

# 基于大数据的旅游聊天机器人应用研究

文 / 王新宇

**摘要:** 本文对旅游聊天机器人进行了研究,设计了一个旅游聊天机器人,建立了相关的旅游业务知识库,编写了聊天机器人的APP软件,通过实验证明将聊天机器人应用于旅游业,可以更好地推动旅游业的发展。

**关键词:** 旅游顾问;聊天机器人;旅游大数据

DOI:10.14011/j.cnki.dzsw.2016.11.022

## 引言

近年来,国家及各地方政府纷纷出台政策措施,以促进电子商务持续健康发展为目标,着力完善发展环境,重点加快应用推广,电子商务得到蓬勃发展,据《2014年度中国电子商务市场数据监测报告》中的数据显示:截止2014年12月,中国电子商务市场交易规模达13.4万亿,同比增长31.4%。其中,B2B电子商务市场交易额达10万亿元,同比增长21.9%。网络零售市场交易规模达2.82万亿元,同比增长49.7%。电子商务服务企业直接从业人员超过250万人,而由电子商务间接带动的就业人数,已超过1800万人。移动电子商务交易规模达到9285亿元(2013年达2731亿元,同比增长240%)。

易观国际的大数据也显示,2014年中国在线旅游移动端市场规模达到1247.3亿元(在线旅游总规模2798.2亿元)。根据数据,易观国际认为:2015年在线旅游行业PC端和移动端将出现逆转,移动旅游客户规模将首超PC端客户,用户规模约占全部交易的64%,而到了2017年,所其预测,在线旅游移动端占比将达到76%。另外,从中国电子商务研究中心的数据可以看出,2015年第一个季度,中国在线旅游市场通过移动端完成的交易额达到501.5亿元,移动客户端在线旅游市场的渗透率达到52.9%。

本文提出通过将聊天机器人应用于旅游电子商务,促使旅游电子商务更好地推动旅游业的发展。

## 1、聊天机器人

聊天机器人,是一种通过人类的语言同用户进行交流和对话的智能系统,目前广泛应用于诸多领域,如在银行的在线客服领域,客服机器人可以为客户解答一些简单的业务问题;在电商领域,机器人可以为客户推荐商品,为食客寻找美食,为游客查询最佳游玩路线等等。一款优秀的聊天机器人不只是仅仅回答用户提出的某个问题,还可以进一步和用户进行人性化的交流,它能够像朋友一样理解用户的需求,按Facebook的设想,在未来,聊天机器人完全可能取代企业客服和各类APP软件。

目前的聊天机器人根据商业应用目的不同,可以分为两种:纯粹以陪用户聊天打发时间为目的的娱乐型机器人、帮助用户获得所需信息和知识的信息获取型机器人。对于前者来说,它的唯一目的就是陪聊,在聊天的过程中,例如:SimSimi

机器人会讲出各种卖萌、搞笑的话语,以“语出惊人”来娱乐用户。而对于后者,一般来说,用户对它的功能要求更为严格,主要表现在:机器人提供信息的相关性、可靠性和准确性,有时,用户可能还要求机器人去直接完成某项任务,如:医院挂号、订购商品、接通某位朋友的电话等等。本文主要研究的是信息获取工具型聊天机器人在智慧旅游中的应用。

## 2、打造旅游聊天机器人

### 2.1 旅游聊天机器人设计

设计旅游聊天机器人,并不一定从头开始编写程序,目前有许多开放平台供企业进行二次开发,图灵机器人就是一个免费的智能机器人开放平台(每天可以免费提供5000条问答,完全满足一般实验要求),我们做旅游聊天机器人实验,可以直接调用该平台的API接口,具体的方法可以参考图灵机器人的相关开发手册。

我们课题组实现了一个旅游聊天机器人“游游”,尝试将人工智能引入旅游咨询系统,“游游”从设计初,课题组就放弃了“一切从头开始,完全自行开发”的办法,而是利用成熟免费的第三方平台(图灵机器人平台),将研究工作的重点放在如何建立知识库,如何从网络上提取旅游信息等重点内容上。另外,知识库中添加了当前的网络热词和一些俏皮话,让“游游”一改常规旅游客服专员严肃商务形象。旅游聊天机器人“游游”实验成功,为课题增加了亮点。

### 2.2 知识库的建立

通过调用图灵机器人的API接口,就可以实现聊天机器人和用户的正常对话,因为图灵机器人本身已经提供了智能聊天的功能,但图灵机器人缺少专业的旅游业务知识,所以我们必需通过建立旅游业务知识库的方式,实现专属旅游的聊天机器人。

在图灵机器人平台,系统提供了自定义NLP知识库的功能,实际上就是以<模板,输出>这种二元组的形式存储知识。如果用户的输入匹配了某种特定的模板,机器人就会输出相应的内容。由于图灵机器人已经对匹配的问题进行了中文分词及关键词的提取,可以进行模糊匹配,用户提出的问题只需要和固定的问题模板相似,机器人就会给出相应的内容,不需要和固定问题模板一字不差。所以,我们在建立知识库问题模板时,需要对问题模板进行反复测试,用中文分词软件先进行分词处理,找出关键词,发现规律,这样就可以匹配更多的问题。

旅游知识库中的业务知识,主要来自两个方面:

第一个方面是来自课题组中有旅游专业教师,他们不仅有丰富的旅游业务知识,而且还拥有长期在一线旅行社从事工作的经验,能够灵活及时地完成游客关于游、购、娱、吃、住、行的信息和旅游线路的整合,安排出较好的旅游线路,将专业

★基金项目:南京旅游职业学院基金资助项目(2015YKT10),大数据时代旅游数据挖掘与应用研究;南京旅游职业学院“青蓝工程”基金资助(2013YQL008)。

教师的业务知识采集,并录入到知识库中,可以为机器人提供专业的业务知识。当然,专家也不是万能的,专家有其熟悉的领域,在如今“互联网+”时代,专家旅游知识也不可能及时覆盖新出现的旅游热点信息。

所以,第二个方面就是从网上收集整理,进入大众旅游时代后,互联网上发布的各种旅游信息大量涌现。旅游网络论坛、旅游微博等平台存在由游客撰写的各种旅游数据,其中包含了对旅游景点的讨论、对旅游线路问题的问答,还有个人的游记等等,这些内容都可以作为知识库中的数据,利用这些网上已经存在的知识,可以大大提高知识库构建速度,同时互联网活跃性,也可以确保聊天机器人的聊天用语紧随时代发展,学会一些新词、热词、俏皮话。

本次实验,课题组通过人工的方法,从各大旅游网站和旅游论坛收集、整理出有用的业务知识,最终将这些业务知识录入知识库。人工从网上收集数据是比较费时费力的,今后将研究让程序从网上自动收集和提取有用信息。

## 2.3 应用方法

### 2.3.1 微信公众号中嵌入机器人

如果已经拥有了微信公众号,就可以很方便的将旅游机器人直接嵌入到微信公众号应用中,这种方法可以省去独立开发APP程序和APP程序的宣传和推广。非常方便和实用,可以快速的构造出自己的机器人,缺点是有时不满足企业的特殊需要。

### 2.3.2 单独设计APP程序

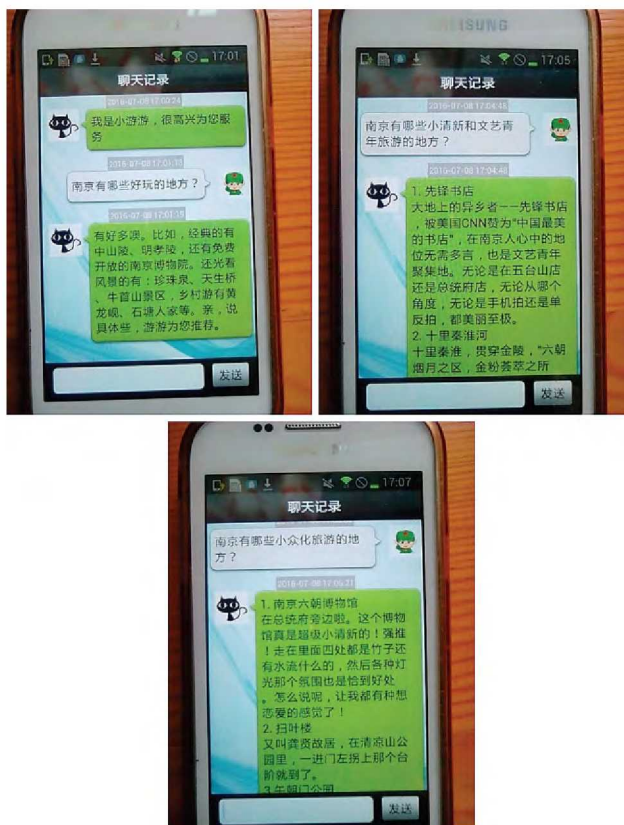


图1 旅游聊天机器人APP界面

如果不用微信公众号平台,在手机上应用就需要另外编写APP程序了,因为使用的智能手机大多数是基于Android平台智能手机,为了便于实际运行和测试,本次实验的是Android平台,使用的编程环境用是Eclipse+ADT插件,ADT全称是Android Development Tools,是Google提供的一个Eclipse插

件,用于在Eclipse中提供一个强大的,高度集成的Android开发环境。Eclipse中安装了该插件后,不仅可以联机调试,还可以用模拟器模拟各种手机的事件、分析程序的性能等。如图1所示,为本次实验的APP程序演示结果。

## 3、旅游聊天机器人应用

### 3.1 旅游电商客服及旅游顾问

旅游电商客服,是伴随着旅游电子商务的开展而形成的,最早是以呼叫中心(CCT)的形式出现的,主要进行帮助游客进行酒店、机票、景区门票预订,并接受客人的投诉等工作,呼叫中心最鼎盛时,一个大型旅游网络公司可能拥有几千个CCT客服专员的岗位,近两年来,随着网上自助预订和手机APP预订的兴起,简单的呼叫中心客服岗位需求有所下降,但旅游电商客服又有一个新的岗位:旅游顾问。这是旅游电商客服发展到一定程度,而细分出来的一个岗位,旅游顾问可以为游客出行前提供咨询服务、客人在旅游过程中如有问题和困难可以和旅游顾问沟通,能得到旅游顾问的解答,并可以及时解决问题等等。一般来说,旅游顾问的旅游知识储备和技能要求比普通电商客服专员要高。

目前电商客服专员一般由人担任,也有部分企业开始使用智能聊天机器人来代替人工客服。

### 3.2 当前旅游电商人工客服的不足

旅游业是一个复杂的行业,与不少行业密切联系,概括来说,旅游就是:游、购、娱、吃、住、行六个要素,这些要素涉及了景区、商店、酒店、餐馆、客运公司等众多企业,它们提供的产品和服务对大旅游来说是不可缺少的。作为专职旅游电商客服,不但需要掌握丰富的旅游业务知识,还需要有从事过计调工作的经验,好的旅游电商客服,能够灵活及时地完成游客在旅游目的地关于旅游六要素的信息和旅游行程的整合,并在实际行程中安排出较佳的线路。但是,目前的旅游电商客服仅在某一旅游企业工作,能获取的旅游资讯和提供的服务一般局限在本企业的旅游产品范围内,在这种情况下,游客不容易获得经济实惠的行程方案,为了得到好的方案,游客们不得不在出行前自己做好“功课”,花费时间和精力在网上搜索旅游目的地信息,如:景区附近有哪些好的酒店和餐馆,查看网上对景区、住宿、餐馆的点评,临时“客串”一回旅行社的计调的角色,整合自认为最优的“出行方案”,因为人的精力和时间有限,自助游的游客无法通过网络获得完整、详细的信息,只能获得一些零散的,不成体系的旅游信息,除此之外,现在网上的旅游营销软文较多,如果游客在互联网中仅仅通过看帖子、微博、微信等信息,难以客观地甄别旅游产品的质量,所以普通游客试图通过搜索浩瀚的互联网的信息设计最佳旅游行程的想法,并不实际。

### 3.3 旅游聊天机器人的优势

旅游聊天机器人可以有效的解决人工旅游电商客服的不足之处,通过收集、整理、分析、归纳旅游大数据,并且建立相应的知识库,旅游聊天机器人可以很好的将游、购、娱、吃、住、行六要素涉及的旅游信息整合,只要建立的旅游数据库足够大,机器人就可以回答所有的旅游问题,目前,现成的知识库如:百度知道、知乎,新浪爱问,维基百科等等,为此提供了数据上的支持,事实上,现有的不少聊天机器人的知识库就引用上以上的数据库。同时在与游客聊天和交流过程中,机器人也可以不断的学习新知识,及时补充新信息,这样就可以为游客提供优质的旅游顾问服务。

(下转第87页)



重要的IT实践,详细描述了可适用于任何组织的全面的清单、任务、规程和职责,这些实践在ITSM中被定义为覆盖IT服务组织大部分活动的流程体系,使之成为IT组织经常性的有益参考,组织可以通过使用ITSM来为其自身设定改进目标,并得以不断成长和逐渐成熟<sup>[5]</sup>。实施ITSM涉及到很多管理模式和管理理论,包括现代企业管理、人力资源管理、系统工程、信息技术、风险管理、过程管理和优化、客户关系管理、质量工程、工程技术、度量分析技术、统计技术、项目管理、创新管理、战略管理等多个学科和领域。可以说,IT运维服务管理已经成为IT服务企业的核心管理技术,从初期的概念提出到实施标准的成功开发和在世界范围内的广泛应用,已经成为受到重视、相对独立和发展迅速的管理领域。目前,已经形成了包括IT服务管理标准制定、咨询、培训、认证、软件开发和服务实施的完整产业价值链。

随着现代信息服务业的高速发展,目前国内IT运维服务企业的发展也非常迅速,对IT运维服务工程师和IT运维服务管理人员的需求一直较大,在笔者以前工作过的两家IT运维服务企业,平均每年对相关人才的需求总量都在50~200人之间,加上市场上众多的IT运维服务企业以及其他类型的IT服务企业(如软件技术服务企业和系统集成服务企业等),每年的人才需求总量很大。另外,IT运维服务管理的知识也是传统非IT企业在信息化建设中所需必需的,如果加上这部分人才需求,总量更大。因此,IT运维服务管理在逐渐发展成为一个独立的管理领域后,其人才需求总量一直呈现不断放大的趋势。

### 3、IT运维服务管理课程的定位

尽管IT运维服务管理在企业界受到越来越多的重视和应用,但是相关的课程开发却一直局限于企业界,主要由专门从事IT服务管理咨询和培训的公司完成相关培训课程的开发和推广。根据我们的调查,在各大中专院校里并没有及时开发出《IT运维服务管理》这样一门紧跟时代发展步伐的专业课程,从而导致市场上急需的IT运维服务管理专业人员主要是通过市场上的一些培训机构完成《IT运维服务管理》的课程培训,培训合格后上岗工作。因此,如果在高校研究并开设这方面的课程,并保证教学达到市场培训机构的同等效果,将能够以更低的社会成本较好地满足这方面的人才需求。

在此背景下,我们专门组织教研团队进行了《IT运维服务管理》的课程开发研究和试验教学,较成功地将这一管理知识



图2 IT运维服务管理课程的定位

引入到了工商企业管理专业和计算机信息管理专业相关专业方向的课程体系中,充实并完善了相关专业的知识结构,并为学生提供了更多的就业方向选择,取得了较好的效果。在课程开发过程中,我们从一开始就明确将课程定位于“IT服务企业的核心管理技术”这一位置,由于IT服务企业在现代服务业和现代信息服务业中的重要地位,所以这一课程的研发和教学便具有了很高的重要性和定位(如图2所示),也就受到了教师和学生的高度重视。目前,在我校的工商企业管理(ERP应用师)和计算机应用(企业信息管理师)两个专业方向中,本课程的内容均已成为其知识体系的有机组成部分。

本文通过理论和实践两个方面论述了,IT运维服务管理课程是IT服务业从业人员必须掌握的一门核心管理课程,它是IT服务企业的核心管理技术,进而也成为了现代信息服务业和现代服务业的重要基础支撑技术之一。因此,针对本课程的研究与实践对于产业发展和职业教育都具有非常重要的意义。

#### 参考文献:

- [1] 李华北,刘颖. IT服务管理及CMMI-SVC实施[M]. 北京:电子工业出版社,2012.
- [2] 匡佩远. 信息服务业:定义和统计框架. 统计教育,2009,第5期(总第116期).
- [3] 国家信息技术服务标准工作组. ITSS系列培训(IT服务工程师)[M]. 北京:电子工业出版社,2012.
- [4] 国家信息技术服务标准工作组. ITSS系列培训(IT服务项目经理)[M]. 北京:电子工业出版社,2012.
- [5] (荷兰)博恩(Jan van Bon). IT服务管理——基于ITIL的全球最佳实践[M],章斌译. 北京:清华大学出版社,2006.

作者简介:冯万贵,广州工程技术职业学院专任教师,副教授、高级工程师,工商管理硕士,工商企业管理(ERP应用师)专业负责人,研究方向:企业信息化管理、战略管理和运营管理。

(上接第62页)

### 4、进一步研究的展望

在本次研究中,知识库的建立,主要是靠人工收集和编写,这样做的主要问题是,人工在网上收集、整理、录入知识库,需要大量的时间和精力,所以构建成本很高,到目前为止,课题组构建的知识库还很小,完成的功能也有限,所以课题组将进一步研究如何自动从网上提取知识,让聊天机器人进行自我学习。另外,目前课题组实验中的聊天机器人,还只能进行聊天、知识问答,线路推荐,不能进行诸如:酒店预订、从网上订景点门票等工作,这也是课题今后研究的方向。

#### 参考文献:

- [1] 王新宇. 基于Android系统的移动学习平台的设计与实现[J]. 电脑知识与技术,2015(19):70-73
- [2] 图灵机器人开发者中心. 图灵机器人接入文档[EB/OL]. <http://www.tuling123.com/html/doc/docs.html>
- [3] 杨凯翔. 聊天机器人知识挖掘方法研究[D]. 华南理工大学,2015(3)
- [4] 姚飞,张成里,陈武. 清华智能聊天机器人“小图”的移动应用[J]. 现代图书情报技术,2014,(Z1):120-126.
- [5] 冯德虎. 基于ALICE的研究生招生咨询智能聊天机器人研究与实现[D]. 西南交通大学,2013
- [6] 蔡静. 网络客服——贵州发展智慧旅游的关键[J]. 价值工程,2015(12):200-201.

作者简介:王新宇,硕士研究生,南京旅游职业学院讲师,研究方向为旅游电子商务、软件工程。