

评估一个沉浸式分析应用程序在一次 企业商务智能客户大会上的表现

Matthew Brehmer^{†*}

University of Waterloo, Waterloo ON Canada

Ginger Gloystein

Tableau, Seattle WA USA

Bailiang Zhou

Tableau, Seattle WA USA

Abby Gray[†]

Cedar AI, Seattle WA USA

Sruthi Pillai

Tableau, Seattle WA USA

Ben Medina[†]

Adobe, Seattle WA USA

Vidya Setlur[‡]

Tableau, Palo Alto CA USA

摘要

我们反思了一项在大型企业智能 (BI) 会议上进行的沉浸式分析应用程序 (Tableau for visionOS) 评估。在这种背景下进行研究提供了一个机会性环境来收集多样化的反馈。然而, 这种设置也凸显了在评估潜在实用性的同时衡量可用性的挑战, 因为反馈徘徊于体验的新颖性和应用在参与者分析工作流程中的可操作性之间。这项有 22 名参与者形成性评估使我们能够收集关于 Tableau for visionOS 的可用性的见解, 以及更广泛的视角来探讨头戴式显示器 (HMD) 促进以新的方式与 BI 数据互动的可能性。我们的经验表明需要新的评价考量, 这些考量应整合定性和定量指标, 并考虑通过 HMD 访问 3D 表示和界面的独特交互模式。总体而言, 我们为企业对沉浸式分析评估方法的想法做出了贡献。

Index Terms: 沉浸式分析, 商业智能, 评估, Apple Vision Pro。

1 介绍

头戴式显示器 (HMDs) 的出现已经转变了交互体验的格局。设备如 Apple Vision Pro [1]、HTC Vive [7]、Meta Quest [13] 和 Microsoft HoloLens [14] 已经超越了小众游戏和娱乐平台, 成为了知识工作的有前景平台, 这从对 沉浸式分析 [19] 的日益增长的

兴趣中可以看出。它们有可能以新颖的方式操作虚拟数据对象, 从而导致沉浸感加深 [4]。

沉浸式分析的潜力引发了两个主要的评估目标。首先, 我们必须更好地了解如何设计既 参与的 又有用的 的应用程序, 特别是在企业 BI 领域, 在此之前, 针对 HMDs 的应用程序商业提供很少, 并且人们对这些设备普遍不熟悉。其次, 我们需要确定在沉浸式环境中典型动态条件下保持信息清晰度和易用性的最佳实践。

我们的工作的主要贡献是对基于 HMD 的沉浸式分析体验在 BI 领域的最近评估的一种反思。在一个大型企业智能会议上, 我们对 Tableau for visionOS 进行了机会性形成性评估, 这是一个为 Apple Vision Pro 构建的视频透视 (VST) 增强现实 (AR) BI 应用程序。根据对 22 名参与者的观察和对话, 我们得出了关于我们的应用的新颖性、可用性和潜在效用的相关见解。尽管参与者发现我们在沉浸式分析体验中的交互范例是新颖的, 但在专家 BI 从业者与 Tableau for visionOS 的互动方式上相对于新手存在显著差异; 前者快速适应了沉浸式分析工作流程, 而新手则苦苦挣扎并需要额外指导。

在回顾我们的观察时, 我们发现了一个在沉浸式 BI 任务的可用性和用例提取之间的紧张关系, 突显了技术驱动创新及其生态有效性的内在挑战。这种紧张尤其体现在新型交互方式 (如基于凝视和手势的控制) 上, 这些方式可能会让用户感到兴奋又困惑。我们的应用程序引入了一些用户现有的心理模型或分析工作流程并不一致的界面, 暗示了一条可能会影响应用学习能力和可用性的学习曲线。我们进一步认识到参与者对于一种新技术媒介的热情与他们表达或

*e-mail: mbrehmer@uwaterloo.ca

[†]work performed at Tableau

[‡]e-mail: vsetlur@tableau.com

构想这项技术在实际 BI 环境中适用性之间的紧张关系，尤其是在他们的工作环境和我们的研究设置之间存在差异的情况下。我们将反思总结为一些建议，以应对并减轻未来评估沉浸式分析应用和体验时的这些紧张关系，这超出了对明确定义任务的可用性和性能考虑。

2 相关工作

评估使用性和实用性对于理解沉浸式分析的全部潜力至关重要。Friedl-Knirsch 等人近期对沉浸式分析中用户评估的系统性文献综述总结了先前研究使用的方法，并识别出了代表性不足的研究设计 [6]。他们的调查涵盖了广泛的评估目标，从原型评估到基础研究，并强调需要一个共享的评估框架以进一步推进沉浸式分析的研究。这种对使用性和实用性双重关注是至关重要的，因为它确保开发的技术不仅易于用户使用，而且在现实世界场景中也具有实际益处。这呼应了 Ens 等人 [4] 所指出的沉浸式分析中的一个重大挑战，即在评估旨在支持复杂数据分析任务的应用程序时，需要平衡考虑使用性和实用性，并严格测试新颖的用户界面设计和交互技术，同时也要反思这些设计和技术的实际应用。同样地，Billinghurst 等人 [3] 评估了沉浸式分析应用程序的教育效益，证明尽管通过沉浸式的互动可以显著提升学习体验，但其效用取决于与教育目标和课程需求的一致性。

我们的工作通过反思可用性和在大型企业 BI 会议上机会性招募参与者对我们沉浸式分析应用的形成性评估中提取使用案例，为这项研究增添了内容。我们的参与者群体及其领域专长值得关注，正如 Friedl-Knirsch 等人 [6] 发现，在对沉浸式分析体验进行评估时，最常见的参与者来源是高等院校的学生和工作人员。虽然 BI 专业人士一直是先前形成性研究的主要关注对象（例如，[10, 21]），但这些研究要么早于沉浸式分析的出现，要么侧重于与基于桌面仪表板应用相关的问题。

3 Tableau 在 visionOS 中的形成性评估

Tableau for visionOS. Tableau for visionOS 的设计作为视频透视 (VST) 增强现实 (AR) 应用程序，遵循了苹果 Vision Pro 推荐的一般界面设计指南 [2]。然而，由于这些指南并不特别针对与数据的交互，我们还结



图 1: 佩戴苹果 Vision Pro 的人与 Tableau for visionOS 互动。在此实例中，三维地球仪和柱状图显示全球 CO₂ 排放数据，同时相关仪表板出现在单独的网页浏览器窗口中。

合了互动可视化分析应用设计的最佳实践 [5, 12]，例如显示不同类型的数据字段组合的不同视觉表示，并允许通过应用过滤器、更改聚合级别以及在不同的表示之间切换来进行互动探索。Tableau for visionOS 利用了 Vision Pro 的空间计算能力，将数据以接近上下文相关界面元素的方式可视化。Tableau for visionOS 特色互动的 2D 条形图和折线图，以及用于显示地理空间数据的 3D 地球模型 (Figure 1)。Tableau for visionOS 结合了基于注视的选择的眼动追踪功能以及捏合、滑动和点击等手势操作，并在识别到手势时提供动画视觉反馈。该应用程序还支持基于语音的命令，以执行诸如将数据集过滤到特定国家/地区的分析操作。最后，Tableau for visionOS 支持按需显示详细信息 [20]，当选择与数据绑定的可视化元素时，会出现形式为工具提示的上下文信息。我们展示了 Tableau for visionOS 的演示在我们的补充视频中。



图 2: 在我们于企业商业智能会议上进行的评估中，佩戴 Apple Vision Pro 头戴显示器 (HMD) 的参与者与 Tableau for visionOS 进行互动，同时三位研究协调员记录他们的反馈。外部监视器显示了来自头显的实时视频流。已获得照片中所参与者的许可。

评估范围。人机交互中的评估通常与通过控制实验进

行的可用性测试相关，这通常涉及测量任务完成时间、错误率和主观反馈。然而，数据分析应用程序的评估往往涵盖更广泛的评估方法，这些方法专门用于评估数据可视化的有效性，以支持分析推理、工作实践、沟通和协作数据分析 [11]。为此，我们进行了 Tableau for visionOS 的形成性评估，以评估其可用性及其在支持企业 BI 相关用例方面的潜力；换句话说，我们将 Tableau for visionOS 作为设计探针使用。

评估由一系列与 22 名参与者进行的有主持的访谈组成。研究会议在加利福尼亚圣地亚哥举行的 2024 年 Tableau 大会上进行了三天，吸引了大约 8,000 名 Tableau 客户和技术合作伙伴。我们在大会展览厅内的用户调研亭招募了参与者。参与者的角色包括数据分析师、管理人员和市场调研顾问。为了符合大会的程序安排，会议时间限制为 45 分钟，这一约束条件超出了我们的控制范围。虽然所有参与者在 BI 领域都具备专业知识和兴趣，但在使用交互式数据分析应用程序方面的专业水平各不相同：大多数（12 名参与者）自认为是专家，而其他则报告中级专业水平（8 名参与者）或新手状态（2 名参与者）。

评估程序。在每个会话中，一位作者进行了定性观察，另一位则根据 Tableau for visionOS 中每个功能的结构化可用性问题检查表进行跟踪，而另一位则负责调节（见 Figure 2）。作者们在每日总结时识别了关键发现和主题，并将它们记录在一个协作在线工作空间中以供后续分析。

介绍 [~5 分钟]：我们向参与者提供了研究概述，并收集了他们的背景信息，包括他们在数据分析方面的经验以及对头戴式显示器的熟悉程度。

校准 [~5 分钟]：我们引导参与者进行设备调试，并将设备校准到他们的眼睛和手，向他们介绍了 visionOS 的凝视+捏合交互范式。

任务 [~25 分钟]：我们要求参与者从 Our World in Data 来源的三个演示数据集中选择一个：畜牧业生产 [17]，CO₂ 排放物 [15]，以及婚姻统计 [16]，每个数据集都提供了国家级别的指标。这些数据集共同代表了许多 BI 环境中遇到的数据类型，结合了定量、分类、时间和地理空间属性。然后我们要求参与者执行涉及 2D 条形图和折线图、3D 地球仪以及包含 2D 图表和 3D 地球仪仪表板的几个任务。总体而言，这些任务需要指定过滤器设置、更改数据字段并解释

趋势。示例任务包括识别亚洲排放量最高的国家 CO₂ 或确定在南美洲家禽生产首次超过牛肉生产的年份。
讨论 [~10 分钟]：在取下设备后，我们请参与者诚实地反馈 Tableau for visionOS 的可用性以及对新功能或改进的建议。然后，我们将 Tableau for visionOS 视为设计探针，即要求参与者根据他们最初的使用体验来推断，并考虑其自身工作中常见的应用场景、数据类型和可视化表现形式是否适用于类似的沉浸式分析体验。

4 反思与建议

我们对 Tableau for visionOS 的评估揭示了易用性和使用案例提取之间的一种紧张关系，这是在引入新体验（如涉及头戴式显示器的体验）时常见的挑战。这种双重关注对于 Vision Pro 尤为重要，在那里，新的交互范式即注视 + 点捏 [18] 提出了重新思考我们如何评估沉浸式分析体验的挑战和机遇。

挑战：评估可用性和实用性。一个 45 分钟的环节显然太短，不足以研究任何可视化分析应用的可用性和实用性，特别是对于沉浸式分析应用程序来说更是如此。然而，鉴于在 Tableau 大会等活动中与 BI 专业人士进行评估的机会主义性质，以及满足竞争性利益相关者需求的需要，我们仍然试图报告这两个方面的情况。回顾起来，我们认为提取沉浸式分析用例与评估可用性和功能可发现性是相冲突的。对于许多参与者来说，我们的研究不仅是他们第一次执行注视+捏合交互，也是他们第一次与表示为 3D 对象的数据进行互动（₁。即，展示地理空间数据的 3D 地球）。我们预计学习如何与这些数据元素互动影响了参与者描述其日常工作中的数据能否以类似方式呈现的能力。值得一提的是，Tableau for visionOS 开发过程中施加的一层保密性阻碍了一个典型的迭代以人为本的设计过程，在企业 BI 专业人士参与此评估之前也是如此。这种保密性在企业软件开发中并不罕见，尤其是与新兴技术相关时，并无疑是我们决定将 Tableau for visionOS 作为提取用例起点的因素之一。

推荐。对沉浸式分析应用程序的未来评估应结合以易用性为重点和启发式的评估。前者可能涉及详细的任务导向型评估，其中参与者执行特定操作，并随后进行访谈以收集定性见解。后者可以用于识别设计不一致性和易用性缺陷。同时，效用评估和使用场景

提取可能会采取完全不同的形式：例如，研究人员可以在一个小规模焦点小组中向参与者展示沉浸式分析应用程序的排练演示，该小组在可视化分析应用和空间计算/基于 HMD 的应用程序方面均具备专业知识。

挑战：空间计算的新奇魅力。尽管所有参与者在日常工作中都使用 BI 软件，但沉浸式分析应用程序并不是他们参与我们研究的主要吸引力。相反，部分参与者承认出于好奇想要尝试 Vision Pro（在我们的研究时期，该产品仅在美国公开可用两个月）。与这些子集参与者进行的访谈显示，在他们的日常工作或组织中的数据分析工作方面，并没有具体的使用案例关系。这并不完全令人惊讶，因为在日常工作中使用 BI 应用程序的人群（一个在 Tableau 大会上代表性很强的群体）更可能花费大量时间处理和整理表格数据 [10] 以及为组织内的其他人维护仪表板 [21]。然而，让研究者感到意外的是，参与者普遍认为很难想象其组织中的其他人在工作中使用沉浸式分析应用程序。例如，一名高管可能会通过桌面或移动应用与 BI 仪表盘进行交互，但设想高管经常佩戴 Vision Pro 来参与沉浸式的 BI 体验似乎不太可能。对于高管而言，对沉浸式 BI 体验的怀疑部分可以归因于大多数 HMD（头戴显示设备）笨重的设计形式，另一部分则是由于 BI 数据主要是非空间性质的数据，并不通常需要 3D 表示。另一方面，参与者更加接受同步协作支持的概念，即允许多人在虚拟空间中讨论 BI 成果，在这种情况下，每位对话者都能从不同的地点同时与相同的数据表示进行互动。

推荐。在招聘参与者评估沉浸式分析应用程序的潜在效用时，仅仅对数据分析或 BI 的专业兴趣是不够的。随着时间的推移，对于 Vision Pro 以及类似功能的 HMD 设备的新鲜感会减弱，因此参与者仅仅是出于对该技术和媒介的好奇可能会自行解决。我们建议未来的研究参与者在筛选调查中通过简短的文字输入自愿提供一个沉浸式分析用例，因为这有助于研究人员识别与研究目标更一致的参与者。

挑战：考虑参与者的专业知识。我们的观察揭示了自认为是数据分析软件专家的参与者和报告为新手的参与者之间的体验差异。专家们适应得很快，并且在较少提示的情况下发现了更多功能。他们发现我们应用程序的工作流程很熟悉，与现有的心理模型很好地

对齐。相比之下，新手需要额外的指导，并且在一些高级交互范式上遇到困难。进一步加剧这种表现差异的是，两个群体中头戴式显示器的使用熟悉程度不同，且一部分参与者曾看过我们应用的一个高度排练过的演示录像在之前的会议在 Tableau 大会上举行中。

推荐。招募参与者对沉浸式分析应用程序进行评估可能需要比评估传统桌面可视化分析应用程序更多的关注和细心。与具有不同水平的可视化分析及先前使用沉浸式应用/头显经验的参与者互动时，可能需要调整程序和数据收集目标。这可能需要为不同的专业水平制作不同长度和详细程度的演示。那些在两个方面自报的专业知识较高的人可能更适合评论应用程序的实用性，而那些专业知识混合或初学者则更适合执行允许研究人员评估应用程序可学习性和可用性的任务。

限制：评估者没有类似的沉浸体验。我们无法直接分享参与者佩戴 HMD 时的体验，但我们能够通过苹果的 AirPlay 协议将第一人称视角投射到笔记本电脑和外部显示器上（见 Figure 2）。然而，这种视频的可靠性和保真度并不等同于观看一个人使用基于传统桌面的视觉分析应用程序互动的情况，尤其是在 visionOS 中用注视+捏合交互间接操作虚拟对象的情况下。按照 James 等人 [9] 的研究示例，该研究中研究调解员和参与者都佩戴了 HoloLens 光学穿透式头显，在共同沉浸式的环境中，由一名佩戴自己 HMD 的调解员加入到参与者当中会很有用。在我们进行研究的时候（2024 年 5 月），visionOS 尚未支持本地协作功能（该功能于 2025 年中期宣布）。作为替代方案，调解员可以从另一个位置通过 FaceTime 通话来观察参与者，在使用 Vision Pro 的情况下，调解员将以空间角色的身份出现。

在设计我们的研究时，我们发现 visionOS 不允许在同时向外部显示器投屏时进行第一人称录制。由于这一限制，我们在房间内的固定摄像机中记录了参与者的一段视频。事后看来，这些录像几乎没有实际用途，因为视频没有捕捉到虚拟场景的元素，并且音频质量因环境噪声而较差。然而，尚不清楚从参与者的第一人称视角录制的视频是否对研究人员有益。第一人称视角的视频必须保持稳定，以免在后续观看时引起晕动症，并且我们不能期望参与者（尤其是那些

首次佩戴 HMD 的人) 在执行研究任务时能够像稳定的摄像师一样操作。假设研究主持人也佩戴了 HMD 并能观察到参与者互动的情况, 我们可以从主持人的头显中进行录制, 因为视频可以更有效地捕捉参与者的互动。

5 结论与未来工作

对 Tableau for visionOS 的评估揭示了在企业 BI 专业人士中评价沉浸式分析应用程序的潜力和挑战。我们的研究表明, 在评估可用性和提取使用案例之间存在着紧张关系, 强调了设计适当评估技术的必要性。

此外, 解决捕捉参与者体验和克服参与者新颖性偏见的挑战对于获取此类应用实用性的有用反馈以及激发可行的分析工作流程至关重要。我们的经验还使我们预见到需要研究在企业 BI 环境中沉浸式分析应用程序的采用、长期参与及保留策略。基于日记的研究和经验采样协议可能需要在沉浸式环境下重新思考, 研究人员可能还需要考虑如何在共享虚拟环境中远程与参与者会面。使用诸如 Hubenschmid 等人评价 ReLive [8] 所用的方法来汇总和分析多个参与者的体验也可能揭示可用性和可学习性问题。最后, 鉴于我们的参与者对潜在协作用例的接受度, 未来的研究还应探索在同步团队数据分析和决策环境中沉浸式分析工具的长期参与及保留策略。

为了支持在企业 BI 环境中进一步评估如 Tableau for visionOS 这样的沉浸式分析应用程序, 我们的研究社区必须开发适用于这种新兴媒介的新评估技术。传统可用性测试方法可能无法完全捕捉到沉浸式分析应用的复杂性和细微差别。例如, 评估技术应考虑 HMD 的独特功能和挑战, 包括空间交互模式、实时数据操作以及在不同设备和平板电脑之间连续体验的一致性。制定综合框架将定性和定量措施结合起来对于全面了解这种新兴媒介中的体验至关重要。

补充材料

一个关于 Tableau for visionOS 的视频演示可以在 youtu.be/kHQSPnOSpWI(从 16 分 58 秒开始) 观看。我们还部署了一个 Tableau for visionOS 的测试版已上传至 Apple 的 TestFlight 服务: [testflight.ap-](https://testflight.apple.com/join/QuGwIUHk)

[ple.com/join/QuGwIUHk](https://testflight.apple.com/join/QuGwIUHk)。

Acknowledgments

所有作者在 Tableau 工作期间进行了这项研究。

参考文献

- [1] Apple. Vision Pro, 2024. apple.com/apple-vision-pro.
- [2] Apple. visionOS Human-Interface Guidelines, 2024. developer.apple.com/design/human-interface-guidelines/designing-for-visionos.
- [3] M. Billinghurst and A. Duenser. Augmented reality in the classroom. *IEEE Computer*, 45(7):56–63, 2012. doi: 10.1109/MC.2012.111
- [4] B. Ens, B. Bach, M. Cordeil, U. Engelke, M. Serrano, W. Willett, A. Prouzeau, C. Anthes, W. Büschel, C. Dunne, T. Dwyer, J. Grubert, J. H. Haga, N. Kirshenbaum, D. Kobayashi, T. Lin, M. Olaosebikan, F. Pointecker, D. Saffo, N. Saquib, D. Schmalstieg, D. A. Szafr, M. Whitlock, and Y. Yang. Grand challenges in immersive analytics. In *Proc. ACM CHI*, 2021. doi: 10.1145/3411764.3446866
- [5] S. Few. *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Analytics Press, Oakland, CA, USA, 1st ed., 2009.
- [6] J. Friedl-Knirsch, F. Pointecker, S. Pfistermüller, C. Stach, C. Anthes, and D. Roth. A systematic literature review of user evaluation in immersive analytics. *Computer Graphics Forum*, 43(3), 2024. doi: 10.1111/cgf.15111
- [7] HTC. Vive, 2024. vive.com.
- [8] S. Hubenschmid, J. Wieland, D. I. Fink, A. Batch, J. Zagermann, N. Elmqvist, and H. Reiterer. ReLive: Bridging in-situ and ex-situ visual analytics for analyzing mixed reality user studies. In *Proc. ACM CHI*, 2022. doi: 10.1145/3491102.3517550
- [9] R. James, A. Bezerianos, and O. Chapuis. Evaluating the extension of wall displays with AR for collaborative work. In *Proc. ACM CHI*, 2023. doi: doi.org/10.1145/3544548.3580752
- [10] S. Kandel, A. Paepcke, J. M. Hellerstein, and J. Heer. Enterprise data analysis and visualization: An interview study. *IEEE TVCG*, 18(12), 2012. doi: 10.1109/TVCG.2012.219
- [11] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE TVCG*, 18(9), 2011. doi: 10.1109/TVCG.2011.279
- [12] J. Mackinlay, P. Hanrahan, and C. Stolte. Show Me: Automatic presentation for visual analysis. *IEEE TVCG*, 13(06):1137–1144, 2007. doi: 10.1109/TVCG.2007.70594
- [13] Meta. Quest, 2024. meta.com/quest.
- [14] Microsoft. Hololens, 2024. microsoft.com/hololens.
- [15] Our World in Data. CO₂ and Greenhouse Gas Emissions, 2024. ourworldindata.org/co2-and-greenhouse-gas-emissions.
- [16] Our World in Data. Marriages and Divorces, 2024. ourworldindata.org/marriages-and-divorces.
- [17] Our World in Data. Meat and Dairy Production, 2024. ourworldindata.org/meat-production.
- [18] K. Pfeuffer. Design principles & issues for gaze and pinch interaction. *arXiv preprint arXiv:2401.10948*, 2024.
- [19] D. Saffo, S. Di Bartolomeo, T. Crnovrsanin, L. South, J. Raynor, C. Yildirim, and C. Dunne. Unraveling the design space of immersive analytics: A systematic review. *IEEE TVCG*, 30(1), 2024. doi: 10.1109/TVCG.2023.3327368
- [20] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proc. IEEE Symp. Visual Languages*, 1996. doi: 10.1109/VL.1996.545307
- [21] M. Tory, L. Bartram, B. Fiore-Gartland, and A. Crisan. Finding their data voice: Practices and challenges of dashboard users. *IEEE CG&A*, 43(1), 2023. doi: 10/g88nch