2024 年 6 月法规通讯

附录二　私隐公署建议概要

（注：页面参考按私隐公署文件）

I. AI 策略及管治（第 10 页）

- **AI 策略**
  - **主要原则：问责**
  - 展示**高级管理层有决心**通过**合乎道德标准及负责任**的方式采购、实施及使用 AI
  - 提供**以下指引**
    - 采购 AI 方案的**目的**
    - 如何**实施**和**使用** AI 系统
  - **要素**
    - 如：界定 AI 系统提供的功能、制定道德原则、列明不可接受的用途
- **关于采购 AI 方案的管治考虑**
  - 通常涉及聘请第三方**定制** AI 系统或购买**现成**的 AI 系统
  - 管治考虑（第 12 页）
    - 如：使用 AI 目的、私隐和保安的主要责任及道德规定、技术性和管治方面的国际标准
- **管治架构**
  - **主要原则：问责/人为监督**
  - 建立具有足够资源、专业知识和决策权的**内部管治架构**
  - **AI 管治委员会**
    - **监督**所有 AI 方案的**整个生命周期***（由采购、实施、使用以至终止）*，并**向董事会汇报**
    - 监督应**横跨整个业务**，而不受部门划分*（如：风险和合规）*
    - **高级管理层**参与及**跨专业领域合作**
    - 指派**高级管理人员(C-level)领导**
    - 角色及责任的例子（第 16 页）

- **培训及加强认识**
  - **主要原则：问责**
  - 如 AI 系统使用者：遵从资料保障法律/规例/内部政策、网络保安风险、一般 AI 科技
  - 其他人员的培训，如：法律及合规专业人员（第 18 页）
  - **审查员**的角色（第 18 页）

II. 风险评估及人为监督（第 20 页）

- **总体：以风险为本的方式**
  - 在**采购、使用及管理** AI 系统时
  - **全面的风险评估**
    - **识别**、**分析**及**评估**风险*（包括私隐风险）*
  - 建立/实施/记录存档/维持**风险管理机制**
    - 在 AI 系统的**整个生命周期**内
  - **跨部门团队**进行风险评估
    - 在**采购**过程时或有**重大更新**时
  - 所有风险评估应妥善**记录存档**
- **须考虑的风险因素**
  - **主要原则：有益的 AI/数据私隐/公平**
  - **风险评估**时应考虑的**主要因素**
    - 如：资料的准许用途（**保障资料第 3 原则**）

      （在保障资料第 3 原则下，未得资料当事人的订明同意，个人资料不得用于新目的）
    - 如：个人资料的数量（**保障资料第 1 原则**）

      （在保障资料第 1 原则下，所收集的个人资料就收集目的而言，属足够但不超乎适度）
- **决定人为监督的程度**
  - **主要原则：人为监督(human oversight)**
  - **以风险为本**的方式**：缓减风险**措施的类别/程度与风险**相称**
  - **人为监督**是减低使用 AI 的风险的主要措施
  - 可能带来较高风险的 AI 应用例子（第 25 页）
- **减低风险的权衡**（第 25 页）

III. AI 模型的定制与 AI 系统的实施及管理（第 27 页）

- **为定制及使用 AI 准备数据**
  - **主要原则：数据私隐/公平**
  - **数据准备**的 4 个范畴：
    - 遵循《私隐条例》的规定
    - 尽量减少定制及使用 AI 所涉及的个人资料
    - 管理用以定制及使用 AI 模型的数据
    - 妥善记录处理数据的情况
- **AI 方案的定制及实施**
  - **主要原则：透明度与可解释性/可靠、稳健及安全**
  - **按照**所涉的**风险程度**，**严格测试及验证**
  - 建议采取的措施（第 32 页）
- **AI 系统的管理与持续监察**
  - **主要原则：可靠、稳健及安全/人为监督**
  - **持续监察**及**检视**
    - 风险因素或会改变
  - 考虑制定 **AI 事故应变计划***（第 38 页：详细要素）*

IV. 与持份者的沟通及交流（第 40 页）

- **主要原则：透明度与可解释性**
- **提供资讯**
  - 定期及有效地与持份者联络及交流
    - 尤其是**内部员工**、**AI 供应商**、个别**消费者**及**监管机构**
  - **透明程度**根据持份者而**改变**
  - 有效的沟通对于**建立信任**至关重要
- **资料当事人的权利及反馈**
- **可解释的 AI(Explainable AI)**
  - **建立**持份者**信任**的关键
  - 让 AI 的决策及输出结果达致可解释性

- **语言及方式**
  - 浅白的语言，清楚易明的方式