|  |
| --- |
| 中文專題名稱：模擬銀髮安居 |
| 英文專題名稱：Simulated Aging-Friendly Living |

一、前言

隨著全球人口老化趨勢加劇，如何為高齡者提供適當的居住環境與照護服務，已成為現代社會的重要課題。

高齡安居模擬資料（Aging-Friendly Living Simulation Data）的原始資料具有高度應用價值，能夠提供關鍵資訊，支援個人、家庭、政府及社福機構應對人口老化問題的決策。例如，家屬可以在高齡者尚未出現嚴重健康問題前，預先選擇適當的居住設施或進行居家改造，確保高齡者的生活舒適與安全；亦或是，透過分析資料，政府能更精準地預測高齡者在不同年齡階段的照護需求，從而針對性地提供社會福利和醫療資源分配，進一步改善長期照護制度。

然而，這類資料的公開也帶來隱私保護的挑戰。若資料過於詳盡，可能洩露高齡者的個人資訊，如健康狀況、經濟條件或居住地址等敏感內容，進而使高齡者在尋求照護或資源時感到不安和疑慮。這樣的隱私風險不僅可能影響高齡者的生活質量，還可能阻礙他們獲得所需的社會協助與資源。因此，如何在保障隱私的前提下，有效利用這些資料，成為了我們要去解決的重要議題。

二、創意描述

在這個專案中，我們探索了如何在保護隱私的同時，有效利用銀髮族生活資料來解決社會日益加劇的人口老齡化問題。我們所設計的隱私強化技術，不僅僅是單純的資料加密，更是一次資料保護與分析價值的雙重提升。透過將相似屬性資料進行巧妙的合併，例如將年齡、殘疾等級合併成一個綜合指標，我們不僅降低了資料的辨識性，還保留了重要的統計意義。

我們運用了創新的資料交換技術，將相鄰的資料隨機交換，讓每個個體變得更加模糊，卻不改變整體資料的分佈特徵，這種技術類似於在不影響整體畫作的前提下，隨機打亂畫布上的筆觸，讓每一筆看起來不同，但整幅畫的意境仍能被保留。再加上對部分資料添加隨機噪聲，進一步增加資料的不可預測性，從而確難以還原出個體的具體資訊。

為了評估隱私保護技術，我們的設計了測試指標，從欄位相似度到指認攻擊指標針對 action 欄位做長條圖分析。結果顯示，隱私保護處理後的資料分佈與原始資料幾乎沒有差異，證明我們的方法不僅能有效保護隱私，還能保留資料的保真度。

這項技術更具實際應用價值，無論是政府部門用於制定更加精準的社會政策，還是健康機構針對老年群體設計定制化的照護方案，亦或是智慧城市規劃中需要更適合銀髮族的居住環境，這項技術能提供強有力的資料支援。

三、共享資料目的簡介

1. 下游任務：銀髮安居需求指數差異分析

* 應用場景
  + 社會政策制定

政府或社會組織可以依據加密後的資料制定更具針對性的政策，以改善銀髮族的居住環境和生活品質。

* + 學術研究

研究人員可以利用加密資料進行社會學或人口學研究，分析銀髮族生活水準的變化而不侵犯個人隱私。

* + 健康服務規劃

健康機構可依據分析結果設計適合銀髮族的健康照護計劃，提升服務的針對性和有效性。

* + 社區服務優化

社區組織可以依據資料來優化長照服務，如交通、餐飲或社交活動，提升銀髮族生活的便利性。

* + 智能城市建設

在智慧城市的規劃中，分析銀髮族居住環境的資料可以幫助設計更友善的居住空間，改善公共設施。

1. 使用對象

* 政府機關——衛生福利部
  + 【模擬銀髮安居】提供銀髮安居需求指數，可給予政策規劃單位參考，除了能夠安全利用的資源，更將有限資源做最有效配置推動更高效的政策制定，提升銀髮族照顧福祉。
* 社會大眾
  + 【模擬銀髮安居】可向社會大眾開放，提供自我評估是否向政府申請相關福利與照護補助。

四、應用技術簡介

1. 隱私強化演算法

* 資料前處理

將屬性相似的資料合併為大項以概化資料，將age、disability\_lv合併為action；將family\_type、child\_cnt、is\_living\_same\_county合併為Nurse；將low\_type\_cd、having\_house\_type合併為gold；將build\_age、is\_apartment、合併為entity；將bus、store、hospital合併為liberty；將lique設為security。

* 資料交換

透過資料交換的方式增強資料的隱私保護力，可以獲得的優點如下：

* + 資料匿名化

透過隨機交換相鄰的資料，無法直接識別各行所代表的具體個體，使得個體資料的被識別難度增加。

* + 保持統計特性

雖然資料進行了交換，但整體的分佈特性（例如均值、標準差等）基本保持不變，這樣可以在某些分析中仍然獲得有效的資訊。

* + 降低重識別風險

對於敏感資訊來說，隨機的資料交換可以減少資料被用來重新識別個人的風險，特別是在合併多個資料集的情況下。

* 加入噪音

為了進一步增強隱私保護，我們選擇對三類多樣性較高的資料action、gold和entity採用加入雜訊的技術。這些資料的多樣性使得在其加入隨機擾動後，混淆性顯著提升，而透過將雜訊引入原始資料，使得每個資料點都變得不那麼可預測，從而降低了攻擊者根據統計資料回推原始值的可能性。

1. 隱私評測

* 在隱私強化後，我們會自動產生隱私保護力與資料保真度自我評測報告，方便監管單位以及使用者了解，使用我們的服務是合乎個資保護的。
* 隱私保護力

針對隱私保護力我們提出以下三個指標：

* + 原始資料可識別率
    - 原始資料可識別率衡量的是在未經處理的資料中，個體被識別的可能性。這一指標的降低意味著資料在未經授權的情況下，無法輕易被還原或識別。
    - 我們將欄位合併概化，因此與原始資料欄位不一樣，達到無法識別特定人的資料。
  + 指認性攻擊指標
    - 指認性攻擊指標評估的是對資料進行識別的攻擊方式的有效性，攻擊者在面對經過處理的資料時，能否成功識別出特定個體。
    - 在特定組合下，某些個體會容易被辨識出並被攻擊。

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| security | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| liberty | 0 | 0.017318 | 0.021028 | 0.038346 | 0.042508 | 0.059826 | 0.063536 | 0.080854 |
| is\_use\_long\_term\_care=0 | 291783 | 12613 | 3566 | 5554 | 14697 | 32935 | 1024 | 19235 |
| is\_use\_long\_term\_care=1 | 9632 | 613 | 167 | 282 | 552 | 1838 | 42 | 1058 |
| security | 0.013669 | 0.013669 | 0.013669 | 0.013669 | 0.013669 | 0.013669 | 0.013669 |  |
| liberty | 0 | 0.017318 | 0.021028 | 0.038346 | 0.042508 | 0.059826 | 0.080854 |  |
| is\_use\_long\_term\_care=0 | 1454 | 185 | 33 | 64 | 139 | 199 | 165 |  |
| is\_use\_long\_term\_care=1 | 53 | 17 | 2 | 0 | 4 | 8 | 16 |  |

* 資料保真度

資料保真度我們提出以下指標：

* + 欄位相似度

我們針對action這個欄位做分布圖，可知資料做隱私強化之後，此欄位的分布不變。

|  |
| --- |
|  |
| 隱私處理前後對比圖 |

五、結語

在這項專案中，我們專注於銀髮族居住環境的資料，透過加密和隱私保護來評估他們的生活水準。首先使用自訂的資料交換函數，有效地對資料進行交換，增強了隱私保護的同時，保留了統計特性，減少了個體識別的風險；此外，針對三項特別資料，添加了雜訊，使資料更具多樣性和不可預測性。這些處理不僅提升了資料的安全性，也確保了分析的有效性，最終的結果顯示，經過加密和雜訊處理後的預測結果與未處理資料差異不大，表明我們的方法在保護隱私的同時，依然能夠有效支持生活水準的評估。這項專案展現了如何在保護個人隱私的背景下，利用資料為社會問題提供見解，更為改善銀髮族的生活品質提供資料支撐。