arXiv:2302.14472v3 [cs.MM] 3 Mar 2023

电视陪伴机器人：用户体验分析

DONGHUO ZENG* 和 JIANMING WU*，KDDI Research, Inc.，日本
GEN HATTORI，KDDI Research, Inc.，日本
YASUHIRO TAKISHIMA，KDDI Research, Inc.，日本

看电视不仅能获取新闻信息，也为不同世代的人们提供了交流的机会。随着智能手机、个人计算机和互联网的普及，在电视机前交流的机会也可能随之减少。这导致了一些问题，例如缺乏自我控制能力，以及沟通技能发展不足等，使人们进一步远离面对面交流。本文提出了一种具备开放域聊天能力的电视陪伴机器人，可围绕 50 个日常生活话题进行交流。该机器人包含两种模式："电视观看模式"和"对话模式"。在"电视观看模式"下，机器人首先从电视节目中提取关键词，然后根据提取出的关键词生成自我披露话语，仿佛正在享受电视节目。在"对话模式"下，机器人以相同方式利用关键词生成提问话语，随后采用由多个对话引擎组成的基于主题的对话管理方法，围绕电视节目展开丰富的对话。我们进行了初步实验，结果显示，来自三个组别的所有参与者都乐于与机器人交谈；对于"他们对机器人有多感兴趣"这一问题，评分为 6.5/七级量表。这表明，电视陪伴机器人的对话功能有望让我们的日常生活更加愉快。在初步实验分析的基础上，我们进一步开展了参与者人数更多的实验，并将参与者分为两组：不使用机器人的对照组和使用机器人的干预组。结果显示，人们更喜欢与机器人交谈，因为机器人会带来更多"愉悦""放松"和"有趣"的体验。

CCS概念： • **计算机系统组织** → **嵌入式系统**； 冗余；机器人学； • **网络** → 网络可靠性。

附加关键词 W ords 和短语：电视陪伴机器人、KACTUS、基于话题的对话管理、提升交流 n

ACM参考文献格式：
Donghuo Zeng、Jianming Wu、GenHattori 和 Yasuhiro Takishima。2022。电视陪伴机器人：用户体验分析。载于 *AMX '22*：*ACM*交互媒体体验，*2022*年*6*月*22*日至*24*日，葡萄牙阿威罗。 ACM，纽约州纽约市，15 页。
https://doi.org/XXXXXXX.XXXXXXX

1 引言

电视已成为日常生活中无处不在的一部分，大多数家庭的客厅里都摆着一台电视。在客厅看电视，不仅可以获取新闻信息、享受娱乐节目，还能了解一些日常生活知识，同时也能促进不同世代的家庭成员或朋友之间的交流。在这种情况下，人们有机会围绕所看的电视节目展开对话，并体验喜怒哀乐。另一方面，随着智能手机、个人电脑和互联网的普及，如今人们可以从电视以外的其他来源获取信息，因此

*两位作者对本研究的贡献相同。

特此授权免费制作本作品全部或部分内容的数字副本或纸质副本，供个人或课堂使用，但不得为营利或商业目的制作或分发副本，且副本必须包含本声明及首页上的完整引文。必须尊重本作品中由 ACM 以外的其他方拥有版权的部分内容。注明出处后允许摘录。其他复制、再版、发布到服务器或重新分发至列表的行为，须事先获得特定许可和/或支付费用。请向 permissions@acm.org 申请许可。© 2022 美国计算机协会。稿件已提交至 ACM

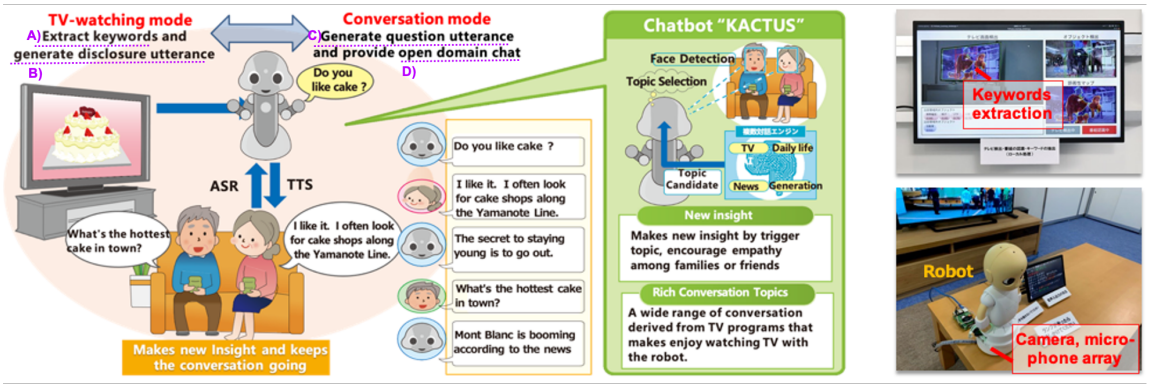


图1. 电视陪伴机器人的概念。

在电视以外的其他来源获取信息，所以人们在电视机前交流的机会可能减少。交流减少会导致与他人互动的机会变少，进而引发我们如今观察到的问题，例如健康风险 [5], 自我控制能力发展不足和沟通技能欠佳 [26]。

随着人工智能技术的发展，近年来社交机器人已日益普及。预计未来它们将进一步融入我们的生活。然而，目前机器人还无法与人类流畅交流。造成这一问题的原因之一是，机器人无法根据对话中不明显的语境做出相应的行为或回应，而人们则能够基于对话背后的语境进行交流。当我们与家人或朋友一起看电视时，通常会围绕电视节目内容以及他人讨论的话题展开对话、分享情感，从而享受观看电视的乐趣。因此，我们认为，可以利用电视节目中的信息作为机器人与人交流的共享语境。此外，这也可能有助于促进人们在观看电视时的交流。尤其是，对话机器人／聊天机器人有望帮助老年人提高互动效果，减轻他们的孤独感和孤立感。这有助于延缓老年人失能的进程，并降低社会成本及照护者的负担 [27]

鉴于先前研究 [32] 已证明，独自居住在家中的人能够接受对话机器人，因此我们希望聚焦于陪伴家庭或朋友的看电视机器人的对话功能，以评估未来用户与机器人一起看电视的用户体验。因此，本研究围绕以下两个研究问题展开。

- 研究问题1：家庭或朋友与对话机器人一起看电视时，用户体验是否可以接受?
- 研究问题 2：机器人的对话功能能否提升家庭或朋友的幸福感并改善他们的人际关系?

在本研究中，我们基于对观看电视时个人对话行为的对话分析，提出了一种电视陪伴机器人 [2, 4]。我们在机器人上采用开放域聊天机器人“KACTUS” [30?]，并期望它能让与朋友一起观看电视屏幕上的节目变得更加有趣，同时起到促进交流的作用。相关工作将在第 2 节介绍。第 3 节介绍机器人原型的系统设计，第 4 节介绍使用机器人原型进行电视观看实验的结果。第 5 节对本研究进行总结。

2 相关研究

2.1 电视机器人

尽管与人互动的虚拟代理可以采用计算机生成（CG）显示的形式，但据报道，在现实空间中的机器人更适合就实际存在的物体进行对话，而不是显示屏上的虚拟代理 [8]。因此，人们一直在尝试将机器人应用于电视观看环境。

Ogawa等人提出了一种通过在电视节目中同步分发元数据来操作物联网设备（包括机器人）的方法 [18]。这种方法的优点在于，机器人可以配合用户正在观看的电视节目运行。然而，广播机构可能需要创建用于操作机器人的元数据，这会增加电视节目的制作成本。Nishimura等人提出了一种机器人，它利用用户在Twitter等社交媒体上发布的评论与用户聊天[25]。这种方法不需要广播机构创建用于操作机器人的元数据。然而，当用户观看收视率较低的电视节目时，由于社交媒体上的评论较少，机器人的话语数量也会减少。Minami等人提出了一种观看电视时进行聊天的机器人，该机器人利用社交网络服务（SNS）上的评论融入幽默，使用户不会厌倦机器人并停止使用它 [3]。Muto等人提出了一种机器人系统架构和算法，可根据用户的兴趣提供电视节目及相关信息 [16]。然而，这些机器人系统都不是基于开放域聊天机器人构建的。

另一方面，已有研究提出了能够为现有电视节目提供额外信息的头戴式显示器方法 [13, 21]。据报道，在现实空间中的机器人比显示器上的虚拟代理更擅长在现实空间中进行对话 [8]。然而，额外设备必须佩戴在人头部或身体上，这不仅会造成负担，还会干扰人与人之间的对话。

2.2 观看电视时的个人对话行为 g TV

为了理解观看电视的人与人之间自然对话的结构，我们记录并分析了两名彼此熟悉的人组成的小组自由观看自己喜爱的电视节目时的对话 [4]。我们通过对人们观看电视时彼此进行的对话类型进行分类，开展了对话分析，以确定他们的行为。研究发现，在观看电视时，“披露”话语（个人表达其情感或想法的话语）以及对话中的“提问”所占比例较高。

2.3 与老年人进行人机对话 e

鉴于日本等国家65岁及以上人口占比较高，开发面向老年人的人机对话系统已成为当务之急 [19], , 以帮助独居者增加在日常生活中进行对话的机会。Pearl [20] 会提醒老年人完成日常活动，例如每三小时去一次洗手间以及服用药物。此外，老年人的照护者可以提前录入某人的日常活动，Pearl 会根据日程安排提醒此人。第一个试验系统 [15] 旨在支持用户一边观看电视节目一边与机器人聊天；该系统在同时响应方面实现了自然的时机控制，并支持基于用户正在观看的电视节目相关社交媒体网络内容而展开的有趣对话。人机对话系统开发面临的障碍之一是语音识别经常失败。为解决这一问题，有研究提出了问答响应对话模型 [6]，使机器人能够主动向用户提出各种问题，并让两个机器人参与对话。另一项研究 [7] 旨在通过对话建模增强用户与机器人之间的交流能力

以改善为老年人提供的老年人护理服务。此外，他们还引入了一种对话系统架构，该架构能够基于照护任务与用户进行交流。

过去的对话机器人大多进行定向对话或有限主题对话。我们提出了一种适用于开放领域的对话机器人，它能够自由地谈论更多丰富的话题，并表现得像人类一样。

3 系统设计

我们的对话功能按照以下原则设计，如图所示：1。机器人通过在以下两种模式——“电视观看模式”和“对话模式”之间切换，来模拟人类看电视时的对话行为 [4]。机器人原型可以在“电视观看模式”和“对话模式”之间反复切换。

1）“电视观看模式”： 在此模式下，机器人会谈论电视内容中的评论和自己的感受，就好像它正在欣赏电视节目一样。我们认为，机器人可能会发起一个对话话题，在家庭成员或朋友一起看电视时创造交流机会。我们实现了以下两项功能。

- 特征A：从电视节目中提取关键词
- 特征B：根据提取的关键词生成自我披露话语

2) “对话模式”： 在此模式下，机器人以相同方式生成包含关键词的提问话语，然后围绕电视节目相关内容开展开放领域聊天。对话主题起初与电视节目相关，随后会扩展到日常生活以及最新的新闻/社交网络服务。我们认为，机器人或许能提供更多交流机会，从而提升家庭成员或朋友的幸福感并改善彼此关系。我们实现了以下两项必要功能：

- 特征C：根据电视节目中的关键词生成提问话语
- 特征D：采用由多个对话引擎组成的基于主题的对话管理方法，以开展丰富的、与电视节目相关的对话

我们的机器人原型配备了所提出的开放领域对话功能，并基于开发机器人 CommU 进行了扩展 [28]，，如图1所示。我们制作了一个摄像头-麦克风阵列原型，并将其作为附加设备安装在 CommU 底部（图1 右下）。通过使用摄像头-麦克风阵列，机器人能够检测阵列约 200 度范围内附近的人，并借助 HARK 的声音定位与分离方法区分人声和电视声音 [17]。识别出的人声通过微软语音识别转换为文本 [1]。此外，尽管本研究以英语介绍各项功能，机器人实际使用日语运行。

3.1 电视观看模式（包括功能 A 和 B）

功能 A：首先，机器人从电视节目中获取视频、音频和字幕数据，以提取用于话语生成的关键词。字幕数据中的句子通过 MeCab [9] 与 mecab-ipadic-NEologd [22]转换为分词形式。如果电视节目不包含字幕数据，我们会在 AWS [23], 上实施视频目标检测方法 [12, 24] ，从电视节目内容中提取关键词；该方法通过 MS COCO [11] 进行学习，并采用 [1]提供的识别 API。

检测到的物体的分类结果被用作提取出的关键词。图中展示了一个在本地计算机上使用目标检测进行关键词提取的示例：1。绿色矩形表示检测到的物体，而

红色区域表示显著性估计结果。在图中示例中，“大象”被提取为关键词。我们使用了 Microsoft Azure [1] 和 Amazon Web Service [1] 云服务 [23]。图 1 ）右上方展示了 “ 大象 ” 的关键词提取结果，作为示例。

功能B：然后，使用提取出的关键词生成表达机器人情感的自我披露话语。我们使用过去七年频道10的电视节目字幕作为自我披露话语生成的模板句语料库。由于电视节目中的全部话语有79%少于20个字符，因此这些模板句少于20个字符 [4]。通过Word2Vec [14]，利用过去的字幕模板句学习词语的分布式表示。机器人通过计算模板句与提取出的关键词之间的高余弦相似度，选择要使用的模板句。然后，将新的关键词插入模板句或替换其中的词语，从而生成话语。例如，在上述示例中，将生成“I like elephants”这样的自我披露话语。部分模板样例如下。

- 自我披露：生成的自我披露话语采用包含情感表达的模板句（例如：我想吃***、我想去***：关键词以***表示）
- 提问：生成的自我披露话语采用包含提问的模板句（例如：你想吃***吗？你喜欢***吗？关键词以***表示）

接下来，机器人会判断是否切换到对话模式。自我披露话语和提问话语的比例预先设定为参数，并根据该比例随机选择其中一种。话语间隔也可以调整，以避免机器人持续说话的情况。话语频率预先设定，话语间隔则根据泊松分布确定。此外，为了生成多样化的话语，在话语生成中使用过一次的关键词会暂时不再重复使用。

最后，机器人转向电视，并通过 CommU TTS API [28]切换到对话模式，说出生成的自我披露话语。在进行上述处理的同时，面向电视的机器人会持续随机重复“眨眼”“点头”“转动脖子”“移动上半身”等动作，表现得仿佛正在看电视。

3.2 对话模式（包括功能 C 和 D）

功能 C：在此模式下，提问话语的生成方式与上述自我披露话语的生成方式类似。我们使用相同的模板句语料库来生成提问话语。机器人通过计算提取出的关键词的余弦相似度，来选择要使用的模板句。如图1中的上述示例，生成的提问话语将是“你喜欢大象吗？”。

自我披露话语和提问话语的比例预先设定为参数，并根据该比例随机选择其中一种。还可以调整话语间隔，以避免机器人持续不断地说话。此外，话语生成中使用过的关键词会暂时停用，以便生成多种不同的话语。

特征D：此时，如果麦克风阵列接收到人类的回应，机器人便会转向观众，说出生成的提问话语。随后，开放域聊天机器人“KACTUS”会根据语音识别结果生成下一条回应话语。为了提升家庭或朋友之间的幸福感与人际关系，聊天机器人不仅提供与电视节目直接相关的愉快对话，还

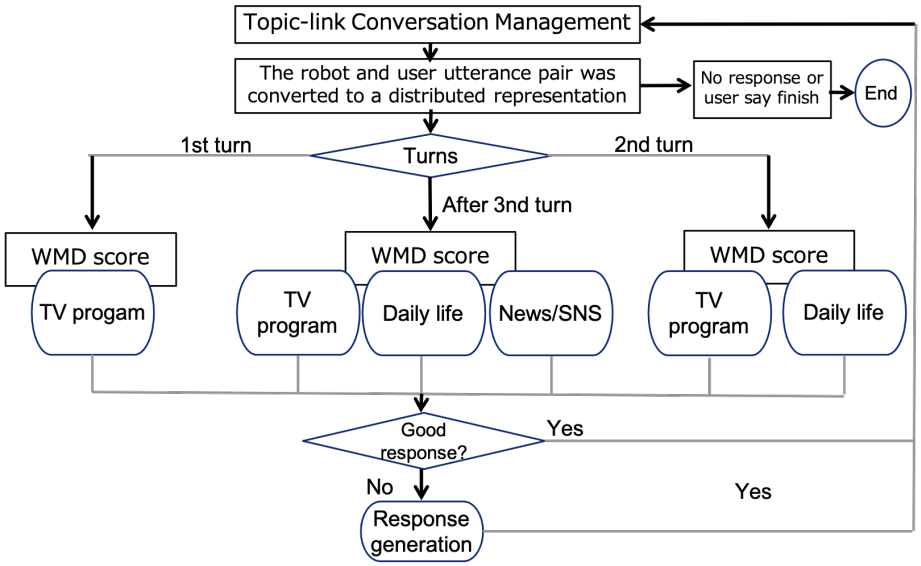


图2。函数 D：主题关联对话管理。

通过多重对话引擎，促进用户与机器人围绕各种相关话题进行聊天。以下四种对话引擎被结合用于响应生成。

- 电视节目：电视节目对话包括机器人提问话语之后的回应。
- 日常生活：针对一般内容的广泛对话。
- 新闻/社交网络：使用相关互联网新闻和会随时间变化的 Twitter 评论进行对话。
- 响应生成：基于 Transformer 的模型 [29]使用大规模对话对数据集。

主题关联对话管理的聊天机器人策略如图 2 所示。首先，将机器人和用户的话语对转换为分布式表示。接下来，机器人从不同的对话引擎中选出与该分布式表示最相似、词移距离最高的句子 [10], ，并将其作为话语。第 1 轮仅使用电视节目的对话引擎；第 2 轮使用电视节目和日常生活的对话引擎；从第 3 轮开始，则使用电视节目、日常生活和新闻/社交网络服务的对话引擎。也就是说，随着对话推进，话题预计会逐渐从电视节目转向日常生活、最新新闻/社交网络服务等更广泛的话题。在这些情况下，机器人会在上述三个对话引擎之间切换，以继续对话。当用户的话语与对话引擎集合之间的词移距离相似度低于阈值时，机器人会使用基于 Transformer 的响应生成引擎生成回应话语，以避免对话中断。最后，满足以下任一条件时，对话便会结束，随后进入电视观看模式。

- 如果根据用户的话语判断其想要结束对话。
- 如果判断用户连续两次以上未作答。

对话继续时，机器人会再次等待语音输入。对话结束时，机器人转向电视，当前状态切换为状态 1：观看电视。



（a）在聆听室中



（b）实验中的样子?

图3. 视听实验进行中

4 实验结果

为了检验为家庭或朋友设计的对话功能的用户接受度，我们分两步开展实验。虽然使用的是机器人的同一套对话功能，但第一项实验聚焦于“研究问题1”：在观看电视时，机器人的对话功能能否促进家庭成员或朋友之间的交流？随后开展的第二项扩展实验则聚焦于“研究问题 2”：机器人的对话功能能否提升家庭成员或朋友的幸福感并改善彼此的人际关系?

4.1 初步实验结果

4.1.1 实验设置。

参与者被分为三组，每组两人，如表1所示。我们假设初始使用场景为参与者在周末观看电视节目，因此实验时间约为六小时，从10:00到16:00，其中包括1小时的午餐时间。参与者必须全程观看，但可以选择过去七年内的任何电视节目。实验结束时，参与者填写了一份七级量表问卷，选项从“非常不同意”到“非常同意”（见表 2 ），并接受了访谈。每组的实验持续两天：第一天在没有机器人的情况下进行，第二天则在有机器人的情况下进行。

为了在轻松的状态下进行实验，我们使用了一间布置得仿佛客厅的房间作为实验场地。机器人原型放置在桌子上（见图 3 (a)），并独立运行，不受实验人员控制。机器人话语中自我披露与提问的比例设为 3:1，平均话语频率设为每隔 80 秒一次。

表1. 参与者的属性

参与者	性别，年龄	关系	交谈	机器人
G1（无机器人）	女性，三十多岁	家庭	673	-
G1（有机器人）	女性，三十多岁	家庭	555	37%
G2（无机器人）	女性，二十多岁	朋友	859	-
G2（有机器人）	女性，二十多岁	朋友	1446	12%
G3（无机器人）	男性，二十多岁	朋友	1337	-
G3（与机器人一起）	男性，二十多岁	朋友	960	35%

表2。结果

ID	内容	得分	SD
Q1	你觉得对话机器人有趣吗?	6.5	0.46
Q2	你觉得机器人对话怎么样?	5.0	1.51
Q3-1	话语长度合适吗?	5.0	1.77
Q3-2	话语频率合适吗?	3.5	1.67
Q3-3	话语时机合适吗?	3.5	1.75
Q3-4	话语内容合适吗?	2.5	1.16

4.1.2 分析结果。

4.1.2.1 人类话语（交谈）的数量。见表1，G1 和 G3 组的人类话语数量在有机器人时有所减少，而 G2 组则有所增加。由机器人引发的人类话语所占比例分别为 37%、12% 和 35%。尽管我们无法断定机器人能够增加交流机会，但从录制的视频中可以看出，所有小组都饶有兴趣地倾听机器人的话语，而且比没有机器人时笑得更加频繁。例如，当有人回答“我经常在池袋购物”时，机器人回应了池袋发生车祸的新闻（池袋是东京著名的商业和娱乐区），随后两个人不禁好奇机器人怎么会知道这起车祸。未来，我们计划使用 FER [33]进行定量分析，以展示机器人如何改变交流情绪。

4.1.2.2 问卷。我们在设计问卷时考虑了用户兴趣和话语的可接受性，采用从“非常不同意”到“非常同意”的七级量表。关于用户兴趣，表2中问题1和问题2分别询问整体体验和对话功能，平均评分为6.5和5.0，标准差分别为0.46和1.51。在访谈中的话语接受度测试中，所有参与者都表示，与机器人的对话足够有趣。随后，参与者对机器人话语的长度、频率、时机和内容分别给出了5.0、3.5、3.5和2.5的评分。访谈中，尽管~~大多数参与者认~~为话语长度（Q3-1）“合适”，但他们对其他三项功能的印象较差，这是因为机器人在参与者专注观看电视节目时仍会说话。他们还指出，有时生成的话语听起来不自然（例如，“你吃床吗？”）。最后，我们分析了对话日志，发现人机对话平均包含7.4条话语，这与观看电视时人际对话的平均水平相近 [4]。与此同时，对于与披露话语内容相关的问卷项目（Q3-4），六名受试者中有四名作出了否定回答；第2组的一名受试者在访谈中表示：“机器人有时会说一些与电视节目无关的内容。”这也证实了机器人在延续对话方面的有效性。考虑到这些问题，我们正在开发一种参与度人工智能方法 [31]，以便机器人能够识别用户观看电视时的专注程度。此外，我们计划在未来的研究中提高话语生成方法的准确性。

4.1.2.3 受试者与机器人之间对话的样本和轮次数。最后，我们参考受试者与机器人之间对话的轮次数，使用“KACTUS”验证了机器人响应话语的质量。将受试者和机器人围绕单一话题进行连续对话时所产生的话语数量定义为轮次数，并进行了统计。表3 展示了连续对话样本，表4展示了轮次统计。人际对话由大约3至6轮组成，而另一方面，80%的人机对话由大约5至8轮组成。尤其是第3组，据推测该组参与者享受了与机器人的对话，因此与其他组相比，他们在单次对话中的轮次数往往更高；同时也确认了机器人通过切换话题对话来持续进行对话的有效性。

表4. 与机器人对话的轮次

受试者与机器人对话的轮次数			
平均值	4.71	5.26	6.88
最大值	10	11	14

表3. 对话示例

说话者	话语内容	对话引擎
S:	你看过互联网吗?	电视节目
U:	Yes, 我看过!	-
S:	哦，我明白了。其实，我也看过。你不想什么时候也看看吗?	电视节目
U:	Oh, 我也看过?	-
S:	毕竟是互联网啊。不知道会是什么样呢?	日常生活
U:	一定很不可思议 , 不是吗, 互联网?	-
S:	据 CNET Japan 报道，英伟达接近达成收购协议 软银旗下的 Arm。"	新闻/社交网络
U:	哦，真的吗? 这是怎么回事?	-
S:	软银集团想以4.2万亿日元将 ARM 卖给英伟达。这笔交易 预计将在18个月内完成。	新闻/社交网络
U:	那可真花了不少时间!	-
S:	TikTok 将成为甲骨文收购案中的一大漏洞。	新闻/社交网络
U:	TikTok 怎么样?	-
S:	英伟达宣布与软银达成协议，将以 400 亿美元收购 ARM。	新闻/社交网络
U:	$40B 太便宜了!	-
S:	我知道，对吧?	响应生成

4.1.2.3 扩展实验的结论及改进思考。尽管我们无法根据结果得出机器人能够促进家庭成员和朋友之间交流的结论，但结果显示，所有参与者都喜欢与机器人交谈；关于他们对机器人感兴趣程度的提问，得分为6.5（七级量表：从“非常不同意”到“非常同意”递增）。这表明，与机器人一起观看电视有可能使我们的日常生活更加愉快。另一方面，考虑到Q3-2至Q3-4的得分相对较低，我们为下一次扩展实验进行了如下三项调整。

- 我们添加了一个取消按钮，让用户专注于电视剧或电影等电视节目时，机器人可以立即停止说话。
- 我们向词典中添加了更多电视节目相关词汇，以便更好地匹配。
- 我们将 WMD阈值设得更高，以防止机器人谈论与电视节目无关的话题。

4.2 扩展实验结果

为探究该方案对“研究问题 2”的影响——用户是否认为机器人的对话功能能够改善家庭或朋友之间的幸福感和人际关系，我们开展了以下扩展实验。我们还修订了问卷，以便评估电视陪伴机器人对幸福感和人际关系的影响。

4.2.1 实验设置。
共有64名受试者，由32组两人组成（家庭成员或朋友），彼此关系融洽且喜欢观看电视。他们均由临时就业机构招募，不限年龄或性别。实验概要如表9 所示，受试者属性如表**??**所示。与初始实验一样，实验时长约为六小时，其中包括一小时的午餐休息时间、说明时间和填写问卷的时间。具体时间分配见表9。受试者可以观看任何电视节目

他们可以从内容服务器中自由选择电视节目；该服务器包含过去七年录制的电视节目，共十个频道。实验结束时进行了问卷调查和访谈。

为探究观看电视时有无机器人的影响，我们开展了一项随机对照实验。干预组为与机器人一起观看电视的组，对照组为仅与人类一起观看电视的组。

为了使参与者在放松状态下进行实验，我们使用了一间模拟客厅的房间开展实验。机器人原型被放置在桌上（见图3（a）），并在无需实验人员控制的情况下独立运行。机器人的“电视观看模式”与“对话模式”的比例设为3:1，话语频率平均设为每80秒一次。我们使用AWS收集对话数据。

实验前，我们说明了实验的目的和流程。实验结束后，我们进行了问卷和访谈，以了解参与者对实验的印象。

实验开始时，向受试者给出了以下指示。指示内容：“今天是假期。你总是很忙，但今天你决定放松一天，和你的好朋友（家人）一起看你最喜欢的电视节目，聊聊天。

表5. 实验概要

受试者	64人（每组2人，共32组） 年龄： 10~70s 性别：男性（9人），女性（55人）
实验周期	2021年9月~11月（每组进行一天）
实验时间	上午（10:00-12:00） 1) 实验知情同意书的说明与确认 2) 不使用/使用机器人观看电视
	下午（13:00-18:00） 1）观看电视（无机器人/有机器人） 2）事后问卷

4.2.2 问卷。

作为观看实验一部分所进行的问卷包含以下两个问卷题项：I–II。对照组填写了问卷一，而干预组在实验结束后填写了问卷一和问卷二（表6）。问卷一和问卷二的问题见表7–8。

- 问卷一：用于考察对一般对话机器人的接受度/印象的问卷题项。
- 问卷二：用于考察所提出对话功能印象的问卷题项。

表6。实施时间和条件

	对照组	**干预组**
上午（10:00-11:00）	实验知情同意书的说明与确认	
上午（11:00-12:00）	不与机器人一起看电视	不与机器人一起看电视
下午（13:00-17:00）	不与机器人一起看电视	与机器人一起看电视
下午（17:00-18:00）实验结束后	I	I、 II

表7. 问卷一。

No.	提问
I-1	您愿意在家里拥有一台对话机器人吗?

表8. 问卷二。

No.	问题
II-1	对话机器人增加了两人之间的交流。
II-2	对话机器人让这里的氛围更加轻松。
II-3	对话机器人让我把注意力转向了电视。
II-4	对话机器人让它比平时更有趣。
II-5	你认为，如果家里有一台对话机器人，你的日常生活会变得更好吗?
II-6	当对话机器人喃喃自语时，你有什么感受?
II-7	与对话机器人互动后，你有什么感受?
II-8	如果我们改进这款对话机器人，你希望家里有一台吗?
II-9-1	预期改进：围绕电视节目的对话越来越多
II-9-2	预期改善：对话机器人能够共情人类情感

4.2.3 分析结果。
根据问卷结果，对以下两个因素进行了分析。

- 因子（1）：对家中通用对话机器人的接受度。
- 因子（2）：实验结束后，电视观看机器人对幸福感和人际关系的影响／印象

在因子（1）的分析中，我们比较了对照组和干预组在问卷一中的回答结果；因子（2）的分析则是干预组在问卷二中的回答结果。

4.2.4 问卷一分析（因子（*1*））。
在本研究中，我们调查了哪些问卷题项会根据是否使用沟通机器人而改变对沟通机器人的总体印象。表 9 展示了问卷一第 1 题“您想在家里拥有一个会说话的机器人吗？”（两级制：1：我认为是，2：我不这么认为）的回答结果。我们发现，干预组对话机器人的接受度为 43.75%，高于对照组的 18.75%。我们对问卷一第 1 题进行了卡方检验，以在 5% 的显著性水平下检验两组之间是否存在显著差异。

- （问卷一第 1 题）：您想在家里拥有一个对话机器人吗?

表 9。问卷一第 1 题“您想在家里拥有一个会说话的机器人吗？”的回答

	对照组	干预组
我认为是	6	14
我不这么认为	26	18
接受率	18.75%	43.75%

4.2.5 问卷*II*分析（因子（*2*））。
为了解用户印象并验证所设计的对话功能 A-D 的效果，我们分析了表10中干预组填写的问卷II答复。

首先，问卷二中问题 II -1 至 II -5 的答案直接反映了受访者对对话功能能否提升家庭或朋友之间的幸福感和人际关系的看法。回答选项采用 7点评特量表，范围从“7：非常多”到“1：非常少”。可以看出，II -1、II -2 和 II -4 的得分相对较高，而 II -3 的得分相对较低。从补充意见来看，关于“机器人的存在增进了两个人之间的对话”和“机器人的存在让场所氛围更轻松”的正面评价相对较多。另一方面，也有一些关于“机器人的说话声让我注意到了电视”的负面评价。当前对话功能的表现并未起到在观看电视或与他人互动时将人们的注意力引向电视的作用。至于 II-5：你认为如果家里有一台机器人，你的日常生活会

变得更好吗？约 75% 的受访者总计表示，他们喜欢与机器人互动。他们中的大多数人给出了正面评价，包括“我能够和朋友及家人一起愉快地看电视节目”。

其次，问题 II -6- II -7 的答案直接反映了用户对对话机器人的两种模式的印象。

–II-6：机器人自言自语（“电视观看模式”）时，你有什么感受？超过72%的受访者认为这很有趣。也有一些负面评论，例如“有时它说的内容与节目不符”和“有时它会重复说同样的话”。

–II-7：与机器人互动后，你有什么感受（“对话模式”）？约76%的受访者喜欢与机器人互动。正面评论包括“他们喜欢机器人与家人或朋友之间的对话”。负面评论包括“有时它说的内容与电视节目不符”和“有时它会反复谈论同一个话题”。如上所述，约70%的受访者认为机器人的对话很有趣，但也存在一些问题，例如内容重复。

第三，对这些问题的回答提供了有关用户接受度以及未来改进因素的信息。II-8显示，如果改进对话功能，72%的用户愿意使用对话机器人，接受度高于I-1中的43.75%。II-9-1和II-9-2显示了用户对未来对话功能改进因素的前两项预期。

表10. 问卷问题II-1至II-9的回答结果

问卷问题	回答“是”的比例	平均得分	SD
II-1：对话机器人是否增进了两人之间的交流?	-	4.78	1.38
II-2：对话机器人让这个地方变得更加轻松自在。	-	5.06	1.29
II-3：对话机器人让我把注意力转向了电视。	-	3.16	1.43
II-4：对话机器人让它比平时更有趣。	-	4.5	1.39
II-5：你认为，如果家里有一台对话机器人，你的日常生活会变得更好吗?	75%	-	-
II-6：对话机器人喃喃自语时，你有什么感受?	72%	-	-
II-7：与对话机器人互动时，你有什么感受?	76%	-	-
II-8：如果我们改进得更好，你希望家里有对话机器人吗?	72%	-	-
II-9-1：期望的改进：增加更多与电视节目相关的对话	-	5.03	1.23
II-9-2：期望的改进：对话机器人能够了解并体察人类情感	-	5.22	1.61

最后，表11。展示了对话轮次的分析。

表 11。对话轮次的统计数据

轮次	总计	9/1	9/8	9/14	9/18	9/23	9/30	10/5	10/13	10/16	10/19	10/20	10/23	10/27	10/28	11/9	11/12	11/18	11/19
Mean	5.54	4.92	4.15	7.00	6.32	5.96	5.45	5.38	6.45	5.12	5.67	6.83	5.91	5.40	6.21	5.88	6.52	4.09	4.61
Max	16	8	8	10	11	8	10	9	10	12	8	13	16	9	9	9	11	7	9

与初始实验一样，表格11 展示了扩展实验中的轮次统计。我们发现，对话的平均轮次数和最大轮次数没有显著变化，这也表明干预组享受与机器人进行的对话，并证实了主题关联对话管理方法能够使机器人持续进行对话。

4.2.6 考量。

结果显示，与对照组相比，干预组对“如果家里有一个看电视机器人就好了”的需求更高。结果还表明，所提出的对话功能有望改善家庭成员或朋友的幸福感和人际关系。此外，我们还发现了未来可改进的方面，以提高用户对对话机器人的接受度，例如对话

时机——用户专注于电视时、与电视节目不一致的对话主题，或不恰当的主题转换。我们将改进机器人以解决这些问题，并在未来对更多受试者开展实验，以验证人们在看电视时我们的机器人所产生的影响。

5 结论

本研究基于人们观看电视时的对话分析，提出了机器人与人们一同观看电视所需的功能要求。此外，我们依据这些功能要求评估了作为实现该机器人第一步而开发的机器人原型，并根据电视观看实验的结果进行了运行验证。受到人们观看电视时个人对话行为的对话分析启发，该机器人通过在以下两种模式之间切换来工作：“电视观看模式”和“对话模式”。在“电视观看模式”下，机器人首先从电视节目中提取关键词，然后根据提取出的关键词生成自我披露话语，表现得仿佛正在欣赏电视节目。在“对话模式”下，机器人以相同的方式围绕关键词生成提问话语，并采用由多个对话引擎组成的基于主题的对话管理方法，开展与电视节目相关的丰富对话。

初步实验中，所有参与者被分为三组，每组两人。虽然我们无法得出结论，认为机器人能在家庭成员或朋友一起看电视时增加彼此交流的机会，但结果显示，三组中的所有参与者都喜欢与机器人交谈。这也表明，所提出的看电视陪伴机器人的对话功能有望让我们的日常生活更加愉快。

此外，我们进一步开展了扩展实验，共有32组参与者，每组两人；其中一半不使用机器人看电视，作为对照组，另一半使用机器人看电视，作为干预组。与对照组相比，干预组对“如果家里有一台看电视机器人就好了”的需求更高，且差异显著。我们还发现，许多干预组参与者认为，机器人的对话功能“让氛围更放松”“增加了我们两人之间的对话”，并且“比平时更有趣”。结果表明，所提出的对话功能有望提升家庭成员或朋友的幸福感并改善彼此关系。本研究还根据问卷和操作评估总结了改进要点。

与此同时，我们根据运行评估收集了需要改进之处，例如用户专注于电视时的对话时机、与电视节目内容不一致的对话主题，或不恰当的主题转换。我们将改进机器人以解决这些问题，并在未来让更多受试者参与实验，以验证我们的机器人对观看电视的人类产生的影响。

致谢

作者谨向 Systemsoft Inc 的专家 Zhiguang Zhou、Zheshuang Lyu、Wataru Nishioka 和 Megumi Komiya 致以谢意，感谢他们对我们所提出的方法及其优化提供的大力支持与建议。

参考文献

[1] Microsoft Azure. 2021. *Azure* 认知服务. Microsoft. 检索于2021年10月3日，网址： https://azure.microsoft.com/en-us/services/cognitive-services/[2] Yuta Hagio、Marina Kamimura、Yuta Hoshi、Yutaka Kaneko 和 Masao Yamamoto。2021。看电视机器人：丰富媒体体验与促进人类交流。收录于 *2021*年国际广播大会。[3] Masayuki Kanbara、Hidekazu Minami、Hiromichi Kawanami 和 Norihiro Hagita。2016。结合机器响应与社交媒体评论、实现连续对话的聊天机器人。收录于 *2016*年*IEEE*多媒体博览会国际会议研讨会（*ICMEW*）。IEEE出版社，1–6。

[4] Yuta Hoshi、Yutaka Kaneko、Michihiro Uehara、Yuta Hagio、Yasuhiro Murasaki、Satoshi Nishimura 和 Masao Yamamoto。2020。面向看电视人群的陪伴机器人话语功能。收录于2020年IEEE国际消费电子会议（ICCE），美国内华达州拉斯维加斯，2020年1月4日至6日。IEEE，美国内华达州拉斯维加斯，1–5。[5] James S House、Karl R Landis 和 Debra Umberson。1988。社会关系与健康。《科学》241，4865（1988），540–545。[6] Takamasa Iio、Yuichiro Yoshikawa、Mariko Chiba、Taichi Asami、Yoshinori Isoda 和 Hiroshi Ishiguro。2020。用于老年人与机器人对话、可抵御语音识别失败的双机器人对话系统。《应用科学》10，4（2020），1522。[7] Kristiina Jokinen、Satoshi Nishimura、Kentaro Watanabe 和 Takuichi Nishimura。2019。用于解释活动的人机对话。收录于第九届口语对话系统技术国际研讨会。Springer，239–251。[8] Cory D. Kidd 和 Cynthia Breazeal。2004。机器人对用户感知的影响。收录于2004年IEEE/RSJ智能机器人与系统国际会议，日本仙台，2004年9月28日至10月2日。IEEE，日本仙台，3559–3564。[9] Taku Kudo。2006。MeCab：又一个词性与形态分析器。NAIST。检索于2021年10月3日，网址：http://taku910.github.io/mecab/[10] Matt J. Kusner、Yu Sun、Nicholas I. Kolkin 和 Kilian Q. Weinberger。2015。从词嵌入到文档距离。收录于第32届国际机器学习会议论文集，ICML 6-11（JMLR研讨会与会议论文集，第37卷）。JMLR.org，法国里尔，957–966。[11] Tsung-Yi Lin、Michael Maire、Serge J. Belongie、James Hays、Pietro Perona、Deva Ramanan、Piotr Dollár 和 C. Lawrence Zitnick。2014。Microsoft COCO：语境中的常见物体。收录于《计算机视觉——ECCV 2014——第13届欧洲计算机视觉会议》，瑞士苏黎世，2014年9月6日至12日，论文集第五部分（《计算机科学讲义》，第8693卷），David J. Fleet、Tomás Pajdla、Bernt Schiele 和 Tinne Tuytelaars（编）。Springer，瑞士苏黎世，740–755。https://doi.org/10.1007/978-3-319-10602-1_48[12] Tsung-Yi Lin、Michael Maire、Serge Belongie、James Hays、Pietro Perona、Deva Ramanan、Piotr Dollár 和 C Lawrence Zitnick。2014。Microsoft COCO：语境中的常见物体。收录于《欧洲计算机视觉会议》。Springer，瑞士苏黎世，740–755。[13] Florian Mathis、Xuesong Zhang、Mark McGill、Adalberto Simeone 和 Mohamed Khamis。2020。评估混合现实电视中的社交文本布局。收录于ACM交互媒体体验国际会议（IMX’20）。ACM，西班牙巴塞罗那，205–211。[14] Tomas Mikolov、Ilya Sutskever、Kai Chen、Greg S Corrado 和 Jeff Dean。2013。词与短语的分布式表示及其组合性。收录于《神经信息处理系统进展》。载于第26届神经信息处理系统国际会议论文集，美国内华达州太浩湖，3111–3119。[15] Hidekazu Minami、Hiromichi Kawanami、Masayuki Kanbara 和 Norihiro Hagita。2016。结合机器响应与社交媒体评论、实现连续对话的聊天机器人。收录于2016年IEEE多媒体博览会国际会议研讨会（ICMEW）。IEEE，1–6。[16] Yuki Muto、Yoshihiro Iwase、Shunichi Hattori、Yasufumi Takama 和 Kaoru Hirota。2006。面向看电视环境下人机交流的网络智能方法。收录于SCIS与ISIS 2006。日本模糊理论与智能信息学会、ACM，日本，426–429。[17] Kazuhiro Nakadai、Hiroshi G. Okuno 和 Takeshi Mizumoto。2017。机器人听觉开源软件HARK的开发、部署与应用。《机器人学与机电学》29，1（2017），16–25。[18] Hiromu Ogawa、Masaya Ikeo、Hisayuki Ohmata、Chigusa Yamamura 和 Hiroshi Fujisawa。2018。广播内容物联网服务的系统架构。收录于IEEE国际消费电子会议，ICCE 2018，美国内华达州拉斯维加斯，2018年1月12日至14日。IEEE，美国内华达州拉斯维加斯，1–2。[19] 日本内阁府：《老龄社会年度报告：2020》（摘要）<财年2020>。2020。通过科学技术在老龄社会中建立健康新范式。https://www8.cao.go.jp/kourei/english/annualreport/2020/pdf/2020.pdf（访问日期：2022年1月23日）[20] Martha E Pollack、Laura Brown、Dirk Colbry、Cheryl Orosz、Bart Peintner、Sailesh Ramakrishnan、Sandra Engberg、Judith T Matthews、Jacqueline Dunbar-Jacob、Colleen E McCarthy 等。2002。Pearl：面向老年人的移动机器人助理。收录于《自动化作为老年护理》AAAI研讨会，第2002卷。85–91。[21] Pejman Saeghe、Sarah Clinc、Bruce Weir、Maxine Glancy、Vinoba Vinayagamoorthy、Ollie Pattinson、Stephen Pettifer 和 Robert Stevens。2019。通过增强现实扩展电视体验。收录于2019年ACM电视与在线视频互动体验国际会议（TVX’19）。ACM，英国索尔福德（曼彻斯特），255–261。[22] Toshinori Sato、Taiichi Hashimoto 和 Manabu Okumura。2017。名为mecab-ipadic-NEologd的分词词典的实现及其在信息检索中的有效使用方法研究。收录于第二十三届自然语言处理学会年会论文集。自然语言处理学会，载于自然语言处理学会年会，日本，NLP2017–B6。[23] Amazon Web Service。2021。。Amazon。检索于2021年10月3日，网址：https://aws.amazon.com/[24] Ren Shaoqing、He Kaiming、Girshick Ross 和 Sun Jian。2017。Faster R-CNN：利用区域提议网络实现实时目标检测。《IEEE模式分析与机器智能汇刊》39，6（2017），1137–1149。[25] Masayuki Kanbara、Shogo Nishimura、Hiromichi Kawanami 和 Norihiro Hagita。2017。具备时间平移功能、用于日常交流的电视聊天机器人。收录于第九届国际社交机器人会议（ICSR）。Springer，日本筑波，516–525。[26] Sohyune R Sok、Mi Hyeon Seong 和 Mi Hye Ryu。2019。韩国护理专业学生中智能手机成瘾风险组与普通组在自我控制、日常生活压力和沟通技能方面的差异。《精神病学季刊》90，1（2019），1–9。[27] Ming-Hsiang Su、Chung-Hsien Wu、Kun-Yi Huang、Qian-Bei Hong 和 Hsin-Min Wang。2017。采用基于LSTM的多层嵌入、用于老年人护理的聊天机器人。收录于2017年国际橙色技术会议（ICOT）。IEEE，70–74。[28] 大阪大学。2015。社交对话机器人“CommU”和“Sota”的开发。大阪大学。检索于2021年10月3日，网址：https://resou.osaka- u.ac.jp/en/research/2015/20150120_2

[29] Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N Gomez、Łukasz Kaiser 和 Illia Polosukhin。2017。《Attention Is All You Need》。载于神经信息处理系统进展。Curran Associates Inc.，加利福尼亚州长滩，5998–6008。[30] Jianming Wu、Bo Yang、GEN HATTORI 和 YASUHIRO TAKISHIMA。2020。《利用多模态识别技术的聊天机器人 AI“KACTUS”》。日本信息处理学会《SIG技术报告》，AVM 2021(7)，1-5。[31] Jianming Wu、Bo Yang、Yanan Wang 和 GEN HATTORI。2020。《用于参与强度预测的先进多实例学习方法：多特征工程与保守优化》。载于*ICMI '20*：国际多模态交互会议，线上会议，荷兰，*2020*年*10*月*25*日至*29*日。ACM，荷兰，777–783。[32] Daisuke Yamamoto、Yuka Kobayashi、Sachie Yokoyama 和 Miwako Doi。2009。《面向老年人的对话界面——机器人化身伙伴，让老年人快乐》。日本电子信息通信学会技术报告；*IEICE*技术报告 109，224（2009），47–51。[33] Bo Yang、Wu Jianming 和 GEN HATTORI。2021。《新冠肺炎疫情期间的口罩感知鲁棒面部表情识别》。载于*2021 IEEE*国际图像处理会议（*ICIP*）。IEEE，美国阿拉斯加州安克雷奇，240–244。