

類 科：圖書資訊管理（選試英文）、圖書資訊管理（選試日文）

科 目：資訊系統與資訊檢索

考試時間：2 小時

座號：_____

※注意：(一)禁止使用電子計算器。

(二)不必抄題，作答時請將試題題號及答案依照順序寫在試卷上，於本試題上作答者，不予計分。

- 一、資訊檢索系統可供檢索文件的來源是非常多元的，可能是原生數位全文（Born Digital），可能是紙本文件經過重新打字成為數位全文文件，也可能是採用光學字元辨識（Optical Character Recognition, OCR）的作法，將紙本文件轉變為數位文件。OCR 的作法顯然會有辨識錯誤的情形，請說明對於資訊檢索系統而言，OCR 的文件是否需要進行校正，如果需要是基於何種使用需求，何種考量因素？如果不需要又是基於何種使用情境，何種考量因素？（25 分）
- 二、近年來，鏈結開放資料（Linked Open Data, 簡稱 LOD）受到大家的重視，國內外圖書館界亦陸續討論圖書館書目資料轉換為鏈結開放資料的可行性。請敘述何謂 LOD，並說明圖書館書目資料轉換為 LOD 後，對於使用者而言，相較於使用傳統的線上公用目錄（Online Public Access Catalogue, OPAC），有何重大的不同？（25 分）
- 三、當使用者應用資訊檢索系統或是網路搜尋引擎時，可能因為種種原因，輸入錯誤的詞彙或是不合用的詞彙，某些資訊檢索系統或是網路搜尋引擎會提供建議的詞彙或是相關的詞彙。請說明這些資訊檢索系統或是網路搜尋引擎是如何判斷使用者輸入錯誤的詞彙與不合用的詞彙，以及如何產生前述的建議詞彙與相關詞彙？（25 分）
- 四、目前資訊檢索系統多為單語資訊檢索系統（Monolingual Information Retrieval System，或稱為 Single Language Information Retrieval System），跨語資訊檢索系統（Cross Language Information Retrieval System）較為罕見。請說明建構跨語資訊檢索系統時，必須處理那些難題，以及處理這些難題可以應用的技術。（25 分）