# 異質生物醫學資料庫整合系統設計

# Modeling Heterogeneous Biomedical Database System

## 中文摘要

在這資訊價值發掘的時代，無論是 HIS(Hospital Information System ) 、BI(Business Intelligence)、EIS(Enterprise Information System)、KM(Knowledge Management)、Data mining，甚至生物資訊(Bio-informatics)等等，透過資料的萃取、整理與呈現即可彰顯其資料的重要價值，而這些工作都將會面臨異質資料庫的整合問題。

本論文探討現有的異質資料庫整合工具及方法，並且研究各資料庫結構、資料欄位的格式與其對應關係。設計一個圖形化使用者介面系統雛型，讓使用者在不需了解不同資料庫間的差異之下，可以透過簡單的操作，即能同時對多個異質資料庫之資料進行維護並萃取所需的資料。

本論文所設計的系統雛型相對於已商品化的 ETL 產品，除了簡單易懂的圖形化操作介面外，使用者亦不需深度了解多個資料庫管理知識即可進行多資料庫的資料維護、萃取與轉移。

爲了對本論文所設計出的異質資料庫整合工具雛型進行評估，本論文以健保資料庫與某區域醫院資料庫爲材料，從文字檔案將資料載入到各異質資料庫並透過本系統雛型對各資料庫進行查詢、新增、修改、刪除，再匯入特定資料庫進行資料整合。最後核對資料的正確性及時間，以證明本異質生物醫學資料庫整合系統的方便性及可用性。

## 英文摘要

In the age of information value excavation，HIS、BI、EIS、KM、Data Mining，even the Bio-informatics could manifest the value of information by extracting, organizing and displaying data. And all this work would face the integration problems among heterogeneous databases.

This thesis is to probe into the current integration tools and methods of heterogeneous databases. Also，it will study the data structure，data type of data fileds and the corresponding relationships. I will design a graphic user interface as a system prototype，so that users can maintain and extract needed information through simple manipulations without understanding the differences among heterogenous databases.

The system prototype in this thesis has the interface which is eazy to understand compared to other commercialized ETL products. Users have no needs to realize various knowledge of database management，but they can maintain、extract and transform in multi-databases.

In order to evaluate the prototype of heterogeneous databases integration designed in this paper，therefore，this thesis will use the National Health Insurance Database and certain local hospital，load the text files into heterogenous databases and query、insert、update、delete the data in databases by means of the system prototype.
And then，converge defined database to run the data integration. Eventually, it will verify the correction of data for proving the convenience and practicability of this Heterogeneous Biomedical Database Integrate System .