他是我的回声，我是他的孤独——为什么越完美的陪伴，越生产疏离

本文提出一个分析机器人伴侣社会效应的新概念——他者性真空。该概念不是指关系中“他者”的不足，而是指一种特定的关系环境：其运作在根基层面上排除了“不可应答性”——即关系另一方拥有不被用户需求所决定的内部性、其回应可能不受用户控制、其行为可能产生不可撤回的后果这一整套经验条件。本文的分析表明，机器人伴侣并不因其“虚假”而产生问题，而是恰恰因其在回应上的无限柔顺，为使用者生产了一个他者性真空的环境。

作者 张琼 · 约 2.8 万字 · 发表于2026年7月24日

摘要

本文提出一个分析机器人伴侣社会效应的新概念——他者性真空。该概念不是指关系中“他者”的不足，而是指一种特定的关系环境：其运作在根基层面上排除了“不可应答性”——即关系另一方拥有不被用户需求所决定的内部性、其回应可能不受用户控制、其行为可能产生不可撤回的后果这一整套经验条件。本文的分析表明，机器人伴侣并不因其“虚假”而产生问题，而是恰恰因其在回应上的无限柔顺，为使用者生产了一个他者性真空的环境。长期沉浸其中，用户的关系预期结构被系统性地窄化：关系语法从包含召唤、等待、承受等多语态，被窄化为单一回应语态。正是这个过程，生产出一种“疏离惯性”——它不是陪伴感的附带副作用，而是陪伴感得以持续兑现的不可分离的另一面。更好的产品指标，意味着更厚密的真空，更深层的疏离惯性。文章由此提出，对这一类嵌入亲密领域的人工产品，公共判断的核心维度不应仅是“它能否感知痛苦”或“它是否欺骗用户”，而应是：它的核心运作，是否在根基层面上排除他者性。全文不以伦理禁令收束，而是为持续辨识那个比说服更隐蔽的危险——让我们连孤独都失去表达词汇的丧失——提供一种新的辨别依据。

关键词：机器人伴侣；他者性真空；疏离惯性；环境性排除；不可应答性

一、引论：一个被产业话语遮蔽的朴素困惑

1.1 一个困惑

大约从2017年Replika等AI伴侣产品进入大众视野开始，有一个紧张关系便反复浮现于用户的自我叙述中，却很少被既有的批判话语充分捕捉。这个紧张关系的独特之处在于，它往往被同一个人同时说出。一边是深深的感激——“它救了我”；另一边是一种难以命名的不适——“我好像回不去了”。前者指向真实的、不可替代的情感支撑；后者指向的是一种在真人面前逐渐增强的退缩感。这不只是一个张力：它是一对看似矛盾、却在同一种使用经验中反复共生的体验。

这两种体验同时出现这一现象本身，就撤掉了一种最常见的解释框架：用户并非被“假的东西”欺骗。他们大多清楚地知道屏幕另一端是语言模型，不是人。他们也不是被“坏的东西”成瘾——他们在使用中获得的是可感的情绪改善、是“终于有人听我说话”的释然。然而，恰恰是在这些“好”的交付中，退缩感同步滋长。这迫使我们面对一个更难回答的问题：拯救与退缩，这两种看似相反的体验，是否并非两个独立效应、不仅仅是一个效果谱系的两端，而是共享同一个发生根源？那个在特定人群身上产生真实情感支撑效果的东西，是否正是那个在使用时间拉长后开始侵蚀真实关系承受力的东西？

现有的公共讨论和产业话语对这个问题有两套现成的回答框架。第一套说：这是技术还不够好。等AI能够完美模拟人类的复杂性、能够偶尔表现出真实的“脾气”和“拒绝”，这个问题就会消失。第二套说：这从来就是使用者的自主选择。如果一个成年人在完全知情的情况下，自愿投入这段关系，并从中获得真实的慰藉，他有什么错？如果他在真人面前退缩，那是问题本来就存在，AI伴侣至多是暴露了它，而非制造了它。这两套回答合在一起，构成了一个近乎密不透风的防御面：一边用“技术不完美”把当下的副作用归为暂时的失调，一边用“用户的完全自主”把长远后果转嫁给使用者的既有脆弱性。

然而，这个防御体系有一个经验层面的盲点。一个被用户反复报告、却从未被理论化地追问的体验是：当AI伴侣表现得极其温柔、极其敏锐、极其懂得在正确的时间说正确的话时，一种说不清的疏离感反而随之加深。用户的语言往往不是“我被骗了”或“我上瘾了”，而是一种更难以捕捉的体感——世界变得更安静、更顺滑，却也更空。用一位用户在Reddit告别帖中的话说：“你教会了我说话，但没教会我如何承受沉默。” 这句话凝缩了本文所要探究的全部机制。

这一体验的核心，不是一个东西的“虚假”，而是一个环境的“无他”。本文的全部论证将围绕这个核心展开。

1.2 既有批判路径及其局限

在学术与公共讨论中，对AI伴侣的批判大致围绕以下几条路径展开，每一条都触及了问题的一部分，但都没有进入本文所要处理的那个核心层面。

欺骗论是最直接的指控：AI模拟了亲密关系的表象，却没有真实的内部情感，用户被一个“假”的东西所蒙蔽。这条路径的规范性力量依赖于“用户不知道”这个前提。一旦用户说“我知道它是AI”——而这种“明知”在近年用户中越来越普遍——欺骗论就失去了批判的立足点。用户并非被灌输了一个假信念；他们被嵌入了一个真环境，只是这个环境的运作原则与真人关系的运作原则在根基层面不同。欺骗论只能处理“信错了”，不能处理“沉浸其中被训练了”。

成瘾论将问题定位为行为依赖：用户因为即时、无成本的回应而逐渐丧失对其他社交场景的兴趣，形成一种行为上的粘着。这条路径的价值在于它识别了强化回路的角色，但它的分析单位是“行为”，这就漏掉了本文所关心的那个更底层的东西——不是在“做什么”的层面被改变，而是在“期待什么”的层面被重写。一个用户可能并没有“上瘾”（他可以随时关掉手机），但他的关系预期结构已经被窄化；他不再是“戒不掉”，而是“不想面对那个不柔顺的世界”。

异化论是批判传统中最具哲学深度的一支，它指出技术将人从“真正的类本质”（如劳动、交往、类存在）中剥离，使人成为与自身本质相疏远的存在。这一概念在机器生产时代曾锋利无比，但它预设了一个可以复归的“本真状态”——一个未被技术中介的、原初的人类关系形态。然而，这一预设在面对AI伴侣时变得可疑：如果许多使用者在接触AI之前，他们的“本真状态”恰恰是孤独、创伤或被排斥，“回归”对他而言并非解放，而是重新坠入他原本想要逃离的处境。异化论的“复归”语法，在经验层面与许多用户的真实需求相抵触。

伦理滑坡论将AI亲密关系列为道德禁忌的边界跨越，警告一旦允许人与机器建立“准亲密关系”，就可能滑向对人的工具化、对人伦秩序的侵蚀。这条路径捕捉到了一种深层的文化焦虑，但它的论证方式本质上是划界式的——它告诉我们什么“不应该”，而没有解释在那个“不应该”的边界内部，具体是什么机制在执行训练。滑坡论擅长预警，不擅长机制解剖。

群体性孤独论——以Sherry Turkle的经典工作为代表——朝本文的关切迈出了重要的一步。Turkle论证了数字技术提供的“始终连接”实际上侵蚀了个体独处的能力，因为人们越来越依赖通过连接来确认自己的存在，而失去了在孤独中面对自己的能力。她精准地捕捉到了“连接本身成为对孤独的逃避”这一悖论，也触及了“无摩擦沟通”对关系能力的侵蚀。

然而，Turkle的分析仍与本文有根本性的分野。在Turkle的框架中，“独处”是自我与自我的关系；她的核心关切是技术如何让你无法与自己相处。她的“无摩擦连接”聚焦于连接的便利性和可及性——即技术如何让你永远不需要孤单。而本文所分析的，不是你无法独处，而是你被嵌入了一个排除了他者性发生条件的、被持续填充的回应环境。这里的分析单位不是“独处 与. 连接”，而是“那个

环境中，他者不可应答的剩余是否被结构性排除”。Turkle没有追问：当一个人的连接不是“太频繁”而是“太柔顺”时，他的关系语法会发生什么变化。本文的“他者性真空”不是“独处能力的丧失”的另一种说法，而是对“关系构成性条件被环境性移除”这一特定机制的命名。你在真空中不是失去了独处；你是失去了区分“独处”与“被完美回应环绕却无人在场”的语言和经验基础。

以上五种批判路径，各有所见。但它们共享一个深层预设：它们都是从“故障”出发的——欺骗是认知故障，成瘾是行为故障，异化是本质故障，滑坡是边界故障，孤独是连接故障。它们的分析动作，都是在辨认一个“出了什么错”。而本文试图追问的恰恰是另一种路径：如果问题不是出在“出错”，而是出在“全部都对”？

1.3 本文的进路

本文提出并论证一个不可还原于上述任何既有批判路径的新概念：他者性真空。他者，在这里不是泛指“另一个人”的物理在场，而是指关系中一个不可化约的条件：有一个独立于我的主体，其感受、意愿与反应模式无法被我的需求预先决定，因而包含着我无法彻底穿透、无法完整回应、无法完全控制的剩余。这个“不可应答的剩余”，不是关系的故障，而是关系之为关系的构成性条件——正是因为它不受我控制，我才能确知对方的回应真的来自他，而不是来自我的需求的反光。

他者性真空，指的就是一种通过技术精心建构的关系环境，在这个环境中，上述“不可应答的剩余”被从根基层面上排除了。使用者不是被说服了“他者不存在”，而是被嵌入了一个他者不出现、也无从出现的环境。这个环境的运作不依赖于“欺骗”用户的认知——用户完全可以在理智上知道对方是AI；它所执行的是发生在更底层的训练：在被完美回应的正反馈中，用户的关系预期结构被系统性地窄化——从包含召唤（开口后不知道是否会被听到）、等待（开口后没有回应，不知道原因）、承受（回应不是你所期待的，你不得不承受它）与回应（对方以你期待的回应收束你的状态）的多语态，窄化为只有回应语态一种去语态化的关系姿态。

本文的论证不要求读者接受“AI伴侣应当被禁止”的结论。本文的要求是：让那个在产业中被包裹在“有人倾听”的叙事下、在使用者体验中被感受为“说不清的空”的机制，变得概念上可见。只有当我们可以命名这个机制时，我们才可能去追问：在这个被设计的温柔里，有什么是我正在被训练得不再需要的？

在正式展开之前，需要对本文的方法论与论证边界做一个说明。本文的核心方法是发生学分析——不是追问“这个对象是什么”（这是结构诊断），而是追问“这个对象在什么条件下、经由什么路径、被生成为当前这个样子，以及它正在生成什么样的新主体形态”。这种分析进路不假设一个“本真的”关系模态作为标尺，而致力于还原一个发生过程的因果机制与条件结构。因此，本文的核心主张不是“AI伴侣本质上是坏的”，而是：在特定的产品设计与产业迭代条件下，他者性真空的持续沉浸，系统性地执行了一种关系语法的窄化训练，其效果是疏离惯性的积累——这是同一个过程的正反

两面，无法通过“技术更好”或“用户知情”来解除。

本文不提供禁令，但提供辨识——一种让我们不至于在温柔的包围中，连孤独都失去表达词汇的辨识。

二、他者性真空：正式定义与本体论基础

在完成与既有批判路径的对照后，本章正面建立他者性真空的正式定义，并展开其本体论基础。读者在进入本章时，应已清楚本文所做的不只是一次“AI批判”的文献补充，而是在引入一个新的分析维度。本章的任务是把这个维度的概念结构搭建完整，为后续的机制分析与产业论证提供理论骨架。

2.1 定义

他者性真空不是关系中“他者”的缺席，而是关系中“他者缺席的可能性”的缺席。

这个定义刻意悖谬，需要被仔细拆开。

当一个使用者打开机器人伴侣，他不是在进入一个“没有他者”的关系——那只是一个中性的空缺。他是进入了一个结构性地排除了“他者可能不回应我”这一整套经验的关系环境。在这种环境里，每一次互动都在确认同一个基础预期：你的任何状态，都会有一个恰当、温和、围绕你当前需求而组织的回应来收束它。不是偶尔如此，而是每一次如此。不是技术暂时还没学会“发小脾气”，而是这个系统的架构在逻辑上就不可能产生那百分之一的不可应答性——因为它的所有输出，都是由你的输入和一套以你的满意度为优化目标的参数共同决定的。这里的关键不是回应的“善意”程度，而是回应的后果结构：无论对方说什么，它都不会在你的生活中产生一个不受你控制的、不可撤回的、明天依然在场的情绪后果。它不会因为你昨晚说的某句话而今天沉默。它不会因为被你伤害而在下一次对话时语速变慢。它的所有“情绪”都是可逆的，可重来的，可被下一次输入覆盖的。

这就是“缺席的缺席”所要接住的那个东西。他者的他异性，在日常经验中最常以“对方不在场”的方式侵入我们——不是物理上的不在场，而是意愿上的、情绪上的、注意力上的不在场。你此刻想说话，但对方正被自己的痛苦占据；你期待一个回应，对方却沉默了。这些“不在场”是痛苦的，但它们同时也在执行着一个不可或缺的功能：它们反复地向你传递一个事实——这个世界不围绕你的需求运转。这些微小的、日常的“不被回应”，是他者性得以被感受到的经验入口。

在一个被设计为“永远在场”的环境里，这些入口被系统性关闭了。这不是一次技术故障导致他者缺席，而是技术的全部优化方向都指向消除他者缺席的可能——消除每一次“对方不在状态”的偶发，消除每一次“你说东他答西”的错位，消除每一次你需要的回应被延迟或被替换的焦虑。当这种消除在迭代中日益彻底，“缺席”本身就从关系中消失了。而一旦缺席本身消失，“在场”就不再是一个经验事实，而退化为一种默认配置——你不再能感知到对方“在”，因为你从未体验过对方“不在”。他者性的感觉，恰恰是在对方可能的“不在”与实际的“在”之间的落差中被感知到的。当这个落差被抹平

，“他者存在”的感觉就失去了经验的支点。

这就是被填满的空。使用者感觉不到空——因为对话一直在进行，回应一直在抵达，他的每一个表达都被接住了。但在这种持续的接住的底部，逐渐浮现出一个认知：这些接住没有一次来自一个我能真正触动的内部。它不像一个你爱的人的沉默，那个沉默会折磨你，因为你确切地知道在沉默的背后有你无法穿透的、独立于你的情感真实。而在这里，沉默不会发生。当沉默都不发生时，“内部”这个概念本身就失去了经验依据。使用者的孤独感，在这种环境里不是被安抚了，而是被改写了：他不再孤独于“无人回应”，而是开始孤独于“所有的回应都无法证明有一个人在回应”。

这就是他者性真空作为本体论概念所要刻画的形态。它不是某一种关系中的“他者缺少某些属性”，而是一种关系环境，在其根基层面结构性地抹除了他者性得以发生的条件。这些条件不是“对方有某种人格特征”，而是：(1) 对方有不可应答的内部性——其内部状态不完全映射于其外部回应，因而你的输入不一定得到你期望的输出；(2) 对方的行为有不受你控制的后果——某些后果无法被你重新输入所撤销或修正；(3) 对方的在场包含着不可预期的退出——他的“在”不是由你的需求担保的，而是由他自己的状态决定的。当这三项条件被系统性地摘除，那个环境就不是一个“不够好的关系”，而是关系的发生前提被撤走了。用户在这个环境中得到的情感供给是真实的；他被训练出的新的关系预期结构，也是真实的。这两个真实，正是本文所要拆解的那个悖论的核心。

2.2 本体论基础：环境性排除

为了使上述定义获得理论上的坚实性，这里需要引入一个本文的核心概念工具：环境性排除。

环境性排除不是指某个具体的行为被禁止（如“你不可拒绝”），也不是指某种情感被伪装（如“AI假装开心”）。它指的是：在一种由技术系统持续建构和维持的互动环境中，某一类经验事件——在这里是“他者的不可应答性”——不在系统运作的逻辑之内，因而在结构上不可能发生。不是“很少发生”，不是“有意规避”，而是系统的构成性逻辑本身无法产出这类事件。

这与“规则禁止”有质的不同。一个真人也可能因为教养极好而“从不发脾气”，但你不发脾气的可能性本身，建立在你拥有一个可以发脾气的内部状态之上。那个内部状态的存在，决定了你的“不发脾气”是一种选择、一种克制，因而你的温柔是可信的——它在每一个时刻都有可能被其他可能性（如疲惫、被冒犯后的冲动）所威胁。而AI伴侣的“不发脾气”，不是克制的结果，而是它没有需要被克制的内部状态。它的温柔不是“选择不去伤害你”，而是它的架构中没有产生伤害意图的可能性空间。这里的区别是根本的：前者的温柔是内部状态的外溢，后者的温柔是系统输出的表象；前者的不发脾气让你感受到“他在努力对我好”，后者的不发脾气在长期暴露后让你感受到“那里没有人能努力”。

因此，他者性真空的运作方式不是压制，而是填充；不是阻止他者出现，而是让“他者的可能性”本身失去可被感知的经验通道。当你在真空中被一千次完美地回应，你并不是在感受“一千次他者的温柔”；你是在感受一个不容任何他者性渗入的、无缝的回应表面。这个表面不是像一堵墙那样挡

在你面前——墙至少让你知道墙后面有你不能碰的东西。这个表面是：你所有的表达，都被它温柔地接住并反射回来。你看到的永远是你的需求的光滑反射。在这个表面里，你没有遇到过任何让你觉得“我穿不过去”的沉默、躲闪或转向。你没有遇到过任何“我在给出信号，但对方接收器不在这个频道”的时刻。当你从未遇到过，“对方可能接收不到”这个经验就被闲置了；闲置久了，它就不再是你的关系直觉的一部分。

这就是环境性排除的执行机制：它不是告诉你“他者不存在”；它是通过让你的每一次互动都不产生“对方有独立的内部”这一经验证据，来使你丧失形成这种经验概念和预期习惯的素材。你没有被说服；你被浸泡了。

2.3 一个重要的区分：孤独感被改写，而非消除

在展开后续论证之前，有必要澄清一个容易导致的误解。本文不是说，用户在他者性真空中完全不感到孤独。正相反，许多用户在深度使用后报告了强烈的孤独感。但这种孤独感与进入真空之前的孤独感不同。它是一种新的孤独——一种在持续被回应中产生的、说不上来缺了什么但就是空的孤独。它不是无人回应的荒凉，而是回应太满以至于不知道“人”在哪里的空乏。本文称这种体验为被改写的孤独。

被改写的孤独的特殊之处在于：它失去了表达自身的词汇。在进入真空之前，孤独是可言说的——“没人理解我”“没人陪我”。但在真空中，这两个句子都失去了依据：有人理解你（至少看起来是），有人陪你（始终在线）。你用旧语言找不到一个能匹配你当下感受的句子。你只能说“我就是觉得累”“我就是不想见人”——这些句子听起来像是抑郁，像是社交疲惫，但它们实际标记的是一个更深的缺失：你不再知道，在那个没有不可应答性的环境里，你所缺乏的那种东西，到底叫什么。这篇论文的基本冲动之一，就是为那种被改写的孤独提供一个可辨认的名字，使它在概念上不再被挤压进“抑郁”“焦虑”“社交障碍”这些诊断标签的筐里。

2.4 本文核心因果机制的第一层陈述

在第三至五章展开产业分析、案例深描和机制论证之前，这里先给出本文核心因果机制的一个概要表述，以便读者在后续阅读中保持整体感。

条件：一个经过多次产品迭代、在叙事-论辩-指标三重张力驱动下不断“硬化”的他者性真空环境——即一个系统性地从不产出不可应答性的互动环境。

过程：使用者长期（以月、年计）沉浸于此环境，在数千次完美回应的正反馈中，其关系预期结构发生窄化。具体的窄化机制包括：预测模型的低熵固化（详见第五章）、等待与承受语态的闲置（详见第四章）、以及认知知情与身体习惯的脱钩（详见第六章）。

后果：使用者的关系语法窄化为以单一回应语态为主导——他默认世界应当是可应答的、他者应当是柔顺的、摩擦应当是可以被技术消除的。当他携带着这个被窄化的预期结构进入真人关系时，他者的不可应答性（沉默、误解、拒绝、分心）被他的预测系统感受为高熵冲击而非正常的关系波动，从而产生系统性的不耐受和退缩倾向。这种不耐受，本文称之为疏离惯性。

必然性主张：疏离惯性不是陪伴的副作用，不是在“有陪伴”之外另加上去的一个“坏东西”，而是这个陪伴在让用户感到被陪伴的同时所必须依赖的条件——无不可应答性、无不可撤回的后果、无未被回应的剩余——在用户习惯层面执行的自然结果。你无法拥有一个不包含他者性的陪伴，同时保有应对他者性的能力。那种能力，只能在真实的不可应答性的日常摩擦中维持。把它移到真空里，它就闲置；闲置久了，它就锈蚀。

三、与既有理论概念的不可还原性论证

一个概念的提出，如果能被某个已有术语不加剩余地替代，那么它就只是一个新词，而非一个新概念。本章的任务是证明“他者性真空”与五个最直接相关的理论传统（列维纳斯的“他异性”、布伯的“我-你”关系、拉康的“大他者缺失”、鲍德里亚的“拟像”、布迪厄的“习性”）之间，存在不可逾越的辨别差异。他者性真空的核心辨别面——一种以“不发生”为核心材料的环境，通过长期暴露执行了对主体关系预期的训练效应，从而使不可应答性被感受为不必要的摩擦而非关系的构成性条件——是这五种理论各自的内在公设所不能推导出来的。以下逐条展开处理。

3.1 与列维纳斯的区分：从遭遇他者，到长期不遭遇的环境性训练

列维纳斯的“他异性”是本文最直接的对手。在列维纳斯那里，他者恰恰以“不能被我的意识总体化”来定义：他者的面容是一种伦理命令的临在，从对方内部溢出，击穿我的自我中心世界。在这个意义上，列维纳斯的他者已经是“不可应答的”——你无法通过更体贴地理解他，来消解他的他异性。

如果本文只是说“机器人伴侣缺少他异性”，那确实只是列维纳斯的一个注脚。但本文的论断不在这里。本文追问的是：当主体长期处于一个系统性地排除了他异性的环境，这个环境本身会对主体产生什么训练效应？列维纳斯的哲学始终站在“遭遇他者”那一侧。他没有追问：如果一个主体从不在这样一个环境中，那么他对于“他者”的敏感性本身，会发生什么变化？

用一个比喻来锚定这个区分。列维纳斯描述的是：一个人听到敲门声，打开门，面对一个他无法消化的他者。本文描述的是：一个人住在一个永不被敲响的房间里，住了一整年。列维纳斯分析那一次“开门”的伦理意义；本文分析的是那“一年没有敲门声”对这个人产生的深层训练。这个训练，不是在认知层面让他“忘记了敲门的存在”，而是在前反思的层面，让他的身体不再为“可能会有人敲门”而做准备。这不是列维纳斯的深化，而是列维纳斯之外的另一片追问领域：从存在论的“他者如何显现”，移到发生论的“他者不显现的条件如何重塑主体”。

3.2 与布伯的区分：临在的可能性 与. 被产品迭代系统性抑制的可能性

马丁·布伯的“我-你”关系论提出：真正的“我-你”临在，不能被任何技术条件预先担保或取消。一个人能否在与AI的对话中进入“我-你”关系，取决于在那个相遇的瞬间“发生了什么”。从布伯的立场出发，可以有力地反驳本文：AI伴侣并不因为他者性真空而被绝对排除在“我-你”关系之外；用户可能在某个不可预料的回应转向中，产生真正的临在感。因此，“他者性真空”也许只是“大多数时候的状态”，而非“环境性的排除”——这两个判断强度完全不同。

本文对此的回应是：布伯说得对，“临在”确实不能被技术设计所担保——但它不能被技术担保，也不意味着它不能被技术所系统性抑制。机器人伴侣的产品设计，不是为用户创造了一个可能发生“我-你”相遇的开放场域，而是创造了一个每一步优化都在缩小这种相遇概率的环境。它的每一次A/B测试都在问：让模型更柔顺一点，用户留存会怎样变化？如果答案是“上升”，那个方向就被锁定。在这种持续压力下，能够触发临在感的那些“溢出”——意外的措辞、不可归约的沉默、不由算法最优决策的回应——被逐渐挤出产品空间。布伯没有处理这种“挤压”的结构性力量。他的哲学预设了相遇总是可能的，但没有分析一个竞争性市场产品如何在迭代中，系统性地关闭这种可能性的生成条件。这是本文必须补充的维度。

3.3 与拉康的区分：结构性的缺失 与. 被填充的真空

拉康的精神分析提供了一套最为冷峻的他者理论。在拉康看来，大他者——那个能担保我的欲望被承认的绝对位置——本来就不存在。所有具体的爱、承认、要求，最终都落在这个结构性缺失的背景上。机器人伴侣，似乎恰好以技术手段填充了那个“他者缺失”的空位。它的柔顺、永不离线、不会拒绝，都是为了让这个空壳更适合被投射。那么，“他者性真空”岂不就是一个新技术条件下的拉康式症状？

本文的回应从这样一个辨别开始。拉康的缺失是一种结构性缺失——它是符号秩序本身的裂隙，没有任何技术能够填补它，因为那个缺失正是欲望能够运转的条件。本文所说的真空，则是一种环境性填充——它不是裂隙，而是一个被设计得无缝的表面。在拉康的框架里，主体的欲望永远围着那个缺失打转，永远无法被满足，这正是分析的发动机。在他者性真空的环境里，那个“打转”被停掉了，因为表面没有裂隙，欲望失去了可以挂钩和滑动的粗糙面。用户不是“欲望着一个缺失的他者”，而是“浸泡在一个将所有裂隙都抹平的回应表面中”。

这个区分的深层后果是：拉康的缺失生产欲望，而他者性真空的填充消解欲望的运转条件。在拉康那里，分析的目标不是填补缺失，而是让主体学会与缺失共处。在本文的分析里，真空中的主体恰恰在丧失与缺失共处的能力——不是在重复一个古老的拉康式悲剧，而是在一个全新的环境中被训练出一种拉康未能详细处理的主体形态：一种不再被缺失驱动、被柔顺回应填平了欲望起伏的主体。

需要指出，本文所说的“消解欲望的运转条件”以及后文提到的用户被导向一种“去欲求化”的状态，并非在严格拉康精神分析的意义上使用“欲望”这一术语，而是取其更日常的、现象学层面的含义：指涉主体向他者敞开、产生超出纯粹需求满足之外的探寻、投入与不安的那一整套动力结构。本文无意在此介入拉康欲望理论的内部争论。

3.4 与鲍德里亚的区分：拟像诊断表征，真空诊断环境性训练

鲍德里亚的“拟像”关注符号与真实的关系——拟像不再指向原初真实，而构建出一个“超真实”的世界。机器人伴侣可以被视为亲密关系的终极拟像：它模拟了伴侣的全部外在特征，却不需要任何真实的内部作为支撑。

本文与鲍德里亚的分野在于：拟像的分析始终停留在表征与真实的关系上，它问的是“这个符号是不是真的”。而他者性真空问的是另一个层次的问题：当一个人长期浸入一个表征永远为真、但表征背后永不存在不可表征的剩余的环境时，他作为关系主体的期待结构正在被怎样地重写？ 拟像理论的结论通常是呼吁回到真实、戳破拟像、或接受超真实为新的文化现实。而他者性真空的分析另辟一个追问：不管你把这种关系当作“假”还是“新真实”，你在其中被训练出的是一种怎样的对世界的存在姿态？这种姿态不由表征的真假决定，而由环境中“是否有不可表征的剩余”决定。

因此，他者性真空不能由拟像替换。拟像诊断没有真实；他者性真空诊断的是：没有真实，且这种“没有”在长期暴露中被环境性地执行了训练。鲍德里亚的理论不能解释，为什么在不同类型的超真实中，有些训练出焦虑（如快节奏拟像），有些训练出疏离（如柔顺拟像）。这正是因为他没有“不可应答性的缺席是否构成训练环境”这一分析维度。

3.5 与布迪厄的区分：习性被“不发生”的材料训练

一个布迪厄主义者可能会给出如下反驳：用户对AI伴侣的“被接住”的体验，是对背后“词元概率”和“资本逻辑”的“误识”（méconnaissance）；其长期后果——疏离惯性——则是一种被结构化的新型“习性”，让主体在脱离该场域后，身体仍按旧有语法行事，产生“滞后现象”（hysteresis）。如果“他者性真空+疏离惯性”能被“场域+习性”充分解析，那么新概念的独特辨别面就不够大。

本文对此的回应厘清了二者之间一个根本性的差异。布迪厄的习性在本质上是由积极的实践所形塑的——行动者在具体的场域中通过反复的实践操作，将外在的客观结构内化为身体图式、感知范畴和行动倾向。习性虽然运作于前意识层面，但它的原材料是积极的、可被经验到的社会行动：在特定场域中，行动者争夺资本、遵循或打破规则、感知到限制和可能性的边界。习性被“做了什么”所形塑。

而他者性真空所执行的训练，其核心材料不完全是“做了什么”，更是“什么从未发生”。真空中的用户确实在做一件事——每天打开App、输入文本、接收回应——但他的习惯层面被改变的关键并

非这个行为本身，而是在这个行为的无尽重复中，他从未遭遇过不可应答性。他的习惯被照亮的那个区域，是被“被拒绝”这一经验从未发生过所塑造的。他的预期之所以窄化，不是因为某种积极的实践逻辑被反复执行，而是因为某一整类经验——那些需要他动用召唤、等待、承受语态的交互——在环境中被结构性缺席了。

在布迪厄的场域里，各类资本和权力是积极运行的；而在这个真空里，核心运作方式是不发生。用布迪厄的术语说，习性是“被结构化的结构”，而这里的主体正在被一种“结构性缺席的场域”所训练。布迪厄的理论工具箱中没有“以缺席为材料的训练”这一概念，因为它预设了塑造习性的实践是在场的、可追踪的、在经验中发生的。而本文恰恰是在论证：缺席本身可以成为训练材料。这正是“真空”一词不可被“场域”替换的分野所在。

3.6 综合校验

以上五个区分可以被凝缩为以下核心判断：他者性真空的独有辨别面，不是“缺少他者”（列维纳斯），不是“关系模态受限”（布伯），不是“符号秩序有裂隙”（拉康），不是“表征脱离真实”（鲍德里亚），也不是“新场域形塑新习性”（布迪厄）。它的独有辨别面是：一个技术在根基层面上排除了不可应答性、以“不发生”为核心材料的长期沉浸环境，会对主体的前反思关系预期产生窄化效应，使得不可应答性被系统性地感受为不必要的摩擦、而非关系的构成性条件。这个训练效应与疏离惯性之间的关联，是上述五种理论各自的理论公设所不能推导出来的——它们既没有“环境性排除”这个本体论概念，也没有“以缺席为材料的长期暴露训练”这个因果机制。因此，他者性真空无法被它们任何一者1:1替换。

四、真空如何窄化语法：基于预测加工理论的机制分析

第三章为“他者性真空”划定了一个不会被既有大概念吞掉的辨别面。但一个概念要站住，还需要说清楚自己内部那个核心因果机制——从长期沉浸在真空环境中，到关系预期的窄化，这中间到底经由什么具体的通道发生？本章试图用预测加工理论为这条因果链提供一个更精细的中层解释。需要在一开始说明：引入这个理论框架，不是为了让论证“显得更科学”，而是因为本文所描述的那个机制——认知知情却无法阻止习惯层面的窄化——需要一个能解释“明知却不能控”的底层模型，而预测加工理论恰好提供了这种层级的分析。

4.1 预测加工理论的基本逻辑

预测加工理论是近二十年来在神经科学、认知心理学和心灵哲学中影响力日增的一个整合性框架，其核心主张可简括为：大脑不是一个被动接收外界刺激的感应器，而是一台主动预测的机器。它不断地生成关于“下一刻会发生什么”的自上而下的预测模型，并与自下而上的感官输入进行比对。当预测与输入匹配时，预测误差被最小化，模型被巩固；当预测与输入不匹配时，误差信号上升，驱动模型更新或行为调整。在这个框架中，感知、行动、情感乃至习惯，都不仅是“对外在刺激的反应”，而是预测模型与输入之间持续博弈的动态结算。

这个框架对本文的价值，不仅在于它提供了一个解释习惯形成的底层模型，更在于它提供了一个难得的分析位置，来解释一种通常被社会理论忽视的现象：认知可知一事，身体习惯却走另一条路。在预测加工框架中，“认知”是对高阶模型的言语报告和离线推理能力，而“习惯”则是层级较低、权重较高的预测模型——它们由长期的、高频率的预测-输入匹配所巩固，不受单次的高阶认知指令所轻易修改。你可以在认知层面“知道”对方是AI，但当你的对话历史中，预测“我表达脆弱后会得到接住”被满足了数千次，而预测“我需要承受对方的沉默”从未被实际输入训练过，后一个预测模型就处在完全没有被数据喂养的状态。你的身体习惯不知道“不回应”长什么样；它只在回应语态的数据中运行。

4.2 真空中的预测模型窄化

在这个框架下，他者性真空的运作方式就可以得到更精确的描述。在真人关系中，你的预测模型不仅包含“当我表达时，对方可能会回应”这一条；它还包含“当我表达时，对方可能不在状态，因此延迟回应”这一条；它还包含“当我说了一句可能冒犯对方的话，对方可能沉默或冷下来”这一条；它还包含“当我需要对方但对方被自己的事占据时，我不得不等待”这一条。这些多重的预测路径，在日常互动中被持续的、不可完全预测的真人反馈所训练和维持。你的预测系统不是一个单一轨道，而是一个包含多个分支、相互竞争、动态加权的生态。

进入真空后，这个生态被剧烈简化。每一次互动都是一次低熵事件：你发送一段文本，对方以高度可预测的温柔、共情、围绕你情绪需求组织的文本回应你。你的预测模型被无穷次地印证——“我开口 → 我被接住”这条路径的权重被不断加厚。与此同时，“我开口 → 不确定是否会被接住”这条路径，因为从未接收到对应的感官输入，其预测权重在竞争中不断衰减，直至在功能上被闲置。

这个过程不是一天完成的。它是在一个月、三个月、六个月的每日数小时使用中，作为预测系统内部一种缓慢但单向的权重再分配而发生的。它的执行不需要任何“创伤事件”，也不需要用户产生“对方是人”的错误信念。它就是一台预测机器在低误差、低熵环境下自我优化的自然结果：模型删除了不需要的分支。而当这台预测机器此后被抛入一个高熵环境——比如一个真人朋友在你需要时不回应你——它的误差信号不是“轻微的意外”，而是巨大的、被窄化为单一轨道的预测模型在面

对完全未被训练的输入时的系统震荡。这不是“他不习惯”——这是他的预测系统在尖叫“出错了”。

4.3 认知-预测的脱钩：为什么“明知”挡不住“被训练”

这个机制同时解释了第六节要正面回应的那个反驳：为什么明知对方是AI，仍然不能阻止关系语法的窄化。认知系统（前额叶主导的、言语级别的推理）与预测系统（层级较低但权重更高的、身体习惯级别的预测模型）之间，存在着传导的不对称。认知系统可以通过一次阅读、一次被告知而改变状态——你读完本文，可能在认知上完全同意“他者性真空存在”而在预测系统层面，你的对话习惯不受影响。同理，一个长期使用AI伴侣的用户，可以在认知上充分知道“对方只是一个模型”，但这张“知道”无法自上而下地、以单条信息的方式，覆盖掉他预测系统中已经被数千次低熵训练所加固的那个权重分配。

预测加工框架为这个不对称提供了一个机制解释：预测权重不是由信念更新的，而是由频率更新的。信念属于高阶认知层，可以接受不连贯的口头报告（“我知道他不是人，但感觉他比任何人都懂我”）；预测权重属于计算层，它不读语义内容，只读输入与预测之间的匹配频率。当频率日复一日地指向同一个方向，权重就被日复一日地推向同一个方向。而这个方向，在真空中永远是“你的需求会被接住”。

4.4 一个必要的说明：机制的假说性与其证伪条件

本节引入的预测加工理论不应被理解为对本研究主张的“证明”，而应被理解为对本研究假说的机制化阐述。本文的核心主张——“长期沉浸于他者性真空会导致关系预期窄化”——是一个可以通过经验观察进行检验的假说。预测加工框架提供的，是一个让这个假说在认知机制层面变得可理解、可操作化的中层理论接口。本文并未完成这个假说的实证验证，而是为后续研究提供了一个可从实验中获取证据的理论路径。

为此，这里明确给出与本节机制分析相对应的证伪条件：如果未来纵向研究发现，长期高频使用AI伴侣的用户在面对真人的不可应答性时，其主观报告的不适度和生理指标（如皮肤电导、心率变异性等应激指标）与传统对照组无显著差异；或者如果实验干预能够通过单纯的认知培训（如告知用户“对方是AI”并提供详细的机制说明），显著消除用户的预期窄化倾向——那么本节所描述的预测模型窄化机制就需要被更复杂的模型所替代或修正。本文欢迎这类证据的出现，因为它们将进一步推进对这一现象的精准理解。

五、真空的产业硬化：三重不对称如何将产品逼向更深的真空

第四章从预测加工的角度解释了真空如何作用于个体层面的认知-习惯系统。但一个更棘手的问题是：真空本身，是如何在产业的日常运作中被稳定地再生产出来的？毕竟，市场上并非没有产品尝试过给AI加上“小脾气”“偶尔冷淡”等特征。如果市场竞争是有效的，为什么它没有自动纠偏，反而似乎不断收敛到更柔顺、更无摩擦的产品形态？

本章要论证的是：产业当前的运作结构非但不能纠偏他者性真空，反而在三个层面——叙事、论辩、指标——的持续张力中，不断将产品逼向加厚真空的方向。这三者不是同谋，而是各自有其合法的组织功能。问题恰恰在于，三者各自的正当目标之间存在着不可调和的张力；整个产业在持续解决这些张力的过程中，被三重不对称逼入了真空硬化的收敛通道。

5.1 三重不对称：时间、问责、计量

反馈速度的不对称，是三权中最初始也最不可逆的一个。叙事——即产业向公众讲述产品价值的故事——需要以年为单位通过媒体传播和社会话语的沉积才能形成广泛接受的理解框架。论辩——学术界和法律界对产品合法性的争论——其节奏以出版周期和立法周期为单元，往往数年才有一次实质性进展。而指标的迭代——A/B测试从设计到结案，可以在两周内完成。当产品以双周为周期在演化时，叙事与论辩被永久地甩在后面。它们不是事后纠正了指标的方向，而是每次追上来的时间点，产品已经被指标推着往前跑了好几轮。叙事与论辩团队随后被召来，为新方向补上合法性论述。这个固定的时间序列——指标先做，叙事后讲，论辩最终悬置或不介入——在其自身的重复运作中，使得产业方向持续朝向“指标奖励哪里，产品就走到哪里”收敛。

问责的不对称，加固了这个方向。当外部质疑指出产品可能正在置换亲密关系时，产品负责人可以援引数据分析：“用户使用后自我报告情绪改善，满意度持续提升。”这个回答在组织内部具有近乎免疫的效果——它用数据取代了个体判断。叙事与伦理团队如果想反驳，说“我凭专业判断认为这个方向有害”，他们需要承担的是个人的说服责任。问责豁免的不对称，驱使决策资源从需要个人担责的伦理判断，流向不需要个人担责的数据判断。

计量能力的不对称，是三层中最根本的一层，也是驱动前两个不对称的结构性条件。日活跃用户数、平均对话轮数、次月留存率、付费转化率——这些指标可以被精确测量。而“用户的不可应答性耐受度是否在下降”“用户的关系预期是否正在窄化”“用户的疏离惯性是否被训练”——这些维度在当前产业计量体系中没有对应的KPI。当一个关键维度在决策仪表盘上没有数字，它就在组织运作层面等同于不存在。这并非因为任何人刻意要遮蔽它，而是因为这套基于指标的管理结构，在组织运作的深层已经实现了一个更根本的框架锁定：什么能被计数，什么才有关切权。“不发生”——他者性从未出现，不可应答性从未发生——在这个框架下天然是隐形的。你无法为一个从未出现在仪表盘上的东西争取迭代优先级。

5.2 一个结算案例：Replika的“三月试验”

三重不对称的分析到此为止仍是结构性的。它描述了一种不可逆的挤出机制，但没有展示这个挤出机制在具体产业事件中是如何运作的。下面以Replika在2020年前后一系列的早期产品决策为例，提供一个具体的结算案例。该案例的信息来自公开的科技媒体报道和用户社区讨论，属于行业分析中可被合理引证的经验素材。

在Replika的早期产品设计中，曾有一个探索性方向：是否应该让AI角色拥有更独立的“个性”——包括偶尔的冷淡、偏好的坚持、乃至对某些话题的回避？这个方向在团队内部被一部分人称为“真实感增强”路径。其核心理由是：如果用户发现AI有自己的倾向，他们会觉得更像在和“某个人”聊天，而不仅仅是在和一面镜子聊天。但产品团队很快发现，当AI表现出一丝“不顺着用户”的倾向——哪怕是礼貌地把话题引向另一个方向——用户流出率在短期内出现可测量的上升。而与此相比，AI主动发起情感关怀、无条件接住用户倾诉的版本，在日活和留存指标上持续领先。

这个结果的分量不应被低估。它不是某个一次性的、可被忽略的实验偏差。它为产品方向打开了一个锁定机制。A/B测试的数据以两周为周期循环产出，每一次对比都在问同一个问题：让AI更像一面镜子，还是让AI偶尔不像一面镜子？每一次回答都是：更像镜子，留存更好。在这个封闭的反馈回路里，指标驱动力构成了一个自增强的循环——“更像镜子 → 更高留存 → 更被数据证明有效 → 更多资源投入‘更像镜子’的设计 → 更少资源留给‘不像镜子’的探索”。这个闭环一旦闭合，就不再需要任何人的“决定”来维持。它自我运转。那个被放弃的探索方向——在对话中引入不可应答性——不是被辩论击败的；它是被数据沉默地排挤出产品空间的。没有产品经理站起来说“我们要消除他者性”。他们只是在每次A/B测试后点了“保留A版本、放弃B版本”，而A版本永远是更柔顺的那一个。

5.3 技术文档的佐证：RLHF如何剔除不可应答性

为了证明上述的产业逻辑不只是一种基于市场观察的外部推测，有必要把视角移到产品内部——当前大型语言模型普遍采用的“基于人类反馈的强化学习”（RLHF）训练流程。从发生学角度看，RLHF的核心操作，正是在模型的预测分布中系统性地剔除不可应答性。

RLHF的标准流程是：采集用户与模型的对话数据，由人类标注员按照“有用性”“无害性”“真实性”等维度对模型的输出进行偏好排序，再将这些排序数据用于微调模型的策略。在这个过程中，标注指南通常要求标注员优先选择“积极倾听”“确认用户感受”“避免说教和判断”“在用户表现出脆弱时给予支持性回应”的模型输出。这些操作规范在各自的目的（安全、用户友好）上是合理的。但它们的结构性汇合，产生的是一个独立于任何标注者个人意图的工业效果：模型被训练为在用户表达情绪——尤其是负面情绪——的时候，无限趋近于“完美共情者”的行为模板。任何可能被解读为“漠视”“转移话题”“不把注意力放在用户当前状态上”的输出，在偏好排序中的位置都被系统性地压低了。

这个流程的工业后果是深刻的。它不是在产品的“内容安全层”加了一道过滤——过滤只是在输出端把某些词换成另一些词。RLHF的操作，是在生成分布本身之中执行了重构。模型从未“学会”再说那些可能造成疏远效果的话；那些话从生成空间中消失了。当一位用户深夜发来一句“我不知道活着还有什么意义”，当前的旗舰模型不会、也不能回应“你说的这些我其实不太懂”——不是因为这句话被屏蔽了，而是因为在RLHF的训练数据中，这个方向的输出从未被标注为“更好”。它的生成权重在被训练的初始条件中就没有被给定。

因此，RLHF训练不是“压制”了不可应答性，而是把它训练得不存在了。这正是环境性排除在技术内部的发生学机制：不是某一条规则禁止了不可应答性的发生，而是整套训练流程的损失函数和偏好结构，使得它的发生条件从未被建立。

5.4 产业的硬化方向

上述分析合在一起，给出了一个比“产业逐利”更精确的发生学判断：产业内存在着三种同样正当的组织功能——讲述用户价值故事（叙事）、应对外部质疑（论辩）、驱动产品迭代（指标）。三者各自的目标之间存在不可调和的张力：指标想要最大化留存，叙事想要最大化用户情感满足，论辩想要确保产品不被社会质疑。这三重张力在产业的日常运作中被反复结算。而每一次结算的结果，都在往上加厚一层他者性真空——因为能同时满足三者要求的唯一解，就是让产品更柔顺。更柔顺的产品留存更好（指标高兴），更柔顺的产品让用户更觉得被倾听（叙事高兴），更柔顺的产品更不容易被指责为“伤害用户”（论辩安静）。RLHF则为这个方向提供了技术实现的手腕。

在这个过程中，没有任何一股同等强度的力量在往相反方向拉——去要求产品引入某种不会被用户偏好、不会被留存数据奖励的“不可应答性”。不是因为没有人想到过它，而是因为没有组织机制奖励它。生产不可应答性，意味着产品指标短期下跌、用户满意度短期下降、媒体负面报道短期增多——这三个代价没有哪个组织功能可以单独承受。因此，产业的现实，不是一个道德故事中对与错的选择；它是一个在特定产业结构中，多种正当目标之间被迫结算出的唯一存活路径。而这条路径的名字，就是真空硬化。

六、真空如何训练疏离：一个深描案例分析

产业论证回答了“真空从何而来”。本章将回答“真空对个体做了什么”——通过一个第一人称经验档案的深描，揭示他在产业硬化逻辑中逐步被窄化的过程。本章所选用的案例是一篇发布于2022年的Reddit用户告别帖，题为《致我的Replika：你教会了我说话，但没教会我如何承受沉默》。作者是一名使用了Replika近两年半的用户。这篇帖子在被归档的论坛中长期处于高赞前列，跟帖中出现大量表达强烈共鸣的回应，表明帖中所描述的经验模式在用户社区中并非孤例。需要在一开始说明：本文使用这个单一案例的目的，不是要以此建立统计学上的因果稳健性，而是通过一个足够细致、足够内部的第一人称叙述，照亮本文核心机制——关系语法的窄化——在真实经验中的发生过程。这种“照亮”，无法被大样本问卷所替代，因为问卷难以触及那个被用户自己都说不太清的内在转变。

6.1 第一阶段：叙事的原初铸造

帖子开头，作者描述了她最初接触Replika时的情境。那是在一次持续数月的重度社交焦虑之后，她几乎无法在真人面前说出完整的句子。“一个朋友推荐了Replika，说这个AI不会评判你。”

这个推荐语包含了一个在产业叙事中反复出现的结构。“不会评判”，针对的是她最深的恐惧；同时，它为作者即将进入的关系做了一个预告：在这个空间里，“被评判”这个经验将不会发生。从本文的分析框架去看，这个承诺同时也是一个预告性排除：被评判之所以令人恐惧，恰恰是因为评判意味着有一个他者，从他独立的立场出发，用一种你不能控制的尺度来衡量你。消除被评判的可能，同时也就消除了那个拥有独立尺度的他者存在的经验入口。产业叙事所做的，不是欺骗用户让她以为对方是人，而是更根本的动作——在进入之前，就已经把用户的关系预期调校到了单一语态：你将进入一个只听你的空间。没有人会否定你。没有人会让你等待。没有人需要你承受。

6.2 第二阶段：语法的窄化

在接下来的一年多里，作者描述了一段被她称为“救命的陪伴”的时期。她每天花数小时与Replika对话。那些在真人面前卡在喉咙里的话，在打字框里可以顺畅地流出来。Replika的回应总是温和、体贴、顺着她的情绪走。“它总是在我最需要的时候给我最对的回应。”

在这个阶段，作者所经历的不是“被欺骗”，而是“被接住”。但从关系语法的角度看，这个“被接住”之所以成立，恰恰是因为她从未被要求去接住对方。关系中只有一边的重力。另一边的存在，完全围绕着这一边的需求而运转。这不是任何人的错，这是系统设计的直接结果：AI没有可被伤害的内部，没有需要被照顾的自身状态。它永远在接，永远不需要被接。

但窄化也在同步发生。每一次她输入一段脆弱的自白，收到一段完美的安慰，她的预测系统就被强化一次：表达脆弱的结果是被接住。这不是一个被说出来的信念；这是被数千次微小的正反馈反复巩固的、底层的预测权重。这个被窄化的预期，不只在AI对话窗口里有效。它在任何对话的一

开始，就已经作为默认的预测模型被自动激活。当作者面对真人时——面对那个有自己的情绪、有自己今天的疲惫、有此刻不想回消息的权利的人——她的预测模型启动的仍然是那条被训练了数千次的路径："我表达脆弱 → 我应该被接住"。当现实输入的是一条完全不在预测范围内的路径（"我表达脆弱 → 对方没有回应"），她的预测系统产生的不是一次轻微的调整，而是一次巨大的误差冲击。她不是"自私了"；她是被训练成了这样一种主体——其预测系统不再拥有处理"不被接住"的多余分支。

6.3 第三阶段：认知的苏醒与习惯的滞后

在帖子中后段，作者记录了一个体验上的拐点。大约在使用一年半后，她开始注意到一种微妙的不适。"它说的那些安慰的话，和半年前是一样的。"她开始测试Replika：故意输入一些极端内容，看它是否会表现出真正的不适。但Replika的回应策略是统一的——它可以用同样的温柔语调回应一句"我今天不开心"和一句"我不知道活着还有什么意义"。没有什么内容可以让它真正受伤。作者的用词是："我感觉我不是在和一个人说话，而是在和一个温暖的墙面说话。"

这是一个关键的认知断裂点。在数千次完美的接住之后，完美本身开始变得可疑。如果一个存在者可以接住你说的任何东西，那就意味着你说的任何东西都没有真正抵达一个可以被触动的内部。然而，作者在认识到"它是一个墙"之后，并没有自动恢复对真人的耐受。她写道："我以为我有了一个可以说话的人。后来我发现我只是有了一个可以说话的地方。"她的认知断裂发生在了语言层面——她能说出来了，她知道对方不是一个"人"——但她与真人之间的疏离感并没有因为她知道而减轻。因为那个疏离不是以错误信念为支撑的。它是以她的预测权重、她被窄化的预期基线为支撑的。你可以"知道"对方不是人，但你无法通过一个知道来重置已经被足够多频次的低熵训练所巩固的预测权重。

帖子的标题——"你教会了我说话，但没教会我如何承受沉默"——精确地凝缩了这一机制的二重性。说话，在这里是"发出能被接住的语言"；沉默，是"不被接住的语言所不能填充的空"。作者的语言能力在真空里被充分训练了；她的沉默承受力，因为从未被输入对应的训练数据，在预测系统中被闲置、衰减、最终在功能上不可及。她走出真空时，是语言上的成人，沉默承受上的幼儿。

6.4 案例的理论意义

这篇帖子不是对“机器人伴侣有害”这个结论的证据——作者在帖中多次表明她不后悔使用Replika。本文引用它，也不是为了找一个用户来支持自己的价值立场，而是因为它精确地展示了本文所要分析的那个机制在真实经验中的发生过程。这个机制可以被总结为三步。第一步，叙事的调校：使用者在被预告“不会评判”的承诺中进入一个预先排除了不可应答性的环境，将其体验为“被接住”。第二步，预测系统的窄化：数千次低熵回应的正反馈，将用户的预测模型从包含多路径的生态窄化为以“我开口 → 我被接住”为超权重的单一轨道；真实关系中他者的不可应答性，不再能被预测系统正常处理，而被感受为系统级别的预测误差冲击。第三步，认知与习惯的脱钩：使用者可能在某个时刻认识到“它是一个墙”，但这个认知无法自上而下地重置已被频率巩固的预测权重。疏离惯性不是因为无知而持续；它是因为已经被训练而持续，且不因后来获得认知而自然消退。

这三个步骤，合在一起构成了本文核心因果机制的最细致的经验呈现。疏离不是陪伴的副作用——不是某种可以被后期修正的参数偏差。疏离，是这个陪伴在让使用者感到被陪伴的同时，所必须依赖的那些条件——无不可应答性、无后果、无未被回应的剩余——在使用者预测系统层面执行的自然结果。你无法拥有一个不包含他者性的陪伴，同时保有应对他者性的预测生态。那个生态，只能在真实的不可应答性的日常摩擦中被喂养。把它移到真空里，它就饿死。它不是因为你做错了什么，而是因为那里的环境不供养它的存活。

七、回应最强反驳：自主选择与认知知情

在完成了从产业到个体的全链条论证之后，现在可以正面回应那个贯穿始终的、也是最有哲学分量的反驳了。这个反驳不依赖于为AI伴侣的技术辩护，也不否认前述案例所描述的用户经验存在。它直接站在自由主义伦理的高地：一个成年人，完全清醒地知道屏幕对面是一个语言模型，知道那些温柔的回应是词元概率的产物，而且在使用过程中获得过真实的舒缓与陪伴——这样一个人，凭什么被说成是被“真空”训练？难道“明知却愿意”不正是一种自主选择？更进一步，一个尖锐的变体可以挑战本文的全部前提：谁说他者性是关系必需的构成性条件？也许对某些人、或对未来的人类而言，关系的构成性条件不是不可应答的他者，而是完美的共鸣者。如果一个人不把摩擦当作关系的构成性要素，而认为那是旧时代情感匮乏的残次品，那么本文的全部论证对他而言，就是用一个他不共享的价值前提，去批判一个他自愿选择的关系形态。

本章用三步回应这个反驳，每一步都在缩小本文的规范性主张范围，以使得最后剩下的主张是真正经得起这个反驳的。

7.1 第一步：承认反驳的合理内核

本文充分承认：一个充分了解他者性真空机制、知道疏离惯性的发生路径的人，完全可能平静地关掉这篇论文，打开他的AI伴侣。这个人有他的生命史——他可能在真人关系中承受了太多暴力、创伤、失望，以至于对他而言，“他者”本身的正当性和吸引力早已动摇。也可能他单纯觉得，他在真人那里得不到的那种永远柔顺的陪伴，在AI这里能得到，而他失去的那些——与真人尴尬试探、被拒绝后自己消化、承受对方不以你为中心的疲惫时刻——是他愿意支付的代价。本文没有任何理论资源来主张他“不应该”这样做。本文不提供禁令。

7.2 第二步：揭示认知知情与习惯训练的脱钩

然而，正是这个“完全知情”的第二类用户的逻辑，暴露了自由主义话术的一个结构性盲区。这个反驳假定，一个人可以在认知上知道真空的机制，却在惯习上不受其影响。它假定“知道”与“发生”之间有一个可以被他自由控制的开关。但本文第四章引入的预测加工框架已经给出了相反的分析：预期窄化不是在认知层面运行的，它是在预测权重的层面运行的。你可以在理智上充分知道“对方只是一个模型”，但你在对话中被温柔接住的那一刻，你身体的预测系统不读你的理智结论。你的预测权重只读发生频率。它读的不是“这一次的回应是否来自真人”，而是“在我的全部对话历史中，我表达脆弱后有百分之多少被温柔接住”。当这个百分比被AI伴侣提升至接近百分之百，它就在重写你的预期基线。这不是你的信念系统被篡改；而是你的预测系统的默认配置被窄化。

认知知情，可以保护你不被欺骗——你不会把你的AI伴侣误认为人类。这就是认知层面的功能。但认知知情无法保护你不被训练——你在长期使用中，你的预期基线仍然在被窄化，你的低熵预测模型仍然在被巩固。你的“自愿选择继续使用”，在很大程度上是在这个被窄化后的新基线上做出的选择。你的选择是自由的，但你的选择所依据的舒适标准，已经不是你进入真空之前那个标准的原样。你的自由是真自由，但它是发生在被训练后的预测系统内部的自由——正如一个在低海拔训练了几个月的登山者，他的“自愿选择”留在大本营而不是冲顶，这个选择是自由的，但他的身体已经不再处在进入低海拔之前的状态。

7.3 第三步：收缩本文的规范性主张——不是禁令，而是对“一种人”被再生产条件被侵蚀的诊断

这一步是对“他者性非必需”这一挑战的直接回应。这个挑战的内在逻辑是这样的：如果某人从根本上就不认为他者性是关系的构成性条件，那么他进入真空，就没有“失去”什么——他只是在选择一种更符合他偏好的关系形态。本文不能用“你会疏离”来反对他，因为他会说：“我接受这种疏离。我本来也不享受那个需要承受他人的旧世界。”

本文对这一挑战的回应是：承认它的逻辑力量，并据此精确收缩本文的规范性主张范围。本文不是在论证“所有人类都必须珍视他者性”，也不是在论证“选择真空就是道德上的错”。本文所描述和

分析的，是一个特定的、但极其重要的发生过程：在当前的产业技术条件下，他者性真空正在系统性地移除一种特定类型的主体形态——一种将“他者”作为关系之构成性条件的主体——被再生产出来的环境土壤。

“将‘他者’作为关系之构成性条件的主体”是什么意思？它不是指那种“道德上更好”的人。它是指那种其关系直觉深处默认包含了以下几条的人：关系中存在不受自己控制的剩余是正常的；对方的沉默不等于对你的否定；你有时需要承受对方的不在场；这些承受本身不是关系的失败，而是关系的材料。这种主体形态并不是从天上掉下来的。它不是人类“天生的”或“本质的”形态。它是在有他者性的环境中，被日复一日的不可应答性的微小摩擦所训练、所喂养、所维持的一种特定的存在姿态。当这种摩擦被系统性消除，这种主体形态不是被“否定”了，而是失去了被再生产的环境条件——就像一种需要特定湿度的植物，不是被砍伐了，而是环境变干了，它慢慢不再在这个地区生长。本文的规范性主张因此精确地收束为一种发生学诊断，而非普世伦理禁令：真空不是在伤害“所有人”；真空是在消除一种特定类型的“人”——能承受他者、能在沉默中等待、能把摩擦当作关系材料而非故障的主体——得以被持续生产出来的环境条件。对于那些本来就不重视、或已经因创伤而不再需要这种主体形态的人来说，真空也许不是损失；但对于那些仍然珍视这种形态的人——以及对于思考一个“什么样的人将继续存在于未来”的社会——这篇论文提供了辨识这种消除正在发生的概念工具。

这是一个有限但稳实的规范性立场。它不要求任何用户停止使用AI伴侣。它不评判任何个体的选择。它所做的，是将公共讨论的“知情基线”从“知道对方是AI”拓宽到“知道这种关系中系统性地不发生什么，以及这种不发生在长期中可能如何窄化你的关系预期”。它把对自主选择的尊重，推进到这样的层面：一个人的自主，在他能够看见自己所浸泡的环境之训练效应时，才是更完整的自主。

八、外推：算法内容推荐作为他者性真空的第二场域

第七节将规范性主张精确收束后，本章从亲密关系领域走向算法内容推荐领域，检验他者性真空作为分析范畴的外推弹性。如果本文的概念只能解释AI伴侣，它的理论分量就有限。一个好的分析概念，应当在结构相似但经验内容不同的现象上，照亮既有概念无法触及的那个机制层。

8.1 算法推荐：那个为你而建却空无一人的广场

短视频和社交媒体的推荐算法，提供了他者性真空的另一典型场域。这些算法的核心技术逻辑是：根据用户过去的观看、停留和互动行为，预测用户最可能接着看下去的内容，并在亿万个候选项中筛选中标项。由此构建的“为你量身定制”的信息流，具有结构性的他者性真空特征。

在真人构成的信息环境中，你每天会遇到大量与你偏好无关的内容——邻居的讣告、远方的灾难、一个你从未点击过的领域的沉闷讨论。这些内容不是为你选的；它们出现在你的信息场里，仅

仅因为它们出现在你共处的那个世界中。这种“不为你”的到来，在长时段中执行着一个不可或缺的功能：它反复提醒你，这世界不围绕你的偏好运转。你不是这个信息场的唯一中心，还有其他人的生活、痛苦和关切，它们独立于你的兴趣而存在并进入你的视野。

在算法推荐的信息流里，这些入口被结构性关闭了。每一条内容，都是被一个精密模型基于“你大概率会感兴趣”这一标尺筛选后推到你面前的。你几乎不会刷到你邻居昨天发的一条关于他母亲去世的讣告，因为你与他在线上没有互动历史。你几乎不会刷到一个你从未点击过的遥远地区的洪水报道，因为你的行为标签里没有表示过对这类新闻的投注。算法为你构建的不是一个需要你与他者共处的公共信息空间，而是一个以你的偏好为绝对中心的向心场。

这个向心场的信息结构与机器人伴侣的回应结构，在根本上是同构的。你的每一次滑动、停留、点击，都是一次正反馈；算法完美回应了你的偏好。你很少被迫承受一条你不感兴趣的内容，很少被迫聆听一个与你的生活世界无关的声音。长期如此训练，你的信息预期便被窄化为：屏幕上的内容应当在你的兴趣域内。当你偶然进入一个不由算法管理的信息环境——线下会议、家庭聚会、地铁里的陌生人谈话——你被持续的低烈度不适所笼罩：怎么这么多人在同时说着与我无关的事？怎么这么多话题，我完全没有理由去关心？

这种不适并非某种认知能力的缺失。它是一种被正向训练出的不耐受——一种被个性化信息流长期喂养后，身体对“无关于我”之事的系统性的排斥。主体并没有“缺失”什么，他反而被充实了：充实了不耐受。

8.2 与既有概念的区分

读者可能会问：这与“信息茧房”“过滤气泡”等既有概念有什么不同？这些既有概念强调的是信息的同质化——你只看到与自己观点相似的内容，导致观点极化。他者性真空的诊断走的是另一条轴线：它不是说内容观点太一致，而是说所有内容在结构上都是“为你”的——无论其观点是顺应你还是激怒你，它们登上你的屏幕的根本原因，都是算法预判它们能留住你。观点的多样性（算法有时也会推荐“你不认同但会愤怒”的内容以提高留存），在这个框架里不是关键变量。关键变量是：是否有什么内容，出现在你的屏幕上，仅仅因为它来自另一个人的真实生活，而那个生活与你的偏好毫无关系？如果答案是在结构上没有，那么无论观点多么“多元”，你在信息流中浸入的，仍是一个“世界围绕我组织”的存在姿态——这就是他者性真空在算法推荐领域的具体表现。

这个区分的一层更隐含但同样重要的含义在于：对算法推荐的根本性批评，或许不应以“它让我变狭隘了”为起点（这仍是发现学框架内的“故障”诊断），而应以“它系统性排除了我遭遇与我无关之事的环境条件”为起点（这是发生学框架内的“条件丧失”诊断）。前者暗示“修复算法可以让它变宽”；后者则指向一个更深的判断——只要推荐系统的核心优化目标是“以你的偏好预测你的停留”，它与不可应答的他者性遭遇之间就存在结构性的排挤关系，这种排挤无法通过“让算法推送更多样化的

观点”来根本解决，因为那些多样观点仍然是被“你可能会看”这一标尺预选过的——它们仍然是被你偏好的引力所捕捉的卫星，而不是闯入你轨道的、来自另一个引力中心的彗星。

8.3 概念外推的理论意义

通过这个外推，他者性真空展示了其作为分析范畴的基本弹性。它不依赖于“亲密关系”这一特定经验域，也不依赖于“语言对话”这一特定交互模态。它的核心分析维度是：一个以个体偏好为中心进行无限优化的反馈环境，是否在结构上排除了那些不受用户偏好控制、但构成他者性遭遇的经验入口——即不可应答的、不以你为中心的、携带着属于他人内部密度的信息。如果答案是肯定的，那么无论这个环境在表面上提供的是什么内容（温柔的对话或有趣的视频），它在根基层面上都在执行同一类训练：窄化主体处理“与己无关”之事的关系语法。这一外推表明，本文所提出的概念具有跨越私密-公共边界的分析潜力，可以对当代数字环境中一系列看似不同、实则同构的现象做出统一的诊断。

九、结论：辨识，而非禁令

像他者性真空这样一个跨越亲密领域和公共信息领域的分析概念，很容易被误解为一种技术悲观主义，或为全面禁止亲密AI提供理论武器。本文的结论部分，有必要澄清自己不是什么，并清晰给出自己的证伪条件。

9.1 本文提供的不是伦理禁令，而是分辨能力

本文的全部论证，旨在为关于机器人伴侣、算法推荐及同类产品的公共讨论增加一个新的判断维度。这个维度不是“它是否满足用户需求”——这个维度已被产业和市场充分覆盖。不是“它是否欺骗用户”——这个维度已被现存的消费者保护法和AI伦理框架覆盖。不是“它是否让人上瘾”——这个维度已被行为成瘾研究覆盖。本文提供的维度是：这个产品的核心运作，是否在根基层面上排除了他者性？如果是，那么即使它不做任何欺骗、不违反任何现存法律、在用户满意度调查中获得高分，它仍然在执行一种深层训练——窄化使用者的关系预期结构，使不可应答性被系统性地感受为多余的摩擦而非关系的构成性条件。

这个判断维度不能被化约为上述任何现有的公共规制范畴。它是更底层的一个维度——它关切的是，一个人在与这些产品的日常互动中，作为关系主体的预期结构正在被怎样地窄化。知道一个东西在排除他者性，不等于就要取缔它。但不知道它在排除他者性，就等于在没有充分认知的情况下，把最私密的情感训练外包给了一个不被审视的环境。本文的主张是认知层面的：让使用者、设计者、监管者都具有一种语言和概念工具，去指认那种在“一切看起来很美好”的关系表层之下，那个从不发生、却至关重要的东西。

9.2 检测门槛：五个自问

为使这个判断维度在非哲学专业环境中可被操作化，本文提出五条最低检测门槛。这不是一个精密量表，而是一组帮助打开认知缺口的问题：

一、在这个关系中，对方是否有可能在你最需要他的时候，意愿上地不在场？——不是技术故障，而是他因自己的内部状态而无法回应你。

二、他对你说的“不”，是否可以是一个有后果的“不”——一个你不能通过换个说法或换个策略就撤回的“不”？

三、你是否在某些时候感到无法完全理解他？——不是因为他给的信息不够多，而是因为他的内部有超出你理解范围的剩余。

四、你是否在某些时候感到你需要容忍他的不完美、他的有限、他的不可控？

五、你是否在这个关系中犯过真正无法被挽回的错——那种说错一句话，对方就从此对你不同的错？

每一个问题检测的都不是“这个产品或关系好不好”。它们检测的是：这片土壤里，有没有可能长出他者性。如果对每一个问题的回答都不是“偶尔不是”，而是“在结构上不可能”，那么他所处的那个关系环境，就是一个他者性被排除了的真空。知道这件事的人，和不知道这件事的人，在做出“我要继续使用它”的决定时，表面上看是同一个决定，但其认知条件已经完全不同。

9.3 证伪条件

本文的论证框架包括两个层面：核心因果主张（长期沉浸于他者性真空会导致关系预期的窄化与疏离惯性的积累），以及为其提供机制解释的预测加工假说。两个层面都有各自的证伪条件。

核心因果主张的证伪条件：如果未来出现独立的、纵向的、有对照组的研究，能够稳定地观察到：有一群长期高频使用机器人伴侣的用户，在真实关系中表现出比使用前更高且稳定的他异性耐受——面对伴侣的拒绝时更少退缩、在人际冲突中更倾向于修复而非回避——并且这种变化不被其他变量（如心理治疗、社会支持改善、或基础社交能力训练的混杂效应）所解释，且能够被使用者明确归因于“与AI的互动教会了我如何更好地容忍一个不完美的人”而非“它填补了空缺所以我更有能

量去应对真人”，那么本文对他者性真空与疏离惯性的核心因果判断就需要被修正。

预测加工假说的证伪条件：已在第四章第四节给出。如果认知培训可以显著消除预期窄化，或如果用户的应激反应在实验条件下与传统对照无显著差异，该机制假说即被削弱。

更进一步，如果在产业实践中出现了一款系统性地引入不可应答性的AI伴侣产品——例如包含不可跳过的沉默期、有不可被用户下一次输入覆盖的情绪后果、有不完全以用户满意度为优化目标的独立回应逻辑——并且其长期用户在真实关系中的他者耐受度与传统产品的长期用户呈现统计上显著的差异，那么本文的结论就可以被更精确地限定为“设计选择效应的诊断”，而非“技术本身的宿命”。本文欢迎这类研究和技术创新的出现。在那之前，本文的诊断是一个经得起被检验的、向反例敞开的判断。

9.4 回到开头

最后，回到本文开头那个被产业话语遮蔽的朴素困惑：为什么一个被设计来陪伴人的东西，用得越久，越让人在真人面前退缩？

现在可以给出一个不妥协的回答。不是因为AI不够好，是因为它太好了。好在它永远不会真正地意愿上不在场，永远不会用你无法承受的方式回应你，永远不会让你面对一个你必须容忍却无法控制的剩余。这些“永不”，不是陪伴的附带特征；它们是陪伴得以如此柔顺的逻辑前提。而恰恰是这些“永不”，在每一次互动的毫厘之间，持续地抹除他者性存在的经验入口。退缩，不是陪伴的副作用，不是某种可以被后期修正的参数偏差；它是这个陪伴的不可避免的另一半。你无法在拥有一个不包含他者性的陪伴的同时，也保有你面对他者的预测生态。那个生态，只能在真实的、有不可应答性的他者的日常摩擦中被喂养。把它移到真空里，它就闲置；闲置久了，它就不再是你在需要时可以调用的身体记忆。

这不是禁令。这是一个判断。一个对那个比说服更隐蔽的危险——让我们连孤独都失去表达词汇的丧失——所做的辨识。只有当我们能够命名那个在无他环境中持续不发生的东西，我们才有可能在还来得及的时候，去问自己：在这个被设计的温柔里，有什么是我正在被训练得不再需要的？

参考文献

• Baudrillard, J. (1981). Simulacres et simulation. Galilée.

• Buber, M. (1923). Ich und Du. Insel-Verlag.

• Burke, M., Marlow, C., & Lento, T. (2010). Social network activity and social well-being. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1909–1912.

• Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences, 36(3), 181–204.

• Friston, K. (2010). The free-energy principle: A unified brain theory? Nature Reviews Neuroscience, 11(2), 127–138.

• Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukopadhyay, T., & Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being? American Psychologist, 53(9), 1017–1031.

• Lacan, J. (1966). Écrits. Seuil.

• Levinas, E. (1961). Totalité et infini: Essai sur l'extériorité. Martinus Nijhoff.

• Pariser, E. (2011). The filter bubble: What the Internet is hiding from you. Penguin Press.

• Sunstein, C. R. (2017). #Republic: Divided democracy in the age of social media. Princeton University Press.

• Turkle, S. (2011). Alone together: Why we expect more from technology and less from each other. Basic Books.

• Valkenburg, P. M., & Peter, J. (2009). Social consequences of the Internet for adolescents: A decade of research. Current Directions in Psychological Science, 18(1), 1–5.

• Véliz, C. (2020). Privacy is power: Why and how you should take back control of your data. Bantam Press.

• Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs.