Paper 1-B Mock Exam 4 — Q2

學生版 5 精簡 Study Notes：AI 生成內容 × Deepfake × 強制標示**

主要單元：媒體倫理
次要單元：人權 / 商業及經濟倫理
分數：20 分

本題問：香港是否應強制所有主要由 AI 生成或大幅修改的公開媒體內容，清楚標示「AI 生成 / AI 修改」。核心不是「AI 好不好」，而是：

當影像、聲音和影片已不能自動證明事情真實時，媒體製作者和平台應承擔甚麼透明責任?

1. Truthfulness 真確性

媒體倫理要求：

不要故意令受眾對重要事實產生錯誤理解。

AI 虛構本身不一定欺騙，例如清楚標示的模擬影片。

但若 AI 令真人看似說過從未說過的話，又沒有標示：

Fiction concealed as fact may become deception.

---

2. Accuracy ≠ Authenticity

Accuracy 準確性：資訊是否正確。
Authenticity 真實來源：影像、聲音是否真的來自該人物或事件。

例如 AI 主播報道正確天氣：

資訊可以 accurate，但主播並不 authentic。

因此：

Accurate content ≠ Authentic content

---

3. Right to Know 知情權

受眾需要知道：

©PETER HSU 2026

- 內容是否 AI 生成;
- 是否重大修改;
- 是否真實人物;
- 是否真實事件。

清楚標示不是禁止內容，而是讓受眾作:

Informed Choice 知情選擇

所以標示可以增強自主，而非削弱自主。

---

4. Freedom of Expression 表達自由

AI 可以用於:

- 藝術
- 教育
- 電影
- 廣告
- 翻譯
- 虛擬角色
- 特效

因此不能認為:

AI use = deception

如果只是 AI 字幕、修光、背景修改，未必有欺騙問題。

真正應問:

Could a reasonable audience be misled about reality?

---

5. 支持強制標示

©PETER HSU 2026

Protect Truth

AI 愈逼真，觀眾愈難自行辨識，所以創作者透明責任愈高。

Protect Autonomy

標示讓觀眾知道內容性質，再自行決定是否相信或分享。

Protect Reputation

Deepfake 可令人看似：

- 支持某產品；
- 說過某些話；
- 參與某事件。

因此可能損害名譽、私隱及人格尊嚴。

Preserve Media Trust

如果人人相信「任何影片都可能是假」，最後真實影片也會失去可信度。

---

6. 反對一刀切標示

Overbreadth

如果 AI 只用來：

- 字幕
- 修圖
- 文法修改
- 色彩調整

都必須大型標示，可能過度。

AI Is a Tool

傳媒一直使用 CGI、剪接、Photoshop 和特效，並非所有修改都需要警告。

Labelling ≠ Truth

©PETER HSU 2026

必記：

AI-generated ≠ false
Human-generated ≠ true

真人影片也可以被剪輯、斷章取義或誤導。

Innovation Cost

過度規管會增加創作者、平台和小企業成本。

---

7. 5概念：Liar's Dividend**

AI 不只令：

假的看似真的。

也可能令：

真的被說成假的。

例如某人真實影片曝光後，可以聲稱：

「那是 AI 做的。」

這叫：

Liar's Dividend

因此 AI 時代要保護的不只是「避免假資訊」，也是整個社會對真實證據的信任。

---

8. Responsibility Triangle

Creator 創作者

- 不故意欺騙
- 適當標示

Platform 平台

- 提供標示功能
- 對高風險 deepfake 加強監管

©PETER HSU 2026

- 不刻意放大欺騙內容

Audience 觀眾

- 查證來源
- 培養 media literacy
- 不輕率分享

但:

Audience responsibility does not erase creator responsibility.

---

9. Proportionality 比例原則

Legitimate Aim

保護真實資訊和知情權。

Rational Connection

AI 標示有助辨識合成內容。

Necessity

是否所有 AI 使用都需要標示?

→ 未必。

Balance

透明度的利益是否值得限制創作自由?

→ 應按風險決定。

因此 5**答案通常不是:

Label everything

而是:

Risk-based Labelling

©PETER HSU 2026

10. 最佳政策：Risk-based Labelling

Low Risk

AI 字幕、文法、光線修正
→ 不必強制標示。

Medium Risk

主要人物或場景由 AI 生成，但明顯屬藝術 / 娛樂
→ 簡單標示。

High Risk

AI 模擬真人：

- 聲音
- 新聞
- 公眾事件
- 商業代言
- 專業意見

→ **強制清楚及不可移除標示。**

---

11. Gospel-Centric Reasoning

Truth

不能刻意把虛構冒充真實。

Imago Dei

人的肖像、聲音和身份不能只被視為可任意操控的數碼素材。

Creativity

人的創意和科技能力可以是恩賜。

Fall

好的工具也可能被欺騙、貪婪和操控扭曲。

©PETER HSU 2026

Love of Neighbour

不應利用別人的信任取得流量或利益。

Stewardship

平台和創作者要對自己的傳播能力負責。

人的創造自由不能脫離真理、愛鄰舍和尊重他人的責任。

---

12. 5 Skeleton Answer**

Introduction

我支持高誤導風險的 AI 內容必須清楚標示，但不支持所有 AI 使用都受相同規管。

P1 — Truth

未標示 deepfake
→ 虛構被當成事實
→ 違反 truthfulness。

P2 — Autonomy

標示
→ 受眾知道內容性質
→ 可以作 informed judgment。

P3 — Harm

冒充真人可能損害名譽、私隱及尊嚴。

P4 — Freedom

AI 字幕、修圖、藝術等未必欺騙，不應過度限制。

P5 — Risk-based Approach

高風險 deepfake → 強制標示;
低風險技術輔助 → 可不標示。

Conclusion

真正標準應是:

©PETER HSU 2026

Risk of deception，而不只是 presence of AI。

---

13. 常見失分陷阱

❌ AI 內容都是假的
→ **AI-generated ≠ false**

❌ 真人影片一定是真的
→ **Human-generated ≠ trustworthy**

❌ 只談科技進步
→ 本題核心是媒體倫理

❌ 只談言論自由
→ Freedom 必須配合 responsibility

❌ 觀眾自己小心便可以
→ 不能取消 creator responsibility

❌ 所有 AI 修改都當成 deepfake
→ 要分風險程度

考前必記

Accuracy ≠ Authenticity
AI Use ≠ Deception
AI-generated ≠ False
Human-generated ≠ True
Audience Responsibility ≠ No Creator Responsibility
Best Policy = Risk-based Labelling

一句總結

AI 時代的媒體倫理不是反對科技，而是確保創作者在擁有更強的製作能力時，也承擔更大的真實和透明責任；規管應針對欺騙風險，而不是把所有 AI 使用一刀切視為問題。

©PETER HSU 2026