

# TAICA\_生成式 AI：文字與圖像生成的原理與實務

## 第四週作業 03/11: 建立自己的 Benchmarks

臺師大 613K0019C AI 所碩一 劉思廷

### 一、作業設計

**Prompt:** 設計一個不存在於現實世界的社會制度，並描述其運作方式與優缺點

**比較模型種類與對話連結:**

ChatGPT: <https://chatgpt.com/share/67dfc030-8f44-8009-976e-6b9bf2ffb082>

Grok 3: [https://grok.com/share/bGVnYWN5\\_c883b8cc-f7c0-42d6-966f-dae7ea98f1ac](https://grok.com/share/bGVnYWN5_c883b8cc-f7c0-42d6-966f-dae7ea98f1ac)

Claude: <https://claude.ai/share/92f3e02c-e331-4546-8adb-b764d8bce9fc>

### 二、各模型回應

#### ChatGPT 提出的制度: 共感信用社會 (Empathic Credit Society)

**制度概述:** 在共感信用社會中，每位公民的社會地位與資源分配取決於共感指數 (Empathy Index)。此指數衡量個體對他人情感、需求的理解與回應能力，並根據此指數來決定其在社會中的權利、義務與福利。

**運作方式:**

##### 1. 共感指數評估系統

- 日常行為監測：透過智慧裝置與社群平台，持續監測個體的行為，例如：
  - 幫助他人（如志工服務、援助弱勢）
  - 情緒理解與回應（如安慰他人、處理衝突的方式）
  - 社群參與（如公共討論、共識決策）
- 情感分析技術：透過自然語言處理與面部表情識別，分析個體的情感反應與他人互動時的態度。
- 群體回饋機制：任何人都可對他人進行「共感回饋」，包括正向評價（例如：善行）或負向評價（如冷漠、傷害他人）。

2. 等級與權利: 根據個體的共感指數，社會將個體分為五個等級，每級享有不同的社會資源與權利。

- S 級（共感指數 90-100）
  - 免費享有高等教育與醫療
  - 參與政策制定的優先權
  - 可擔任社會要職（如法官、教師、醫師）

- A 級（共感指數 70-89）
  - 稅收減免與住房補貼
  - 可參與重要社會討論並提出政策建議
- B 級（共感指數 50-69）
  - 基本社會保障（醫療、教育、住房補助）
  - 可申請升級計劃，參與社會服務提升共感指數
- C 級（共感指數 30-49）
  - 基本生存權（最低生活保障）
  - 需參加強制共感培訓，通過測驗才能獲得更多權利
- D 級（共感指數 0-29）
  - 嚴格限制社會資源使用（如無法參與政策投票）
  - 需完成長期公共服務以提升共感指數

3. 共感修復計劃: 低共感指數者可透過「修復計劃」提升指數，方法包括：

- 參與社區服務
- 接受心理輔導
- 完成同理心教育與考核

優點:

- 1) 促進社會互助與和諧: 共感指數鼓勵人們主動幫助他人，減少社會冷漠，營造友善、合作的社會氛圍。
- 2) 獎勵利他行為: 以共感指數作為社會資源分配的核心標準，確保善行者得到相應回報，打破財富或出身的不公平結構。
- 3) 減少權力濫用: 重要社會職位（如政治家、法官）需具備高共感指數，避免冷血決策與腐敗行為。

缺點:

- 1) 隱私侵犯風險: 持續監測個體行為與情感反應可能侵犯個人隱私，造成過度監控的社會環境。
- 2) 共感操控與虛偽行為: 個體可能為了提升指數而刻意表現共感，導致虛偽行為與道德綁架，扭曲真正的情感價值。
- 3) 情感多樣性忽視: 共感的衡量標準可能無法完全捕捉個體間情感的複雜性，忽視內向或表達方式不同者的價值。
- 4) 社會階層固化: 低共感指數者可能陷入社會底層，若修復計劃無法有效運行，將導致社會流動性減弱、弱勢群體被邊緣化。

結論:

共感信用社會強調個體對他人情感與需求的理解作為社會地位的核心指標，試圖建立更具關懷與合作精神的社會。然而，該制度也面臨隱私、操控、與社會分化的挑戰，需謹慎平衡技術監控與人性自由，才能實現真正的社會公平與共融。

## Grok 提出的制度: 共識流動社會 (Consensus Fluid Society)

**制度概述:** 共識流動社會是一種基於動態共識和資源共享的社會結構。這個社會沒有固定的政府或階級，而是由所有成員通過即時數位平台共同決策。社會的核心是一個名為「流動核心」(Fluid Core) 的 AI 系統，負責收集意見、分析需求並協調資源分配。每個成員都擁有平等的發言權，但權重會根據個人貢獻和專業知識動態調整。

**運作方式:**

### 1. 即時共識決策

- 每個公民配備一個個人裝置，連接到「流動核心」。無論是日常事務（如資源分配）還是重大政策（如環境保護），所有決定都通過平台投票完成。
- 投票不是單純的多數決，而是採用「加權共識模型」。例如，一位長期研究氣候變化的科學家的意見在環境議題上會獲得更高權重，但權重會隨時間和議題變化而流動。
- 平台每隔 24 小時更新一次決策，確保社會能快速適應新情況。

### 2. 資源共享與動態分配

- 社會沒有貨幣，而是以「貢獻點」(Contribution Points) 作為衡量個人對社會價值的單位。貢獻點由「流動核心」根據工作、教育、創新等動態計算。
- 資源（如食物、住房、科技設備）根據需求和貢獻點分配，但基本生活保障無條件提供給所有人。
- 若某人長期不參與社會活動，系統會發出「參與邀請」，提供教育或工作機會，而不是懲罰。

### 3. 教育與角色流動

- 教育是終身制且免費，重點在於培養多領域能力。個人可以隨時轉換職業或角色，無需固定身份。
- 「流動核心」會根據社會需求推薦學習方向，但選擇權完全屬於個人。

### 4. 衝突解決

- 若出現爭議，系統會召集相關方進行公開討論，並由隨機選出的公民小組（權重調整後）裁決。

- 極端情況下，「流動核心」會模擬衝突的長期影響，供所有人參考。

#### 優點:

- 1) 高度靈活性: 社會能迅速適應危機（如自然災害或技術變革），因為決策和資源分配是即時且動態的。
- 2) 公平性與參與感: 每個人都能參與決策，避免權力集中於少數人手中。權重調整機制確保專業知識得到尊重，又不形成永久精英階層。
- 3) 資源效率: 無貨幣經濟減少浪費，AI 系統能精準匹配供需，避免過剩或短缺。
- 4) 個人自由: 教育和角色的流動性讓人們能追求興趣，而非被固定在單一身份中。

#### 缺點:

- 1) 技術依賴風險: 若「流動核心」系統故障或被操控，整個社會可能陷入混亂。
- 2) 決策疲勞: 公民需要頻繁參與投票和討論，可能導致部分人感到壓力或選擇退出參與。
- 3) 權重爭議: 雖然權重動態調整旨在公平，但可能有人質疑其公正性，認為系統偏袒某些群體。
- 4) 缺乏長期穩定性: 過於頻繁的決策更新可能導致政策缺乏連貫性，影響長期規劃（如大型基礎建設）。

#### 結論:

共識流動社會是一個高度民主、靈活但技術依賴的制度。它適合一個重視個人自由和集體協作的未來世界，但其成功取決於 AI 系統的可靠性和公民的積極參與。若技術與人性都能完美配合，它可能創造一個近乎理想的社會；反之，若信任崩潰，它也可能迅速瓦解。

### **Claude 提出的制度: 動態優先權治理系統 (Dynamic Prioritization Governance)**

**制度概述:** 動態優先權治理系統是一種以議題為中心、專業為導向的社會治理模式，其核心理念是根據個人在不同議題上的專業程度、參與度和貢獻度，動態分配決策權重。

#### 運作方式:

##### 1. 專業資格認證

- 社會中的每個公民可以在各種領域（教育、醫療、環保、經濟等）建立專業資格
- 專業資格基於以下因素獲得：
  - 正規教育和專業訓練

- 實踐經驗和工作成果
- 同行評估
- 公開可驗證的貢獻

## 2. 議題分類與決策分配

- 社會議題依性質分類（環保、教育、醫療等）
- 不同議題的決策權重根據以下公式分配：
  - 相關專業資格 (60%)
  - 直接利害關係 (25%)
  - 基本公民權 (15%)

## 3. 動態投票系統

- 公民在議題上的投票權重根據其專業資格和利害關係動態調整
- 專業者在其領域有較高權重，但從不達到可獨斷決策的程度
- 所有公民保留基本投票權，確保基本平等

## 4. 透明度與問責制

- 所有決策過程、專業資格評估和權重分配完全透明
- 區塊鏈技術記錄所有決策和權重變化，確保不可篡改
- 定期審查機制評估系統公平性和有效性

## 5. 知識與專業培養

- 系統鼓勵公民持續學習和專業成長
- 公共資源投入建立開放教育平台
- 跨領域知識整合獲得額外權重獎勵

### 優點:

- 1) 專業導向決策 - 確保決策由最了解相關領域的人主導，提高決策品質
- 2) 減少民粹主義風險 - 避免情緒化、非理性決策主導社會發展
- 3) 鼓勵公民學習成長 - 激勵人們持續學習、擴展知識領域以獲得更多參與權
- 4) 彈性且適應性強 - 權重隨社會需求和個人發展動態調整，適應變化
- 5) 平衡專業與公平 - 保留基本公民權，同時尊重專業知識
- 6) 促進跨領域合作 - 鼓勵不同專業背景的人合作，獲得整合性解決方案

### 缺點:

- 1) 專業評估的複雜性 - 客觀評估專業程度極具挑戰，可能引發爭議
- 2) 潛在精英主義風險 - 若實施不當，可能導致知識精英階層的形成
- 3) 資源差距擴大可能 - 教育資源不平等可能限制某些群體獲取專業資格的機會
- 4) 系統操縱風險 - 專業認證可能被有心人士操縱或偽造
- 5) 決策過程延緩 - 權重計算和專業評估可能導致決策過程緩慢
- 6) 社會分化潛在風險 - 可能造成「專家」與「非專家」之間的社會分化

#### 緩解措施

- 建立強大的防操縱機制和專業評估標準
- 投入大量資源確保教育機會平等
- 設計有效的跨專業溝通平台，促進不同領域間的理解
- 建立申訴和審查機制，防止精英壟斷
- 確保基本公民權重足以保護弱勢群體利益

我設計了一個稱為「動態優先權治理系統」的社會制度，這是一種以專業和議題為中心的治理模式，在其中每個公民的決策權重會根據其專業知識、直接利害關係和基本公民權動態調整。這個制度的核心優點包括能夠讓專業知識在決策中發揮更大作用、減少非理性決策的風險、鼓勵公民持續學習成長，以及在保持適應性的同時平衡專業與公平。然而，這種制度也面臨著一些挑戰，如專業評估的複雜性、可能出現的精英主義風險、教育資源不平等導致的機會差距，以及系統被操縱的潛在可能。我在設計中也納入了一些緩解措施，如建立防操縱機制、確保教育機會平等、設計跨專業溝通平台等，以減少這些潛在負面影響。

### 三、三種回答的相同、相異處比較 (使用 Grok 協助整理)

#### 相同處:

- 技術依賴:

三種制度都高度依賴技術系統來運作。例如，「共感信用社會」使用情感分析技術評估共感指數，「共識流動社會」依靠 AI 系統「流動核心」進行決策協調，而「動態優先權治理系統」則需評估專業資格和調整權重。
- 動態調整:

這三者都會根據個人的表現或條件（如共感能力、貢獻度或專業知識）調整權利、資源或決策權重，與傳統靜態結構不同。

#### 相異處:

- 核心衡量標準:

- 共感信用社會：以「共感指數」為核心，衡量個人對他人情感和需求的理解與回應能力。
- 共識流動社會：以個人貢獻度和專業知識為基礎，決定資源分配和決策影響力。
- 動態優先權治理系統：以專業資格和議題相關性為標準，動態分配決策權重。
- 決策方式:
  - 共感信用社會：基於共感指數的分級制度，權利與義務按等級分配。
  - 共識流動社會：透過即時共識投票，公民在數位平台參與，AI 根據貢獻與專業調整權重。
  - 動態優先權治理系統：決策權重根據個人專業程度和議題相關性動態分配，強調專業導向。
- 社會結構:
  - 共感信用社會：具有明確的等級制度，根據共感指數劃分社會階層。
  - 共識流動社會：無固定階級，強調流動性和平等參與。
  - 動態優先權治理系統：以專業為導向的治理結構，權力分佈因議題而異。
- 主要挑戰:
  - 共感信用社會：隱私侵犯（行為監測）、虛偽行為增加，以及忽視情感多樣性的問題。
  - 共識流動社會：過度依賴技術系統、決策疲勞（頻繁投票），以及權重分配爭議。
  - 動態優先權治理系統：專業資格評估複雜、潛在精英主義風險，以及決策過程可能過於緩慢。

#### 四、我對這三種回答的看法

##### **(ChatGPT): 共感信用社會 (Empathic Credit Society)**

這個制度以共感指數作為社會地位和資源分配的核心，透過日常行為監測、情感分析技術和群體回饋評估每個人的共感能力。這種試圖量化情感並促進互助的設計表面上看起來頗具吸引力，然而，深入分析後，我認為這個制度存在倫理與實踐上的多種問題。首先，這個制度要求持續監測行為與情感，這明顯侵犯個人隱私。人們的言行被無時無刻記錄並分析，令人感到壓迫。更嚴重的是，這種監控可能導致人們為了提升自己的共感指數，而刻意表演善良或共感，扭曲了真實情感的價值，侵蝕彼此間的信任。其次，共感標準難以客觀量化。情感表達因個人性格與文化差異而異，例如內向者可能不善於外顯情感，但並非缺乏共感。然而，系統可能誤判

他們為低共感者，剝奪其應有資源。這忽略了人性的多樣性，顯然將導致一些問題。再者，按共感指數分級並限制低指數者的資源，可能導致階級固化。低共感者若因資源不足無法改善生活，難以提升指數，形成惡性循環。長期下來，這不僅無法促進互助，反而加劇不平等，背離制度初衷。最後，技術局限使制度難以實現。情感分析無法精準區分真誠與表演，且易受數據偏誤影響。群體回饋也可能被個人恩怨或偏見左右，難以保證公正性。這些缺陷讓制度的實踐效果大打折扣。「共感信用社會」雖然理想化，但過於簡化情感的複雜性，忽略技術局限與倫理挑戰。從隱私侵犯、虛偽行為，到標準偏誤與階級固化風險，這些問題顯示它在現實中難以有效運作。

### **(Grok 3): 共識流動社會 (Consensus Fluid Society)**

這個制度透過 AI 系統「流動核心」實現動態共識和資源共享，強調去中心化與靈活性。我認同這個制度能夠促進民主參與也擁有高適應性，但也發現一些問題值得關注。首先，制度過度依賴 AI 技術。如果「流動核心」故障或被操控，技術穩定性若無法保證，整個系統的運作將岌岌可危。其次，頻繁的決策更新雖然能快速應對危機，但可能使政策缺乏連貫性。這對需要長期規劃的議題不利。短期的靈活調整或許有效，但長期目標可能因此受損，影響社會的穩定發展。制度中還有權重調整機制，旨在提升專業人士的影響力。然而，若評估過程不透明或不公平，可能引發爭議並削弱公民信任。公正性若無法保障，這個設計反而會成為衝突的來源。最後，制度假設公民能持續參與頻繁的決策，但現實中並非如此。工作壓力、資訊過載等因素可能導致人們退出參與，產生「決策疲勞」。這不僅降低決策品質，還可能使系統無法反映多數人的意願，與去中心化的初衷背道而馳。

### **(Claude): 動態優先權治理系統 (Dynamic Prioritization Governance)**

這個制度以專業知識為基礎，根據不同議題動態分配決策權，試圖在專業性與公平性之間找到平衡。我認同它的設計理念，但在實踐中可能存在一些問題。首先，評估專業資格的過程過於複雜且難以保持客觀。比如在醫療領域，醫生的能力標準可能因人而異，導致爭議。這類主觀性可能造成不公平，讓人們對制度失去信任。其次，教育資源的不平等可能限制某些群體參與決策。低收入家庭的孩子往往難以獲得專業資格，從而被排除在外。這不僅加深社會差距，還可能讓決策權集中在少數人手中。另外，過於強調專業知識可能導致精英主義。當決策權向知識精英傾斜時，普通人的聲音容易被忽視，引發不滿，損害公平性。制度提出了一些防操縱和促進教育平等的措施，但實際操作起來很困難。監管機制可能因為權力集中而失去效果，而教育資源的公平分配也受限於現實條件，短期內難以實現。我認為這個制度在專業性與公平性之間尋求平衡的目標有價值，但其複雜性和潛在的不公問



題與我重視的邏輯與平等相違背。若要改進，應該提高資格評估的透明度，並致力於教育資源的公平分配，以減少不公和社會分化的風險。

## 結論

「共感信用社會」、「共識流動社會」和「動態優先權治理系統」三種制度都試圖透過技術實現更公平、專業的社會秩序，但各有獨特魅力與挑戰。「共感信用社會」以共感指數促進互助，卻因隱私侵犯與階級固化風險顯得過於理想化；「動態優先權治理系統」強調專業決策，雖具邏輯性，但專業評估的複雜性與精英主義隱憂使其難以完美平衡公平與效率。相較之下，我最喜歡 Grok 3 提出的「共識流動社會」，因為它透過 AI 系統實現動態共識與資源共享，既靈活又能實現民主。它讓每個人都能參與決策，並以貢獻與專業調整權重，這在我看來能提升決策品質，尤其在專業議題上。最吸引我的是終身免費教育，這對熱愛學習的我來說非常具有吸引力，在激發個人潛能的同時也推動社會進步。當然，它對技術系統的依賴讓我有些擔心，若系統故障或被操控，可能帶來混亂；頻繁的決策也可能影響長期穩定性。因此，我認為若能將 AI 定位為輔助工具，而非完全依賴，並確保公民積極參與，這個制度會是相較之下最理想的選擇。然而，透過這次作業的實測，我意外發現大家現在最常用的 ChatGPT 提出的「共感信用社會」是我認為最不理想的設計。雖然這個模型相對強調人與人之間情感互動的重要性，顯示出對人性關懷的重視，但它在制度設計上卻選擇犧牲個人隱私來實現目標，這不僅顯得自相矛盾，也讓人覺得不太實際。透過持續監測行為與情感來評估共感指數的做法，可能會讓人們感到壓迫，甚至誘發虛偽行為，這與其追求互助與情感連結的初衷背道而馳。因此，我認為這個設計在理念與執行之間存在明顯落差，實用性令人質疑。