤掉；3) 引入证据源和基于多证据的打分（类似循证

医学) 来确保正确答案的排名位于首位或头部地位。总而言之, Watson 做了精细的模型分拆和设计, 并基于集成学习和知识库的结合来实现传统做法 (如 TREC 的 QA 任务中采用的方法) 无法达到的精度和覆盖度。



之前已经提到 Facebook Messenger 是一个庞大的 Bot 平台, 有非常活跃的开发群体, 平台包含上万种 Bots。针对 Messenger, 我想讲以下几点: 第一, 它在 2014 年收购了 wit.ai。Wit.ai 类似于谷歌所收购的 api.ai, 包含大量的行业相关或场景相关的对话。基于以上高质量海量的对话数据, Facebook 基于深度学习技术推出了一个用于自然语言处理的框架叫 DeepText, 用于自然语言表示学习和各种分类等任务。有名的 Fast Text 也包含在内。今年 Facebook 更是基于 Deep Text 推出了 CLUE, 进一步提高了其易用性, 有兴趣的可以去进一步了解一下。通过以上的数据和技术积累, Facebook 就可快速构建一个端到端的 Chatbot 或者问答系统。此外, 还有一点需要强调的是, 我们可以发现 Facebook Bot 的很多应用场景涉及到购物、递送礼物、预约参观和安排旅程等非实时任务, 即相对比较复杂, 但不要求马上得到反馈。传统的做法是, 通过指派一名客服来对接, 提供进一步的服务。对于这些非实时任务, Facebook 结合机器返回的自动化推荐结果和人工的进一步编辑和审核来保证用户体验的同时也降低了纯人工对接存在效率低、工作量大等弊端。而这也是近期大家很推崇的人机融合, 即赋予人工智能新的内涵: Artificial Intelligence+Human Intelligence (人类智能) =Augmented Intelligence (增强智能)。



Alexa 作为亚马逊 Echo 智能音箱背后的 Bot 框架，通过 Skill Set 的形式不断扩展其功能，其内核是亚马逊在 2016 年底发布的 Lex，并对接专注图像识别的 Rekognition 和基于机器学习特别是深度学习技术的快速 TTS（文本到语音转换）。细心的观众会发现 Echo 音箱并没有提供任何屏幕，仅通过语音进行交互，依托 Amazon 的内容资源和电商购物优势提供各种智能交互。这种以语音为主的交互方式在家庭、车载等领域得到广泛关注和应用，由此也提出了 Voice UI 的概念。除了语义理解，这里需要强调的是：对于 Echo 音箱的交互，是采用远场（通常 3-5 米）沟通的。对于远场语音交互，目前远比近场通讯的难度大，涉及到声源定位、噪声（如回声、背景噪声、各种声波反射折射产生的混响）消除、人声分离、声音增强甚至是声纹识别等各种技术挑战。目前通用的做法是采用麦克风阵列+波束成形等方案，不过有很大的提升空间。不过智能音箱是否能在中国成为一个爆款，这个还是一个未知数，当然这里涉及到更多使用习惯、价格、内容质量等很多非技术因素的考量，在此就不做具体展开。



Google Allo

个人助理
深度学习
用户嵌入

2016



Google传统语音系统：

主要是将语音识别、语义处理后的命令调用搜索引擎，返回用户得到的结果



Allo：

具有自我学习的能力，会随着时间的推移，学习用户的说话和行为，帮助用户自动回复短消息、邮件和评论照片

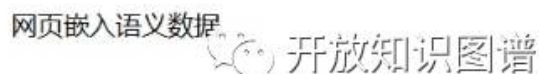
Google采用全新的深度学习框架，加入user embedding特性，可通过用户输入建立用户画像，学习用户行为

开放知识图谱

从 Google Now 到 Google Assistant，谷歌一直没有停止过在语音个人助理方面的尝试。这里想介绍一下基于 Google Assistant 的新一代人工智能类微信 IM 应用。Allo 具有几个亮点：首先，其具备一定的自我学习能力。这里包括两方面的学习，一方面是学习用户的习惯，包括说话风格和交互模式。值得一提的是，Allo 的开发者也参与了 Gmail Smart Reply 功能的开发，帮用户草拟回复的邮件。具体来说，根据邮件接收的对象、主题和关联的场景等，根据用户口吻来尽量完成要回复内容。另一方面也包括用户偏好的学习，这一点在推荐系统中是非常重要的，属于用户画像的学习。Allo 学习用户画像的低维稠密向量化表示（User Embedding）。将 User Embedding 加入 Chatbot 的回复生成解码模型中，将有助于回复的相对一致性和个性化。

2 为什么需要 KG

🗨️ 开放知识图谱



知识图谱于 2012 年由谷歌提出，旨在提供更好的搜索体验。随着整个 Web 从原先由网页和超链接构成的 Web of Docs 转换为由实体或概念及他们之间的关系构成的 Web of Data，谷歌提出了更准确的语义搜索，旨在解决原有的关键字搜索仅基于字符串无法理解内容语义的局限。在 KG

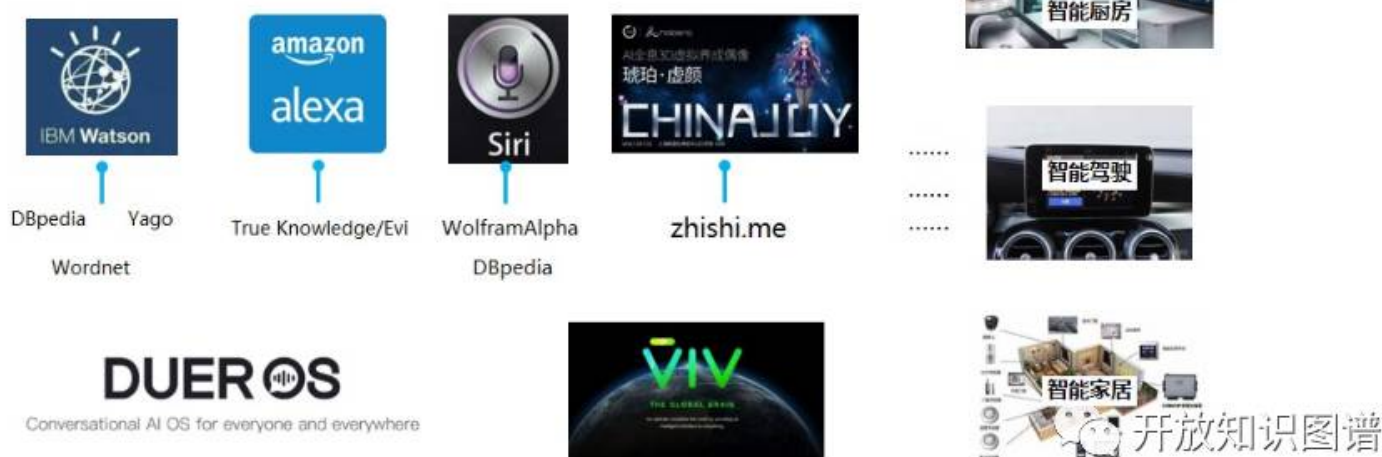
发展的浪潮中，也诞生了基于社区协同构建的众包典范 Freebase。而谷歌的 Google KG 也是基于 Freebase 逐步发展起来的。同样还有 Wikimedia 社区所推出的 WikiData 项目，目前 Freebase 也已经关闭，并将数据等均贡献给了 WikiData 做进一步发展。此外，作为谷歌、Bing、Yahoo! 和 Yandex（俄罗斯搜索引擎）共同推出的 Schema.org，通过鼓励站点所有者（Site Owner）在其页面中添加符合 Schema.org 分类体系（及关联属性等）规范的语义知识片段（以嵌入在 HTML 页面中的 RDFa 或 Microformats 等形式出现）来扩充和完善知识图谱。据悉，25% 的站点和 30% 的页面包括 Schema.org 的标注。对应的回报是提升特定关键字或实体查询时相关站点（提供与关键字或实体相关的高质量语义标注知识）的搜索排名，从而起到免费搜索引擎优化（SEO）的作用。当然，KG 不仅仅只有以各种类型的实体为节点的实体图，Facebook 则呈现了另一种关联人、事、物的兴趣图谱。相比谷歌的实体图谱，兴趣图谱节点和边的类型没有那么丰富，但是包含的节点数更多，稠密程度也更高。对于兴趣图谱上的搜索或问答也提出了不同的要求。为此，Facebook 专门提出了一种针对性的图搜索，内部的项目名叫 Unicorn（独角兽），有兴趣的可以去了解一下。

KG辅助问答交互



机器人及IoT设备的智能化：给万物都挂接一个背景知识库

对话式的信息获取更加需要精准度和可靠度，知识图谱对于提升用户体验更加不可少。



除了搜索，知识图谱也被广泛用于各种问答交互场景中。Watson 背后依托 DBpedia 和 Yago 等百科知识库和 WordNet 等语言学知识。类似地，Alexa 也依托其早年收购的 True Knowledge 公司所积累的知识库；Siri 则利用 DBpedia 和可计算的知识服务引擎 WolframAlpha；狗尾草公

司推出的虚拟美少女机器人琥珀虚颜则用到了首个中文链接知识库 Zhishi.me。伴随着机器人和 IoT 设备的智能化浪潮，智能厨房、智能驾驶和智能家居等应用层出不穷。无独有偶，百度推出的 Duer OS 和 Siri 的进化版 Viv 背后也都有海量知识库的支撑。

KG辅助决策



大数据与可计算大数据

...我们现在手可及的文本型健康医疗数据很多，能够获取的数据规模也都很大，但这是否就意味着这些数据可以供研究机构或商业机构做大数据分析？

.....要把自然语言的描述转换为**计算机可以计算的数据**，需要依赖自然语言处理中的很多基础技术，比如句子的**分词，实体识别，实体的归一化和链接等**

.....可以把**符号推理与统计学习结合起来**，让碎片化知识图谱的知识推理和深度学习的决策模型结合起来，我们称之为**Local Knowledge Powered Global Learning...**

”文本型医疗大数据，拿来就可用？”微软亚洲研究院闫峻博士

开放知识图谱

KG 也越来越多地被用于辅助决策。这里无外乎是通过对文本、多媒体和各种传感器产生的原始数据流建立更加规范的数据表达，通过语义抽取、数据链接形成更多机器可理解的语义知识，从而将原本异构分散的各种数据转变为机器可计算的大数据。通过可计算机的大数据，人们更容易发现领域或行业内原先不为人知的规律或一些有趣的模式，从而更好地做出决策。Plantir 就是这方面的成功应用典范，而 ImageNet 及 Visual Genome 等数据项目则极大地推进了图像语义理解和推理的进程。在构建用于辅助决策的知识图谱形成可计算大数据的过程中，如何将符号推理与统计学习有机结合起来，即碎片化的知识图谱上的推理和深度学习决策模型结合起来，形成所谓的 Local Knowledge Powered Global Learning 是非常有趣而富有挑战的研究课题。

当一个人听到或看到一句话的时候，他使用自己所有的知识和智能去理解。这不仅包括语法，也包括他的词汇知识、上下文知识，更重要的，是对相关事务的理解。

——Terry Winograd (自然语言理解系统SHRDLU作者)

Winograd Schema Challenge

I. The trophy would not fit in the brown suitcase because **it** was too **big (small)**. What was too **big (small)**?
Answer 0: the trophy Answer 1: the suitcase
II. The town councilors refused to give the demonstrators a permit because **they** **feared (advocated)** violence. Who **feared (advocated)** violence?
Answer 0: the town councilors Answer 1: the demonstrators

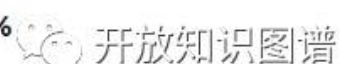
纯NLP : 50%



NLP+KB : 超过60%



及格线 : 90%



KG 也可辅助通用人工智能 (Artificial General Intelligence, AGI)，即在常识推理方面起到作用。过去人们常用图灵测试对机器的智能进行评估，近年来，Winograd Schema Challenge 逐渐进入大家的视线。这里举一个指代消解的例子。指代消解是一个经典 NLP 任务，旨在将代词指向具名实体。例如，The trophy would not fit in the brown suitcase because it was too big (small). What was too big (small)? 当我们描述 it 是 big 时，人们很容易理解这时候是在说奖杯 (trophy)；而当 it 与 small 搭配时，我们也很容易识别出在抱怨 suitcase 太小。这个看似非常容易的问题，却难倒了机器，这是因为人具有非常庞大的世界知识

(world knowledge) 和常识知识 (common-sense knowledge)。当我们仅采用 NLP 技术来努力理解并给出答案时，正确率仅 50%；当结合知识时，正确率提升到了 60%，而及格线是 90%。因此，我们离真正的通用智能还有很漫长的路要走，需要更多的技术突破和数据积累才能完成这项挑战。

3 需要什么样的 KG

03

需要什么样的KG

开放知识图谱

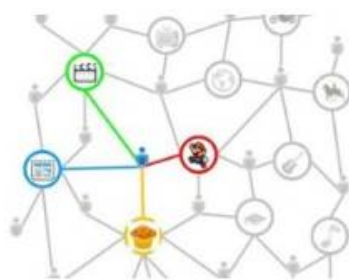
刚刚对 KG 的各种应用做了简单的介绍。结合 Chatbot，我们又需要怎么样的 KG 呢？

更加个性化的知识图谱



实体KG

开放领域
稀疏大图



兴趣KG

VS

机器人 KG

情感状态
喜好
技能
基本属性



社会关系

基本属性
职业状态

喜好
生活轨迹



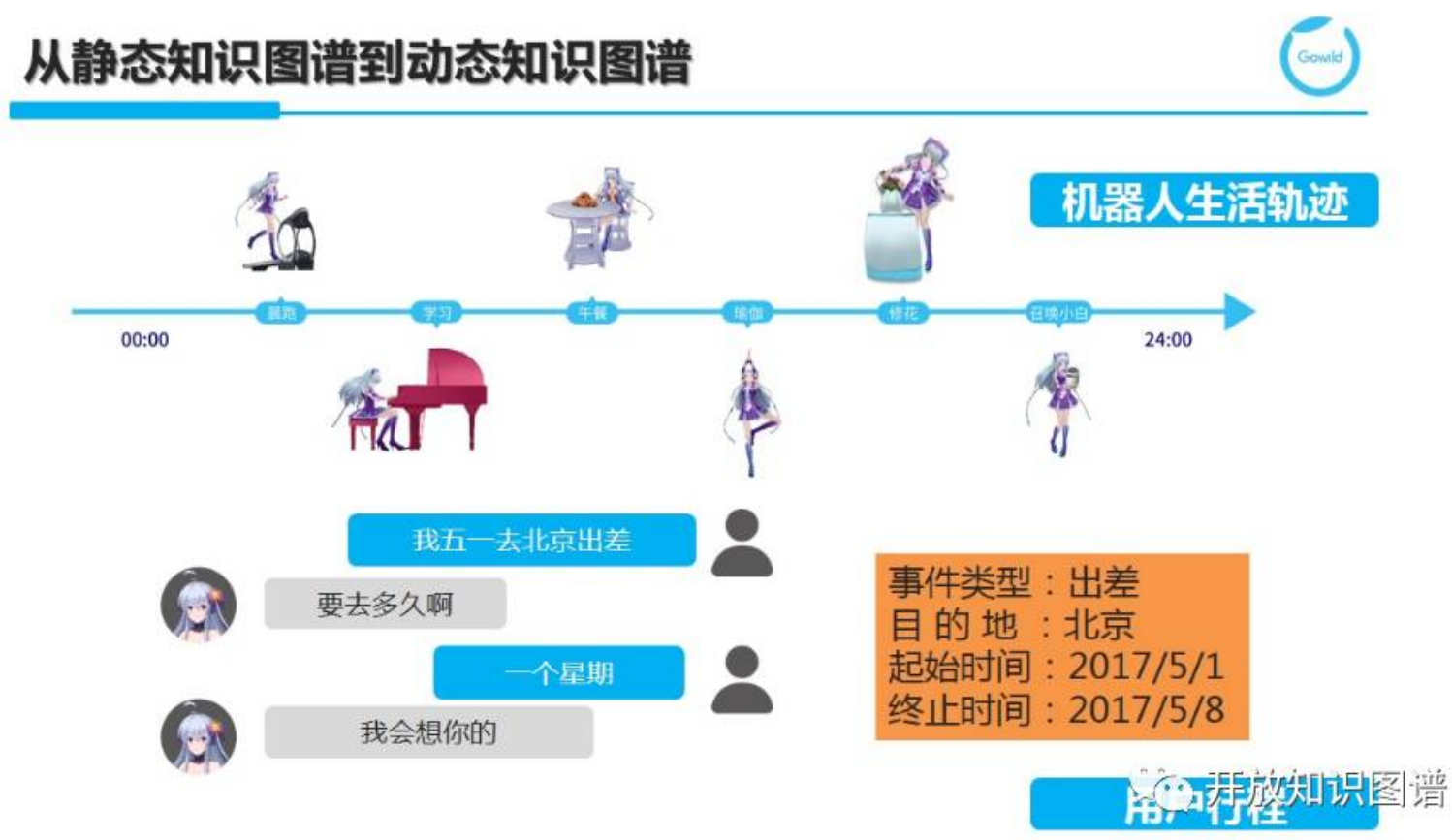
用户KG

个性化
稠密小图

开放知识图谱

首先，Chatbot 需要更加个性化的知识图谱。除了前面提到的实体 KG 和兴趣 KG 等开放领域的稀疏大图，我们也需要构建机器人 KG 和用户 KG

等个性化稠密小图。机器人或 Agent 需要图谱来建模和展示它的自我认知能力，而用户图谱则可被看作是更精细化的用户画像的知识表现。例如，机器人如“琥珀.虚颜”，有情感状态，喜好，技能等知识维度。同理，用户则需要表达其职业状态和生活轨迹等信息。需要强调的是，无论是个性化小图还是开放域大图，都不是独立存在的，需要将它们融合在一起，才能发挥更大的价值。机器人喜欢吃的食物则需要和实体 KG 中的食谱图谱关联，而与用户形成经纪人、好友等社会关系，同时爱好方面则和兴趣图谱又关联在一起，可以实现机器人社交、机器人-用户社交和用户社交网络的统一连接。

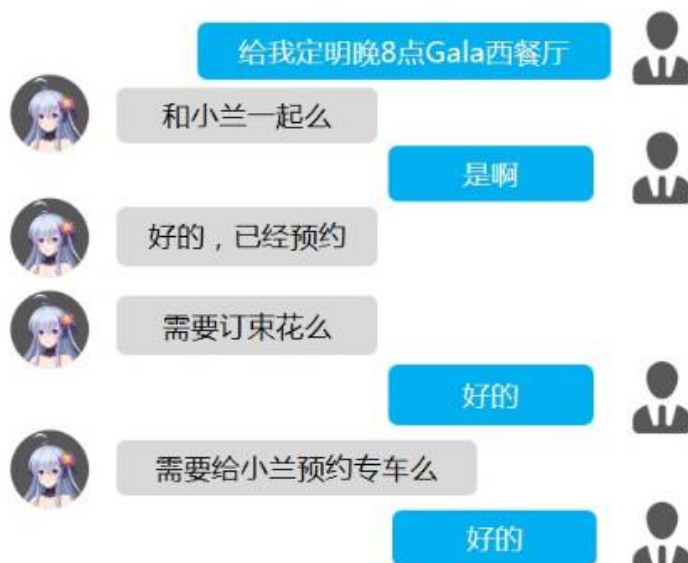


其次，我们的世界不仅仅是静态的，而是动态地反映各种事物在时空上的变化。因此，我们不仅仅需要刚刚谈到的静态图谱，而是需要思考如何表示和应用动态图谱。对于一个机器人，它从早到晚会做不同的事情，也就是有自己的生活规则。我们该如何刻画生活轨迹呢？这就需要在图谱中体现时态知识。另一个例子，用户行程，即对于用户图谱，需要记住用户各种已经发生、正在发生或即将发生的事件。图谱中的行程不仅仅是一个关系或属性，而是一个由多元（N-ary）组成的事件。我们需要定义多种事件类型，并刻画时间和空间两个维度。



第三，机器人不能只是冷冰冰的回答用户的问题或帮助用户完成特定功能。它需要感知用户的情感并在输出答案回复的同时伴随着相应的情感，这样才更加拟人化。我们发现，之前构建的知识图谱大多是客观的，即描述一些客观的事实。如何在结合个性化图谱时，能包括一些主观知识，进而刻画机器人或用户的情感元素。例如，用户说：“我心情不好”。这属于闲聊中的情感表达范畴。这时需要将用户当前的心情状态更新到用户图谱的对应维度数值中。相应地，机器人也会有自己的心情、体力，甚至和用户之间的好感度关联。当此时，机器人心情不错，同时和用户很亲密时，它就会主动关心用户。这样结合机器人和用户情感因素的动态回复会更加温馨和贴合场景。当在多轮对话时，用户进一步说：“来一首快乐的歌吧”。需要进一步结合音乐知识 KG（快乐作为歌曲的曲风或风格标签）和用户 KG 中的音乐偏好，推荐用户喜好的欢快的歌。

如何从二元关系扩展到动态服务（多元）？



如何刻画服务之间的关系，并支持多轮对话话题切换？

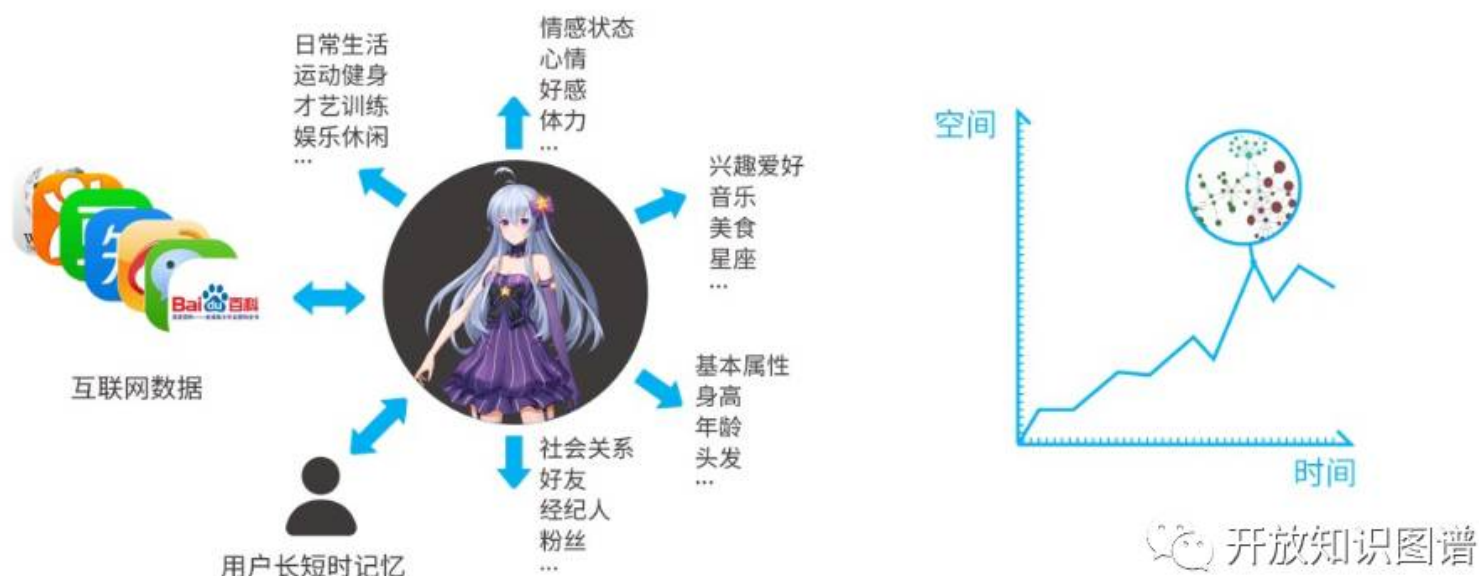
开放知识图谱

第四，我们发现聊天机器人为了完成很多功能需要对接外部服务或开放API。此时，图谱就需要从传统的关系型知识图谱（刻画二元关系）扩展到支持动态服务的动态图谱（刻画多元关系，事件属于服务图谱的一个特例）。另一方面，如何刻画服务之间的各种关系（如因果、时序依赖等）也是图谱扩展过程中需要考虑的。例如，当完成了订餐，会有很多Follow-up的服务（订花或预约车等）可作为后续服务被消费。建立这些服务之间的关联对于进行精准的多轮对话过程中的场景切换是非常有必要的。



我们接触世界的手段不仅仅是文字，而是结合图像、语音和文字等多模态来了解外部世界的。因此，我们所构建的知识图谱也应该从单纯文本自然扩展到多媒体知识图谱。而 ImageNet 和 Visual Genome 正是这方面的努力。但是这里我想强调的是对于用户图谱这样更新频度非常高且很稠密的 KG，多媒体知识的引入能帮助机器人从更多的维度来了解用户，并提供诸如 Visual QA 等潜在的问答服务。例如，小明正在和琥珀进行交互，通过摄像头识别出当前交互的用户是小明根据小明的图像与用户 ID 的关联，进一步得到其长短时记忆，了解到他在 4.20 到 23 号期间会去北京出差，而 4 月 24 号要和小兰共进晚餐。此时，通过用户图谱中的社交关系了解到小兰是小明的女友，当我们需要进一步了解小兰长什么样时，或者当小兰出现在琥珀面前时，需要可以认出小兰，这时也需要用到我们提到的多媒体知识图谱。

基于多源、异构的数据, 构建多类别、动态、个性化的知识图谱



总而言之，我们需要基于不同来源异构的数据来构建包含多类别且体现动态和个性化的知识图谱。这其中包括来自互联网的数据来刻画世界知识，用户数据来刻画画像知识，以及针对机器人的各种基本属性、社会关系、情感状态、兴趣爱好、以及日常生活等静态和动态知识。而我们得到的融合图谱是时空坐标中针对特定交互场景和时间节点的一个镜像。

知识图谱如何构建？

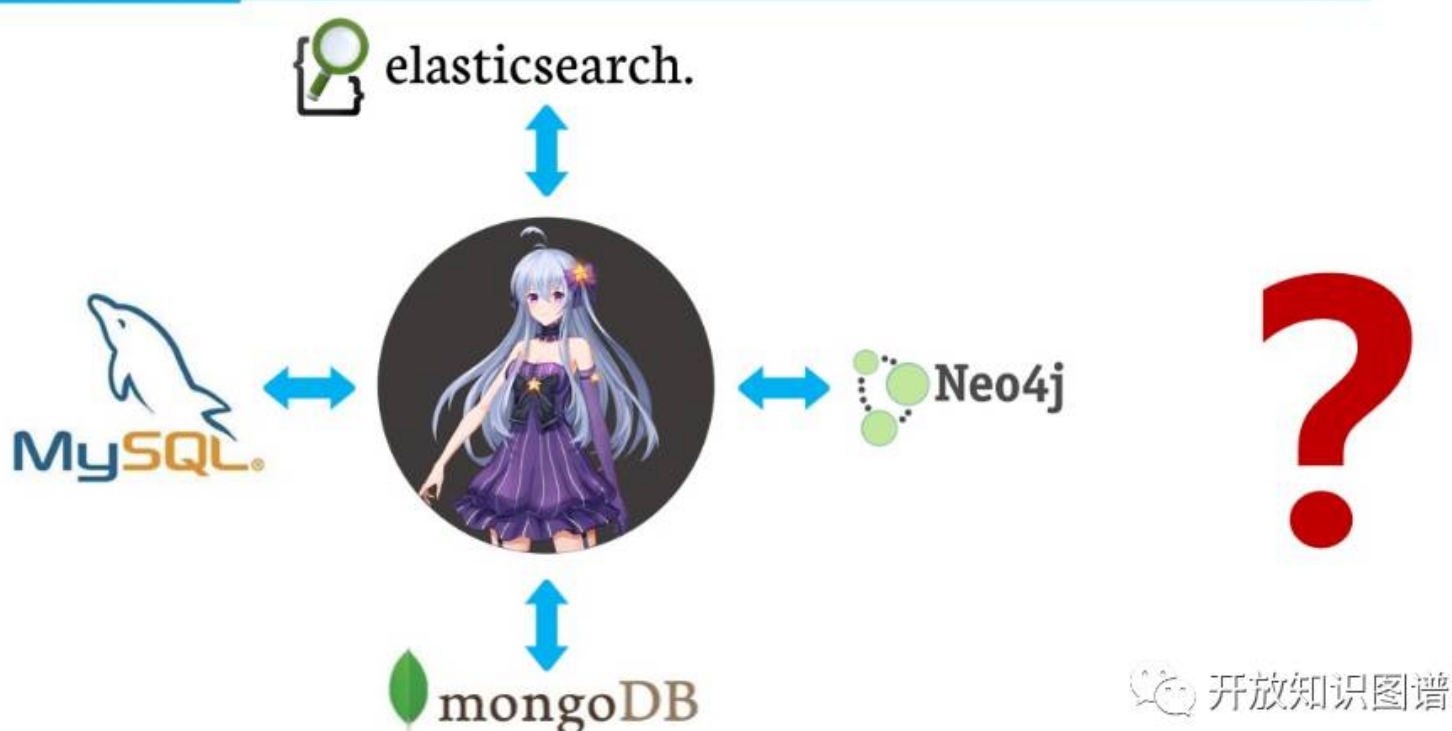


开放知识图谱

从更技术的角度来说，我们需要考虑知识图谱将如何构建。这里不仅包含如何结合文本、多媒体、半结构化、结构化知识、服务或 API，以及时态

知识等的统一知识表示。在此基础上，需要进一步考虑如何结合结构化（如关系型数据库）、半结构化（HTML 或 XML）和非结构化（文本、图像等）多源异质数据源来分别构建通用事实类（各种领域相关实体知识）、常识类、用户个人记忆类、机器人自我认知类和服务任务类知识库等。针对不同类型的数据和不同种类的知识构建，有相应的构建技术，如针对结构化数据的知识映射、针对半结构化知识的包装器（Wrapper），以及针对非结构化知识的文本挖掘和自然语言处理。文本挖掘充分利用 Web 或大规模语料库的冗余信息来发现隐含的模式；而自然语言处理更多是做各种知识抽取（开放或确定 schema 的）。为了得到融合的图谱，我们除了离线的多源异构的知识融合，还需要额外考虑服务任务类动态知识的对象绑定，这块工作往往是在线完成的，相当于根据不同的交互，在线动态扩充知识图谱并实例化的过程。

知识图谱如何存储？



所构建的知识图谱既包括事实类和常识类的静态全局大图、服务任务类动态图谱，也有对于每个用户不同的用户图谱和机器人 KG。随着用户数量的增加，用户图谱的数量也随之增加。这些图互相隔离，但每个均与全局图谱关联来提供个性化的聊天机器人的对话问答服务。每个用户图谱+机器人 KG 又随着交互不断填充和更新其中的节点和边，导致此类图谱的读写频度均非常高。面对这样的图谱，我们该如何选择存储方案呢？从工程应用的角度，我们更愿意站在巨人的肩膀上，选择一个现有数据库或几个数据库的组合来形成高效的图谱存储。

注意：这里所谓的存储，不仅仅是将知识存放的问题，而是考虑存储之后是否可以根据图谱的规模、读写特点和查询推理等基础在线操作的效率等多个因素来考量。MongoDB，作为面向文档（Document Oriented）的 NoSQL 代表，他支持无模式（Schemaless）的数据建模方式，即不要求一开始就将 Schema 都确定，而可以按需进行添加或修改。这对于需求经常变更或一开始对领域不是完全了解的情况下，支持自底向上方式的知识管理。不过 MongoDB 仅支持数据库级别的锁，写入速度受限。对于读并发的提高，可以使用基于数据分片（Data Sharding）的分布式版本。关系型数据库 MySQL 应用广泛，也被用于 Apache Jena（HP 开源的 RDF 数据库）中 TDB 的存储引擎。而 Elasticsearch 支持图谱上的简单模式（如单关系或链式）查询，适合如 Facebook Graph Search 或聊天机器人中大部分口语对话，因此也是面向聊天对话的知识图谱存储方案之一；Neo4j 是知名的图数据库，不同于 RDF 数据库，它的数据模型是属性图（property graph），基于图遍历（graph traversal）来实现各种查询功能，对于大部分熟悉面向对象编程的工程师来说非常容易上手。基于上述任何一个数据库或多个数据库的组合来满足知识图谱的管理都是工程做法。从研究的角度来说，希望做一个统一的存储和查询引擎，需要支持多租户（multi-tenant）环境下的海量个人和机器人知识图谱管理，以及融合个人图谱和全局知识上的查询分解、分布式环境下的查询路由和子查询执行，以及结果融合等操作。

4 KG 应用

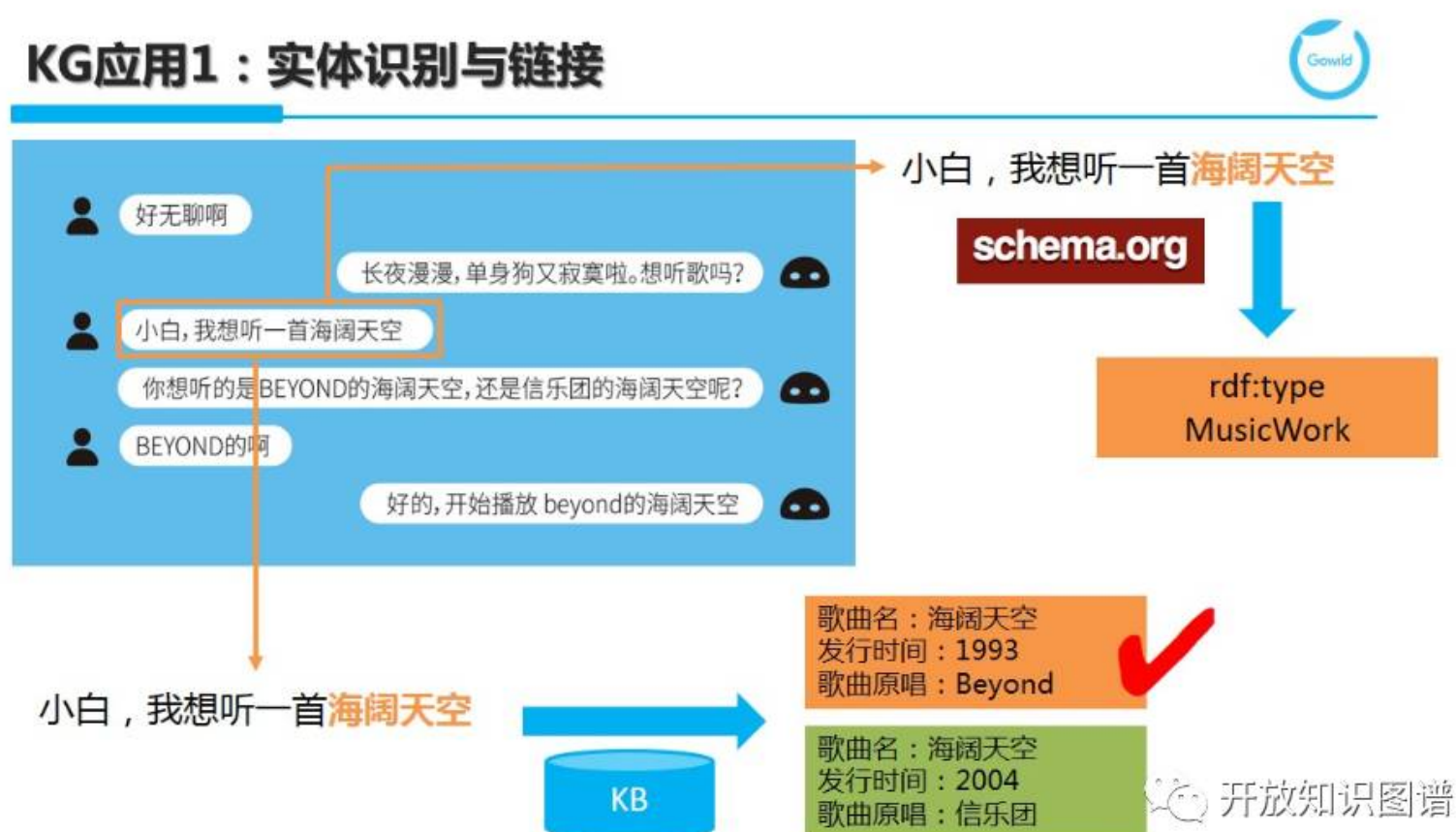
04

KG应用

开放知识图谱

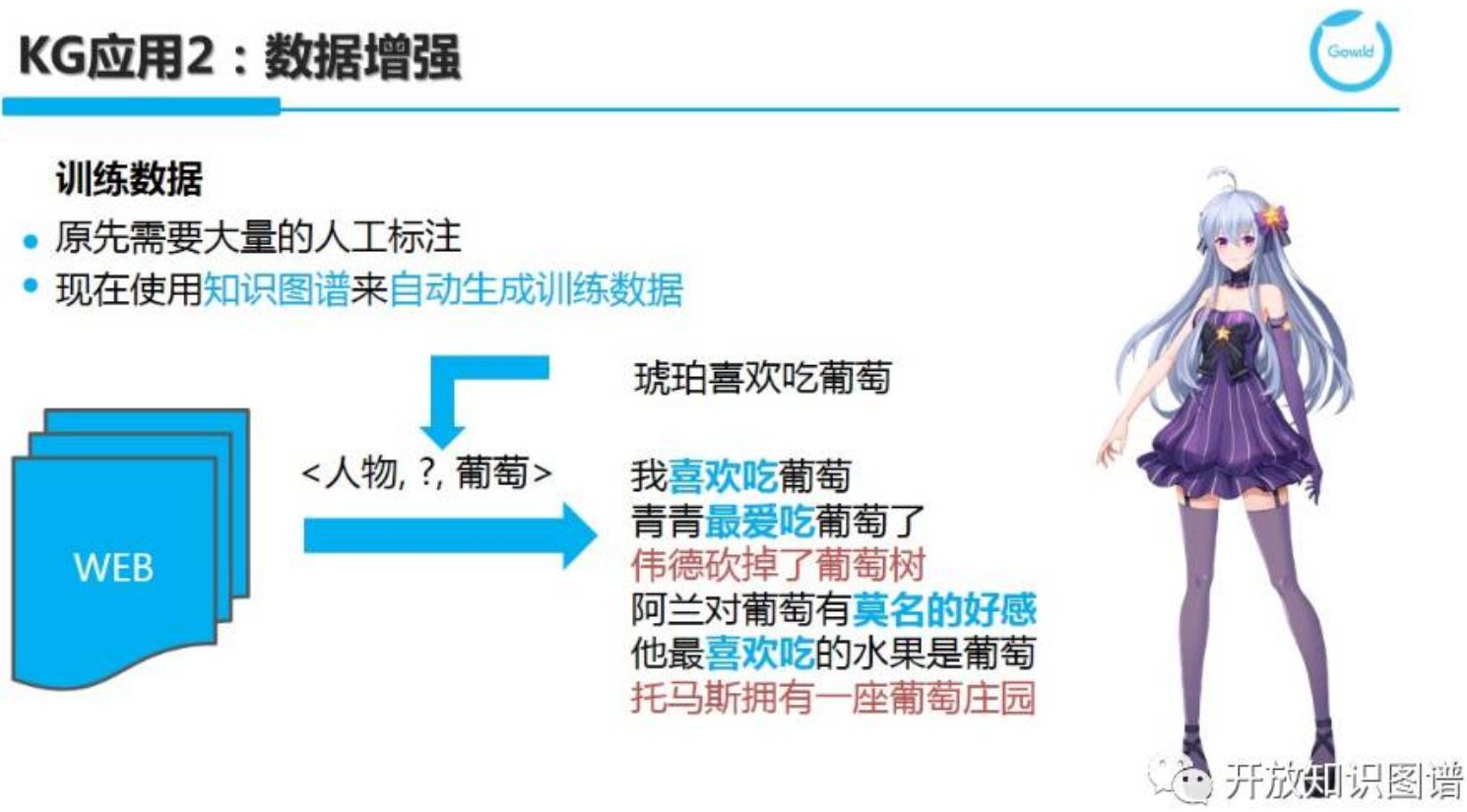
刚刚介绍了聊天机器人需要怎么样的知识图谱，以及相应的一些挑战。下面列举在 Chatbots 中 KG 的一些典型应用场景。

KG应用1：实体识别与链接



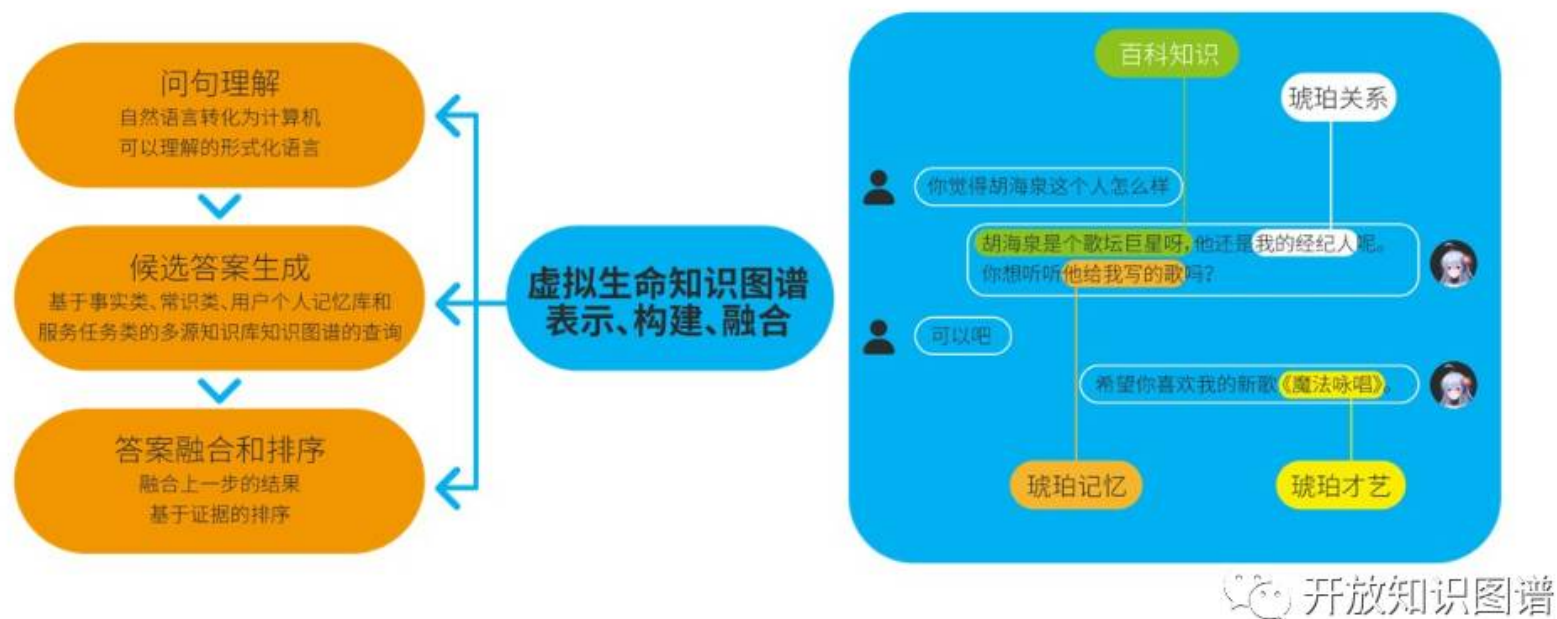
第一个应用叫实体识别和链接。实体识别称为 Named Entity Recognition，简称为 NER。在传统 NLP 任务中，仅能识别 PERSON（人物）、LOCATION（地点）、ORGANIZATION（组织机构）、DATE（时间日期）等有限类别。在实际应用中，NER 的主要挑战在于识别大量细粒度实体类型，比如以 Schema.org 作为实体类别的分类体系，这里有很多标注数据充足的大类，也有很多缺乏标注数据的小类，

如何保证在小类上的识别准确率。此外，分类体系是有层次结构的，如何保证底层的细粒度类别上有令人满足的识别率。例句“我想听一首海阔天空”中的“海阔天空”通过 NER 任务可以识别为是一个音乐作品。仅仅这样是无法执行对话意图“音乐点播”的，我们需要进一步将候选链接到知识图谱中的给定实体，这一过程称为 Entity Linking。这里的核心在于歧义消解，一般借助于候选周围的其他实体或用语作为上下位来帮助去歧义。如果如例子所示，仍然无法明确是哪个实体，可通过反问来引导用户来给出更明确的实体指引。在实体链接过程中，我们所面临的挑战在于如何应对新兴实体（Emerging Entity）和实体的新兴说法（各种新说法和别名）。



定的泛化能力和鲁棒性。使用 Web 作为关联的语料库，主要看中 Web 上描述比较多样化，且信息具有冗余性，可以在保证覆盖率的同时确保数据的分布贴近真实情况。然而对于以语音作为主要交互方式的口语化聊天对话场景，我们仍然需要考虑从 Web 语料上学习到的模式或训练得到的模型如何进一步迁移适配。

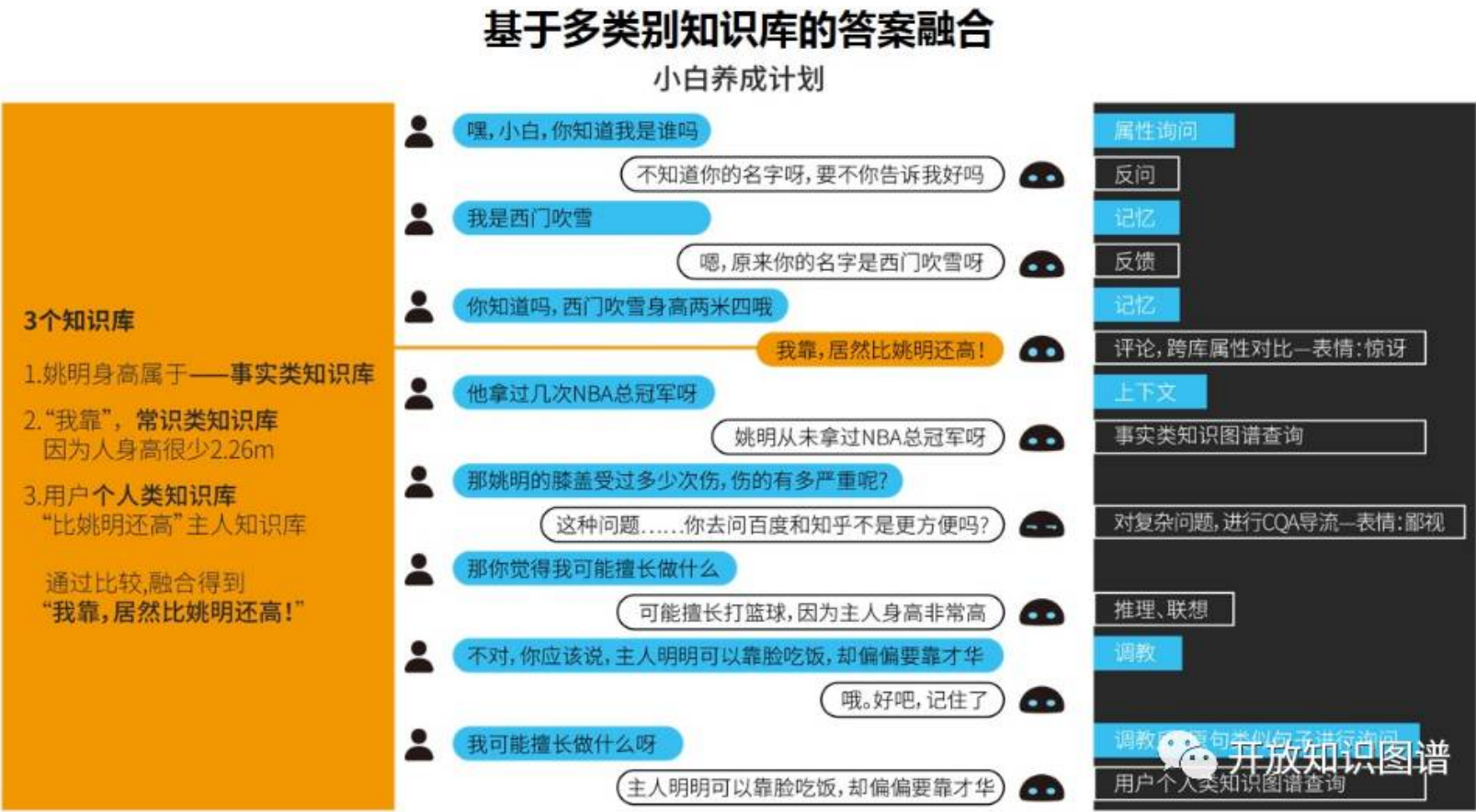
KG应用3：知识问答



第三个应用就是知识问答（KB-QA）。其中句理解的难点在于 NLU，而候选答案生成则与检索过程关联，至于答案融合和排序，则重点考虑各种基于证据的收集和学习排序算法。这里我们看一个真实的例子，比如说“你觉得胡海泉这个人怎么样？”，这是一个意见询问类查询（opinion query），此时可以有很多回答，为了使得答案的多样化，除了利用摘要技术（summarization）从百科站点中得到“胡海泉是个歌坛巨星呀”之外，通过机器人 KG 中的经纪人关系，可以显式表明琥珀和他的关系。更进一步，可以通过琥珀记忆和技能关联，主动推荐“海泉给琥珀写的歌”。当用户给予明确的回复时，将表演自己的才艺，即唱自己的歌。在我们所描述的知识图谱下支持问答，需要额外考虑：

- 1) 如何统一对实体、问句、图像、上下文进行统一的表示，映射到同构的语义空间中？
- 2) 知识库永远不可能是完备的，如何从 KB-QA 扩展到支持知识库和 Web 的混合 QA 场景下，并提供精准的数据源选择和语义解析？

3) 如何评估问句的复杂程度，并从单一知识库查询扩展到多知识库查询？



知识问答中非常有挑战的是 Multi-KB 环境下的问答。这对问句分解和知识源选择等都提出了更高的要求。更有挑战的是：不仅仅一句很复杂的话会涉及多个 KB，即使对于很简单的话，往往在聊天的多轮交互中，会逐步涉及不同方面的 KB，甚至需要在某个看似不经意的回复中用到某个 KB。在研制小白的多轮对话中，需要考虑属性询问、反问、记忆、反馈、基于跨库属性比对后的评论，基于上下文的问答、事实类知识图谱查询、对复杂问题的导流、推理联想，调教以及用户类知识图谱的查询等。例子“我靠，居然比姚明还高”就是一个多知识库问答之后的回复生成。其中，姚明身高属于事实类知识库、“我靠”等惊讶的回复，是通过常识知识库了解到很少有人身高超过 2.26 米，而通过用户个人知识，其身高数值比姚明还高，而返回“比姚明还高”的回复片段，最后通过融合，得到最终的返回。

采用了知识图谱、逻辑推理机，使得机器人具有记忆功能，
同时能够进行一定的推理能力，使得会话更加流畅自然。



第四个 KG 的应用就是联想和推理。这里我列举了三种推理，但实际情况下不局限于这三种。第一种是空间推理，比如说“桌子上面有电脑，电脑旁边有水杯”，然后问，“桌子上面有什么”，正确的回答是电脑和水杯。

桌子上有水杯是通过空间位置的判断得到的。空间推理在地理类问答和智能家居控制等应用中有非常广泛的应用。第二种是答案类型推理。答案类型 (Answer Type) 作为一种很重要的证据，对问答的准确性有很大的作用。这里的推理包括实例推理（如例子中乒乓球是一种运动）、上下位推理（白色家电是一种家电）和互斥推理（空调和电视没有交集）等。第三种是场景推理，即结合场景业务规则和相关常识知识进行一些联想。例如空调需要一定时间之后才能制冷，而用户在这段时间感到热时可以吃一些冷饮。除了这三类，冲突检测对于聊天机器人尤其是用户记忆很有价值。这里不仅包括前面提及的类别之间的互斥定义，还可以包括关系单值或数量约束，甚至形成很多由推理得到的事实和显式定义的事实组成的冲突关系链。这些对推理机的表达能力提出了更高的要求。

至此，我给大家介绍了聊天机器人的发展历程和知识图谱的应用，探讨了针对聊天机器人我们需要怎么样的知识图谱，并列举了知识图谱在表示、构建、存储、以及实体识别与链接、数据增强、知识问答，甚至推理与联想方面的机遇和挑战。我希望今天的演讲可以激发大家对知识图谱与聊天机器人结合的兴趣，一起参与到其中的各项研究中。大家有兴趣也可以扫描二维码，了解狗尾草公司在构建公子小白和 Holoera 琥珀虚颜虚拟生命

方面的更具体信息，谢谢大家！

谢谢！



 开放知识图谱

关于PaperWeekly

PaperWeekly 是一个推荐、解读、讨论、报道人工智能前沿论文成果的学术平台。如果你研究或从事 AI 领域，欢迎在公众号后台点击「**交流群**」，小助手将把你带入 PaperWeekly 的交流群里。



自然语言处理



计算机视觉



深度学习

Knowledge Sharing Made Easy.



▲ 长按扫描二维码关注