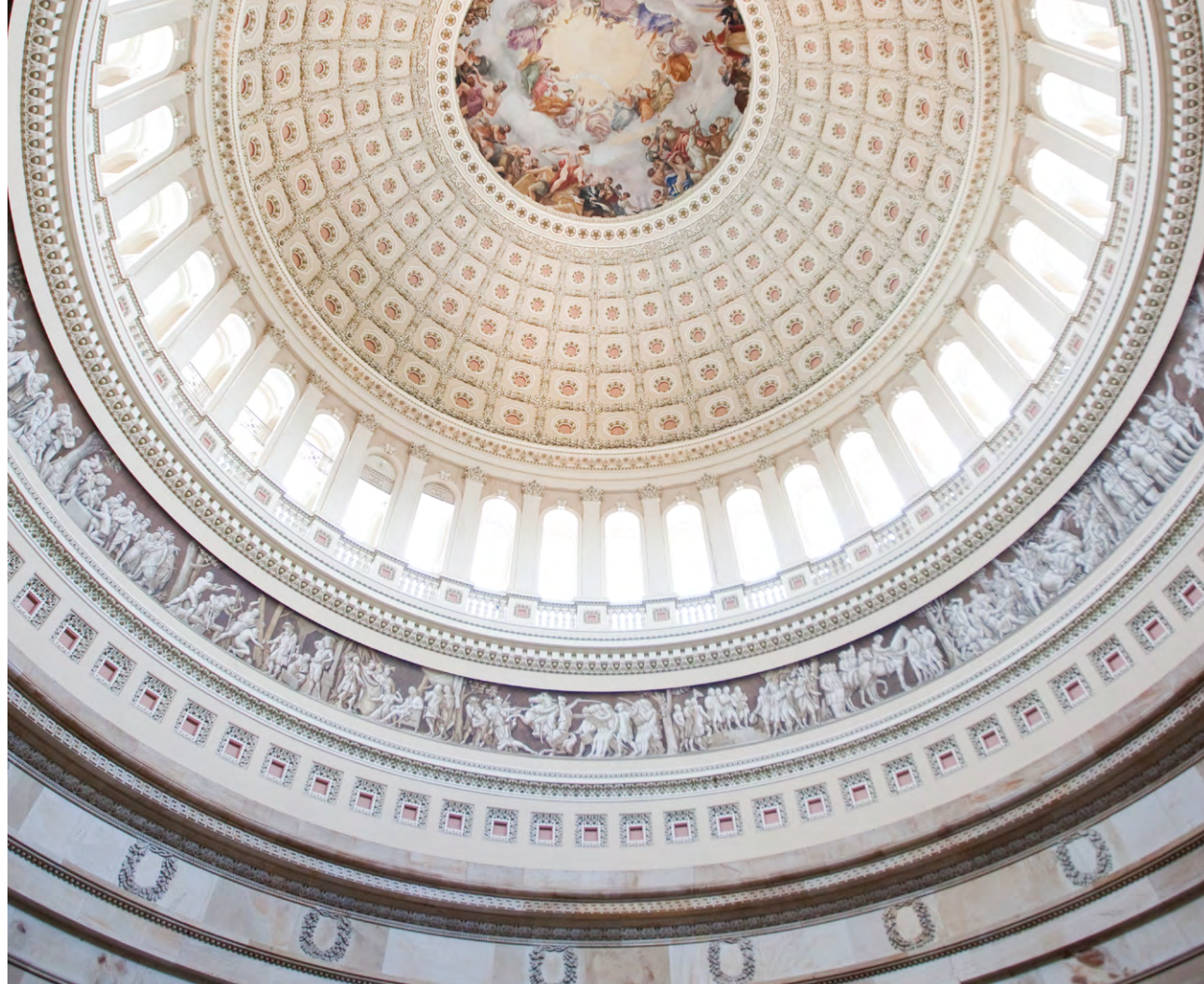




重新思考人工智能时代的 政策制定

政府如何运用人工智能强化政策设计、提升服务交付
水平及管理风险

OLIVER WYMAN
A MARSH BUSINESS

引言

为什么人工智能在政策制定中现在很重要

政策制定是政府最根本的职能之一，它将领导层的愿景和优先事项转化为塑造社会的决策。在海湾合作委员会（GCC）中，随着各国政府建立更强大的机构以将政策与国家规划和转型议程相衔接，其重要性日益凸显。整个地区范围内，政策职能正在得到加强，流程正在规范化，政策专业人员的期望值也在提高。

与此同时，人工智能（AI）正在为政策制定和实施方式创造新的可能性。它已经能够处理海量信息、综合复杂证据，并自动化耗时分析。随着AI从大型语言模型发展到多模态、具主动性以及日益通用化的系统，它有望从支持离散任务转变为赋能政府跨领域的持续感知、模拟、咨询和决策支持。在未来十年，这一转变有可能使政策制定更加快速、严谨和具有前瞻性——帮助政府改进决策，为公民带来更好的成果。

迄今为止，人工智能的采用却参差不齐。一些政府仍在早期、试探性的使用阶段，而另一些政府则积极推动人工智能在政策制定中的前沿应用。本文阐述了人工智能如何帮助政府加强政策设计、提升政策执行效率以及管理相关风险。通过借鉴全球范围内人工智能在政策制定中的早期应用案例，本文为政府如何有效且负责任地采用人工智能提供了实践视角。

人工智能为何对政策制定至关重要

传统政策制定的局限性以及人工智能带来的机遇

传统的政策制定是一项复杂的任务。它所应对的挑战在复杂性上差异很大：从相对简单的，例如推行禁烟令，到复杂的，例如养老金改革，再到真正复杂的，例如扭转生育率下降。每个案例都需要从多个来源收集充分证据，将其综合为可靠的见解，并将这些见解转化为站得住脚的政策建议。这项工作几乎在政策周期的每个阶段都依赖于大量的人力投入和专业知识：界定问题、收集证据、设计方案、咨询利益相关者以及评估结果。

人工智能可以支持更高效的政策制定。它能够自动化耗时分析，快速处理大量信息，从而减轻政策团队的手工负担，让他们能专注于最需要人类洞察力、判断力和审慎讨论的工作。潜在收益的规模已经可见：在一项评估中，一个大型语言模型平均每项研究53秒内完成了30项随机对照试验的偏见评估，而传统审查方法下估计每项研究需要28分钟。

然而，人工智能的价值超越了效率。通过使政府能够审查更多证据、测试更多选项以及更有效地识别风险，人工智能可以加强政策设计本身，并支持为公民带来更好的结果。

¹ 2024年，LAI 等人，使用大型语言模型评估随机临床试验中的偏倚风险

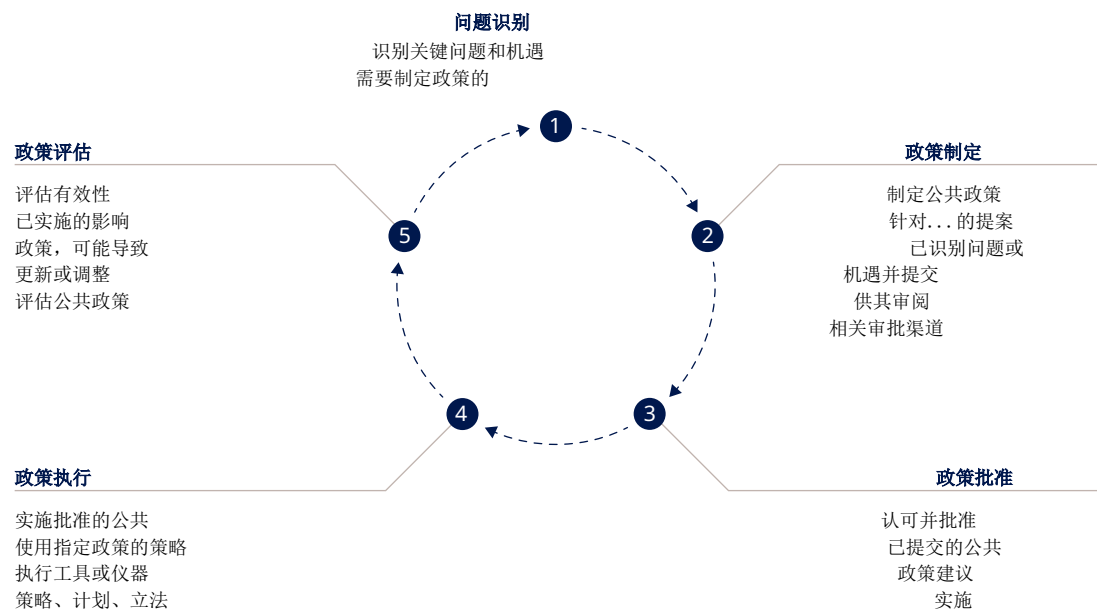
政策制定周期

一个展示人工智能如何提供帮助的五阶段模型

政策制定是指政府将其意图转化为决策、机制和行动，以实现实际成果的过程。政策是指政府为实现实际成果、协调领导层优先事项与公众需求、平衡长期目标与短期需求所采取的一系列决策、原则和措施。

本文提出一个五阶段政策制定周期，旨在明确人工智能如何强化政府能力。该周期涵盖政策制定的全过程，从问题识别到评估，并提供一个实用的框架，用于评估人工智能如何支持每个阶段。

图1：最佳实践政策制定周期



奥利弗·怀特咨询公司分析

政策制定全周期中的人工智能

如何加强问题发现、政策设计、审批、实施和评估

第一阶段 问题识别

人工智能可以通过及早揭示新兴挑战与机遇、增强远见以及支持更明智的优先级排序，来提升政府识别需要政策关注的议题的能力。通过扫描大量结构化与非结构化数据（例如媒体来源、从市场捕获的实时数据），人工智能能够检测到传统分析可能忽略的模式，包括新兴风险、情绪变化以及社会与经济变化的早期迹象。

此外，人工智能引入了一种真正全新的能力：实时预见。预测分析、预测和情景建模已经是政府中成熟的工具；人工智能带来的能力是能够持续运行这些工具，整合实时经济、市场和运营数据的信号，并在情况演变时生成洞见。这不仅仅是分析能力的提升，更代表了政策制定方式的范式转变，从周期性的、通常在问题出现后才做出反应的练习，转变为一种持续的能力，使政府能够预见、识别和在问题变得重大之前解决这些问题。然后，这种能力可以应用于综合领导层优先事项、国家承诺和新兴问题，形成关于中期应制定哪些政策的统一观点。

人工智能驱动的议会搜索与议题检测

欧洲议会的人工智能赋能搜索工具，使公民和政策制定者能够分析超过20年的议会材料，包括38,000项决议动议和议会质询，从而提高机构知识的可及性，并支持更快、更精准的问题识别。

新加坡风险评估与前瞻扫描系统

由总理办公室的国家安全协调秘书处开发，RAHS运用先进分析技术和人工智能驱动的情感分析，整合政府机构、全球数据源和研究成果的信息。通过评估新兴战略威胁与机遇，它帮助高级决策者识别哪些问题应提升至国家政策议程。

今天与明天

今天

一个部际团队每周追踪媒体报道、部际函件、公民投诉以及滞后的市场指标，以发现新兴的公共关注点和问题。

明天

人工智能系统持续扫描社交媒体、公民反馈平台、区域新闻、跨部门报告和实时市场数据，自动将新兴信号与相关政策领域进行关联。例如，在问题通常通过传统渠道呈报至部长办公桌前的几周内（例如48小时内），该系统可在短时间内识别出某特定州/省青年失业率上升的担忧，同时发现劳动力市场指标和商业活动趋势的恶化。政策制定者能在问题尚处萌芽阶段时就采取行动。

第二阶段政策制定

人工智能可在四个关键维度支持政策的制定与测试：证据综合与方案生成、跨政策及立法一致性检查、起草辅助，以及咨询意见和利益相关方反馈的分析。

政策选项设计与权衡分析

人工智能可以生成和构建政策选项，模拟替代情景，并揭示在成本、影响、可行性和公平性方面的可能权衡。

2. 经济合作与发展组织(OECD)，《运用人工智能治理》(Governing with Artificial Intelligence)，2025。

阿联酋的人工智能赋能场景建模与权衡分析

阿联酋的Arqam平台通过人工智能驱动力的预测、模拟和情景建模支持政策规划。它允许决策者使用整合的国家数据测试政策选项、评估潜在影响，并比较权衡，从而支持更具循证的政策决策。

芬兰的AI生成城市规划方案

赫尔辛基市利用UrbanistAI为关键街道生成替代性城市规划方案的可视化图。这些AI生成的渲染图帮助市民和当地企业更具体地参与到讨论中，并支持利益相关者的共识建立。

跨政策与立法影响扫描

人工智能可以在新政策最终确定之前，审查现有的政策、战略和立法，以识别其中的重叠、矛盾、空白或需要修订之处。

英国人工智能赋能的立法互引与议会回应分析

英国政府人工智能孵化器开发的Lex和Parlex工具，展示了人工智能如何已支持政策制定。Lex通过将草拟文本与现有法律进行交叉引用来识别冲突并提高起草质量，而Parlex则通过分析过去的议会辩论来预测新政策可能获得的支持，并在最终提交前识别出可能受到挑战的要点。

今天与明天

今天

一个政策团队花费三到四周时间收集证据、委托法律意见、接触少数利益相关者并起草一份初步选项文件——很大程度上依赖一个小组的分析能力，但在权衡取舍建模、测试立法一致性或处理大量公众意见方面能力有限。

明天

一个负责制定新劳动市场干预政策的政策团队在制定过程的每个阶段都使用人工智能。一款人工智能工具在夜间模拟了五种政策情景，揭示了就业、财政成本和公平性之间的权衡。一款交叉引用工具扫描了400部现行法律，并标记出与现有市一级法规的三个潜在冲突。一个咨询平台分析了14000份公众提交意见，将其聚类为八个不同的公民观点，并确定了最有可能面临有组织抵抗的两项条款。该团队带着更完善的文件、更丰富的证据和已知的风险进入高级别审查阶段——所需时间大约是传统流程的一半。

4 联邦竞争与统计中心，Arqam，2026年。5 经济合作与发展组织，《以人工智能治理》，2025年。

策略稳健性测试

人工智能可以在政策实施前帮助政府识别漏洞、漏洞、未预料到的后果以及潜在的合规规避措施。通过模拟不同参与者可能如何利用政策条款，政府可以测试政策设计的抗压能力，并增强政策的稳健性。

一次性塑料政策

在引入一次性塑料限制之前，一个AI系统会识别潜在的规避方法，包括将产品重新归类为可重复使用，以及转向豁免的塑料类别。政策制定者随后可以在实施前完善产品定义和资格标准，从而减少规避的机会。

起草与分析副驾驶支持

人工智能可在基础分析完成后，作为副驾驶加速研究、整合证据并起草政策文件，主要作为政策团队提升效率的工具。

新泽西州的人工智能支持政策研究、起草与提案制定

新泽西州的AI任务小组使用Policy Synth进行了大规模研究，整合证据和利益相关者的意见，并在20个议题领域迭代起草了1,451项政策建议。在每个阶段都经过专家评审，该工具不仅支持了备选方案的生成，还促进了全面政策建议的快速制定和优化。

咨询

人工智能可以在四个关键领域显著扩大政策咨询的规模、深度和质量：

咨询意见综合

人工智能能够处理来自参与式过程、公共服务反馈和正式咨询的大量输入，将其转化为可操作的建议。它可以识别重复出现的主题、聚类观点、检测异常值、执行情感分析，并根据既定偏好对政策选项进行排序，帮助政府揭示新兴问题、回应利益相关者的关切，并预见潜在的政策影响。

大规模网络审议与共识图谱

波利斯是一款人工智能驱动的开源软件，旨在通过聚类反馈和识别多数支持领域，使大量人群能够讨论复杂议题。它已在奥地利用于环境议题，在乌拉圭用于国家公投，在新西兰用于促进政府政策发展，以及在菲律宾用于生成市镇政策等。

参与过程中的公民引导

人工智能驱动和政府代理可以帮助公民理解咨询流程，并指导他们使用政府参与平台。

马德里参与式流程中的人工智能赋能公民指导

2021年，西班牙马德里市将其AI助手克拉拉嵌入其参与平台，以解答用户关于平台及城市参与流程如何运作的问题。这有助于使公民参与更加便捷，便于市民理解和操作。

大规模审议的节制

人工智能可以大规模地管理在线咨询流程，即使在参与量很高的情况下，也能保持其质量和包容性。

智利利用人工智能对大规模网络审议进行调解

智利参议院利用斯坦福大学的在线审议平台，支持了一场关于医疗卫生和养老金改革的全国公民咨询过程。人工智能帮助管理发言时间、组织发言队列，并围绕定时议程构建小组讨论。这使得更大规模的咨询能够以一种更有序和包容的方式进行。

合成目标片段

人工智能可以生成公民群体的合成表示，使政府能够在实施前预先测试拟议干预措施可能受到的接受程度以及不同群体可能的反应。

8 经济合作与发展组织，《运用人工智能治理》，2025年。9 经济合作与发展组织，《运用人工智能治理》，2025年。

阿联酋合成人群用于行为干预的预测试

阿联酋的行为科学小组（BSG）利用大型政府数据源中的7,652条记录以及全国代表性在线实验的4,577份回应，构建并测试了合成AI生成的参与者。这为行为干预措施的预测试、评估潜在公众反应以及尽早识别有前景的方案提供了坚实的证据基础，同时利用了一个开源模型，该模型提供了更大的控制力、灵活性以及本地部署的潜力。

第三阶段政策审批

人工智能可以通过总结要点、揭示影响与权衡，以及识别与更广泛政策环境的不一致之处，来强化政策提案的审查。它还可以支持审批机构自身的审议和决策。

政策审查支持

人工智能可以通过结构化分析和对现有政策、战略及立法的合规性检查，帮助审批者和评审机构评估提案。

美国人工智能赋能政策审查与冲突检测

美国国防部开发了GAMECHANGER，旨在帮助用户导航、比较和交叉引用大量的政策、指令和指南。通过在几分钟内而非数小时中识别重叠、矛盾、过时的条款和相关先例，它加强了政策审查、协调和审批流程。

审批机构决策赋能

人工智能可以直接嵌入审批机构的流程中，在审议过程中实时支持成员解读提案、质疑假设以及阐明权衡。

11 《行为经济学指南》，萨姆森，2025年。12 美国国防情报局，《变革者：政策与人工智能的交汇点》，2022年。

阿布扎比穆巴达拉公司推出人工智能赋能的董事会成员。

阿布扎比的穆巴达拉将MAIA正式纳入其投资委员会，以挑战提案、运行反事实场景，并揭示人类偏见可能忽略的风险，从而提升机构严谨性。穆巴达拉的长远目标是让MAIA最终成为投票成员。虽然该模型应用于投资环境，但它直接说明了人工智能如何通过强化审查、揭示权衡关系以及收紧审批前的决策来支持政策审批机构。

会前成员支持

AI代理可以被分配给个别董事会成员，用于解释提案、解答成员特定问题，并在正式会议前将观点或初步立场反馈至公共界面，从而在委员会召开时促成更实质性的审议。

今天与明天

今天

资深审阅员收到一份60页的政策草案，并花费数日时间阅读、检查其一致性，并将其与相关策略进行比较——通常在多个委员会之间并行进行，且协调有限。

明天

人工智能系统在夜间预先处理提交内容，并向每位委员会成员分发个性化简报——根据其职责范围量身定制——同时标记出与国家数字战略的矛盾之处、执法条款中的法律空白，并以通俗易懂的语言总结关键权衡。在会议期间，嵌入委员会工作流程的人工智能工具实时记录澄清问题，对有争议的假设进行反事实分析，并从可比司法管辖区中提取先例。委员会将时间用于决策，而非文件。

人工智能可以通过及早揭示新兴挑战与机遇，强化远见，并支持更明智的优先级排序，从而提升政府识别需要政策关注的议题的能力。

第四阶段政策实施

一旦一项政策获得批准，人工智能可以通过两种截然不同的方式来强化其执行：一是支持政府进行政策沟通和调整政策传递方式，二是改变公民在日常生活中的政策体验和互动方式。

沟通与响应

人工智能可以通过针对不同受众定制指导、回应公民咨询以及使政策信息更易于获取和理解，帮助政府更有效地传达政策。它还可以在政策实施过程中实时监测公共辩论和舆情，使政府能够在出现困惑、阻力或非预期影响时调整信息。这些功能共同提升了公众的理解，促进了更顺畅的执行，并加强了政府在实施阶段的责任响应能力。

政策推行期间的公共舆情AI监测

某美国地方政府在执行新冠公园使用及社交距离措施期间，使用了一个人工智能驱动的情绪监测平台。在识别出公众对某项执法措施的负面反应后，官员们通过政府自有渠道完善了信息传递，以增进公众的理解和支持。

政策体验

人工智能也在重塑公民体验、理解及参与政策的方式。个人AI代理可以向个人提醒相关政策与更新，将法律和技术语言翻译成通俗易懂的表述，并帮助人们理解政策选择中蕴含的权衡。在系统层面，AI驱动政府界面可以为公民提供统一、个性化的公共服务接入。

爱沙尼亚的Bürokratt虚拟助手

为公民提供了一个单一的AI驱动界面，用于访问公共服务，通过日常生活中标准的爱沙尼亚语的自然语言交互，帮助他们查找信息并导航政府流程。

今天

当一项新的住房法规生效时，政府通讯团队会发布一份单一的新闻稿，然后被动地监测媒体报道——只有在出现困惑或反馈后才调整信息。

明天

当一项新的住房法规生效时，人工智能工具会自动为受影响的各方——房东、租户、开发商和地方当局——生成多语言、多渠道的定制化指导。情绪监测在两周内检测到小房东中日益增长的困惑，并在合规差距出现前触发针对性的澄清活动。与此同时，公民可以通过统一的政府人工智能助手用平实的语言进行查询，获得关于法规如何影响他们的个性化答复，并通过单一界面完成所有必要步骤，无需自行导航多个机构或解读法律文本。

第五阶段政策评估

人工智能可以通过实时监控绩效、对照目标评估成果、检测非预期影响以及分析公众反馈以提供纠偏信息，从而实现更持续、更有依据的评估。

实时监控

人工智能实时追踪政策表现，评估预期成果是否达成，并识别非预期影响。

政策合规管理

人工智能可以通过识别潜在的非合规行为、检测规避模式以及实现更具针对性的执法，来强化政策合规管理。随着人工智能能力的普及，这些工具对于维持政策有效性以及减少政策体系中的漏洞将变得越来越重要。

土耳其的KURGAN平台

库尔干平台展示了高级分析和人工智能如何通过识别可疑行为、标记潜在违规行为以及实现更具针对性的执法活动，来支持大规模合规管理。

实施后评估与课程纠正

人工智能通过分析结果、将其与目标进行比较，并推荐改进或调整来支持评估。

经合组织的人工智能赋能政策评估

经合组织（OECD）运用人工智能，分析了1998年至2022年间41个国家超过1500项环境政策干预措施，以识别哪些措施最有效地减少了排放。其核心发现是，政策组合而非单一措施更能产生显著效果，这进一步凸显了战略政策设计的重要性。

世界银行的AI驱动项目评估

世界银行采用无监督机器学习方法，分析了64个国家的392份项目报告，以揭示影响项目成功与失败的因素。该案例展示了人工智能如何从大型评估数据集中发掘出仅通过人工审查可能遗漏的洞见。

今天与明天

今天

项目启动后 12 至 18 个月，会委派政策评估小组。该小组需手动从多个机构收集数据，协调不一致的格式，并产出一份回顾性评估报告，但这份报告往往在需要调整时已过时，无法提供有效指导。

明天

评估从项目启动之初就融入了实施过程。人工智能系统实时追踪结果指标——包括参与率、获取公平性、投诉模式及支出效率——并生成月度仪表盘，附带异常警报。当数据显示一项职业培训项目在实施六个月后表现不佳时，该部门会在下一个预算周期开始前介入。在项目组合层面，人工智能系统扫描数百项过往干预措施的结果，以识别哪些政策组合始终能带来更优成果，从而揭示出任何人工审查流程都难以实际产生的跨项目洞察。随后，它会将这些发现直接输入下一轮政策设计中。

¹⁶ 经济合作与发展组织，《运用人工智能治理》，2025年。¹⁷ 经济合作与发展组织，《运用人工智能治理》，2025年。

人工智能可以通过总结要点、揭示影响与权衡，以及识别与更广泛政策环境的不一致之处，来强化政策建议的审查。

人工智能进入政策制定需管理的五项风险

展望未来，随着人工智能系统在关键分析任务上的能力超越人类，人类判断虽然在形式上仍会参与其中，但在实践中其作用将逐渐缩小。政策制定者可能仍保留决策权，但会失去亲自测试人工智能生成建议的逻辑和影响的能力。随着时间的推移，这可能会削弱监管，并促使政策制定从人类判断转向对机器输出的验证。因此，政府需要决定人类判断必须在哪些方面保持决定性作用，应如何对先进系统设定限制，以及如何在人工智能塑造更大比例的政策选择时维持公众信任。

与此同时，各国政府在当前利用人工智能制定政策时，也必须管理一系列更紧迫的风险：

有偏见的或低质量的输入数据

当AI在训练数据不完整、过时、偏差或不具代表性时，它可能会产生误导性的政策见解。在政策制定中，这是危险的，因为薄弱的输入可能会在看似严谨和有据可依的同时，强化盲点。因此，政府需要强大的数据治理，这包括具有代表性的数据集、部署前的数据质量测试，以及诸如检索增强生成（RAG）框架等安全措施，这些措施在生成回应前会参考外部可信知识库。

英国关于公共部门人工智能可靠数据的指南

英国政府的AI实施指南，由政府数字服务部门制作，旨在帮助公共部门安全、合乎道德地采用人工智能，建议在事实必须保持最新和准确的应用场景中使用RAG框架。该指南还提供了关于公共部门人工智能数据集准备方面的指导。这些措施共同有助于提升人工智能支持的分析和可靠性。¹⁸

自动化偏见、幻觉与现实错误

决策者可能因为人工智能的输出看似中立、技术化或客观，而对其过度信任，即使这些输出是不准确的、具有误导性或被伪造的。在政策制定中，这种情况会减少审查，排挤人类判断，并扭曲复杂的权衡。因此，政府需要为高风险决策保持人工监督，将人工智能视为副驾驶而非自主作者，确保输出基于经批准和受信任的来源，要求进行事实核查，并将挑战和复核步骤纳入工作流程，以便用户对输出进行检验而非被动接受。

欧盟关于高风险人工智能人类监督的规定

欧盟正帮助其成员国政府防范在重大决策中对人工智能的过度依赖。欧盟人工智能法案第26条规定，高风险人工智能系统的部署者必须指定具备必要能力、培训和权限的人员进行人工监督。其目的是确保在决策具有重大后果的情况下，人类判断仍然处于活跃和负责的状态。

不透明模型和有限的透明度

一些人工智能工具无法解释其生成内容的推理过程，这给用户理解输出是如何产生带来了困难。在政策制定领域，这是一个严重的问题，因为决策通常需要在众多不同的受众面前进行解释和辩护。因此，政府应要求透明度、文档记录、审计追踪，以及与具体应用场景重要性相匹配的、关于如何解释输出的标准。

英国算法透明度标准

算法透明度记录标准是一项在英国中央政府强制推行的规范，它为公共部门组织提供了一个清晰且一致的框架，用以说明其使用人工智能及其他算法工具的方式与原因。

数据安全、保密性及数据主权风险

人工智能在政策制定中的使用可能会泄露敏感政策材料以及公民的机密数据。当用户将敏感信息输入外部工具或依赖外国拥有的平台时，可能会发生泄露。内部人工智能系统也可能扩大威胁的攻击面，为网络钓鱼、身份冒充、深度伪造和有组织的操纵创造新的漏洞。这些违规行为可能会扰乱内部工作流程并损害公众信任。

19 欧洲委员会，《人工智能素养：问答》，2025年。²⁰ Gov.uk，《算法透明度记录标准中心》，2025年。

因此，政府应在使用前对数据进行分类，将敏感工作限制在批准的环境中，并明确主权和供应商控制要求。它们还应将人工智能视为一个网络风险领域，通过保护系统和供应链、测试在线服务是否可能被欺骗、操纵或滥用，以及建立深度伪造检测能力来应对。

新加坡对安全生成式人工智能使用的标准

新加坡政府已发布标准，旨在提升公共部门对人工智能相关安全风险的管理。该标准指出，在新加坡托管的生成式人工智能API服务，其使用仅限于特定的敏感级别，并需进行分类检查和授权环境要求。

排斥与参与不平等

利用人工智能进行政策制定可能会无意中偏袒那些精通数字技术并拥有技术接入能力的人，而将那些不具备这些条件的人排除在外。这会扭曲参与度并削弱其合法性。因此，政府从一开始就应致力于包容性建设，提供多语言接入、低摩擦界面、语音通道以及支持广泛参与的安全保障措施。

马德里AI支持参与平台

由马德里市议会发起的“决定马德里”（Decide Madrid）是一个利用人工智能鼓励公众广泛参与决策的在线平台。该平台允许市民提交提案、就城市事务进行辩论和投票，并使用人工智能虚拟助手为用户提供建议。该平台也已集成到市议会的移动应用程序中，以便市民更便捷地参与公共事务。

21 新加坡政府技术部门，[生成式人工智能系统安全计划](#)，2026年。22 马德里市政府，[推出“克拉拉”，Decide Madrid虚拟助手](#)，2021年。

将人工智能融入政策制定过程中的关键考量因素

审计你所拥有的数据

人工智能工具的效用取决于其所依据的数据。在部署任何工具之前，政府都应梳理相关实体所持有的数据：哪些是结构化的，哪些是非结构化的，哪些是当前的，以及哪些仍因机构边界而形成数据孤岛。应尽早识别最显著的数据差距和质量问题。若缺乏这一基础，人工智能则可能自动化低质输入，并生成看似严谨但基于不可靠证据的政策分析。

在部署任何工具之前，定义人类角色

对于每种应用场景，政府都应明确：哪些决策权保留给政策专业人员、哪些AI输出需要在进入政策流程前进行审核，以及当输出结果出错时由谁负责。这不是一项合规步骤，而是负责任部署的基础。那些回避此问题的机构往往只有在出现有缺陷的输出并已影响政策决策后，才会面临这一问题。

在可控的低风险政策环境下试飞

人工智能工具应首先在特定环境中进行测试，例如单一政策团队、非敏感用例以及限定的时间范围内。其目的并非证明该工具在原则上有效，而是了解其在机构自身数据、工作流程和决策文化中的表现。在试点项目结束后，应进行结构化评估，涵盖输出质量、用户采纳和潜在风险，此评估应先于任何规模化决策。

在扩大规模之前设定质量和透明度标准

在人工智能工具从试点阶段转向更广泛使用之前，政府应明确界定如何验证其输出结果、需要什么样的审计追踪记录，以及如何披露人工智能的使用情况，这既包括向决策者披露，也应在适当的情况下向公众披露。每个应用场景都应指定一个有名称的机构负责人，并授予其监控性能和干预的权力，以防输出质量恶化。在没有建立这些标准的情况下进行规模化推广，一旦嵌入后，将产生难以撤销的机构和声誉风险。

为人工智能辅助的决策制定调整运营模式

人工智能赋能的决策制定可能需要政府调整其运营模式、治理结构和决策流程，以支持更快速、更持续的政策周期。若缺乏相应的制度变革，政府将面临难以跟上日益人工智能化环境的速度和复杂性的风险。

同步提升内部素养与部署

如果使用AI工具的政策专业人士无法批判性地评估这些工具生成的内容，那么这些工具的表现就会不佳。政府应投资于其团队质疑AI生成分析的能力，以理解结果是如何得出的，识别不合理的输出，并在必要时判断是否应推翻系统。这种能力不能通过一次性培训来建立；它必须作为政策团队运作的持续特征来发展和维护。

建立反馈回路并规划迭代

任何人工智能部署在初次引入时都不会表现最佳。政府应建立结构化的机制，以捕捉哪些有效、哪些无效以及哪些需要调整，并借鉴政策专业人员的意见。这些反馈应被用于持续优化工具、调整工作流程以及随时间更新监管框架。那些在政策制定中从人工智能中获得最大价值的政府，是将采用视为一项持续的制度能力，而非一次性的实施。

结论

人工智能代表了政府在制定和实施政策方面的一种范式转变。

在人类现代史的大部分时间里，政策制定的步伐和质量都受到政策专业人员的分析能力、他们所能审查的证据数量、所能测试的选项数量、他们预测变化的速度以及他们所能综合的反馈广度的限制。人工智能消除了许多这些限制。它使政府能够更早地发现问题，以更严谨的方式设计政策，以更精准的分析支持决策，以更高的精度进行沟通和实施，并能够持续而非定期地评估结果。其结果并非对现有流程的边际改进，而是在政策周期每个阶段所能实现的事情上的根本性变革。

本文中的案例表明，各国政府正开始在整个政策周期中应用人工智能，但大多数应用仍处于早期阶段且范围狭窄。要充分释放人工智能的潜力，需要强有力的治理、可靠的数据、明确的规范、合适的制度能力，以及在高风险决策中坚守人类判断的不渝承诺。

政府无需等到技术完全成熟才采取行动。更有效的做法是：从现在开始，在明确领域内，以清晰的规则、适当的监管以及务实地看待收益与风险为前提，着手推进。人工智能正重塑政策的设计、实施与评估方式。对于海湾合作委员会成员国及更广泛地区而言，当务之急是主动介入，构建制度基础，以塑造而非仅仅应对下一阶段的政策制定时代。

威灵顿咨询公司（Oliver Wyman）， Marsh（纽约证券交易所代码：MRSH）旗下的企业，是一家深谙行业洞察、勇于创新、并采用协作方式简化复杂问题，以帮助组织应对其最具定义性的转型时刻的管理咨询公司。Marsh是全球在风险、再保险和资本、人力资源与投资以及管理咨询领域的领导者，为130个国家的客户提供建议。凭借270亿美元的年营收和超过9.5万名同事，Marsh帮助客户借助其独特的视角建立信心，从而蓬勃发展。

欲了解更多信息，请访问 oliverwyman.com，或在领英（LinkedIn）和推特（X）上关注我们。

美洲
+1 212 541 8100

欧洲
+44 20 7333 8333

亚太地区
+65 6510 9700

印度、中东和非洲
+971 (0) 4 425 7000

作者

勒亚·夏尔

合作伙伴、政府及公共机构

南·马修

合作伙伴、政府及公共机构

阿比希克·夏尔马

合作伙伴、政府及公共机构

丹妮·萨亚尔

校长、政府及公共机构

版权所有 2026 OliverWyman

版权所有。未经 Oliver Wyman 书面许可，本报告不得全部或部分转载或重新分发。
且， Oliver Wyman 对此方面第三方行为的任何责任概不承担。

本报告中的信息及观点由 Oliver Wyman 整理。本报告不构成投资建议，亦不应被视为投资建议或替代就相关事宜咨询专业会计师、税务师、律师或财务顾问。Oliver Wyman 已尽合理努力使用可靠、最新及全面的信息和分析，但所有信息均未经保证，亦不构成任何明示或暗示的保证。Oliver Wyman 不承担更新本报告中信息或结论的责任。对于因依据本报告或本报告中引用的任何其他报告或信息来源所采取或未采取的任何行动而导致的任何损失，或对于任何后果性、特殊或类似损害，即使已被告知此类损害的可能性，Oliver Wyman 亦不承担责任。本报告不构成买卖证券的要约或招揽买卖证券。未经 Oliver Wyman 书面同意，本报告不得出售。

奥利弗·怀特曼，马什业务

www.oliverwyman.com