

# 開發及使用 人工智能道德標準指引



PCPD



[PCPD.org.hk](http://PCPD.org.hk)

香港個人資料私隱專員公署  
Office of the Privacy Commissioner  
for Personal Data, Hong Kong



1996-2021  
守護・私隱・廿五載  
GUARDIAN · PRIVACY · 25 YEARS

# 合乎道德標準的 人工智能

《開發及使用人工智能道德標準指引》為開發人工智能或在營運中使用人工智能的機構而編製。該指引旨在促進人工智能在香港健康發展和使用，並協助機構在開發及使用人工智能時遵從《個人資料(私隱)條例》(第486章)的規定。本小冊子重點列出該指引的核心建議。

三項數據  
管理價值

七項人工智能的  
道德原則

實務指引：  
四項主要業務流程

二項數據管理價值

**尊重**

持份者的權利、  
利益及合理期望



**互惠**

為持份者帶來益處  
及減低對持份者  
造成的傷害



**公平**

避免不義的偏見及  
非法的歧視



# 七項人工智能的道德原則

## 問責

機構應對其作為負責，並能夠就其行為提供充分的理據

## 透明度與可解釋性

機構應向持份者披露它們使用人工智能及相關的政策，並致力改善自動化和人工智能輔助作出的決定的可解釋性

## 公平

使用人工智能時應避免不義的偏見及非法的歧視

## 可靠、穩健及安全

人工智能系統應能可靠地運作，能夠應付錯誤，並得到保護而免遭攻擊



## 人為監督

在人工智能的運作過程中，人為參與的程度應與使用人工智能的風險及影響相稱

## 數據私隱

機構應具備有效的數據管治，在開發及使用人工智能時保障個人的個人資料私隱

## 有益的人工智能

使用人工智能應為持份者帶來益處及減低對持份者造成的傷害

## 實務指引：四項主要業務流程

合乎道德標準地開發及使用人工智能的四項主要業務流程



### 制定

人工智能策略及管治



### 進行

風險評估及人為監督



### 實行

人工智能模型的開發及  
人工智能系統的管理



### 促進

與持份者的溝通及交流

出現重大改變時  
重新評估風險

調整人工智能系統  
以應對持份者的關注



## 制定

人工智能策略及管治

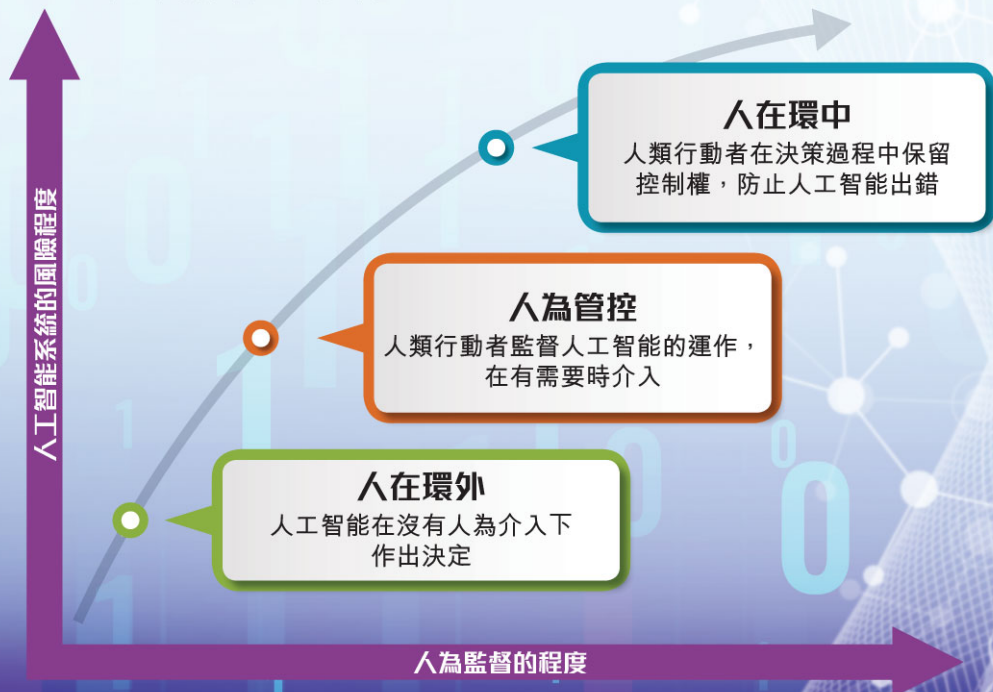
- 由高層作出指導
- 人工智能管治委員會監督人工智能的開發及使用
- 職員培訓及加強認識



## 進行

風險評估及人為監督

- 及早進行全面的風險評估，考慮：
  - ▶ 人工智能對持份者的潛在影響是甚麼？
  - ▶ 如影響出現，有多嚴重？
  - ▶ 影響出現的可能性有多大？
- 以風險為本的方式採取適當的風險緩減措施
- 決定適當程度的人為監督





## 實行

人工智能模型的開發及人工智能系統的管理

- 確保用以訓練人工智能模型的數據質素及合法使用數據
- 測試人工智能模型，以確保可靠、穩健及公平
- 持續監察及檢視人工智能系統





## 促進

與持份者的溝通及交流

- 與持份者保持具透明度的溝通
- 設立與持份者交流的途徑
- 通訊以淺白語言、清楚易明的方式書寫



與持份者的  
溝通



披露使用  
人工智能系統



提供充足  
資訊



披露風險



與持份者  
的交流



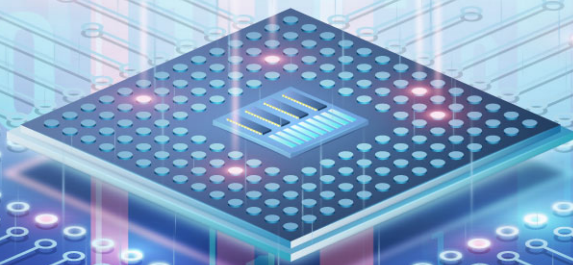
容許退出及  
改正不準確之處



按要求  
提供解釋



提供人為介入  
的選擇





私隱公署網頁  
PCPD website  
pcpd.org.hk

查詢熱線 Enquiry Hotline : (852) 2827 2827  
傳真 Fax : (852) 2877 7026  
地址 Address : 香港灣仔皇后大道東248號大新金融中心13樓1303室  
Room 1303, 13/F, Dah Sing Financial Centre,  
248 Queen's Road East, Wanchai, Hong Kong  
電郵 Email : [communications@pcpd.org.hk](mailto:communications@pcpd.org.hk)



下載本刊物  
Download  
this publication



本刊物使用署名4.0國際(CC BY 4.0)的授權條款，只要你註明原創者為香港個人資料私隱專員，便可自由分享或修改本刊物。詳情請瀏覽[creativecommons.org/licenses/by/4.0/deed.zh](https://creativecommons.org/licenses/by/4.0/deed.zh)。

This publication is licensed under Attribution 4.0 International (CC BY 4.0) licence. In essence, you are free to share and adapt this publication, as long as you attribute the work to the Privacy Commissioner for Personal Data, Hong Kong. For details, please visit [creativecommons.org/licenses/by/4.0](https://creativecommons.org/licenses/by/4.0).

#### 免責聲明 Disclaimer

本刊物所載的資訊和建議只作一般參考用途，並非為法例的應用提供詳盡指引，亦不構成法律或其他專業意見。私隱專員並沒有就本刊物內所載的資訊和建議的準確性或個別目的或使用的適用性作出明示或隱含保證。相關資訊和建議不會影響私隱專員在《個人資料(私隱)條例》下獲賦予的職能及權力。

The information and suggestions provided in this publication are for general reference only. They do not serve as an exhaustive guide to the application of the law and do not constitute legal or other professional advice. The Privacy Commissioner makes no express or implied warranties of accuracy or fitness for a particular purpose or use with respect to the information and suggestions set out in this publication. The information and suggestions provided will not affect the functions and powers conferred upon the Privacy Commissioner under the Personal Data (Privacy) Ordinance.