



(19)中華民國智慧財產局

(12)新型說明書公告本

(11)證書號數：TW M613338 U

(45)公告日：中華民國 110 (2021) 年 06 月 11 日

(21)申請案號：110203088

(22)申請日：中華民國 110 (2021) 年 03 月 23 日

(51)Int. Cl. : **G06Q50/00 (2012.01)**

(71)申請人：中國信託商業銀行股份有限公司(中華民國) CTBC BANK CO., LTD. (TW)

臺北市南港區經貿二路 166、168、170、186、188 號

(72)新型創作人：林志豪 (TW)；邱國豪 (TW)；江侑倫 (TW)；曾文忻 (TW)；陳岳榆 (TW)；符妤琪 (TW)；劉家騫 (TW)；邱順源 (TW)；宋政隆 (TW)；王俊權 (TW)

(74)代理人：高玉駿；楊祺雄

(NOTE)備註：相同的創作已於同日申請發明專利(Another patent application for invention in respect of the same creation has been filed on the same date)

申請專利範圍項數：7 項 圖式數：1 共 15 頁

(54)名稱

醫囑資訊擷取系統

(57)摘要

一種醫囑資訊擷取系統包含一儲存模組、一指代消解模組、一文字替換模組、一實體抽取模組、及一處理模組。該指代消解模組包括一第一機器閱讀理解模型及一第一輸入問句。該實體抽取模組包括一第二機器閱讀理解模型及多個第二輸入問句。該處理模組電連接該儲存模組、該指代消解模組、該文字替換模組、及該實體抽取模組，並將一醫囑文件輸入至該指代消解模組，以獲得一同日日期及一同日位置，再將該醫囑文件、該同日日期、及該同日位置輸入至該文字替換模組，以獲得一取代文件，再將該取代文件輸入至該實體抽取模組，以獲得多個醫囑資訊。

指定代表圖：

符號簡單說明：

1:醫囑資訊擷取系統

11:處理模組

12:儲存模組

13:指代消解模組

14:實體抽取模組

15:文字替換模組

2:資料庫

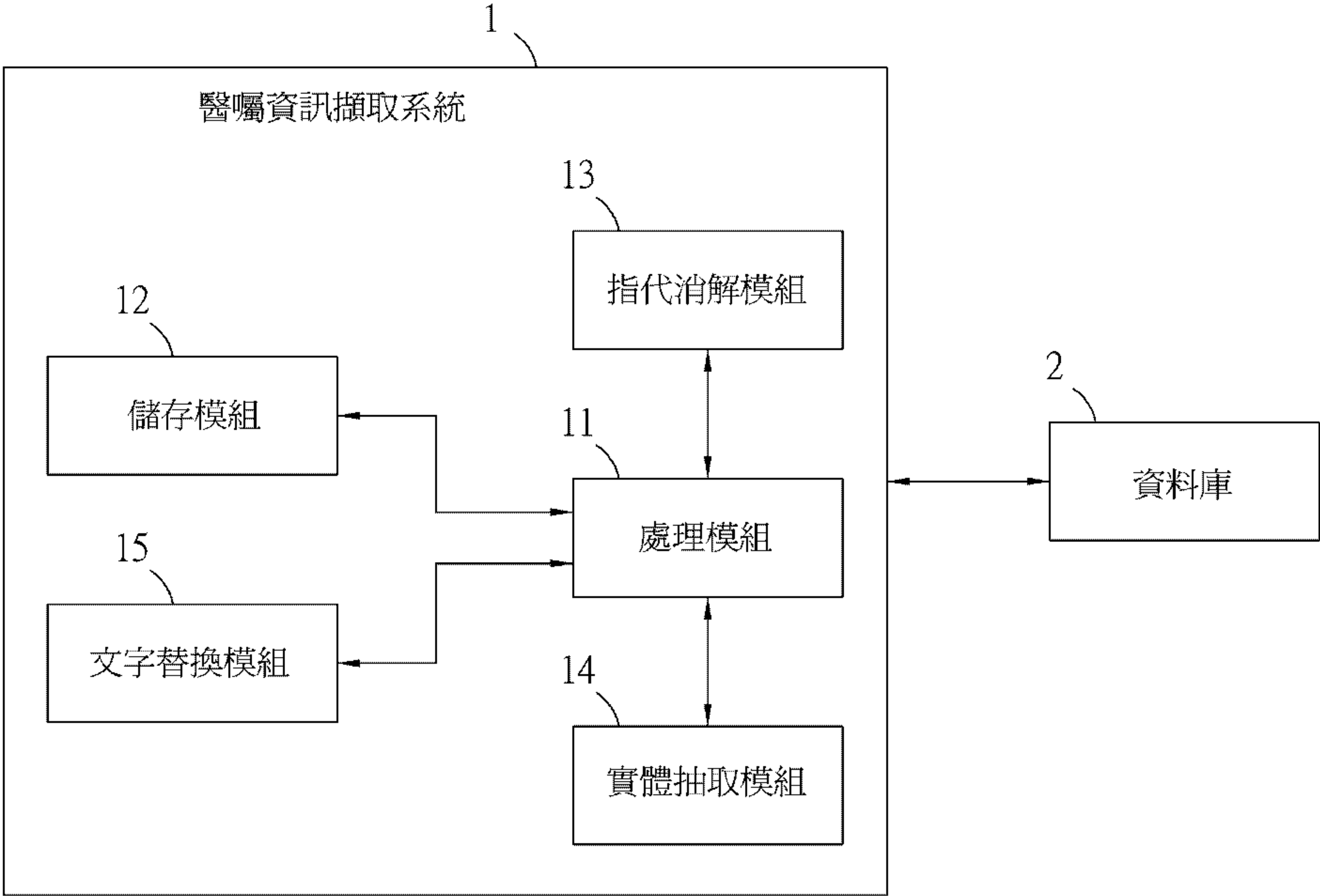


圖1



公告本

【新型摘要】

M613338

【中文新型名稱】 醫囑資訊擷取系統

【中文】

一種醫囑資訊擷取系統包含一儲存模組、一指代消解模組、一文字替換模組、一實體抽取模組、及一處理模組。該指代消解模組包括一第一機器閱讀理解模型及一第一輸入問句。該實體抽取模組包括一第二機器閱讀理解模型及多個第二輸入問句。該處理模組電連接該儲存模組、該指代消解模組、該文字替換模組、及該實體抽取模組，並將一醫囑文件輸入至該指代消解模組，以獲得一同日日期及一同日位置，再將該醫囑文件、該同日日期、及該同日位置輸入至該文字替換模組，以獲得一取代文件，再將該取代文件輸入至該實體抽取模組，以獲得多個醫囑資訊。

【指定代表圖】：圖（1）。

【代表圖之符號簡單說明】

1:醫囑資訊擷取系統

11:處理模組

12:儲存模組

13:指代消解模組

14:實體抽取模組

15:文字替換模組

2:資料庫

【新型說明書】

【中文新型名稱】 醫囑資訊擷取系統

【技術領域】

【0001】本新型是有關於一種資訊的擷取系統，特別是指一種用於擷取醫囑文件之資訊的醫囑資訊擷取系統。

【先前技術】

【0002】傳統的資訊檢索系統僅能針對字面上的字義進行提取，同時，對於每個詞組也只能進行一次性的提取。這會造成所提取的資訊在相關於日期時間的部分會具有以下的問題，舉例來說，在醫囑文件「病人於民國109年4月22日至本院急診就診，同日住院接受藥物治療，4月30日出院」中，傳統的演算法所提取的結果包含：「急診日期：民國109年4月22日」、「入院日期：同日」、及「出院日期：4月30日」。也就是說，在入院日期的資訊上，只能提取到「同日」，而非「民國109年4月22日」。此外，因為「同日」已經被歸類為「入院日期」，所以無法再把「治療處置日期」的資訊(即民國109年4月22日)提取出來。因此，是否存有其他更有效辨別日期時間的資訊擷取系統便成為一個待解決的問題。

【新型內容】

【0003】因此，本新型的目的，即在提供一種有效判定文件內容中關於日期時間的醫囑資訊擷取系統。

【0004】於是，本新型提供一種醫囑資訊擷取系統，包含一儲存模組、一指代消解模組、一文字替換模組、一實體抽取模組、及一處理模組。該儲存模組儲存一醫囑文件。

【0005】該指代消解模組包括一第一機器閱讀理解模型及一第一輸入問句，並用於將一第一輸入文件及該第一輸入問句輸入該第一機器閱讀理解模型，以獲得該第一輸入文件中的「同日」所表示的一同日日期及所處的一同日位置。該第一輸入問句相關於詢問「同日」所表示的日期。

【0006】該文字替換模組用於接收一第二輸入文件、一取代字串、及對應的一取代位置，以將該第二輸入文件中的該取代位置取代為該取代字串，而獲得一取代文件。

【0007】該實體抽取模組包括一第二機器閱讀理解模型及多個第二輸入問句，並用於將一第三輸入文件及該等第二輸入問句之其中任一者輸入該第二機器閱讀理解模型，以獲得該第三輸入文件中對應該等第二輸入問句之解答的多個醫囑資訊。該等第二輸入問句相關於詢問醫療行為或醫療時間。

【0008】該處理模組電連接該儲存模組、該指代消解模組、該文字替換模組、及該實體抽取模組，並將該醫囑文件輸入至該指代消

解模組以作為該第一輸入文件，進而獲得該同日日期及該同日位置，再將該醫囑文件、該同日日期、及該同日位置輸入至該文字替換模組，以分別作為該第二輸入文件、該取代字串、及該取代位置，進而獲得該取代文件，再將該取代文件輸入至該實體抽取模組以作為該第三輸入文件，進而獲得該等醫囑資訊。

【0009】在一些實施態樣中，其中，該第一機器閱讀理解模型及該第二機器閱讀理解模型都是屬於一種深度閱讀理解神經網路。

【0010】在一些實施態樣中，其中，該第一機器閱讀理解模型是一種長短期記憶(Long Short-Term Memory，LSTM)模型或一種雙向編碼文字向量轉換器(Bidirectional Encoder Representations from Transformers，BERT)。該第二機器閱讀理解模型是一種長短期記憶(LSTM)模型或一種雙向編碼文字向量轉換器(BERT)。

【0011】在另一些實施態樣中，其中，該同日位置包含一第一座標及一第二座標。該第一座標及該第二座標分別是「同」及「日」在該第一輸入文件中的第幾個字。該取代位置包含一第三座標及一第四座標。該文字替換模組根據該取代位置，刪除該第二輸入文件中該第四座標的字，並在該第三座標插入要取代的該取代字串，且將該第四座標之後的文字根據該取代字串的字數，作對應的座標位置調整。

【0012】在另一些實施態樣中，其中，該等第二輸入問句包含詢問一入院日期、一出院日期、一手術日期、一治療處置日期、一急診日期、一門診日期、一入加護病房日期、一出加護病房日期、一入燒燙傷病房日期、一出燒燙傷病房日期、一入負壓病房日期、一出負壓病房日期、一放療日期、一化療日期、一手術名稱、一治療名稱、及一藥物名稱之其中部分或全部者。

【0013】在另一些實施態樣中，該醫囑資訊擷取系統適用於一資料庫。其中，該資料庫預先儲存該醫囑文件，使得該處理模組將該資料庫的該醫囑文件儲存至該儲存模組。

【0014】在一些實施態樣中，其中，該資料庫包括一字元辨識模組，該資料庫先儲存有一手寫文件電子檔，使得該字元辨識模組對該手寫文件電子檔作字元辨識，以產生對應的文字內容的該醫囑文件。

【0015】本新型的功效在於：藉由該指代消解模組先獲得該醫囑文件中的該同日日期及該同日位置，再藉由該文字替換模組根據該醫囑文件、該同日日期、及該同日位置，獲得該取代文件。最後藉由該實體抽取模組根據該取代文件獲得對應該等第二輸入問句之解答的該等醫囑資訊，以解決習知技術所存在的問題。

【圖式簡單說明】

【0016】本新型的其他的特徵及功效，將於參照圖式的實施方式中清楚地呈現，其中：

圖 1 是一方塊圖，說明本新型醫囑資訊擷取系統的一實施例。

【實施方式】

【0017】在本新型被詳細描述之前，應當注意在以下的說明內容中，類似的元件是以相同的編號來表示。

【0018】參閱圖1，本新型醫囑資訊擷取系統1之一實施例，適用於一資料庫2，並包含一儲存模組12、一指代消解模組13、一文字替換模組15、一實體抽取模組14、及一處理模組11。該處理模組11電連接該儲存模組12、該指代消解模組13、該文字替換模組15、及該實體抽取模組14。該儲存模組12例如是一記憶體或一硬碟機，該處理模組11例如是一處理器，該指代消解模組13、該文字替換模組15、及該實體抽取模組14例如都是一種特殊應用積體電路(Application Specific Integrated Circuit，ASIC)。

【0019】在本實施例中，該資料庫2包括一字元辨識模組，該資料庫2儲存一手寫文件電子檔。舉例來說，該手寫文件電子檔是藉由掃描器掃描或相機拍攝醫療人員手寫的醫囑說明文件而產生。該字元辨識模組對該手寫文件電子檔作字元辨識(即文字辨識)，以產生對應的文字內容的一醫囑文件。該醫囑文件例如是一種文字檔案

(text file)的格式。

【0020】該處理模組11讀取該資料庫2所儲存的該醫囑文件，並將該醫囑文件儲存至該儲存模組12。該指代消解模組13包括一第一機器閱讀理解模型及一第一輸入問句。該第一機器閱讀理解模型屬於一種深度閱讀理解神經網路，並例如是一種長短期記憶(Long Short-Term Memory, LSTM)模型或一種雙向編碼文字向量轉換器(Bidirectional Encoder Representations from Transformers, BERT)，但不以此為限。該第一輸入問句相關於詢問「同日」所表示的日期，舉例來說，該第一輸入問句是「請幫我找出同日的代表的日期」，或「請幫我找出同日的同義詞彙」等等。

【0021】該指代消解模組13用於將一第一輸入文件及該第一輸入問句輸入該第一機器閱讀理解模型，以獲得該第一輸入文件中的「同日」所表示的一同日日期及所處的一同日位置。該同日位置包含一第一座標及一第二座標，該第一座標及該第二座標分別是「同」及「日」在該第一輸入文件中的第幾個字。該處理模組11將該醫囑文件輸入至該指代消解模組13以作為該第一輸入文件，進而獲得該同日日期及該同日位置。舉例來說，該醫囑文件的內容是「病人於民國109年4月22日至本院急診就診同日住院接受藥物治療4月30日出院」，定義第一個字(即「病」)的座標位置是0，則該同日位置=[該第一座標, 該第二座標]=[18, 19]，該同日日期是「民國109

年4月22日」。另外要特別說明的是：該醫囑文件的內容也可以是有包含標點符號。

【0022】該文字替換模組15用於接收一第二輸入文件、一取代字串、及對應的一取代位置，以將該第二輸入文件中的該取代位置取代為該取代字串，而獲得一取代文件。該取代位置包含一第三座標及一第四座標，該文字替換模組15根據該取代位置，刪除該第二輸入文件中該第四座標的字，並在該第三座標插入要取代的該取代字串，且將該第四座標之後的文字根據該取代字串的字數，作對應的座標位置調整。該處理模組11將該醫囑文件、該同日日期、及該同日位置輸入至該文字替換模組15，以分別作為該第二輸入文件、該取代字串、及該取代位置，進而獲得該取代文件。

【0023】承續前例，該取代文件的內容是「病人於民國109年4月22日至本院急診就診民國109年4月22日住院接受藥物治療4月30日出院」。也就是說，該醫囑文件的座標位置19的「日」被刪除，且座標位置18的「同」被「民國109年4月22日」共8個字取代，且「住院接受藥物治療4月30日出院」根據「民國109年4月22日」的字數向後調整座標位置。

【0024】該實體抽取模組14包括一第二機器閱讀理解模型及多個第二輸入問句。該第二機器閱讀理解模型也屬於一種深度閱讀理解神經網路，並例如是一種長短期記憶(LSTM)模型或一種雙向編碼

文字向量轉換器(BERT)，但不以此為限。在本實施例中，該等第二輸入問句包含詢問一入院日期、一出院日期、一手術日期、一治療處置日期、一急診日期、一門診日期、一入加護病房日期、一出加護病房日期、一入燒燙傷病房日期、一出燒燙傷病房日期、一入負壓病房日期、一出負壓病房日期、一放療日期、一化療日期、一手術名稱、一治療名稱、及一藥物名稱。而在其他的實施例中，待等第二輸入問句也可以包含較少數量的問句，或其他相關於詢問醫療行為或醫療時間的問句。

【0025】該實體抽取模組14用於將一第三輸入文件及該等第二輸入問句之其中任一者輪流輸入該第二機器閱讀理解模型，以獲得該第三輸入文件中對應該等第二輸入問句之解答的多個醫囑資訊。該處理模組11將該取代文件輸入至該實體抽取模組14以作為該第三輸入文件，進而獲得該等醫囑資訊。承續前例，該等醫囑資訊分別是「急診日期：民國109年4月22日」、「入院日期：民國109年4月22日」、「治療處置日期：民國109年4月22日」、「出院日期：4月30日」。

【0026】另外要補充說明的是：該實體抽取模組14將每次接收到的該第三輸入文件會輪流與每一該第二輸入問句輸入至該第二機器閱讀理解模型，但並一定會獲得對應每一該第二輸入問句的該醫囑資訊，例如在本實施例中，該等第二輸入問句的數量是17個，而

該等醫囑資訊的數量是4個。此外，在其他的實施例中，該指代消解模組13、該文字替換模組15、及該實體抽取模組14也可以採用軟體的方式來實現，例如以程式碼執行在記憶體中執行對應的功能。

【0027】綜上所述，藉由該處理模組11先將該醫囑文件輸入至該指代消解模組13，以獲得該醫囑文件中的該同日日期及該同日位置，再藉由該文字替換模組15根據該醫囑文件、該同日日期、及該同日位置，獲得該取代文件。接著，該處理模組11再藉由該實體抽取模組14根據該取代文件獲得對應該等第二輸入問句之解答的該等醫囑資訊，故確實能達成本新型的目的。

【0028】惟以上所述者，僅為本新型的實施例而已，當不能以此限定本新型實施的範圍，凡是依本新型申請專利範圍及專利說明書內容所作的簡單的等效變化與修飾，皆仍屬本新型專利涵蓋的範圍內。

【符號說明】

【0029】

1:醫囑資訊擷取系統

11:處理模組

12:儲存模組

13:指代消解模組

14:實體抽取模組

15:文字替換模組

2:資料庫

【新型申請專利範圍】

【請求項1】一種醫囑資訊擷取系統，包含：

一儲存模組，儲存一醫囑文件；

一指代消解模組，包括一第一機器閱讀理解模型及一第一輸入問句，並用於將一第一輸入文件及該第一輸入問句輸入該第一機器閱讀理解模型，以獲得該第一輸入文件中的「同日」所表示的一同日日期及所處的一同日位置，該第一輸入問句相關於詢問「同日」所表示的日期；

一文字替換模組，用於接收一第二輸入文件、一取代字串、及對應的一取代位置，以將該第二輸入文件中的該取代位置取代為該取代字串，而獲得一取代文件；

一實體抽取模組，包括一第二機器閱讀理解模型及多個第二輸入問句，並用於將一第三輸入文件及該等第二輸入問句之其中任一者輸入該第二機器閱讀理解模型，以獲得該第三輸入文件中對應該等第二輸入問句之解答的多個醫囑資訊，該等第二輸入問句相關於詢問醫療行為或醫療時間；及

一處理模組，電連接該儲存模組、該指代消解模組、該文字替換模組、及該實體抽取模組，並將該醫囑文件輸入至該指代消解模組以作為該第一輸入文件，進而獲得該同日日期及該同日位置，再將該醫囑文件、該同日日期、及該同日位置輸入至該文字替換模組，以分別作為該第二輸入文件、該取代字串、及該取代位置，進而獲得該取代文件，再將該取代文件輸入至該實體抽取模組以作為該第

第1頁，共3頁(新型申請專利範圍)

三輸入文件，進而獲得該等醫囑資訊。

【請求項2】如請求項1所述的醫囑資訊擷取系統，其中，該第一機器閱讀理解模型及該第二機器閱讀理解模型都是屬於一種深度閱讀理解神經網路。

【請求項3】如請求項2所述的醫囑資訊擷取系統，其中，該第一機器閱讀理解模型是一種長短期記憶(Long Short-Term Memory，LSTM)模型或一種雙向編碼文字向量轉換器(Bidirectional Encoder Representations from Transformers，BERT)，該第二機器閱讀理解模型是一種長短期記憶(LSTM)模型或一種雙向編碼文字向量轉換器(BERT)。

【請求項4】如請求項2所述的醫囑資訊擷取系統，其中，該同日位置包含一第一座標及一第二座標，該第一座標及該第二座標分別是「同」及「日」在該第一輸入文件中的第幾個字，該取代位置包含一第三座標及一第四座標，該文字替換模組根據該取代位置，刪除該第二輸入文件中該第四座標的字，並在該第三座標插入要取代的該取代字串，且將該第四座標之後的文字根據該取代字串的字數，作對應的座標位置調整。

【請求項5】如請求項2所述的醫囑資訊擷取系統，其中，該等第二輸入問句包含詢問一入院日期、一出院日期、一手術日期、一治療處置日期、一急診日期、一門診日期、一入加護病房日期、一出加護病房日期、一入燒燙傷病房日期、一出燒燙傷病房日期、一入負壓病房日期、一出負壓病房日

第2頁，共3頁(新型申請專利範圍)

期、一放療日期、一化療日期、一手術名稱、一治療名稱、及一藥物名稱之其中部分或全部者。

【請求項6】如請求項1所述的醫囑資訊擷取系統，適用於一資料庫，其中，該資料庫預先儲存該醫囑文件，使得該處理模組將該資料庫的該醫囑文件儲存至該儲存模組。

【請求項7】如請求項6所述的醫囑資訊擷取系統，其中，該資料庫包括一字元辨識模組，該資料庫先儲存有一手寫文件電子檔，使得該字元辨識模組對該手寫文件電子檔作字元辨識，以產生對應的文字內容的該醫囑文件。

【新型圖式】

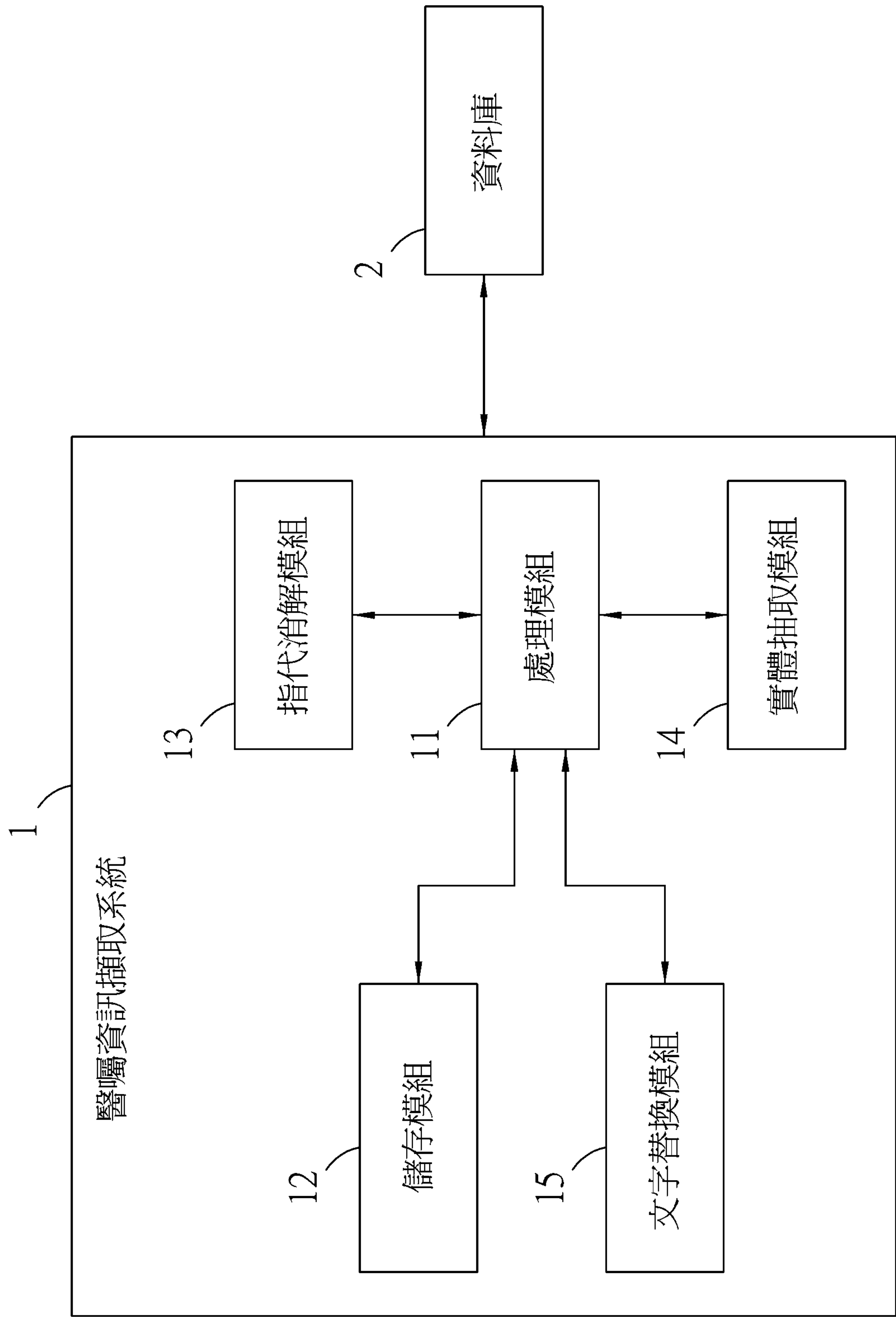


圖1