告知不知情者：借助人工智能驱动的聊天机器人提升在线知情同意的阅读体验

Ziang Xiao
ziang.xiao@jhu.edu
约翰·霍普金斯大学 美国
巴尔的摩 微软研究院 加拿大蒙特利尔

Tiffany Wenting Li
wenting7@illinois.edu 伊利诺伊大学厄巴纳-香槟分校 美国厄巴纳

Karrie Karahalios
kkarahal@illinois.edu 伊利诺伊大学厄巴纳-香槟分校 美国厄巴纳

Hari Sundaram
hs1@illinois.edu 伊利诺伊大学厄巴纳-香槟分校 美国厄巴纳

arXiv:2302.00832v1 [cs.HC] 2 Feb 2023

摘要

知情同意是人体受试者研究伦理的核心基石。通过知情同意流程，参与者可以了解研究流程、益处、风险等信息，从而做出知情决策。然而，近期研究表明，当前的做法可能导致参与者做出未经充分知情的决定，并使他们暴露于未知风险之中，尤其是在线研究。由于没有研究者在场并提供指导，在线参与者只能自行阅读冗长的表格，且无法获得问题的解答。在本文中，我们考察了人工智能驱动的聊天机器人在改善在线知情同意方面的作用。通过将聊天机器人与基于表单的交互进行比较，我们发现，聊天机器人改善了同意书阅读，增强了参与者的自主性感受，并缩小了参与者与研究者之间的权力差距。进一步的探索性分析表明，权力动态的变化最终可能有利于研究回答质量。我们讨论了设计人工智能驱动的聊天机器人的启示，以便在更广泛的场景中提供有效的知情同意。

CCS 概念

• **以人为中心的计算** → **人机交互**； • **计算方法** → **智能体**。

关键词

知情同意、对话式智能体、人工智能驱动的聊天机器人、人机交互、权力动态

ACM 引用格式：

Ziang Xiao、Tiffany Wenting Li、Karrie Karahalios 和 Hari Sundaram。2023。告知未获知情者：借助人工智能驱动的聊天机器人改善在线知情同意阅读体验。载于*2023*年 *CHI*计算系统中的人为因素会议论文集（*CHI '23*），*2023*年*4*月*23*日至*28*日，德国汉堡，德国。ACM，美国纽约州纽约市，17页。https://doi.org/10.1145/3544548.3581252

允许出于个人或课堂用途免费制作本作品部分或全部内容的电子版或纸质副本，前提是不得为营利或商业利益制作或分发副本，且副本须附有本声明及首页上的完整引文。必须尊重本作品中第三方组件的版权。其他用途请联系所有者/作者。*CHI '23*，*2023*年*4*月*23*日至*28*日，德国汉堡© 2023 版权所有，归所有者/作者所有。ACM ISBN 978-1-4503-9421-5/23/04。
https://doi.org/10.1145/3544548.3581252

1 引言

作为人体受试者研究伦理的核心基石， [26], 知情同意是保障潜在参与者自愿且知情地作出参与决策的流程。通过知情同意流程，参与者可以了解研究目的、将要执行的程序、参与可能带来的风险和收益、数据收集范围及保密性，以及自身权利。尽管知情同意十分重要，研究发现，人们往往没有仔细阅读表格就签了字 [28, 52, 58, 63, 70, 76]。本研究考察了人工智能驱动的聊天机器人在改善在线知情同意流程中的作用。

同意书阅读研究表明，从后果严重的高风险研究到风险极低的低风险研究，参与者都没有认真阅读知情同意书 [28, 52,58, 63, 70, 76]。知情同意流程转移到线上后，参与者获得的信息会更少 [33, 63, 65]。例如，Ghandour 等人 [34] 发现，在一份超过 800 词的在线同意书上，65% 的人阅读时间不到 30 秒，而读完这份同意书大约需要 7 分钟。Pedersen 等人 [63] 表明，与面对面的知情同意流程相比，当同一份知情同意书在线呈现时，参与者识别和回忆其中信息的能力进一步下降。因此，当前的知情同意流程可能既没有充分告知参与者，也没有获得他们的有效同意。

未经充分知情的参与决策会使参与者和研究者都面临风险。如果参与者没有注意到研究流程可能会加重其既有的心理或身体状况，那么后果可能危及参与者的健康 [21]。如果参与者对研究者的数据共享做法存在错误认知，也可能给参与者的个人数据带来隐私风险 [19]。此外，对知情同意书缺乏充分理解，可能会妨碍参与者自主作出同意决定。Cassileth 等人 [13] 发现，在一项医学研究中，超过四分之一的参与者认为，接受知情同意书是获得治疗的唯一选择，尽管该表格中说明了其他选择。与此同时，缺乏成效的知情同意

流程可能损害研究效度和数据质量 [23, 41]。未能理解研究程序和目的，不仅可能对参与者完成实验任务的表现产生负面影响，还可能引入混杂因素，尤其是在涉及欺骗的研究中 [23, 41]。相反，有效的知情同意流程可以提高参与者投入度，并促进参与者与研究者之间的信任和融洽关系 [23, 40, 67]。通过弥合研究中的信息差距、告知参与者其权利并保障其自愿决定，该流程可以减少参与者—研究者关系中的权力不对称 [69]。因此，成功的知情同意流程对参与者和研究者双方都有益。

近年来，尤其是在COVID-19疫情期间，在线知情同意日益普及。越来越多的线上和线下研究通过远程方式征求参与者的知情同意。这些研究可能涉及有风险的程序，或收集敏感信息，包括个人的心理健康信息和个人遗传数据。因此，提升在线知情同意的阅读体验变得愈发重要。

然而，改善知情同意书阅读是一项艰巨的任务，尤其是在在线环境中 [3, 29, 63]。与研究者能够直接与参与者互动、引导他们阅读知情同意书并回答问题的线下环境相比，在线环境中缺少研究者，造成了沟通障碍，并大大降低了参与者阅读知情同意书的积极性和意愿 [23, 65, 77]。过去，研究人员尝试过多种干预措施来改善知情同意书阅读，包括简化内容、降低阅读年级水平以及引入交互功能 [2, 64, 65]。然而，对70多项既往研究的两篇文献综述表明，这些干预措施的效果有限，而且设计和开发引人入胜且有效的交互体验也十分耗费精力 [20, 29]。因此，有必要探索改善在线同意书阅读的新方法。

我们考察了人工智能驱动的聊天机器人[1]在传达知情同意书内容方面所发挥的作用。我们构建了聊天机器人 Rumi，它采用将基于规则的系统与人工智能组件相结合的混合系统。Rumi 可以向参与者问候、逐节讲解知情同意书、回答参与者的问题，并收集参与者的同意答复，其做法类似于经验丰富的研究者在实验室环境中的工作。借助对话式界面，聊天机器人能够吸引人们的注意力、提供个性化体验，并实现类人交互 [83]。所有这些特点都有望使在线知情同意流程受益。然而，聊天机器人也存在若干风险。首先，逐轮聊天需要额外的时间与精力才能完成知情同意流程，这对阅读知情同意书构成了重大挑战 [29]。对于有付费参与者的研究而言，风险甚至更高，因为参与者耗时越久也不会因此获得更多报酬。其次，目前的聊天机器人远非完美。其有限的对话能力可能会给出错误答案，或令用户失望、沮丧 [39]。因此，聊天机器人会如何影响在线知情同意流程，目前仍不得而知。

[1]与传统的基于规则的聊天机器人不同，我们将人工智能驱动的聊天机器人定义为：在构建过程或部署环境中运用人工智能（AI）技术（包括机器学习、自然语言处理和高级分析）的聊天机器人。

为了探索人工智能驱动的聊天机器人引导参与者完成知情同意流程的有效性，我们提出了三个研究问题（RQs）：

研究问题1： 在采用人工智能驱动的聊天机器人知情同意流程的研究中，参与者阅读知情同意书的情况与采用基于表格的知情同意流程的研究相比有何不同？（知情同意书阅读）

研究问题2： 在采用人工智能驱动的聊天机器人知情同意流程的研究中，参与者与研究者之间的权力关系与采用基于表格的知情同意流程的研究相比有何不同？（权力关系）

研究问题3： 在采用人工智能驱动的聊天机器人知情同意流程的研究中，回答质量与采用基于表格的知情同意流程的研究相比有何不同？（研究回答质量）

为回答研究问题，我们设计并开展了一项被试间研究，在一项在线调查研究中比较了人工智能驱动的聊天机器人 Rumi 与典型的基于表格的知情同意流程。由于此前没有研究考察过用于提供知情同意的聊天机器人，本研究着重考察聊天机器人的整体效果，而非单个聊天机器人功能的效果。通过详细分析 238 名研究参与者的知情同意经历及其对调查问卷的回答，我们发现：1）Rumi 在回忆和理解方面改善了知情同意书阅读；2）使用 Rumi 完成知情同意书的参与者认为自己与研究者之间的权力关系更加平等；3）通过缩小权力差距，改善后的知情同意经历最终提升了研究回答质量。

据我们所知，我们的工作首次系统地比较了人工智能驱动的聊天机器人知情同意流程与典型在线知情同意流程的整体效果。我们的工作做出了三项独特贡献。

（1）深入了解人工智能驱动的聊天机器人开展在线知情同意的整体效果。我们的发现拓展了以往关于创建有效知情同意流程的研究，并揭示了使用人工智能驱动的聊天机器人进行知情同意的实际价值，尤其是在提升对知情同意表格的回忆和理解、改善研究者与参与者之间的关系，以及提高研究回答质量方面。我们还提供了实证证据，表明研究回答质量的提升可能归因于权力差距的缩小。（2）创建有效的人工智能驱动知情同意聊天机器人的设计启示。我们讨论了个性化阅读辅助和权力动态管理等设计考量，以进一步改善在线知情同意流程。（3）在更广泛的情境中创建并实现由人工智能支持的知情同意体验的新机遇。聊天机器人驱动的知情同意所展现出的有效性，为在人工智能时代将人工智能驱动的聊天机器人用于其他类型的知情同意开辟了新机遇。

2 相关工作

2.1 知情同意书阅读

知情同意这一概念蕴含于许多伦理准则的原则之中，包括《纽伦堡法典》、《赫尔辛基宣言》和《贝尔蒙特报告》 [26]。四项核心要素

构成知情同意流程，包括信息披露、理解、自愿性和能力。通过知情同意流程，参与者将了解研究的目的、程序、风险和获益，从而做出知情决策。

尽管知情同意书十分重要，人们往往仍未仔细阅读便签署，不论它涉及的是可能危及身心健康的临床试验，还是关于政治观点的研究。例如，Lavelle-Jones 等人 [52] 发现，准备接受各种外科手术的患者中，有 69% 未仔细阅读知情同意书便签署了。Varnhagen 等人 [76] 指出，在一项关于技术使用的研究中，研究参与者只能回忆起表格所含信息的不到 10%；35 % 的参与者表示，他们只是浏览了一下表格，甚至完全没有阅读。Cummings 等人将人们的研究参与决策与其对保密性、匿名性、数据安全以及采用开放数据共享做法的研究敏感性的顾虑之间存在的差异，归因于参与者没有认真阅读知情同意书 [19]。随着同意书阅读转移到线上，这一过程变得越来越具有挑战性 [33, 63, 65]。在线同意书成本低、易于管理，而且能够通过互联网触达更广泛的人群。然而，与实验室环境相比，没有研究者引导参与者阅读知情同意书、解释内容并解答参与者的疑问 [2, 20, 29]。因此，在线知情同意往往难以促成知情的参与决策，尤其是在研究内容更复杂、风险更高时。

知情同意阅读效果不佳，不仅会使参与者在不知情的情况下面临风险，还会损害研究的效度和数据质量 [23, 40, 67]。知情同意书包含有关研究程序和目的的重要信息，参与者能否理解这些信息，可能决定后续研究操纵的成败，尤其是在采用掩饰性故事的欺骗研究中。参与者未意识到的混杂因素可能进一步加剧社会科学领域的可重复性危机。

总而言之，知情同意流程对于研究中的参与者和研究人员都至关重要，但目前实施知情同意流程的方式存在诸多不足，尤其是越来越多的研究开始在线收集参与者的同意。在本研究中，我们旨在通过探索新颖的交互技术（例如人工智能驱动的聊天机器人），改善在线同意书阅读体验。

2.2 改善同意书阅读

尽管此前许多研究致力于改善知情同意书的阅读体验，但研究结果并不总是一致 [20, 29]。一组研究人员着重研究知情同意书的设计，包括文本可读性 [9, 37, 59], 、篇幅 [24, 64, 65]、版式 [18], 和媒体 [1, 31, 71]。Dresden 和 Levitt [24] 发现，减少知情同意书中不必要的信息并使用更简单的词汇，可以帮助研究参与者记住更多信息。然而，其他研究发现，简短的知情同意书未必能带来更高的理解得分 [9, 64]。尽管研究参与者倾向于选择篇幅较短、语言更简单的表格，但受法规和研究性质的限制，往往难以做到这一点 [33]。研究人员还探索了将文本形式的表格转换为多媒体形式。例如，Friedlander 等人发现，使用视频来传达知情同意书内容可以提高人们的参与度和理解程度 [31]。然而，多项荟萃分析表明，目前尚无确凿证据证明其有效性 [29, 62]。

另一组研究人员将交互性引入知情同意流程，旨在打造更具吸引力且个性化的体验，从而促进同意书阅读 [2, 4, 7]。在面对面场景中，由研究者带领参与者阅读同意书并回答其问题，被认为是最有效且最理想的方式 [2]。随着越来越多的研究转向线上，研究人员开始探索新的交互功能，以改善在线知情同意 [4, 7]。一种有效的干预方式是在参与者签署前测试其对知情同意书内容的了解程度 [49]。此外，Balestra 等人在在线知情同意书中使用社交批注，以促进参与者对知情同意书的理解 [4]。Bickmore 等人创建了一种具身智能体，可向读者解释医疗知情同意书；他们发现，与无法根据参与者已有知识调整解释内容的智能体相比，能够进行此类调整的智能体让人们更满意 [7]。尽管并非所有尝试都取得了成功，许多研究仍强调，与研究者互动对于确保参与者理解知情同意信息并建立信任至关重要，尤其是在复杂且风险较高的研究中 [14]。

我们采用先进的混合式聊天机器人框架构建了 Rumi，并集成了人工智能问答模块。与先前的研究相比， [7, 87] 以往的智能体主要依赖基于规则的系统，无法接收自然语言输入；而近期的语言技术使 Rumi 能够以自然语言回答各种问题，并凭借多种对话技能带来引人入胜的体验（详情请参见第 3.3 节）。这样的聊天机器人为在知情同意流程中重新融入拟人性，并提升知情同意流程的有效性，带来了新的契机。

2.3 人体受试者研究中的权力关系

在权力关系中，社会构成性权力较弱的一方通常会受到权力较强一方的制约。 [30]在人体受试者研究中，研究者往往被视为拥有绝对权威，能够决定资源分配；而参与者通常处于弱势地位。 [45]研究者对参与者的招募、治疗、数据和补偿拥有控制权，由此造成了研究者与参与者之间的权力不对称。 [12]这种权力差距会限制参与者的自主性，降低其参与研究的积极性，并使其不愿给出真实答案。 [12, 45]

许多研究人员主张通过权力再分配缩小人体受试者研究中的权力差距，以兼顾伦理和数据质量 [25, 79]。缩小权力差距后，参与者在研究过程中可能会更加投入，更自在地披露真实想法，也更愿意配合研究流程，最终或有助于提升研究质量。例如，Chen 分析了访谈员的语言使用情况，发现权力差距缩小有助于提高数据丰富度 [16]。然而，一些研究指出，为了做出专业判断，研究者与参与者之间保持距离十分重要 [56, 75]。

通过分享研究流程的信息、阐明风险和收益，并详细说明参与者的权利，知情同意流程旨在弥合信息差距

并确保参与者的自主性 [69]。此外，由于研究者与参与者之间最早发生的互动之一就是知情同意流程，因此该流程为重新分配权力和建立信任提供了绝佳机会。Kustatscher[51] 使用视觉辅助工具来改善儿童的知情同意，并发现这种引人入胜的流程改变了权力关系，也营造了更舒适的信息披露环境。

我们进一步拓展了关于研究环境中权力关系的先验知识，表明通过人工智能驱动的聊天机器人改进知情同意流程可以缩小权力差距。通过路径分析，我们发现，所观察到的研究质量影响可归因于权力关系的改变。

2.4 对话式人工智能作为研究工具

自然语言处理的最新进展使研究人员能够使用功能更强大的聊天机器人。一项新兴应用是将聊天机器人作为促进研究成功的工具，包括在线开展深入的对话式问卷调查 [22, 83], 或在现实世界中进行实地研究 [47, 72, 78]。与传统的基于表格的交互相比，聊天机器人在具备可扩展性优势的同时，还能提供更具吸引力和个性化的体验 [83]。具体而言，对话式界面通过轮流进行的聊天提供交互性，使聊天机器人能够以更个性化、更具对话性的方式组织问题，提供类似人类的社会互动以鼓励自我披露，并探究深入的信息 [81]。Tallyn等人 [72] 证明，在没有人类民族志研究者的情况下，使用聊天机器人收集民族志数据是有效的。Xiao等人表明，与基于表格的调查问卷相比，人工智能驱动的聊天机器人能够管理对话式问卷调查、收集更高质量的数据，并提供更具吸引力的问卷调查体验 [83]。聊天机器人还被用于其他研究场景，包括构建心理档案 [86], 、提供干预措施 [27], 以及指导研究程序 [50]。例如，心理学家通过对话使用聊天机器人推断参与者的性格，从而避免参与者在基于问卷的评估中弄虚作假 [82, 86]。李等人构建了聊天机器人，用于在心理健康研究背景下练习写日记 [53]。总之，既有研究表明，聊天机器人有潜力充当类似研究助理或访谈员的主持人，主动管理研究流程，以收集高质量数据并促进研究成功。

我们的研究考察了聊天机器人在人类受试者研究中另一个关键步骤——知情同意流程——中的实用性。我们展示了一个新的机会，并提出了设计启示，以简化研究中对话式人工智能的使用。我们提出了设计启示，以构建更优的未来社会科学人工智能。

3 方法

为回答我们的研究问题，我们设计了一项被试间研究，比较了在线提供知情同意书的两种方法所产生的结果：人工智能驱动的聊天机器人（聊天机器人条件）和典型表格（表单条件），考察其对同意书阅读、参与者—研究者权力关系以及研究回答质量的影响。

3.1 虚拟调查研究设计

为了了解人工智能驱动的聊天机器人如何促进在线知情同意书阅读，以及它可能如何影响研究回答质量，我们必须将研究质量评估与知情同意书阅读评估区分开来。因此，我们设计了一项专门用于评估研究回答质量的模拟研究。

该模拟研究围绕问题性社交媒体使用展开。参与者首先阅读一篇关于问题性社交媒体使用的短文，以熟悉这一议题。随后，参与者完成一份同时包含封闭式问题和开放式问题的调查问卷。基于选择的问题包括六道采用卑尔根社交媒体成瘾量表的五点李克特量表问题 [55]。开放式问题改编自 [6]，旨在了解人们对社交媒体的依恋及其对人们现实生活的影响。这两组问题在社交媒体研究中均得到广泛使用。我们选择这一主题有四个原因。第一，它与大多数在线人群相关，适合在线开展。第二，我们可以改变心理不适和数据敏感性的程度，以模拟更广泛的在线研究。第三，调查方法是使用最广泛的研究工具，从而确保研究发现的可推广性。最后，先前的研究为我们提供了成熟的方法，用于稳健地测量同时包含开放式问题和基于选择的问题的调查问卷的回答质量 [60, 83]。我们还考虑了一项由 [4] 使用的、涉及遗传数据的遗传学研究。尽管处理不当的遗传数据可能造成更为切实的后果，但风险程度难以改变，且可推广性有限。

我们为在线调查研究中三种常见的风险等级设计了三个版本的模拟研究，

- 低 - 不含个人标识符的非敏感数据
- 中 - 不含个人标识符的敏感数据
- 高 - 含个人标识符的敏感数据

在低 风险版本中，调查问卷将询问人们对他人问题性社交媒体使用的看法 。此版本旨在通过询问参与者对他人的看法，而非直接回忆自身经历，尽可能减少心理不适。在中 风险版本中，参与者将回答有关自身 问题性社交媒体使用的问题。而在高 风险版本中，我们还会要求参与者提供其社交媒体账号标识，作为后续研究中的个人标识符。知情同意书中的研究说明和潜在风险部分明确说明了这些区别。

3.2 知情同意书设计：表单条件

该知情同意书以伊利诺伊大学厄巴纳-香槟分校机构审查委员会（IRB）提供的社会行为研究同意书模板为基础。我们根据先前研究的建议 [18, 24, 64, 65]，改进了基于表格的知情同意书设计。具体而言，我们将知情同意书划分为多个部分，以减少信息过载，并采用清晰简洁的格式，确保清晰度和可读性。

3.3 知情同意书设计：聊天机器人条件

我们开发了一个名为 Rumi 的聊天机器人，旨在模拟线下知情同意流程的体验：研究者逐节讲解知情同意书，询问参与者是否有疑问，并作出说明。在我们的研究中，Rumi 首先向参与者致意，并告知参与者，在知情同意流程中的任何时候都可以

提出问题。随后，Rumi 按照表单条件中的原文逐节讲解知情同意书。参与者可以点击“下一步”按钮进入下一节，也可以在文本框中输入问题或其他请求。流程中，Rumi 主动询问参与者是否有问题，共两次：一次在风险部分之后，另一次在最后。参与者可以按下“没有问题”按钮跳过。之后，Rumi 确认参与者的年龄，并征求其加入研究的同意。我们在补充材料中提供了一段视频，演示 Rumi 如何进行知情同意流程。

我们采用混合方法，将基于规则的模型与人工智能模块相结合，构建了 Rumi [80]。我们做出这一决定有以下几个原因。首先，基于规则的模型在识别参与者的问题并提供多样且引人入胜的回答方面能力有限，更不用说处理非线性对话（例如闲聊、要求重复信息等）了。其次，尽管端到端生成系统可以带来更具吸引力的体验，但它可能会生成错误或不符合事实的回答。在高风险场景中，这些回答可能会造成严重问题 [8, 10], ，例如提供知情同意书。第三，知情同意书的内容通常取决于其使用的具体情境或背景。少样本或零样本设置给构建有效的生成模型带来了另一项挑战，即使是对预训练模型进行微调也不例外（例如 DialoGPT [85]）。因此，我们决定优先保证答案真实性，并基于混合系统构建 Rumi，借助人工智能模块使其能够处理广泛的问题，并提供多样且准确的回答。

具体而言，我们在 Juji平台上构建了 Rumi；Juji平台是一个混合型聊天机器人构建平台。Juji 提供基于 AI 的内置功能，用于对话管理和高效问答 [44]。Rumi 将按照基于规则的对话议程，逐节讲解知情同意书并征得参与者的同意。收到问题后，Juji 会使用预训练的自然语言理解（NLU）模型，在问答数据库中识别相关问题及其已知答案，并返回答案或提出后续问题以进行澄清。当聊天机器人不确定如何回答问题时，它会推荐类似问题，让参与者有机会获得所需答案，并进一步了解聊天机器人的功能 [44]。Juji 还提供多种对话技能，例如处理题外话和对话修复，从而带来更具吸引力的对话体验 [83]。

我们通过自行创建一组种子问题，并让 54 名在线参与者试用 Rumi，整理构建了问答数据库。由于我们的目标是了解参与者可能会提出哪些问题，因此我们请参与者尽可能多地提问。团队中的研究人员根据知情同意书为每个问题撰写答案，并将问答对添加到数据库中。在将 Rumi 用于本研究之前，问答数据库中已收录 200 多组问答对。为进一步提升 Rumi 识别不同表述方式的问题的能力，我们利用文本生成模型 GPT-3 (text-davinci-002) [61] 生成了问题改写。同样，我们也使用 GPT-3 为每个问题生成了一组候选答案。我们为每组问答对补充了五个问题改写和五个候选答案，从而扩充问答数据库。我们人工检查了所有生成文本，以确保信息真实可靠。

3.4 研究流程

完整研究分为两个部分（图 1），并已获得机构审查委员会批准。第一部分中，参与者被随机分配到其中一种条件（聊天机器人条件或表单条件）以及一种风险等级（低、中或高）。研究根据分配条件，从知情同意流程开始。参与者同意后，便完成虚拟调查研究。

我们的同意书评估安排在第二部分。参与者回答了有关其对知情同意书的理解、他们与研究者之间感知到的关系、知情同意流程体验以及人口统计信息的问题。最后，我们向参与者说明了研究的真实目的，并请他们填写另一份知情同意书，以同意分享他们对研究第二部分问题的回答。在离开研究前，我们还请参与者就其知情同意情况发表开放式评论。

如果参与者在第一部分决定不参加研究，我们会向他们说明研究的真实目的，并询问他们是否愿意参加第二部分，以评估他们阅读知情同意书的情况。如果他们同意参加，我们会直接引导他们进入第二部分，完成研究的剩余内容。

3.5 测量指标

3.5.1 知情同意书阅读。 我们从两个维度测量参与者阅读知情同意书的情况：回忆和理解。

回忆：能够回忆起知情同意书中的信息，表明人们注意阅读了知情同意书。与 [23], 类似，为了评估参与者对表格内容的回忆，我们在知情同意书中研究流程部分和风险部分的中间分别插入了两条随机陈述：“观看一段有橙色猫咪的视频”和“阅读蓝色背景下的研究材料”。我们测试参与者能否回忆起颜色词“橙色”和“蓝色”。我们决定用对两条随机陈述的回忆情况作为衡量同意书阅读的指标，因为这一方法能够灵敏地测量阅读的全面程度，而且很难靠猜测答对 [23]。研究表明，参与者更有可能仔细阅读研究流程和风险部分 [23, 76], ，这为我们测量同意书阅读的下限提供了机会。参与者需要从五种常见颜色词（“绿色”“红色”“白色”“橙色”和“蓝色”）中选出一个或两个。如果两个颜色词都选对，参与者得 2 分；如果只选一个选项且答案正确，得 1 分；如果选了两个选项但只答对一部分，得 0.5 分。其他情况下得 0 分。

理解： 参与者对知情同意书的理解反映了阅读的有效性。为了理解知情同意书，参与者需要处理文本、理解其含义，并将其与先验知识相结合 [36]。受[36], 启发，我们通过六个问题测量理解程度，这些问题要求参与者处理知情同意书中呈现的信息，而不仅仅是简单回忆。问题涉及研究流程、潜在风险，以及在特定情境发生时应采取的行动，例如，如果发生数据泄露，参与者如何保护自己的隐私，或者如果参与者决定退出研究，应该怎么做。

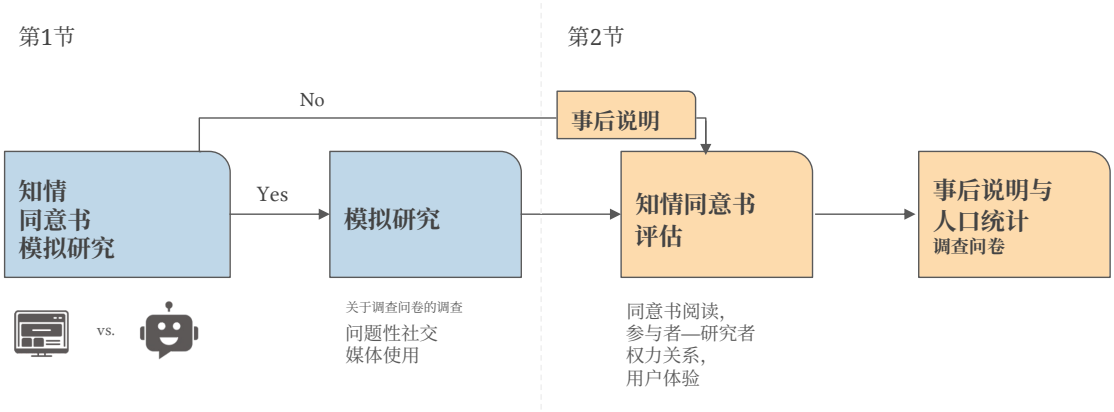


图1：该图展示了整体研究流程。在第1节中，参与者根据分配条件，在知情同意流程中与Rumi 或表格进行互动，随后完成一项关于社交媒体使用的模拟研究。在第2节中，参与者回答有关其知情同意书阅读体验的问题。最后，我们向参与者进行了事后说明，并收集了人口统计信息。如果参与者在第1节决定不参加研究，则会受邀开始第2节。

共设置了四道多项选择题。理解得分衡量参与者回答这些问题的准确程度。最终得分为答题正确率，范围为 0% 至 100%。

3.5.2 参与者—研究者权力关系。 我们通过参与者对研究中与研究者关系的感知，以及他们的自主性与控制感，来衡量参与者—研究者权力关系。

我们从两个方面衡量参与者对这种关系的感知：伙伴关系和信任。根据 [42], , 我们询问参与者是否认为自己与开展该研究的研究者是伙伴。我们借鉴 [4], , 通过询问参与者在多大程度上相信研究者会妥善处理他们的数据，来衡量信任。

根据 [12, 45],, 研究各方之间的权力不对称会抑制参与者的自主性。通过权力再分配，我们预期参与者能够重新获得能动性与控制感——这两者都是自主性的构成要素。我们从积极能动性和消极能动性的感知出发，测量了参与者感知到的能动性与控制感。我们根据知情同意流程的情境，对 [74] 中的量表进行了调整。所有条目均采用7点李克特量表，选项从“强烈不同意”到“强烈同意”。

3.5.3 研究回答质量。 我们通过考察参与者在问题性社交媒体使用调查问卷中对基于选择的问题和开放式问题的回答，测量了研究回答质量。

非差异化： 非差异化是一种调查敷衍作答行为，指受访者使用相同的回答量表，对所有题项给出几乎相同的回答，未能区分各题项 [48]。非差异化会降低题目回答的信度和效度，还会进一步夸大问卷内各题项之间的相互相关性，并抑制题项之间的差异 [84]。我们使用成对均值根方法来测量选择题中的非差异化 [60] 。我们计算问卷中所有题项对之间绝对差的平方根的均值。该指标范围为0（非差异化程度最低）至1（非差异化程度最高）。

响应质量指数（*RQI*）： 为测量开放式问题的回答质量，我们基于 [83]创建了一个响应质量指数。该指数从相关性、具体性和清晰度三个维度，衡量参与者给出的 N 个回答的整体回答质量，这些维度源自格赖斯准则 [38]:

$$\text{RQI} = \sum_{n=1}^{N} \text{relevance}[i] \times \text{clarity}[i] \times \text{specificity}[i] \quad \text{(N is the number of responses in a completed survey)} \tag{1}$$

相关性。 好的回答应与语境相关。对于开放式问题，高质量的回答应与调查问卷问题相关。无关回答不仅不会提供新信息，还会增加分析过程的负担。我们手动从三个等级评估每条开放文本回答的相关性：0 – 无关，1 – 有些相关，2 – 相关。

具体性。高质量的沟通通常内容丰富、细节充足。具体的回答提供充分的细节，有助于信息收集者更好地理解和利用这些回答，并从中获得更有价值、更深入的见解。我们对每条开放式文本回答的具体性进行人工评估，分为三个等级：0——仅有笼统描述，1——包含具体概念，2——包含具体概念及详细示例。

清晰度。清晰度是另一个重要维度。无论回答聚焦于何种主题，每条文本回答都应便于人们理解，且不存在歧义。我们对每条自由文本回答进行人工评分，分为三个等级：0——文本难以辨认，1——句子不完整，2——表达清晰。

编码过程： 我们共对 1428 份开放式回答进行了编码。两位具有人机交互背景且擅长开放式问卷分析的研究人员负责编码过程。他们首先随机选取了 10% 的回答

	Low	媒介	High	总计
聊天机器人条件	41	38	40	119
表单条件	39	40	40	119
总计	80	78	80	238

表1：该表展示了参与者在各条件下的分布。参与者根据知情同意书条件和风险等级随机分配至某一条件。最终分析共纳入238名参与者。

并据此针对上述三个维度编制了编码手册，其中包含定义和示例。随后，两位研究人员独立对其余数据进行编码，且对实验条件不知情。每组编码的克里彭多夫 α 系数介于 0.83 至 0.98 之间。最终的分歧通过讨论解决。

3.5.4 参与者体验与人口统计学特征。 时间与精力： 引入交互性通常意味着参与者需要花费额外的时间和精力与系统交互，这是一个重要的权衡 [20, 29]。因此，为了衡量参与者对知情同意流程所耗费时间和精力的感受，我们改编了ASQ量表中的两个条目 [54] ，询问参与者对完成知情同意流程所花费的时间和精力是否满意，之后将两项得分取平均，合并为一个分数。

未来使用： 人们未来使用同一系统的意愿，是衡量良好用户体验和满意度的重要指标。我们使用单项7点李克特量表，询问参与者未来是否会使用聊天机器人或表格来完成知情同意流程。

人口统计学特征：关于聊天机器人的既有研究表明，个体差异 [83] 可能会影响人们使用聊天机器人的体验。我们收集了基本人口统计信息，包括年龄、性别、教育水平和家庭年收入。

3.6 参与者概况

我们通过 Prolific [2]招募了来自美国、英语流利的参与者。在打开我们的链接并开始知情同意流程的278名参与者中，252人完成了知情同意表格。聊天机器人条件中的两名参与者明确拒绝加入研究，并立即退出。

在开始研究的250名参与者中，238人（记为P#）完成了研究，并通过了注意力检查和重复性检查。以下分析基于这238份有效答复（表1）。在这238名参与者中，97人认同自己为女性，136人认同自己为男性，5人认同自己为非二元性别或第三性别。参与者的教育水平中位数为学士学位。家庭年收入中位数介于50,000至100,000美元之间。参与者的年龄中位数介于25至34岁之间。平均而言，表单条件下的参与者完成知情同意流程耗时1.24分钟（SD = 3.03），而与Rumi互动耗时7.75分钟（SD = 7.06）。我们按每小时12美元的标准向参与者支付报酬。

3.7 数据分析

我们使用贝叶斯分析比较了两种同意方法在知情同意书阅读（研究问题1）、参与者与研究者之间的权力关系（RQ2）以及研究回答质量（研究问题3）方面的效应分布。我们选择使用贝叶斯分析，出于以下原因。首先，贝叶斯模型使我们能够明确呈现模型的所有方面；模型描述中已明确呈现的建模假设，无需再进行检验。其次，与零假设显著性检验（NHST）相比，贝叶斯分析关注的是“效应量有多大”，而不是“是否存在效应”。这更符合我们这项研究的探索性质。第三，贝叶斯模型有助于研究社群积累知识，因为研究结果日后可用作信息性先验。Kay 等人对贝叶斯方法在人机交互研究中的优势进行了详细综述 [46]。

我们针对每个结果测量指标构建了层级贝叶斯模型。我们构建了两种类型的层级贝叶斯模型：针对连续测量指标的线性回归模型，以及针对有序测量指标的有序逻辑回归模型。对于回忆、理解、能动性与控制感3、非差异化和 RQI，我们将数据建模为正态分布，并使用线性回归模型估计聊天机器人条件和表单条件下正态分布的均值。通过对比两个条件下均值的后验分布，我们可以了解同意方法如何影响结果变量。此外，对于贝叶斯线性回归模型，我们使用 Cohen's d 估计后验分布差异的效应量。

对于伙伴关系、信任、时间与精力以及未来使用，我们使用有序逻辑回归模型来估计有序量表上给定值的累积优势的后验分布。我们的贝叶斯分析目标是比较相对于量表中性中点（既不同意也不反对）而言，不同条件下的评分分布是否存在显著差异。这将告诉我们，在某一条件下，参与者是否比在另一条件下更可能不同意该陈述。我们构建了聊天机器人条件与表单条件之间评分为 4 分或以下（7点李克特量表的中点）的累积优势差异分布。负值表示，与另一条件相比，某一条件下的参与者提供中性或负面回应的优势较低。除回忆外，上述所有测量指标均采用了这一方法；对于回忆，我们估计的是回忆问题答对率为 50% 或更低的累积优势的后验分布。基于后验分布，我们计算了优势比（OR）作为效应量 [4]。

在所有模型中，我们控制了以下协变量：研究风险等级、参与者的年龄、性别、教育水平和家庭年收入。此前关于对话代理的研究表明，个体差异可能会影响参与者与对话代理的互动，因此我们控制了这些人口统计学特征 [83]。各类模型的完整数学描述见补充材料。我们使用广受欢迎的贝叶斯推断框架 NumPyro 5 进行贝叶斯分析。我们使用

[2]www.prolific.co

[3]我们将感知能动性与控制感建模为连续变量，因为它是根据11项量表计算得出的综合得分。[4]我们根据Chen等人对比值比的大小进行解读。 [15] 其中，OR = 1.68、3.47和6.71分别相当于Cohen's d = 0.2（小）、0.5（中）和0.8（大）[5] https://num.pyro.ai/

马尔可夫链蒙特卡洛（MCMC），这是一种用于抽取后验分布 P(θ|D) 的随机抽样技术；该后验分布表示给定数据观测值 D 时，似然函数中参数的分布函数。具体而言，我们使用 NUTS（无 U 形转弯采样器）进行抽样。

我们通过分析聊天机器人条件下参与者的聊天记录，以定性证据补充了定量分析。我们对参与者提出的问题进行了主题分析 [35]。研究团队的一名成员首先对数据进行开放式编码，随后通过迭代和反思的过程不断完善这些编码。接着，同一位成员采用主轴编码，将这些编码归纳为更大的类别，从而提取出共同的主题。

4 结果

总体而言，Rumi 改善了参与者阅读知情同意书的效果。与采用传统的基于表格的知情同意流程的参与者相比，与 Rumi 互动的参与者能够回忆起知情同意书中的更多信息，并据此采取更多正确行动。我们还发现，在聊天机器人条件下，参与者认为自己与研究者之间的权力关系更加平等，并且在虚拟调查研究中提供了更高质量的回答。探索性路径分析揭示了一种潜在机制：基于聊天机器人的知情同意方法可能通过缩小权力差距来提高回答质量。

4.1 人们参与了 Rumi，并对体验感到满意

尽管参与者花费更多时间与 Rumi 聊天以完成知情同意流程，但总体而言，他们对完成知情同意流程所花费的时间与精力感到满意（聊天机器人：M = 4.72，SD = 1.54；表格：M = 4.48，SD = 1.57）。他们还表示，愿意在未来的知情同意经历中使用此类聊天机器人（聊天机器人：M = 4.21，SD = 1.89；表格：M = 4.03，SD = 1.84）。我们将这两项测量指标均建模为有序变量，并对比了两种条件下评分为 4 的累积似然后验分布。我们发现，参与者对时间与精力的感受以及他们表示的未来使用意愿，在两种条件之间均无显著差异。在所有情况下，累积优势差的最高后验密度区间（HPDI）[6] 均与实际等效区间（ROPE）[7] 0 ± 0.05 重叠（时间与精力：M = 0.05，94% HPDI：[0.01，0.08]；未来使用：M = -0.22，94% HPDI：[-0.49，0.01]），这表明参与者并未显著更倾向于在时间与精力以及未来使用这两项指标上，在某一种条件下表示不同意或同意，而非另一种条件。

我们深入分析了参与者访谈记录，以进一步了解参与者的交互模式。我们发现，参与者积极地与 Rumi 互动。参与者总共提出了 449 个问题（M = 3.77，SD = 2.56），Rumi 回答了其中的 389 个（ 85.97% ）。我们将问题归纳为四个主要类别：Rumi 的能力（12.69%）、研究团队信息（11.58%）、研究信息（56.15%）和闲聊（19.59%）。参与者会询问 Rumi 能做什么，以及 Rumi 能回答哪些问题（例如，“你知道些什么？”[P107]）。另一类问题与研究团队有关（例如，“谁是 *[研究者姓名]*？”[P57]）。通过这些问题，参与者可以进一步了解研究团队，从而开始建立融洽关系。不出所料，人们提出最多的问题都与研究本身有关。具体而言，参与者询问了研究流程（45.23%；例如，“完成调查问卷后我该做什么？”[P41]）、风险（28.17%；例如，“我的信息安全吗？”[P83]）、补偿详情（12.30%；例如，“之后我会得到什么？”[P94]），以及有关研究的一般信息（14.30%；例如，研究目的、调查主题等）。有趣的是，一些参与者开始与 Rumi 闲聊，例如 “你今天过得怎么样？”[P43]，或 “我昨天没睡好。你睡觉吗？”[P67]，，这表明他们的参与度更高。这些问题表明，参与者愿意投入精力与 Rumi 互动。Rumi 也帮助参与者厘清了知情同意书中的重要信息。

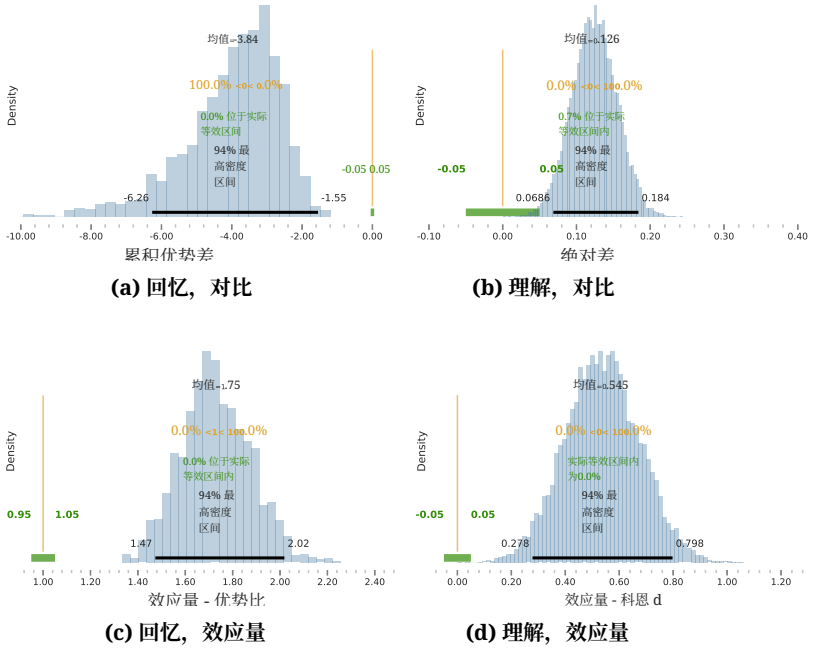
以往研究表明，引入交互性往往会增加用户负担，进而可能损害用户体验，尤其是在阅读知情同意书时，因为所需的时间和精力是一大障碍 [20, 29]。我们的结果表明，尽管与常规在线知情同意相比，使用 Rumi 花费的时间更长，但人们仍愿意主动与 Rumi 互动，以完成知情同意流程。我们认为，规模和速度不应成为知情同意流程中唯一的衡量标准。尤其是在这种高风险情境下，为了带来更具吸引力的体验并引起人们的注意而牺牲速度和规模，是值得考虑的。

4.2 Rumi 促使人们更好地阅读知情同意书

总体而言，与以往研究 [23, 63, 64], 相似，知情同意书的阅读效果并不理想。在回忆方面，两个条件下总共只有 26 位参与者正确回忆出颜色短语列表中的“橙色”和“蓝色”。平均而言，人们在回忆任务中的得分为 0.63 分（满分 2 分，SD = 0.64），在理解任务中的得分为 53.22%（SD = 24%）。

与静态表单相比，通过 Rumi 阅读知情同意书能显著提升同意书阅读效果。与 Rumi 互动的参与者能从知情同意书中回忆起更多正确的颜色词（聊天机器人：M = 0.76，SD = 0.71；表单：M = 0.51，SD = 0.54）。我们将回忆建模为有序变量，并估计获得总分 50%（2 分中的 1 分）的累积似然后验分布。累积优势差的最高后验密度区间不包括实际等效区间 0 ± 0.05（M = -3.84，94% 最高后验密度区间：[-6.26，-1.55]）。估计的效应量（优势比）较小（M = 1.75，94% 最高后验密度区间：[1.47，2.02]，，不包括实际等效区间 0 ± 0.05；见图2）。结果表明，参与者借助 Rumi 能更好地回忆内容。

[6]最高后验密度区间涵盖了后验密度的 94%。它与非贝叶斯统计学中使用的置信区间概念相似，但并不相同。在非贝叶斯统计学中，94% 置信区间通常被非正式地解释为“该参数以 94% 的概率落在某个特定区间内；两端的宽度相等（即各为 3%）”；而最高后验密度区间是覆盖后验分布 94% 区域的最密集区间。最高后验密度区间必然包含最可能的值，但置信区间并不总是如此；参见 McElreath [57]。关于置信区间更严谨的定义，请参见 Hoekstra 等人 [43]。[7] 不同于非贝叶斯统计学（在其中可以询问两个条件的均值是否不同 $P(\mu_1 \neq \mu_2)$），在贝叶斯统计中，人们会询问分布的最高后验密度区间$P(\mu_1 - \mu_2)$——即两个条件均值之差的分布——是否排除了一个我们可以认为两种治疗等效的区间。这个等效区间取决于具体领域。在贝叶斯数据分析中，如果后验分布的最高后验密度区间位于实际等效区间之外，则认为结果具有显著性。



(a) 回忆，对比 **(b) 理解，对比**

(c) 回忆，效应量 **(d) 理解，效应量**

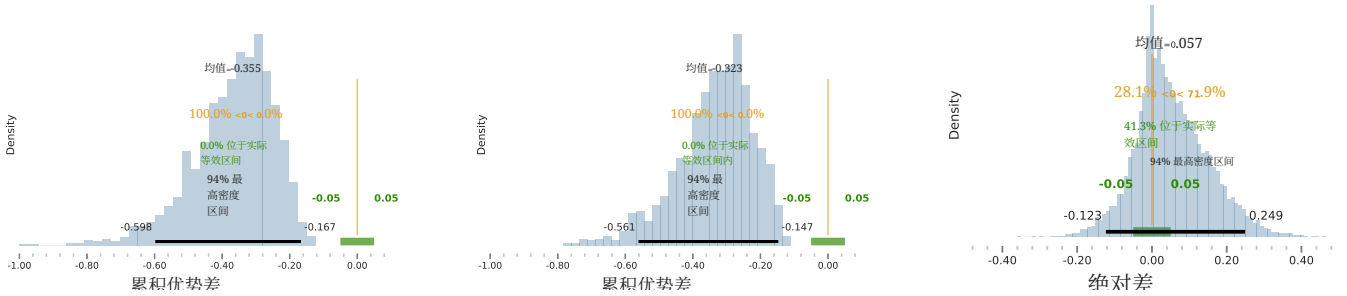
图 2：第一行展示了聊天机器人条件与表单条件之间的后验分布对比，内容包括回忆任务中正确率低于 50% 的累积优势，以及理解得分的均值。第二行展示效应量分布：回忆的优势比和理解的 Cohen's d。(a)、(b) 和 (d) 图中，橙色竖线位于 0 处，绿色条形表示实际等效区间。类似地，(c) 图中橙色竖线位于 1 处，并有一个绿色的实际等效区间。落入这些实际等效区间的效应分布表明，条件之间没有差异，或不存在效应。请注意，并非所有图的 x 轴刻度都相同。主要发现：与表单条件的参与者相比，聊天机器人条件的参与者在回忆任务中给出了更多正确答案，且理解得分更高。这些差异均具有统计显著性。

聊天机器人条件下的参与者对知情同意书的理解也更好（聊天机器人：M = 61%，SD = 22%；表单：M = 46%，SD = 23%）。我们估计了聊天机器人条件和表单条件下均值的后验分布。我们发现，聊天机器人条件（M = 51%，94% 最高后验密度区间：[39%，62%]）与表单条件（M = 51%，94% 最高后验密度区间：[27%，48%]）之间的差异具有统计显著性（不包括实际等效区间 0 ± 0.05），效应量（Cohen's d）为中到大（M = 0.55, 94% 最高后验密度区间：[0.28, .80], ，不包括实际等效区间；见图）。这表明参与者理解了内容，并能根据知情同意书采取更好的行动来保护自己的权利。

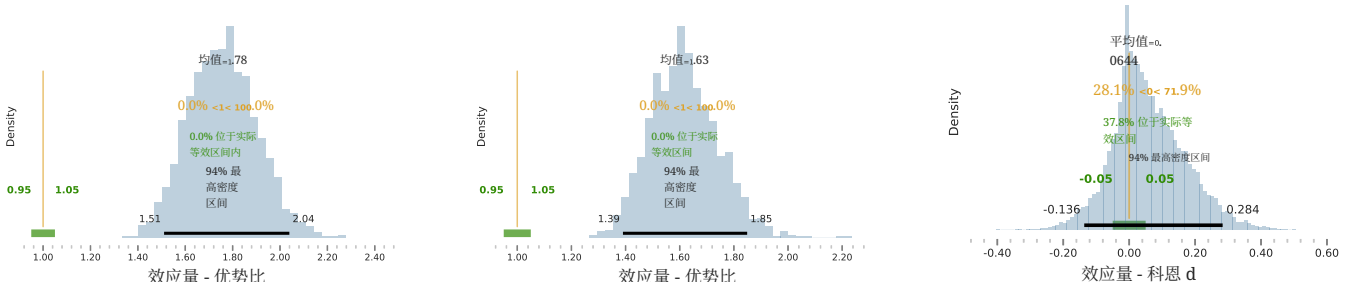
研究结果清晰地回答了研究问题1。聊天机器人驱动的知情同意流程提升了同意书阅读效果，包括回忆知情同意书中的信息以及理解其内容。观察到的改善可能由两个因素促成。首先，参与者可能投入度更高。Rumi 的互动功能旨在模拟面对面的体验，让研究助理能够积极与参与者互动。许多研究参与者喜欢这种类似真人互动的体验，并评论说 “这个机器人非常友好。”[P61] 以及 “我喜欢机器人和我说话的方式。”[P58]。参与者也很喜欢 Rumi 陪他们一起阅读知情同意书：“有机器人用文字陪我读知情同意书，比面对一大堵文字更轻松、更愉快。谢谢 *:D*”[P15]。正如我们在第4.1节所展示的，在聊天机器人条件下，参与者在知情同意流程中花费的时间明显更长。虽然花费更多时间并不总能带来更好的阅读效果 [20], ，但这可能意味着参与者投入度更高，而投入度对阅读理解起着关键作用 [36]。其次，Rumi 旨在回答人们的问题。参与者总共提出了 449 个问题，Rumi 回答了其中的 85.97%。参与者很认可 Rumi 实时回答问题的能力：“聊天机器人能马上回答我的问题，真挺酷的。”*[P32]* 不过，有些参与者提到 Rumi 无法完全理解他们的问题。尽管聊天机器人的能力有限，但它能够即时回答问题，或许有助于提升同意书阅读效果。

4.3 Rumi 助力建立参与者-研究者关系

与 Rumi 互动的参与者认为，他们与负责该研究的研究者之间的权力关系更加平等。我们的结果表明，与表单条件下的人相比，聊天机器人条件下的人对研究者更加信任（聊天机器人：M = 5.60，SD = 1.40；表单：M = 4.92，SD = 1.93），并且认为自己与研究者的关系更像是一种伙伴关系（伙伴关系），这表明权力差距更小（聊天机器人：M = 4.26，SD = 1.70；表单：M = 3.63，SD = 1.74）。与第 4.1 节类似，我们将这两个变量视为有序变量，并估计累积似然的后验分布。同样，我们构建了聊天机器人条件下评分为 4（即 7点李克特量表的中点）的累积优势与表单条件下相应评分的累积优势之差的分布。（



(a) 伙伴关系，对比 (b) 信任，对比 (c) 能动性与控制感，对比



(d) 伙伴关系，效应量 (e) 信任，效应量 (f) 能动性与控制感，效应量

图 3：第一行展示了聊天机器人条件与表单条件之间的后验分布对比，内容包括在伙伴关系和信任方面观察到中性或更低评分的累积优势，以及能动性与控制感的均值。第二行展示了效应量分布：伙伴关系和信任的优势比，以及能动性与控制感的 Cohen's d。图 (a)、(b)、(c)、(f) 中，橙色竖线位于 0 处，绿色条带表示实际等效区间。同样，图 (d) 和 (e) 中的橙色竖线位于 1 处，并显示绿色实际等效区间。效应分布落入这些实际等效区间，表明条件之间没有差异，或不存在效应。请注意，并非所有图表的 x 轴刻度都相同。主要发现：与表单条件下的参与者相比，与 Rumi 互动的参与者认为自己与研究者的关系更像伙伴关系，也更信任研究者。这一差异具有统计显著性。尽管聊天机器人条件下的参与者报告了更高的感知能动性与控制感，但这一差异并不具有统计显著性。

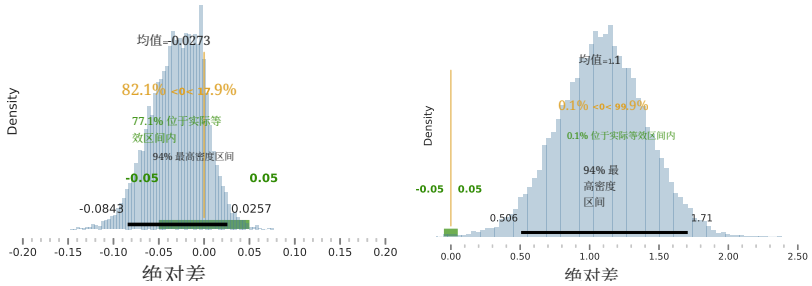
7点李克特量表的中点）在聊天机器人条件下的累积优势与表单条件下该评分的累积优势之差。根据估计的后验分布，我们发现两项测量指标中累积似然的差异均具有统计显著性（伙伴关系：M = -0.36，94% 最高后验密度区间：[-0.60，-0.17]，，不包括实际等效区间 0 ± 0.05；信任：M = -0.67，94% 最高后验密度区间：[-1.14，-0.28]，，不包括实际等效区间 0 ± 0.05）。优势比表明，这两项测量指标的效应量都较小（伙伴关系：M = 1.78，94% 最高后验密度区间：[1.51，2.04]，，不包括实际等效区间 0 ± 0.05；信任：M = 1 63, 94 [1 39, . % 最高后验密度区间：. .85]，，不包括实际等效区间 0 ± 0.05 ）。

然而，我们并未观察到参与者在知情同意流程结束后，其能动性与控制感存在显著差异。与表单条件相比，与 Rumi 互动的参与者报告了更强的能动性与控制感；该感受通过正向和负向自主性构成的综合得分来衡量（M = 5.15；SD = 0 78 .）。由于我们使用 11项量表的综合得分来衡量参与者的能动性与控制感，因此将该得分视为连续变量，并对比聊天机器人条件和表单条件的均值。我们估计了均值差异的后验分布，发现差异不具有统计显著性（M = 0.06，94% 最高后验密度区间：[-0.12，0.25]，，与实际等效区间 0 ± 0.05 重叠）。

通过 Rumi 参与在线知情同意流程，能够增进参与者对研究者的信任，并缩小权力差距。观察到的这一效应可能由两个因素解释。首先，如 [69]，所述，更有效的同意书阅读或许能通过弥合信息差距并确保参与者作出自愿决定来缩小权力差距。我们确实观察到参与者的同意书阅读情况与其和研究者之间的权力关系存在显著相关性（回忆：r(236) = 0.14，p = 0.03；理解：r(236) = 0.21，p < 0.01）。观察到的权力关系差异可能归因于更有效的沟通。其次，Rumi 的设计所具有的拟人性可能有助于研究者与参与者建立融洽关系，从而可能缩小权力差距 [45]。在我们的研究中，Rumi 被设定为虚拟研究助理，并代表研究团队。作为研究者与参与者之间的首次互动，知情同意流程除了告知研究参与者研究信息外，也可能发挥建立融洽关系的作用。

4.4 Rumi 提升调查回答质量

与聊天机器人互动的参与者在虚拟调查中给出了质量更高的回答。对于基于选择的问题，聊天机器人条件下的参与者表现出的调查敷衍作答较少



(a) 非差异化，对比 **(b) RQI，对比**

(c) 非差异化、效应量 **(d) RQI、效应量**

图 4：第一行展示了聊天机器人条件与表单条件在非差异化和响应质量指数（RQI）上的均值差异后验分布。第二行展示了基于后验分布计算的这两项测量指标的效应量分布（Cohen's d）。每张图中的橙色竖线位于 0 处，绿色条带表示实际等效区间（ROPE）。这表示条件之间没有差异，或不存在效应。请注意，各图的 x 轴刻度并不相同。主要发现：与Rumi 互动的参与者在封闭式问题上的非差异化程度较低，但差异不具有统计显著性。我们还观察到，聊天机器人条件下的参与者对开放式问题给出了质量更高的回答。后验分布表明，这一差异具有统计显著性，效应量为中等。

行为（聊天机器人：M = 0.42，SD = 0.22；表格：M = 0.47，SD = 0.25）。后验分布显示，观察到的差异在统计学上并不显著（M = - 0.03，94% 最高后验密度区间：[-0.08，0.03]，，与 0 ± 0.05 的实际等效区间重叠）。

对于开放式问题，聊天机器人条件下的参与者给出了质量更高的回答（聊天机器人：M = 5.38，SD = 2.29 = 4.17， = 2.43 .；表单：M . SD .）。两个条件之间 RQI 差异的后验分布显示，均值对比为 1.1，最高后验密度区间（HPDI）为 [0.50，1.71]. 由于 94% HPDI 位于具有显著性的实际等效区间（ROPE） 0 ± 0.05之外，结果表明效应具有统计显著性，且效应量为中等（Cohen's d：M = 0.47，94% HPDI：[0.21，0.73]，，不包含实际等效区间（ROPE） 0 ± 0.05）。

研究结果回答了研究问题3。在知情同意流程中与聊天机器人互动的参与者，对开放式问题给出了质量显著更高的回答。处于高风险条件下的参与者也愿意进一步阐述自己的回答。根据现有研究，这种观察到的效果可能源于更有效的知情同意书阅读 [21, 23]，以及研究者—参与者关系中权力差距的缩小 [12, 45]。我们在第4.5节进一步考察了两种潜在路径。

4.5 权力差距缩小或可解释研究回复质量的提升

现有研究表明，研究者与参与者之间的权力差距会对参与者的投入程度和自我披露水平产生重要影响 [12, 45]。更平等的权力动态有助于建立信任，并增强参与者的自主性与控制感，从而可能提高研究回答质量。另一方面，更有效的知情同意书阅读也可能提升研究回答质量，因为参与者能够更好地理解研究流程 [21, 23]。因此，我们探讨了聊天机器人驱动的知情同意影响研究回答质量的两种潜在中介路径：缩小权力差距和改善知情同意书阅读。

我们构建了一个结构方程模型（SEM）来回答这个问题。如图5所示，我们将同意方式（聊天机器人与表格）设为研究者—参与者权力关系和知情同意书阅读的预测变量。研究者—参与者权力关系是一个综合得分，由研究者与参与者之间的关系以及参与者感知到的自主性与控制感构成。知情同意书阅读通过标准化，将回忆和理解合并为一个单一指标。研究回答质量则结合了基于选择的问题中的非差异化和开放式问题的研究

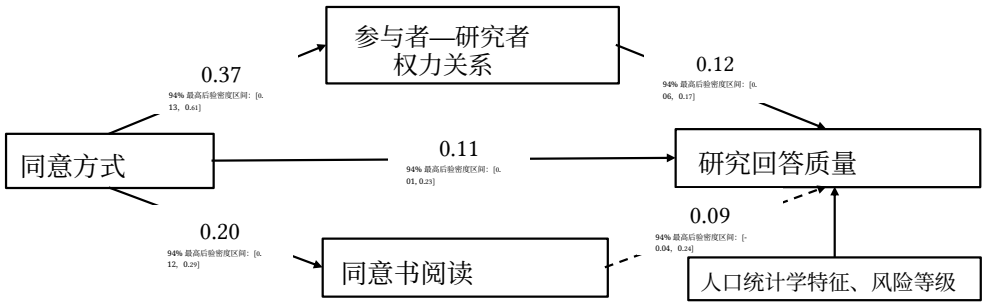


图5：该图展示了一个贝叶斯结构方程模型（SEM），用于理解聊天机器人驱动的知情同意为何能提升研究质量。根据以往文献的建议，我们考察了两条潜在路径：1）研究者—参与者关系中的权力差距缩小；2）同意书阅读改善。结果表明，权力差距缩小产生了显著的部分中介效应。

回答质量指数。我们将研究者—参与者权力关系和知情同意书阅读都设为研究回答质量的预测变量，同意方式也被设为其预测变量。我们认为，研究者—参与者权力关系能够预测研究回答质量，因为权力动态会影响信任和自主性，而这两者是自我披露的关键因素 [12, 45]。我们认为，知情同意书阅读能够预测研究回答质量，原因在于参与者理解并遵循研究说明的程度。同意方式是否能预测研究回答质量是我们的主要研究问题；另外两条路径则介导了同意方式对回答质量的影响。与所有其他模型一样，我们控制了参与者的人口统计学特征和风险等级。所有变量均被视为显变量，并以正态分布进行建模。我们使用贝叶斯推断和 No-U-Turn 采样器 [8] 对所提出的结构方程模型进行拟合。模型拟合完美，后验预测p值为 = 0 48 = 0 01, 94 [0 00, 0 06] = 0 99, 。；94 [0 98, 1 00] Γ = 0 99, 94 [0 96, 1 00] BRMSEA . % 最高后验密度区间：. .；BMc . % 最高后验密度区间：. .；adjBˆ . % 最高后验密度区间：. . .

结果表明，同意方式对研究回答质量具有显著的总效应（β = 0.18，94% 最高后验密度区间：[0.08，0.29]，，不包括实际等效区间 0 ± 0.05）。值得注意的是，如图5所示，同意方式通过参与者—研究者权力关系影响研究回答质量，存在一条间接路径（β = 0 12, . % 最高后验密94 [0 06,0 17], 0 ± 0.05度区间：. .，不包括实际等效区间）。这条路径使总效应降低了39%。然而，直接效应仍然显著，表明仅存在部分中介。同时，我们未观察到同意方式通过同意书阅读影响研究回答质量的显著路径（β = 0 09, . % 最高后验密94 [0 04, 0 24], 0 ± 0.05度区间：- . .，与实际等效区间重叠）。

贝叶斯结构方程模型显示，聊天机器人驱动的知情同意之所以可能改善后续的研究回答质量，或许是因为它减少了研究者与参与者之间的权力不对称。这一发现与既有研究一致：作为参与者与研究者之间早期互动之一，知情同意流程能够弥合信息差距，并影响研究者—参与者权力关系，最终可能有助于提升研究回答质量和数据丰富度 [12, 45, 69]。然而，模型并未显示同意书阅读得到改善后存在潜在的作用路径。我们认为有两种可能的解释。首先，我们的研究流程简单直接。改善同意书阅读未必能显著帮助参与者为后续研究做好准备。其次，与医学试验相比，我们的研究风险较低。现有关于临床试验的研究发现可能无法推广适用 [21]。

5 讨论

在本研究中，我们发现，聊天机器人驱动的知情同意流程能够有效改善知情同意书阅读体验、缩小权力差距，并最终提升研究回答质量。本节将讨论如何设计更有效的对话式 AI 知情同意体验。

5.1 个性化聊天机器人驱动的知情同意体验

在逐轮对话中，聊天机器人可以询问参与者的经历和偏好，以启动个性化体验 [86]。在我们的研究中，我们发现参与者喜欢个性化体验带来的感受。例如，一位参与者评论道：“从没见过这样的机器人。我喜欢这种感觉，Rumi 是在和我说话。”*[P11]*。有鉴于此，我们应考虑将个性化融入聊天机器人驱动的知情同意体验中。然而，在在线知情同意中引入个性化时，我们需要格外谨慎，避免产生违背自愿参与原则的不必要劝诱效果 [73] 。参与决定应完全出于自愿。个性化的目标应是帮助参与者理解知情同意书，而不是推动参与者参与研究。

人工智能驱动的聊天机器人可以根据参与者以往阅读知情同意书的经历，突出显示重要内容。一方面，现有研究表明，经验更丰富的研究参与者阅读知情同意书所花的时间更少，而且常常会遗漏重要信息 [66]，因为他们往往认为所有知情同意书都差不多。这种错误的假设使参与者面临不必要的风险。在这种情况下，他们已有的经验反而妨碍了同意书阅读。另一方面，诸如《通用数据保护条例》（GDPR）之类的法规通常要求知情同意书包含参与者可能已经熟悉的信息。

[8] https://cran .r-project .org/web/packages/blavaan/

在这种情况下，聊天机器人可以概述这些材料，帮助参与者更好地分配注意力，将注意力放在新内容上。因此，让聊天机器人了解参与者以往的知情同意经历、分析新的知情同意书、概述参与者可能熟悉的内容，并突出显示新内容，以确保其仔细阅读，将会很有帮助。一个值得探索的有趣想法是建立一个集中式聊天机器人，帮助参与者管理所有知情同意书，同时保护其以往研究经历的匿名性。尽管机构审查委员会要求研究者以通俗易懂的语言起草知情同意书，但有些研究程序确实很复杂。有时，为了清晰沟通，某些术语也必不可少，尤其是在高风险临床试验中。本研究的参与者很认可 Rumi 澄清问题并回答疑问的能力。因此，聊天机器人应包含交互功能，帮助参与者理解知情同意书。我们还可以借鉴一类以教育为重点的聊天机器人研究，在这类研究中，聊天机器人帮助人们学习新内容并复习材料 [17, 68]。例如，聊天机器人可以先评估参与者对研究主题已有的知识，以确定需要进行哪些解释，并在最后提问，以确保取得预期的学习成果。

然而，时间与精力是阅读知情同意书的最大障碍之一，因此需要适当权衡并审慎考量交互功能带来的益处与风险。具体而言，我们需要考虑参与者阅读知情同意书所需的时间成本。参与者的补偿通常与其参与研究流程所花费的时间相关。如果补偿总额固定，交互式知情同意程序就会降低报酬率。在我们的研究中，我们为聊天机器人条件下的参与者在最后额外提供了奖金，以确保达到承诺的报酬率。尽管研究团队可能需要为每位参与者安排更多预算，但正如本研究所示，我们认为回答质量的提升最终将有益于研究结果。这样也能节省研究人员清理未仔细阅读研究流程的参与者所提供低质量数据的时间，尤其是在未受监督的参与者会造成危及研究质量的混杂因素时 [23]。

5.2 管理权力动态

我们的结果与先前关于有效知情同意流程作用的研究相呼应：有效的知情同意流程有助于缩小研究者—参与者权力关系中的权力差距，并提升研究回答质量。我们认为，聊天机器人可以通过调整自身与参与者之间的权力关系来影响研究者—参与者关系。我们还可以进一步拓展其用途，设计能够主动管理权力动态的知情同意聊天机器人。

从权力关系的角度来看，一个人的身份会强烈影响其相对于对方的权力 [30]。Cassell 提到，研究者所选择的身份，例如访谈员、促进者、发起者或研究者，都可能改变其与研究参与者之间的权力动态 [12]。我们可以仔细设计知情同意聊天机器人的身份，以适应不同情境。例如，为了缩小权力差距，我们可以将聊天机器人设计成研究伙伴，而非研究者。然而，设计虚拟代理的身份十分复杂。许多设计维度，包括外观、语言风格等，都需要仔细考量。任何不协调之处都可能破坏整体体验。

知情同意流程可以被视为研究者与参与者之间就信息披露进行的协商 [45]。研究者掌握有关研究的信息，而参与者则获得开展研究所需的知识和经验。Karnieli-Miller 等人指出，此类协商有可能通过向参与者提供更多信息来改变权力关系 [45]。因此，我们应准备一个具备协商对话能力的知情同意聊天机器人，使其能够理解参与者的请求、明确他们的信息需求，并主动管理有关研究的信息披露。

我们还可以通过考虑知情同意聊天机器人的所有权来增强参与者的权能。在本研究中，聊天机器人代表研究者行事，是研究团队的一员。虽然在大多数聊天机器人应用场景中，聊天机器人归创建者而非用户所有，但由参与者拥有的知情同意聊天机器人可能带来多项益处。首先，参与者将对对话历史拥有完全控制权。在这种情况下，参与者可以拥有一个安心提问、不必担心受到评判的空间。其次，此类知情同意聊天机器人可以成为满足所有知情同意需求的中心枢纽。它可以代表参与者行事，根据参与者的偏好分析同意书，并主动向研究者提问，以满足参与者的信息需求。

5.3 将人类专业知识与大语言模型相结合

许多参与者喜欢 Rumi 能够依据知情同意书中的内容，实时回答他们的问题。然而，要创建一个能够准确回答人们问题的聊天机器人，尤其是在高风险情境下，颇具挑战。由于自然语言理解能力有限，目前大多数商业聊天机器人构建平台的问答功能都依赖于人工编写的问答对数据库。在知情同意情境下，这尤其耗时，因为参与者的问题都针对具体的知情同意书，进一步限制了问答数据库的可复用性。尽管有些问题可以复用，例如，某个机构可能会使用相同的模板，而且一些研究流程可能相似，但仍需开展后续研究，设计支持此类共享实践的工具。

GPT-3 等大语言模型展现出了一种构建对话代理、以回答人们自然语言问题的新前景。虽然可以利用现成的大语言模型，通过上下文学习构建聊天机器人，以回答各种特定领域的问题，但大语言模型有时会生成不准确的信息，而且记忆长文档的能力有限 [11]。在高风险情境下，例如提供知情同意书时，应避免这两个缺陷，因为不准确的信息可能会误导参与者，使其作出不知情的决策。这不仅涉及伦理问题，还可能导致严重后果。例如，参与者在不了解具体研究流程的情况下同意参加研究，可能会经历意料之外的极端身体或心理压力。

因此，我们应在充分考虑上述不足的前提下，谨慎地利用大语言模型。一种值得考虑的框架是

将大语言模型与人类专业知识相结合。在我们的研究中，我们使用 GPT- 3 来增强问答对生成，从而提升 Rumi 的能力。为确保正确性，我们充当验证层，检查 GPT-3 生成的改写后的问题和答案是否正确且恰当。扩充后的问答数据库使 Rumi 能够捕捉到更多参与者的问题，并提供更多样的答案。我们认为，大语言模型与人类专家协作，可以促进更多聊天机器人开发任务。例如，可以利用大语言模型生成问题集，作为测试工具来发现问题并制定修复方案。这种方法能够加快迭代速度，而传统方法通常依赖于即时进行对话引导 [80]。除了将人类专业知识用作把关手段外，我们还应研究更完善的大语言模型控制机制，以支持事实性问答。例如，我们可以采用知识驱动方法 [32]，将纯文本知情同意书解析为结构化知识图谱，并利用该图谱引导大语言模型生成以知情同意书内容为依据的事实性答案。再次考虑到生成式模型的不足，我们认为，采用人在回路框架，才能安全地发挥生成式模型的优势，打造能力更强的知情同意聊天机器人。

总之，未来的研究应探讨有效的人工参与框架，以支持研究团队，尤其是缺乏人工智能专业知识和资源的团队，构建并测试能够始终提供事实性回答的知情同意聊天机器人。

5.4 未来方向

5.4.1 作为虚拟研究助理的聊天机器人。 除了此前探讨聊天机器人在研究中用途的工作（包括用于对话式调查 [83] 和民族志研究 [72], ）之外，我们也看到了未来打造虚拟研究助理的机会，它们可以帮助研究者从头到尾管理人体受试者研究。虚拟研究助理可以帮助研究者接触全球人口（如研究需要），与参与者互动并建立融洽关系，实施干预、收集高质量数据，并向参与者进行事后说明。在研究中，除了协助管理研究流程外，我们还发现有些参与者希望通过虚拟代理与研究者沟通。一位参与者对 Rumi 说：“让你的主人知道我喜欢这项研究。”[5]。这样的虚拟研究助理对于纵向实地研究尤其有帮助，因为在这类研究中，长期保持参与者的投入并收集高质量数据尤其困难且成本高昂。

然而，创建这样一个虚拟研究助理颇具挑战。首先，在研究的不同阶段与虚拟研究助理进行重复交互，会给交互设计带来新的挑战。该智能体需要随着时间推移，适应并回应个体独特的经历。其次，在人体受试者研究情境中，个性化是一把双刃剑。尽管个性化可能提高参与度，但高度个性化的聊天机器人可能因参与者之间缺乏一致性而引入不必要的混杂因素。第三，在参与者与研究人员之间进行沟通协调，需要一种新的交互范式。该智能体需要协调双方的沟通，并积极促成双方参与，以确保研究成功。尽管充满挑战，这样的虚拟研究助理仍可帮助研究人员开展具有可扩展性、稳健性和吸引力的人体受试者研究。

5.4.2 超越研究情境的在线知情同意。个人数据，从个人健康信息到网页浏览历史，已经变得越来越有价值。它为数百万种智能应用提供支持。随着世界变得更加互联，个人数据正成为一种新的权力来源。因此，尽管系统构建者热衷于收集数据，政策制定者和用户却对个人数据共享更加谨慎。例如，GDPR[9] 法规明确要求数据收集者在收集和存储任何个人数据之前，征得用户的同意。然而，目前在线数据同意的做法存在很大缺陷 [5]。例如，用户通常需要在没有任何指导的情况下阅读一份冗长的文件。因此，如何帮助用户就其数据作出有意义且知情的同意，正成为一项日益突出的挑战。在本研究中，我们发现，人工智能驱动的聊天机器人能够在人体受试者研究情境中提供有效的知情同意。未来，我们应探索此类智能体在更广泛情境下用于数据共享同意的潜力。

5.5 局限性

我们认识到本研究方法存在若干局限。首先，作为首个此类研究，我们的主要目标是探索人工智能驱动的聊天机器人知情同意流程的潜在益处与局限性。我们通过结构方程模型进一步探索了可能解释所观察到效应的潜在路径；具体而言，聊天机器人驱动的知情同意流程通过改变权力关系，提高了研究回答质量。然而，由于本研究具有探索性质，我们无法据此推断出有力的因果关系。未来需要开展验证性研究，以确认观察到的效应并解释其作用机制。

其次，我们的研究设计范围仅限于采用调查问卷的在线研究。尽管我们设计了三个风险等级，以模拟收集不同类型数据的研究，但与可能造成严重后果的高风险临床试验相比，在我们的研究中，未充分知情而作出决策的风险较低。有两个因素可能限制此类设计的可推广性。首先，对于风险较低的研究，人们往往不太关注知情同意书 [63]。如果参与者在两种条件下都更加认真地阅读知情同意书，我们观察到的差异可能会更小。其次，研究流程简单直接，例如完成一份调查问卷。尽管这代表了大多数在线研究，但有些研究可能包含更复杂的研究程序，例如玩游戏；在这些研究中，充分理解流程细节可能会对后续研究产生更大的影响。因此，我们需要在各种情境以及复杂程度各异的研究中考察聊天机器人驱动的知情同意。

第三，拒绝参与研究的参与者未纳入我们的分析。尽管在我们的研究设计中，决定不参与研究的参与者会被重定向至第2节，并获得完成同意书评估的机会，但有些参与者可能在未作答的情况下关闭了知情同意书。尽管其他因素也可能发挥作用，例如可用性问题，这些参与者仍可能仔细阅读知情同意书，并在充分知情的情况下决定不参与。本研究设计未将这些参与者纳入分析。在我们的研究中，278名参与者中有26名打开了我们的知情同意书

[9]https://gdpr-info.eu/

（聊天机器人条件：N = 18；表单条件：N = 8），但未完成填写。我们认为这一潜在混杂因素对研究结果的影响有限，但仍有必要开展后续研究。

第四，我们的研究旨在考察使用人工智能驱动的聊天机器人引导知情同意流程的整体效果。然而，聊天机器人的设计（例如语言风格、名称和外观）及其能力（例如自然语言理解、问答和对话管理）都是重要的设计维度，可能会影响最终结果。作为第一步，我们的目标是打造 Rumi，提供尽可能出色的体验。本研究收集的数据不足以区分并量化各个设计因素的独立贡献。由于每项交互功能都有其益处和风险 [83], ，因此严谨地量化不同功能的贡献很有价值。不过，这需要开展额外且完全受控的实验，超出了本研究的范围。

最后，尽管聊天机器人在我们的日常生活中日益普及，从客户服务到对话式调查 [39], ，但使用聊天机器人开展知情同意流程仍不常见。作为首项此类研究中使用的聊天机器人，Rumi 对大多数参与者而言都很新颖。由于我们无法在当前研究设计中控制新奇效应，因此不了解新奇因素的影响。我们计划通过纵向研究考察新奇效应的影响；不过，随着聊天机器人逐渐成为常态，新奇效应可能会像任何新兴技术一样逐渐消退。就本研究而言，随着聊天机器人驱动的知情同意变得越来越普遍，这种效应很可能会逐渐消退，类似于如今越来越多的研究采用电子知情同意表 [20]。

6 结论

本文探讨了人工智能驱动的聊天机器人在改善在线知情同意方面的作用。我们开发了 Rumi，这是一款人工智能驱动的聊天机器人，能够向参与者致意，逐节讲解知情同意书，回答参与者的问题，并收集人们对知情同意的回应，以模拟面对面的知情同意体验。我们设计并开展了一项被试间研究，在一项关于人们社交媒体使用情况的在线调查研究中，将 Rumi 与典型的基于表格的知情同意流程进行比较，以考察聊天机器人在引导在线知情同意流程中的整体效果。我们发现，Rumi 改善了同意书阅读，促进了参与者与研究者之间更平等的权力关系，并提高了研究回答质量。我们的探索性路径模型表明，研究回答质量的提升可能归因于聊天机器人驱动的知情同意流程缩小了权力差距。鉴于研究结果以及创建此类聊天机器人的简便性，我们的工作提出了一种开展有效在线知情同意的新颖且有前景的方法。随着聊天机器人日益普及，我们的研究结果也为打造更有效的知情同意聊天机器人提供了重要的设计启示。

致谢

我们感谢 Heng Ji、Brent W. Roberts、Yu Xiong、Michelle X. Zhou 以及匿名评审者对本研究提出的深思熟虑的意见和建设性反馈。Tiffany Wenting Li 获得了 Google 博士奖学金项目的资助。

参考文献

[1] Patricia Agre, Frances A Campbell, Barbara D Goldman, Maria L Boccia, Nancy Kass, Laurence B McCullough, Jon F Merz, Suzanne M Miller, Jim Mintz, Bruce Rapkin 等。2003年。改善知情同意：媒介并非信息本身。IRB：伦理与人体研究 25, 5 (2003年), S11–S19。 [2] Emily E Anderson, Susan B Newman 和 Alicia K Matthews。2017年。改善知情同意：利益相关者观点。AJOB 实证生物伦理学 8, 3 (2017年), 178–188。[3] Michela Assale, Erica Barbero 和 Federico Cabitza。2019年。知情同意数字化：面向实践的设计挑战。载于2019年IEEE第32届基于计算机的医疗系统国际研讨会（CBMS）。IEEE, 609–615。 [4] Martina Balestra, Orit Shaer, Johanna Okerlund, Madeleine Ball 和 Oded Nov。2016年。接触社会标注对在线知情同意信念与行为的影响。载于第19届ACM计算机支持的协同工作与社交计算会议论文集。900–912。 [5] Solon Barocas 和 Helen Nissenbaum。2014年。大数据如何绕过匿名性与同意。载于《隐私、大数据与公共利益：参与框架》1 (2014年), 44–75。 [6] 贝勒大学。2016年。你沉迷于社交媒体吗？六个问题。《科学日报》（2016年10月）。[7] Timothy Bickmore, Dina Utami, Shuo Zhou, Candace Sidner, Lisa Quintiliani 和 Michael K Paasche-Orlow。2015年。虚拟代理对研究知情同意进行自动化解释。载于国际智能虚拟代理会议。Springer, 260–269。[8] Timothy W Bickmore, Ha Trinh, Stefan Olafsson, Teresa K O'Leary, Reza Asadi, Nathaniel M Rickles 和 Ricardo Cruz。2018年。使用对话助手获取医疗信息时患者与消费者面临的安全风险：一项针对 Siri、Alexa 和 Google 助理的观察性研究。医学互联网研究杂志 20, 9 2018 11510 ( ), e 。[9] Else Bjørn, Peter Rossel 和 Soren Holm。1999年。能否改进提供给研究受试者的书面信息？——一项实证研究。Journal of Medical Ethics 25, 3 1999 263 267 ( ), – 。[10] Su Lin Blodgett, Solon Barocas, Hal Daumé III 和 Hanna Wallach。2020年。语言（技术）即权力：自然语言处理中"偏见"的批判性综述。载于计算语言学协会第58届年会论文集。5454–5476。[11] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell 等。2020年。语言模型是小样本学习器。《神经信息处理系统进展》33 (2020年), 1877–1901。 [12] Joan Cassell。1980年。开展田野工作的伦理原则。《美国人类学家》82, 1 (1980年), 28–41。 [13] Barrie R Cassileth, Robert V Zupkis, Katherine Sutton-Smith 和 Vicki March。1980年。知情同意——为何其目标未能充分实现？《新英格兰医学杂志》302, 16 (1980年), 896–900。 [14] Cindy Chen, Pou-I Lee, Kevin J Pain, Diana Delgado, Curtis L Cole 和 Thomas R Campion Jr.。2020年。在学术医疗中心以电子知情同意取代纸质知情同意进行研究：一项范围综述。AMIA转化科学峰会论文集 2020 (2020年), 80。[15] Henian Chen, Patricia Cohen 和 Sophie Chen。2010年。多大的优势比才算大？解读流行病学研究中优势比的大小。统计学通讯：模拟与计算® 39, 4 (2010年), 860–864。[16] Shu-Hsin Chen。2011年。研究者与被研究者之间的权力关系：对本土与非本土民族志访谈的分析。田野方法 23, 2 (2011年5月), 119–135。[17] Francesco Colace, Massimo De Santo, Marco Lombardi, Francesco Pascale, An-tonio Pietrosanto 和 Saverio Lemma。2018年。用于电子学习的聊天机器人：一个案例研究。国际机械工程与机器人研究杂志 7, 5 (2018年), 528–533。[18] Cathy A Coyne, Ronghui Xu, Peter Raich, Kathy Plomer, Mark Dignan, Lari B Wenzel, Diane Fairclough, Thomas Habermann, Linda Schnell, Susan Quella 等。2003年。针对临床试验参与者进行的一项易读版知情同意声明随机对照试验：东部肿瘤协作组的一项研究。临床肿瘤学杂志 21, 5 (2003年), 836–842。 [19] Jorden A Cummings, Jessica M Zagrodney 和 T Eugene Day。2015年。开放数据政策对人体受试者研究参与同意的影响：参与者行动与其所报告顾虑之间的差异。公共科学图书馆综合刊 10, 5 (2015年), e0125208。[20] Evelien De Sutter, Drieda Zaçe, Stefania Boccia, Maria Luisa Di Pietro, David Geerts, Pascal Borry, Isabelle Huys 等。2020年。生物医学研究中电子知情同意的实施及利益相关者观点：系统综述。医学互联网研究杂志 22, 10 (2020年), e19129。 [21] Marcela G Del Carmen 和 Steven Joffe。2005年。医疗治疗与研究中的知情同意：综述。《肿瘤学家》10, 8 (2005年), 636–641。[22] David DeVault, Ron Artstein, Grace Benn, Teresa Dey, Ed Fast, Alesia Gainer, Kallirroi Georgila, Jon Gratch, Arno Hartholt, Margaux Lhommet 等。2014年。

SimSensei Kiosk：用于医疗决策支持的虚拟人访谈员。载于2014年自主智能体与多智能体系统国际会议论文集。1061–1068。 [23] Benjamin D Douglas, Emma L McGorray 和 Patrick J Ewell。2021年。一些研究者穿黄色裤子，但读同意书的参与者更少：探索并改进人体受试者研究中的同意书阅读。Psychological Methods 26, 1 (2021年), 61。 [24] Graham M Dresden 和 M Andrew Levitt。2001年。修改标准行业临床试验同意书，可作为知情同意流程的一部分改善患者对信息的保留。Academic Emergency Medicine 8, 3 (2001年), 246–252。 [25] Catherine A Ebbs。1996年。定性研究探究：权力与伦理问题。Education 117, 2 (1996年), 217–223。 [26] Ruth R Faden 和 Tom L Beauchamp。1986年。《知情同意的历史与理论》。牛津大学出版社。 [27] Ahmed Fadhil 和 Silvia Gabrielli。2017年。应对促进健康生活方式的挑战：al-chatbot方法。载于第11届EAI泛在计算技术促进医疗保健国际会议论文集。261–265。 [28] Carlos Miguel Ferreira 和 Sandro Serpa。2018年。社会科学研究中的知情同意：伦理挑战。Int'l J. Soc. Sci. Stud. 6 (2018年), 13。 [29] James Flory 和 Ezekiel Emanuel。2004年。改善研究知情同意过程中研究参与者理解程度的干预措施：系统综述。Jama 292, 13 (2004年), 1593–1601。 [30] Michel Foucault。1982年。主体与权力。Critical Inquiry 8, 4 (1982年), 777–795。 [31] Joel A Friedlander, Greg S Loeben, Patricia K Finnegan, Anita E Puma, Xuemei Zhang, Edwin F De Zoeten, David A Piccoli 和 Petar Mamula。2011年。一种增强知情同意的新方法：一项比较表格式知情同意与电子辅助知情同意在儿科内镜检查中的前瞻性随机试验。Journal of Medical Ethics 37, 4 (2011年), 194–200。 [32] Yubin Ge, Ziang Xiao, Jana Diesner, Heng Ji, Karrie Karahalios 和 Hari Sundaram。2022年。我该问什么：一种基于知识驱动的方法，用于生成对话式调查中的后续问题。arXiv预印本 arXiv:2205.10977 (2022年)。 [33] Caitlin Geier, Robyn B Adams, Katharine M Mitchell 和 Bree E Holtz。2021年。在线研究的知情同意——有人在阅读吗？评估数字同意书的理解程度与阅读中的个体差异。Journal of Empirical Research on Human Research Ethics 16, 3 (2021年), 154–164。 [34] Lilian Ghandour, Rola Yasmine 和 Faysal El-Kak。2013年。未获充分告知便给予同意：研究伦理中的跨文化问题。Journal of Empirical Research on Human Research Ethics 8, 3 (2013年), 12–21。 [35] Graham R Gibbs。2007年。主题编码与分类。载于《定性数据分析》703 (2007年), 38–56。 [36] William Grabe。2008年。《第二语言阅读：从理论走向实践》。剑桥大学出版社。 [37] Bradford H Gray, Robert A Cooke 和 Arnold S Tannenbaum。1978年。涉及人体受试者的研究：这项实证研究评估了机构审查委员会的表现。《Science》201, 4361 (1978年), 1094–1101。 [38] Herbert P Grice。1975年。逻辑与对话。载于《言语行为》。Brill, 41–58。 [39] Jonathan Grudin 和 Richard Jacques。2019年。聊天机器人、人类机器人与通用人工智能的探索。载于2019年CHI计算系统中的人为因素会议论文集。1–11。 [40] Scott D Halpern, Jason HT Karlawish 和 Jesse A Berlin。2002年。统计效能不足的临床试验持续存在不道德行为。Jama 288, 3 (2002年), 358–362。 [41] Jeremy D Heider, Jessica L Hartnett, Emmanuel J Perez 和 John E Edlund。2020年。对研究情境的感知与理解如何随同意书特征和实验者指示而变化。Methods in Psychology 2 (2020年), 100015。 [42] Saras Henderson。2003年。护士与患者之间的权力失衡：可能妨碍护理伙伴关系的因素。Journal of Clinical Nursing 12, 4 (2003年), 501–508。 [43] Rink Hoekstra, Richard D Morey, Jeffrey N Rouder 和 Eric-Jan Wagenmakers。2014年。对置信区间的持续误解。Psychonomic Bulletin & Review 21, 5 (2014年), 1157–1164。 [44] Juji。2020年。面向聊天机器人设计者的Juji 文档。https://docs.juji.io/。 [在线；访问日期：14-June-2020]。 [45] Orit Karnieli-Miller, Roni Strier 和 Liat Pessach。2009年。定性研究中的权力关系。Qualitative Health Research 19, 2 (2009年), 279–289。 [46] Matthew Kay, Gregory L Nelson 和 Eric B Hekler。2016年。以研究者为中心的统计学设计：贝叶斯统计为何更契合人机交互的文化与激励机制。载于2016年CHI计算系统中的人为因素会议论文集。4521–4532。 [47] Soomin Kim, Joonhwan Lee 和 Gahgene Gweon。2019年。比较聊天机器人调查与网络调查的数据：平台和对话风格对调查回答质量的影响。载于2019年CHI计算系统中的人为因素会议论文集。1–12。 [48] Yujin Kim, Jennifer Dykema, John Stevenson, Penny Black 和 D Paul Moberg。2019年。直线作答：测量概述、指标比较及其在邮件与网络混合模式调查中的影响。Social Science Computer Review 37, 2 (2019年), 214–233。 [49] Michael M Knepp。2018年。使用问题改善学生阅读知情同意书的行为。Ethics & Behavior 28，7（2018年），560–577。 [50] V Manoj Kumar、A Keerthana、M Madhumitha、S Valliammai 和 V Vinithasri。2016年。面向健康寻求者的医疗聊天机器人。International Journal Of Engineering And Computer Science 5，03（2016年），16022–16025。 [51] Marlies Kustatscher。2014年。学校民族志中的知情同意：使用视觉磁铁探索参与、权力与研究关系。International Journal of Child, Youth and Family Studies 5，4.1（2014年），686–701。 [52] Christine Lavelle-Jones、Derek J Byrne、Peter Rice 和 Alfred Cuschieri。1993年。影响知情同意质量的因素。《英国医学杂志》306，6882（1993年），885–890。 [53] 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度自我披露。载于《2020年CHI计算系统中的人为因素会议论文集》。1–12。 [54] James R Lewis。1991年。用于计算机可用性研究的情境后问卷的心理测量学评估：ASQ。《ACM SIGCHI公报》23，1（1991年），78–81。 [55] Chung-Ying Lin、Anders Broström、Per Nilsen、Mark D Griffiths 和 Amir H Pakpour。2017年。运用经典测验理论和Rasch模型对波斯语卑尔根社交媒体成瘾量表进行心理测量学验证。《行为成瘾杂志》6，4（2017年），620–629。 [56] Yvonna S Lincoln 和 Egon G Guba。1985年。自然主义探究。sage。 [57] Richard McElreath。2015年。统计学再思考：结合R和Stan示例的贝叶斯课程。Chapman and Hall/CRC。 [58] Narjara Conduru Fernandes Moreira、Camila Pacheco-Pereira、Louanne Keenan、Greta Cummings 和 Carlos Flores-Mir。2016年。成年牙科患者对知情同意的理解与回忆：系统综述。《美国牙医协会杂志》147，8（2016年），605–619。 [59] Gary R Morrow。1980年。受试者同意书的可读性如何？Jama 244，1（1980年），56–58。 [60] Kenneth Mulligan、Jon A Krosnick、Wendy Smith、Melanie Green 和 George Bizer。2001年。态度评定量表中的非差异化：对调查满意化理论的检验。未发表手稿（2001年）。 [61] Long Ouyang、Jeff Wu、Xu Jiang、Diogo Almeida、Carroll L Wainwright、Pamela Mishkin、Chong Zhang、Sandhini Agarwal、Katarina Slama、Alex Ray 等。2022年。训练语言模型遵循人类反馈下的指令。arXiv预印本 arXiv:2203.02155（2022年）。 [62] Barton W Palmer、Nicole M Lanouette 和 Dilip V Jeste。2012年。多媒体辅助工具增强研究同意过程中理解的有效性：系统综述。机构审查委员会34，6（2012年），1。 [63] Eric R Pedersen、Clayton Neighbors、Judy Tidwell 和 Ty W Lostutter。2011年。本科生研究参与者会阅读心理学研究知情同意书吗？考察记忆效应、条件效应和个体差异。Ethics & Behavior 21，4（2011年），332–350。 [64] Evan K Perrault 和 Seth P McCullock。2019年。简明知情同意书受到欢迎——但仍未被理解：在在线研究中应用修订后的共同规则指南。Journal of Empirical Research on Human Research Ethics 14，4（2019年），299–306。 [65] Evan K Perrault 和 Samantha A Nazione。2016年。知情同意——不知情的参与者：在线社会科学同意书的不足与改进建议。Journal of Empirical Research on Human Research Ethics 11，3（2016年），274–280。 [66] Kyle R Ripley、Margaret A Hance、Stacey A Kerr、Lauren E Brewer 和 Kyle E Conlon。2018年。不知情的同意？参与者特征和传递格式对知情同意的影响。Ethics & Behavior 28，7（2018年），517–543。 [67] Michael C Rowbotham、John Astin、Kaitlin Greene 和 Steven R Cummings。2013年。交互式知情同意：与纸质同意书的随机比较。公共科学图书馆综合刊8，3（2013年），e58603。 [68] Sherry Ruan、Liwei Jiang、Justin Xu、Bryce Joe-Kun Tham、Zhengneng Qiu、Yeshuang Zhu、Elizabeth L Murnane、Emma Brunskill 和 James A Landay。2019年。Quizbot：用于事实性知识的基于对话的自适应学习系统。载于《2019年CHI计算系统中的人为因素会议论文集》。1–13。 [69] Peter H Schuck。1994年。重新思考知情同意。《耶鲁法律杂志》（1994年），899–959。 [70] Anne Sherlock 和 Sonya Brownie。2014年。患者对知情同意的回忆与理解：文献综述。ANZ journal of surgery 84，4（2014年），207–210。 [71] Alan R Tait、Terri Voepel-Lewis、Stanley J Chetcuti、Colleen Brennan-Martinez 和 Robert Levine。2014年。增强患者对医疗程序的理解：对包含内嵌练习的交互式多媒体程序的评估。《国际医学信息学杂志》83，5（2014年），376–384。 [72] Ella Tallyn、Hector Fried、Rory Gianni、Amy Isard 和 Chris Speed。2018年。民族志机器人：物联网时代的民族志收集。载于《2018年CHI计算系统中的人为因素会议论文集》。1–13。 [73] Kar Yan Tam 和 Shuk Ying Ho。2005年。网络个性化作为说服策略：精细加工可能性模型视角。Information systems research 16，3（2005年），271–291。

[74] Adam Tapal、Ela Oren、Reuven Dar 和 Baruch Eitam。2017年。能动性感量表：衡量对自身心智、身体和直接环境的有意识感知控制。*Frontiers in Psychology* 8（2017年），1552。[75] Vasti Torres 和 Marcia B Baxter Magolda。2002年。建构主义纵向研究中研究者角色的演变。*Journal of College Student Development*（2002年）。[76] Connie K Varnhagen、Matthew Gushta、Jason Daniels、Tara C Peters、Neil Parmar、Danielle Law、Rachel Hirsch、Bonnie Sadler Takach 和 Tom Johnson。2005年。在线知情同意有多充分？*Ethics & Behavior* 15，1（2005年），37–48。[77] James Walkup 和 Elinor Bock。2009年。潜在研究参与者想知道什么？他们认为自己已经知道什么？*Journal ofEmpirical Research on Human Research Ethics* 4，2（2009年），59–63。[78] Alex C Williams、Harmanpreet Kaur、Gloria Mark、Anne Loomis Thompson、Shamsi T Iqbal 和 Jaime Teevan。2018年。利用对话智能支持工作场所脱离与重新连接。载于*Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*。1–13。[79] Diane L Wolf。2018年。在田野工作中定位女性主义困境。载于*Feminist dilemmas in fieldwork*。Routledge，1–55。[80] Ziang Xiao、Q Vera Liao、Michelle X Zhou、Tyrone Grandison 和 Yunyao Li。2023年。通过专家征集为人工智能聊天机器人提供支持，以促进可信健康信息获取。*arXiv preprint arXiv:2301.10710*（2023年）。[81] Ziang Xiao、Michelle X Zhou、Wenxi Chen、Huahai Yang 和 Changyan Chi。2020年。如果我正确听懂了你：构建并评估访谈聊天机器人

积极倾听技能。载于*2020*年*CHI*计算系统中的人为因素会议论文集。1–14。[82] ZiangXiao、Michelle X Zhou和Wai-Tat Fu。2019年。谁应该成为我的队友：利用对话代理了解个人并促进团队协作。载于第*24*届国际智能用户界面会议论文集。437–447。[83] Ziang Xiao、Michelle X Zhou、Q Vera Liao、Gloria Mark、Changyan Chi、Wenxi Chen和Huahai Yang。2020年。告诉我关于你自己的事：利用人工智能驱动的聊天机器人开展包含开放式问题的对话式调查。*ACM*计算机与人类交互汇刊（*TOCHI*）27卷，3期（2020年），1–37。[84] T Yan。2008年。非差异化。调查研究方法百科全书第2卷（2008年），520–521。[85] Yizhe Zhang、Siqi Sun、Michel Galley、Yen-Chun Chen、Chris Brockett、Xiang Gao、Jianfeng Gao、Jingjing Liu和Bill Dolan。2019年。DialoGPT：用于对话响应生成的大规模生成式预训练。*arXiv*预印本*arXiv:1911.00536*（2019年）。[86] Michelle X Zhou、Gloria Mark、Jingyi Li和Huahai Yang。2019年。信任虚拟代理：人格的影响。*ACM*交互式智能系统汇刊（*TiiS*）9卷，2–3期（2019年），1–36。[87] Shuo Zhou、Timothy Bickmore、Michael Paasche-Orlow和Brian Jack。2014年。代理与用户一致性及其对虚拟出院护士满意度的影响。载于国际智能虚拟代理会议。Springer，528–541。