# 聊天机器人研究背景及发展现状综述

**摘要**  聊天机器人是自然语言处理和人工智能领域的重要研究方向之一。近年来，聊天机器人系统的研究受到了广泛的关注。本文综述了近年来关于聊天机器人的研究，并从研究的背景和意义到目前国内外相关技术的研究现状。详细介绍了聊天机器人的分类、系统框架、构建方法、技术评价等。总结了目前研究中存在的不足，展望了未来的发展趋势。

**关键词 聊天机器人 生成式模型 检索式模型 深度学习**

**引言**

## 1 国内外研究进展

图灵通过图灵猜想给聊天机器人种下了一颗种子。从上世纪60s到上世纪末，chatbot大约经历了三个重要的历史时期。

      第一个时期（1966年），麻省理工学院（MIT）的约瑟夫·魏泽鲍姆（Joseph Weizenbaum）开发的聊天机器人ELIZA[1]，用于在临床治疗中模仿心理医生。值得注意的是尽管ELIZA的实现技术仅为关键词匹配及人工编写的回复规则，但魏泽鲍姆本人对ELIZA的表现感到吃惊，随后撰写了《Computer Power and Human Reason》这本书，表达他对人工智能的特殊情感。

      第二个时期（1988年），加州大学伯克利分校（UC Berkeley）的罗伯特·威林斯基（Robert Wilensky）等人开发了名为UC（UNIX Consultant）[2]的聊天机器人系统。UC是一款帮助用户学习怎样使用UNIX操作系统的聊天机器人。它具备了分析用户的语言、确定用户操作的目标、给出解决用户需求的规划、决定需要与用户沟通的内容、以英语生成最终的对话内容以及根据用户对UNIX系统的熟悉程度进行建模的功能。

      第三个时期（1995年），受到ELIZA聊天机器人的启发，理查德·华勒斯（Richard S. Wallace）博士在1995年开发了ALICE[3]系统。并于1998年开始开源，目前全世界有超过500个开发者为ALICE项目贡献代码。值得注意的是，随着ALICE一同发布的AIML（Artificial Intelligence Markup Language）目前被广泛应用在移动端虚拟助手的开发中。尽管ALICE采用的是启发式模板匹配的对话策略，但是它仍然被认为是同类型聊天机器人中性能最好的系统之一。

21世纪的到来为聊天机器人的发展创造了更大的发展空间。随着人工智能技术的崛起产生了一种基于深度学习的聊天机器人算法，很多国内外顶级的互联网公司也纷纷推出了自己的产品，比如苹果公司 Siri、亚马逊Alexa、谷歌 、微软小冰、百度的“度秘”等等。 详细

## 2 聊天机器人分类

进入本世纪后，chatbot机器人随着人工智能的兴起有了长足的发展，各大互联网公司、各个高校和研究院都取得了成就。但总的来说，也可以大致的分为三种。检索式、生成式和知识图谱式。其中，知识图谱式的chatbot比较难以实现，但是是有前景的。

       第一种，检索式对话生成chatbot，检索式的代表性技术[4]是在已有的人人对话语料库中通过排序学习技术和深度匹配技术找到适合当前输入的最佳回复。这种方法的局限是仅能以固定的语言模式进行回复，无法实现词语的多样性组合。

       第二种，生成式对话chatbot，生成式的代表性技术[5，6]则是从已有的人人对话中学习语言的组合模式，是通过一种类似机器翻译中常用的“编码-解码”的过程去逐字或逐词地生成一个回复，这种回复有可能是从未在语料库中出现的、由聊天机器人自己“创造”出来的句子。

       第三种，知识图谱式的chatbot目前还没有典型的代表，这一类机器人将最常用的语言和简单逻辑通过AIML模版处理，通过知识图谱（可以自动生长）的逻辑推理来处理复杂的逻辑，运用深度神经网络来将复杂的逻辑用人类的语言表达出来。

总结以上模型均有优缺点。对于基于检索技术的模型，由于使用了知识库且数据为预先定义好的，因此进行回复的内容语法上较为通顺，较少出现语法错误；但是基于检索技术的模型中没有会话概念，不能结合上下文给出更加[智能]的回复。而生成式模型则更加[智能]一些，它能够更加有效地利用上下文信息从而知道你在讨论的东西是什么；然而生成式模型比较难以训练，并且输出的内容经常存在一些语法错误（尤其对于长句子而言），以及模型训练需要大规模的数据。深度学习技术都能够用于基于检索技术的模型和生成式模型中，但是目前的研究热点在生成式模型上。深度学习框架例如Sequence to Sequence非常适合用来生成文本，非常多的研究者希望能够在这个领域取得成功。然而目前这一块的研究还在初期阶段，工业界的产品更多的还是使用基于检索计算的模型。

## 3聊天机器人系统框架

## 四、聊天机器人技术评测及评测指标

[1] Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine[J]. Communications of the ACM, 1966, 9(1): 36-45.

[2] Wilensky R, Chin D N, Luria M, et al. The Berkeley UNIX consultant project[J]. Computational Linguistics, 1988, 14(4): 35-84.

[3] http://www.alicebot.org/

[4] Z. Ji, Z. Lu, H. Li. An information Retrieval Approach to Short-text Conversation. arXiv: 1408.6988, 2014.

[5] Ritter A, Cherry C, Dolan W B. Data-Driven Response Generation in Social Media.[J]. Empirical Methods in Natural Language Processing, 2011:583-593.

[6] L. Shang, Z. Lu, and H. Li. Neural Responding Machine for Short-text Conversation. In Proceedings of ACL, 2015.