



评估执法调查中人工智能技术的有效策略

第一章： 人工智能格局——为何此刻至关重要

人工智能聊天机器人面市仅短短数年后，人工智能已无处不在。它正影响着并改变着我们生活的方方面面。



执法领域自21世纪初起就已使用具有类人工智能特征的自动化工具。而在过去几年中，开发者们创造了众多人工智能驱动的技术，承诺将改变刑事调查的进行方式。

在2025年执法机构认证委员会（CALEA）会议上，217名与会执法人员被问及他们使用人工智能的情况：



国家警务学院，警务人工智能：早期采用者之外，您的机构能做和应该做的事，policinginstitute.org/infocus/ai-in-policing/（访问于2026年7月22日）。

各规模机构都有理由对此感到好奇。证据表明，人工智能辅助调查已显著缩短了警官在研究和文书工作上的时间。在其众多能力中，人工智能可以：

- 利用现有记录和数据源，识别不同司法管辖区的相关主体和关联方。
- 分析跨不同数据源的资金模式和可疑活动
- 快速识别跨司法管辖区的模式与趋势，从而有可能将相关主体与更多案件联系起来



但这里还有另一个重要的事实：并非所有可供执法部门使用的AI工具都能真正帮到他们。我们都曾在搜索引擎、个人电脑和网络上遇到过所谓的“消费级AI”。这种AI在起草基本文件和邮件、总结网络搜索、创建网络内容、管理个人日程以及其他任务方面表现尚可。但它缺乏执法调查人员所要求的可靠性、全面的数据以及在法庭上的可辩护性。

各机构所需要的是专门为调查而构建的AI解决方案。鉴于案件完整性和法庭辩护能力至关重要，调查人员、指挥官和机构技术决策者在评估用于调查的AI工具时，必须进行彻底的尽职调查。

第二章： 人工智能评估 的五大支柱





由于并非所有人工智能工具都能满足执法部门的严格要求，机构领导者需要一套框架来帮助他们评估市场上的工具，并确定哪些工具能够提升而非削弱调查效率。以下五个支柱就提供了这样一种框架。

支柱一：数据来源与可信内容

不言而喻，调查人员需要经过核实和值得信赖的信息。不准确的数据可能会让调查人员误入死胡同、走错方向。这不仅会浪费时间，如果调查最终导致逮捕，还可能危及案件结果。

人工智能的一大优势在于其能够即时访问并筛选海量数字数据，然后迅速从这些数据中提炼出经过验证的洞见。但执法机构必须清楚了解人工智能工具获取其数据的方式和来源。

总的来说，人工智能主要采用以下一种或两种自动化技术：

开放式网络爬取

开放式网络爬虫是指从公共网站自动提取数据。这种技术以用户（如执法调查人员）易于检索和分析的格式收集数据。网络爬虫可以获取大量必要和有用的信息，但它确实有其局限性。

数据管理

这项技术专注于获取更专业的数据，包括公共记录（例如涉及监禁的记录）以及跨司法管辖区的其他优质专有来源。“策展”（Curation）是指主动管理和组织这些数据，使其保持严格更新。策展还可以解决涉及记录冲突或不完整的问题。

网页公开爬取和数据整理对于全面的数据采集都至关重要。结合这些方法有助于确保数据质量最高，并且调查人员能够信任人工智能工具的输出（即工具根据调查人员的查询所提供的信息）。



第二支柱：面向消费者的AI问题

在调查用途方面，消费级人工智能使机构面临诸多风险，最显著的是：

- 安全漏洞
- 数据保留问题
- 缺乏数据保管链
- 合规、隐私和责任问题
- 人工智能输出在法庭上可采性不明朗

简而言之，消费级AI平台缺乏执法部门对其数据所要求的控制水平。这种控制不足会使机构面临风险，因为它会在检方的案件中留下漏洞，供辩护方利用。

这直接引出了下一个支柱。

第三支柱：法庭可辩护性——幻觉与引证

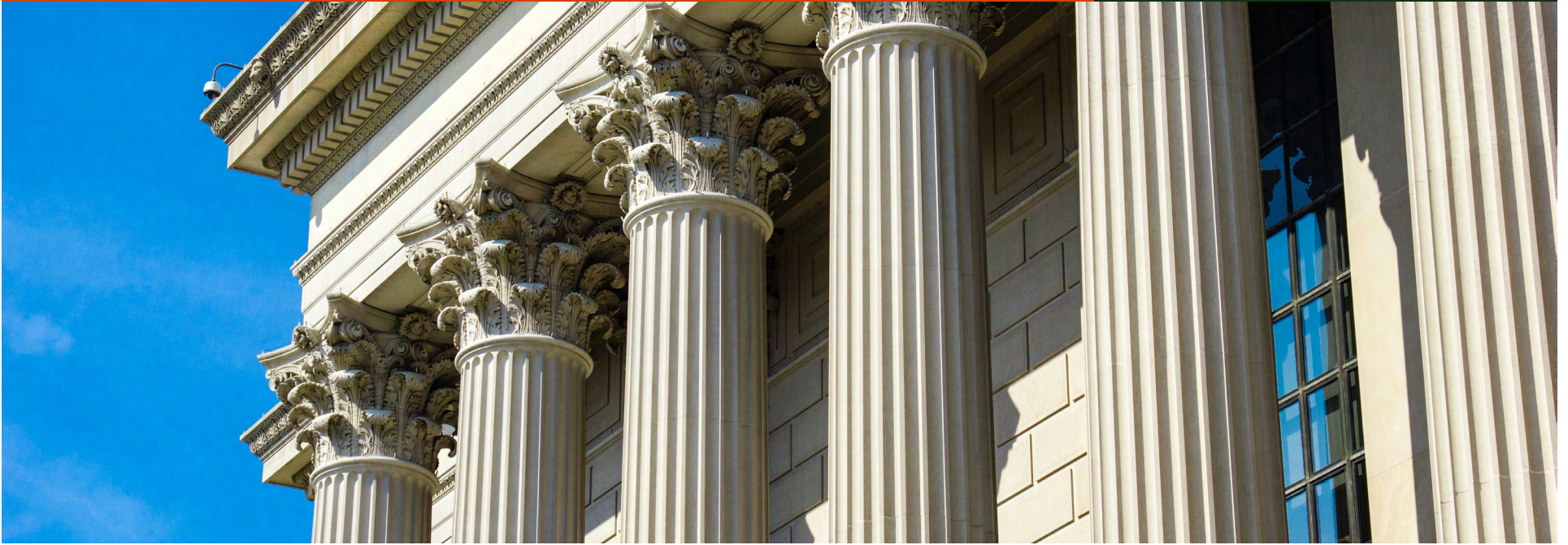
当执法调查人员使用AI工具协助检察官时，该工具提供的输出必须能在法庭上站得住脚。这就是为什么调查人员需要了解幻觉和引用错误。

• **当人工智能工具生成虚假、误导性或完全捏造的信息时，就会产生幻觉。**换句话说，该工具提供了虚假的输出。工具“产生幻觉”有几种原因。这些原因主要归结于缺乏可靠数据。简单来说，如果工具找不到用户请求的信息，它通常会编造出来。

• **引证是指能够为所提供的信息或数据提供可验证来源的AI输出。**每一条数据都可以追溯到可靠的权威来源。引证输出具有可辩护性——它可以在法庭上站得住脚。

过去几年里，已有大量案例显示AI幻觉破坏了法律诉讼。在这些大多数情况下，研究人员援引了实际上并不存在的法律案件。显然，如果一份包含幻觉的调查报告被提交为官方记录或用于获取搜查令，它可能会破坏证据链。

专门为调查设计的AI工具可以从多个方面帮助机构调查员规避这些风险。最显著的是，专业级的工具能够（或应该）创建审计追踪记录。审计追踪记录有助于调查员亲自确认数据来源的可靠性。通过确保透明度、问责制和证据完整性，审计追踪记录提供了强大的抗辩依据。



第四支柱：现实世界中的时间节省

调查犯罪所需的时间取决于各种情况。简单的财产犯罪可能只需要几个小时，而更复杂的暴力犯罪可能需要几周或几个月。调查人员知道，调查的前48小时至关重要。证据和信息越快被发现和审查，案件就越有可能被快速侦破（并起诉）。

这正是为什么被称为人工智能压缩的技术对调查人员来说可能极具价值。调查事件的标准“人工”方法可能需要调查人员花费数小时的时间。人工智能自动化可以将该调查的某些环节从数小时压缩到数分钟。这使得调查人员可以将更多的时间用于其他与案件相关的任务，例如现场跟进和访谈。

第五支柱：文件编制与报告撰写支持

与调查的研究部分类似，“手写”报告撰写需要警员和调查人员投入数小时的时间。这并不仅仅是因为写作本身。收集并整合用于指挥官报告、CompStat会议的文件，以及（对于高关注度案件）向市县官员进行的简报所需的文件，可能会令人不堪重负。而且，始终存在报告可能无意中遗漏关键文件的担忧，这可能会在后续阶段造成证据风险。同样，检察官和辩护律师都希望证明，调查报告中包含的信息已经过充分核查以确保其可靠性。

人工智能自动化可以通过创建相关文档的摘要来简化报告和文档流程。但执法调查人员需要来源的摘要。工具应该能够生成带有内联引文的草稿摘要。这在整个文档过程中保持了证据的完整性，以便检察官能够用核实的信息支持他们案件的各个方面。

一项注意事项：调查人员应审阅并最终确定这些引用，以确保信息和来源准确无误，并删除任何AI生成的幻觉内容。

第三章： 向AI供应商提 出的5个问题



前一章中的评估框架可以转化为五大基本且可操作的问题，供执法机构向潜在供应商提出。

您的数据来源是什么？

换句话说，您工具所访问和使用作为其输出基础的数据有多可靠？您如何证明其可靠性？

您工具的输出能否轻易追溯到那些来源？

一款用于调查的专业级AI工具，应当能够提供审计追踪，使调查人员能够确认信息来源的可靠性。

如何确保执法数据安全？

如果数据没有得到保护，可能会对调查和诉讼程序造成问题。

您有多少已记录的时间节省？

一个无法证明它能节省代理公司开支的人工智能工具，绝非值得投资。

您的工具如何支持法庭报告撰写和文件制作？

例如，其来源总结是否可靠且没有幻觉？



第四章： 清晰调查——符合 每项标准

鉴于风险如此之高，执法调查人员需要一款极其可靠的AI驱动工具。汤森路透® CLEAR Investigate平台满足了这五项评估支柱的所有要求。

该调查型AI平台由汤姆森路透技术CLEAR提供支持，该技术能够快速分析多种高质量数据集。CLEAR调查平台的 premium 和公开网络数据采集功能，使执法调查人员能够快速查找失踪儿童、锁定相关人员、抓捕罪犯，并执行其他关键公共安全任务。通过自动化数小时繁琐的工作（例如，总结冗长的起诉书、案件和各种内部及外部文件），CLEAR调查平台使调查人员能够将时间和精力集中于关键分析和高优先级线索。

此外，CLEAR Investigate 提供了首款专为执法部门设计的自主式AI调查解决方案。简单来说，自主式AI就像一个自治代理，能够在极少人工监督的情况下，从开始到结束执行多步骤流程。它超越了标准的搜索驱动流程，整合了海量的公共和专有记录，并通过详尽的审计追踪，提供可靠的摘要和可辩护性。

值得注意的最后一项尽职调查最佳实践：人工智能技术正在持续演进和改进。随着执法机构对人工智能的了解日益加深，他们必须确定供应商的解决方案是否具备与其共同学习和发展的能力。

了解更多关于CLEAR调查系统及其如何加速和优化执法调查。

访问 legal.thomsonreuters.com/en/c/clear-investigate/ai-powered-investigations。

路透社并非消费者报告机构，其提供的任何服务或其中包含的数据均不构成《联邦公平信用报告法》（FCRA，15 U.S.C. § 1681及以下条款）中所定义的“消费者报告”。提供给你的数据不得用作消费者债务催收决策、确定消费者信贷、保险、就业、政府福利或住房资格的因素，也不得用于FCRA授权的任何其他目的。通过访问我们的任一服务，您同意不得将该服务或数据用于FCRA授权的任何目的，或用于与消费者申请相关的任何不利行动。