Under theBridge：招募软件开发人员参与实证研究所面临的网络挑衅与挑战

1st Ella Kokinda 克莱姆森大学计算学院 克莱姆森大学 美国克莱姆森
ekokind@clemson.edu

2nd Makayla Moster 克莱姆森大学计算学院 克莱姆森大学 美国克莱姆森
mmoster@clemson.edu

3rd James Dominic 克莱姆森大学计算学院 克莱姆森大学 美国克莱姆森
do
mini44@clemson.edu

4th Paige Rodeghero 克莱姆森大学计算学院 克莱姆森大学 美国克莱姆森
prodegh@clemson.edu

arXiv:2302.00174v2 [cs.SE] 6 Feb 2023

摘要——软件工程研究大多聚焦于软件工具、算法及其优化。近年来，作为一个研究社群，我们逐渐认识到，元研究以及软件工程研究中的人为因素仍存在不足。通过元研究，我们旨在加深对在线参与者招募和涉及人类受试者的软件工程研究的理解。本文阐述了为何有必要考虑人体研究在软件工程研究中带来的独特挑战。我们介绍了研究团队在若干不同研究项目中遇到的挑战，以及这些挑战对研究的影响，并探讨作为研究人员，我们可以如何应对这些挑战。我们展示了一项试点研究的结果，并将遇到的问题归纳为三大类：参与者招募、社区参与和数据投毒。我们还讨论了应对这些挑战的方法，并概述开展完整研究项目可能为软件工程研究社群带来的益处。

索引词——软件开发、元研究、软件工程研究、研究方法

I. 引言

作为软件工程研究社群的一员，我们需要研究软件开发人员及其在这一职业中的行为。对于职业生涯早期研究人员和研究生而言，招募软件开发人员是一项艰巨的任务，因此研究人员转向在线参与者招募。在线招募方法包括发送冷邮件或私信、在 Facebook、Twitter 或 LinkedIn 等社交媒体上发帖，以及使用 Slack、Discord、Gitter 和 Reddit 等社区社交平台。然而，无论是新手研究人员还是资深研究人员，在招募过程中都可能意外遭遇来自社交媒体用户的有害回应，这些回应似乎凭空出现，而研究人员也没有途径来纠正自己所处的困境。有害行为在社交媒体使用和领导情境中已经得到充分研究 [1]–[6]。然而，在研究情境中，人们很少展示或谈及这种现象。研究人员该如何应对自己的工作在发表前就遭到破坏？对于目标参与者所在的社区所提出的反对意见，我们该如何纠正或应对？我们应如何接触软件开发人员和社区，同时避免疏远对方或越过边界？最后，研究社区应如何报告这些挑战？

近期的元研究在理解和界定改进参与者招募策略方面取得了进展，然而，许多研究并未涉及如何应对难以相处的参与者和意外数据 [7]。我们初步开展的研究以及拟议中的研究旨在与近期有关参与者招募的研究协同推进，帮助应对参与者回应和行为所带来的挑战，并提高研究过程的透明度。我们的目标是在这项工作的初始阶段邀请研究社区提供反馈。我们希望拓展这项工作，找到应对这些挑战的成功方案，并鼓励面临类似困境的人与我们联系。通过了解研究人员在在线招募和人体研究的数据收集过程中面临的挑战，我们可以为参与者招募的方法和途径提供信息与影响，并更好地理解实证软件工程研究的方法。

II. 背景

本节讨论元研究的重要性及其对改进研究方法的影响。此外，我们还将讨论社会科学中的参与者招募。

A. 元研究的重要性

元研究“旨在评估并改进研究实践”，我们认为，审视自身的研究实践有助于确保结果有效且可信。 [8]。了解适当的研究方法是研究成功的基础，但在跨研究学科应用时，这一点往往被忽视 [8]。通过分析并借鉴其他研究学科（如社会科学和心理学）的方法，软件工程研究人员可以调整研究方法，以适用于自己的研究项目。

B. 软件工程研究中的参与者招募

历史研究表明，便利抽样是软件工程研究中最常见的抽样策略，且许多参与者来自产业界的少数几家公司 [9]。这些历史上的抽样策略引发了人们对可推广性和总体

样本量适切性的担忧。近期，许多研究人员使用付费平台或其他托管抽样服务，并发现抽样和招募合适且相关的参与者，以及筛选问题方面，仍然存在许多相同的挑战 [10], [11]。近期的其他研究旨在阐明如何应对围绕抽样、参与者筛选，以及使用学生软件工程师作为全职软件开发人员代理所产生的招募挑战 [12]–[15]。最后，需要指出一个突出的招募矛盾案例，以及研究人员与软件开发人员之间围绕招募实践的持续讨论：研究人员使用第三方网站 GHTorrent 获取开发人员的姓名和电子邮件 [16], [17]。

III. 研究设计

本节概述我们针对试点研究和后续研究的研究目标，以及完整研究的方法方案。

A. 研究目标

基于我们为研究项目招募软件开发人员的经验，我们认为，以下研究问题（RQ）可以为了解招募挑战及其对研究的影响奠定基础。我们在软件工程领域开展元研究，旨在加深对在线开发者参与者招募方法的理解，并提高人们对软件工程研究人员所面临挑战的认识；这些挑战或许并不容易被察觉。

RQ_1：软件工程研究人员在为研究项目招募人类参与者时会遇到哪些挑战？RQ_2：这些挑战对研究造成了什么影响？RQ_3：这些挑战对研究人员的个人福祉造成了什么影响？RQ_4：软件工程研究人员从遭遇的招募挑战中汲取了哪些经验？

RQ_1 的依据是了解涉及人类受试者的软件工程研究中存在的挑战的范围和程度。在此基础上，RQ_2 和 RQ_3 旨在了解这些挑战对研究人员的影响。最后，RQ_4 探讨研究人员如何减轻挑战带来的影响，以及在研究过程中遇到挑战时，哪些做法有所帮助。

B. 定性研究方法

1) 试点研究：在我们的试点研究中，作者和研究小组的前成员齐聚一堂，分享了在招募软件开发研究参与者时遇到的经历与挑战。我们总共分析了 7 项不同的研究项目，并发现这些研究项目中既有共同的挑战，也有各自独特的挑战。随后，团队将这些挑战归为 3 个大类：参与者招募、社区参与和数据投毒。表 I 列出了这些研究项目及其带来的挑战。

2) 拟议研究：基于试点研究中的研究问题，我们拟采用问卷调查和基于访谈的方法，在软件工程研究人员群体中开展元研究。我们拟议的问卷将收集有关挑战的定量和定性数据，并量化这些挑战对招募人类参与者的软件工程研究人员所开展研究的影响。对于定量问题，我们拟采用标准统计分析方法，具体方法见[20]。对于问卷调查和访谈收集的定性数据，我们拟采用开放式编码技术 [21]。

IV. 初步结果

本节概述作者在软件工程研究中招募人类受试者时所面临的经历与挑战。此外，我们还着重说明这些挑战对研究和福祉的影响。

A. 挑战

作者的经历中浮现出多种挑战，其中许多与接洽社区以征集研究参与者、社区的回应、对从小众参与者群体中过度抽样的担忧，以及参与者故意篡改其数据有关。

参与者招募——人类受试者研究离不开参与者，而作为研究人员，我们在招募过程中会遇到若干问题，尤其是在在线招募方面。研究人员在所有研究中举办面对面活动或访谈时，遇到的最大问题之一就是参与者未到场或突然断联[1]。研究 S3、S4、S5.1 和 S7 都面临参与者未到场，或在研究人员联系他们重新安排时间后突然断联的情况。参与者未按约定时间到场会妨碍研究进展，例如导致研究取消、打击未来研究项目的积极性，还可能使研究人员与准时到场的参与者陷入不愉快或尴尬的处境。在研究 5.1 中，由于参与者不足，面对面研究不得不取消；研究人员提前一周通过电子邮件联系报名者时，收到了一位参与者数小时内发来的邮件，对方对此表示强烈不满：

> “你们在为一所经常做研究的大学做研究。你却告诉我你们没钱。哈——才怪。去你妈的。”

> “我会亲自确保所有人都知道，别把时间浪费在 *[大学]* 那点小编程项目上。要是你想补偿我，我的 Venmo 账号是 @*[VenmoHandle]*。”

> “其实我是 *[竞争对手大学]* 的粉丝，而这就是我能想到的、最符合特朗普风格的‘老哥们儿’行事方式。”

> “别想着把你们那些狗屁问卷调查带到 *[城市]* 来。我们这里有海滩，可没空理会那些脆弱的人。”此外，在 S5.1 和 S5.2 中，招募具有适当经验的付费参与者参与软件开发研究也颇具挑战。参与者表示对这项研究感兴趣

[1]突然断联是指在此前有过沟通或达成约定后，毫无理由地中断与某人的联系，之后也不再跟进或作出回应。

表 I B参与者招募研究所用工具和方法的分类。* - 表示研究的现场部分，** - 表示研究的远程部分

研究编号	Ref.	招募详情	挑战
1	[18]	Reddit	参与者招募、社区参与
2	审核中	Reddit、Discord、其他社交媒体	参与者招募、社区参与、数据投毒
3	修订中	电子邮件、Twitter 私信	社区参与
4	审核中	个人联系	参与者招募
5.1*	[19]	传单、社交媒体	参与者招募
5.2**	[19]	社交媒体	参与者招募
6	保密研究	社交媒体、传单	数据投毒
7	已录用，预计于2023年发表	Reddit、LinkedIn、Twitter	参与者招募、社区参与、数据投毒

但缺乏参与所需的编程经验，最初的 89 名参与者中有 9 人（10%）最终被排除在研究之外。由于参与者要求较为小众或研究兴趣独特，每位研究者都面临过度招募问题（反复向同一批社区招募参与者），并且最终还要招募足够多的参与者，才能使研究结果具有普遍性。

社区参与——以 S1、S2 和 S6 为例，通过在线社区进行参与者招募时，会遇到一些与社区规则相关的问题，包括研究招募请求、社区版主回应以及在线平台自动垃圾信息过滤器等方面的问题。联系在线社区时，最好先了解社区规范和社区规则；有些社区对此有明确规定，而另一些则没有。我们发现，许多适合邀请参与研究的社区不允许发布此类招募请求，也不允许联系版主团队寻求许可。在社区要求联系版主团队，或规则不明确时，我们会联系版主团队寻求许可，但面临了诸多挑战——无人回应以及充满敌意的反应。在 S7 中，研究团队联系的 1000 个子版块版主团队中，有 202 个（20.2%）愿意允许发布广告或帖子。向版主介绍时，我们的团队会概述研究及其所属大学、我们联系对方的原因、为参与者提供的益处和保护措施，说明已获得机构审查委员会批准，并通常附上问卷调查链接，供其预览。大多数社区不会回应或会礼貌地拒绝，但也有一些采取敌对态度，进行言语辱骂和辱骂：

> “你就是个狗屎。”
> “你们是小偷，在这个网站上偷东西，这是事实。”
> “拿出更酷的点子再来，也许我会重新考虑。”
> “不行，而且 *[如果]* 你发帖，我就会把它删掉。”

此外，我们发现，只因联系版主收件箱，就会被社区封禁，而版主团队既不作进一步解释，也不与我们沟通。

数据投毒——虽然并非每个人都会遇到这个问题，但有几位作者在研究参与者回答问卷问题和回应访谈问题时，遭遇了“捣乱者”或破坏性行为。在 S7 中，研究人员遇到有人利用性别自我描述框发表“世界上只有两种性别”之类的观点，或填写“恐龙”这类毫无意义的回答。另一位研究人员在 S2 中遇到 ..... 到来自同意参与的参与者的敌意问卷回答，这些回答意在破坏研究，例如参与者对问题的回答包括“去***的”和“我的回答已经把这份问卷搅得一团糟，以至于我的回答对你的研究可能毫无意义。”

B. 挑战的影响

这些挑战对研究人员的影响各不相同，但所有人都表示对研究感到沮丧、缺乏动力。发生数据破坏时（如 S7 和 S2 的情况），研究人员表示不确定这些回复中的后续数据是否真实，以及是否应该纳入研究。尤其在 S2 的案例中，研究人员报告称，这些参与者提供的数据在某种程度上仍可利用，但他们质疑是否应将这些回复纳入整体分析，因为这些数据对结果的整体可推广性影响不大，而且往往属于离群值。

对于面临参与者招募挑战的研究人员来说，在相同社区中过度招募、寻找合适的参与者，以及对样本量过小的担忧，都让他们开始怀疑在线招募方法是否值得费心，以及这些渠道是否适合用于研究。此外，招募挑战带来的挫折感使他们整体上失去继续研究的动力，开始犹豫是否使用某些社交媒体平台，也损害了他们的个人福祉。

V. 讨论

研究强调有必要接触更广泛且多样化的参与者，并获得更大的代表性样本量；研究人员必须了解如何筛选、处理和/或使用可能提供难以处理的数据的参与者所提交的数据[7]。在我们的示例研究中，对于那些被证明具有挑战性的参与者响应数据，我们不希望将其完全排除，但这些数据通常除了作为离群数据之外，并未呈现出任何意义。

必须认识到，研究人员可能正是自身问题的根源；参与者或社区表现出的敌意行为，可能源于我们作为研究人员与参与者和社区互动、沟通以及随后对待他们的方式。在邀请开发者社区参与研究时，我们建议向社区联系时提供简洁易懂的研究信息，并明确说明开发者为何应参与。我们发现，在询问研究招募请求是否适合该社区，或联系个人时，先联系版主团队既有帮助，也体现了尊重

。下文提供了联系社区版主和个人时所用文本的范例：

社区版主通用模板

“嘿，[社区]版主！

我是一名来自[大学]的人为因素与软件开发研究人员，正在招募参与者，参加一项经机构审查委员会批准的研究项目，[详情见此处]。

这项研究聚焦于[补充研究详情，以及您联系该社区的原因]。通过这项研究，我希望了解[请说明研究的益处，以及这里的参与者为何应该参加]

我非常乐意回答大家可能有的任何问题，也感谢大家抽出时间！

所有数据均为匿名。您可以在此处找到问卷调查：[问卷调查链接]。”

个人模板

“嗨，[个人姓名]!

我叫[姓名]，，目前在[所属大学/公司]担任[当前职位]，由[导师/研究小组]指导（[人员或研究小组主页链接]）。我正在招募参与者，邀请他们参加一场[访谈/问卷调查]，预计需要[所需时间]，内容与[主题]相关。我特别希望了解[说明为何特别联系此人]！我们正在研究[你认为有助于你在此人的帮助下推进研究的事项]。

这次访谈将包括一场简短的问答环节，内容涉及[你计划询问内容的概述]。

感谢您抽出时间阅读本文，我很乐意回答您可能有的任何问题！”

基于观察结果提出的软件开发人员研究招募建议

- 用简洁易懂的语言介绍这项工作
- 确定研究人员及其所属机构
- 说明您为何联系此人或社群
- 说明参与的益处
- 清晰简洁地说明参与者可能受到的影响（如有）
- 说明您如何处理或匿名化数据
- 提供联系方式，以便对方提出问题或疑虑
- 提供问卷调查或访谈问题的链接，供主持人预览（如适用）

这些关于参与者招募的建议，部分用于研究 2、3 和 7；而这些建议也完全源于我们在为这些研究招募参与者时收到的参与者或社区版主的反馈。

此外，我们认为，努力扩大样本量并从多样化的参与者群体中招募，或许有助于缓解联系社区时遇到的挑战，获得在线软件开发社区更广泛的参与和支持，并使研究结果更具普适性。

Brown 的研究提出建立一个独立网站，供研究人员发布研究机会，并提供独特的参与激励措施[22]。此外，作为未来工作的一部分，我们建议扩大在现有社交媒体平台上的影响力，并创建一个供软件开发人员与研究人员就拟议和即将开展的工作进行交流互动的社区。我们希望借此促进开发人员与研究人员之间富有意义的互动，并应对当前软件工程研究招募方法中已发现的挑战。

尽管每个人面临挑战的方式各不相同，但我们认为，公开研究中面临的挑战，将鼓励新的研究人员，重新激励那些感到气馁的人，并为近期有关参与者抽样方法的工作锦上添花。通过探索参与者招募的替代方法，我们或许也能减轻这些已观察到的挑战所造成的影响。

VI. 未来计划

对于未来的工作，我们建议开展一项由两部分组成的研究，通过访谈和调查，在软件工程研究社区中继续进行荟萃分析，探究研究方法以及研究人员接触各个社区的方式；同时与软件开发者社群版主和领导者合作，了解研究人员如何接触社区并与之互动。我们希望借此了解参与者和社区为何以及如何会产生敌意回应，并探讨如果有机会，研究人员能否采取措施修复与参与者和社区的关系。

继续进行荟萃分析时，应探究参与疲劳：研究人员多频繁地联系社区？社区是否觉得自己被刷屏，或因不断被请求抽出时间而感到疲惫？与这些社区合作，了解它们收到的各类请求或许会有所裨益；这些请求既来自研究人员，也来自转向这些社区征求意见的非研究人员。了解这些社区在参与研究方面被要求做些什么，有助于研究人员与社区建立值得信赖、互惠互利的关系。然而，作为研究人员，我们需要明白，有些社区可能不愿与我们合作；它们没有义务贡献自己的参与或时间，我们必须尊重这一点。

未来的工作将对我们的参与者招募方法以及与开发者社区的关系进行自我反思，并与其他研究学科进行比较。参与者招募方法是否清晰、简洁，且与我们希望接触的开发者相关？招募邀请是否坦诚说明对参与研究的开发者的影响，并说明参与所花费时间的价值？

VII. 结论

通过分享研究团队的经验，我们着重指出软件工程研究人员面临的挑战，并展示这些挑战如何影响研究。我们相信，分享这些经验将促使软件开发研究社区携手交流在线招募和人体受试者研究方面的经验，并为推进未来研究的研究方法与策略作出贡献。通过后续研究，我们的工作将有助于重新界定招募方法，并推动采用更具普适性、也更能反映我们所研究群体的研究实践。最后，我们欢迎软件工程研究社区中曾遇到类似招募困难、并希望携手作出有意义贡献和寻找解决方案的研究人员提供反馈并开展合作。

致谢

感谢研究小组过去和现在的成员抽出时间，帮助汇总以往研究中的经验，并反思那些困难的研究。

参考文献

[1] H. Almerekhi, S. b. B. J. Jansen 和 c.-s. b. H. Kwak，《调查多个 Reddit 社区、用户和版主中的有害性》，载于《2020年万维网大会伴随论文集》，2020年，第294–298页。[2] S. Mohan、A. Guha、M. Harris、F. Popowich、A. Schuster 和 C. Priebe，《有害语言对 Reddit 社区健康的影响》，载于加拿大人工智能会议。Springer，2017年，第51–56页。[3] J. Blackburn 和 H. Kwak，《闭嘴吧，新手！预测在线游戏中关于有害行为的众包决策》，载于第23届万维网国际会议论文集，2014年，第877–888页。[4] K. L. Pelletier，《领导者有害性：对有害行为与言辞的实证调查》，《Leadership》，第6卷，第4期，第373–389页，2010年。[5] C. J. Carpenter，《Facebook 上的自恋：自我推销与反社会行为》，《Personality and Individual Differences》，第52卷，第4期，第482–486页，2012年。

[6] J. Golbeck、Z. Ashktorab、R. O. Banjo、A. Berlinger、S. Bhagwan、C. Buntain、P. Cheakalos、A. A. Geller、R. K. Gnanasekaran、R. R. Gunasekaran 等，《网络骚扰研究的大型标注语料库》，载于2017年 ACM 网络科学会议论文集，2017年，第229–233页。[7] A. Rainer 和 C. Wohlin，《为软件工程研究中的田野研究招募可信参与者》，Information and Software Technology，第151卷，文章编号107002，2022年。[8] J. P. Ioannidis、D. Fanelli、D. D. Dunne 和 S. N. Goodman，《元研究：研究方法与实践的评估和改进》，《PLoS Biology》，第13卷，第10期，文章编号e1002264，2015年。[9] A. Bouraffa 和 W. Maalej，《开发者信息需求实证研究的二十年：初步分析》，载于IEEE/ACM 第42届国际软件工程会议研讨会论文集，2020年，第71–77页。[10] F. Ebert、A. Serebrenik、C. Treude、N. Novielli 和 F. Castor，《在 Prolific 上招募经验丰富的 GitHub 贡献者进行访谈和调查》，载于实证软件工程参与者招募国际研讨会，2022年。[11] D. Russo，《在 Prolific 上招募软件工程师》，arXiv 预印本 arXiv:2203.14695，2022年。[12] M. Endres、W. Weimer 和 A. Kamil，《冒险一试：在预算有限的情况下招募软件工程参与者》，2022年。[13] A. Danilova、A. Naiakshina、S. Horstmann 和 M. Smith，《你真的会编程吗？为程序员在线调查设计并评估筛选问题》，载于2021年 IEEE/ACM 第43届国际软件工程会议（ICSE），2021年5月，第537–548页，ISSN：1558-1225。[14] M. Smith、A. Danilova 和 A. Naiakshima，《开发者研究中招募与研究设计的元研究议程》，2022年。[15] C. Brandt 和 A. Zaidman，《定性评估访谈参与者招募的策略与挑战》，2022年。[16] G. Gousios，《issue32 事件》，2016年5月。 [在线]。网址：https://speakerdeck.com/gousiosg/the-number-issue32-incident [17] ——，《issue 32 事件——最新进展》，2016年3月。 [在线]。网址：https://gousios.org/blog/Issue-thirty-two.html [18] M. Moster、A. Chandra、C. Chu、W. Liu 和 P. Rodeghero，《进入状态：远程软件开发者音乐实践分析》，载于第16届 ACM/IEEE 实证软件工程与测量国际研讨会论文集，2022年，第313–318页。[19] J. Dominic、B. Tubre、C. Ritter、J. Houser、C. Smith 和 P. Rodeghero，《虚拟现实中的远程结对编程》，载于2020年 IEEE 国际软件维护与演进会议（ICSME）。IEEE，2020年，第406–417页。[20] Z. Ali 和 S. B. Bhaskar，《研究与数据分析中的基本统计工具》，《Indian Journal of Anaesthesia》，第60卷，第9期，文章编号662，2016年。[21] J. Corbin 和 A. Strauss，《定性研究基础：发展扎根理论的技术与程序》。Sage Publications，2014年。[22] C. Brown，《助推开发者参与软件工程研究》，2022年。