

新住民相關親權酌定裁判書的文字探勘： 對「平等」問題的法實證研究嘗試*

邵軒磊**、黃詩淳***

<摘要>

本研究關心新住民與我國籍配偶離婚後，進入法院爭取未成年子女親權時，是否獲得平等對待之問題。本文使用資訊科學技術，驗證「法院是否會因父母的國籍而在裁判書中使用不同的詞彙」，分析 116 件親權酌定事件之裁判書。其中有 31 件之一方當事人為外籍、外裔人士（簡稱外籍裁判），85 件為雙方當事人皆為我國籍（簡稱雙臺籍裁判）。首先，將裁判書中的詞彙在分詞後，轉化為 TF-IDF 的文本詞彙矩陣，再透過主成分分析投影在二維平面上，將每則裁判書的詞彙特徵（所在位置）可視化。雙臺籍裁判與外籍裁判並未分為不同的兩群，而是位置重合，意味著法院在兩種當事人的裁判並未使用不同的詞彙，因此無法直接主張司法體系對新住民不平等。其次，再使用機器學習中的梯度提升法，發現外籍裁判中，「是否出境」以及「居住狀況」相關詞彙特徵，較雙臺籍裁判為高，表示法官在外籍裁判中較

* 感謝二位審查人對本文提供諸多寶貴意見，使作者受益甚多。感謝法律扶助基金會提供相關資料，以及工作人員的詳細說明。本研究部分成果完成於筆者與臺大法律資料分析研究室成員一起參與 D4SG (Data for Social Good) 資料英雄計畫 2019 冬季班。本研究為科技部「從知識系譜到知識地圖：大數據與機器學習下的『中國研究』(MOST-107-2410-H-003 -058 -MY3)」與教育部 (NTU-109L900203) 與科技部 (MOST 109-3017-F-002-003) 高教深耕計畫之研究成果。

** 國立臺灣師範大學東亞學系副教授。E-mail: hlshao@ntnu.edu.tw。

*** 國立臺灣大學法律學院教授、國立臺灣大學計量理論與應用研究中心兼任特約研究員。E-mail: schhuang@ntu.edu.tw。

• 投稿日：07/02/2020；接受刊登日：09/25/2020。

• 責任校對：何思奕、陳莉蕻、簡凱葳。

• DOI:10.6199/NTULJ.202011/SP_49.0001

關心程序參與以及居留問題。藉由文字探勘與機器學習來分析裁判書文本，除了檢驗平等問題外，還可將抽象的法律概念（子女最佳利益）具象化為裁判書中提及之事實（出境、租屋），供實務參考，結合了資訊與法律兩個領域，呼應本刊「跨域、科技與法律」的主題。

關鍵詞：法律資料分析、自然語言處理、詞頻—倒數文件頻率法、平等、新住民、法實證研究、主成分分析

◆目次◆

壹、研究背景

貳、文獻回顧

參、研究設計

一、資料來源

二、研究方法

三、研究步驟

肆、研究發現與討論

一、外籍裁判與雙臺籍裁判之詞彙並無不同

二、外籍裁判詞彙之特徵

伍、結論

一、研究結論

二、延伸討論與研究方法之意義