SLAUGHTER AND MAY/

2024 年 7 月 2 日

香港私隐专员公署发布采购、实施及使用人工智能的新模范框架

香港个人资料私隐专员公署（**私隐专员公署**）最近公布了《人工智能(AI)：个人资料保障模范框架》（**《模范框架》**），提出了最佳行事规范的建议做法，协助机构保护个人资料私隐，并确保安全、合乎道德及负责任地使用创新科技。

私隐专员公署于 2021 年发布的《开发及使用人工智能道德标准指引》（**《AI 指引》**）与《模范框架》有所不同，其侧重对象是人工智能（**AI** 或**人工智能**）系统开发商及供应商（**AI 供应商**），而《模范框架》针对的则是有意在其业务运营中采购、实施及使用涉及个人资料的 AI 系统的本地机构（**机构**）。因此，《模范框架》所涉及的香港企业群体极为广泛。

机构应根据自身的风险状况，了解并尽力遵循《模范框架》中的建议做法，即使是有意从第三方供应商处采购 AI 方案（无论是现成的还是定制的）用于其业务中。

总体原则

《模范框架》与《AI 指引》保持一致，并继续应用《AI 指引》中倡议的三项数据管理价值和七项 AI 道德原则：

	数据管理价值	AI 道德原则
1	*尊重*	1. **问责：**机构有责任就其 AI 做出的行为负责。 2. **人为监督：**人为参与的程度应与使用 AI 系统的风险及影响相称。 3. **透明度与可解释性：**机构应披露它们使用 AI 以及相关的数据私隐措施，同时致力改善 AI 辅助决策的可解释性。 4. **数据私隐：**机构应根据六项保障资料原则[1]（**保障资料原则**）实施有效的数据管理。
2	*互惠*	5. **有益的人工智能：**AI 应当带来益处，同时应避免或减低造成的伤害。 6. **可靠、稳健及安全：**AI 应当运行可靠，能够抵御错误和攻击，并制定应变计划，在 AI 系统不能正常运作时作好准备。
3	*公平*	7. **公平：**个人应受到合理的平等对待。差异化对待都应有正当理由支持。

《模范框架》

私隐专员公署建议在以下四个领域采取措施：**(1) AI 策略及管治架构**；**(2) 风险评估及人为监督**；**(3) AI 模型的定制及实施 AI 系统的管理**；以及**(4) 与利益相关方的沟通及交流**：

AI 策略及管治架构

机构应制定内部的 AI 管治策略，包括：

[1] 《个人资料（私隐）条例》（第 486 章）（**《私隐条例》**）附表 1。

- 制定**AI策略**，界定AI系统在机构中提供的功能，为采购、实施及使用AI提供指引，建立制度化的决策过程和上报准则，确保有适当的基础技术设施来支持AI的发展；

- 采购AI方案的**管治考虑**，包括对潜在AI供应商的主要私隐和安全义务、能力以及对技术性和管治方面的国际标准的遵守情况进行尽职调查。如果AI的定制和使用涉及在数据中心分布于多个司法管辖区的云平台上处理个人资料，机构应考虑跨境转移资料的合法性，并采用合约规范手段防止资料处理者未获准许或意外查阅、处理、丢失或使用个人资料[2]；以及

- 建立具足够资源、专业知识和决策权的**内部管治架构**，并明确内部汇报关系，以引领AI策略的实施，监督AI方案的整个生命周期。其中包括：有高级管理层参与和跨专业领域合作的AI管治委员会，向董事会进行汇报。此外，还应为所有相关人员提供充分的培训，以确保他们具有适当的知识、技能和认识，以便使用AI系统工作。

风险评估及人为监督

机构需要进行全面的风险评估，有系统地识别、分析及评估AI系统整个生命周期的风险。

相关评估，包括制定作出最终决定的理由，应由一个跨部门团队进行，并根据本机构的AI政策进行审查，同时要有妥善的记录存档，及应涵盖任何从数据私隐角度来看的风险以及对个人法律权利、人权（包括私隐权）、就业或教育前景以及他们使用服务的资格的影响。一旦发现潜在风险，机构应采取相应的风险缓解和管理措施，包括决定在使用AI系统时所需的适当人为监督的程度。

AI模型的定制及实施AI系统的管理

在准备定制和使用AI的数据集时，机构应采取措施确保遵守《私隐条例》，包括尽量减少用于此类目的的个人资料数量，妥善记录任何此类资料的处理情况，并实行规定可输入AI系统的内容及提示的员工内部指引。

机构应测试AI模型，确保其稳健、可靠和安全，尽量降低攻击、错误或故障风险，并验证其可靠性、公平性、可解释性以及是否符合私隐和道德要求。机构还应审查AI生成的内容（例如，标记和过滤有害内容），并应在将AI方案与机构系统整合之前，进行严格的用户接受度测试。

AI方案的服务器托管地点不容忽视。机构应就自身是否有足够的专业知识安全地运作及保护内部服务器上的系统与在第三方云服务器上处理个人资料的相关风险作出权衡。

由于与使用AI有关的风险因素可能会随着时间的推移而发生变化，因此必须通过定期进行内部审核并向高级管理层汇报审核结果，对AI系统进行持续监控和管理。

实施具有开源组件的AI方案的机构，还应该在维护代码和管理安全风险方面持续遵守行业最佳安全规范的做法，并留意安全建议和警报。

最后，机构应制定AI事故应变计划，以监测和应对可能发生的事件。该计划应界定应报告的事件，并制定有关通报、遏止和调查事件的政策、程序和准则。

与利益相关方的沟通及交流

机构应就AI的使用与利益相关方，尤其是内部员工、AI供应商、个人客户和监管机构进行清晰、有效和定期的沟通和交流。这包括向利益相关方解释AI作出的决策和生成的结果，以及在此过程中是否使用了他们的个人资料。此外，还应建立用户反馈途径，以帮助改进AI系统。

[2]保障资料原则之第4(2)原则的规定。根据《私隐条例》，资料处理者是指符合以下两项说明的人：(a)代另一人处理个人资料及(b)并不为该人本身目的而处理该资料。

使用 AI 处理个人资料的机构，应考虑 AI 供应商是否会更适合满足任何数据查阅或改正请求以及向资料当事人解释 AI 的决策和生成结果。

关注要点与相关建议

尽管《模范框架》不具有法律效力，但仍建议机构尽可能严格遵守，以管理与 AI 技术使用相关的风险。

有意在其业务运营中采用 AI 技术并获得相应利益的企业，应在采购、实施及使用该技术的过程中对个人资料的处理负责。有鉴于此，《模范框架》倡导企业制定政策和程序，以确保合法、负责任和高质量地使用 AI 方案。

机构可咨询相关法律专家，以确保遵守新出台的适用于采购、实施及使用人工智能的法律法规。如果 AI 技术的实施和使用涉及跨不同司法管辖区的数据处理和资料转移，获得法律人士的协助也尤为重要。

联系人

莫宜咏
合伙人
T: +852 2901 7201
E: wynne.mok@slaughterandmay.com

郑诺铭
律师
T: +852 2901 7211
E: jason.cheng@slaughterandmay.com