

[01 課程總表 >](#)

本體結構與文史研究專題

課程資訊：

•

科號 Course No.	科目名稱 Course Title	上課時間 Time	授課老師 Teacher	學分 Credit	教室 Room	修課人數 Enrollment
10520CL 533900	本體結構與文史研究專題 Topics on Ontology and Researches in Humanities	T6T7T8	祝平次	3	計中電二	0

•

- 助教：許漢祥
- 辦公室時間：1:20-3:20(一)，人社B508；12:20-2:20(二)，計中電二(請先以電郵或辦公室電話42742確認)

課程說明：

- 優越的文史學科研究，常常需要植基於對於文本資料的掌握。以前的文史研究者對於文本的掌握，常常依賴超人的記憶力，以及卡片檢索系統。但這樣的情形，由於教育培養的方式差異，今天的文史學者很少接受上面兩者的訓練，導致對於資料掌握的情形並不理想。這種情況更因為文史研究能力需要長時期的培育，因而更形嚴重。但由於數位資訊技術的發達，利用數位技術來彌補現今學者缺乏傳統訓練，成為人文學界必須嚴肅面對的問題。本課程的設計，有兩個重點。第一個重點，在於學習如何利史丹福大學所開發的 [Protégé](http://protege.stanford.edu/) (<http://protege.stanford.edu/>) 數位工具把文本變成知識庫(knowledge base)；第二個重點，則在第一個重點的基礎上，利用文本分析工具(庫博與 [AntConc](#))，來進行文本分析。每位修習課程的同學，需在最後一週，在「本體知識工作坊」上呈現自己一學期實做的成果，並完成一篇依據自己所擇取的研究文本的學期研究報告。報告的成果，必需與課堂上學習的研究方法相關。
- 以下再對兩個課程重點詳加說明。
- 知識庫的結構(ontology)，在台灣譯為「知識本體」，原為西方哲學中的「存有論」，討論物的存在以及物與物之間的關係；後來被資訊學者轉用到知識庫的建立上。而建立知識庫的一大目的，則在於讓電腦可以自動判讀資料，以及在電腦與電腦之間進行資料分享。[\(Semantic Web\)](#) 簡單來講，知識本體就是讓電腦可以認識「世界」或「世界」的某個特別的部分。而所謂的「認識」，和人類的語言的運用有密不可分的關係，因而知識本體也和讓電腦能自動判讀語意有關係，最著名的例子之一，就是美國普林斯頓大學發展出來的 [WordNet](#)。中文的詞彙，也有一些嘗試，如中研院語言學 [中文詞彙網路](#)。本課程的目的，並不在於學習建立一個抽象的知識本體，而在於為現成的文本編造一個知識本體。藉由建立知識本體的過程，將文本後面所隱含的世界突顯出來。並藉由這個過程，加強對於文本的掌握。亦即，我們要藉由建立知識本體的軟體來突顯資料的結構。例如，在《論語》裏，我們可以質問下列的問題：有那些弟子問過孔子話；有那些弟子被孔子評論過；有那些弟子是孔子主動跟他講話？當我們這樣結構化《論語》裏的文本以後，我們很容易就可以利用同樣的問題去質問其它的文本，從而比較在先秦文本問答體的一些相關問題。當然，我們也容易可以去問，什麼樣的「器物」以什麼樣的分類出現在一個文本裏，然後比較在不同文本裏所呈現的「器物」世界的不同。然後，可以進一步思考結構化資料與我們研究的關係，期待在這樣的操作下，對於一些經典議題可以有新的發掘方式。
- 文本分析的重點，則著重在利用一些軟體，進行關鍵詞分析、n-gram分析、正則表示式的使用等等，並將分析的結果利用MS Excel轉為圖表、圖象的報告。
- 修習本課程結束後，同學將對於資料／知識結構、資料庫以及文本分析，以及資料的形式轉換可以有基礎的掌握，並運用於自己的研究上。對於文本資料的理解深度與整體性的掌握，因為藉著比較具體可視的方性進行，應該會加強很多。在課程的實做中，同學也可以從中學會如何與別人進行協同研究。

- 本課程的重點在於從實做中學習，沒有時間進行每週作業、不願意跟同學互動的同學，請勿修習本課程。

教學進度(Syllabus)：

•

週次	內容
1 (2-14)	課程介紹：課程目的與進行方式。(程式安裝與玩玩看) 作業：選定自己想要整理的文本。
2 (2-21)	本體結構(知識本體、知識庫)與Protégé：示範與討論。
3 (2-28)	和平紀念日
4 (3-7)	webProtégé 和：建置本體結構的軟體與線上協作平台介紹。
5 (3-14)	Protégé：實做與討論一
6 (3-21)	Protégé：實做與討論二
7 (3-28)	Protégé：實做與討論三
8 (4-4)	兒童節
9 (4-11)	Protégé：實做與討論四
10 (4-18)	Protégé：實做與討論五
11 (4-25)	文本分析：庫博與Antconc
12 (5-2)	文本分析：實做與討論一
13 (5-9)	文本分析：實做與討論二
14 (5-16)	文本分析：實做與討論三
15 (5-23)	文本分析：實做與討論四
16 (5-30)	端午節

17 (6-6)	期末報告
18 (6-13)	期末報告

-

線上資源：

- [Protégé OWL Tutorial](#)
- [WebProtege](#)

教學影片：

- 台灣地理資訊中心某工程師：[Ontology簡介](#)
- 台灣地理資訊中心某工程師：[Protégé工具簡介](#)

參考資料：

- [中文詞彙網路](#)
- [protégé 使用手冊](#)