

中文新聞語料庫

本院覽號

05T-1070910

公告日期

智財權狀態

know-how

摘要

中文新聞語料庫，為一包含約一千四百萬字的語料庫，主要資料來源擷取自七十九年至八十年間聯合報、中國時報、自由時報及天下雜誌之文章。本語料庫得蔣經國國際學術交流基金會補助，由中央研究院資訊所中文詞知識庫小組執行及研究。

技術優勢

本新聞語料庫內容龐大，涵蓋完整一年的新聞，且來自不同媒體來源，計有聯合報、中國時報、自由時報及天下雜誌，提供中文語言處理技術所需的豐富材料。

應用範圍

資訊檢索 建立辭典 語言分析 語言理解 訊息抽取 媒體比較

創作人

馬偉雲、陳克健



中央研究院
ACADEMIA SINICA