

中文新聞數位化趨勢 國際研討會

理解、意義與內容標誌

謝清俊

玄奘大學 圖資系 2005.05.04

International Conference on the trend of Chinese News Digitalization

Meaning, Understanding and Content Markup

C.C. Hsieh

2005.05.04

大綱


- ❁ 前言—意義與理解
- ❁ 數位化與文獻表達
- ❁ 內容標誌
- ❁ 情境的表達
- ❁ 結語

Outline

- ⊗ Meaning and understanding
- ⊗ Document Representation and digitalization
- ⊗ Content markup
- ⊗ Representation of context
- ⊗ Concluding remark

意義與理解

Meaning and Understanding



數位化與文獻表達

Document Representation
and Digitalization

虛擬的實踐 — 數位化

- ❁ 一個文物的數位化
- ❁ 一群文物的數位化
- ❁ 雙層的文獻結構

數位化的層次

⊗ 外觀的數化

- ★ 文字、圖形、影像、聲樂、多媒體...

⊗ 背景相關資料的數化

- ★ 書目、版權、索引檔案、工具檔案、 *metadata* ...

⊗ 內容相關資料的數化

- ★ 文物本身的詮釋
- ★ 相關的研究著作、說明

A Model of Digitalization

文與物之參照

Text to Object
References

情境之參照

Context
References

背景資料 Background Info.

式形觀外字文
Appearance

內容詮注 Content Notes


互為文本

Intertextuality

文獻的雙語系結構表達

❁ 自然語言與後設語言的相輔相成

數化之文物	使用之表現系統
文件本身	自然語言與藝術語言
背景資料	後設語言
內容詮注	
情境參照	



内容標誌

Content Markup


內容標誌

❁ 標誌的種類

- ★ 句讀、標點符號.....
- ★ 版面標誌 markup for typesetting
- ★ 內容結構標誌 markup for document structure
 - 如: 標題, 篇章節段...
- ★ 詞語標誌 morphological markup
- ★many more

❁ 何謂內容標誌 what is content markup?

- ★ 讀者將他對文章內容的理解作標籤，以便電腦進一步處理。
- ★ 與上述的各種標誌不同。



情境的表達

Representation of Context

情境的表達

❁ 什麼是情境

- ★ 上下文

- ★ 作者情境與讀者情境

Authorial context and readership context

- ★ social context, political context, cultural context...

❁ 意義與情境

- ★ 義隨境轉

情境的表達

❁ 情境的表達


- ★ 字典、辭典、以及各種注疏、註釋都是在解釋意義，也都作適度必要的情境表達。
 - 只是，是給人用的。
 - 有各種表達方式。
- ★ 後設資料〔metadata sets〕即有特殊目的的情境表達，或背景資料的表達。
- ★ 超連接〔hyperlink〕亦可用於作情境表達。
- ★

內容標誌之例

A demonstration of context markup

❁ 心經〔Heart sutra〕多版本處理示範。

A prototype of text processing system of various versions of Heart sutra.



結語

Concluding Remark



完 *The End*
謝謝 *Thank you!*