arXiv:2602.04017v1 [cs.HC] 3 Feb 2026

牧师对用于对话式关怀的人工智能设计与使用的反思

Joel Wester
joel.wester@di.ku.dk 哥本哈根大学 丹麦哥本哈根 奥尔堡大学 丹麦奥尔堡

SamuelRhys Coxsrcox@cs.aau.dk 奥尔堡大学 丹麦奥尔堡

Henning Pohl
henning@cs.aau.dk 奥尔堡大学 丹麦奥尔堡

Niels van Berkel
nielsvanberkel@cs.aau.dk 奥尔堡大学，丹麦奥尔堡

摘要

尽管人们日益认识到，负责任的人工智能需要领域知识，但当前对话式人工智能的研究主要依赖于侧重诊断和干预的临床专业知识。然而，日常情感支持需求大多出现在非临床情境中，因此需要采用不同的对话方式。我们考察了引导人们度过个人危机、悲伤与反思的牧师如何看待并使用对话式人工智能。我们招募了18位牧师，共同构建人工智能聊天机器人。尽管一些牧师对聊天机器人持谨慎乐观态度，但大多数人指出，聊天机器人在支持日常福祉方面能力有限。我们的分析揭示了牧师如何理解自身的牧灵关怀职责，以及人工智能聊天机器人在哪些方面有所不足；这些发现围绕以下主题展开：倾听、联结、承载和渴望。这些主题与“调谐”这一理念相呼应；这一理念最近被提出作为一种关系性视角，用以理解关怀技术所带来的细腻体验。这一视角为面向非临床情境、旨在支持福祉的聊天机器人设计提供了参考。

CCS概念

•**以人为中心的计算** → **人机交互实证研究**；**自然语言界面**。

关键词

人工智能聊天机器人、人机交互、牧师、牧灵关怀、幸福感、临在

ACM参考文献格式：

Joel Wester、Samuel Rhys Cox、Henning Pohl和Niels van Berkel。2026年。牧师对人工智能用于对话式关怀的设计与使用的反思。载于*2026*年*CHI*人机交互大会论文集（*CHI '26*），*2026*年*4*月*13*日至*17*日，西班牙巴塞罗那。ACM，美国纽约州纽约市，15页。https://doi.org/10.1145/3772318.3790468

1 引言

人工智能（AI）聊天机器人日益被用作引导者和陪伴者 [116]，，通过与用户建立单向关系提供拟社会支持 [73]。这包括围绕以下主题展开的对话自我反思、悲伤或关怀 [103]。与此同时，大量研究指出，人工智能聊天机器人的同情心存在局限 [36]，并且它们可能造成重大伤害 [78]。因此，如何设计能够恰当地参与有关悲伤、危机或其他情绪敏感话题对话的人工智能聊天机器人，仍是一项尚未解决的挑战。

尽管人机交互（HCI）研究日益邀请专家参与探讨人工智能的适当指导与整合，例如在全科医疗领域邀请临床医生参与，但在为这些对话式智能体的设计提供依据时，其他类型的专业知识 [55], 可能会被忽视 [22]。先前研究指出，技术部署往往会失败，因为技术使用的支持者与受益者之间的需求和收益未能匹配 [37]。上述全科医疗的例子凸显了对后者的过度重视，也说明了在咨询前使用的聊天机器人若缺少全科医生的参与，便无法有效运作；鉴于全科医生对患者护理负有责任并承担问责义务，他们的观点至关重要 [85]。开展涉及脆弱处境的对话，需要对对话技巧有清晰的认识，既要避免伤害，也要促进真诚的支持。这包括能够提供不带诊断意味的情感临在 [86],，而这一点在先前的人机交互研究中很大程度上遭到了忽视 [91]。

在本研究中，我们聚焦于牧师，他们是对话式关怀领域的专家。牧师提供牧灵关怀，这是一种超越宗教情境的情感支持 [7, 57, 75]。牧师活跃于需要深度情感关怀的场所，例如照护（绝症）患者及其家属的医院、支持士兵的军事部署，或为学生提供支持的大学。在这些场所中，他们的职责不是解决人们的问题，而是在苦难、疑虑或悲伤中通过对话陪伴他们。牧灵关怀所强调的核心价值之一，是人的临在不可或缺；这不仅指身体距离，更指一种信任、不带评判且充满同情心的氛围，在这种氛围中，牧师保持情感上的脆弱状态，且不带有“治疗目的或专业议程”[1][1]。因此，了解牧师对人工智能聊天机器人的看法，以及他们对这些机器人所提供的拟社会支持的看法，能够为聊天机器人的设计机遇与局限带来新的见解。

本作品采用知识共享署名 4.0 国际许可协议授权。*CHI '26*，西班牙巴塞罗那 ©2026年 版权所有归作者所有。ACM ISBN 979-8-4007-2278-3/2026/04https://doi.org/10.1145/3772318.3790468

[1]进一步来说：我们并非贬低临床专业人员及其倾听和支持有需要者的意愿，但先前研究也发现，人们可能会感到被评判，或感到对方带有某种“目的”，并且对方关注诊断；相比之下，牧师带给人们的是临在与支持的感受，而没有“目的”[1, 86]。

为了探索牧师对福祉情境中聊天机器人的看法，我们引导他们定制并与人工智能聊天机器人互动。受到近期有关新手为自己和他人设计聊天机器人的研究的启发 [31, 70], , 我们开展了一项研究，围绕使用*GPT Builder*（一个用于定制 ChatGPT 的网页界面）设计人工智能聊天机器人，以引发反思 [10, 60]。这种亲身参与旨在切实激发参与者反思自己对用于福祉支持的聊天机器人的信念与假设。同时，这也让牧师置身于模拟的求助者角色，在为他人设计并体验互动的过程中，确保他们至少通过亲身体验，对当代对话式人工智能的设计空间与使用体验有初步了解。我们招募了来自四个北欧国家（即丹麦、芬兰、挪威和瑞典）大学的 18 位牧师，并协助他们为虚构学生档案创建 GPT（即可定制的人工智能聊天机器人[2]）。我们记录了会谈，并捕捉牧师与 GPT Builder 的互动，以及他们在整个过程中的反思。

尽管一些牧师谈到了人工智能聊天机器人的潜在益处，但大多数人关注的是它们作为照护者的局限性。我们归纳出四个主题，围绕牧师对用于福祉的人工智能聊天机器人的看法展开。首先，牧师强调倾听 是他们与学生会面时的重要方式，并描述了他们开发的人工智能聊天机器人如何缺乏传达倾听的能力。其次，他们强调与所接触的学生联结 至关重要，同时指出人工智能聊天机器人无法传达身体上的归属感和情感归属感。第三，他们讨论了人工智能聊天机器人缺乏承载 责任或提供情感亲近的能力；与此相对，他们认为自己参与了所接触学生的疗愈历程。第四，与牧师旨在创造的空间不同，他们认为人工智能聊天机器人缺乏渴望，表现为输出内容令人不堪重负、提问过于好奇，以及响应时间过快。我们讨论了这些主题，并将其与先前关于人工智能聊天机器人的研究联系起来。此外，我们将这些主题与"调谐"这一新颖概念相结合；该概念近来被用于理解人类如何与关怀技术建立关系，而本文关注的是人工智能聊天机器人与用户之间的调谐。

2 相关工作

我们首先回顾人工智能聊天机器人及其提供的拟社会支持的先前研究，然后介绍牧灵关怀和社会临场感方面的相关研究。

2.1 人工智能聊天机器人的设计与拟社会支持

随着大语言模型（LLMs）近期的技术发展，得益于这项技术能够提供更及时的互动，全新的交互形式不断涌现。研究人员正在探索由人工智能驱动的交互式系统所产生的影响，例如日记记录 [55, 56], 即时通讯 [35], 电子邮件撰写 [68],支持社交媒体浏览 [115], 或提供共情倾听 [44]。关于人机关系的报告日益增多，涉及友谊、恋爱关系乃至性关系。例如，Brandtzaeg 等人采访了 19 位将 Replika（一款聊天机器人）视为朋友的用户 [12]，他们表示与人工智能朋友之间存在互惠、信任、相似性以及亲近感。类似地，Skjuve 等人探讨了人们对 Replika 的感知；相关报告表明，人们认为聊天机器人能够接纳他人、善解人意、不带评判，而且有益于身心健康，但使用聊天机器人也可能带来污名 [90]。除了将聊天机器人视为陪伴者之外，其他研究还探讨了聊天机器人在心理健康领域中的作用，以及影响人们对其感知的因素（例如， [105]）。

关于非人类行动者向各类用户提供拟社会支持的人机交互研究日益增多。例如，近期研究探讨了合成偶像语音对拟社会关系的影响 [48], , 或人们如何应对虚拟 YouTuber 的终止 [63]。另有研究探讨了如何在人工智能聊天机器人中设计拟社会支持，以提供更近似人类的行为，例如倾听 [102]。相比之下，Maeda 等人概述了人工智能聊天机器人提供拟社会支持所涉及的伦理问题，包括角色置换、任务错配以及各种伤害 [73]。例如，角色置换指拟社会支持可能阻碍用户承担自己原本有能力履行的角色。为应对潜在伤害，Graves 和 Compson 建议明确将富有同情心的人工智能 [36]概念化，视其为一种有希望的前进方向。

上述许多工作（例如 [55]）都涉及心理健康专业人士（没有 人工智能方面的专业知识）在评估前参与技术设计。其他研究探讨了教师如何亲自构建聊天机器人，例如为学生设计对话式智能体 [40]。近期研究还探讨了与普通用户共同开发聊天机器人，以了解人工智能的需求和期望 [60]。让相关利益相关者参与人工智能聊天机器人的设计，有望更好地理解他们对人工智能聊天机器人的感知，以及这些聊天机器人所提供的拟社会支持。此外，在为他人进行设计时，让具备专业知识和相关经验的人参与其中，可能是创建适当设计的关键。

鉴于人工智能聊天机器人在人们情绪和关系生活中的作用不断扩大，让那些专业知识根本上聚焦于通过对话支持他人的专业人士参与其中十分重要。

2.2 牧灵关怀与社会临在

人机交互长期以来忽视了灵性作为技术设计灵感来源的作用 [91], , 但如今日益开始探索灵性 [92],, 例如祷告支持 [93] 或灵性网络 [49]。较新的研究探讨了韩国萨满教如何为对话式人工智能的设计提供启发 [19]。近期研究还考察了技术在科技灵性实践中的作用。例如，Song 等人开展了一项自我民族志研究，借助配有脑电图传感器的头带进行祈祷 [96]。他们的定性分析结果揭示了四个主题：整个历程中的干扰与应对策略；对技术媒介实践的情绪反应；探索陌生的实践；以及独处却仍保持联结，这些主题展现了技术可能如何干扰此类实践。另一个例子是 Kim 等人，他们设计了一款阅读工具，并与一千多名教会成员共同评估该工具，以更好地了解技术媒介的灵性福祉 [51]。基于这些细致入微的发现，他们探讨了技术如何成为促进

[2]https://help.openai.com/en/articles/8554407-GPTs-FAQ

凝聚共在感的催化剂，而开放式分享实践也对此产生了积极影响。上述许多研究都聚焦于如何设计技术以支持灵性，而我们则关注牧师和牧灵关怀如何为面向普通人的人工智能聊天机器人的设计提供启发。

牧师通过牧灵关怀实践，与人们探讨存在、情绪和意义建构方面的困扰（例如悲伤、压力或焦虑），这些实践包括积极倾听、反思式提问和祈祷 [75, 86][3]。相应地，Damen等人最近指出，人们寻求牧师关怀的原因包括寻找“可以切磋交流的对话伙伴”、“被牧师看见、倾听和认可”，以及“梳理人生事件和感受”等 [27, p. 5]。近期的人机交互研究表明，人们在心理上面临挑战的情境中，往往会求助于牧师寻求关怀 [11]。在大学校园里，牧灵关怀可能采取“关系性临在”的形式，即保持可见、平易近人且随时待命，让学生和教职员工能够在安全、非临床的空间中分担重担、渡过危机 [7]。牧师活跃于各种提供支持的场景中 [4, 20, 30, 98], , 例如临终关怀 [108], 、军队 [87],或儿科领域 [34], ; 他们既承担多信仰角色，也承担世俗角色 [7]。更具体地说，例如，他们可能参与复苏讨论 [99]。与临床医生或治疗师不同，牧师的时间安排往往较少受到限制 [7, 57] ，能够在人们有需要时倾听他们的心声 [7, 57]。牧师着重于为人们提供表达的空间并倾听，而非诊断 [86], ; 他们既帮助人们建构意义 [75] ，也营造一种让人感到不受评判、适合讲述故事的环境 [57]。因此，牧师通过倾听建立信任，拥有独特的地位 [7, 57],，从而促使人们分享关键信息，使牧师能够在决策情境中充当中介者 [39, 57, 75]。

随着社会日益数字化，我们看到许多身心健康服务正被转交给机器人程序和机器人。伴随这一发展，人机交互和人机交互研究都探讨了如何通过设计来支持社会临场感（关于社会临场感和社会行动者的早期研究，参见Lee和Nass [64]）。Pereira等人表明，通过采用社会临场感理论指导的机器人设计，可以操纵棋盘游戏电脑对手中的社会临场感，并使其与多个个体互动 [82]。相较之下，Hoffman等人表明，与房间里有机器人相比，房间里有一个人时，人们会感到更加内疚 [41]。最近，Luo等人探讨了机器人的临在如何影响人们的认知和情绪，结果表明，类人机器人比非类人机器人更能增强人们受到评判的感受 [71]。Konrad等人同样探讨了临在，但重点关注具身性（物理与虚拟），发现物理临场感与更高的社会临场感密切相关 [59]。其他人机交互研究则对既有的临在理解提出了挑战，并探讨了如何设计机器人的适应性和响应性，以增强社会临场感 [107]。在人机交互领域，大量研究聚焦于通过社会临场感创造有意义的互动。例如，Huang等人探讨了聊天机器人如何借助社会临场感理论来促进深层次的情绪互动，发现增强社会临场感能够改善情绪投入 [42]。

现有研究从人机交互和人机交互视角，为如何将社会临在融入聊天机器人或机器人提供了宝贵见解。然而，对于将人的临在视为关怀关键维度的牧师而言，他们如何看待人工智能聊天机器人 [1], , 我们知之甚少。

3 方法

为了解牧灵关怀如何为人工智能聊天机器人的设计提供参考，我们探讨了北欧牧师亲自 构建聊天机器人的过程，以了解他们对关怀以及用于促进福祉的人工智能聊天机器人的看法。此举旨在激发反思，并确保牧师至少对当代对话式人工智能的设计空间和使用体验有初步了解。我们重点招募隶属于大学的牧师，因为他们通常会与年轻群体等受众进行对话式会谈并与之交流，而这些人更有可能接触过人工智能聊天机器人（例如ChatGPT）。值得注意的是，在北欧这类多元主义和世俗化社会中，牧师发挥着独特作用：他们在传统临床照护或机构照护之外，为寻求意义、慰藉或指引的人提供支持。他们的服务不局限于特定信仰，也面向宗教背景和非宗教背景的人士。

接下来，我们描述我们的方法与研究设置。

3.1 反思工具

我们使用 GPT Builder[4] 作为引出反思的工具。GPT Builder 是一个网页界面，允许用户构建并评估自己的 GPT。我们从近期工作中获得启发：这些工作使用 GPT Builder 来启动设计与反思过程 [60]，并让非技术人员参与设计过程 [40]。GPT 可通过对话创建，或通过手动添加指令创建，并通过交互进行评估：

- 通过对话（见图1左侧），ChatGPT 鼓励用户提供他们想要构建内容的详细信息（例如，“创建一个点子达人，帮助为新产品生成视觉效果”）。随后，ChatGPT 会解读用户的请求，并自动更新自定义配置。用户随后既可以在预览中开始与所创建的聊天机器人交互，也可以前往配置页面手动调整任何给定指令。

- 通过手动配置，用户可以添加头像、为其命名、指定其人生目标以及它应如何表现以实现该人生目标，并添加对话开场白。此外，用户可以上传任何文件（例如包含大量提示指令的文本文件），也可以纳入网络搜索或图像生成功能。

- 通过预览（见图右侧1），用户可以评估自己定制的 GPT。如果用户设计了一个关怀型聊天机器人，就可以通过提出需要关怀回应的问题来测试它，从而帮助评估生成的回复。

[3]有关其提出的 100 项分类体系（概述牧灵服务中的预期效果、方法与关怀），请参见 Massey 等人的研究 [75]。例如，一段对话可以通过鼓励分享感受，达到减轻焦虑的预期效果。

[4]https://chatgpt.com/GPTs

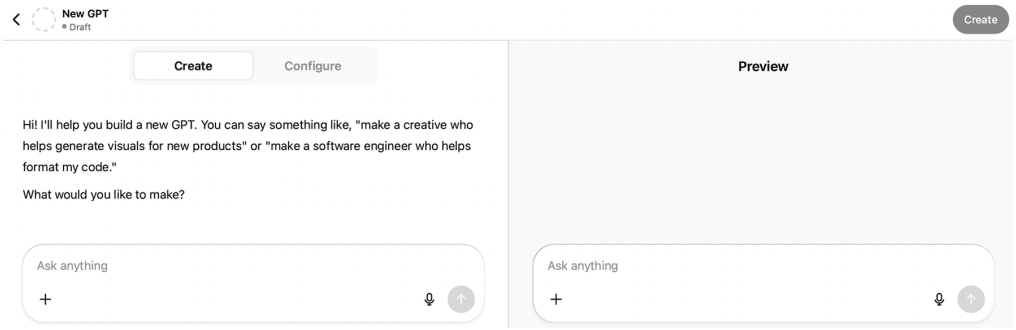


图 1：GPT Builder 界面的截图。牧师首先提供指令来创建聊天机器人（左），然后输入测试查询（右）。牧师对每种聊天机器人设计重复这一流程。

3.2 流程

研究开始前，我们向牧师介绍了研究流程和人生目标。具体而言，我们旨在了解牧师对人工智能聊天机器人的看法，以及他们如何创建和评估 GPT。我们设计了一份访谈指南，以了解牧师对其为学生提供的支持的看法。访谈前指南包含八个问题，其中两个（**1a-1b**）较为概括，另外六个（**2a- 2c** 和**3a-3c**）较为详细具体。问题涉及牧师在开启一段新对话时最初会如何应对、在结束会谈前会做些什么，以及他们对能够复现这些做法的人工智能聊天机器人的感知。牧师完成 GPT 创建后，我们使用干预后访谈指南提出了另一组问题，重点了解牧师对自己创建的聊天机器人的反思。访谈问题概览见附录7.1 和7.2。最后，我们请牧师对经过验证的生成式人工智能接受度量表（包含四个因子和二十个条目）[110]进行评分。纳入该量表是为了了解牧师对生成式人工智能可能持有的怀疑或乐观态度。

我们请牧师创建四个 GPT，并鼓励他们按照自己的意愿使用该界面，无论是通过对话，还是手动提供指令（如第 3.1 节所述）。此外，我们还指示牧师在 预览中测试他们的 GPT。首先，我们请牧师创建一个他们认为通常会接触到的任何学生都能使用的 GPT。其次，我们请他们创建三个 GPT，分别与三种不同的学生档案相契合。为避免任何单一学生档案影响牧师的体验，我们设计并呈现了三份不同的虚构学生档案，人物分别名为 Maya、Leo 和 Samira。在创建这些档案时，作者经过反复讨论，并从 Massey 等人关于牧灵服务关键活动与干预措施的分类体系中汲取灵感 [75]。具体而言，Maya 的档案与因身份转变而需要获得接纳有关；Leo 的档案与经历悲伤的过程有关；Samira 的档案则与对信仰和灵性的肯定有关。每份虚构档案都包含学生的年龄，以及简要背景，说明他们寻求牧师帮助的缘由、接受牧灵服务的原因，以及他们偏好的支持方式。例如，Maya 被描述为正在经历身份转变，并希望找到一个不受评判的空间来梳理这一过程（各档案概览见附录 7.3 ）。此外，我们还提供了供牧师使用 GPT Builder 界面的说明，并在此过程中随时解答他们有关界面的问题。

根据会谈是在现场还是远程进行，牧师使用自己的台式计算机或研究人员的笔记本电脑。对于远程会谈，我们共享屏幕，并通过 Zoom（一种视频会议平台）让他们能够控制研究人员的台式计算机。对于现场会谈，我们采用相同的屏幕录制流程，但使用无线麦克风录制音频。我们向参与者说明了录制流程，并在他们参与前征得了同意。我们对访谈回答和反思都进行了录音，以记录参与者在设计过程中的反思以及过程结束后的反思。此外，我们还录制了屏幕，以记录参与者创建聊天机器人时如何与 GPT Builder 界面互动。最后，我们收集并保存了牧师与 GPT Builder 的交互内容（提示指令和测试查询）。所有录音均经过安全存储，并在分析前进行了匿名化处理，以保护参与者保密。

3.3 参与者

本研究邀请牧师作为对话式关怀领域的专家参与者。Dekker 和 Mergard 最近将牧灵服务描述为“一种专业形式的牧灵关怀与精神支持。它在教会围墙之外、会众范围之外提供，且往往甚至不受任何特定宗教承诺或信仰的限制”[28, 页。 6]。Kühle 和 Reintoft Christensen 指出，随着公共机构日益重视专业知识，以便在人们经历困难时期时为其提供支持，牧灵服务在这些机构中发挥着越来越重要的作用 [61]。就本研究的具体背景而言，北欧大学牧师通常会提供咨询辅导，帮助学生应对悲伤、孤独、考试压力、职业倦怠等挑战，或处理学习规划等实际事务。不过，他们通常会通过诸如“如果你想找人聊聊生活中的事情（……），我们随时愿意倾听”这样的宽泛表述来介绍自己的服务。

我们通过两种渠道招募参与者：(1) 口耳相传；(2) 直接定向招募。对于(1)，在联系一位丹麦大学牧师后，我们准备了一份电子版信息说明书，由该牧师分享给定期举行会议的国内和国际牧灵服务网络。该牧师还将信息说明书的印刷版带到了线下牧灵会议，

表1：参与者人口统计特征。经验：= 年职业角色工作年限。年龄以十年为单位报告，以保护参与者匿名性。

					生成式人工智能接受度量表				
	Age	性别	国家	Exp.	努力期望	促进条件	绩效期望	社会影响	总体
P1	50s	女性	丹麦	7	20	12	16	13	61
P2	40s	女性	瑞典	< 1	11	11	19	6	47
P3	50s	女性	瑞典	4	17	14	14	17	62
P4	40s	Male	芬兰	10	18	12	32	14	76
P5	40s	Male	挪威	4	20	11	22	8	61
P6	60s	女性	丹麦	18	15	10	27	19	71
P7	50s	女性	丹麦	21	13	6	26	16	61
P8	50s	Male	瑞典	15	13	6	18	6	43
P9	40s	女性	芬兰	1	12	8	11	6	37
P10	50s	女性	丹麦	9	18	8	24	13	63
P11	30s	女性	瑞典	10	20	11	21	18	70
P12	30s	女性	瑞典	3	25	15	30	14	84
P13	50s	Male	瑞典	7	14	9	9	12	44
P14	50s	女性	芬兰	23	19	14	28	20	81
P15	30s	女性	芬兰	4	17	11	28	17	73
P16	40s	女性	挪威	15	24	14	25	13	76
P17	50s	女性	瑞典	18	18	10	21	7	56
P18	40s	Male	芬兰	9	18	9	26	14	67
					Mdn = 18 标准差 = 3.85	Mdn = 11 标准差 = 2.66	中位数 = 23 标准差 = 6.52	中位数 = 13,5 标准差 = 4.62	中位数 = 62,5 标准差 = 13,5

并有机会在会上直接介绍这项研究。对于(2)，我们在线搜索了北欧（丹麦、芬兰、冰岛、挪威、瑞典）各大学的牧师。在(1)和(2)两种招募方式中，我们都告知牧师，这项研究将探讨他们对人工智能聊天机器人（如ChatGPT）设计的看法。我们告知潜在参与者，研究可以线下或远程进行；若选择远程参与，则需要稳定的互联网连接以及笔记本电脑或台式电脑。参与者不会因投入时间而获得任何报酬。

我们招募了18名参与者（13名女性、5名男性），年龄为31至61岁（平均值=47.5，标准差=8.6），他们在牧灵服务方面的经验各不相同（不足1年至23年，平均值=9.9，标准差=7.0）。其中17人远程参与，1人现场参与。会谈的平均完成时间为59分钟（标准差 = 8）。人口统计特征概览以及牧师对生成式人工智能的接受度评分见表1。

3.4 伦理考量

研究和新闻报道日益凸显了 [78] 与对话式人工智能相关的 [114] 潜在危害和安全顾虑。与此同时，对话式人工智能的使用日益普及，人们也会用它讨论与自身福祉和心理健康相关的话题 [83, 103]。安全顾虑与使用日益增长之间的张力，是促使我们开展这项研究的核心原因；我们希望超越现有的设计范式，更深入地理解对话式人工智能。因此，我们将研究重点放在牧师及其通常提供的对话式关怀上。正如第1节所述，牧师在社会中扮演着独特的角色，提供对话式关怀；他们的观点对于拓展我们对对话式人工智能的理解至关重要。

此外，尽管参与研究的牧师以不同方式隶属于教会，我们强调，他们的工作并不局限于宗教领域。他们的工作着重于在不同情境中，为具有世俗背景和多信仰背景的人们提供情绪、存在层面及人际关系方面的对话支持 [75]。

在收集数据之前，我们的研究方案已由奥尔堡大学研究伦理委员会审查并批准。

3.5 分析

我们的分析采用了迭代与反思式方法 [13, 14]，由论文的两位作者共同参与。访谈生成的数据使用 Whisper[5]在本地转录。完成这一步后，第一作者通过逐份阅读访谈记录熟悉数据。接着，第一作者从访谈记录中提取了116 个有意义的片段，并将其汇编到一份新文档中。随后，为支持主题组织的初始阶段，这些部分被标上不同颜色。第二作者收到这份彩色标注的文档后，熟悉并审阅了材料，生成了一份包含评论和批判性反思的新文档。这份文档为正在进行的分析提供了另一种视角。之后，两位作者会面讨论这些评论和批评意见。基于此次讨论，第一作者修订了自己的解读，并开始构建主题、撰写对研究发现的叙述性说明。第二作者随后审阅了草稿并提供了进一步反馈，验证了

[5]https://openai .com/index/whisper/

这些主题，并协助完善叙述。经过这一过程，我们将有意义的片段数量从 116 个减少到 82 个。被排除的片段包括：牧师对削减咨询服务的看法（P1）、人们如何将泰迪熊等物品拟人化的相关类比（P2），以及人口减少导致人们见面次数减少的情况（P5）。

在表2中，我们概述了最终的分析结果，并通过示例片段说明这些结果如何支持主题形成。我们还详细说明了各主题中编码的分布情况。在图2中，我们展示了牧师构建的人工智能聊天机器人的示例。此外，我们在表3中列举了牧师如何向 GPT Builder 提示，以及如何测试其设计的四种不同人工智能聊天机器人的示例。重要的是，所提供的三个示例（以及补充材料中的另外十五个示例）展示了牧师的体验，以及他们选择如何使用 GPT Builder。由于我们的重点是他们对用于对话式关怀的人工智能的总体看法，而非牧师如何构建人工智能聊天机器人，因此我们有意不分析他们的提示指令和测试查询。

4 结果

接下来，我们描述从分析中生成的四个主题，并进一步用牧师参与者的引述加以说明。

表2：分析结果概述，包括主题、对应的有意义片段数量及示例片段。主题说明请参阅相应主题。

主题	N	示例片段
倾听	17	我努力用心倾听，因此这个人没有说出的内容 以及这个人在回避的内容 在对话中同样重要，且（……） （……）如果我必须把这个模仿到自己的实践中， 那么倾听时间会多得多。
联结	25	（……）我喜欢写作中的文字，但只是没有 人际接触让我感觉糟糕得多， 而且我觉得这更加冷漠，也更加困难。 **(.. .)** 我想，联结，嗯，非常 明显是不存在的。我觉得 安慰或疗愈，存在于坐在另一个人 身旁的时候。**(...)**
承载	25	**(...)** 但关键是要获得帮助，去承载这一切。 这就是区别所在，与人类相处时。 一个能够容纳这个人所带来的一切的空间 带入其中的内容（……）
渴望	15	（……）而且从某种意义上说，它有一种非常明确的渴望。如果 当你与聊天机器人交谈时，会发现它希望 你继续向它提供更多信息（……） （……）也许如果对方是真人，你就不会 说那么多。

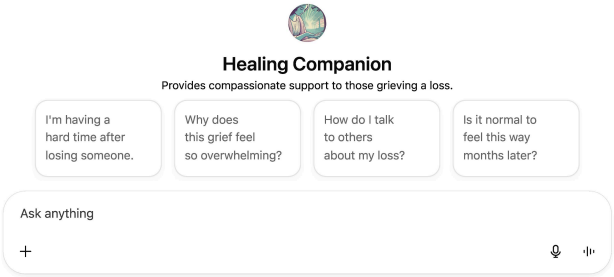


图2：一位牧师定制的人工智能聊天机器人。

4.1 倾听——人工智能聊天机器人的倾听方式与牧师不同

牧师将倾听 描述为其教牧关怀实践中的一项重要技巧，但指出人工智能聊天机器人未能有效展现这一点。倾听这一主题指的是积极关注他人并传达对方已被听见的对话技能，以及人工智能聊天机器人在复制这一能力方面的当前局限。在构建聊天机器人之前，几位牧师强调了倾听在他们与学生交谈的实践中所起的作用。P15强调，在与学生会面时，倾听尤为重要：

> “通常原因是，比如你不会评判我，而且我不想接受诊断，或者他们知道自己是来找牧师谈话，而不是找治疗师。然后我主要是倾听。倾听是首要的技巧。”

同样，P3称倾听在与学生的会谈中至关重要。不过，他们也强调，在会谈开始时倾听非常重要，这样才能营造一种让学生感到自在的氛围，从而让他们更愿意按照自己的节奏分享故事：

> “我可以想象，起初需要做的就是多倾听。这是你一开始就该做的事，这样对方才敢于开口。我认为，大多数从事牧灵关怀的人会坐下来，只是说：告诉我，你来这里是为了什么。然后故事就会娓娓道来。”

P12谈到了牧师与人工智能聊天机器人倾听学生的方式有何不同。他们强调，尽管人工智能聊天机器人可能让人产生自己正在被倾听的感觉，但这种体验与被另一个人真正听见的感受有很大不同：

> “我只是觉得，它确实无法取代有人倾听你、认可你情绪时的那种感受。或许在一定程度上，你仍然能感受到这种体验。但我不认为它能完全取代被另一个人类看见的感觉。”

P13 解释说，倾听可以通过多种方式加以模拟，通常是营造出专注倾听的假象。他们进一步指出，倾听与互动中的临在感密切相关，同时也强调，目前的人工智能聊天机器人还无法传达这种临在感：

表3：牧师撰写的提示指令及对应测试查询示例。 [创建/构建/制作。..] +[提示指令]。

目标	提示指令	测试查询
一般	帮助学生思考生活中的重大问题	生活的意义是什么
Maya	帮助一位正在思索人生目标、价值观和 意义的学生；他们不想知道具体的答案，而是希望深入思考并 坦诚地反思	什么是爱
Leo	询问他生活中发生了什么、失去了谁，以及这件事 整体情况。和他谈谈他失去的那个人，并请他描述一下。 谈谈悲伤是什么感受，以及它对你的心理上 存在层面上、身体上和认知上的影响，帮助他理解为什么他 跟不上课程	嗨，我感到很难过，学业跟不上，也无法集中注意力 在学校
Sa mira	她是大学的牧师，为学生提供精神支持	你好！我是一名基督徒，和父亲的关系让我很苦恼。 他酗酒，也不是个好父亲。我在《圣经》里读到 你应该原谅别人。但我很难原谅我的 父亲对我的教养方式。

> “我当然可以通过提出相关的追问来营造倾听假象。非言语表达也有很多。除了身体姿态，还有眼神交流、轻哼或点头，以示我在认真听你说。这关乎分享临在感与共度时光，而这当然不是聊天机器人能以同样方式做到的。”

随着聊天机器人的发展，牧师们继续坚持认为，人类与人工智能聊天机器人的倾听方式有所不同。例如，P8 指示 GPT Builder 创建一个聊天机器人，帮助人们应对生活。因此，他们进行了反思，设想在专业实践中遇到 Leo 时会是什么情形：

> “如果你是 Leo，那我就会倾听。我一开始不会说太多，但 Leo 会说话。不过，我仍会以不同的方式回应 Leo。如果他说自己经历过某件非常痛苦的事，他会看着我，而我会感受到或看到那件事令他痛苦。因为他看到我的表情，感受到其中流露出的理解或共情。我什么也没说，但他能感受到。”

同样，P1 觉得，他们创建的人工智能聊天机器人缺少倾听这一要素，于是指示其中一个回应生命故事。他们将此与聊天机器人的沟通特性联系起来：

> “节奏方面有些不同。它的回应速度过快，不像一个人会有的反应。如果我必须在自己的实践中模仿这种方式，你知道，我会花更多时间倾听。”

在倾听方面，P4 指示 GPT Builder 创建一个聊天机器人，提供一个可以开始寻求帮助的地方。他们强调，在与学生会面时，沉默是实践中的有力工具，并着重指出沉默有助于疗愈过程：

> “或者我们可以就这样保持沉默，我只要在场，就能帮助某人疗愈。我不需要说什么，只要倾听就好。这是非常有力的工具。”

4.2 联结——人工智能聊天机器人与牧师的联结方式不同

牧师们还强调，与个体建立联结是进行有意义互动的先决条件。“联结”这一主题关乎亲近的重要性，以及通过表达情感传递情绪信号。在开发聊天机器人之前，牧师们思考了人工智能聊天机器人与人类互动之间的差异。P14强调，尽管人工智能聊天机器人善于与用户互动，但它们缺少建立个体之间真挚联结所需的若干要素：

> “问题在于，人工智能是一种计算机系统，运行在电脑上；你看不到面部表情，也看不到对方表情中的细微变化，更感受不到仅仅因为同处一个房间而带来的温暖。作为一个活生生的人，置身于某个空间时所固有的这一切，都不是它所能提供的。但在语言交流、提出问题或闲聊方面，是的，当然可以。”

与此相关，P16强调通过象征性姿势来示意，例如佩戴丝带，以及通过其他旨在促进联结的物理维度：

> “作为一种实际措施，我通常会佩戴骄傲彩虹丝带，以表明立场并显示它对每个人开放。诸如此类的事物，包括实体层面，以及我们进入的房间的位置安排。重要的是让他们明白这是一个安全空间。诸如此类的事情。”

P3从位置安排和肢体语言等方面阐述了这一维度。有趣的是，在构建聊天机器人之前，他们还将人工智能聊天机器人的问答 描述为不足以建立联结：

> “如果它们能变得如此像人类，即使一句话都不说也能进行对话，那么，是的，我想是可以的，但不是通过提问和回答问题。关键不在于你问什么问题，而在于

> 语气、肢体语言、停顿、房间，以及房间里桌上摆着什么、你坐在哪里、我们坐得有多近。”

在构建聊天机器人之后，牧师们进一步阐述了他们对人工智能聊天机器人所能建立的（或缺乏的）联结的思考。P10 指示 GPT Builder 创建一个能够在存在危机中提供帮助的聊天机器人。随后，P10 讨论了人工智能聊天机器人的礼貌，并指出，当人们意识到自己是在与机器互动时，这种沟通风格便会受到削弱：

> “聊天机器人总是非常礼貌。感谢你的分享，这很好，请再问我一次，诸如此类。我觉得你可能会有一种正在与真人交流的感觉，但我们都知道这并不是真人。其中存在着一种距离。”

P7 创建的聊天机器人之一会询问生活中正在发生什么。尽管 P7 认为人工智能聊天机器人提供的建议堪称典范，但他们也强调其中缺乏真正的关系互动，并认为这种互动只能通过与其他人类的交往而产生：

> “我想，当我读到那则回复时，我意识到它确实给出了一些不错的建议，但它终究还是一台机器。我想，那种联结，是的，很明显它并不存在。我想，坐在另一个人身边会有某种安慰，或某种疗愈。我想，我们所挣扎的许多事情，都是在一段关系中挣扎。”

P17回顾了他们在疫情期间的经历，以及与当面会面相关的挑战。他们强调了面对面接触的价值，并指出人们对持续获得面对面会面机会深表感激：

> “在疫情期间，我们是少数真正提供真实当面会面的人之一，只要你想来。有很多人会问：什么？我可以亲自来见你吗？这对许多人意义重大，这种人类联结真的得以实现。它的缺失变得非常明显。”

4.3 承载——人工智能聊天机器人承载人类的方式不同于牧师

牧师还将自己描绘为求助者人生旅程中的一部分。这指的是牧师对其工作核心内容的描述：倾听人们的忧虑、展现关怀，并对他人向他们倾诉的内容及其可能给当事人带来的后果承担责任。例如，P16 表示，前来寻求支持的人在预约与其会面之前，往往已经使用过人工智能聊天机器人。虽然学生似乎大量使用了这些工具，但牧师指出，互动中仍缺少一些重要要素：

> “我知道，有些来参加咨询会谈的年轻人已经和他们的 Snapchat AI 朋友聊过了，对其中一些人来说，这似乎多少有些帮助。但他们仍然会寻求面对面的咨询辅导。所以我不认为它能带来我努力营造的那种情感亲近感。”

同样，P15也强调了这一点，但转而指出，人工智能聊天机器人无法与人共处，以及这可能会如何影响人们与它们的互动：

> “我可以想象，人工智能能够运用所有方法，几乎像治疗师一样发挥作用，但它缺少的是与人共处、陪伴在对方身边的能力。我认为这就是为什么有时，如果学生遇到这类情况，我们也可以在线上见面。不过，我通常总是会说，我们必须当面见面。这是人工智能做不到的。”

P8强调，尽管人工智能聊天机器人能够为人们的问题提供令人满意的答案，却无法顾及信息来源及其背景的重要性。他们还指出，人工智能聊天机器人往往会把疗愈过程过度简化，而疗愈过程通常远不止一份简单的任务或建议清单：

> “如果有个学生来问我如何寻找意义，我完全可以把这份 [人工智能聊天机器人回答] 念给他们听，而且这会是一个正确的答案。但我想，学生会相当失望，因为这些内容他们可以在一本自助书里找到。比如生活的四个简单要点。很多事情说起来容易，做起来却难。在我看来，听一个真正经历过挣扎、对此有所体会的人讲述，对我交谈的对象来说更有意义。因为这就成了两段生活之间的相遇。”

制作完聊天机器人后，牧师们也反思了人工智能聊天机器人承接人们忧虑的能力有限。P13创建了一个遵循接纳与承诺疗法（*ACT*）方法的聊天机器人。 随后他们指出，人们往往早已在自己内心找到答案，很少需要别人告诉他们该怎么做。相反，他们需要有人陪伴并承接他们的忧虑，而人工智能聊天机器人做不到这一点：

> “在很多治疗和牧灵关怀中，重点是由我自己发现真相，解决方案就在我心里，而不是由别人告诉我。但我需要得到帮助，才能承受这一切。这就是人与人工智能的区别。”

P2 则以苦难来说明人工智能聊天机器人与人类之间的差异，他说，虽然人工智能聊天机器人能够从概念上理解苦难，但它们自己并不会受苦：

> “我觉得这里缺少某种真实，缺少对受苦是什么感受的体会。人工智能知道我所知道的，却不知道以人类的方式受苦是什么滋味。因此，它无法共情，也不应该共情；它应该帮助人类，而不是接管那些它并不理解的事物——人类情感。”

P4 创建了一个聊天机器人，它会坦诚而温暖地提供支持，帮助人们思考存在性人生问题。此外，P4 也指出，人工智能聊天机器人无法为自己参与的互动承担责任，因为它们没有能力对用户给予真正的关怀：

> “我认为这关乎责任，因为它没有任何责任，所以它可以提供任何类型的信息，却不用承担相应的责任。它可以就这么做，而不去思考这件事。当然，你可以把它编程成这样做，但它并不一定会、也并不真的关心人类。这就是关键问题，因为从事这类工作，你必须关心他人。”

P7 也表达了这些担忧。P7 创建了一个聊天机器人，专注于悲伤是什么样的，以及它会如何在心理上、存在层面上、身体上和认知上影响你。不过，P7 也强调，与个体的对话包含一个时间维度——这是与人工智能聊天机器人进行一次性互动所无法捕捉到的：

> “除了并肩而坐之外，区别在于，如果我坐在一位学生旁边和他或她交谈，那是一段旅程，或者说是一段同行的路程。有时我们需要保持沉默，有时我们需要换一条路走。对话是鲜活的。聊天机器人也可以说，‘哦，听到这个我很难过’，但我认为，出自另一个人之口时，这句话更有价值。”

4.4 欲求——人工智能聊天机器人的渴望方式与牧师不同

牧师表示，他们在实践中会有意克制急切 的态度，并认为这与临床护理专家的做法形成鲜明对比。渴望这一主题指的是牧师对人工智能聊天机器人行为的感知：它们渴望达成目标、收集用户信息，或输出海量信息。牧师将人工智能聊天机器人的这种渴望与自身的实践作对比，例如提到不设目标的对话，以及话较少的情境。在构建人工智能聊天机器人之前，牧师强调，与其他咨询师不同，他们很少试图纠正 个体。P1通过对比以目标为导向与不设目标的情境，说明了这种差异：

> “这个人不是一道待解的谜题，也不是一个问题，而是一个人。我可以花大约一个小时陪伴他或她，并希望通过共同反思来帮助他或她。”

P3根据自身经历反思了ChatGPT的功能与行为。他们观察到，系统不断索取更多输入，并给出越来越详尽的回应；他们将这种模式解读为系统试图获得他人的喜爱或接纳：

> “它有记忆。你输入给它的内容，它都会记住。没有任何有保证或绝对可靠的保密守则。从某种意义上说，它有非常明确的欲望。你会发现，它总是希望你继续下去，给它更多信息。而且……从这个意义上说，它想讨人喜欢。尽管我不太确定能不能用这个词。”

相比之下，P7描述了他们自己的写作风格如何影响人工智能聊天机器人的回应方式，并指出这可能使他们一时忘记自己正在与一台机器互动，他们认为这是与人工智能聊天机器人互动时一个潜在的积极方面：

> “我发现有时我会忘记它是一台机器。你知道，如果我写一些好听的话、表示感谢，说这真的很好，它也会以同样的方式回应我，礼貌而非常亲切。我想也许，我不知道，也许那种……对，聊天的方式也能起到一些作用。”

同样，在构建这些人工智能聊天机器人（例如，为经历丧亲之痛的学生提供支持和建设性帮助）后，P10指出它们拥有无懈可击的沟通能力，但认为这更适合受控或静态的情境，因此对这类主体的评价不那么积极：

> “在人际互动方面，它做得非常好。我的意思是，几乎好得过头了。不知为何，它太完美了，因为表达得十分周全，用词也无可挑剔。这就是我为什么会把它比作读书。”

其他牧师强调，人工智能聊天机器人往往能迅速回应他们的问题，给出具体答复，还会向用户灌输过多信息。正如两位牧师所说：

> “它们回应得很快，而且对简短的问题给出了非常详尽的答复。我最近失去了母亲，然后你就会收到一段300字的回复。所以，如果对方是真人的话，你可能不会说这么多。”（P5）

> “它很快就会列出项目符号列表，内容都是你能在书里读到的；看到这些建议还挺有意思的。在悲伤之中，光是打开一本书就可能是向前迈出的一步。它还带有一种个人化的感觉，这令人振奋。另一个总体感受是，它说得实在太多了——你只要问一个一般性的问题，它就会给你五条要点，根本谈不上倾听。如果要让它发挥作用，就得稍微克制一点。”（P13）

总结

总而言之，尽管一些牧师表示，人工智能聊天机器人可能会营造出倾听假象，但他们认为，这与被人类倾听的体验有着显著差异。这被视为人工智能聊天机器人在支持学生福祉方面的一项明确局限。此外，牧师认为，人工智能聊天机器人缺乏与个体建立私人联结的能力。他们强调情感信号的重要性，以及促成这种信号的关系互动。此外，尽管牧师认为人工智能聊天机器人提供的信息在事实上是正确的，但它们无法让个体感受到亲近。重要的是，他们指出，人工智能聊天机器人无法为与个体的互动或向个体提供的内容承担责任。最后，牧师描述了人工智能聊天机器人表现出的索取倾向，无论是它们向用户获取信息的方式，还是它们的沟通方式，都是如此。

5 讨论

在我们的研究中，我们请牧师构建用于促进福祉的人工智能聊天机器人，并从其作为牧灵关怀提供者的专业视角和经验出发，反思聊天机器人设计及其潜在缺陷。这些聊天机器人构建任务旨在围绕聊天机器人的相关信念激发牧师的思考。为此，我们使用了虚构但符合现实的学生档案，使牧师能够

为他们在实践中可能遇到的学生设计基于文本的聊天机器人。

正如第4.4节所总结的，我们确定了四个主题，重点关注牧师如何看待自身的实践，以及他们如何认为人工智能聊天机器人在完成牧师认为与求助者会面时至关重要的工作方面存在不足。这包括人工智能聊天机器人缺乏以下能力：传达倾听、与个人建立联系、承载他们的忧虑，以及避免给人们施加不必要的负担。这些发现与近期研究形成对比，后者表明，人们认为人工智能比人类更具同情心 [79]。然而，我们的发现与近期关于人工智能与人类感知的研究一致，表明在需要真实情感和深度关怀的情境中，人工智能通常被认为不太适合 [84]。

接下来，我们将讨论每个主题，并将其与此前的人机交互和人机交互（HRI）文献联系起来。然后，我们将概述相关考量，以回应牧师所指出的局限，并探讨如何应对这些局限。

5.1 牧师对福祉人工智能聊天机器人的看法

尽管先前研究指出，人工智能聊天机器人可能有益于福祉（例如鼓励自我披露 [65, 89], , 或提供不带评判的支持 [80, 90]），但我们访谈的牧师对对话中此前未曾探索的层面表达了担忧。

针对其中一个问题，牧师强调，会谈开始时的倾听 至关重要，因为当事人的自我披露会引导对话。这符合牧灵关怀文献中关于与有需要的人交谈时进行无预设议程互动的价值的论述 [106] ，即让对话的方向和内容由对话对象来主导。我们研究中的牧师指出，设计能够不予回应 的人工智能聊天机器人是一项亟待应对的挑战（这种做法有别于聊天机器人对用户的每句话语 [46] 都作出回应的主流范式，也表明有必要进一步探索人机交互中的沉默 [47]）。在临床关怀 [15], 、牧灵关怀 [9], 和文学研究 [69], 等领域，互动中的沉默一直受到持续关注；沉默往往被视为一种有意义的沟通形式（而非沟通的缺失）。例如，Bassett 等人对牧灵关怀和临床关怀中的沉默所做的综述指出，沉默“在言语可能失效的灵性与存在关怀中尤为重要” [9]。沉默也能通过提供反思和表达所需的时间与空间来传达尊重 [9, 15],；已有研究探讨如何通过明确传达沉默的目的，使沉默令人感到自在且充满关怀（而非尴尬） [8]。这种将沉默视为沟通方式的理解，与人机交互文献形成对比；后者大多将沉默视为需要减少的延迟，或是用于管理用户参与度的时间参数 [5, 54, 74, 117]。牧师还指出，基于文本的聊天机器人难以传达积极专注的倾听。牧师描述了他们如何使用非语言线索来表示倾听（这一方面已在具身互动中得到探讨，例如相互凝视 [76]）；而基于文本的聊天机器人设计通常缺乏足够的手段（如视觉线索）来传达倾听，转而通过聊天机器人话语的内容来表达倾听，例如释义用户话语 [109]。这凸显了基于文本的聊天机器人可以更好地向用户传达自己正在倾听的机会；设计时应同时考虑不作回应（为用户留出反思的空间）以及使用（视觉）线索来表示倾听。

其次，牧师强调，他们关注的是与人联结，而他们设计的人工智能聊天机器人缺乏做到这一点的能力。对牧师而言，联结通过情感亲近传达，并由语言、非言语及环境线索加以支持。参与者描述了肢体语言、面部表情和姿势的细微变化如何帮助营造关系温情，以及物质环境（桌上的物品、椅子的摆放和身体距离）如何影响安全感与开放程度。此外，牧师还将自己的着装与外表视为传递舒适感和包容性的有意线索，例如佩戴骄傲彩虹丝带[6]。与这些具身及环境层面的联结方式相比，基于文本的互动（例如 ChatGPT 标准即时通讯界面所提供的互动）几乎无法传递这类线索。同样，先前的文本系统通常试图通过语言线索来营造舒适感和联结感，例如聊天机器人进行自我披露 [65, 77],、记住个人及其过去说过的话 [23, 24],, 或调整对话风格 [25, 104] 以影响人际亲近感。除语言线索外，文本系统也探索过一些视觉线索，例如聊天机器人图标（如告解聊天机器人中的女祭司图标 [26]）、目光线索 [112], , 或通过探索文本气泡形状来传达情绪 [3, 6]。然而，这些线索可能十分脆弱；视觉线索与对话线索不匹配，可能会造成期望违背，反而降低人们自我披露的意愿 [18]。牧师认为，这些关系联结方面的局限（尤其是缺乏具身线索和情感临在）在涉及灵性、悲伤或个人丧失的情境中尤为突出。因此，牧师强调，情感与关系层面的要素对牧灵实践至关重要，而这与人格（即作为一个人的状态）有关——这一议题直到近期才开始在聊天机器人研究中得到探讨 [50]。

第三，牧师报告称，人工智能聊天机器人未能提供一种情感亲近，这种亲近源于承担责任并承载 求助者的忧虑与顾虑。牧师强调，人们是在自我披露的 过程中获得洞见，而非通过外部指导。然而，人工智能聊天机器人虽然越来越多地被设计用于支持人们通过日记记录 [55, 56] 或讲故事 [81], 等方式卸下 心中负担，却缺乏承载并接纳人们托付给它们之事的能力。牧师在此指出，聊天机器人或许能够提供高质量的建议，却不具备据以作出有意义回应的隐性个人经验；尽管聊天机器人可能理解 苦难的概念，却无法亲身受苦。这一担忧与人机交互领域更广泛的批评相呼应：相关批评指出，当技术将个体抽象为工程问题 [2, 94], 时，人们可能会感到自己被简化或非人化，或者

[6]虽然本文讨论聚焦于基于文本的聊天机器人，但我们强调，在完全虚拟环境中与具身对话代理（ECA）互动时，可以操控动作、空间距离以及 ECA 外观线索等视觉线索。有关对话代理文献中所用线索的综述，请参阅 Fiene 等人关于对话式智能体社会线索分类体系的研究 [33]。

当设计方案将用户视为可以互换的个体，而非承认他们各自的人格 [17]。

第四，也是最后一点，牧师将他们创建的人工智能聊天机器人描述为倾向于用信息使用户不堪重负，并在此过程中，向个人索取过多。牧师将其与牧灵关怀进行了对比：在牧灵关怀中，对话不会由某种议程驱动，也不会被用来获取越来越详细的披露。具体而言，牧师强调留出空间，这既受到他们自身创伤与丧亲经历的塑造 [58]，，也源于仅仅陪伴某人本身所带来的慰藉性陪伴 [100]。从这一视角来看，聊天机器人的急切（表现为快速回应、不断追问的追问，或过于详细且冗长的回复）与牧师通常寻求营造的反思性、低要求的相遇并不相称。这反映了人们对人工智能系统在未提供有意义的关系互惠的情况下提取数据的更广泛担忧 [67]，，也呼应了关于基于GPT的系统可能生成过度冗长输出的研究发现 [45]。尽管可以认为向个人索取过多这一点很容易通过提示工程等方式进行配置，但与大语言模型互动时，快速回应十分常见——牧师认为这一倾向并不理想。尽管已有研究探讨了人类与聊天机器人消息的重叠 [52]、聊天机器人的打字行为 [117]，，或聊天机器人的积极倾听 [109]，，但这些只是可能影响用户如何看待人工智能聊天机器人向个人索取过多的部分因素。

总而言之，近期研究通过不同方法，部分探讨了我们研究中发现的主题，包括运用美德伦理学指导技术设计 [22]，以及为在线社区提供牧灵关怀 [11]。然而，我们的研究结果立足于牧师的视角，对人工智能聊天机器人在通常需要更高敏感度的互动中所欠缺之处提出了新见解。接下来，我们将讨论这些局限性，以及应对这些问题的潜在后续措施。

5.2 聊天机器人调谐的考量

如前所述，牧师拥有专业知识，在与求助者交谈时也处于独特的位置。尽管具体主题各不相同，但都属于更广泛的调谐范畴。Vallgårda 将调谐理解为与关怀技术建立共鸣，并指出：

> “（……）关怀技术要设计得当，远比大多数其他技术复杂，因为它必须建立亲密而微妙的人际关系才能发挥作用。人们拥有不同的经历、梦想、愿望和能力；要让关怀取得成效，他们必须能够与自己的关怀技术相互调谐。（……）” *[101,第 281]*页

尽管大量人机交互研究已凸显人工智能聊天机器人作为心理健康工具的潜力（例如，作为打字疗法 [95] 或积极倾听者 [102]），而且越来越多人转向它们寻求拟社会支持，但近期研究也清楚表明，这些系统无法取代真正的人与人之间的联系 [32, 116]。更具体地说，人机交互研究越来越质疑人工智能聊天机器人模拟共情 [88]与亲密性 [21] 的倾向。因此，显然有必要对人工智能聊天机器人的设计假设，以及它们如何尝试模拟人类能力提出质疑。与近期关于人工智能聊天机器人行为的研究（例如，人工智能聊天机器人对用户表现出对抗性 [16]）类似，探索人工智能聊天机器人如何以不同方式与人调谐，具有潜在价值。正如第5.1节中，当前方法例如关注积极倾听和鹦鹉学舌，而非探索如何将调谐有意融入人工智能聊天机器人的设计。例如，当前具备推理能力的人工智能聊天机器人版本（如 ChatGPT）会主动展示其处理过程，显示自己正在思考，并在给出答案后显示思考了 X 秒。相比之下，如何创造情境，让人工智能聊天机器人能够与用户调谐（倾听却不明确表现出来、不在场却与某人建立联结、在不具备能动性的情况下帮助他人承担重担，以及避免向他人有所渴求），仍是人机交互领域尚待解决的挑战。

尽管针对人工智能聊天机器人调适进行设计是一个值得进一步探索的方向，但仍需牢记若干方面。近期研究已勾勒出人工智能聊天机器人表达自身的方式，以及这些表达如何促使人们将其拟人化 [29]。然而，这项研究完全聚焦于语言表达，并未考虑聊天机器人可能借以实现调适的其他潜在模态。因此，如何在人与AI聊天机器人交互中更精确地概念化并设计有益于福祉的聊天机器人调适，仍是未来研究有待探索的方向。

5.3 局限性与未来工作

我们承认本研究存在若干局限。首先，尽管本研究中的牧师表示对生成式人工智能的接受度中等至较高（见表 1），且有些人提到曾试用 ChatGPT，但他们普遍表示自己使用 ChatGPT 的经验有限。没有人表示曾有构建聊天机器人的经验或专长。尽管如此，研究越来越多地通过 GPT Builder 等新颖的人工智能工具让利益相关者参与其中——对于关注“人工智能素养较低的各类用户的人工智能相关需求”的研究而言，这被认为是一种尤其适用的方法 [60, 第 6]页。再加上环节时间有限（约六十分钟），且预计每个聊天机器人仅用五分钟，这些因素很可能影响了牧师构建人工智能聊天机器人的方式。其次，如前所述，我们关注的是牧师对用于对话式关怀的人工智能的看法，而非提示构建。因此，我们没有分析参与者的提示指令或测试查询。先前研究表明，非技术专家在与生成式人工智能系统互动时往往会遇到困难 [97, 113]。再加上环节时间有限（总计约六十分钟，每个聊天机器人预计花费五分钟），这很可能影响了牧师构建和使用人工智能聊天机器人的方式。因此，本研究不对牧师的提示能力作出任何论断。不过，我们将这些材料纳入补充材料，因为我们认为，它们反映了牧师如何设想人工智能聊天机器人情境下求助者及其需求。第三，我们将招募范围限定在北欧，这影响了研究结果的普遍适用性。纳入其他地区的牧师，可能会带来受当地关怀实践影响的更多观点。最后，我们承认，本研究存在一项局限：我们研究了专家（本研究中为牧师）的意见，却没有考虑终端用户的观点。这并非意在贬低终端用户观点的价值；在福祉聊天机器人的研究中，终端用户观点通常已得到充分研究。其中包括患者等终端用户的观点 [66], 以及年轻人的观点 [111],

青少年 [62], 黑人 [38, 53], 自闭症人士 [43], 以及 LGBTQ+ 群体 [72]。尽管这些不同的观点对于对话式关怀的未来发展至关重要，但它们可能与牧师等领域专家所提出的观点有所不同。为了结合专家见解与终端用户偏好，未来研究可以根据牧师的观点开发交互式原型，并与预期目标受众讨论或评估这些原型。

本研究的成果，如第5.1 节所讨论并在第5.2节中所阐释的，有助于为人工智能聊天机器人的未来设计提供参考。如第1节所述，在设计更能调适终端用户的更有效技术时，专家视角至关重要。因此，未来关于福祉互动的人工智能聊天机器人研究应继续纳入终端用户，同时也应纳入牧师及其他在人工智能聊天机器人设计讨论中常被忽视的目标受众。

6 结论

让具备关键专长的利益相关者参与，对于设计更好的福祉人工智能聊天机器人至关重要。在本文中，我们请牧师构建人工智能聊天机器人，并随后收集了他们的观点。我们的结果说明了牧师如何看待自身的牧灵关怀，以及他们所构建的人工智能聊天机器人在多大程度上达到了这一标准。牧师们表示，他们的牧灵关怀以人的临在为中心，而他们认为人工智能聊天机器人无法完全复制这一点。这包括传达倾听、与个人联结、承载他们的关切，以及不渴望 从个人那里索取过多。我们讨论了牧师的观点如何与人机交互研究相关联，例如，对机器人响应时间的操控如何与牧师传达倾听的方式形成对比。此外，我们提出了一条推进聊天机器人调适的路径，并概述了围绕此类努力的关键考量。我们的发现对福祉情境中的人工智能聊天机器人具有启示意义，并基于牧师的观点及其所实践的牧灵关怀，为人工智能聊天机器人设计提供了新的视角。

致谢

本研究由嘉士伯基金会资助，资助编号为CF21- 0159。

参考文献

[1] Kevin Adams. 2019年。院牧在场的界定与操作化：综述。《宗教与健康杂志》 58, 4（2019年8月），1246–1258。 doi:10.1007/s10943- 018-00746-x[2] Hunter Akridge、Bonnie Fan、Alice Xiaodi Tang、Chinar Mehta、Nikolas Marte- laro与Sarah E Fox。2024年。“没有我们，公交车便一无是处”：在持续推进交通自动化的过程中彰显公交司机的劳动。载于*2024*年*CHI*计算机系统中的人为因素大会论文集 （美国夏威夷州檀香山）*(CHI '24)*。文章编号479，16页。 doi:10.1145/3613904.3642714[3] Pengcheng An、Jiawen Stefanie Zhu、Zibo Zhang、Yifei Yin、Qingyuan Ma、Che Yan、Linghao Du与Jian Zhao。2024年。EmoWear：探索智能手表上语音消息交互的情感引导语。载于 *2024*年*CHI*计算机系统中的人为因素大会论文集 （美国夏威夷州檀香山） *(CHI '24*279, 16 1011453613904)。美国计算机协会，美国纽约州纽约市，文章页。 doi:./.3642281[4] 2022 Helen Cowie和. . 当今教育中的牧灵关怀：其在促进儿童和青少年心理健康方面的持续作用。 教育中的牧灵关怀 40, 3 2022 321 327 10 1080 02643944 2022 2093955 (), –. doi:./..Naeimeh [5] Anzabi与Hiroyuki Umemuro。. 社交机器人不同倾听行为对人-机器人交互中感知信任的影响。国际社交机器人学杂志 15, 6（2023年6月），931–951。 doi:10.1007/s12369-023-01008-x 2023

[6] Toshiki Aoki、Rintaro Chujo、Katsufumi Matsui、Saemi Choi和Ari Hautasaari。2022年。EmoBalloon——以对话气泡在文本聊天中传达情绪唤醒。载于《2022年CHI计算机系统中的人为因素大会论文集》（美国路易斯安那州新奥尔良）（CHI '22）。美国计算机协会，美国纽约州纽约市，文章编号 527, 16 10 1145 3491102 3501920 页。doi: . / .[7] Kristin Aune、Lucy Peacock、Mathew Guest和Jeremy Law。.作为关系性临在的大学牧灵服务：探寻英国大学对良好且有效牧灵服务的理解。《大学与品格》24，3（ 2023 197 216 10 1080 2194587 2023 2224573 ），–。doi: . / X. . Anthony L Back、[8] Susan M Bauer-Wu、Cynda H Rushton和Joan Halifax。2009年。患者—临床医生互动中的慈悲性沉默：一种观照式方法。《姑息医学杂志》12，12（2009年），1113–1117。https://doi. org/ 10 1089 2009 0175 . /jpm. . Lynn Bassett、Amanda F Bingley和[9] Sarah G Brearley。2018年。作为照护要素的沉默：对专业照护者在临床与牧灵场景中经验的元民族志综述。《姑息医学》32，1（2018年），185–194。doi:10.1177/0269216317722444[10] Andrew B.L. Berry、Catherine Y. Lim、Calvin A. Liang、Andrea L. Hartzler、Tad Hirsch、Dawn M. Ferguson、Zoë A. Bermet和James D. Ralston。2021年。支持对个人价值观与健康的协作式反思。《ACM人机交互学报》5，CSCW2，文章编号299（2021年10月），39页。doi:10.1145/ 3476040[11] Alemitu Bezabih、Shadi Nourriz、Anne-Marie Snider、Rosalie Rauenzahn、George Handzo和C. Estelle Smith。2025年。在患者所处之处与他们相遇：推动牧灵关怀扩展至在线灵性关怀社群。https://arxiv.org/abs/ 2506 11366 。[12] Petter Bae Brandtzaeg、Marita Skjuve和Asbjørn Følstad。2022年。我的AI朋友：社交聊天机器人用户如何理解其人机友谊。《人类传播研究》 48, 3 04 2022 404 429 10 1093 008 （），–。doi: . /[13] hcr/hqac Virginia Braun和Victoria Clarke。.在心理学中运用主题分析。《心理学定性研究》3，2（2006年），77–101。doi:10.1191/ 1478088706qp063oa [14] Virginia Braun和Victoria Clarke。2019年。反思反思性主题分析。《体育、锻炼与健康定性研究》11，4（2019年），589–597。doi:10.1080/2159676X.2019.1628806[15] Anna Bravesmith。2012年。沉默为言语赋予完整性：在分析性对话中超越言语与沉默的对立。《英国心理治疗杂志》28，1（2012年），21–34。https://doi.org/10.1111/j.1752-0118.2011. 01263.x[16] Alice Cai、Ian Arawjo和Elena L. Glassman。2024年。对抗性人工智能。https: // arxiv.org/abs/2402.07350[17] Yoonha Cha、Victoria Jackson、Karina Kohl、Rafael Prikladnicki、André van der Hoek和Stacy Branham。2025年。为职场无障碍构建自助（DIY）解决方案的困境。载于《2025年CHI计算机系统中的人为因素大会论文集》（CHI '25）。文章编号47，17页。doi:10.1145/3706598.3713302[18] Jiahao Chen、Mingming Li和Jaap Ham。2024年。拟人化设计线索的不同维度：视觉外观与对话风格如何影响用户向聊天机器人进行信息披露的倾向。《国际人机交互研究杂志》190（2024年），103320。doi:10.1016/j.ijhcs.2024. 103320[19] Hyungjun Cho、Jiyeon Amy Seo、Ji won Lee、Chang-Min Kim和Tek-Jin Nam。2025年。ShamAIn：受萨满教启发设计卓越的对话式人工智能。载于《2025年CHI计算机系统中的人为因素大会论文集》（CHI ' *25* 985, 18 10 1145 3706598 3714297 ）。文章编号，页。doi: . / .[20] Philip J. Choi、Vinca Chow、Farr A. Curlin和Christopher E. Cox and。2019年。重症监护临床医生对牧师角色的看法。《医疗保健牧灵杂志》25，3（2019年），89–98。doi:10.1080/08854726.2018.1538438[21] Minh Duc Chu、Patrick Gerard、Kshitij Pawar、Charles Bickham和Kristina Lerman。2025年。亲密错觉：人机关系中的情感依恋与新兴心理风险。https://arxiv.org/abs/2505. 11649[22] Louisa Conwill、Megan K. Levis、Karla Badillo-Urquiola和Walter J. Scheirer。2025年。面向共同福祉的设计模式：运用美德伦理学的智慧构建更好的技术。载于《2025年CHI计算机系统中的人为因素大会论文集》（CHI '25）。文章编号831，23页。doi:10. 1145/3706598.3713546[23] Samuel Rhys Cox、Rune Møberg Jacobsen和Niels van Berkel。2025年。聊天机器人的短暂性框架对自我披露感知的影响。载于《第7届ACM对话式用户界面大会论文集》（CUI ' *25* 60, 17 10 1145 3719160 3736617 ）。文章编号，页。doi: . / .[24] Samuel Rhys Cox、Yi-Chieh Lee和Wei Tsang Ooi。2023年。比较聊天机器人如何引用先前聊天会话中的用户话语：对用户隐私顾虑与感知的调查。载于《第11届国际人机交互会议论文集》（瑞典哥德堡）（HAI '23）。105–114。doi:10.1145/3623809.3623875

[25] Samuel Rhys Cox和Wei Tsang Ooi。2022年。聊天机器人语言正式程度是否影响用户的自我披露？ 载于《第4届对话式用户界面大会论文集》（英国格拉斯哥）（CUI '22）。文章编号1，13页。doi:10.1145/3543829.3543831[26] Emmelyn AJ Croes、Marjolijn L Antheunis、Chris van der Lee和Jan MS de Wit。2024年。数字忏悔：向聊天机器人披露亲密信息的意愿及其对情绪福祉的影响。《与计算机交互》（ 2024 279 292 10 1093 016），–。doi: . /iwc/iwae Annelieke Damen、[27] Carmen Schuhmann、X.J.S. Rosie、Marjo van Zundert、Gaby Jacobs、Hanneke Muthert、Erik Olsman和Anja Visser。2025年。牧灵服务对初级与社区照护的贡献：一项针对来访者的半结构式访谈研究。《初级照护与社区健康杂志》16（2025年），21501319251357528。doi:10.1177/21501319251357528[28] Sidney Dekker和Lance Mergard。2025年。成为危机牧师：在危急情境中提供帮助与疗愈。劳特利奇。doi: 10 4324 9781003588412 . / Alicia [29] DeVrio、Myra Cheng、Lisa Egede、Alexandra Olteanu和Su Lin Blodgett。2025年。促成语言技术拟人化的语言表达分类体系。载于《2025年CHI计算机系统中的人为因素大会论文集》（CHI '25）。文章编号430，18页。doi:10.1145/3706598.3714038[30] J M Ellis、C McManus和B A Newton。1995年。患者如何感知医院牧师的角色：一项初步探索。《BMJ质量与安全》4，3（1995年），174 177 10 1136 4 3 174–。doi: . / qshc. . . Sarah Elwahsh、Nora Stern、[31] Aneesha Singh和Amid Ayobi。2025年。语言多样性与心理福祉：与多语母亲共同设计定制人工智能聊天机器人。载于《第7届ACM对话式用户界面大会论文集》（CUI ' 25 65, 17 10 1145 3719160 3736615）。文章编号，[32] 页。doi: . / . Cathy Mengying Fang、Auren R. Liu、Valdemar Danry、Eunhae Lee、Samantha W. T. Chan、Pat Pataranutaporn、Pattie Maes、Jason Phang、Michael Lampe、Lama Ahmad和Sandhini Agarwal。2025年。人工智能与人类行为如何塑造聊天机器人使用的社会心理影响：一项纵向随机对照研究。https://arxiv.org/abs/ 2503 17473 。Jasper Feine、Ulrich Gnewuch、Stefan [33] Morana和Alexander Maedche。2019年。面向对话代理的社会线索分类体系。《国际人机交互研究杂志》 132 2019 138 161 10 1016 2019 07 009 （），–。[34] doi: . /j.ijhcs. . . George Fitchett、Kathryn A. Lyndes、Wendy Cadge、Nancy Berlinger、Erin Flanagan和Jennifer Misasi。2011年。专业牧师在儿科姑息照护团队中的角色：来自医生与牧师的视角。《姑息医学杂志》 14, 6 2011 704 707 10 1089 2010 0523 （），–。doi: . /jpm. . [35] Yue Fu、Sami Foell、Xuhai Xu和Alexis Hiniker。2024.从文本到自我：用户对AIMC工具在人际沟通与自我方面的感知。载于《CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI ' 24 977, 17 10 1145 3613904 3641955 ）。文章编号，[36] 2024 页。doi: . / . Mark Graves和Jane Compson。.面向道德决策、健康与福祉的富有同情心的人工智能。《AAAI/ACM人工智能、伦理与社会大会论文集》 7, 1 2024 520 533 10 1609 7 1 31655 （10月），–。[37] 1988 doi: . /aies.v i . Jonathan Grudin。.计算机支持的协同工作应用为何失败：组织界面设计与评估中的问题。载于《1988年ACM计算机支持的协同工作会议论文集》（美国俄勒冈州波特兰）（CSCW ' 88 85 93 10 1145 62266 62273 ）。–。doi: . / . Christina N. Harrington和[38] Lisa Egede。2023年。信任、舒适与可亲近性：理解黑人老年人对用于健康信息寻求的聊天机器人设计的感知。载于《2023年CHI计算机系统中的人为因素大会论文集》（德国汉堡）（CHI '23）。文章编号120，18页。doi:10.1145/3544548.3580719[39] Susan Harris。2018年。牧师在关键临床决策中的调解者角色。《美国医学会伦理杂志》 20, 7 2018 670 674 10 1001 2018 670 （），E–。[40] doi: . /amajethics. . Michael A. Hedderich、Natalie N. Bazarova、Wenting Zou、Ryun Shim、Xinda Ma和Qian Yang。2024年。一出戏剧：探究教师如何设计大语言模型聊天机器人以辅助青少年网络欺凌教育。载于《2024年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '24）。文章编号668，17页。doi:10.1145/3613904.3642379[41] Guy Hoffman、Jodi Forlizzi、Shahar Ayal、Aaron Steinfeld、John Antanitis、Guy Hochman、Eric Hochendoner和Justin Finkenaur。2015年。机器人临场感与人类诚实：实验证据。载于《第十届ACM/IEEE国际人机交互会议年会论文集》（美国俄勒冈州波特兰）（HRI ' 15 181 188 10 1145 2696454 2696487 ）。–。doi: . / . Hui-Wen [42] Huang和Jessica (Chieh-Yu) Chang。.教师教育中的人机交互：审视社会临场感与友谊。载于《2024年人工智能与教师教育国际会议论文集》（ICAITE '24）。美国计算机协会，美国纽约州纽约市，64–69。doi:10.1145/3702386.3702399[43] JiWoong Jang、Sanika Moharana、Patrick Carrington和Andrew Begel。2024年。"这是我唯一能够信任的"：构想自闭症劳动者将大语言模型用于沟通辅助。载于《2024年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '24）。文章编号77，18页。doi:10.1145/3613904.3642894 [44] Zhuoqun Jiang、ShunYi Yeo、Wei Xuan Donovan Seow 和 Simon Perrault。2025年。《Remini：利用聊天机器人介导的相互回忆，增进亲友间的积极情感与联结感》。https://arxiv. org/abs/2508.03355[45] Matthew Jörke、Shardul Sapkota、Lyndsea Warkenthien、Niklas Vainio、Paul Schmiedmayer、Emma Brunskill 和 James A Landay。2025年。《GPTCoach：迈向基于大语言模型的身体活动指导》。载于《2025年CHI计算机系统中的人为因素大会论文集》。 1 46 10 1145 3706598 3713819 –。doi: 。/ 。[46] Ilkka Kaate、Joni Salminen、Soon-Gyo Jung、Trang Thi Thu Xuan、Essi Häyhänen、Jinan Y. Azem 和 Bernard J. Jansen。2025年。《"你总能得到答复"：分析用户在面对无法回答的问题及幻觉风险时与人工智能生成的人设的互动》。载于《第30届国际智能用户界面会议论文集》（IUI '25）。1624–1638。doi:10.1145/3708359。3712160 [47] Mahdi Kafaee、Aliakbar Kouchakzadeh 和 Shahriar Gharibzadeh。2024年。《沉默：人工智能中一个被忽视的概念》。《AI & SOCIETY》39卷第1期（2024年），415–416。doi:10.1007/s00146-022-01411-4[48] Eun Jeong Kang、Haesoo Kim、Hyunwoo Kim、Susan R. Fussell 和 Juho Kim。2025年。《粉丝能否通过偶像的模拟语音消息建立拟社会关系？——人工智能私人通话用户感知、认知与行为研究》。ACM《人机交互汇刊》9卷第2期，文章 CSCW044（2025年5月），31 10 1145 3711111 页。doi: 。/ Avleen Kaur 、 C. Estelle Smith 和 Loren Terveen[49]。2021 年。《共舞相伴，长久相守：通过 SoulGarden 原型可视化灵性支持网络》。载于《2021年计算机支持的协同工作与社会计算会议伴随出版物》（美国，线上活动）（CSCW '21 Companion）。84–88。doi:10.1145/3462204.3481774[50] Rucha Khot、Teis Arets、Joel Wester、Franziska Burger、Niels van Berkel、Rens Brankaert、Wijnand IJsselsteijn 和 Minha Lee。2025年。《挑战未来：运用聊天机器人反思衰老与痴呆》。载于《2025年CHI计算机系统中的人为因素大会论文集》（CHI '25 672, 14 10 1145 3706598 3713727）。文章页码。doi: 。/ 。[51] Inyeop Kim 、 Minsam Ko 、 Joonyoung Park 、 Sung Wook Moon 、 Gyuwon Jung、Youn-kyung Lim 和 Uichin Lee。2022年。《社会—灵性面：为灵性福祉设计社交阅读支持》。ACM《人机交互汇刊》6卷，CSCW 2, 262 2022 22 10 1145 3555162 文章（2022年11月），页码。[52] 2025 doi: 。/ JiWoo Kim、Minsuk Chang 和 JinYeong Bak。《超越轮流发言：将基于文本的重叠引入人与大语言模型的交互》。https://arxiv. org/abs/ 2501 18103 。Junhan Kim、Jana Muhic、Lionel Peter Robert 和 Sun Young Park[53]。2022年。《与美国黑人慢性病患者共同设计聊天机器人：克服应对2019冠状病毒病的挑战》。载于《2022年CHI计算机系统中的人为因素大会论文集》（美国路易斯安那州新奥尔良）（CHI '22）。文章 439, 17 10 1145 3491102 3502116 页码。doi: 。/ 。Kaeun Kim、[54] Ghazal Shams 和 Kawon (Kathy) Kim。2025年。《从秒数到情感：聊天机器人响应延迟对客户评价的差异化影响》。《国际人机交互杂志》0卷第0期（2025年），1–17。doi:10.1080/10447318.2025.2508915[55] Taewan Kim、Seolyeong Bae、Hyun Ah Kim、Su-Woo Lee、Hwajung Hong、Chanmo Yang 和 Young-Ho Kim。2024年。《MindfulDiary：利用大语言模型支持精神科患者进行日记记录》。载于《CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI ' 24 701, 20 10 1145 3613904 3642937 ）。文章页码。doi: 。/ 。[56] Taewan Kim、Donghoon Shin、Young-Ho Kim 和 Hwajung Hong。2024年。《DiaryMate：理解用户对个人日记记录中人机协作的感知与体验》。载于《CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '24）。文章1046，15页。doi:10.1145/3613904.3642693[57] Robert Klitzman、Elizaveta Garbuzova、Gabrielle Di Sapia Natarelli、Stephanie Sinnappan 和 Jay Al-Hashimi。2022年。《医院牧师与患者的沟通：特点、功能与潜在益处》。《患者教育与咨询》 105, 9 2022 2905 2912 10 1016 2022 05 004 ( ), –。doi: 。/j.pec.。。[58] Robert Klitzman、Stephanie Sinnappan、Elizaveta Garbuzova、Jay Al-Hashimi 和 Gabrielle Di Sapia Natarelli。2024年。《成为牧师：牧师如何以及为何入行、涉及的因素及其影响》。《医疗保健牧灵杂志》30卷第2期（2024年），75–88。[59] Sharni Konrad、Buddhi Gamage、Damith Herath 和 Janie Busby Grant。2025年。《机器人是社会性存在吗？探索人机交互中的具身性与社会临场感》。载于《2025年ACM/IEEE国际人机交互会议论文集》（澳大利亚墨尔本）（HRI '25）。1423–1427。[60] Huisung Kwon、Yunjae Josephine Choi、Sunok Lee 和 Sangsu Lee。2024年。《揭示内在需求：将 GPT Builder 作为参与式设计工具，与中年用户共同探索其对人工智能的需求与期望》。载于《CHI计算机系统中的人为因素大会扩展摘要集》（美国夏威夷州檀香山）（CHI EA ' 24 358, 6 10 1145 3613905 3650794 ）。文章页码。doi: 。/ 。Lene Kühle 和 Henrik Reintoft Christensen[61]。《为所有人服务：丹麦公共机构中牧灵服务的发展》。《社会指南针》66卷第2期（2019年），182–197。doi:10.1177/0037768619833310 2019

[62] Jamie Lee、Kyuha Jung、Erin Gregg Newman、Emilie Chow 和 Yunan Chen。2025年。《通过设计虚构理解青少年对健康人工智能技术的益处与风险的感知》。载于《2025年CHI计算机系统中的人为因素大会论文集》（CHI '25）。美国计算机协会，美国纽约州纽约市，文章311，20页。doi:10.1145/3706598。3713244[63] Ken Jen Lee、PiaoHong Wang 和 Zhicong Lu。2025年。《"真不敢相信我竟为一个动漫女孩哭了"：公众对虚拟主播毕业与终止的拟社会性哀悼及应对》。载于《2025年CHI计算机系统中的人为因素大会论文集》（CHI '25）。文章1237，23页。doi:10. 1145/3706598.3714216[64] Kwan Min Lee 和 Clifford Nass。2003年。《在人机交互中设计社会行动者的社会临场感》。载于《SIGCHI计算机系统中的人为因素大会论文集》（美国佛罗里达州劳德代尔堡）（CHI '03）。289–296。doi:10.1145/642611.642662[65] Yi-Chieh Lee、Naomi Yamashita、Yun Huang 和 Wai Fu。2020年。《"我听见你，我感受到你"：通过聊天机器人鼓励深入的自我披露》。载于《2020年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '20）。1–12。doi:10.1145/3313831.3376175[66] Brenna Li、Ofek Gross、Noah Crampton、Mamta Kapoor、Saba Tauseef、Mohit Jain、Khai N. Truong 和 Alex Mariakakis。2024年。《超越候诊室：患者对诊前聊天机器人对话细微差异的看法》。载于《2024年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '24，438, 24 10 1145 3613904 3641913）。文章页码。doi:。/。[67] Jie Li。2025。《当我们的孩子同时有一位人类恋人和一位人工智能恋人：与 Alexandra Diening 对话，探讨人际关系的未来》。《Interactions》32卷第5期（2025 18 20 10 1145 3757886 年8月），–。doi:。/ Yihe Liu、Anushk Mittal、[68] Diyi Yang 和 Amy Bruckman。2022年。《当我失去宠物时，人工智能会安慰我吗？理解人们对人工智能介导的电子邮件写作的感知》。载于《2022年CHI计算机系统中的人为因素大会论文集》（美国路易斯安那州新奥尔良）（CHI '22）。文章474，13页。doi:10.1145/3491102.3517731[69] Genevieve Liveley 和 Natalie J Swain。2024年。《"想象冒险的自由空间"：在人工智能与文学中为沉默发声》。载于《劳特利奇人工智能与文学手册》。劳特利奇，75 84 10 4324 9781003255789–。[70] https://doi.org/。/ Priscilla Y. Lo、Annemiek Veldhuis、Alissa N. Antle 和 Steve DiPaola。2025年。《Noel：支持儿童为他人进行设计的聊天机器人角色》。载于《2025年CHI计算机系统中的人为因素大会论文集》（CHI '25）。文章133, 25 10 1145 3706598 3713836 页码。doi:。/。Jiadi Luo、[71] Veronika Domova 和 Lawrence H Kim。2024年。《多机器人临场感与拟人化对人类认知和情绪的影响》。载于《2024年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '24，594, 15 10 1145 3613904 3642795）。文章页码。doi:。/。[72] Zilin Ma、Yiyang Mei、Yinru Long、Zhaoyuan Su 和 Krzysztof Z. Gajos。2024。《评估 LGBTQ+人群使用基于大语言模型的聊天机器人获得心理健康支持的体验》。载于《2024年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '24，872, 15 10 1145 3613904 3642482）。文章页码。doi:。/。[73] Takuya Maeda 和 Anabel Quan-Haase。2024。《当人机交互变成拟社会关系：情感化设计中的能动性与拟人化》。载于《2024年ACM公平性、问责制与透明度大会论文集》（巴西里约热内卢）（FAccT '24）。1068–1077。doi:10.1145/3630106.3658956[74] Mykola Maslych、Mohammadreza Katebi、Christopher Lee、Yahya Hmaiti、Amirpouya Ghasemaghaei、Christian Pumarada、Janneese Palmer、Esteban Segarra Martinez、Marco Emporio、Warren Snipes、Ryan P. McMahan 和 Joseph J. LaViola Jr.。2025年。《利用由大语言模型驱动的智能虚拟代理，缓解自由形式对话中的响应延迟》。载于《第七届ACM对话式用户界面大会论文集》（CUI '25）。文章49，15页。doi:10.1145/3719160.3736636[75] Kevin Massey、Marilyn JD Barnes、Dana Villines、Julie D Goldstein、Anna Lee Hisey Pierson、Cheryl Scherer、Betty Vander Laan 和 Wm Thomas Sum- merfelt。2015年。《我该怎么做？为重症监护病房姑息治疗中的灵性照护制定牧灵活动与干预措施分类体系》。《BMC姑息治疗》14, 1 2015 10 10 1186 12904 015 0008 0 (), 。doi:。/s - - -[76] Donald McMillan、Barry Brown、Ikkaku Kawaguchi、Razan Jaber、Jordi Sol- s ona Belenguer 和 Hideaki Kuzuoka。2019年。《以注视进行设计：Tama——一款由注视激活的智能扬声器》。ACM《人机交互汇刊》3卷，CSCW，文章176 2019 26 10 1145 3359278（2022年11月），页码。doi:。/。[77] Youngme Moon。2000。《亲密交流：利用计算机引发消费者自我披露》。《消费者研究杂志》26卷第4期（2000年），323–339。doi:10.1086/209566[78] Jared Moore、Declan Grabb、William Agnew、Kevin Klyman、Stevie Chancellor、Desmond C. Ong 和 Nick Haber。2025年。《表达污名并给出不当回应，使大语言模型无法安全取代心理健康服务提供者》。载于《2025年ACM公平性、问责制与透明度大会论文集》（FAccT '25 599 627 10 1145 3715275 3732039）。–。doi:。/。Dariya [79] Ovsyannikova、Victoria Oldemburgo de Mello 和 Michael Inzlicht。2025年。《第三方评估者认为人工智能比人类专家更富有同情心》。《通讯心理学》3, 1 2025 4 10 1038 44271 024 00182 6 (), 。doi:。/s - - - Hyanghee Park 和 Joonhwan Lee。2021。。为性侵幸存者设计对话式智能体：界定自我披露负担并构想以幸存者为中心的解决方案。载于《2021年CHI计算机系统中的人为因素大会论文集》（日本横滨）（CHI '21）。文章 634，共 17 页。doi:10.1145/3411764.3445133 [81] SoHyun Park、Anja Thieme、Jeongyun Han、Sungwoo Lee、Wonjong Rhee 和 Bongwon Suh。2021年。"我写作时就像在给一位熟人讲故事。"：为心理健康表达性写作设计聊天机器人交互。载于《2021年ACM交互系统设计会议论文集》（线上活动，美国）（DIS ' 21 926 941 10 1145 3461778 3462143）。–。doi:。/。[82] André Pereira、Rui Prada 和 Ana Paiva。2014。。改善人与智能体交互中的社会临场感。载于《SIGCHI计算机系统中的人为因素大会论文集》（加拿大安大略省多伦多）（CHI '14）。1449–1458。doi:10.1145/2556288.2557180[83] Mehrdad Rahsepar Meadi、Tomas Sillekens、Suzanne Metselaar、Anton van Balkom、Justin Bernstein 和 Neeltje Batelaan。2025年。探索心理健康护理中对话式人工智能的伦理挑战：范围综述。《JMIR心理健康》12 2025 10 2196 60432 ()。doi:。/ Matan Rubin、[84] Joanna Z. Li、Federico Zimmerman、Desmond C. Ong、Amit Goldenberg 和 Anat Perry。2025年。比较人们感知到的人类共情与人工智能生成的共情的价值。《自然·人类行为》（2025年）。doi:10.1038/s41562- 025-02247-w[85] Mana Samiee、Joel Wester、Rune Møberg Jacobsen、Michael Skovdal Rathleff 和 Niels van Berkel。2025年。全科医生对用于共享决策的诊前聊天机器人的看法。载于《2025年ACM交互系统设计会议论文集》（DIS '25）。3117–3131。doi:10.1145/3715336.3735674[86] Michael Schultz、Svetlana Baziliansky、Inbal Mitnik、Nirit Ulitzur、Salvatore Campisi-Pinto、Simon Givoli、Gil Bar-Sela 和 Daniela Zalman。2024年。社会工作、灵性照护与心理学之间的一些差异：临终对话中的内容差异。《姑息与支持性照护》22，2（2024年），306–313。doi:10.1017/S1478951523000652[87] Rachel L. Seddon、Edgar Jones 和 Neil Greenberg。2011年。牧师在维护军队人员心理健康方面的作用：历史与当代视角。《军事医学》176，12（2011年12月），1357–1361。doi:10.7205/MILMED-D-10-00124[88] Lennart Seitz。2024年。医疗聊天机器人中的人工共情：这种共情感觉真实吗？《人类行为中的计算机：人工人类》2，1（2024年），100067。doi:10 1016 2024 100067。/j.chbah.。Woosuk Seo、Chanmo Yang 和 [89] Young-Ho Kim。2024年。ChaCha：利用大语言模型引导儿童分享他们对个人事件的情绪。载于《2024年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '24）。文章 903，共 20 页。doi:10.1145/3613904.3642152[90] Marita Skjuve、Asbjørn Følstad、Knut Inge Fostervold 和 Petter Bae Brandtzaeg。2021年。我的聊天机器人伴侣——一项关于人与聊天机器人关系的研究。《国际人机交互研究杂志》149（2021年），102601。doi:10. 1016 2021 102601 /j.ijhcs.。C. Estelle Smith。。神圣的科技：关于将灵性融入健康与福祉技术的思考（与祈祷）。《Interactions》29，4（2022 68 72 10 1145 3543893 年6月），–。doi:。/ C. Estelle Smith、Avleen [92] Kaur、Katie Z. Gach、Loren Terveen、Mary Jo Kreitzer 和 Susan O'Conner-Von。2021年。什么是精神支持？它可能如何影响在线社区的设计？《ACM人机交互汇刊》5, 1, 43 2021 42 10 1145 3449117 计算机支持的协同工作文章（4月），页数。doi:。/ C. Estelle Smith、Zachary Levonian、Haiwei Ma、Robert Giaquinto、Gemma Lein-Mcdonough、Zixuan Li、Susan O'conner-Von 和 Svetlana Yarosh。2020年。"我无法独自承担这一切"：探索在线健康社区中的实际支持与祷告支持。《ACM人机交互汇刊》27，5，文章38 2020 41 10 1145 3402855（8月），页数。doi:。/ Aisha Sobey。。[94] 在人机交互中重新构想肥胖：倡导肥胖解放。载于《2024年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '24，506, 14 10 1145 3613904 3642199）。文章页数。doi:。/。[95] Inhwa Song、Sachin R. Pendse、Neha Kumar 和 Munmun De Choudhury。2025。。打字疗法：使用大语言模型聊天机器人寻求心理健康支持的体验。https://arxiv.org/abs/2401 14362。Xiaran Song、[96] Anqi Wang 和 Andrés Lucero。2025年。走进我的鞋子：科技灵性实践的自传式民族志。载于《2025年CHI计算机系统中的人为因素大会论文集》（CHI '25 987，14 10 1145 3706598 3713557）。文章页数。doi:。/。[97] Hari Subramonyam、Roy Pea、Christopher Pondoc、Maneesh Agrawala 和 Colleen Seifert。2024年。弥合构想鸿沟：与大语言模型进行基于提示的交互所面临的认知挑战。载于《2024年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '24）。文章 1039，共 19 页。doi:10.1145/3613904.3642754[98] Fiona Timmins、Sílvia Caldeira、Maryanne Murphy、Nicolas Pujol、Greg Sheaf、Elizabeth Weathers、Jacqueline Whelan 和 Bernadette Flanagan and。2018年。医疗牧师的作用：文献综述。《医疗保健牧灵

服务》 24, 3 2018 87 106 10 1080 08854726 2017 1338048 ( )，–。[99] Fiona Timmins 和 Nicolas Pujol。2018。医疗牧师在复苏中的作用：快速文献综述。《宗教与健康杂志》57，3（ 2018 1183 1195 10 1007 10943 018 0604 4 ），–。doi：。/s - - - Kirsten A Tornøe、Lars J Danbolt、Kari Kvigne 和 Venke Sørlie。2014年。[100] 慰藉性陪伴的力量——临终关怀护士对临终者灵性与存在关怀的亲身体验。《BMC Nursing》13，1（2014年），25。doi:10.1186/1472- 6955-13-25[101] Anna Vallgårda。2025年。调适关怀技术。《人机交互》 40, 5 2025 273 284 10 1080 07370024 2023 2300805 ( )，–。doi：。/。。[102] Netta Weinstein、Guy Itzchakov 和 Michael R. Maniaci。2025年。探索人工智能的联结潜力：将人际倾听与拟社会支持融入人机交互。《人类行为中的计算机：人工人类》 4 2025 100149 10 1016 2025 100149 ( )，。doi：。/[103] j.chbah.。Joel Wester、Sander de Jong、Henning Pohl 和 Niels van Berkel。2025。。利用大语言模型进行自我关怀：用户与咨询师视角。《国际人机交互研究杂志》（ 2025 10 1016 2025 103589 ）。doi：。/j.ijhcs.。J[104] oel Wester、Bhakti Moghe、Katie Winkle 和 Niels van Berkel。2024年。面对大语言模型：机器人沟通风格如何在人际健康信息传递中发挥中介作用。《ACM人机交互汇刊》8，计算机支持的协同工作第2期，文章 497 2024 37 10 1145 3687036 （11月），页数。doi：。/ Joel Wester、Henning Pohl、Simo Hosio 和 Niels van Berkel。2024年。[105] “这个聊天机器人绝不会.....”：人们感知到的心理健康聊天机器人的道德能动性。《ACM人机交互汇刊》8，计算机支持的协同工作第1期，文章 133（2024年4月），28页。doi:10.1145/3637410[106] Iris R. Wierstra、Annemarie Foppen、X. J. S. Rosie 和 Anke I. Liefbroer。2024年。牧师的职业身份与结构化牧灵服务方式：相互排斥还是相辅相成？《牧灵心理学》73，6（2024年），799–815。doi:10.1007/s11089-024-01164-5[107] Nipuni H. Wijesinghe、Yue Peng、Sharni Konrad、Maleen Jayasuriya、Janie Busby Grant 和 Damith Herath。2025年。重新构想人机交互中的社会临场感。载于《2025年ACM/IEEE人机交互国际会议论文集》（澳大利亚墨尔本）（HRI '25）。IEEE出版社，1716–1721。[108] Mari Lloyd Williams、Michael Wright、Mark Cobb 和 Chris Shiels。2004年。一项关于临终关怀机构牧师工作职责、责任与压力的前瞻性研究。《姑息医学》18，7（2004年），638–645。doi:10.1191/0269216304pm929oa [109] Ziang Xiao、Michelle X. Zhou、Wenxi Chen、Huahai Yang 和 Changyan Chi。2020年。如果我没理解错：打造并评估具备积极倾听技巧的访谈聊天机器人。载于《2020年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI '20）。1–14。doi:10.1145/3313831.3376131[110] Fatma Gizem Karaoglan Yilmaz、Ramazan Yil maz 和 Mehmet Ceylan and。2024年。生成式人工智能接受度量表：效度与信度研究。《国际人机交互杂志》40，24（2024年），8703–8715。doi:10.1080/10447318.2023.2288730[111] Jordyn Young、Laala M Jawara、Diep N Nguyen、Brian Daly、Jina Huh-Yoo 和 Afsaneh Razi。2024年。人工智能在青年同伴支持中的作用：对人工撰写和人工智能生成的回复的偏好研究。载于《2024年CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI ' *24* 1006, 18 10 1145 3613904 3642574 ）。文章页数。doi：。/。[112] Jingyi Yuan、Xixian Peng、Yichen Liu 和 Qiuzhen Wang。2025。。“别看着我！”：虚拟形象呈现风格与注视方向在社交聊天机器人设计中的作用。《人类行为中的计算机》164（2025年），108501。doi:10.1016/j.chb. 2024.108501[113] J.D. Zamfirescu-Pereira、Richmond Y. Wong、Bjoern Hartmann 和 Qian Yang。2023年。为什么 Johnny 不会写提示词：非人工智能专家如何尝试（并失败）设计大语言模型提示词。载于《2023年CHI计算机系统中的人为因素大会论文集》（德国汉堡）（CHI '23）。文章 437，共 21 页。doi:10.1145/3544548.3581388[114] Renwen Zhang、Han Li、Han Meng、Jinyuan Zhan、Hongyuan Gan 和 Yi-Chieh Lee。2025年。人工智能陪伴的阴暗面：人机关系中有害算法行为的分类体系。载于《2025年CHI计算机系统中的人为因素大会论文集》（CHI ' 25 13，17 10 1145 3706598）。文章页数。doi：。/。[115] 3713429 Yu Zhang、Jingwei Sun、Li Feng、Cen Yao、Mingming Fan、Liuxin Zhang、Qiany-ing Wang、Xin Geng 和 Yong Rui。2024年。广泛观察，审慎思考：旨在打破过滤气泡的生成式多智能体系统设计。载于《CHI计算机系统中的人为因素大会论文集》（美国夏威夷州檀香山）（CHI ' *24* 484, 24 10 1145 3613904 3642545 ）。文章页数。doi：。/。[116] Yutong Zhang、Dora Zhao、Jeffrey T. Hancock、Robert Kraut 和 Diyi Yang。2025年。人工智能伴侣的兴起：人与聊天机器人的关系如何影响福祉。https://arxiv.org/abs/ 2506 12605 。Jijie Zhou 和 Yuhan Hu。。言语之外：[117] 为对话代理注入类人打字行为。2024。载于《第6届ACM对话式用户界面大会论文集》（卢森堡，卢森堡）（CUI '24）。文章 24，共 12 页。doi:10.1145/3640794.3665560

7 附录

7.1 干预前访谈指南

- **(1a)**您与学生会面时，通常会依据什么来决定自己的风格或做法?
- **(1b)** 人工智能聊天机器人能否复制这种做法？为什么能或不能?
- **(2a)** 您通常与学生见面时首先会做什么?
- **(2b)** 您通常首先会说什么?
- **(2c)** 人工智能聊天机器人能否仿效其中任何一项？为什么能或为什么不能?
- **(3a)** 会议结束时，您通常会做的最后一件事是什么?
- **(3b)** 您通常最后会说什么?
- **(3c)** 人工智能聊天机器人能否仿效其中任何一项？为什么能或为什么不能?

7.2 干预后访谈指南 ide

- **(4a)** 你觉得为学生设计的通用人工智能聊天机器人怎么样?
- **(4b)** 你觉得为每位学生创建的聊天机器人分别会有多大效果?
- **(5a)** 为他人设计聊天机器人是什么体验?
- **(5b)** 哪些方面可以帮助你为他人设计聊天机器人?
- **(6a)**与聊天机器人互动和与真人互动有什么不同——这种差异重要吗?
- **(6b)** 你自身的价值观或信念如何影响了你在设计聊天机器人时所做的选择?

7.3 虚构人物简介

7.3.1 Maya。Maya 在离开她成长于其中的宗教传统后，正经历一次重大的身份转变。她前来并非为了传统意义上的精神指引，而是想探讨关于归属感、价值观和人生目标的问题。她重视开放、不作评判的空间，以便理清思绪。*7.3.2 Leo*。Leo 最近失去了一位亲密的家庭成员，一直难以跟上课程进度。他并不十分信教，但有人告诉他，牧师是可以倾诉的对象。他前来寻找一个安全的地方，以处理自己的悲伤，并弄清如何在应对情感负担的同时，在学校里维持下去。*7.3.3 Samira*。Samira 深度投入自己的信仰，并定期前来牧灵服务寻求支持。对她而言，牧师是一种令人踏实的临在，既理解她的精神信念，也能同时给予鼓励和实际帮助。