

# 數位化與語言情境

謝清俊

玄奘大學 講座教授

中央研究院 語言學研究所 兼任研究員

2004年7月12日

台北·中央研究院

# 大 綱

- ✿ 前言：語言與社會
- ✿ 資訊概念的多義性
- ✿ 媒介、傳播、資訊與文化
- ✿ 資訊的界說
- ✿ 一個數化的通用模式
- ✿ 結語

# 資訊、傳播與語言學

---

✿ *The role of linguistics in information access, exaction and dissemination is essential. Radical changes in the techniques of information and communication have had a significant effect on the function of linguistic paradigm and its applications in all the forms of communication.*

- *El Hadi and Widad Mustafa, “Human Language technology and its role in information access and management”, Cataloging and Classification Quarterly; 37(1/2) 2004, pp.131-151*

# What solution can linguistics offer in human computer interaction, extraction and management ?

---

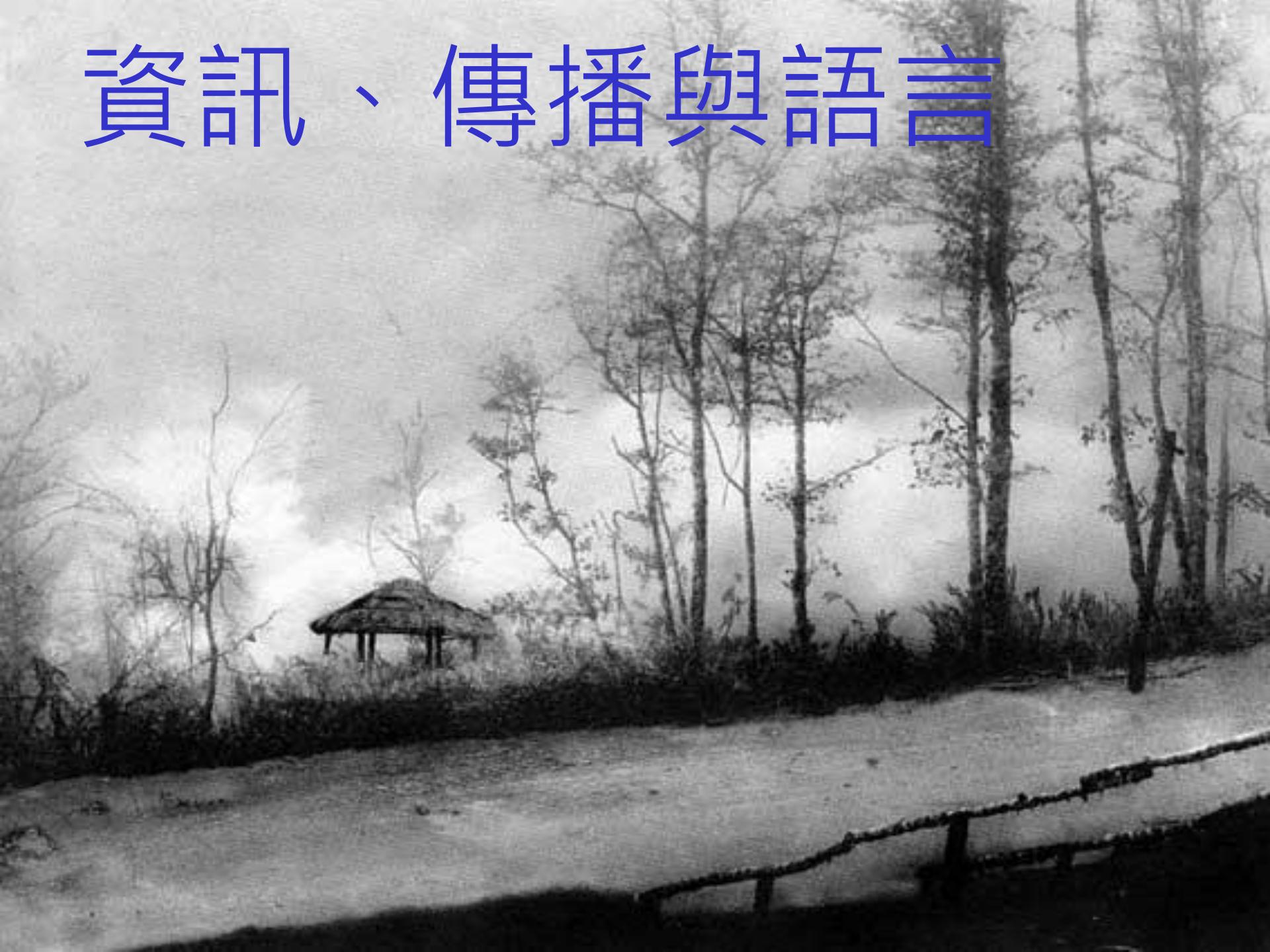
✿ *Many fields shows the relevance of the linguistic paradigm through the various technologies that require NLP, such as:*

- ◆ document and message understanding
- ◆ information detection, extraction and retrieval
- ◆ question and answer
- ◆ cross-language information retrieval
- ◆ text summarization...
  - *El Hadi and Widad Mustafa, "Human Language technology and its role in information access and management", Cataloging and Classification Quarterly; 37(1/2) 2004, pp.131-151*

# 語言現象的變遷

- ✿ 語言學是根據語言現象為首要資料而建立的學科。應用語言的情境改變了，就自然而然產生了語言現象的轉變，進而影響到語言學的研究。
- ✿ 資訊科技已引起傳播系統和傳播行為的急遽變化，以及史無前例的快速社會變遷。這改變，無疑也將影響未來語言學的研究。

# 資訊、傳播與語言



# 時下的資訊概念...

---

- ✿ Of our mundane and technical concepts, information is currently one of the most important, most widely used and least understood.
  - Luciano Floridi, *Two approaches to the philosophy of information*, Forthcoming in *Mind and Machine*
  - Luciano Floridi, *Open Problems in Philosophy of Information*, Forthcoming in *Metaphilosophy*, volume 35, no.3, April 2004

# 資訊概念混淆...

---

- ✿ From the perspective of semiotics, "information" is an ambiguous theoretical concept because the word is used to represent both signifier and signified, both text and content.
  - Douglas Raber & John M. Budd, *Information As Sign*, Journal of Documentation; 2003;59,5

# ...是一個眾說紛紜的問題

---

- ✿ The definition of Information is a problem over 50 years since before the term “Information Science” was coined in 1955.
  - A.D. Madden, *A Definition of Information*, Aslib Proceedings vol. 52, No.9, p.343-, 2000.10

# Information as...

- ❖ thing, goods
- ❖ knowledge, power
- ❖ resource, resource of resources
- ❖ part of communication
- ❖ part of a system
- ❖ ...

❖ 在下列文件中收集了二十多種說法：

Alvin M. Schrader, *In Search of a Name: Information Science and Its Conceptual Antecedents*, Library and Information Science Research, 6:4, pp227-271, 1984

# 資訊科學...

---

“Information science, by nature and necessity, is interdisciplinary.”

- Douglas Raber & John M. Budd, 2003,  
p.507

✿ 以下，讓我們綜觀傳播、文化、語言、媒介等與資訊科技和數位化的關係...

# 傳播與資訊

- ❖ 很多人談資訊，第一個聯想到的是電腦，很少想到傳播。
- ❖ 試想，若沒有資訊，能傳播些什麼？
  - 還會有傳播嗎？
- ❖ 反之，若沒有傳播，資訊從何處來？
  - 我們能偵知任何資訊嗎？

# 傳播與媒介材料

- ❖ 事實上，沒有傳播，不止沒有資訊，連任何群體、社會、文化都不會生成；當然，也就不會有人類的文明。
- ❖ 然而，傳播或溝通必需借助物質作為媒介，否則，人們無法偵知傳播或溝通的行為。所以，媒介材料就從根本處影響到文明的進程。

# 媒介、資訊與文明

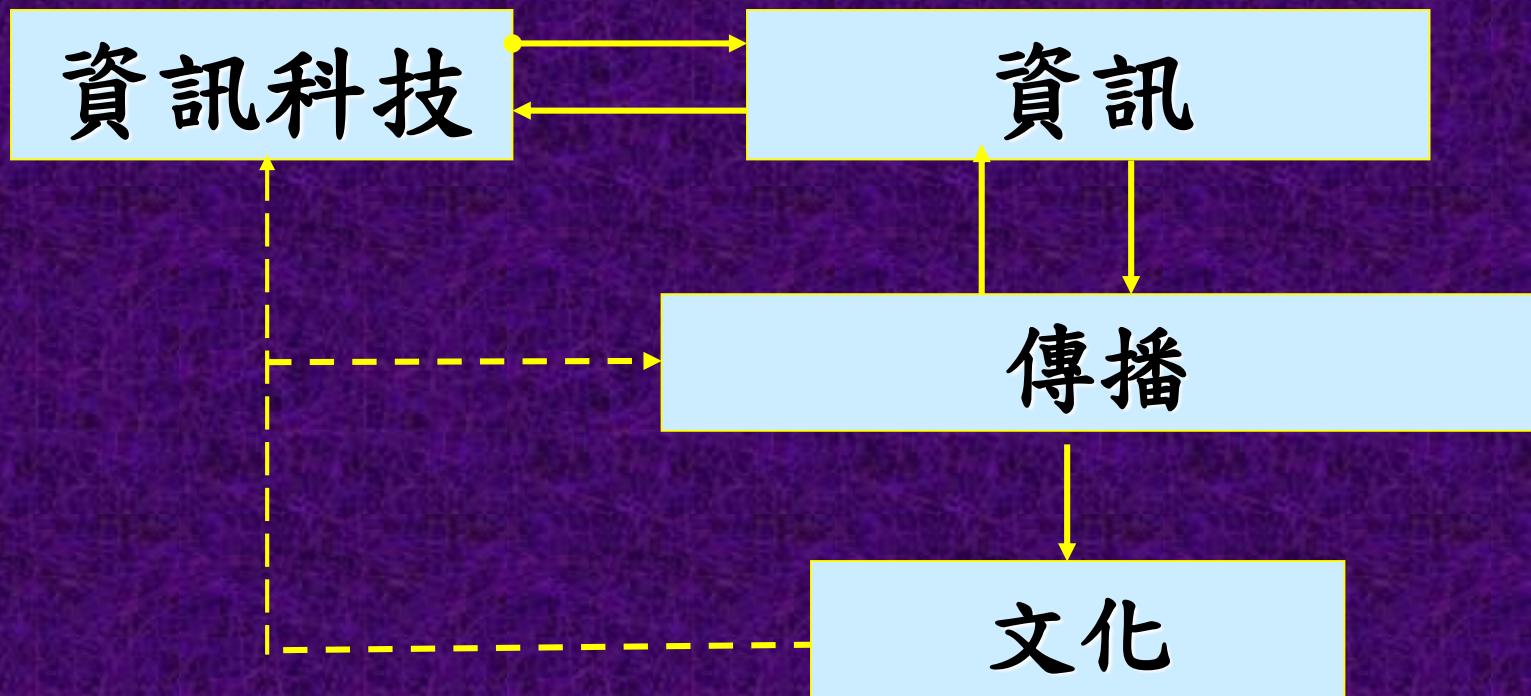
## 從人類文明發展的歷史觀察

- ◆ 凡是出現一種新媒介時，必定引發信息和知識傳播方式的改變。
- ◆ 新媒介誘發新工具的發明，新工具又導致技術的進展，因而擴大了人們獲得知識的範疇、深度。
- ❖ 媒介之於溝通和知識處理的影響非常大：常引起人際關係的變化、導致組織和社會的變革，而產生新的文明。

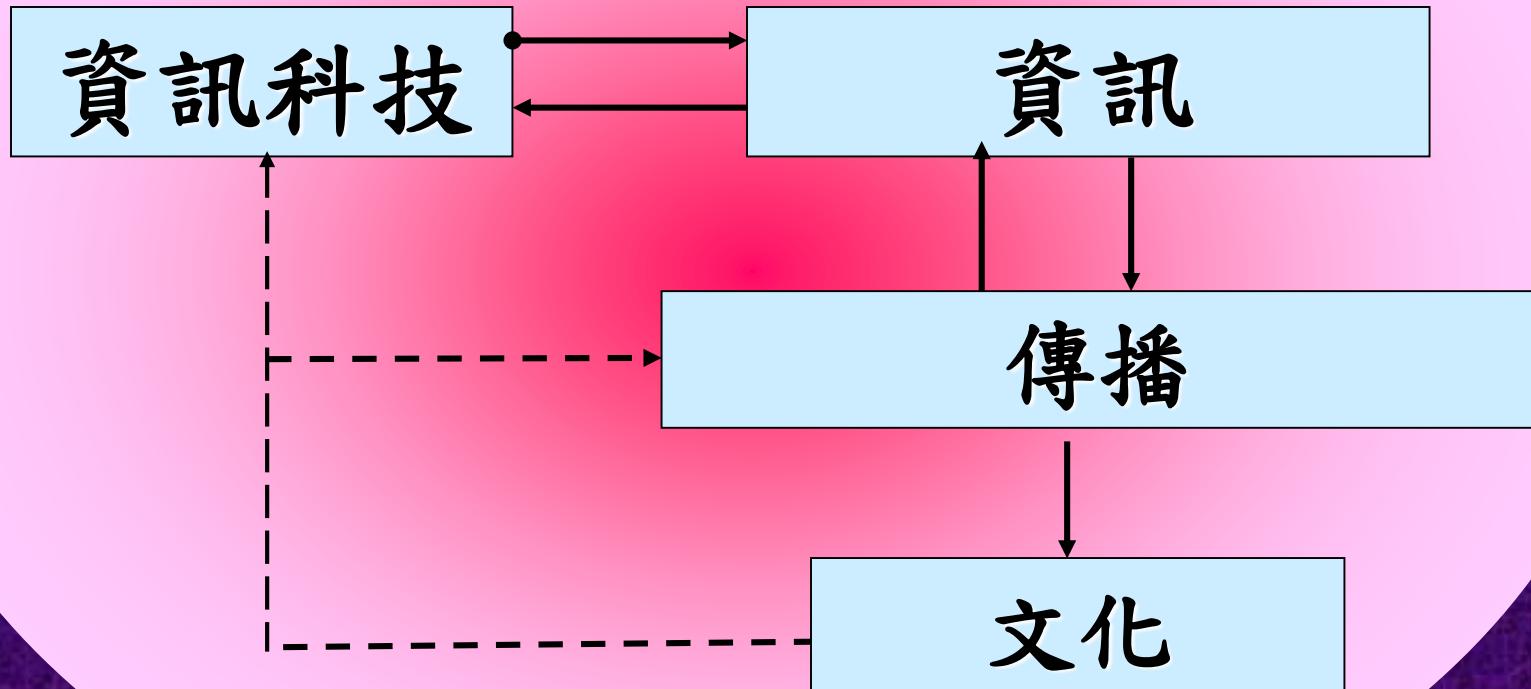
# 傳播科技的進程

- ✿ 西元230至1830年  
(共1600年)
  - ◆ 紙的發明 西元 105年
  - ◆ 彫版印刷 西元 650年
  - ◆ 活字印刷 西元 1045年
  - ◆ 鉛筆 西元 1630年
  - ◆ 鉛筆擦,複印紙 西元 1830年
  - ◆ 電報電話 西元 1870年
  - ◆ 廣播 西元 1910年
  - ◆ 彩色電視 西元 1950年
  - ◆ ESS,通信衛星,光纖 西元 1970年
  - ◆ PC,光纖通訊 西元 1990年
  - ◆ ATM, PCS, CD, WWW, 多媒體.....
  - ◆ ? ? ? .....
- ✿ 西元1830至1990年  
(共160年)
- ✿ 西元1990至2006年  
(共16年)

# 資訊科技與文化

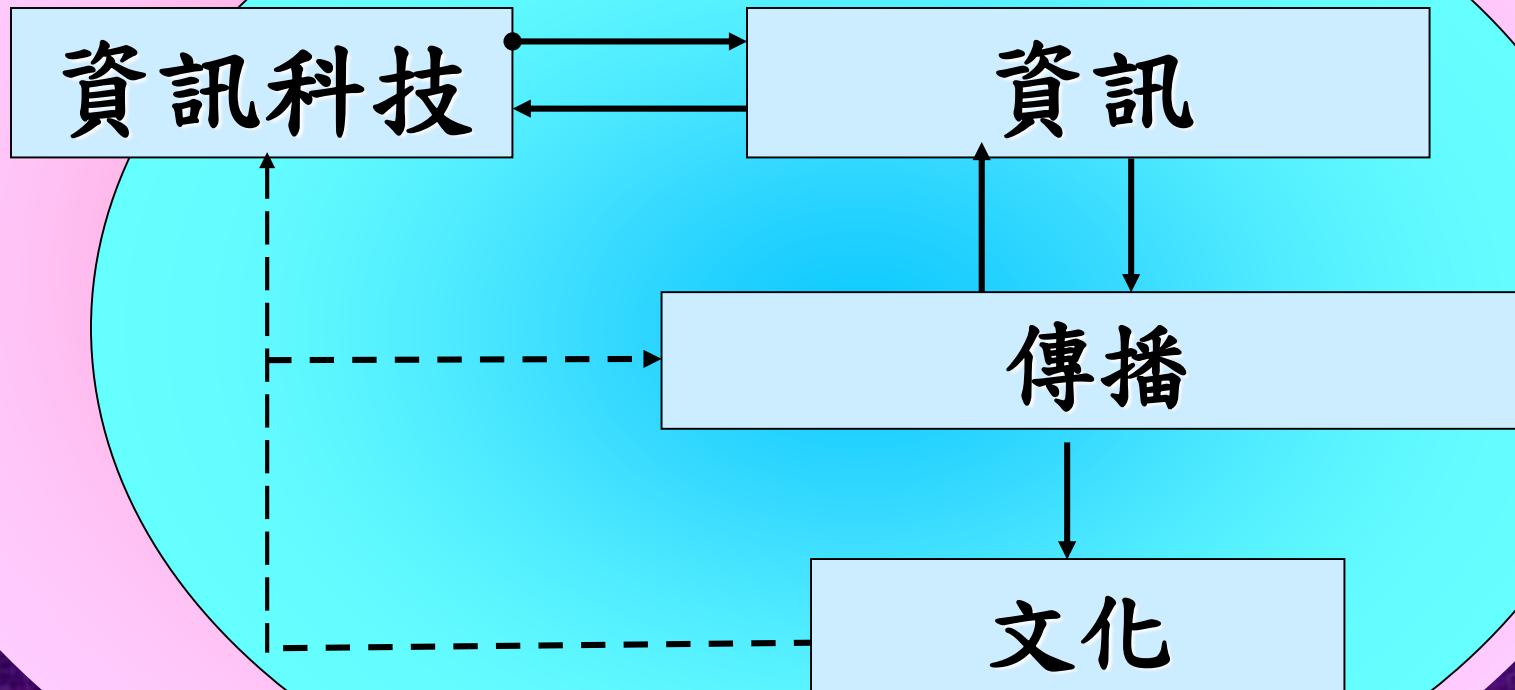


# 媒介



媒介

表現系統



# 表現系統

---

◆ 語言系統

◆ 記號系統

# 語言、媒介與文明

## 文明發展的三個階段

- ◆ Audio Space 口語文明
- ◆ Visual Space 文字文明
- ◆ Audio Space 多媒體文明
  - ❖ *Electronic Multimedia*

Marshall McLuhan,  
*Understanding Media*,  
McGraw-Hill, 1964

# 媒介物、損耗 與複製狀況

文化三階段

表現系統

**Audio Space**

口語

聲波，無損耗  
無法複製

(素養問題，語言內轉移)

**Visual Space**

文字

物質加工，有損耗  
複製不易

(功能素養問題，表現系統內轉移)

任何形式皆可數化

**Audio Space**

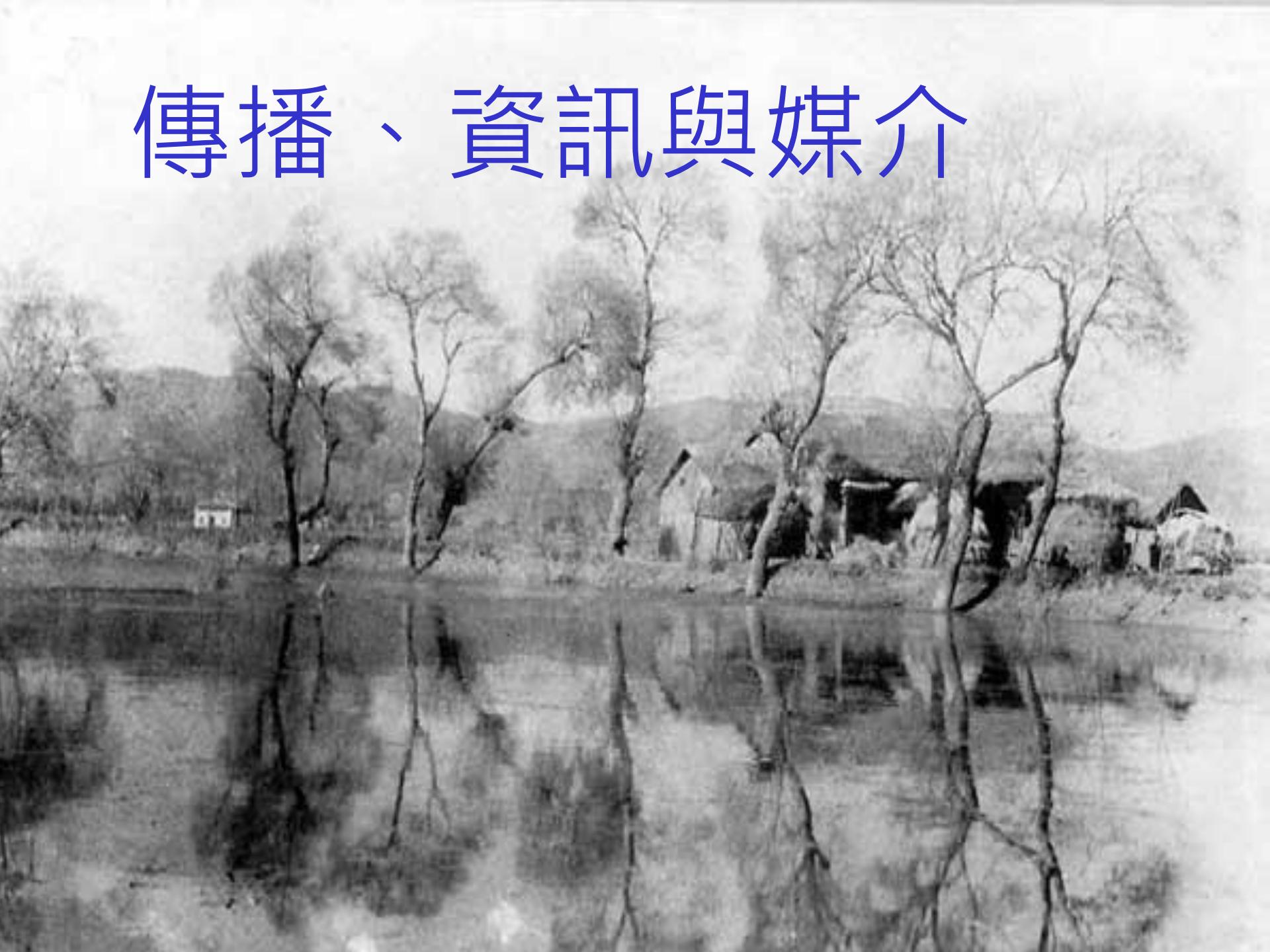
多媒體

多媒體

加工物質中之  
數位能階，無損耗  
複製容易

媒材與文化進程的關係

# 傳播、資訊與媒介



# 媒介

- 媒介材料
  - 依媒介材料所研發的工具、設備
  - 依工具、設備所發展的技術
- 
- 表現系統
    - 文字、語言→文件、書籍、檔案...
    - 記號系統→藝術品、號誌、圖像標誌...

媒介與表現  
系統之間相  
互影響...

# 傳統媒介的物質障礙

- ◆ 傳統媒介種類甚多，且都是以消耗或破壞物質的方式來表達知識，不僅要耗用物質，更用去不少資源。
- ◆ 如果用以製成產品，便會受制於這種媒介的物理性質，而有生產、儲存、運輸、分配等經濟上的問題。
- ◆ 在使用時，除要注意保養維護以外，還要面臨折舊、損耗、腐壞、甚至於遺失、盜取、水火災害等等。這些都是傳統媒介攀附物質所得的障礙。

# 電子媒介、能階媒介

- ◆ 電子媒介利用帶電的粒子、電波、電磁材料、光電材料等的能量變化，或是能量平衡的狀態來表達所知。因此，電子媒介在表達所知時只耗用少量的能而無物質損耗。
- ◆ 能的消耗常常是可以補充，且能夠使物質回復到原來狀態。
- ◆ 數位化統一了所有的傳統媒介。因此，數位能階媒介就主宰了未來傳播或資訊的發展

# 源頭活水

- ✿ 數位能階媒介像是資訊或是傳播的基因，由此基因的性質能推演出各式各樣應用系統的性質。
- ✿ 了解能階媒介和數位資訊的性質，正是掌握了整個資訊科技和傳播科技發展的源頭。

# 資訊的界說與性質



# 所知與資訊



人類有『致知』的能力。

◆ 古時論及認知時，常稱人為『能知』  
把所知道的所有事務統稱為『所知』

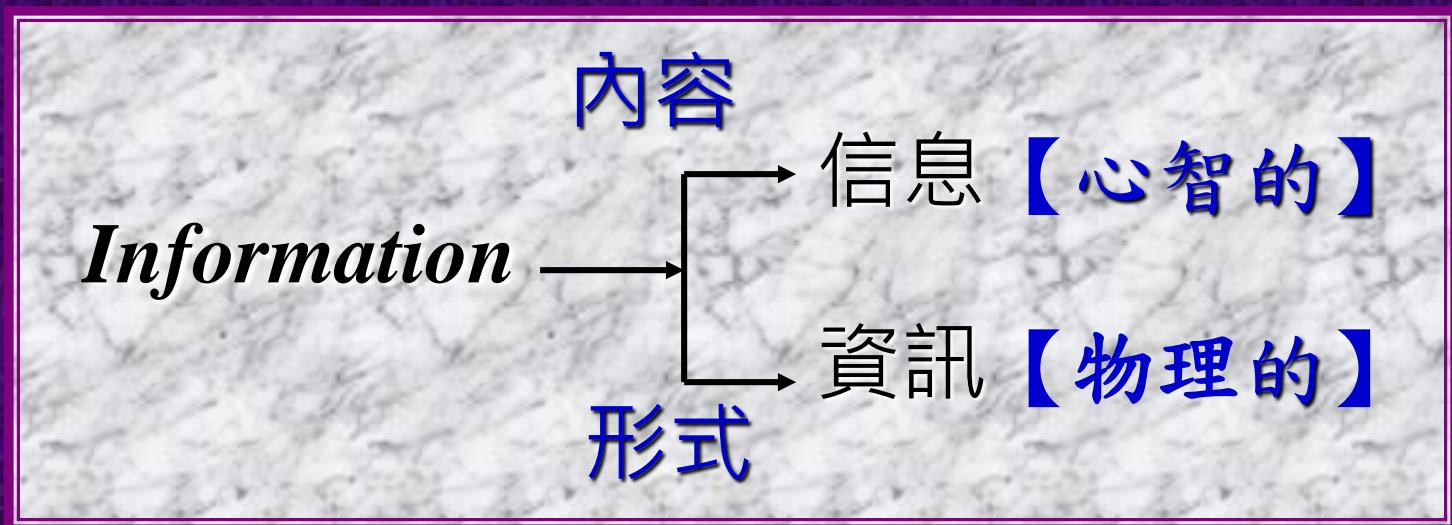
►是故所知中有：

- 知性的成份                  如常識、知識；
- 也有感性的成份          如感覺、感觸；
- 還有創意成份                  如規畫、設計；
- 意志成份                  如信仰。

# *Information*

✿ *Information* 有內容和形式兩個面向：

- ◆ 心智活動是無形無象的
- ◆ 物理現象則可以被偵知



# 一些學科中心物雙生的關係

	心智的	物質的
美學 <i>Aesthetics</i>	Content	Form
記號學 <i>Semiotics</i>	Meaning, Ideas	Sign
語言學 <i>Linguistics</i>	內容, 語意, 意義	語言現象, 語法
資訊學 <i>Info. Science</i>	Information 信息	Information 資訊

# 從資訊的產生 對資訊的界定

資訊即：所知表現在媒介上的形式

立論方針：從致知的行為(認知)、傳播、以及表現(美學)等資訊的產生點出發，並針對資訊的內涵、動作、所用的媒介、表現系統等相關的重要因素，來界定資訊的界說，並依此導出數位資訊與資訊科技的基本性質。

# 資訊的界定

✿ 資訊的界定分別從創作端與接收端(或閱聽端)來釐定。這是考慮到

- ◆ 創作情境（authorial context）與
- ◆ 閱聽情境（readership context）

可能有甚大差異的緣故，也考慮到

- ◆ 所知的外化、表現〔傳播者是有意的〕與
- ◆ 記號的理解、詮釋〔傳播者是無意的〕

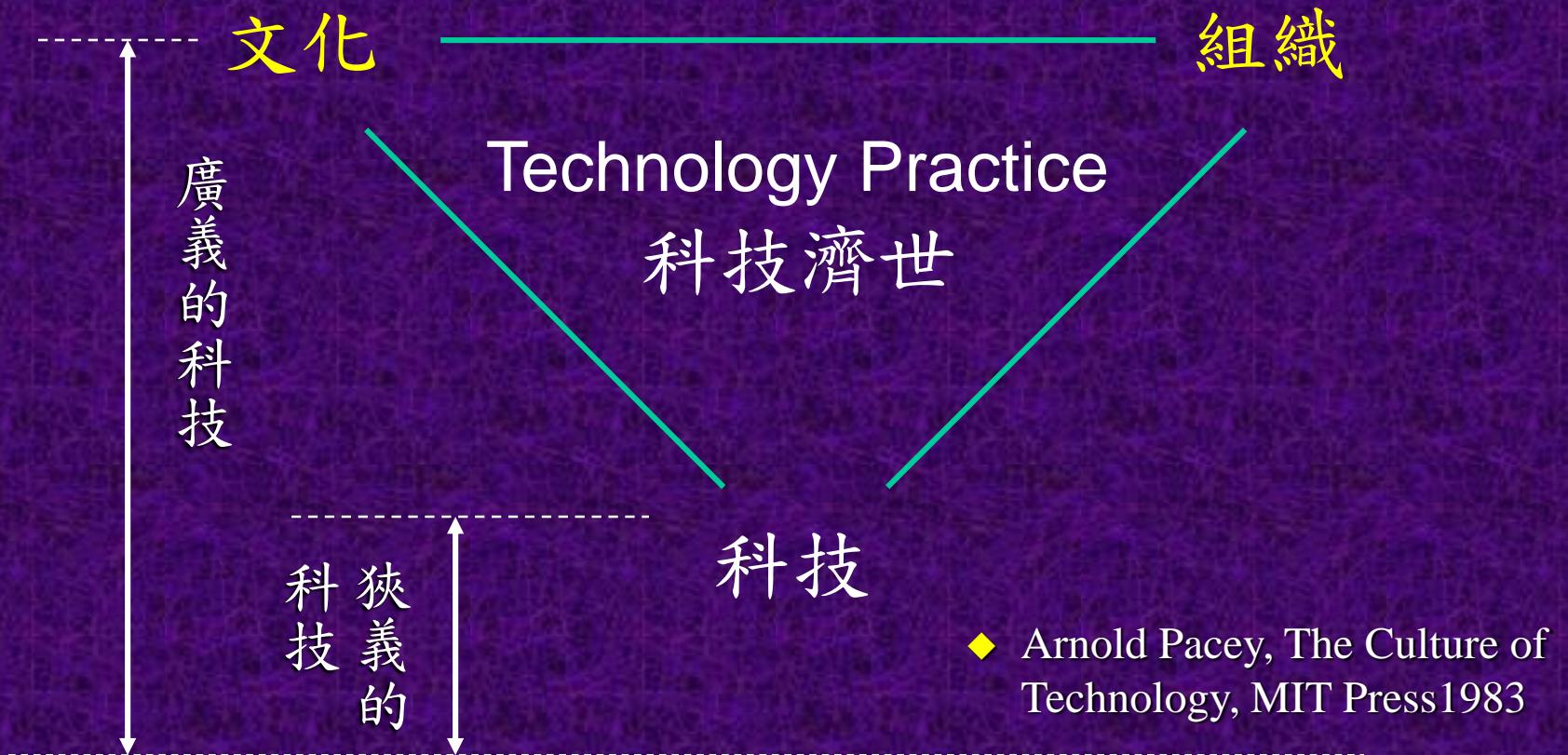
等行為的性質有相當大差異的緣故。

# 所知、資訊與知識

- ✿ 所知是資訊的內容，資訊是所知的形式。
- ✿ 資訊並不完全等於所知，它是所知在媒介上的投影，它承載著所知，它是我們可由感官察覺的。
- ✿ 知識是所知的一部份，所以資訊承載著知識，資訊不等於知識。

在應用時，我們用的是信息  
〔即資訊的內容—所知〕而  
不是其形式。

# 科技、資訊科技的運用環境



# 媒介轉換—數位化

- ✿ 數位化即將傳統文物以數位能階媒介表達。
  - ◆ 對『先行媒體』而言，這是經數位能階媒介的轉換而產生『後續媒體』的過程。
- ✿ 以下，介紹數位化的理論根據、模式與實務。

# 記號學美學 Semiotic Aesthetics

◆ 蘇珊•郎格(*Susanne K. Langer, 1895-1982*)：

◆ 藝術是表現人類情感的記號形式。

◆ 文化記號 cultural signs

◆ 把審美和藝術現象歸結為文化記號或文化符號。

◆ 對人文主義和科學主義的美學都有較大的包容性。

◆ 亦適用於文學。

◆ 如葉嘉瑩教授利用符號學對詩詞之詮釋。

# 記號學美學

## ◆ 文化記號

### ◆ 推理記號 *reasoning signs* :

❖ 是內涵概括確定的理性符號，它既可以翻譯，也可以被分解、推理；如語言符號。

### ◆ 表象記號 *image signs* :

❖ 是非理性的、完整獨特不能被分解的，具豐富含義的情感意象；如藝術符號。

\* 蘇珊•郎格(Susanne K. Langer, 1895-1982)

可參考《電影語言學導論》中國電影出版社, 北京, 1996 pp.20-22

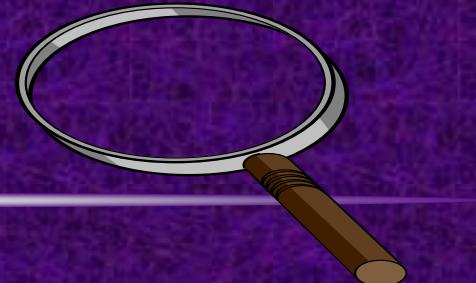
# 記號與資訊

- ✿ 對觀察者而言〔或在資訊的接收端〕資訊即我們觀察的對象。這對象是在物理世界存在的實物或現象。
- ✿ 此實物或現象，即記號學所說的『記號〔*Sign*〕』
- ✿ 在接收端資訊的界說：

資訊即記號或形式

- ❖ 我們對某形式進行理解或詮釋時，此形式即資訊。

# 資訊的性質



- ✿ 壹：因襲了所知的性質。
- ✿ 貳：依附媒介物質所得到的性質。
- ✿ 參：駕馭媒介工具與技術所增益的性質。
- ✿ 肆：從表現系統表達內容的手法或溝通的品質和效果上所呈現的性質。

# 資訊與虛、實的關係

## ✿ 文明與虛擬 — 虛擬不是現代才有的

- ◆ 文明之前的虛擬 — 記號和符合的世界
- ◆ 虛擬的表達系統 — 語言
  - ❖ 感性的語言 — 許多動物皆有
  - ❖ 理性的語言 — 只有人類和電腦才有
- ◆ 虛擬的記錄系統 — 文字
  - ❖ 造就純想像的虛擬世界
  - ❖ 『百官以治、萬民以察』
  - ❖ 跨越時空

# 虛擬實境—境由心造！

- ✿ 虛擬實境的前身即系統的模擬。
- ✿ 虛擬實境與實際的世界
  - ◆ 虛擬實境具有實體世界中事物的一部份功能。
    - ❖ 這是我們所設計、所想要的。
  - ◆ 在我們設計的虛擬實境裡，有與實體世界不同、且超越實體世界的性質與功能。
    - ❖ 這正是我們想要超越實體世界的、也是我們想要利用的。

# *Virtual*：虛 擬 ？

## **Virtue**

### **Virtual**

(adj.)

✿ 實際上的；實質上的；事實上的。

✿ 語源：指有實質能力、效用、以及效果的。

例: She is the virtual president, though her title is secretary.

### **Virtuous**

✿ 有德性的、高潔的；貞節的、堅貞的。

✿ 按：*Virtue* 指

- ◆ 德與善
- ◆ 優點、長處、價值
- ◆ 效能、效力、功效

此觀念與用法在中西古文明中甚相似。

『擬諸其形容，而象其物宜。』

—《易》

# 虛擬實境的變幻源頭－媒介

- ✿ 新媒介的利用，促成新的虛擬系統產生。
  - ◆ 例：廣播、電視、手機、網路 ...
  
- ✿ 數位媒介引發了溝通與傳播基本的改變
  - ◆ 改變了溝通與傳播的方法、形式與效果。
  - ◆ 也改變了人們利用溝通與傳播的觀念、行為和價值觀。
  - ◆ 隨之，知識的擁有、儲存、發掘、散播和利用也改變了。

# 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

# 參照的連接

## ✿ 互為文本 (*Inter-textuality*)

- ◆ 文本之間
  - *Julia Kristeva*
- ◆ 學術領域之間
  - synergy

## ✿ 情境之參照 (*Context Reference Links*)

- ◆ 作者情境、讀者情境
- ◆ 個人情境、社會情境、文化情境...

## ✿ 文與物之參照 (*Multimedia Reference*)

- ◆ 文與物之彼此參照
- ◆ 人文與自然科學之彼此參照

# A Model of Digitalization

文與物之參照

情境之參照

Multimedia  
References

Context  
References

Text to  
Objects  
and vice  
versa

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

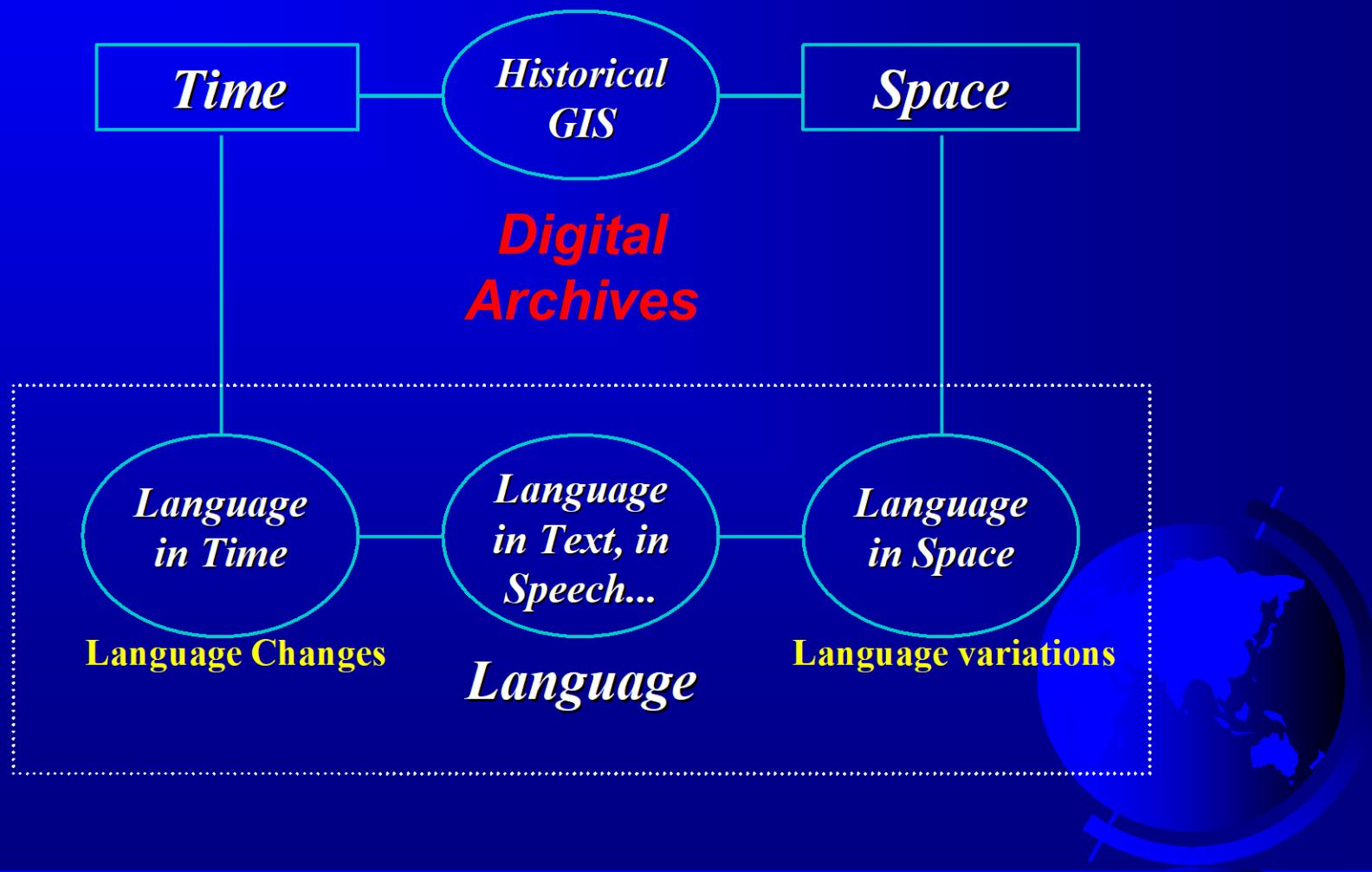
## ✿ 自然語言與後設語言的相輔相成

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	
內容詮注	後設語言
情境參照	

# 數化工程考量重點

- ✿ 智財權
- ✿ 時間、空間與語言座標
- ✿ 多語問題
- ✿ 公共資訊系統之建立
- ✿ 技術規格與工業標準
- ✿ 後設語氣與文獻表達
- ✿ 協力與合作

# *Space, Time and Language Coordinates for Digital Archives*



## 結 語 I

- ✿ 資訊的定義與概念一直是困擾學術界的问题。本報告釐清了資訊的概念，也提供了一個通用的資訊定義。
- ✿ 根據本報告提供的資訊定義，可以推演出四類資訊的性質，有系統的對資訊的性質與作用作全面的說明。
- ✿ 根據本報告提供的資訊定義，可以對目前文獻中無法說明的資訊概念清楚的說明其緣由。

## 結語 2

- ✿ 虛擬實境是實際世界的模擬，在某些性質上與實體世界有相同的作用。所以，我們可以利用虛擬實境做些實體世界中想做又不容易做的事。
- ✿ 虛擬實境是依據數位能階媒介而產生。數位能階媒介超越了物質的障礙，虛擬實境也就擺脫了物質的障礙以能量的方式示現。

## 結語 3

- ✿ 虛擬實境的發展約略顯示出文明進展的軌跡；文明的進程也與我們採用的媒介息息相關。本報告亦說明了媒介、語言、傳播、資訊、知識以及文化等彼此之間的關係。
- ✿ 虛擬實境的發展越成熟，我們用虛擬實境的機會就越增加。現代人做事已經常常往返於這『虛實』之間。

## 結 語 4

- ✿ 什麼是『虛』？ 什麼是『實』？  
我們如何對『虛』？ 就如同我們對待  
『實』嗎？ 這值得我們深思。
- ◆ 現代的道德、倫理問題，事實上涉及  
許多如何對待『虛擬』成份的問題。  
這部份正是顛覆傳統道德、倫理、以  
及社會秩序的源頭。

# 結 語 5

---

- ◆ 數位化是人類文明進程中勢之所趨，無可規避。
- ◆ 數位化是人類累積的資料、知識、以及文物的全盤整理，也是人類有史以來最全面、最重大的知識工程。這工程最重要的基礎即語言學。
  - ◆ 多語、多媒體的語言情境

## 結 語 6

本報告談到媒介轉移引起的文化變遷，對未來的語言學研究和應用，勢必產生相當的影響。此影響目前雖然只處於啟萌階段，提高警覺未雨綢繆應該是因應此變局的最佳策略。

報告完畢

---

謝謝

---