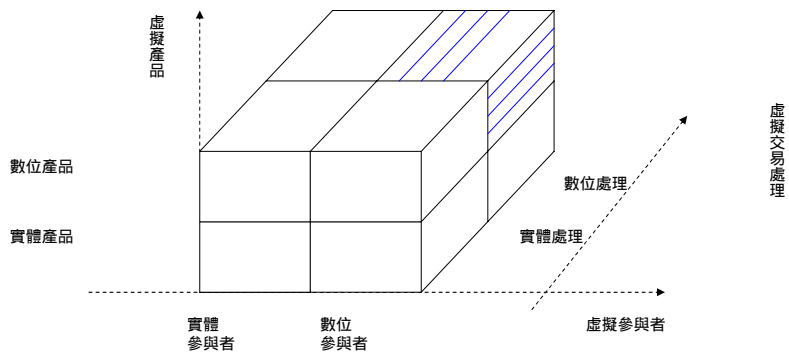


電子資源命名方式

淡江大學：歐陽崇榮

2004年7月

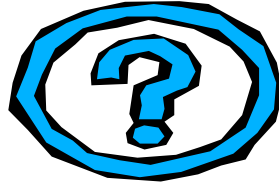
電子商務



電子商務構成要素

資料來源：參考Choi, S., D. O. Stahl, A. B. Whinston, The Economics of Electronic Commerce, 1997.

電子商務



最重要的是什麼？

前言

- ✿ 識別碼使用廣泛，如商品條碼與國際標準書號
- ✿ 數位出版品與網路資源大量興起，對識別碼有新的需求
- ✿ 網路環境的因應：UR*家族
- ✿ 數位典藏與長久保存
- ✿ 目前識別系統無法處理各類型數位出版品

商品條碼

- ✿ 條碼最早起源於1949年
- ✿ 1973年UCC選用UPC碼建立條碼系統
- ✿ 1977年EAN由歐洲12個工業國家共同推廣，條碼系統因此開始步入國際化領域中
- ✿ 1992年改名為 EAN International
- ✿ 區分為：EAN-13碼和EAN- 8碼

ISBN(1)

- ✿ ISO 2108:1992
- ✿ ISBN是因應圖書出版、管理需要，並便於國際間出版品的交流與統計所發展，用以識別出版品所屬國別地區(語言)、出版機構、書名、版本及裝訂方式。
- ✿ 1980年與EAN簽訂合作合約，成為圖書商品條碼
- ✿ 結構

群體識別碼 - 出版者識別碼 - 書名識別碼 - 檢查碼

ISBN 957 - 678 - 000 - 4

ISBN(2) – 目前發展

🌟 ISBN標準

- ISO/TC46/SC9更新標準
- 在2005年1月公布修訂第四版，於2007年1月1日前完成所有改變措施
- 不建議對數位檔案更改ISBN原來的應用原則
- 2005年1月，ISBN將從10碼擴展到13碼
- 使用BICI 作為圖書內容元件的命名方式
- 詮釋資料

🌟 ISBN使用手冊

- 國際書號總部成立手冊更新委員會
- ISTC
- 具有行動/解譯能力

ISSN

🌟 ISO 3297:1998

🌟 用來識別包括電子期刊在內的連續性出版品

🌟 預定於2006年1月完成第四版的修訂工作

🌟 結構

- ISSN的結構
- ISSN register的結構
 - ISSN – 識別題名 – 識別題名的縮寫
 - 補充性資訊
 - MARC

CODEN

✿ CODEN是一組六個符號的代碼，用來識別全世界期刊或非期刊出版品，不只用於化學相關文獻，也適用於所有主題範圍。

✿ 結構

- 連續性刊物：AAAADC，如JUPJ2U
- 非連續性刊物：NNAADC，如56OFAM

SICI(1)

✿ ANSI/NISO Z39.56-1996

✿ 定義期刊內容元素的唯一識別碼，適用於任何連續性出版品的所有格式。

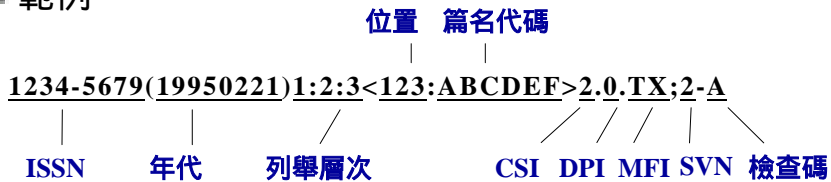
✿ SICI利用ISSN來定義期刊刊名

SICI(2)

- 結構

| 項目部分 | 稿件部分 | 管理部分 |
|----------------|--------------------|---------------------|
| ISSN(年代)項目列舉層次 | <位置: 篇名代碼: 其他編碼綱要> | CSI.DPI.MFI;SVN-檢查碼 |

- 範例



PII

PII的結構與編配主要是為識別出版品項目。

- 結構

| | | | | | | | | | | | | | | | | | |
|-------|----|------|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|
| 範例 | S | 0 | 1 | 6 | 5 | 3 | 8 | 0 | 6 | 9 | 6 | 0 | 0 | 4 | 0 | 3 | 8 |
| PII結構 | N1 | N2 | N3 | N4 | N5 | N6 | N7 | N8 | N9 | N10 | N11 | N12 | N13 | N14 | N15 | N16 | N17 |
| 期刊項目 | S | ISSN | | | | | | | 年代 | 項目 | | | | | | 檢查碼 | |
| 圖書項目 | B | ISBN | | | | | | | 項目 | | | | | | 檢查碼 | | |

BICI

- ✿ 目前處於NISO的標準草案階段
- ✿ 主要用來識別非連續性出版品的項目
- ✿ 利用ISBN來定義圖書
- ✿ 結構

| 項目部分 | 稿件部分 | 管理部分 |
|------------------|-------------------------|---------------------|
| ISBN(年代) 項目列舉層次 | (內容元件列舉;篇名代碼;位置或其他編碼綱要) | CSI.CTI.MFI;SVN-檢查碼 |
| 0910965218(1996) | (5;HTSFIS;) | 2.2.TX;1-H |

ISMN

- ✿ ISO 10957:1993
- ✿ 主要識別印刷型樂譜出版品
- ✿ 結構

區別碼 – 出版者識別碼 – 作品識別碼 – 檢查碼

 M – 2306 – 7118 – 7

ISRC

✿ ISO 3901:2001

✿ 辨識錄音/錄影等音樂資料本身，而非載體

✿ 結構

| ISRC TW-A53-03-42901 | | | | |
|----------------------|-----------|----------|---------|-----------|
| | 國碼 | 登記者代碼 | 編配年代碼 | 錄音/錄影資料代碼 |
| ISRC | TW | A53 | 03 | 42901 |
| | TW-Taiwan | A53-華納音樂 | 03-2003 | |

ISWC

✿ 主要為識別音樂作品

✿ 結構

– T-345246800-1

✿ 註冊需包含詮釋資料

- 作品的題名；
- 所有作曲者、作者、編曲者的CAE/IPI與角色代碼；
- 作品分類代碼；
- 版本

ISAN

- ✿ ISO 15706:2002
- ✿ 對視聽作品本身做唯一且永久的識別
- ✿ ISAN的結構
 - 16個十六進位的字元
 - ISAN 006A-15FA-002B-C95F-A
- ✿ V-ISAN：識別不同的版本
 - 於2002年6月著手進行
 - ISO Project 20925
 - 做為ISAN標準的補充結構
 - RRRR-RRRR-RRRR-EEEE-X-VVVV-VVVV-Y

ISTC

- ✿ ISO CD 2104
- ✿ 識別的對象是作品本身，而非作品的表現形式
- ✿ ISTC的結構
 - 16個16進位的字元組成
 - 共分為四部分：註冊單位、年代、作品與檢查碼，如
0A9 2002 12B4A105 6
- ✿ 經過作品所有權人授權改編、修改或翻譯的作者，
即視為不同的作品

數位出版品

- ✿ 經過出版行為的結果，此出版行為是指資訊經由電腦編碼、存取、可理解以及取用，其具有對公眾發行的特徵，不論是免費或經銷售而取得。
- ✿ 數位出版品所包含的資料可以是靜態或動態，也就是說，可能會有潛在或實際上的持續變化。

數位資源識別碼(1)

- ✿ 識別碼是指一個清楚的字串用來代表一個實體，也就是被識別的對象，可以同時包含物件與概念。
- ✿ 應用
 - 網路文獻的引用
 - 數位資源的典藏
 - 多重複本
 - 權利管理/取用管理/電子商務
 - 目錄、索引與資源探勘服務
 - 數位資源儲存/取用管理
 - 聯合目錄或整合查詢

數位資源識別碼(2)

✿ 編碼情境與目的

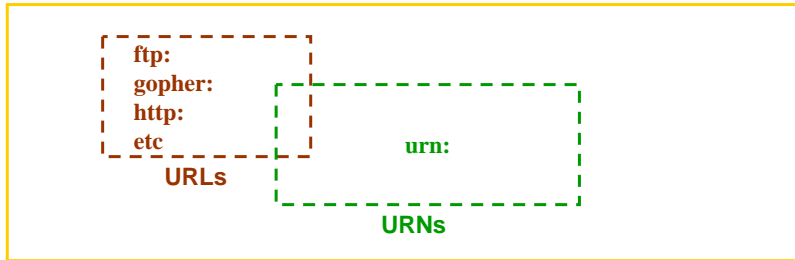
- 全球性網路資源的編碼情境
- 數位出版或電子商務資源的編碼情境
- 數位典藏的編碼情境

數位資源識別碼(3)

✿ 識別碼：命名方式、規格與識別碼系統

- 識別碼如同標記：命名方式
- 識別碼如同基礎建設的規範：讓標記具有行動能力
- 識別碼如同實際系統：在基礎建設環境中執行標記

數位環境資源識別系統



URIs

UR* 家族

參考來源：“Web Naming and Addressing Overview,” in Workstations At Maryland, 1993,
<http://www.wam.umd.edu/~mlhall/keystroke_commands/Addressing.html>(18 January 2004).

URN(1)

- ✿ 依據RFC 1737文件，URN需具有8項功能：
 - 全球性、唯一性、永久性、包容性、相容性、延展性、獨立性、解譯性

✿ 語法結構

- $\langle \text{URN} \rangle ::= \text{“urn:”} \langle \text{NID} \rangle \text{“:”} \langle \text{NSS} \rangle$
- $\langle \text{NID} \rangle$ Namespace Identifier，表示命名方式
- $\langle \text{NSS} \rangle$ Namespace Specific String，由命名單位自行給予唯一的識別碼

URN(2)

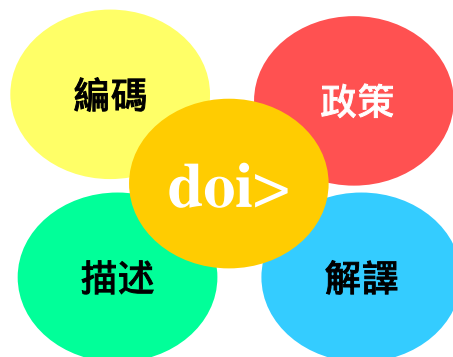
✿ URN的服務功能

- 解譯服務
- 註冊服務
- 管理服務
- 使用者服務
 - 查詢功能
 - 出版與管理
 - 安全控制
- URC服務

DOI

✿ 數位物件識別碼

✿ ANSI/NISO Z39.84- 2000



DOI應用於電子書

🌟 AAP 對電子書編碼的建議

🌟 結構

Prefix Suffix
┌──────────┬──────────┐
10.9999 / 9999999999.99.9999

第一段 第二段

- 方案一：出版社可選擇編配一個ISBN與DOI給電子書的所有版本，並在ONIX中記錄。
- 方案二：電子書的每個版本都編配一個ISBN與DOI，對於可銷售的內容元件亦是如此。

PURL

🌟 PURL的語法結構分為三部分：通訊協定、解譯系統位址與名稱

<http://purl.oclc.org/keith/home>

<http://purl.oclc.org/OCLC/PURL/FAQ>

<http://purl.oclc.org/OCLC/OLUC/32127398/1>

/ | \
通訊協定 解譯系統位址 名稱

Handle System(1)

💡 Handle System

- 分散式永久命名系統
- 主要為數位圖書館應用所發展，由CNRI開發

💡 語法結構

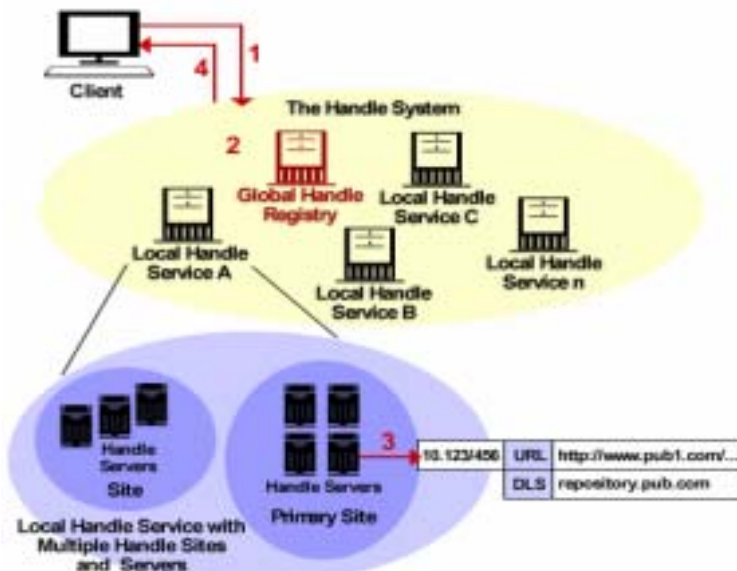
- $\langle \text{Handle} \rangle ::= \langle \text{Handle Naming Authority} \rangle "/" \langle \text{Handle Local Name} \rangle$
- 如DOI

10.1045 / january99-bearman

命名授權
prefix

區域名稱
suffix

Handle System(2)—解譯



數位典藏計畫

✿ LC 的NDLP

✿ 澳洲國家圖書館

- 依據不同藏品類型有不同的命名方式

✿ 國家文化資料庫

✿ 圖書館法技術規範

- 「資料數位化與命名原則規範」

比較與分析(1)--綜合分析

1. 不同的出版品有不同的命名方式
2. 對於出版品衍生物的識別方式不同
3. 各識別碼本身並不作為版權管理之用
4. 多為自願性註冊系統
5. 國際標準識別碼的管理機制，屬分散式的管理架構
6. 國際標準識別碼雖然設計良好，但仍有因其他因素而失敗的情形
7. 國際/國家標準與出版界自行定義
8. 有增加容量的趨勢
9. 增加詮釋資料的著錄

比較與分析(2)--個別分析

- ✿ ISBN與ISSN應用在數位出版品的原則並不相同
- ✿ ISSN與CODEN的比較
 - 管理方式
 - 應用範圍
 - 本身是否有意義，如分類、地區、或記憶的作用
- ✿ SICI、PII與BICI
 - 需已編配ISBN、ISSN
 - PII可與SICI、BICI相結合
 - SICI與BICI：採用文數字；識別不同媒體與內容結構；對於相同媒體但不同格式，並未加以區別

比較與分析(3)--小結

- ✿ 可與解譯系統結合
 - 出版品識別碼具有行動能力
 - 數位資源識別系統可並存各種既有的命名系統
 - 如DOI的suffix、URN的NID與NSS

比較與分析(1)--識別碼的功能需求

1. 唯一性
2. 全球性
3. 一致性/永久性
4. 相容性
5. 包容性
6. 容量足夠
7. 易於使用
8. 解譯性
9. 獨立性
10. 易於自動化系統使用
11. 需連結到詮釋資料
12. 可識別並結合作品的衍生物
13. 支援各種識別層次的基本單元
14. 分散式管理系統
15. 具有註冊與管理機制

比較與分析(2-1)--數位資源識別系統分析

✿ 定址方式不穩定解決方案

1. 採用標準的轉址功能；
2. 採用簡易的解譯系統，如PURL；
3. 採用URN模式，使用代理伺服器；
4. 建置Handle System；
5. 加入DOI系統

比較與分析(2-2)--數位資源識別系統分析

- ✿ 要解決URL不穩定的問題，雖然所花費的資源較多，但採用命名管理系統是較長遠的作法
- ✿ 但命名管理系統目前無法直接使用瀏覽器，解決方案有二：
 - 過渡時期使用代理伺服器
 - 在瀏覽器加上外掛程式
- ✿ 然而不論是採用何種解譯系統與機制，皆需確保系統能長久穩定的運作。

綜合分析(1)--需求分析

- ✿ 數位出版與傳統出版的差異不只在於媒體，包括內容結構與閱讀方式
- ✿ 數位資源識別系統的發展，使得出版品識別碼可具有行動能力
- ✿ 出版社的需求
 - 當出版品從實體延伸到數位時，希望書號中心能夠提供同ISBN識別碼的編配服務
 - 有識別數位出版品內容元件的需求，但對於數位出版品識別系統的參與持保留態度
- ✿ 書號中心
 - 國內目前並無主管單位或註冊中心
 - 研議原生電子書ISBN的編配
 - 先導計畫「數位資源管理中心」

綜合分析(2-1)--現有命名方式彙整

✿數位資源

- 命名方式 + 註冊機關代碼 + 註冊資源代碼
(圖書館法技術規範—資料數位化與命名原則規範)
- 單位代碼 - 物件類別 - 物件代碼 - 多部分之序號
- 使用目的. 附屬檔名
(國家文化資料庫檔案命名原則)

綜合分析(2-2)--現有命名方式彙整

✿電子書

- Prefix/整體作品.第一層基本單元.第二層基本單元.等
(AAP)
- 電子書：ISBN E# C#S#、按需印刷：ISBN D#C#S#
(凱立/凱旋國際資訊)
- /書冊號/登錄號/檔名
(國家圖書館臺灣記憶系統—地方志)

綜合分析(2-3)--現有命名方式彙整

✿ 電子期刊

- 館藏代碼-期刊識別碼-卷期-出版日期-作品影像編號-頁碼-使用目的-原件轉成數位的版本
(澳洲國家圖書館)
- 期刊識別碼/卷期總號/出版日期/頁碼(檔名)編碼
(國家圖書館期刊影像掃描檔案編碼原則)

數位出版品一致性命名原則

- 一、數位出版品命名之應用原則
- 二、數位出版品命名結構與說明
- 三、數位出版品識別系統需求

數位出版品命名之應用原則

- ✿ 唯一性
- ✿ 永久性
- ✿ 包容性
- ✿ 能識別各種類型、不同基本單元與衍生物
- ✿ 檔案名稱應避免超過256個字元
- ✿ 註冊唯一識別碼時，需一併註冊詮釋資料，以使兩者相互連結

數位出版品命名結構與說明(1)

機構代碼-出版品類別代碼-數位物件識別碼-多部份序號-使用目的代碼.副檔名

- 機構代碼：在同一個命名管理系統中，不能重複，一經編配，則永久不再編配給第二個機構
- 出版品類別代碼：代表數位出版品的類型
- 數位物件識別碼：識別該數位出版品，可為單件或內容元件，依各類型出版品有不同的命名方式
- 多部份序號：共四碼，使用流水號編碼
- 使用目的代碼：共兩碼

數位出版品命名結構與說明(2)—— 出版品類別代碼

| 代碼 | 描述 |
|----|--|
| eb | 電子書(ebook) |
| ej | 電子期刊(e journal) |
| td | 數位化學位論文(Electronic Theses and Dissertations) |

數位出版品命名結構與說明(3-1)——數 位物件識別碼

🌟 電子書

- 採用ISBN 對書的定義，因此必須先編配有ISBN
- 單件：ISBN
- 內容元件
 - BICI
 - PII
 - 採AAP的DOI電子書方案

數位出版品命名結構與說明(3-2)—數位物件識別碼

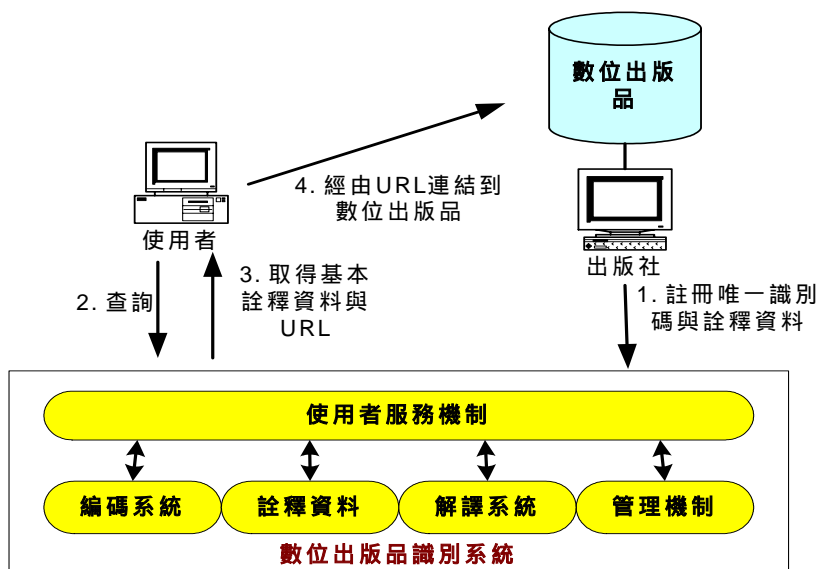
✿ 電子期刊

- 採用ISSN 對期刊的定義，因此必須先編配有ISSN
- 單件：ISSN
- 內容元件
 - SICI
 - PII
 - 國家圖書館期刊影像掃描檔案編碼原則

使用目的代碼表

| 代碼 | 檔案類型 | 用途 |
|----|------|-----------------------|
| tu | 文字 | 永久格式 |
| tp | 文字 | 線上參考/下載用(出版格式) |
| tv | 文字 | 摘要或簡介 |
| iu | 影像 | 永久格式 |
| ip | 影像 | 線上參考/下載用(出版格式) |
| iv | 影像 | 縮圖 |
| su | 音訊 | 永久格式 |
| sl | 音訊 | 區域線上參考/下載用(出版格式) |
| sm | 音訊 | 多重頻寬線上參考/下載用(出版格式) |
| sb | 音訊 | 寬頻線上參考/下載用(出版格式) |
| sn | 音訊 | 窄頻線上參考/下載用 |
| vu | 視訊 | 永久格式 |
| vl | 視訊 | 區域線上參考/下載用(出版格式) |
| vm | 視訊 | 多重頻寬線上參考/下載用(出版格式) |
| vb | 視訊 | 寬頻視訊/音訊線上參考/下載用(出版格式) |
| vn | 視訊 | 窄頻視訊/音訊線上參考/下載用 |
| mu | 多媒體 | 永久格式 |
| mp | 多媒體 | 多媒體線上瀏覽/下載用(出版格式) |
| mv | 多媒體 | 預覽 |

數位出版品識別系統需求



敬請指教

參考資料

1. 歐陽崇榮、歐陽慧，「數位物件識別碼與圖書館相關議題之探討」，[國家圖書館館刊](#)
2. 歐陽慧，台灣地區數位出版品一致性命名原則之研究，未出版碩士論文，2004年

Digital Object Identifier

- ✿ 是一個在數位環境中提供互通識別與交換智慧財產機制的系統，也是一種識別碼。
 - DOI是一個永久不變(persistent)的識別碼
 - DOI是一個具有行動能力(actionable)的識別碼
 - DOI是一個可互通(interoperable)的識別碼
 - 識別適當的層級(appropriate level)
 - 識別複本(copies)與版本(visions)

DOI的起源

- ✿ 為對數位化物件（最主要是出版品）做出唯一的命名與辨識，並藉此保護其智慧財產權，美國出版商協會（AAP）於1994年成立「技術授權委員會」(Enabling Technologies Committee)，進行一種既能保護智慧產權且能保障版權所有者商業利益的系統設計
- ✿ 在1997年法蘭克福圖書博覽會首次亮相，名稱為The Digital Object Identifier，簡稱DOI，正式成為數位化資源命名的一項標準。
- ✿ 由在1998年法蘭克福成立的International DOI Foundation（簡稱IDF）負責政策的制定、技術支援、註冊及繳納規費、維護線上的使用指南等
- ✿ DOI網站: <http://www.doi.org>

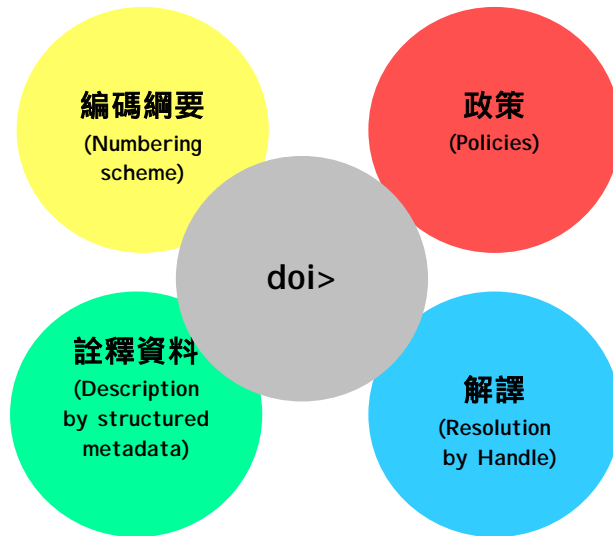
DOI的發展

- ✿ DOI系統發展至今2004年已經有六年歷史了，已給予了千萬個DOI號碼；
- ✿ 參予DOI系統的組織已有國家圖書館、政府文獻機構等的加入；
- ✿ 在語言方面除了英文以外，也陸續有法文、德文、西班牙文、義大利文及韓文等，正逐步的成長中；
- ✿ DOI不僅是號碼，而是包括了編碼、解譯、詮釋資料及與政策結合的系統；
- ✿ DOI系統工作除了發展外尚涵蓋了如何有效的全盤佈置導入及應用此系統。

DOI的特點

- ✿ DOI識別碼本身，即帶有大量的資訊與連結，它可以讓使用者直接由此連結到出版商的資料庫、文獻、摘要甚至是全文，使得識別碼可以直接指引到此份數為出版品的本身，而不再只是出版品的識別碼而已。
- ✿ DOI特點
 - 永久不變的識別碼碼
 - 可運作的識別碼碼
 - 可互通的識別碼碼
 - 可識別出示當的層級

DOI系統的組成



編碼(Numbering)

<DOI>=<DIR>.<REG>/<DSS>

<DIR>=10

<REG>Registrant's Code

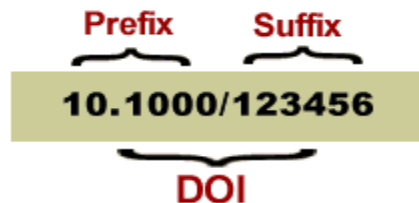
<DSS>DOI Suffix String

Character set is Unicode 2.0

不區分大小寫

<DSS> 的起始字元不能為*/

<REG>碼是由註冊中心發給各註冊單位

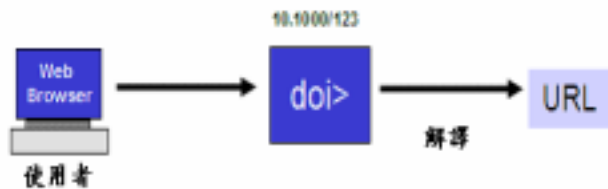


描述(Description)

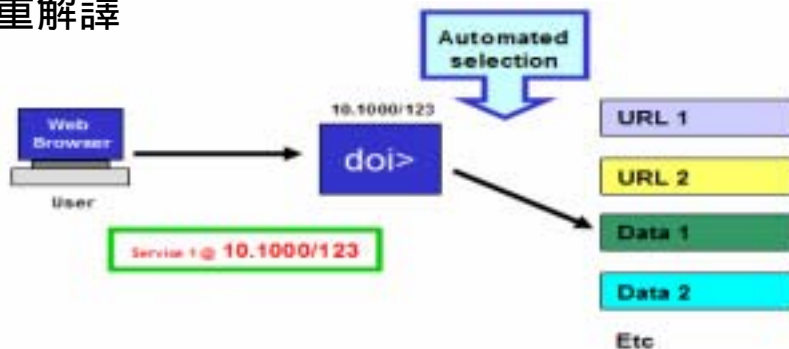
- 💡 詮釋資料
- 💡 可行動的

解譯(Resolution)

💡 單一解譯



💡 多重解譯



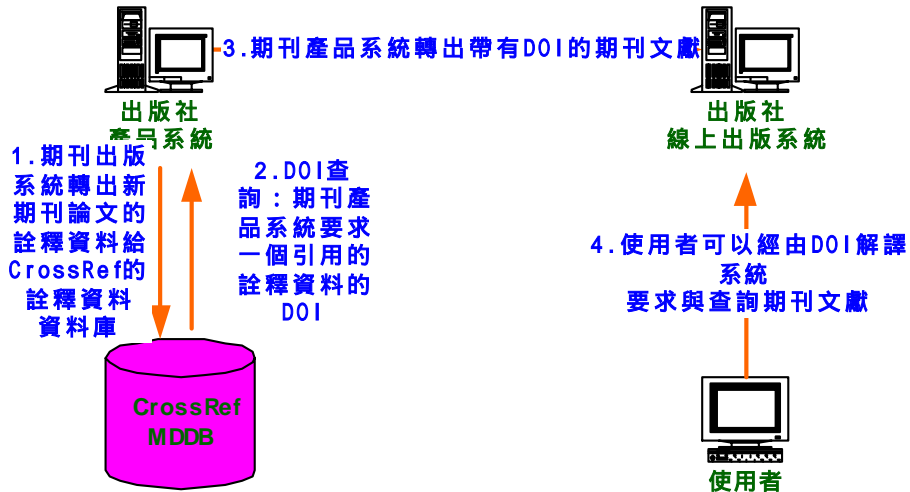
政策(Policy)

☀️ 確保運作

註冊中心(Registration Agencies)

☀️ 註冊中心主要的任務是對註冊者提供服務：分配DOI prefixes、註冊DOI、以及提供必要的設備讓註冊者可以登錄和維護詮釋資料與狀態資料。目前DOI的註冊中心共有九個，CrossRef、Content Directions、Enpia Systems、LON Inc.、Copyright Agency Ltd.、Nielsen BookData、R. R. Bowker、TSO (The Stationery Office)與mEDRA。

CrossRef/DOI運作機制



DOI與E-Commerce

- ✿ DOI可以為一個開放性應用建立一個框架
- ✿ DOI意圖使內容管理更為容易
- ✿ 資訊的管理意味著交易的簡化
- ✿ 交易發生在三件事交互作用的結果：物件（交易的對象）、人（進行交易的當事人）、以及協議（交易的條款）
- ✿ 需要一個共同的框架與詞彙以達到互通性
- ✿ IDF與INDECS計畫合作，發展出支援E-Commerce的共同詮釋資料框架，這個框架會呼叫一個唯一的識別（包含智慧財產、當事人和協議）

