arXiv:2603.26324v1 [cs.DL] 27 Mar 2026

PRISMA：迈向负责任药学知识管理的规范性信息基础设施

Eugenio RodrigoZimmer Neves[1], **Amanda VanonCorrea**, **CamilaCampioni**, **GabielliPareGuglielmi**, **BrunoMorelli**

Pharmdata，巴西阿雷格里港

2026年2月

摘要。现有大多数药学领域的人工智能方法，将三个在认识论上截然不同的操作压缩到单一技术层中：文档保存、语义解读和情境化呈现。这种混淆是反复出现的脆弱性之根源，包括来源追溯丧失、解释不透明性、警报疲劳以及问责能力的弱化。本文提出 PATOS–Lector–PRISMA（PLP）基础设施，作为负责任药学知识管理的规范性信息架构。PATOS 通过明确的版本管理和来源溯源保存监管文件；Lector 以人工策展实施机器辅助阅读，生成锚定于一手来源的类型化断言；PRISMA 则通过 RPDA 框架（监管、处方、调剂、给药）提供情境化呈现，将同一信息核心折射为不同的专业视图。该架构引入证据包，作为可问责断言的正式单元（经过版本管理、可追溯、认识论边界明确且经策展人验证），并依据言外之力对断言进行类型化。一个演示示例使用真实系统数据，追踪一水合安乃近在这三个层中的呈现过程。该架构在巴西的监管背景下开发并验证，其基础是一套实际运行的系统，其中包含超过 16，000 份官方文件和38 个精选证据包，涵盖五种参考药物。本文证明，该方案可与操作型决策支持系统互为补充，提供当前系统所缺乏的基础设施条件：文档锚定、解释透明度和机构问责。

关键词：药房信息系统、知识管理、数字保存、可解释人工智能、临床决策支持、药学情报、社会技术系统

1. 引言

人工智能融入药学实践的步伐正在不断加快。综合性综述记录了日益增长的

[1]通讯作者：eugenio@pharmdata.com.br

药物相互作用检测、处方审核、自动化调剂、不良事件预测和治疗方案核对支持等应用[1–3]。NoHarm.ai 等运行系统在真实医院环境中展现出可衡量的临床影响，显著降低了处方错误率，并大幅提高了药学审核覆盖率 [4,5]。大语言模型的出现为这一领域增添了新的维度；评估结果表明，在特定药物审查任务中，其表现可与人类药师相当 [6, 7], ，但随着临床复杂性增加，其表现会显著下降 [8]。

然而，这一技术进步与持续存在的结构性脆弱性并存。有关药房信息系统的文献提供了实证证据，表明在信息层错配的情况下引入技术，可能会放大潜在故障并产生意想不到的后果：文档碎片化、认知过载、任务重复，以及信息质量逐步下降 [9–11]。因此，问题不在于缺少计算工具，而在于缺少一种能够连贯组织药学知识生产、解释和呈现过程中涉及的不同责任层级的基础设施 。这一诊断并非制药领域独有：近期对更广泛的企业人工智能生态系统的分析也得出了相同结论。MIT NANDA（网络化智能体与去中心化架构）报告基于对 300 项公共人工智能项目的系统综述以及对 52 个组织的访谈，指出尽管企业在生成式人工智能上的投资高达 300 亿至 400 亿美元，95% 的组织仍未获得任何回报，只有 5% 的定制化试点项目投入生产 [12]。失败归因于脆弱的工作流、缺乏上下文学习以及运营失配。与本文尤为相关的是，在所分析的九个行业中，医疗保健与制药 行业的颠覆性评分处于最低水平之一（满分 5.0 分中为 0.5 分），相关应用仅限于文档与转录试点，并未改变底层临床模型。Ramesh [13], 在评论这些发现时指出，薄弱的基础并非根本问题，而是组织在不确定性下进行决策、奖励工作和学习的方式所体现出的症状。

这种错位的根源在于一个明确的概念缺口：大多数现有方法（无论是临床决策支持系统、检索增强生成（RAG）流程、对话代理还是规则引擎）都基于已经转换过的数据运行，却没有保留一手来源、解释与建议之间的正式区分。原始监管来源（药品说明书、专论、临床指南、规范性文件）通常会被摄取、拆分并重新处理，直到其来源溯源无法恢复。Talisman [14] 精准地阐述了这一机制：在训练过程中，数十亿份文档被压缩、合并，

并以统计方式分布在数十亿个参数中；这一过程具有损耗性 且不可逆，从结构上摧毁了来源溯源。这种破坏发生在两个层面：检索层可以引用来源，而参数化层则会不可挽回地丢失来源溯源。由此产生了氛围式引用现象：带有合法性标记的虚构参考文献污染了后续知识库。解释是一种认知行为，需要语境、判断和专业责任，却常常被委托给统计模型，且缺乏明确的可追溯机制。而面向最终用户的呈现也经常将系统输出与来源权威性混为一谈，侵蚀机器、专业人员与机构之间的责任边界。

近期研究彼此独立且未经协调地进一步表明，科学界已认识到这种脆弱性。Sabel 等人证明，即使在高度受控的情境中（就单一种非处方药提供咨询），应用大语言模型也需要逐步引入规范层、控制 YAML 文件、决策图、专用智能体以及明确的人工监督 [15]。概率模型本身无法提供该领域所要求的问责保障。这一实证发现为一个更广泛的结构性问题提供了外部验证：药学信息的核心挑战并非算法层面，而在于架构。

正是在这一领域，本文提出将**PATOS–Lector–PRISMA** 组合打造为负责任药学知识管理的规范性信息基础设施。该方案包含三个功能各异、概念上相辅相成的层。**(i) PATOS** 是一个自主文档保存层，致力于实现明确的版本管理和原始证据完整性，其基础是 OAIS（开放档案信息系统）模型和可信数字存储库的原则 [16, 17]。**Lector**是一个机器辅助阅读与技术策展环境，旨在为保存的源文档与专业人员的解读搭建认知中介空间，并借鉴计算意义建构 [18] 和人在回路系统 [19]的传统。**(iii) PRISMA** 是一个情境化呈现层，将知识组织为带版本的断言，并附有正式证据链（证据包）、监管情境图以及明确的论证结构；其基础是数据溯源 [20, 21]、知识图谱 [22], 和计算论证 [23]领域的相关文献。

这一基础设施的设计遵循一项审慎确立的原则：系统不作决定、不提供处方，也不授予权限。它保存来源、组织解释空间，并记录知识构建的轨迹，将责任牢牢置于人类和机构的范畴之内。这一立场直接根植于社会技术系统文献——

中关于卫生自动化风险的证据 [24]，，以及当代人工智能治理研究 [10, 25], [26, 27]。

本文提出了四项主要贡献：

1. **架构问题的构建**：识别并刻画药学知识管理中保存、解释与呈现之间的结构性缺口，并指出这一缺口是健康信息系统反复出现脆弱性的根本原因。

2. **提出 PATOS–Lector–PRISMA 基础设施**：介绍一种三层架构，其概念基础扎根于成熟的学术传统，旨在实现可追溯性、可审计性，并明确划分职责。

3. **引入证据包这一概念**：提出一种正式的证据包装单元（具有版本信息、可追溯且语境化），作为在监管和临床领域支撑断言的核心机制。

4. **依据言外之力对药学断言进行类型划分**：依据断言的规范性质（适应证、禁忌证、警告、注意事项、规范性沉默等）进行分类；这一分类以言语行为理论和道义逻辑为基础，并将规范性沉默作为一个具有认知相关性的类别进行形式化处理。

本文结构如下。第 2 节将本文提案置于现有文献的脉络中，阐述支撑各层的理论基础。第 3 节介绍所提架构及其设计原则。第 4 节以 NoHarm.ai 为启发性案例，展示该基础设施如何补充而非取代运营决策支持系统。第 5 节探讨架构、监管和制度层面的影响。第 6 节总结本文的贡献并展望未来方向。

2. 相关工作

PATOS / Lector / PRISMA 这一组合所提出的架构，处于多个成熟学术传统的交汇处。这些传统在各自领域内虽已发展成熟，但应用于制药和监管知识管理时，历来仍彼此割裂。这些领域包括数字文档保存、机器辅助阅读、证据来源与可追溯性、情境化知识表示、可解释人工智能，以及对健康信息系统的社会技术研究。该提案的原创性不在于对某个单一

子领域的渐进式推进，而在于将这些基础进行连贯的架构集成，并明确以监管、临床和机构问责为导向。

2.1 文档保存、版本管理与监管证据主权

自主文档保存层与解释或分析层之间的概念区分，直接植根于开放档案信息系统（OAIS）模型。该模型被正式确立为 ISO 14721，为经口摄入、储存、管理和访问已保存数字对象制定了结构原则，并强调真实性、完整性、明确的版本管理，以及与二次使用相独立 [16]。与之相辅相成的是，可信数字存储库模型进一步指出，存储库的可靠性并非源于其应用目的，而是取决于其控制、审计和机构治理机制是否稳健且可验证 [17]。

在制药和监管信息的语境中，这些原则尤为重要，因为药品说明书、专论、临床指南和规范性文件等文档都具有时效性，存在连续更新的版本，并会直接产生法律效力。Lagoze 等人在数字文档工程领域的研究预见了这一需求，将数字对象视为富含元数据、关系和转换历史的实体，超越了将文档看作被动文件的观念 [28]。这些基础支撑着 PATOS 作为规范性文献层的作用，其首要功能是以自主方式保存来源，实行明确的版本管理并确保可追溯性，而不预先引入临床或监管层面的解释。

2.2 机器辅助阅读、策展与计算意义建构

Lector 的构想遵循关于计算意义建构的经典文献，尤其是 Pirolli 和 Russell 的模型。这些模型将复杂信息的分析描述为一种迭代过程：计算系统协助组织、筛选并外化人类推理，而不取代分析者的判断 [18]。这种方法广泛应用于解释取决于语境、例外情况和模糊性的领域，而这些因素都是药学与监管信息的核心特征。

关于探索式搜索 的研究进一步佐证了这一观点：与大型文献语料库的相互作用是一种开放的认知过程，受暂定假设、不断改写以及逐步建构理解的过程引导 [29]。在技术和科学阅读这一特定领域，Marshall 表明，即使在计算环境中，数字阅读本质上也是主动的、带有注释性的和解释性的；这一发现否定了将文档仅仅视为自动生成答案之输入的模型 [30, 31]。

同样，关于 人在回路系统 的文献表明，在解释循环中明确纳入人类，并非技术上的局限，而是在敏感领域运行的系统的一项结构性要求；在这些领域，解释错误可能造成临床、法律或机构层面的后果 [19]。这些研究奠定了 Lector 的基础，使其成为辅助阅读、技术策展和认知中介的环境，而非自主进行规范性解释的代理。

2.3 决策支持系统中的来源溯源、可追溯性和可解释性

PRISMA 提出的证据可追溯性明确要求，在数据溯源文献中有坚实的理论依据。由 W3C 标准化的 PROV 模型，建立了一个正式框架，用于表示参与数据生成、转换和使用的实体、活动和代理，从而能够重建来源链和责任链 [20]。Buneman、Khanna 和 Tan 的经典研究进一步深化了这一概念，区分了 *where* 溯源和原因溯源，并强调对结果的说明既取决于其来源，也取决于证明其存在合理性的条件 [21]。

这些概念与当代关于可解释人工智能（XAI）的文献直接交汇。在相关研究中，可解释性被视为一种系统性且依赖语境的属性，而非孤立统计模型的可选特征 [32]。在健康领域，新兴法规以及相关研究进一步强化了这一要求；这些研究表明，不透明的决策支持系统可能带来重大的临床和法律风险。

这一要求的紧迫性因一项发现而进一步凸显：大语言模型会以结构上不可逆的方式破坏来源溯源。Talisman [14] 表明，大语言模型训练是一种有损压缩过程，它会瓦解将知识与其源头相连接的文献链，使来源溯源在参数层面变得无法恢复。作者追溯了来源溯源作为档案原则从1841年到伯纳斯-李的关联数据系统的发展历程，并指出，生成式人工智能产业忽视了数十年来关于计算来源溯源的研究。在制药领域，这一现象尤其严重：捏造的药物相互作用或虚构的禁忌证，并非学术上的不便，而是对患者安全的直接风险。在此语境下，证据包这一概念虽最初是在 PRISMA 的范围内提出，却在概念上契合这样一项要求：每一条呈现出来的断言都必须附有可追溯、具备语境信息且可审计的证据；同时，它也直接回应了这一挑战：应将来源溯源作为内在的架构属性加以保留，而非作为事后添加的元数据。

2.4 知识图谱、语境与监管论证

关于知识图谱的文献为 PRISMA 的呈现层提供了结构基础，将知识视为由语义类型化的实体和关系构成的网络 [22]。在制药领域，这种方法已被用于表示物质、适应症、禁忌证、特定人群与临床证据之间的关系，但往往与文献来源信息相分离。

此外，Brézillon 提出的情境图 将语境形式化为一个明确的实体，从而能够表示与特定使用条件、临床情景或监管框架相关的有效性变化[33]。在计算论证领域，受 Toulmin 启发的模型为组织主张、证据、限定条件和反驳提供了一个框架；这在监管领域尤为重要，因为断言的力度取决于其支持条件是否明确 [23]。这些基础构成了 PRISMA 的设计根基：它是一个用于呈现有依据断言的情境化呈现层，而非自动响应系统。

与此相呼应，当代对语义层 （面向人类使用的指标治理）和上下文图 （具备可追溯性和推理能力的决策基础设施）的区分，精准阐明了此处所提架构要应对的挑战：人工智能系统需要的不是用于标准化计算的层，而是支持围绕概念、关系和领域推理规则进行推理的形式化表示 [34]。Talisman 指出，业界已投入大量资源建设语义层，以确保指标一致，但“意义并不等同于测量”：指标无法解释因果关系，也不支持推理。这一观点是在企业数据工程背景下提出的，与此处所提架构直接对应：PRISMA 并非药品领域的语义层 （标准化的指标呈现层），而是一个情境化呈现层，由自主文档保存（PATOS）以及经人工策展的合格解读（Lector）提供支撑。

2.5 社会技术系统、系统性风险与治理

该提案与社会技术系统文献直接对话；这类文献反对在复杂环境中作出完全确定性决策的观点。Suchman 指出，人类行动始终具有情境性、依赖语境且部分不可预测，因此将权威来源完全委托给自动化系统并不妥当 [24]。在卫生健康情境中，经典研究表明，信息技术往往会带来意料之外的后果，这些后果源于系统、专业人员、流程与

组织 [10, 11, 25]。

关于机器行为和人工智能治理的较新研究进一步强调，我们需要将智能系统视为机构组成部分，其影响超越了技术领域 [26, 27]。这一观点明确构成 PRISMA 所采纳立场的基础：系统负责构建知识并记录解释轨迹，而非行使决策权，从而维护人类与机构责任。

2.6 认识论基础：信息、解释与呈现

本文提出的保存、解释与呈现之间的架构性分离，并非随意为之，也不只是技术层面的安排；这一构想在信息科学、哲学诠释学和信息哲学的既有认识论传统中，都能找到相互印证的依据。

Buckland 在其奠基性研究中，将常被混为一谈且有害无益的“信息”区分为三种含义：作为事物的信息（文档、数据、有形制品），作为过程的信息（使人获得信息的认知行为），以及作为知识的信息（由此产生的理解，即通过这一过程感知并领会的内容）[35]。这一区分与所提出的架构直接对应：PATOS 保存作为事物的信息（作为制品的监管文件）；Lector 处理作为过程的信息（技术性阅读、辅助解读）；PRISMA 则呈现作为知识的信息（经语境化、为专业决策而组织的理解）。Buckland 的贡献与其说是一种类比，不如说是一种告诫：混淆这三种含义的系统，往往会把文档当作知识，或把知识当作文档；这两者都是截然不同的信息失效形式。

与此相呼应，Capurro 和 Hjørland 将诠释学传统应用于信息科学 [36]，Myers 将其应用于信息系统 [37]，为这一论点提供了支持：保存的文本、解释与应用之间的区分并非技术上的权宜之计，而是认识论上的必要性。Ricoeur 所阐述的诠释学弧包含三个相互构成的时刻：说明（将文本作为自主对象进行结构分析）、理解（通过文本世界与读者世界之间的中介发现意义），以及占有（将解释内化到读者的行动视域中）[37, 38]。Ricoeur 坚持认为，文本一旦通过书写固定下来，便获得了独立于作者及其原始创作语境的自主性，成为一个本身即可解释的对象。

这一立场直接奠定了 PATOS 的文献主权原则：保存的文档并非作者思想的呈现，而是一种自主的人工制品，可以也应当在不同的语境条件下被阅读、重读和解释，而不损害其完整性。

在当代信息哲学中，Floridi 提出了抽象层次（LoA）方法。根据这一方法，任何系统都可以在不同层次上进行合理分析，而每个层次均由其自身的一组可观测量和操作所定义 [39]。同一系统的不同 LoA 并非相互竞争的描述，而是互为补充的视角，各自适用于特定类型的分析或行动。这种形式化为多层设计提供了最严谨的哲学理由说明：PATOS、Lector 和 PRISMA 既非冗余，也不相互竞争，而是在同一药学信息现实之上，分别处于真正不同的抽象层次，各自拥有其可观测量、合理操作和有效性标准。

2.7 规范类型化、言语行为与证据判定

药学信息本质上具有规范性：药品说明书并非只是描述药物的性质；它还规定行为、警示风险、禁止某些用途，并且在某些情况下，对临床相关问题保持沉默。这种言外行为异质性（即同一文档中的不同断言具有截然不同的交际效力这一事实）需要特定的理论基础。

由 Austin 提出并经塞尔形式化的言语行为理论，将每个话语区分为三个维度： 言内 行为（产生有意义的表达）、 言外 行为（说话时所实施的行为， 例如：断言、警告、请求、禁止），以及言后 行为（对交谈对象产生的影响） [40]。塞尔将这一划分进一步细化为五类言外行为： 断言类 （使说话者对命题的真实性作出承诺）、指令类 （试图促使交谈对象采取行动）、 承诺类（说话者的承诺）、 表达类（心理状态的表达），以及宣告类（仅凭宣告行为本身即创造某种事态） [41]。在药学信息的语境中，治疗适应证属于监管 断言性；禁忌证属于指令性 （指示专业人员予以避免）；警告结合了断言性和指令性；而规范性沉默（对于临床相关问题缺乏监管方面的明确表述）则以不作承诺的形式，具有其自身的言外之力。这一分类体系构成了 Lector 所执行的断言类型划分的基础，而 PRISMA 则以差异化形式呈现了这一划分。

此外，应用于规范系统的道义逻辑传统，

由 Navarro 和 Rodríguez 系统化 [42]，其基础源自 Alchourrón 和 Bulygin 的开创性研究 [43]，，将规范性命题与描述性命题之间的区别形式化，表明规范系统（由义务、许可和禁止构成）具有与描述性系统根本不同的逻辑结构。这一传统引入了 规范性缺口这一概念，指规范系统既未许可也未禁止某项行为的情形，从而产生真正的不确定性 [44]。对于药学信息而言，这一概念具有直接的相关性：药品说明书未提及某种药物相互作用，就会造成规范性缺口；保持沉默 并不 等同于许可或安全。未明确禁止与明确许可之间的形式区别，恰恰是基于概率模型的系统往往会混淆的区别，也是本文提出的架构力求保留的区别。

哈贝马斯的交往行为理论，经 Ngwenyama 和 Lee [45]以及 Mingers [46]，应用于信息系统后，增添了一个维度。哈贝马斯认为，每一种以理解为导向的言语行为都会同时提出三项有效性主张：真实性 （命题与外部世界相符）、规范正当性 （行为符合社会认可的规范），以及真诚性 （说话者真实表达其内在状态） [47]。Ngwenyama 和 Lee 指出，信息丰富度并非内容的固有属性，而是取决于提出并证成有效性主张时所处的社会语境 [45]。药品专论同时主张事实真实性（药品数据）、规范正当性（法规符合性）和机构真诚性（制造商的披露义务）。这三重要求解释了为何 RPDA 框架（第 3.4 节）的不同维度需要从不同视角审视同一信息。每个维度都会优先关注不同的有效性主张，以适应决策情境和所涉及的专业责任。

最后，循证医学传统为证据包概念提供了具体依据。Sackett 将循证实践定义为“在做出有关个体患者照护的决策时，审慎、明确且明智地使用当前最佳证据” [48]，这一表述本身就要求将证据基础（保存）、批判性评价（解释）和语境化临床应用（呈现）区分开来。GRADE（建议分级、评价、制定与评估）框架从这一视角出发，进一步明确证据质量与推荐强度是彼此独立的维度：强建议可以基于中等质量证据，反之亦然；这也说明解释层必须对药学信息进行限定，而不能只是简单传递 [49]。

2.8 药学实践中的人工智能

综合性综述表明，人工智能已日益应用于药学实践的多个方面，包括药物相互作用检测、处方审核、自动化调剂、剂量优化和治疗依从性 [1, 2, 50]。较新的评估显示，人工智能已对临床药学和医院药学产生直接影响，减少了错误并提高了运营效率 [3]。

在医院语境下，研究报告的应用主要集中于识别可能不适当的处方、预测不良事件以及高风险患者优先排序 [51, 52]。然而，系统综述指出，这些研究存在较高的方法学偏倚风险，且依赖本地数据库，从而限制了结果的普遍适用性 [53]。通过自然语言处理技术挖掘非结构化电子健康记录数据，进一步凸显了人工智能作为大量临床信息与以患者安全为导向的药学实践之间中介的作用 [54]。值得注意的是，针对儿科等特定语境下临床决策支持系统的综述显示，基于警报的系统的驳回率在63%至95%之间，这表明，当中断与阻断模式缺乏适当的认知中介时，专业人员往往会系统性地忽略警报 [55]。

大语言模型的出现为这一领域带来了新的维度。实证评估表明，在药物审查、相互作用识别和治疗方案核对等特定任务中，尤其是在边界明确的场景下，这些模型的表现可与人类药师相当 [6, 7, 56]。然而，基准测试研究显示，不同模型的表现参差不齐；随着临床复杂性增加，模型稳定性下降；而且，生成的建议始终缺乏明确的理由说明和可追溯性机制 [8]。这些发现进一步表明，尽管语言模型大幅提升了药学实践支持系统的语义能力，但也加剧了与治理、可解释性以及解释责任界定相关的挑战。一项针对患者护理领域89项大语言模型（LLM）研究的系统综述证实了这一情况，并指出其在设计 （数据不透明、缺乏针对医学领域的优化）和输出 （76.4%的研究存在不完整性，87.6%存在错误性，且幻觉与虚构是两种不同而又反复出现的现象）方面均存在结构性局限 [57]。

尽管取得了这些进展，现有文献却惊人地一致地指出了一个关键点：人工智能普遍被视为决策支持工具，而非专业判断的替代品。实证研究显示，人们对人工智能应用的态度总体积极，同时明确要求进行人工监督、确保透明度并实施机构控制 [58]。更具主张性的

研究提出了“药学情报”等概念，并承认核心挑战不仅在于算法性能，还在于如何组织支撑药学实践的信息生态系统[59]。

MIT NANDA 关于商业领域人工智能现状的报告从大规模实证角度佐证了这一发现。报告揭示了作者所称的生成式人工智能鸿沟：一些组织能够从人工智能中提取真正的价值，而绝大多数组织仍困在试点项目中，无法产生可衡量的影响[12]。医疗保健与制药 行业明显处于这道鸿沟的不利一侧，颠覆指数仅为 0.5 分（满分 5.0 分），与金融服务和先进制造业并列最低。观察到的迹象仅限于文档与转录试点，没有证据表明临床模型发生了结构性变化。报告指出，涉及复杂内部逻辑、不透明的决策支持，或基于专有启发式方法进行优化的工具——而这恰恰是临床决策支持系统的特征——属于最难推广应用的类别之一。这一独立于制药领域得出的实证发现，进一步印证了本文的核心假设：问题不在于缺乏能力强大的算法，而在于缺乏能够连贯地组织这些算法所处理信息的基础设施。

2.9 投入生产环境的药学人工智能系统：NoHarm.ai 案例

学术文献之外，我们还可以看到投入生产环境的药学人工智能系统日渐成熟，并产生可衡量的临床和机构影响。一个尤为典型的例子是 NoHarm.ai，这是一个专注于医院药学的巴西临床决策支持系统，已在公立和私立机构中得到广泛采用。该系统可实时自动分析电子医疗处方，识别药物相互作用、剂量不当、治疗重复、过敏、给药途径不兼容以及高风险患者特征，并通过 HL7 和 FHIR 等标准直接与医院系统集成。

与 NoHarm.ai 相关的科学研究表明，处方错误率大幅降低，药学审核覆盖率显著提高，即使临床团队规模没有扩大亦是如此；这证明了人工智能有望增强临床药师的专业能力[4, 5]。该项目还因其运营规模而备受瞩目，每月分析数百万张处方，并获得了国际认可。

从概念层面来看，NoHarm.ai 代表了面向患者

安全与临床效率的运营型药学决策支持系统的先进水平。然而，与文献中描述的大多数解决方案一样，其范围侧重于实时风险检测与缓解，并未明确涉及药品监管知识的结构化组织、版本化文档保存或形式化可解释证据链的构建。

2.10 药房信息系统中系统性风险的证据

除了功能局限之外，健康信息学文献还提供了经验证据，表明若药房信息系统在未采用专为满足社会技术要求而设计的信息架构的情况下融入临床环境，便可能引入意料之外的系统性风险。Rohani 和 Yusof 指出28使用药房信息系统所带来的意外后果，包括系统僵化、文档碎片化、认知过载以及机构内部的变通做法等因素会逐渐降低信息质量，并增加用药错误的风险 [9]。

具体证据进一步证实，电子处方系统不仅会导致未被发现的错误，还会直接引发错误。Brown 等人发现，计算机化医嘱录入系统会诱发八类相互关联的错误，包括下拉菜单中的选择错误、被动接受默认选项、药物信息在多个局部屏幕间显示碎片化，以及变通做法规避安全机制；这表明，以人为中心的设计缺陷会直接转化为临床风险 [60]。此外，Feather 等人指出，在电子处方中记录临床适应证——这是一项卫生主管部门建议采用的做法，旨在将每种药物与其治疗依据相关联——在实践中仍不常见；相关实施尝试还面临冗长的选择列表、处方者抵触，以及初级专业人员采取变通做法等问题，导致适应证错误，削弱处方决策的可追溯性 [61]。处方药物与其记录在案的临床依据之间这种脱节，从操作层面清楚地体现了行动与证据之间的碎片化，而本文提出的基础设施正是要从结构上解决这一问题。

这些发现印证了早期针对健康信息系统诱发错误的研究，并进一步表明，在脆弱的信息层之上引入自动化往往会放大潜在故障，将局部局限转化为系统性风险 [10]。对于基于人工智能的系统而言，这些结构性故障往往会在新的自动化组件与原本已不匹配的信息层耦合时扩散并放大。

这些证据并不构成反对数字化或

在药学实践中使用人工智能的论据，而是指出需要建立能够明确处理信息溯源、证据可追溯性，以及清晰界定人类与系统之间责任的基础设施。

2.11 综合与贡献定位

综观相关文献可以发现，尽管单点人工智能应用领域发展迅速、日趋成熟，但在规范性信息基础设施方面仍存在结构性不足，无法妥善保存文档、版本、监管背景和明确的证据链。大多数现有方法处理的都是已经转换过的数据，无法保证一手来源、解释和建议之间得到正式区分。从 Buckland 的三分法 [35] 到应用于信息科学的诠释学弧 [36–38], 从言语行为理论 [40, 41] 到规范系统的道义逻辑 [42, 43], 从循证医学 [48, 49]到应用于信息系统的交往行为理论 [45, 47],, 本文回顾的认识论基础一致表明，这种区分并非技术上的便利之举，而是认识论和规范上的必要性。

正是在这一领域，PATOS / Lector / PRISMA提案应运而生。该提案并不将人工智能视为监管或临床权威的替代品，而是将其视为一种基础设施的组成部分，旨在以负责任、可追溯且可审计的方式解释药学知识；它并非单点人工智能应用，而是面向新一代卫生信息基础设施的结构性提案。

3. PATOS–Lector–PRISMA 架构

3.1 设计原则与认识论基础

PATOS–Lector–PRISMA 架构的构想源自一项明确原则：将三种认知与机构运作功能分离开来。在当前药房信息系统实践中，这些功能往往被合并到单一技术层中：文档保存、语义解释以及语境呈现。这种分离原则分别得到以下理论的共同支持：Buckland 的三分法，即将信息区分为作为事物的信息、作为过程的信息和作为知识的信息 [35], ；诠释学弧中的说明、理解与占有 [37, 38], ；以及 Floridi 的抽象层次方法 [39]。这些运作各自都有对治理、时间性、责任和可审计性的要求，并且

它们的融合（这是现有大多数系统的典型特征），正是本文所主张的反复出现的脆弱性之根源，例如来源追溯丧失、解释不透明性，以及机器、专业人员与机构之间责任边界的模糊。

该架构采用五项指导原则：

1. **文档主权**：监管、临床和规范性来源均以其原始完整性予以保留，明确标注版本，并受到保护，免遭破坏性转换。任何解释层或呈现层都不得更改、重写或替代源文件。

2. **辅助而非自主阅读**：技术文档的解释（包括适应证、禁忌证、限制、警告、证据和规范性沉默的识别）由技术性阅读环境执行，该环境将计算能力与明确的人工策展相结合。计算模型协助进行文档组织与分解，但绝不承担规范性或临床权威。

3. **受治理的情境化呈现**：任何断言都不会在缺乏有效的权威语境、范围、目的和风险的情况下呈现给最终用户。系统不回答“什么是真实的”，而是回答“此时此地作出什么断言才有效”，并精确区分受监管适应证、有证据支持的标签外用途、获授权的孤儿药用途、关键限制，或仅仅是缺乏依据。

4. **完整可追溯性**：呈现给最终用户的每项断言都附有正式证据链，包括标明一手来源、文件版本、提取信息所依据的具体文本节点，以及得出该断言的解释轨迹。

5. **人类与机构责任**：系统不作决定、不提供处方，也不予授权。它保存来源、组织解释空间，并记录知识建构轨迹，从而将责任牢牢置于人类与机构领域。

总而言之，这五项原则构成了一种可称为治理优先架构的架构：在这种架构中，来源完整性、策展验证和双向可追溯性是首要设计约束，而非下游的合规要求 [62]。与那些将治理作为机构叠加层事后补上的、以输出为导向的系统不同，在这里，治理是组织原则，所有下游呈现均由此衍生。

这一架构的核心突破并非技术性的，而是认识论层面的：该模型能够为复杂的健康决策提供信息支持，同时不取代监管角色或临床判断，而很少有系统

能够在全球范围内连贯、可辩护且大规模地做到这一点。这一立场直接植根于社会技术系统文献，该文献表明，人类行动始终是情境化的、受语境影响的，并且在一定程度上不可预测，因此，将权威完全委托给自动化系统并不妥当[10, 24]。近期研究证实，随着大语言模型的出现，这一担忧日益加剧；研究指出，以技术为中心的方法试图复制而非支持卫生专业人员的认识功能，因此从根本上使人工智能系统与临床实践格格不入[63]。

3.2 PATOS：主权文档记忆

PATOS（官方来源的保存与可追溯性）[2] 构成了该架构的文档保存层，其设计理念是打造一个以 OAIS 模型 [16]和可信数字存储库指南 [17]为指导原则的可信数字存储库。

当前 PATOS 实现采用结构化文档数据湖，分为四个逐步成熟的阶段：**原始**（原始文档，逐位保存）、**已清理**（规范化格式与编码，不改变内容）、**结构化**（元数据提取、章节识别与索引），以及**精选**（在监督下进行丰富处理，并关联至规范实体）。每份文档都会获得唯一标识符、加密校验和（SHA-256）以及来源元数据，以便完整重建经口摄入和转换的历史。

该存储库汇集来自巴西多个监管和机构来源的文档，包括巴西国家卫生监督局（ANVISA）和参考医院数据库，涵盖专业人员和患者版药品说明书、专论、产品特性概要（*Sm-PCs*）、临床方案及规范性文件。该系统的首个运行版本包含超过 17,000 份药学文件，涵盖在巴西注册的全部参考药物，并符合 RDC（集体董事会决议）47/2009 的文档要求。

PATOS 的核心架构原则是原始层不可变性：源文件一经摄取，便不会被修改或替换。同一文档的后续版本会在存储库中共存，从而能够重建历史并对监管变更进行时间分析。这一设计决策使 PATOS 有别于会覆盖先前记录的数据库，并确保

[2]这个名称也让人联想到*Lagoados Patos*，巴西最大的潟湖，位于南里奥格兰德州，本研究正是起源于此。其主要功能是经口摄入、保存、版本管理并提供监管、临床和规范性文档，同时保持其原始完整性，不预设任何解释或语义转换。

每一种后续解释都可以根据其所依据的确切源文件进行核验。

3.3 Lector：辅助技术阅读与证据记忆

Lector 是该架构的第二个功能阶段，定位为机器辅助技术性阅读与明确人工策展的环境。其构想借鉴了计算式意义建构 [18]和探索式搜索 [29], 的传统：它并非将监管文件的解释视为自动提取操作，而是将其视为一种迭代认知过程，结合计算式分解与专业人员的专业判断。

Lector 的操作链包含四个正式概念：

1. **文档引用**：指向保存在 PATOS 中的文档，包括标识符、版本、来源元数据和完整性校验和。此引用确保每项阅读操作均以可验证且不可变的来源为依据。

2. **PageIndex**：将文档层级分解为可导航的单元，由语言模型生成（该架构与模型无关；当前实现使用 Mistral 7B），并在参数化监督下完成。*PageIndex* 生成由各级章节构成的树状结构，并附上语义摘要，从而支持非线性导航，精准定位相关内容，无需通读全文。

3. **证据包**：正式的证据包装单元，详见小节 3.6。每个证据包都封装了一个合格问题、一份有依据的答复、对一手来源的引用（包括特定*PageIndex*节点的标识符）、明确的认知边界以及策展元数据。

4. **CuratorialDecision**：关于人工对证据包作出决定的正式记录，包括接受、拒绝或要求修订，并附有书面理由说明。该产物体现了这一原则：诠释权威仍属于人类领域，而计算模型的作用是辅助阅读，而非充当规范性权威。

Lector 遵循一项明确的内部规范：模型仅从已经获得认可的人工策展中学习。这意味着计算系统绝不会生成可成为参考依据的自主解释；进入系统知识体系的每一项断言，都经过了有文档记录的人工验证。这种刻意的限制从可扩展性的角度看或许显得有所束缚，但实际上，它正是在解释错误可能造成严重临床、法律和监管后果的领域中，保障系统机构可靠性的机制。

Lector 会识别并界定监管文档中出现的不同类型断言，包括：受监管适应证、绝对或相对禁忌证、给药剂量、药物相互作用、警告、注意事项、特殊人群用药，以及规范性沉默（即对临床相关问题明确未作出监管表述）。这种分类至关重要，因为药品陈述的效力和范围，关键取决于其规范性质（按照塞尔的术语，每种类型都对应一种不同的言外行为 [41]）；将这些区别一概归入笼统的“信息”类别，恰恰是会在决策支持系统中引发风险的简化方式。尤其是，将规范性沉默明确作为一种正式类别处理，而非视为单纯的数据缺失，这与道义逻辑中规范系统的规范性缺口概念相一致 [42–44]。在这一框架下，规范系统的不确定性并不等同于许可或禁止，而是一种具有认识论意义、应当如实传达的状态。

3.4 RPDA 框架：药学信息的情境折射

RPDA（监管、处方、调剂与给药）构成了 PRISMA 呈现层的核心信息组织框架。它的作用并非描述诊疗流程（这一职责已由文献中归纳的 用药流程 承担 [55]），而是将一种情境化呈现 原则形式化：应根据决策语境、专业人员画像以及行动相关风险，对相同的药学信息进行不同的组织、筛选与呈现。

区分流程模型 与信息架构 是理解该提案的关键。用药流程描述药物在整个诊疗周期中发生了什么；RPDA 则描述在该周期的每个阶段，应如何构建药物相关信息的结构并加以呈现，以支持专业人员决策，同时不混淆语义、治理或可追溯性。

该框架基于三个形式化概念：

信息核心：一组结构化、规范化且可版本化的数据，用于描述药物，不依赖于任何特定用例，并在时间推移中保持可追溯性和语义一致性。在所提架构中，信息核心由 PATOS 保存的文档以及 Lector 生成的合格语义单元构成。

信息视图：信息核心的一种语境化投影，按特定专业目的组织，仅呈现相关属性，

证据和断言，其粒度与治理方式均与所涉风险相匹配。

风险导向的粒度：指所呈现信息细节的程度应与其所支持决策可能产生的影响成正比，既要避免关键内容的遗漏，也要避免无关信息过多。文献已广泛记录，这种现象会导致警报疲劳和临床注意力退化 [55, 60]。

每个 RPDA 轴都定义了一种独特的决策情境：

- **监管**：信息经过组织，以满足注册、合规、可追溯性和监测方面的正式要求。-其优先呈现官方标识符、版本历史、监管状态、规范性链接和法律框架。典型的专业人员是监管药师或法规事务经理。

- **处方**：重新组织信息，以支持治疗决策，重点呈现适应症、药物剂型、浓度、给药途径、使用限制及相关临床证据，同时确保语义清晰并减少歧义。典型的专业人员是处方医师，或从事用药核对活动的临床药师。

- **调剂**：信息具有操作性，着重呈现市售包装规格、包装、分拆、治疗等效性和安全替代的可能性，同时尽量降低物流风险和供应错误。典型专业人员是调剂药师或医院药房技术员。

- **给药**：信息视图经过细化，以支持安全的药物配制和治疗实施，涵盖给药剂量、复溶、稀释、稳定性、储存、特定给药途径的配制说明以及实用的给药指导。典型专业人员是负责药物配制的药剂师和负责给药的护士。

RPDA 不会创建新数据，而是根据具体情境组织并呈现现有信息，使内容、语境与专业责任相协调。这种差异化处理的必要性，根植于交往行为理论；Ngwenyama 和 Lee [45]以及 Mingers [46], 将该理论应用于信息系统。根据这一理论，每一种以理解为导向的交往行为都会同时提出关于真实性、规范正当性和真诚性的主张 [47]：每个 RPDA 轴都会优先关注不同的有效性主张，并使其与决策情境及所涉风险相适应。此外，言语行为理论 [40, 41] 为依据言外之力对药学断言进行分类的必要性提供了依据，具体分类包括

断言类、指令类和承诺类；RPDA 通过根据与相应语境相关的主要断言性质来组织各个视图，从而将这种分类付诸实践。

从功能上看，RPDA 作为一种折射机制，使 PRISMA 能够将同一信息核心投射到不同视图中，这有意类比于这样一种光学现象：光束穿过棱镜后被分解为各个组成部分，而原始光线并未发生改变或消失。

3.5 PRISMA：情境化呈现与情境图

PRISMA（药品折射与用药保障信息系统）构成该架构的情境化呈现层，负责将药学知识组织为带有版本的断言，并附加正式证据链，同时将其投射到 RPDA 框架所定义的信息视图中。

PRISMA 的设计建立在**上下文图**之上。这些结构明确表示药品陈述有效的条件，包括支持该陈述的规范性来源、适用的监管范围、目标人群、相关临床状况、用于调节其强度的限定条件（例如“适用于成人”“重度肝功能不全者除外”“依机构方案执行”），以及已知的认知边界（图1）。这种方法呼应了 Brézillon 的情境图 [33] 以及 Toulmin的计算论证模型 [23]，，在这些模型中，断言的有效性并非绝对，而是取决于明确的论证依据、限定条件和反驳。

具体而言，当卫生专业人员向 PRISMA 查询某种药物时，系统不会返回笼统、未经区分的答案，而是提供一种语境化呈现，其中：

- 识别当前RPDA轴（监管、处方、调剂或给药），并调整相应的信息视图；
- 针对该语境选择相关的 证据包，按类型（适应证、禁忌证、给药剂量、警告等）及其与决策场景的相关性筛选断言；
- 呈现每项断言时，均附上其溯源链：来源文档、版本、具体章节、断言类型及策展状态；
- 明确指出认知边界：来源之间的分歧、文献缺失、临床情境依赖，以及规范性沉默的领域。

当前处于实施阶段的 PRISMA原型被组织为

五个与 RPDA 轴相对应的界面模块：**搜索** （跨领域搜索）、**监管** （合规与可追溯性）、**处方**（治疗决策支持）、**调剂** （操作与等效关系）以及**给药**（安全执行）。每个模块都以相应视图呈现相同的信息基础，在实践中体现了情境折射原则。

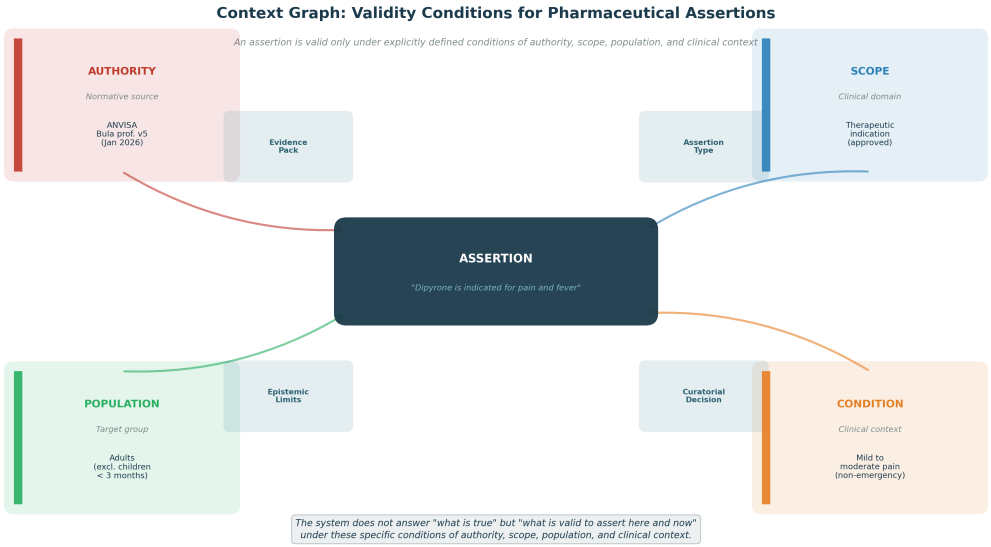


图 1. 情境图结构展示了药学断言的四个有效性维度。每个断言只有在明确界定权威来源（规范性来源）、范围（临床领域）、人群（目标群体）和临床情境的条件下才有效。示例展示了断言“安乃近适用于疼痛和发热”及其具体有效条件。限定节点（EvidencePack、断言类型、认知边界、策展决策）在断言与其语境维度之间发挥中介作用。

3.6 EvidencePack：可问责断言的形式化单元

证据包 概念是连接三个架构层的核心机制。证据包 被定义为一种正式的药学证据封装单元，由以下部分组成：

- **合格问题**：对该包装 所回答问题的类型化表述，依据药学问题类型分类体系进行分类（适应证、禁忌证、给药剂量、相互作用、特殊人群用药等）。问题类型化至关重要，因为它决定了预期的响应类型、适用的验证标准以及相关治理。

- **有依据的答复**：根据对源文件的技术性阅读构建的断言，以专业语言撰写，并附有

明确的限定条件，以界定其范围和有效条件。

- **来源溯源链**：指向 PATOS 中保留的源文件的可追溯引用，包括文档标识符、版本、校验和，以及提取信息所依据的特定 *PageIndex* 节点的标识符，从而能够根据支持该断言的确切来源对任何断言进行验证。

- **认知边界**：对已识别限制的明确声明，包括：不同来源针对同一主题存在的分歧、文档信息缺失（不可获得或未找到的信息）、影响断言有效性的临床情境依赖，以及规范性沉默领域，即监管来源未对临床相关问题作出明确表述的情况。

- **策展状态**：记录包装在策展工作流中的验证阶段（草稿、审核中、已接受、已拒绝），并注明负责的策展人及形成文档的理由说明。

证据包 在操作层面具体体现了数据、解释与呈现之间的分离，这一分离构成整个架构的基础；它将循证医学关于每项临床决策都必须由明确证据支持、经过批判性评估并以情境化方式应用的要求落实到制药领域 [48]。GRADE 框架 [49]，所确立的证据质量与推荐强度之间的正式区分，通过有依据的答复（证据支持的内容）与认知边界（证据具有何种可信度以及适用条件为何）之间的分离，纳入证据包 。

EvidencePack 并非“系统响应”，而是一个可审计的药学知识单元，其构建过程可追溯、验证过程有文档记录，且其局限性明确。这一特性使其与 RAG 系统或对话代理生成的响应有着根本区别，因为后者生成后通常无法恢复信息溯源。

3.7 信息流与架构集成

图2 展示了三层基础设施及其信息流的架构概览。三层之间的集成遵循单向且可审计的信息流：

1. **摄取与保存** (PATOS)：监管和临床文件经摄取后，以其原始完整性予以保存，并进行版本管理和索引。原始层保持不可变；后续转换均记录完整的来源溯源信息。

2. **阅读与策展 (Lector)**：保存的文档被拆解为可导航的树状结构（*PageIndex*），通过计算处理与人工策展相结合的方式进行技术性阅读，并转化为合格的证据包。计算模型辅助阅读；人工验证解释。

3. **折射与呈现 (PRISMA/RPDA)**：精选 证据包 会依据当前启用的轴、专业人员画像及相关风险，整理为情境图，并投射到 RPDA 定义的信息视图中。呈现始终伴随溯源链和明确的认知边界。

这一流程确保呈现给最终用户的所有信息，都有可验证的来源作为依据，都经过有文档记录的阅读阶段，并且都经过明确的语境化处理。颠倒或绕过这一流程（例如，未经整理阶段便直接根据文档生成答复，或呈现缺乏可追溯来源溯源的断言）会被视为架构违规，而不是可接受的简化。

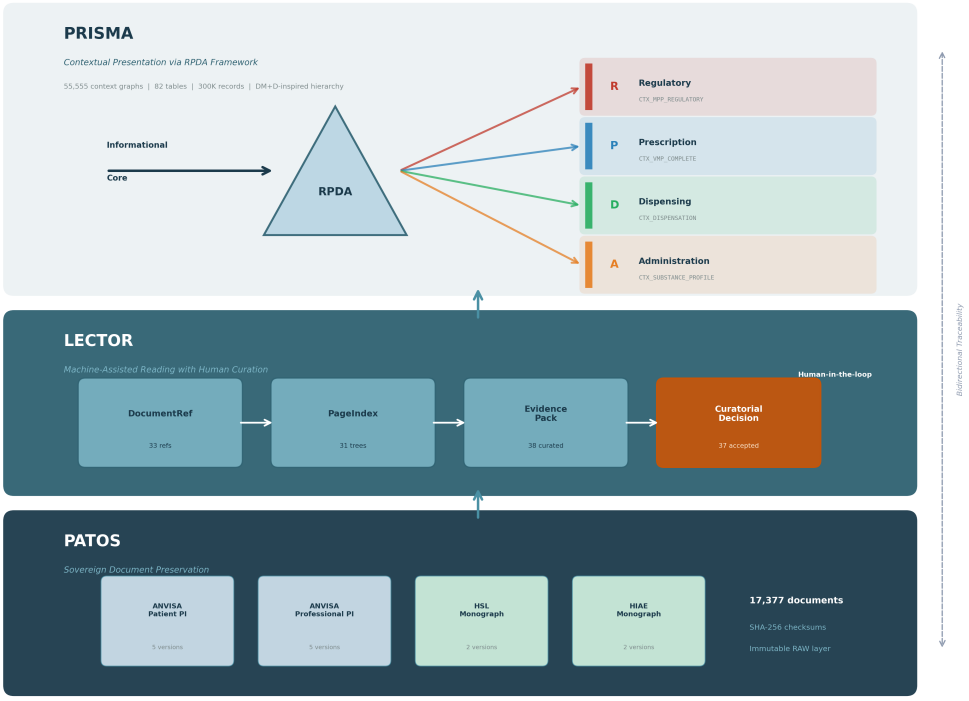


图 2. PATOS–Lector–PRISMA 基础设施的架构概览。底层（PATOS）以不可篡改性、版本管理和密码学校验和保存监管与机构文件。中间层（Lector）实现机器辅助阅读，生成由人工策展人员验证的证据包。顶层（PRISMA）通过 RPDA 框架，将信息核心折射为四种情境视图（监管、处方、调剂、给药）。实线箭头表示单向信息流；虚线箭头表示从任一断言回溯至其源文件的双向可追溯性。

从实现成熟度来看，重要的是透明地说明这三层的成熟度各不相同：PATOS 已投入运行，拥有一个包含 17,000 多份文档的活跃存储库；Lector 正在积极开发中，前两个阶段已经完成，第三个阶段仍在进行；PRISMA 则处于高保真原型设计阶段，已开发出展示情境折射原则的用户界面，但后端集成尚未完成。这种透明度是有意为之：本文提出了一种概念架构及其指导原则，以成熟的学术传统为基础，并通过部分实现得到验证，但并未声称已建成一个完整定型的系统。

3.8 演示示例：一水合安乃近贯穿三层

为使前文所述流程具体可感，本小节将以一种具体药物——一水合安乃近（Novalgina，巴西国家卫生监督局注册号 186200018）为例，结合操作系统中的真实数据，追踪其在三个架构层中的流转过程。

选择安乃近（metamizole）作为说明案例，是经过审慎考虑且具有认识论意义的。它是在巴西广泛使用的一线镇痛药和解热药，无需处方即可购买，每年给药量超过 215 百万剂；仅圣保罗州的年消费量就达 488 吨（ 2016 [64]）。矛盾的是，由于安乃近与粒细胞缺乏症有关，它在美国、英国和法国被禁用，在其他欧洲国家也受到限制。然而，来自巴西的大规模药物警戒数据，包括一项使用真实世界数据、涉及 384,668 名患者的回顾性研究，在队列中未发现粒细胞缺乏症病例，并报告称，安乃近的不良反应发生率比对乙酰氨基酚低38.8%，比乙酰水杨酸低 46.8% [65]。拉丁美洲多中心研究估计，与安乃近相关的粒细胞缺乏症发生率为每百万人年 0.38 例，显著低于促使其他国家实施禁令的历史估计值 [66]。

司法辖区之间的这种监管差异，使安乃近成为所提架构的典型案例：关于这种药物的药学信息，极大程度上取决于其产生和解读时所处的规范性语境。若不加甄别地引入国际知识库，基础设施就会将安乃近归类为高风险或禁用物质，从而否定数十年来的巴西药物流行病学证据以及巴西国家卫生监督局的自主监管决定。PATOS 通过保留具有明确来源溯源信息的巴西监管文件，Lector 则通过生成以这些国家级来源为依据的证据包，确保药学

PRISMA 所呈现的知识能够反映临床决策实际制定时所处的监管和流行病学语境。

3.8.1 第 1 层，PATOS：文档保存

PATOS 存储库收录了 192 份与安乃近相关的文档，来自四个独立来源：巴西国家卫生监督局的官方药品说明书（12 份文档，包括患者版和专业人员版）、Hospital Sírio-Libanês 和 Hospital Albert Einstein 的医院专论（12 份文档），以及公开的药品信息数据库（168 份文档，涵盖多家制造商的 170 种商业包装规格）。

对于 Novalgina 注册号（编号 186200018），PATOS 保存了五个完整的历史版本药品说明书，时间跨度为 2024年2月 至 2026年1月，每个版本均包含患者版和专业人员版。当前版本（`186200018_20260116_pac.pdf`，SHA-256 校验和：d55ebcaf...a7f6）与先前版本并存，且不会覆盖它们。这种保存方式使人们能够追踪禁忌证何时被添加，或巴西国家卫生监督局何时修改警告措辞，从而重建该药品随时间推移的监管演变过程。

每份文档在摄入时都会获得一个唯一标识符（文档引用）。对于当前的 Novalgina 专业版说明书，文档引用记录以下信息：来源（巴西国家卫生监督局）、类型（专业版说明书）、药物（Novalgina）、活性成分（一水合安乃近）、格式（PDF）、采集日期（2026-01-28）、SHA-256 校验和以及时效状态（`is_当前 = 1`）。旧版本标记为 `is_当前 = 0`，但仍可在存储库中完整访问。

3.8.2 Layer2，Lector：辅助阅读与证据包

Lector 从 PATOS 中保存的文档出发，分三个阶段执行技术性阅读流程。

PageIndex. 语言模型（目前通过 Ollama 使用 Mistral 7B，但该流程与模型无关）将专业版说明书分解为分层章节树。对于 Novalgina 药品说明书，此分解会生成如下节点：

- 节点 1.1：“该药物适用于治疗什么？”（治疗适应证）
- 节点 1.3：“我什么时候不应该使用这种药物？”（禁忌证）
- 节点 1.5：“我应该如何使用这种药物？”（剂量与给药）
- 节点 1.7：“这种药物可能引起哪些不良反应？”（不良反应）

每个节点都会收到模型生成的语义摘要，从而支持非线性导航并精准定位相关内容。

证据包。 对于每种具有临床相关性的问题类型，Lector 都会生成结构化的 *EvidencePack*。以下是 Novalgina 的适应证类型证据包：

```
证据包 EP -001（适应证）+- - 问题："该药物的适应证是什么？"+- - 重点：一水
```

对每种断言类型（禁忌证、给药剂量、相互作用、不良反应）都会重复这一流程，从而生成一组证据包，共同构成该药物经策展的证据记忆。在当前实现中，系统针对五种参考药物包含 38 个经策展的证据包，其中 37 个已接受，1 个已拒绝，并关联至 119 个指向 PRISMA本体的规范链接。

将规范性沉默作为正式类别。并非所有临床相关问题都能从监管来源中得到明确答复。例如，Novalgina 专业版说明书并未对葡萄糖-6-磷酸脱氢酶（G6PD）缺乏症患者使用安乃近作出说明；对于其他镇痛药和解热药，这种状况在临床上具有相关性。在 PLP 架构中，这种缺失不会被忽略，而是被正式记录为一种 NORMATIVE_SILENCE 类型的证据包：

```
证据包 EP -039 ( NORMATIVE_SILENCE )+-- 问题 : " 安乃近
```

+- - 关注对象：一水合安乃近 500毫克片剂+- - 结论：未发现监管方面的说明+-

这种对沉默的形式化具有重要的架构意义：它将信息的缺失从不可见的空白转变为明确、可审计且可采取行动的类别，而概率模型往往会将这类细微差别简化抹平（第 2.7 节）。

策展决策。 每个证据包都要经过有文档记录的人类验证。策展人可以接受、拒绝或要求修订，并记录理由说明。原则十分明确：任何断言都必须经过人工验证，才能纳入系统的知识体系。这一限制确保来源溯源既有文档依据（来源确实存在），也经过诠释验证（对来源的解读已由具备资质的专业人员验证）。

3.8.3 第三层，*PRISMA*：规范本体与 *RPDA* 折射

精选证据包通过规范链接与 PRISMA 的药品规范本体相连，从而实现证据与规范实体之间的双向可追溯性。该本体采用了自身的层级结构，其设计受 DM+D（药品与器械词典，英国国家医疗服务体系，英国）模型启发，并扩展以适应巴西的监管实际，同时采用 IDMP（药品标识，ISO 11615/11616）中的通用标识符和概念，以实现国际互操作性。这种调整并非对外国模型的简单翻译，而是主权治理的一项审慎决策：层级的每一层都赋予了功能命名，以反映巴西的监管、临床和物流实践，同时保持与国际标准的正式映射能力。该规范层级包含 300,134 条记录，分布在 82 张表中，按具体化程度递进分为五个层级：

物质：一水合安乃近（SUB -000033943）

```
   +- - CAS: 5907 -38 -0 | UNII: 6429 L0L52Y | DCB: 9564
   +- - 17 个国际标识符（ChEMBL, DrugBank,
        PubChem , MeSH , ...）
   +- - 24 个多语言同义词（安乃近, 安乃近,
        安乃近, ...）
          v
虚拟治疗医药产品 -- 定性成分：一水合安乃近
     ( 虚拟治疗医药产品 -000010750)
   +- - 16种成分组成含有此物质（单方制剂 +
        复方制剂）
          v
虚拟医药产品 -- 剂型：一水合安乃近500毫克片剂
     ( 虚拟医药产品 -000051605)
   +- - 解剖学治疗学及化学分类系统：N02BB02
   +- - 药物剂型：片剂 ( PDF -000002766)
   |    +-- 欧洲药品质量管理局分类：口服 / 固体 / 经口摄入 /
        常规释放
   +- - 定量组成：500 mg 安乃近
        一水合物
   +- - 此物质有 7 种制剂
          v
VMPP -- 包装：安乃近 500 mg x 30 片剂
     ( VMPP -000103766)
   +- - 可处方单位：片剂
   +- - 内包装：泡罩包装
   +- - 13 虚拟 包装 ( x10 , x20 , x30, x50, x100, ...)
          v
安瓿 -- 产品：NOVALGINA 500 mg ( 安瓿, 当量。 IDMP
     药品 )
   +- - 制造商 : Sanofi Medley ( ORG -000000032)
   +- - 巴西国家卫生监督局注册 : PMA 183260351
   +- - 治疗类别 : 非麻醉性镇痛药
          v
AMPP -- 商品包装 : NOVALGINA 500 mg 盒装 x 30
     ( 等同于。 IDMP包装药品）
   +- - 欧洲商品编码：7891058008635
   +- - 标签：非处方药（无需处方）
   +- - 来自 24 家制造商的 92 种商品包装，适用于
        此剂型
```

表1 概述了规范层级、DM+D 模型与 IDMP 概念之间的对应关系：

RPDA 折射。 PRISMA 将这一信息基础投射为五种不同的情境视图，通过具体化的上下文图 呈现（共生成 55,555 个 JSON 图，耗时不到 0.45 秒）：

表1：PRISMA规范层级、DM+D 模型（英国国家医疗服务体系）与 IDMP概念（ISO 11615/11616）之间的对应关系。

PRISMA层级	功能命名	DM+D	IDMP（ISO 11615/11616）
物质	物质	物质	物质
VTM	定性成分	虚拟治疗部分	—
VMP	剂型	虚拟药品	药品
VMPP	Pack	虚拟药品包装	—
AMP	产品	实际药品	药品
AMPP	商品包装	实际药品包装	包装药品

- **监管视图** (CTX_MPP_REGULATORY)：展示贸易产品及其巴西国家卫生监督局注册（PMA 183260351）、授权状态、上市日期、标签分类、治疗分类、监管类别、储存条件（15–30°C，保质期为 24 个月），以及用于物流追溯的欧洲商品编码（EAN）。此视图面向需要核验合规情况的监管药师。

- **处方视图** (CTX_虚拟医药产品_完整)：展示含有物质、浓度（500 mg）、解剖学治疗学及化学分类系统（N02BB02）、DDD（每日规定剂量，0.167）、完整分类体系的药物剂型（口服、固体、经口摄入、常规释放）、定量组成及可用药品规格的虚拟产品。该视图面向需要在不同药物剂型和浓度之间做出选择的处方医师。

- **调剂视图** (CTX_调剂)：展示含有可处方单位（片剂）和包装详情（泡罩包装 × 30）的虚拟包装，并列出此配置下可用的 17 种商品药品及其品牌、制造商、欧洲商品编码、标签和上市状态。该视图面向需要从等效替代品中进行选择的配药人员。

- **给药视图** (CTX_物质_概况)：展示完整的物质概况（17 个国际标识符、24 个多语言同义词、DCB（巴西通用名称）代码，以及与 16 种虚拟治疗药品和 36 种虚拟药品的映射），旨在支持安全的药物配制、给药和治疗方案核对。该视图面向负责药物配制的药剂师、需要在给药前核实物质身份的护士，以及负责治疗方案核对的临床药师。

每一种视图都源自同一本体基础，而这一基础又由 Lector 精选的证据包支撑，这些证据包则以保存在 PATOS 中的文档为依据。追溯链至关重要：从呈现给专业人员的情境视图，到原始监管文件中的特定节点，中间经过将相关

解释形式化的证据包，以及对其进行验证的策展人。

已展示能力摘要。 表2 汇总了演示示例的定量范围，涵盖安乃近案例的三个架构层。

表 2。 演示示例的定量摘要：跨越三个 PLP 层的安乃近一水合物。

指标	层	值
已摄取文档	PATOS	192（4 个来源）
文档类型	PATOS	说明书、专论、数据库
已生成 PageIndex 树	Lector	31
已精选的证据包	Lector	38（37 个已接受，1 个已拒绝）
规范链接	Lector → PRISMA	119
本体记录	PRISMA	300,134（82 个表）
层级	PRISMA	6（物质 → AMPP）
生成的上下文图	PRISMA/RPDA	55,555（在 < 0.45 秒内）
RPDA 视图展示了 PRISMA		4
断言类型已形式化	Lector	7

4. 动机案例：与运营型决策支持系统的互补性

上一节将 PATOS–Lector–PRISMA 架构作为一项概念性提案加以介绍，该提案以成熟的学术传统为基础，并通过实施得到部分验证。为展示这项提案的实际意义，本节将探讨该基础设施如何与已投入生产的运营型药学决策支持系统形成互补，而非取而代之；本文以 NoHarm.ai 为动机案例。NoHarm.ai 是一个巴西系统，在医院药学领域得到广泛采用 [4, 5]。

4.1 运行系统的优势

NoHarm.ai 等系统在实时风险检测与缓解层运行：它们根据知识库分析处方，识别药物相互作用、剂量超出范围、治疗重复、不相容性及高风险特征，并生成警报以指导药学干预。校准得当时，这些系统能显著提升临床药师的专业能力，使人手有限的团队也能审核数量庞大的处方，而这些处方若仅靠人工审核，根本无法完成。

这是一项合理且经临床验证的运营贡献。

然而，它基于一个隐含前提：底层知识库可靠、及时、具有语境信息且可追溯。拟议的架构正是针对这一前提进行干预。

4.2 持续存在的结构性缺失

即使是最优秀的运营型药学决策支持系统，也存在一系列结构性局限。这些局限并非源于工程故障，而是因为缺少底层信息层：

缺少可追溯的文献来源信息。当系统发出药物相互作用警报时，专业人员通常无法查阅支持该警报的监管来源：具体的药品说明书、文件版本、确切章节，以及断言的原始语境。警报作为一项计算结果存在，却未锚定于其原始证据。若专业人员需要核实、说明或质疑该警报，就必须转而查阅外部来源，手动重建本应由系统保留的溯源链。

解释轨迹不透明。在原始监管文档与呈现给专业人员的警报之间，存在一条通常不可见的解释轨迹（信息抽取、术语规范化、严重程度分类、规则应用）。专业人员接收到最终结果（警报），却无法了解生成该结果的推理过程。在一个机构信任取决于可审计性的领域，这种不透明是系统性脆弱因素，而非可以接受的简化。

语境未区分。 同一药学信息（例如，相对禁忌证）在不同决策情境中的相关性可能截然不同：在监管轴上，可能需要正式注册；在处方轴上，可能需要进行风险获益评估；在调剂轴上，可能无关紧要；在给药轴上，则可能需要发出具体的监测指示。以未加区分的方式呈现信息的系统（在所有情境下都采用相同格式显示相同警报），错失了根据相关风险和责任调整呈现方式的机会。

对知识的二元化处理。警报系统主要采用二元逻辑：相互作用存在或不存在；剂量足够或不足；过敏已登记或未登记。这种做法难以自然地容纳监管药品知识的细微差别：有证据支持的超说明书适应证、依情境而定的限制、不同监管来源之间的分歧，或不等同于许可的规范性沉默。文献中记载的63% 至95% 覆盖率 [55] 或许部分源于这种不恰当的二元化

处理本质上具有渐进性且依赖语境的知识。

4.3 PATOS–Lector–PRISMA 基础设施如何补充运营周期

该提议并非要取代 NoHarm.ai 等运营系统，而是为它们提供底层信息基础设施，以弥补上述缺失。这种互补性可从四个维度加以阐述：

文档锚定。 借助 PATOS 作为保存层，运营系统生成的每条警报都可以关联到支持相关规则或所用信息的特定监管来源（文档、版本、章节）。收到警报的专业人员能够直接查阅原始证据，无需手动重建。

解释透明度。 借助 Lector 作为技术性阅读层，源文件与应用规则之间的轨迹将记录在可审计的 *EvidencePack* 中：向文档提出了什么问题、提取了什么答复、识别出哪些认知边界，以及由谁验证了解释。系统的不透明性将被可重建的解释轨迹所取代。

语境化重构。 借助 PRISMA 和 RPDA 框架，同一药学信息将根据当前的决策轴以不同方式呈现。如今，系统会向处方医师、配药人员和护士发出相同警报；而在此框架下，系统将根据专业人员画像、相关风险以及适合决策时刻的粒度，对信息进行语境化处理。

分级限定。 借助 Lector 的断言类型划分和 PRISMA 的上下文图，警报的二元逻辑将得到丰富，从而更忠实地呈现药学知识：受监管适应证与有证据支持的超说明书使用、绝对禁忌证与相对禁忌证、强有力的证据与脆弱的共识，以及监管声明的存在与明确缺失。

为具体说明这种互补性，可以考虑一个已基于第 3.8 节所述数据构建的场景：处方医师为一名正在同时使用华法林的患者开具 500 mg 安乃近。运营中的临床决策支持系统会发出二元药物相互作用警报。采用 PLP 基础设施后，同一警报将附带更丰富的信息：(i) 记录安乃近与华法林相互作用的具体证据包（证据包类型：INTERACTION），其依据为专业版说明书 v5（PATOS 文档 BULA_PROF_NOVALGINA_v5，已通过 SHA-256 验证）；(ii) 记录该相互作用

被分类为中度，并建议监测国际标准化比值，而非将其视为绝对禁忌证；(iii) 经 RPDA 折射、适合具体情境的呈现：处方医师看到药理机制和临床管理建议，配药人员看到调剂注意事项，护士看到监测参数。从警报到源文件的溯源链将可被完整重建。尽管这一场景尚未通过与实际临床决策支持系统集成进行运行验证，但所用数据实体均来自第 3.8 节所述操作系统。

4.4 关于互补性本质的说明

此处所述的互补性并非叠加式的（不是在现有系统上添加“更多功能”），而是基础设施层面的。这个区别很重要：附加功能与其所补充的系统处于同一架构层面；而基础设施则位于底层，提供运行系统可以利用、但无需在内部实现的可能性条件。

从这个意义上说，PRISMA 与 NoHarm.ai 这类系统之间的关系，类似于操作系统与应用之间的关系：应用需要操作系统才能可靠运行，但不会取代操作系统，操作系统也不会取代应用。缺少规范性信息基础设施并不会阻止运行系统运作（NoHarm.ai 就证明了这一点），但会从结构上限制这些系统在可追溯性、可解释性和机构问责方面所能提供的能力。

5. 讨论

5.1 架构层面的影响

本文提出的方案将关于药房中人工智能的讨论，从目前主要聚焦于算法性能、模型准确率和比较基准的 debate 转向对信息架构本身更根本的探讨。这一视角的转变并非无关紧要：它意味着我们必须认识到，许多被归因于人工智能模型或系统界面的局限，实际上是信息基础设施缺失或不完善所导致的。这一观点也在对更广泛的企业人工智能生态系统的分析中得到了相互印证。MIT NANDA 报告 [12] 表明，寄望于足够先进的模型能够弥补脆弱的数据、易出故障的系统以及定义不清的流程，这一设想已被系统性地证明是错误的：对于复杂度高的工作，90% 的专业人员仍然更倾向于选择人类，而非

并非出于对人工智能的普遍不信任，而是因为当前系统无法保留反馈、无法适应语境，也无法随时间推移而改进。值得注意的是，高管在评估人工智能系统时优先考虑的五项标准（信任、深入理解工作流程、最小干扰、明确的数据边界和持续改进能力），几乎与 PLP 架构的五项设计原则直接对应：文献主权、策展式辅助阅读、受治理的情境化呈现、完整可追溯性，以及人类和机构责任。Ramesh [13], 在评论这些发现时，提出了五项供组织进行诊断性自我审视的测试，包括营地测试（“引入人工智能后，系统是否变得更健康？”）以及激励测试（“谁因基础性工作而获得认可？”）。将这些测试应用于制药领域后，它们进一步印证了这一论点：挑战在于基础设施，而非算法。

保存（PATOS）、解释（Lector）与呈现（PRISMA）之间的形式化分离，不仅仅是对功能进行模块化组织；它反映了三种认知与机构运作之间的认识论区分，这三种运作在治理、时间性、可审计性和责任方面各有不同要求。将这些运作融合起来（这是目前大多数现有系统的做法），并非无害的简化，而是系统性脆弱性的一个明确来源：来源溯源丧失、解释不透明性、警报疲劳，以及机器、专业人员与机构之间责任边界的弱化。

与此相呼应，Talisman [34]指出，业界已在语义层上投入大量资源以进行指标治理，但解决的问题可能并不正确：人工智能系统需要的是情境工程（概念、关系、约束和推理能力），而非标准化计算。作者指出，医疗保健和生命科学已经证明，形式本体能够大规模应用；竞争优势属于那些将知识架构视为核心能力的组织。在后续研究中，同一位作者 [14] 进一步深入分析，指出大语言模型会从结构上不可逆地破坏来源溯源，瓦解文明历经数世纪建立起来、用于区分知识与单纯断言的文献链。这两项分析分别独立从数据工程和元数据领域提出，其汇合从两个方面印证了本文的核心主张：药学信息面临的挑战，既关乎知识架构（表示什么以及如何推理），也关乎来源溯源基础设施（如何保留赋予断言认识论合法性的链条）。（我们指出，这些来源 [13, 14, 34] 属于面向从业者的分析，而非同行评审出版物；引用它们是为了说明其独立提出的诊断与本文架构论点相互印证，而非将其作为主要证据。）

在这一语境下，RPDA 框架的贡献在于表明，问题不仅限于各层之间的垂直分离，还包括一个水平维度：同一信息需要根据决策情境以不同方式进行组织、筛选和呈现。RPDA 与文献中已确立的用药流程（MUP）之间的区别至关重要：用药流程描述的是医疗照护流程中的药物流转，而 RPDA 描述的是关于药物的信息架构，并提出使用周期的每个阶段都构成一种独特的信息情境，具有各自的粒度、治理和风险要求。

5.2 与现有方法的关系

有必要将这一提案置于当代相关方法的背景下加以考察，尽管这些方法各自作用于不同层面。图 3从七个关键维度（文档保存、解释、呈现、来源溯源、错误风险、知识模型和问责）进行了系统比较，对照了传统方法与 PLP 架构。

RAG 系统（如 DrugRAG） [67], 通过将语言模型的回答与文献来源关联起来，取得了重大进展。然而，RAG 中的“检索”通常是技术性的向量搜索操作，而非辅助式、精选式的技术性阅读行为；“生成”产出的是看似合理的文本，而非经过类型化且可追溯的断言。RAG 在一定程度上解决了幻觉问题，但未能解决解释问责问题。

药学知识图谱（如 medicX-KG [68], ）以结构化方式表示药物、适应证与相互作用之间的关系。然而，这些图谱通常将所表示的知识与其原始文献来源割裂开来，导致图谱构建完成后无法恢复来源溯源。PRISMA 提出，应从保存并经精选的来源中衍生出上下文图，而不应将其构建为自主实体。

基于代理和规范性 YAML 文件的系统，例如 Sabel 等人所描述的系统 [15], 表明，即使在高度受控的场景中，将大语言模型应用于受监管领域，也逐渐需要引入规范层、流程控制、人工监督和形式化规范。这种经验上的趋同（彼此没有协调的作者得出了相同的结构性需求）构成了对这一问题的外部验证，而 PATOS–Lector–PRISMA 基础设施正是旨在以系统化、可复用的方式解决这一问题。

Bean 等人在一项有 1,298 名参与者参与的对照研究中证明，

大语言模型能够正确识别94.9% 的医疗状况，但前提是单独进行测试；然而，它们并未提升真实用户的表现：与相同模型互动的参与者得分低于对照组（$p < 0.001$） [69]。算法能力与信息效用之间的这种脱节进一步说明，核心问题不在于模型的表现，而在于知识与决策之间中介架构的设计。

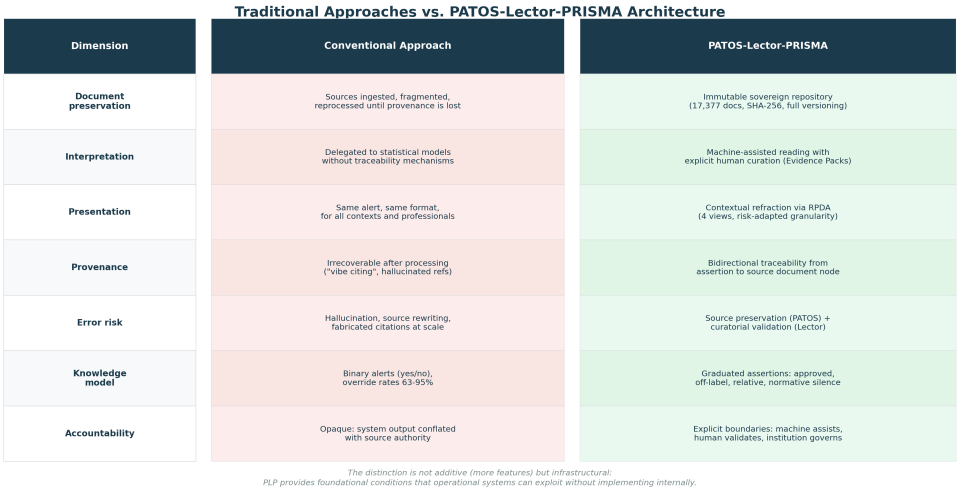


图 3. 从七个关键维度对比分析传统方法与 PATOS–Lector–PRISMA 架构：文档保存、解释、呈现、来源溯源、错误风险、知识模型和问责。两者的区别并非累加性的（功能更多），而是基础设施层面的：PLP 提供基础条件，供运行系统加以利用，而无需在内部实现这些条件。

5.3 监管与制度考量

所提出的架构以巴西监管环境为设计基准，尤其是巴西国家卫生监督局的药品注册制度，以及 RDC 47/2009 对药品说明书的文档要求。然而，其基本原则（文献主权、辅助技术阅读、受治理的情境化呈现和完整可追溯性）在概念上不受特定司法管辖区限制，适用于任何涉及规范性文档、版本、专业人员解读以及具有临床或法律后果的决策的监管背景。

该系统刻意不作决策的立场具有重大的监管影响。在全球监管机构开始要求卫生领域的人工智能系统具备可解释性和可审计性的情况下，能够保留来源、记录解读轨迹并明确责任的基础设施，比以“黑箱”方式运行的系统提供了更稳健的合规模式

，后者虽然能给出看似合理却无法追溯的输出。

从机构角度看，证据包 这一概念引入了一种药剂师熟悉的工作与责任单元：合格问题、有依据的答复、已识别的来源、明确的限制，以及有文档记录的策展决策。这种概念上的熟悉感或许有助于机构采用这一基础设施。尽管其架构复杂，但它基于已确立的专业实践运行。

5.4 评估标准

作为一篇提出规范性架构的立场论文，本文并未声称已完成实证验证。不过，为了能够系统评估当前提案及未来的实施方案，我们提出五项评估标准，用以衡量 PLP 基础设施：

1. **来源完整性**：呈现给最终用户的每项断言是否都能追溯至特定的源文件、版本和文本节点？ 2. **解释过程可追溯性**：从来源到断言的整个过程（包括模型的解读、策展人的决策和理由说明）是否都能完整重现？ 3. **策展覆盖率**：有多大比例的临床相关断言经过了有文档记录的人类验证？ 4. **情境差异化**：系统是否会根据专业语境（RPDA 轴）、相关风险和当前决策，以不同方式呈现相同的信息？ 5. **机构问责**：机器、专业人员和机构之间的责任边界是否得到明确界定并可供审计？

第 3.8 节中的演示示例为初步证据，表明该架构满足单个用药案例的标准 1、2、4 和 5。标准 3 得到了部分展示（为五种参考药物精选了 38 个证据包）。未来实证工作的首要议程，是针对更广泛的药物集进行系统性验证，并开展涉及卫生专业人员的对照研究。

5.5 局限性与未来工作

必须明确承认本研究存在局限性。

首先，该提案在根本上属于架构与概念层面。尽管各个组件处于不同的实施阶段（PATOS 处于运作阶段，Lector 正在开发中，PRISMA 正在进行原型设计），

但这三层之间尚未在生产环境中实现完整集成。依据第 5.4 节所定义标准开展实证验证，包括可用性研究、临床影响测量和可扩展性评估，是一项必要但尚未实现的研究议程。

其次，该模型要求进行明确的人工策展，尤其是在 Lector 的文档引用 → PageIndex → EvidencePack → CuratorialDecision 链条中，这意味着运营成本和成熟所需时间，可能会限制其初期可扩展性。如何在不损害可靠性的前提下扩大策展规模，仍是一项尚待解决的挑战；不过，循序渐进的策略（按参考药物优先排序、借鉴已获认可的策展成果、按治疗类别逐步拓展）提供了可行的路径。

第三，RPDA 框架提出了情境折射的四个维度。尽管这些维度以专业实践和用药流程相关文献为基础，但可能无法涵盖所有相关的决策情境。药物警戒、临床研究或专业教育等其他领域可能需要各自的信息视图，而该框架能否扩展至这些情境，仍有待未来研究。

第四，将语言模型用作技术阅读助手（Lector 的 PageIndex 所承担的角色）仍在迅速发展。模型的具体选择、提示策略，以及监管文档层级分解的质量指标，都是需要专门研究的课题；随着底层技术的发展，研究结果可能会有显著差异。

第五，该架构引入了必须予以承认的运行风险。Lector 人在回路模型所需的策展工作可能造成延迟，无法适应时间紧迫的临床场景；在这些场景中，即时警报优先于可追溯的解读轨迹。此外，尽管系统不作决策是其明确原则，证据包结构化且权威的呈现方式仍可能引发自动化偏见（即专业人员不加批判地接受系统输出的倾向），尤其是在时间压力和认知负荷较大的情况下 [25]。要降低这一风险，必须采用明确的界面设计策略（例如不确定性指标、强制确认认知边界，以及针对高风险断言设置阻尼机制），但这些策略仍有待实证检验。

第六，尽管当前本体已采用 IDMP 概念和标识符（ISO 11615/11616），并借鉴了 DM+D 模型（英国国家医疗服务体系）的层级结构，但要实现与药物识别和临床数据交换国际标准（如 FHIR，即快速医疗保健互操作性资源）的全面互操作，仍是自然的演进方向，其中涉及术语映射、监管协调和机构间互操作等不容忽视的挑战

治理。

最后，实证验证计划必须涵盖信息复杂程度各异的药物：不仅包括五种参考药物（安乃近、头孢曲松、氯硝西泮、华法林、阿达木单抗）——它们是经过刻意挑选的，具有丰富的文献资料且在监管方面意义重大；还包括文献来源较少的常见仿制药、近期获批且上市后证据有限的药物，以及监管状态存在争议或不断变化的药物。在信息稀缺、相互矛盾或快速变化的情况下，该架构的稳健性是一项关键考验，而当前的演示示例按设计并未涉及这一点。

6. 结论

本文提出 PATOS–Lector–PRISMA 体系，作为一种旨在负责任地进行药学知识管理的规范性信息基础设施方案，其基础是对文档保存、辅助解读与受治理的语境化呈现进行形式上的区分。

本文的核心贡献在于识别并处理一个结构性缺口：现有大多数系统（无论是 RAG 流水线、对话代理、警报系统还是规则引擎）都将具有截然不同的治理、时间性和责任要求的认知与机构运作压缩到单一技术层中。本文认为，这种融合是反复出现的脆弱性问题的根源，例如来源追溯丧失、解释不透明性，以及机器、专业人员与机构之间责任边界的模糊。

该方案阐述了四项具体贡献：

1. **PATOS–Lector–PRISMA 三元架构**作为一种具有认识论基础的三层架构，以Buckland的三分法、诠释学弧线和Floridi 的抽象层次方法为根基。2. **RPDA 框架**（监管、处方、调剂和给药）作为制药信息的情境折射机制，有别于用药流程，其作用层面是信息架构，而非照护流程。3. **“证据包”概念**作为可问责断言的正式单元（具有版本记录、可追溯、类型明确，并附有清晰的认知边界），在制药领域将循证医学和计算论证的原则付诸实践。

4. **药学断言的类型划分**依据其规范性质

以及言外之力，包括将规范性沉默作为一个具有认识论相关性的类别进行形式化处理，并以言语行为理论和道义逻辑为基础。

对 No-Harm.ai 等运行系统的互补性分析（第 4 节）表明，这一提案并非与现有解决方案竞争，而是提供了基础设施层面的条件（文档锚定、解释透明度、情境折射和分级鉴定），从而扩展了这些系统在可追溯性和机构问责方面所能提供的功能。

本文明确承认，这是一项概念性和架构性提案，目前仅部分实现，实证验证也尚待开展。然而，本文认为，对问题的清晰表述、严谨的认识论基础，以及通过正在开发的组件展示其可行性，共同构成了对健康信息学领域的一项重要贡献。尤其是在受监管领域对人工智能应用的压力日益增大的当下，这一点尤为重要，因为这种压力要求建立信任基础设施，而现有的大多数方法都无法提供这类基础设施。

本文的根本方向可以归纳为一句话：药学信息面临的核心挑战并非算法层面，而是架构层面。应对这一挑战，不仅需要更好的模型，还需要能够保留来源、记录解释、将药品规格置于语境中，并确保责任始终由人和机构承担的信息基础设施。

作者贡献

E.R.Z.N. 构思了 PATOS–Lector–PRISMA 架构，设计了三层分离结构，开发了 RPDA 框架，主导了系统实施，并撰写了稿件。A.V.C. 参与设计了证据包模型、策展工作流程和 PRISMA本体结构。C.C. 参与了药物警戒分析和监管文件整理。G.P.G. 参与设计了规范层级和数据填充流程。B.M. 参与设计了 PATOS存储库架构和文档摄取基础设施。所有作者均审阅并批准了最终稿件。

利益冲突声明

所有作者均为 Pharmdata 的合作人员。Pharmdata 是一家初创公司，正在开发本文所述的 PRISMA系统。PATOS–Lector–PRISMA 基础设施目前由 Pharmdata 积极开发。作者声明，除已说明的隶属关系外，本研究是在不存在任何可被视为潜在利益冲突的商业或财务关系的情况下开展的。

资金

本研究未获得外部资助。PATOS–Lector–PRISMA 基础设施由 Pharmdata 内部资助开发。

数据可用性

第 3.8 节（演示示例）中报告的数据源自巴西国家卫生监督局（巴西）发布的公开监管文件。PATOS 存储库、Lector整理成果和 PRISMA本体均为 Pharmdata 的专有资产。文稿中提供了汇总统计数据和架构说明，以便评估该提案。

参考文献

[1] Sumanth Chalasani、Juveria Syed、Madhan Ramesh、Varsha Patil 和 T. M. Pramod Kumar。药学实践领域的人工智能：文献综述。《临床与社会药学探索研究》，12：100346，2023年。doi：10.1016/j.rcsop.2023.100346。

[2] Neda Mottaghi-Dastjerdi 和 Mostafa Soltany-Rezaee-Rad。人工智能在药学科学中的进展与应用：综合综述。《伊朗药学研究杂志》，23(1)：e150510，2024年。doi：10.5812/ijpr-150510。

[3] Hadi Hamishehkar 和 Mehrdad Shahidi。人工智能对临床药学和医院环境未来的影响。《药学实践研究杂志》，14(3)：87–97，2025年。doi：10.4103/jrpp.jrpp_51_25。

[4] Henrique D. P. Dos Santos、Ana Helena D. P. S. Ulbrich、Vinicius Woloszyn 和 Renata Vieira。DDC-Outlier：使用

无监督学习预防用药错误。*IEEE*生物医学与卫生信息学期刊，23 (2):874–881，2019。doi: 10.1109/JBHI.2018.2828028。

[5] Cassia Lorena Leitão、Amanda Firmino Medeiros、Eny Florêncio Dias、Rosimeire Pereira Souza 和 Maria Auxiliadora Parreiras Martins。贝洛奥里藏特/MG 一家公立医院临床药学服务中的人工智能。巴西医院药学与卫生服务杂志，14(3):0991，2023。doi: 10.30968/rbfhss.2023.143.0991。

[6] Don Roosan、Patricia Padua、Rafia Khan、Hira Khan、Christine Verzosa 和 Yan Wu。ChatGPT 在临床药学中的有效性及人工智能在药物治疗管理中的作用。美国药师协会期刊，64(2):422–428，2024。doi: 10.1016/j.japh. 2023.11.023。

[7] Eunhye Shin、Megan Hartman 和 Murali Ramanathan。ChatGPT 大语言模型在社区药房决策支持中的表现。英国临床药理学杂志，90(12):3320–3333，2024。doi：10.1111/bcp.16215。

[8] Kaitlin Blotske、Xingmeng Zhao、Kelli Henry、Yanjun Gao、Adeleine Tilley、Moriah Cargile、Brian Murray、Susan E. Smith、Erin F. Barreto、Seth Bauer、Sunghwan Sohn、Tianming Liu、Tell Bennett、Mitch Cohen 和 Andrea Sikora。使用大语言模型识别药物相互作用。*medRxiv*，2025年。doi：10.64898/2025.12.03.25341549。预印本。

[9] Nurkhadija Rohani 和 Maryati Mohd Yusof。药房信息系统的意外后果：案例研究。国际医学信息学杂志，170：104958，2023年。doi: 10.1016/j.ijmedinf.2022.104958。

[10] Michael I. Harrison、Ross Koppel 和 Shirly Bar-Lev。医疗保健领域信息技术的意外后果——一项交互式社会技术分析。《美国医学信息学会杂志》，14(5)：542–549，2007年。doi：10.1197/jamia.M2384。

[11] Emily M. Campbell、Dean F. Sittig、Joan S. Ash、Kenneth P. Guappone 和 Richard H. Dykstra。与计算机化医嘱录入相关的非预期后果类型。《美国医学信息学会杂志》，13(5):547–556，2006。doi: 10.1197/jamia.M2042。

[12] Aditya Challapally、Chris Pease、Ramesh Raskar 和 Pradyumna Chari。《生成式人工智能鸿沟：2025年商业人工智能现状》。技术报告，MIT NANDA（网络化智能体与去中心化架构），2025年7月。多方法研究：系统综述 300+ 公共人工智能项目、对 52 个组织进行结构化访谈，以及对 153 位高级管理者开展调查。医疗保健与制药领域的颠覆评分：0.5/5。

[13]Gaurav Ramesh。《人工智能基础的人的要素》，2026年2月 2026年。网址 https://metadataweekly.substack.com/p/the-human-elements-of-the-ai-foundations。Substack 新闻通讯。引用 MIT NANDA 报告：评估企业级人工智能工具的组织中，只有 5% 最终将其投入生产。

[14]Jessica Talisman。《来源溯源何处终结，知识何以衰减》，2月2026年。网址 https://jessicatalisman.substack.com/p/where-provenance-ends-knowledge-decays。Substack 新闻简报。文章指出，大语言模型会通过对训练数据进行有损且不可逆的压缩，系统性地摧毁来源溯源。文章区分了检索层（可以引用来源）与参数层（来源溯源已不可挽回地丢失），并将“氛围式引用”定义为一种结构性地编造参考文献的现象。文章追溯了来源溯源作为档案原则从1841年至今的发展。

[15] Joar Sabel、Mattias Wingren、Andreas Lundell、Soren Andersson、Sara Rosenberg、Susanne Hagglund、Linda Estman 和 Malin Andtfolk。《使用大语言模型进行用药咨询：在灵活性与严格性之间取得平衡》。2025年。预印本。

[16] 空间数据系统咨询委员会。《开放档案信息系统（OAIS）参考模型》。推荐实践 CCSDS650。0-M-2 (洋红书)，CCSDS，2012年。网址 https://public.ccsds.org/Pubs/650x0m2.pdf。亦以 ISO 14721:2012 发布。

[17] 联机计算机图书馆中心（OCLC）和研究图书馆中心（CRL）。可信存储库审计与认证：标准与检查清单。技术报告，OCLC 和 CRL，2007 年 2 月 2007。网址 https://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf。版本1。0。

[18] Peter Pirolli 和 Daniel M. Russell。关于意义建构专题特刊的导言。人机交互，26(1–2)：1–8，2011。doi：10.1080/07370024.2011.556557。

[19] Aniket Kittur、Jeffrey V. Nickerson、Michael S. Bernstein、Elizabeth Gerber、Aaron Shaw、John Zimmerman、Matt Lease 和 John Horton。众包工作的未来。见《*2013ACM* 计算机支持的协同工作会议（*CSCW '13*），第 1301–1318 页，纽约，纽约州，2013 年。ACM。doi：10.1145/2441776.2441923。

[20]Luc Moreau 和 Paolo Missier。《PROV-DM：PROV 数据模型》。万维网联盟（W3C）建议，万维网联盟（W3C），2013 年 4 月 2013。网址：https://www.w3.org/TR/2013/REC-prov-dm-20130430/。W3C 建议，日期为 302013 年 4 月 2013。

[21] Peter Buneman、Sanjeev Khanna 和 Wang-Chiew Tan。为什么以及何处：数据溯源的特征描述。载于数据库理论 — *ICDT 2001*： 第 *8* 届国际会议，《计算机科学讲义》第 1973 卷，第 316–330 页。施普林格，2001 年。doi: 10.1007/3-540-44503-X_20。

[22] Aidan Hogan、Eva Blomqvist、Michael Cochez、Claudia d'Amato、Gerard de Melo、Claudio Gutierrez、Sabrina Kirrane、José Emilio Labra Gayo、Roberto Navigli、Sebastian Neumaier、Axel-Cyrille Ngonga Ngomo、Axel Polleres、Sabbir M. Rashid、Anisa Rula、Lukas Schmelzeisen、Juan Sequeda、Steffen Staab 和 Antoine Zimmermann。《知识图谱》。《*ACM*计算调查》，54(4)：1–37，2021。doi：10.1145/3447772。

[23] Stephen E. Toulmin。《论证的用途》。剑桥大学出版社 ，英国剑桥，修订版，2003。ISBN 978-0-521-53483-3。初版出版于1958年。

[24] Lucy A. Suchman。《计划与情境行动：人机交流问题》。剑桥大学出版社，英国剑桥，1987。ISBN 978-0-521-33739-8。

[25] Joan S. Ash、Dean F. Sittig、Eric G. Poon、Kenneth Guappone、Emily Campbell 和 Richard H. Dykstra。计算机化医嘱录入相关意外后果的范围与重要性。《美国医学信息学会杂志》，14(4):415–423，2007。doi: 10.1197/jamia.M2373。

[26] Iyad Rahwan、Manuel Cebrian、Nick Obradovich、Josh Bongard、Jean-François Bonnefon、Cynthia Breazeal、Jacob W. Crandall、Nicholas A. Christakis、Iain D. Couzin、Matthew O. Jackson、Nicholas R. Jennings、Ece Kamar、Isabel M. Kloumann、Hugo Larochelle、David Lazer、Richard McElreath、Alan Mislove、David C. Parkes、Alex 'Sandy' Pentland、Margaret E. Roberts、Azim Shariff、Joshua B. Tenenbaum 和 Michael Wellman。机器行为。《自然》，568(7753):477–486，2019。doi: 10.1038/s41586-019-1138-y。

[27] Luciano Floridi、Josh Cowls、Monica Beltrametti、Raja Chatila、Patrice Chazerand、Virginia Dignum、Christoph Luetge、Robert Madelin、Ugo Pagallo、Francesca Rossi、Burkhard Schafer、Peggy Valcke 和 Effy Vayena。《AI4People——面向良好人工智能社会的伦理框架：机遇、风险、原则和建议》。《心灵与机器》，28(4):689–707，2018。doi: 10.1007/s11023-018-9482-5。

[28] Carl Lagoze、Sandra Payette、Edwin Shin 和 Chris Wilper。《Fedora：复杂对象及其关系的架构》。《国际数字图书馆杂志》，6(2):124–138，2006。doi: 10.1007/s00799-005-0130-3。

[29]Marti A. Hearst。《搜索用户界面》。剑桥大学出版社，英国剑桥，2009。ISBN 978-0-521-11379-3。doi: 10.1017/ CBO9781139644082。

[30] Catherine C. Marshall. 注释：从纸质书到数字图书馆。载于《第二届*ACM*数字图书馆国际会议论文集（*DL '97*），第 131–140 页，纽约，纽约州，1997 年。ACM。doi：10.1145/263690.263806。

[31] Catherine C. Marshall. 电子书的阅读与写作。《信息概念、检索与服务综合讲座》。Morgan & Claypool出版社，2009 年。doi：10.2200/S00215ED1V01Y200907ICR009。

[32] Finale Doshi-Velez 和 Been Kim。迈向可解释机器学习的严谨科学。2017 年。预印本。

[33] Patrick Brézillon。情境图中的任务实现模型。载于情境建模与应用（*CONTEXT 2005*），《计算机科学讲义》第 3554 卷，第 55–68 页。施普林格，2005 年。doi：10.1007/11508373_5。

[34] Jessica Talisman. 本体、情境图和语义层：人工智能在2026年年1月2026年。网址：h
t
tps://metadataweekly.substack.com/p/ontologies-context-graphs-and-semantic。Substack 通讯。文章认为，语义层（指标治理）不足以支持人工智能推理；人工智能系统所需的形式化知识表示，应由本体和上下文图提供。文章以医疗保健和生命科学为例，证明了本体在大规模应用中的价值。

[35] Michael K. Buckland. 信息作为事物。美国信息科学学会杂志，42(5):351–360，1991。doi: 10.1002/(SICI) 1097-4571(199106)42:5<351::AID-ASI5>3.0.CO;2-3。

[36] Rafael Capurro 和 Birger Hjørland. 信息概念。信息科学与技术年度评论，37(1):343–411，2003。doi: 10. 1002/aris.1440370109。

[37] Michael D. Myers. 辩证诠释学：信息系统实施的理论框架。《信息系统杂志》，5 (1):51–70，1995。doi: 10.1111/j.1365-2575.1995.tb00089.x。[38] Paul Ricoeur.《诠释理论：话语与意义的剩余》。德克萨斯基督教大学出版社，德克萨斯州沃思堡，1976。ISBN 978-0-912646- 59-6。

[39] Luciano Floridi. 信息哲学。牛津大学出版社，牛津，2011。ISBN 978-0-19-923238-3。

[40] J. L. Austin.《如何以言行事》。哈佛大学出版社，马萨诸塞州剑桥，第二版，1975。ISBN 978-0-674-41152-4。最初于1955年在哈佛大学以威廉·詹姆斯讲座的形式发表。[41] John R. Searle。施为行为分类。载于《表达与意义：言语行为理论研究》，第1–29页。剑桥大学出版社，剑桥，1979。ISBN 978-0-521-31393-3。[42] Pablo E. Navarro 和 Jorge L. Rodríguez。《道义逻辑与法律体系》。剑桥哲学与法律导论丛书。剑桥大学出版社，2014。ISBN 978-0-521-13990-8。doi：10.1017/CBO9781139032711。[43] Carlos E. Alchourrón 和 Eugenio Bulygin。《规范系统》。精确哲学文库第5卷。施普林格出版社，维也纳和纽约，1971。ISBN 978-3-211-81019-4。

[44] Henry Prakken 和 Giovanni Sartor。法律与逻辑：论证视角综述。人工智能，227:214–245，2015。doi: 10.1016/j.artint.2015.06.005。

[45] Ojelanki K. Ngwenyama 和 Allen S. Lee。电子邮件中的沟通丰富度：批判社会理论与意义的情境性。《管理信息系统季刊》，21(2):145–167，1997。doi: 10.2307/249417。

[46] John C. Mingers。信息与意义：交互主体论述的基础。《信息系统杂志》，5(4):285–306，1995。doi: 10.1111/j.1365-2575.1995.tb00100.x。

[47] 尤尔根·哈贝马斯。《交往行为理论》第一卷*1*：理性与社会的理性化。灯塔出版社，波士顿，1984。ISBN 978-0-8070- 1507-0。德文原版：《交往行为理论》，苏尔坎普出版社，1981。

[48] David L. Sackett、William M. C. Rosenberg、J. A. Muir Gray、R. Brian Haynes 和 W. Scott Richardson。循证医学：它是什么，又不是什么。《英国医学杂志》，312(7023):71–72，1996年。doi: 10.1136/bmj.312.7023.71.[49] Gordon H. Guyatt、Andrew D. Oxman、Gunn E. Vist、Regina Kunz、Yngve Falck-Ytter、Pablo Alonso-Coello 和 Holger J. Schünemann。GRADE：关于证据质量和推荐强度评定的逐渐形成的共识。《英国医学杂志》，336(7650):924–926，2008年。doi: 10.1136/bmj.39489. 470347.AD.[50] Muhammad Ahsan Raza、Sadia Aziz、Maria Noreen、Asma Saeed、Ifrah Anjum、Muhammad Ahmed 等。药学中的人工智能（AI）：创新概述。《药学创新》，13(2)，2022年。doi: 10.24926/iip.v13i2.4839.

[51] Laurence Gosselin、Martin Thibault、Denis Lebel 和 Jean-François Bus- sières。药学中的人工智能：一篇叙述性综述。《加拿大医院药学杂志》，74(2):135–143，2021年。

[52] Yolanda González-Pérez、Alejandro Montero Delgado 和 Jose Manuel Martinez Sesmero。人工智能在医院药学中的应用。《*Farmacia Hospitalaria*》，48（增刊 1）：S35–S44，2024。doi：10.1016/j.farma.2024. 02.007。

[53] Emma Johns、Ahmad Alkanj、Morgane Beck、Laura Dal Mas、Bruno Gourieux、Erik Sauleau 等。使用机器学习或深度学习模型，在医院环境中检测不适当处方：系统综述。《欧洲医院药学杂志》，2024。doi：10.1136/ejhpharm-2023-003857。

[54] Carlos Del Rio-Bermudez、Ignacio Medrano、Laura Yebes 和 Jose Luis Poveda。迈向大数据、人工智能与医院药学之间的共生关系。《药学政策与实践杂志》，2020。doi：10.1186/s40545-020-00276-6。

[55] Jeremy S. Stultz 和 Milap C. Nahata。儿科药物处方与使用的计算机化临床决策支持。《美国医学信息学会杂志》，19(6):942–953，2012。doi: 10.1136/amiajnl-2011-000798。

[56] Kannan Sridharan 和 Gowri Sivaramakrishnan。释放大语言模型在药物审查与核对中的潜力：一项概念验证研究。《临床与社会药学探索研究》，2024年。doi: 10.1016/j.rcsop.2024.100492。

[57] Felix Busch、Lena Hoffmann、Christopher Rueger、Elon HC van Dijk、Rawen Kader、Esteban Ortiz-Prado、Marcus R. Makowski、Luca Saba、Mar- tin Hadamitzky、Jakob Nikolas Kather、Daniel Truhn、Renato Cuocolo、Lisa C. Adams 和 Keno K. Bressem。大语言模型在患者护理中的当前应用与挑战：一项系统综述。《*Communi- cations Medicine*》，5:26，2025年。doi: 10.1038/s43856-024-00717-2。一项涵盖29个医学专科、89项研究（2022–2023）的系统综述。研究归纳出大语言模型局限性的两类结构性因素：设计方面（未针对医学领域进行优化、数据不透明、可及性）和输出方面（不可复现性、不完整性76.4%、错误性87.6%、不安全性43.8%）。研究将幻觉、谵妄和虚构区分为不同现象。

[58] Anan Jarab, Walid Al-Qerem, Karem Alzoubi, Hazem Obeidat, Shrouq Abu Heshmeh, Tareq Mukattash 等。人工智能在药学实践中的应用：社区药剂师的态度和意愿，以及

其实施面临的障碍。沙特药学期刊，31(8):101700，2023年。doi: 10.1016/j.jsps.2023.101700。

[59] Nader Hatem. 推进药学实践：情报驱动型药学实践的作用及药学情报的兴起。《综合药学研究与实践》，2024。doi：10.2147/IPRP.S466748。

[60] Clare L. Brown、Helen L. Mulcaster、Katherine L. Triffitt、Dean F. Sittig、Joan S. Ash、Katie Reygate、Andrew K. Husband、David W. Bates 和 Sarah P. Slight。初级和二级医疗中使用计算机化医嘱录入系统所产生的处方错误类型及成因：一项系统综述。《美国医学信息学会杂志》，24(2):432–440，2016。doi：10.1093/jamia/ocw119。

[61]Calandra Feather、Nicholas Appelbaum、Ara Darzi 和 Bryony Dean Franklin。电子处方系统中的适应证记录和基于适应证的处方：一项系统综述与叙述性综合。《英国医学杂志质量与安全》，32(6):357–368，2023。doi：10.1136/bmjqs-2022-015452。

[62] Birger Haarbrandt、Erik Tute 和 Michael Marschollek。全国性数据基础设施项目的临床知识治理框架。载于《健康技术与信息学研究》，第248卷，第196–203页。IOS Press，2018年。PMID：29726437。

[63] Kacper Sokol、James Fackler 和 Julia E. Vogt。人工智能应真正支持临床推理和决策，以弥合转化鸿沟。《*npj*数字医学》，8:345，2025年。doi：10.1038/ s41746-025-01725-9。

[64] José Romério Rabelo de Melo、Elisângela Costa Duarte、Marcelo Vogler de Moraes、Andrew Stephen Fooks 和 Paulo Sérgio Dourado Arrais。新冠肺炎疫情期间，安乃近是巴西消费量最大的镇痛药：这为何重要?《药物流行病学》，4(3)：230–241，2023年。doi：10.3390/pharma4030017。数据显示，巴西每年使用2.15亿剂；仅2016年，圣保罗就消耗了488吨。

[65] Henry Sznejder、Caroline Amand、Andrew Stewart、Ricardo Salazar 和 Wanessa Alessandra Ruiz Scala。巴西人口使用安乃近（dipyrone）的真实世界证据：一项纳入 380,000 多名患者的回顾性队列。《*Einstein*（圣保罗）》，20:eAO6353，2022年。doi: 10.31744/einstein_journal/2022AO6353。回顾性队列研究（384,668 名患者，2015 年 1 月至 2017 年 12 月，Amil 临床数据仓库）。未发现粒细胞缺乏症病例。不良反应发生率比对乙酰氨基酚低 38.8%，比乙酰水杨酸低 46.8%。87% 的患者将其用作单药治疗。

[66] Nelson Hamerschlak、Elenara Maluf、Alexandre Biasi Cavalcanti、Álvaro Avezum Jr、José Eluf-Neto、Roberto Passetto Falcão 等。拉丁美洲国家粒细胞缺乏症的发生率和风险因素：拉丁研究。《欧洲临床药理学杂志》，64:921–929，2008年。doi: 10.1007/ s00228-008-0513-7。在巴西、阿根廷和墨西哥开展的多中心病例对照研究（2002—2005年）。安乃近可作为非处方药购买的地区，粒细胞缺乏症发生率为每百万人年 0.38 例。

[67] Houman Kazemzadeh, Kiarash Mokhtari Dizaji, Seyed Reza Tavakoli, Farbod Davoodi, MohammadReza KarimiNejad, Parham Abed Azad, Ali Sabzi, Armin Khosravi, Siavash Ahmadi, Mohammad Hossein Rohban, Glolamali Aminian 和 Tahereh Javaheri。DrugRAG：通过一种新型检索增强生成流程提升药学大语言模型的性能。2025年。预印本。RAG 流程使 NAPLEX 的准确率提升了 7–21 个百分点。

[68] Lizzy Farrugia、Lilian M. Azzopardi、Jeremy Debattista 和 Charlie Abela。medicX-KG：满足药剂师药物信息需求的知识图谱。《生物医学语义学杂志》，16:11，2025年。doi：10.1186/s13326-025-00332-7。

[69] Andrew M. Bean、Rebecca Elizabeth Payne、Guy Parsons、Hannah Rose Kirk、Juan Ciro、Rafael Mosquera-Gómez、Sara Hincapié M、Aruna S. Ekanayaka、Lionel Tarassenko、Luc Rocher 和 Adam Mahdi。大语言模型作为面向普通公众的医疗助手的可靠性：一项随机预注册研究。《自然·医学》，2026年。doi：10.1038/s41591-025-04074-y。该随机对照研究纳入了 1,298 名参与者。仅使用大语言模型时，94.9% 的病例中的病症被正确识别，但使用大语言模型的参与者表现并不优于对照组（34.5% 对 44.2% ）。基准测试和模拟交互无法很好地预测真实世界表现。

A. EvidencePack：半形式化规范

本附录对证据包进行了半形式化规范。证据包是 PATOS–Lector–PRISMA 架构中可问责断言的核心单元。本规范有意采用概念性表述，旨在确立任何符合规范的实现都必须满足的结构属性和良构条件。机器可读形式化表示（例如采用 JSON-LD 或 RDF）留待未来工作完成。

A.1 定义

证据包是一个五元组：

$$\mathbf{EP} = \langle Q, R, P, L, C \rangle \tag{1}$$

其中：

- **Q** （合格问题）是一个二元组(q, τ)，其中q 是一个自然语言问题，而 $\tau \in T$ 是从受控分类体系T 中选取的断言类型 ={适应证、禁忌证、给药剂量、相互作用、不良_反应、警告、注意事项

- **R** （有依据的答复）是一个三元组(a, V^+, V^-)，其中a 是通过对一个或多个源文档进行技术性阅读而构建的自然语言断言，V^+ 是a 成立时的一组有效条件，而V^- 是a 不成立时的一组无效条件。

- **P**（来源溯源链）是一组非空的来源溯源记录 $\{p1, p2, \ldots, p_n\}$，其中每条$p_i = (d, v, h, N)$ 由文档标识符$d \in$ patos、版本标识符v、密码学哈希h （SHA-256）以及一组 PageIndex 节点标识符$N = \{n1, n2, \ldots\}$ 组成，这些节点标识符对应于提取信息的来源。约束 $|P| \geq 1$ 确保每个证据包至少有一个锚定来源。

- **L**（认知边界）是一个四元组(Δ, G, D, S)，其中：
  - Δ （分歧）是指针对同一主题，不同来源之间已识别出的意见分歧集合；
  - G （缺失）是文档信息缺失集合（信息不可用或未能找到）；
  - D （依赖项）是会调节断言有效性的临床背景条件集合；
  - S （沉默）是一组规范性沉默，即监管来源未明确表态的问题。

- **C**（策展状态）是一个三元组 (σ, c, j)，其中 $\sigma \in \{draft、under_review、已接受、已拒绝\}$表示策展状态，$c$ 表示策展人标识符，j 表示对策展决策的自然语言说明。

A.2 良构条件

证据包 $\mathbf{EP} = \langle Q、R、P、L、C \rangle$ 当且仅当满足以下条件时才是**良构的**：

1. **类型化问题**：$\tau \in T$ (问题类型属于受控分类体系)。

2. 来源锚定： $|\mathrm{P}| \geq 1$ （至少存在一条溯源记录）。3. 完整性可验证：对于每个 $\mathrm{p}_i = (\mathrm{d}, \mathrm{v}, \mathrm{h}, \mathrm{N}) \in \mathrm{P}$，哈希值 h 与 PATOS 中版本 v 的文档 d 相符。4. 节点可追溯性：对于每个 p_i，集合 N 均非空，即该断言可追溯至至少一个具体的文档节点。5. 认知完整性：L 已明确指定；L 的任一组成部分为空集，在语义上不同于未提供 L（前者表示“未发现分歧/缺失/依赖项/沉默”；后者属于规范错误）。6. 策展闭合：如果 $\sigma \in \{$已接受、已拒绝$\}$，那么 $\mathrm{c} \neq \emptyset$ 且 $\mathrm{j} \neq \emptyset$ （已接受或已拒绝的包必须明确策展人，并记录理由说明）。

A.3 言外行为分类

断言类型 τ 依据塞尔的言语行为分类，决定证据包的言外之力 [41]:

表 3. 依据塞尔的言语行为分类，对证据包断言类型进行言外行为分类。

断言类型 τ	言外行为类别	力度
适应证	断言性	承诺治疗适用性的真实性
禁忌证	指令性	指示专业人员避免
给药剂量	指令性 +断言性	基于事实依据规定行为
相互作用	断言性 + 指令性	陈述事实并警告可能造成的后果
不良反应_	断言性	报告观察到或记录在文档中的影响
警告	指令性	提示在特定病症下谨慎使用
注意事项	指令性	建议进行监测或调整用法
特殊人群_	断言性 + 指令性	限定其对某一子群体的适用性
规范性沉默_	不作承诺	表示监管机构未作出声明

类型规范性_沉默在认识论上有别于其他类型：它既不肯定、否定、许可，也不禁止，而是正式记录规范系统未对所提出的问题表明立场。这与道义逻辑中的规范性缺口概念相一致 [42, 43]。

A.4 生命周期

证据包的生命周期严格按正向顺序推进：

```
草稿 -> 审阅中_已接受 | 已拒绝
```

状态转换由以下因素控制：

- **草稿** → **正在_审阅**：由 Lector 阅读流程完成触发。

- **审核中_审核中** → **已接受**：需要 策展人 $\neq \emptyset$、理性 $\neq \emptyset$，以及明确的人工验证。
- **审核中_审核中** → **已拒绝**：需要 策展人 $\neq \emptyset$、理性 $\neq \emptyset$，并记录拒绝理由。
- **不得向后转换**：已拒绝的包不能再次接受；必须创建新的证据包，并建立新的溯源链。这确保了策展轨迹的可审计性。

A.5 组合与派生

证据包具有**组合性**：复杂的药品陈述（例如：“安乃近适用于成人轻至中度疼痛，但已知对其超敏反应或有粒细胞缺乏症病史的患者除外”）可以分解为多个证据包，每个证据包都有自己的类型、溯源链和认知边界。

证据包可以**派生，但不可转换**：可以通过查阅相同或不同的来源创建新的证据包，但已接受的证据包绝不会被修改，而是进行版本管理。这体现了 PATOS 层的不可变性原则。