

行政院及所屬各機關出國報告

(出國類別：研究)

檢驗自動化網路作業系統

出國人服務機關：國立台灣大學醫學院附設醫院資訊室

職 稱：三等程式設計師

姓 名：鄭伯堦

出國地區：美國

出國期間：八十九年五月三日至八十九年十一月二日

報告日期：九十年一月七日

(出國類別：研究)

1. 內容摘要

職於八十九年五月三日至八十九年十一月二日奉派至美國加州大學河濱分校(University of California, Riverside, UCR)，主要修習內容包括 HL7 醫療資訊交換標準，資料倉儲與資料採礦的學習與研究，而上述三種修習的內容均以檢驗自動化網路系統作為學習的主軸，也就是環繞著這個學習主軸並充份應用上三種修習的內容。學習活動可以概略區分為以下兩個時段：前三個月以資料倉儲(Data Warehousing)與資料採礦(Data Mining)在檢驗自動化網路作業系統上的應用為主，後三個月以 HL7 (Health Level 7)醫療資訊交換標準在電子病歷(EPR, Electronic Patient Record)上的應用為主，當然電子病歷就包含了檢驗自動化網路系統在內。

就資料倉儲與資料採礦在檢驗自動化網路作業系統上的應用部份，可以歸納為以下幾個重點：

- 跟隨加州大學河濱分校資訊科學與工程研究所(CSE, Department of Computer Science and Engineering)之 Dimitrios Gunopulos 教授學習 Non-visual 資料採礦技巧。
- 跟隨加州大學河濱分校資訊科學與工程研究所之 Vassilis J. Tsotras 教授學習資料倉儲。
- 旁聽有關 UCR/CSE 每週四下午之資料採礦研討課程。
- 不定期參加 UCR 校內資料倉儲與資料採礦研討會。
- 收集並研讀資料倉儲與資料採的書籍及資料。
- 輔以使用該所資料庫實驗室(DB Lab)設備。
- 資料倉儲主要以系統建置與未來趨勢為學習重點。
- 資料採礦與該校老師與研究生，以 Clustering 方法論為主，套用猛暴性肝炎的醫療應用，企圖找出目前 Clustering 方法論中尚無法實際應用的 Time Series 問題，直到回國為止，因醫療相關病歷報告資料取得有實

際上的困難，該篇研究論文尚未發表，但該校相關研究人員仍持續研究中。

其中就 HL7 醫療資訊交換標準部份，可以歸納為以下幾個重點：

- 主要為自修 HL7 標準規範內容。
- 五月八日隨台灣醫療參訪團參觀(PAMF, Palo Alto Medical Foundation)，了解北加州醫療體系與系統架構。
- 五月九日至十二日隨台灣醫療訪問團參加加州舊金山之 TEPR 2000 年會，取得最新建置電子病歷相關資料。
- 五月十二日隨台灣醫療訪問團參觀加州史坦福大學附設醫院(Hospital of Standard University)之電子病歷系統。
- 五月十二日隨台灣醫療訪問團參觀加大學舊金山分校附設醫院(UCSFH, Hospital of University of California, San Francisco)之電子病歷系統。
- 九月十一日至十五日參加密蘇里州聖路易斯市 HL7 組織第十四屆年會，取得醫療系統資訊交換相關資料。
- 九月十三日於 HL7 組織年會上成為亞洲第一位通過 HL7 組織之醫療資訊交換標準的官方認證者。
- 九月十九日參訪加州大學洛杉磯分校醫學中心(UCLA/MC, Medical Center, University of California, Los Angeles)之兒童醫院(Mattel Children's Hospital)的資訊服務(Computer Service)單位，實際學習與研討該院 HL7 相關建置技術與經驗。

UCR 位於美國加州(California)的河濱縣(Riverside County)，是加州大學之分校之一，該校資訊科學與工程研究所(CSE)主要著重於資料倉儲與資料採礦領域，並設有資料庫研究室。大體上 UCR 非常歡迎訪問學者，職在 UCR 研究期間，前後就有至少七位來自美國以外(台灣、日本、加拿大、大陸等)的訪問學者在該系所從事短期的研究活動。

目前 HL7 醫療資訊交換標準的版本是八十九年十月二十三日通過美國 ANSI 的 2.4 版，其主要的內容分別是

- 簡介 (Introduction)

(出國類別：研究)

- 控制 (Control)
- 病友管理 (Patient Administration)
- 醫令登錄 (Order Entry)
- 查詢 (Query)
- 財務管理 (Financial Management)
- 檢驗報告 (Observation Reporting)
- 參考主檔 (Master Files)
- 病歷資訊管理[文件管理] (Medical Records / Information Management [Document Management])
- 排程 (Scheduling)
- 轉診 (Patient Referral)
- 看護 (Patient Care)

職在研究地點參加 HL7 國際組織認證前所研習的標準內容是以 2.3.1 版本為主，而職在 HL7 國際組織年會中所認證的標準內容為 2.3 版，也是該組織所提供之認證的最新版本。在電子病歷日趨重要的今日，美國以外的其它國家也越來越重視，與會人員也逐年增加。而台灣與大陸同時在八十九年六月加入該國際組織，南韓則在八十九年七月加入，目前的會員國計有十三個。

稍後參訪加州大學洛杉磯分校醫學院之之兒童醫院(Mattel Children's Hospital for Medical Center of University of California, Los Angeles)的電腦服務 (Computer Service)單位，與該單位之主管及技術人員詢問該院之醫療資訊交換知識與經驗，並實際學習該院 HL7 相關建置技術與系統。其中該院的接待技術人員包括 Timothy Carlson (HL7 程式設計師)與 Spencer Sun (HL7 系統管理員)。該醫院位於美國南加州洛杉磯市，共有 500 床病床，該電腦服務單位除支援該院系統外，尚支援另外兩家小醫院，包括 Santa Monica County Hospital 與 UCLA/MC Cancer Hospital。該醫院有不錯的 HL7 電子資訊交換系統，其 HL7 Interface Engine 為 STC 廠商的產品，產品名稱為 DataGate。此外，該院與本院之系統架構非常雷同，同樣擁有 IBM 主機与其它開放式架構之醫療系統，且 IBM 主機與本院一樣使用 ADS+之程式開發環境，因此該院的 HL7 建置經驗對於本院未來建置 HL7 資訊交換系統將有可觀的助益，LDS 的電腦系統是該院資訊部門與 University of

Utah 中的醫療資訊系(Medical Informatics)研究生合作，經過三十年的開發，逐漸建立起來的。PCMC 則是部份購買 3 M 所開發出來的軟體，此軟體協助醫院建立起醫療品管的資料庫，以供進一步分析資料用。

最後衷心感謝台大醫院主管與同事的支持與鼓勵，以及美國主管的協助，讓職能順利在美學習與受訓。

(出國類別：研究)

2. 目次

1. 內容摘要	2
2. 目次	6
3. 目的	11
4. 過程	12
5. 心得報告	14
5.1. 醫療資訊交換標準 HL7 心得	14
5.1.1. HL7 組織	14
5.1.2. HL7 組織歷史	15
5.1.3. HL7 國際組織成員	15
5.1.4. HL7 Technical Committees	16
5.1.5. HL7 Special Interest Groups	16
5.1.6. HL7 版本	17
5.1.6.1. 優點	18
5.1.6.2. 差異性	18
5.1.6.3. 為何需要 HL7 v3.0 ???	18
5.1.6.4. V3.0 vs. V2.x	19
5.1.6.5. V3.0 方法論	19
5.1.6.6. V3.0 模型：抽象層	21
5.1.6.7. V3.0 模型：實體層	21
5.1.6.8. V3.0 的挑戰	21
5.1.6.9. HL7 v3.0 時程	22
5.1.7. HL7 v2.3.1 標準心得	24

5.1.7.1.	觀念部分	24
5.1.7.2.	資料類型	29
5.1.8.	控制(Control)	43
5.1.8.1.	Escape	43
5.1.8.2.	訊息建置規則	45
5.1.8.3.	符號(Notations)	45
5.1.8.4.	Acknowledgement Mode	45
5.1.8.5.	Enhanced Acknowledgement Mode	48
5.1.8.6.	Deferred Processing Mode	49
5.1.8.7.	ACK , QAK vs. MCF	50
5.1.8.8.	Message Data Format	51
5.1.8.9.	Message : ADD 使用時機	51
5.1.9.	查詢(Query)	52
5.1.9.1.	型式	52
5.1.9.2.	Query : Response	52
5.1.9.3.	Original Mode Query	53
5.1.9.4.	Enhanced Mode Query	54
5.1.9.5.	Original Mode Query Segments	54
5.1.9.6.	Enhanced Mode Query Segments	57
5.1.9.7.	Original Mode Query / Immediate Response	58
5.1.9.8.	Query Error	58
5.1.10.	訊息(Message)	59
5.1.10.1.	Message : MSH-20	59
5.1.10.2.	Representation Name Type	59
5.1.10.3.	Z Locally-Defined	59
5.1.10.4.	Data Type Length	59
5.1.10.5.	Message : CD (Channel Definition)	60
5.1.10.6.	RI (Repeat Interval)	61
5.1.10.7.	TQ (Timing / Quantity)	61
5.1.11.	HL7 之會議與認證	62
5.1.11.1.	HL7 國際會議	62

(出國類別：研究)

5.1.11.2.	HL7 認證經驗	64
5.1.12.	HL7 之 Interface Engine	64
5.1.12.1.	如何應用 HL7 ???	64
5.1.12.2.	深入了解 IE	67
5.1.13.	HL7 系統觀摩	69
5.1.13.1.	Palo Alto Medical Foundation	69
5.1.13.2.	Stanford Hospital	71
5.1.13.3.	UCSF Hospital	73
5.1.13.4.	University Hospital of Columbia & Cornell	74
5.1.13.5.	University of Virginia Hospital	75
5.1.13.6.	UCLA Mattel Child Hospital	78
5.1.13.7.	有關 HL7 的建議	81
5.1.13.8.	我對 HL7 的思考方向	82
5.1.13.9.	資訊室對 HL7 的因應計劃	82
5.2.	資料倉儲與資料採礦心得	86
5.2.1.	進修資訊	86
5.2.2.	資料倉儲(Data Warehousing)	87
5.2.2.1.	DW 定義	87
5.2.2.2.	為何需要 DW ???	88
5.2.2.3.	如何成功建置 DW ???	92
5.2.2.4.	DW 系統架構	97
5.2.2.5.	Data Movement Models	102
5.2.2.6.	架構設計考量	102
5.2.2.7.	建置過程	103
5.2.2.8.	SMP vs. MPP	103
5.2.2.9.	Data Modeling	105
5.2.2.10.	Star Schema	105
5.2.2.11.	Star Schema vs. Fully Normalized Relational	108
5.2.2.12.	Snowflake Schema	109
5.2.2.13.	Constellation Schema	111
5.2.2.14.	Snowstorm Schema	112

5.2.2.15.	Aggregate	112
5.2.2.16.	Dimensional Data Modeling (DDM)	114
5.2.2.17.	Logical Data Model vs. DW Data Model	115
5.2.2.18.	Data Granularity vs. Transaction Granularity	116
5.2.2.19.	Multi-dimensional	116
5.2.2.20.	Multi-dimensional vs. Multi-relational	116
5.2.2.21.	Multi-dimensional Analysis	117
5.2.2.22.	資料容量	120
5.2.2.23.	建立第一個 DW	120
5.2.2.24.	DW 工具	122
5.2.2.25.	DW 資料轉換	123
5.2.2.26.	Homonyms vs. Synonyms	123
5.2.2.27.	DW 的延伸應用	124
5.2.2.28.	OLTP vs. DSS	124
5.2.2.29.	OLAP vs. DM	125
5.2.2.30.	DW 的未來	125
5.2.2.31.	決策的進化論	126
5.2.2.32.	Metadata	127
5.2.2.33.	相關參考資源	129
5.2.2.34.	有關 DW 的建議	131
5.2.3.	資料採礦(Data Mining)	131
5.2.3.1.	DM 定義	131
5.2.3.2.	DM 應用範圍	132
5.2.3.3.	DM 之工作	132
5.2.3.4.	DM 方法論(Methodology)	134
5.2.3.5.	DM 之技術	137
5.2.3.6.	DM vs. OLAP	142
5.2.3.7.	DM 的未來	143
5.2.3.8.	DM 相關資源	144
5.2.3.9.	我的 DM 研究題目	149
5.2.3.10.	有關 DM 的建議	154

(出國類別：研究)

6. 建議..... 155

6.1. 資訊部門的挑戰..... 155

3. 目的

- 觀摩與學習電子病歷的實際建置經驗。
- 通過國際醫療資訊交換標準 HL7 (Health Level 7) 官方認證。
- 觀摩與學習當地大型醫院資訊部門，實地瞭解國際醫療資訊交換標準在美國地區的實際應用。
- 思考本院既有醫療系統如何應用國際醫療資訊交換標準-HL7，進而提昇本院醫療資訊系統國際化地位，同時加速其標準化的腳步。
- 學習並模擬本院未來應用資料倉儲與資料採礦的資訊系統面與作業面的需求，並思考如何將其導入本院現行的檢驗自動化網路作業系統。
- 思考檢驗自動化網路作業系統、電子病歷、國際醫療資訊交換標準 HL7、資料倉儲與資料採礦四種層面間的互動關係、影響與應用方式。

(出國類別：研究)

4. 過程

常駐於加州大學河濱分校(University of California, Riverside)資訊科學與工程研究所(Department of Computer Science and Engineering)，主要修習內容包括 HL7 醫療資料交換標準、資料倉儲與資料採礦。

其中就 HL7 醫療資料交換標準部份，主要為自修 HL7 內容，參訪會議與醫院系統，包括

- 五月八日隨台灣醫療參訪團參觀 PAMF (Palo Alto Medical Foundation)，聽取該地區醫療組織相關資料。
- 五月九日至五月十二日隨台灣醫療參訪團參觀加州史丹福大學附設醫院(Stanford University Hospital)，見聞最新電子病歷相關資料。
- 五月十二日隨台灣醫療參訪團參訪加州大學舊金山分校附設醫院(University of California, San Francisco Hospital)，見聞最新電子病歷相關資料。
- 六月一日以個人身份加入 HL7 醫療資訊交換標準組織。
- 九月十一日至九月十五日參加密蘇里州聖路易斯市(St. Louis, Missouri) HL7 組織第十四屆年會，取得醫療資訊交換相關資料
- 九月十三日於 HL7 年會上成為亞洲第一位通過 HL7 官方組織之醫療資訊交換標準認證者。
- 九月十九日參訪加州大學洛杉磯分校醫學院之兒童醫院(Mattel Children's Hospital for Medical Center of University of California, Los Angeles)的電腦服務(Computer Service)單位，實際學習該院 HL7 相關建置技術與經驗。

就資料倉儲與資料採礦部份，主要跟隨加州大學河濱分校資訊科學與工程研究所教授(Prof. Dimitrios Gunopulos)學習，並輔以使用該所資料庫實驗室設備，內容包括

- 資料倉儲主要以系統建置與未來趨勢為學習重點。
- 思考資料倉儲在檢驗自動化網路作業系統中的應用與切入方式。
- 資料採礦部份則與該校老師與研究生，以 Clustering 方法論為主，套用猛暴性肝炎的醫療應用，企圖找出目前 Clustering 方法論中尚無法實際應用在醫療領域內的 Time Series 問題，直到回國為止，因醫療相關病歷報告資料取得有實際上的困難，該篇研究論文尚未發表，但該校相關研究人員仍持續研究中。

5. 心得報告

因為此次研究的目的與檢驗自動化網路作業系統在電子病歷(EPR, Electronic Patient Record)上的應用都有關聯，因此將心得報告分成三大部份，包括醫療資訊交換標準 HL7 心得、資料倉儲與資料採礦心得。

5.1. 醫療資訊交換標準 HL7 心得

醫療資訊交換標準 HL7 (Health Level 7)是一國際性的組織，職司制定國際性的醫療資訊交換標準，標準名稱也叫做 HL7，是目前國內電子病歷建置的必備基石。而在建置電子病歷之前，必須熟悉與建立一套目前與未來共通性高且標準一致的醫療資訊交換標準，再以此標準為電子病歷的基礎，從而達到電子病歷資訊交換、共用、與分享的目的。

行政院衛生署於年前公佈明訂未來國內醫療資訊交換的標準將以 HL7 v2.3 為依據，此外國內也同時有許多醫療資訊單位與研究人員針對此標準加以研究，甚至加以應用。以下是本人在研究地點針對此標準的心得內容。

5.1.1. HL7 組織

1994/06，HL7 被授權為 ANSI 標準開發組織(SDO, Standards Developing Organization)，範圍包括獨立醫療系統下的電子資料交換標準，醫療系統包括 clinical, financial, administrative, laboratory, pharmacy, pathology, radiology 等。

5.1.2. HL7 組織歷史

目前已有六個美國國家標準

- 1996/02，ANSI 通過 HL7 v2.2
- 1997/05，ANSI 通過 HL7 v2.3
- 1999/05，ANSI 通過 HL7 v2.3.1
- 1999/07，ANSI 通過
 - ✓ Arden Syntax for Medical Logic Systems v2.0
 - ✓ Clinical Context Management Specification v1.0

5.1.3. HL7 國際組織成員

目前組織成員至少包括以下次組織團體

- HL7 Australia
- HL7 Canada
- HL7 Finland
- HL7 Germany
- HL7 India
- HL7 Japan
- HL7 The Netherlands
- HL7 New Zealand
- HL7 South Africa
- HL7 United Kingdom
- HL7 Taiwan，2000/06 入會
- HL7 China，2000/06 入會
- HL7 Korea，2000/07 入會

(出國類別：研究)

5.1.4. HL7 Technical Committees

目前其下轄的技術委員會包括

- Architectural Review Board
- Clinical Context Object Workgroup (CCOW)
- Clinical Decision Support and Arden Syntax
- Control Query
- Education / Implementation
- International
- Medical Records/Information Management
- Modeling and Methodology
- Order / Observations
- Patient Administration / Financial Management
- Patient Care
- Scheduling and Logistics
- Vocabulary
- XML

5.1.5. HL7 Special Interest Groups

其轄下之特殊興趣群組包括

- Accountability, Quality and Performance
- Blood Bank
- Claims Attachments
- Community Based Health
- Component Based Messaging
- Conformance
- Data Warehousing
- Government Projects
- Image Management

- Laboratory Automation
- MPI Mediation
- Patient/Provider Messaging
- Personnel Management
- Secure Transactions

5.1.6. HL7 版本

是指 Application Protocol for Electronic Exchange in Healthcare Environments 的標準版本，其所有版本的公佈時程如下

- V1.0：1987 年公佈，並未被實際引用
- V2.0：1988 年公佈，並未被實際引用
 - ✓ V2.1：1990 年公佈，開始備廣泛引用
 - ✓ V2.2：1994 年公佈
 - ✓ V2.3：1997 年公佈且得到 ANSI 認可
 - ◆ V2.3.1：1999 年公佈且得到 ANSI 認可
 - ✓ V2.4：2000/10 年公佈且得到 ANSI 認可
- V3.0：討論中，預計 2002 年公佈

上課講師告知所有學員「不要等版本，因為一個主要版本約十年。」HL7 v2.x 最廣泛被引用於大型系統，而中小型設備，約 90% 以上，其包含更多的功能，包括

- 醫療：laboratory，pharmacy，radiology，dietary，diagnostic services，patient care，public health
- 醫療管理：patient registration，admission，patient accounts，medical document life cycle，master file maintenance，HIPAA attachments
- 引用 ASC-X12 的訊息格式標準

(出國類別：研究)

5.1.6.1. 優點

- 內含具有更多功能
- 高度適用於異質系統環境
- 系統功能的異動
- 醫療字彙獨立性
- 參考較少的共用技術

5.1.6.2. 差異性

- V2.2：新增 ADT、Registration、Orders、Results、Patient Financial、Master Files。
- V2.3：新增 non-ASCII Character Sets、Query Language Support、Immunization Reporting、Adverse Drug Reactions、Clinical Trials、Scheduling、Referrals、Medical Documents、Problems & Goals。
- V2.4：預計新增 Query paradigm、conformance-based queries、Further internationalization support(包括名字，住址，組織名稱，文字)、Personnel management transactions、Clinical laboratory automation transactions、Network application management transactions、Message protocol supports solicited messages (如醫令廣播)、unsolicited messages (如檢驗結果查詢)、支援 XML 編碼規則，產生可讀且不定長度具有分隔符號的 ASCII 訊息字串。

5.1.6.3. 為何需要 HL7 v3.0 ???

因期具有全新的標準類型：Messaging Standards、邏輯上與 HL7 v2.x 一致、在語意上做出重大的改良(如 Document Standards)、引用 SGML/XML 文件(如

Component Standards、Knowledge Representation)、用全新的處理(Process)或方法論(Methodology)來發展此標準 V3.0 的目標包括 Plug-and-play、提供框架(包括事件，訊息，資料元件)、改善規格的清楚度與精細度、改良標準可以快速適應產業的變化。

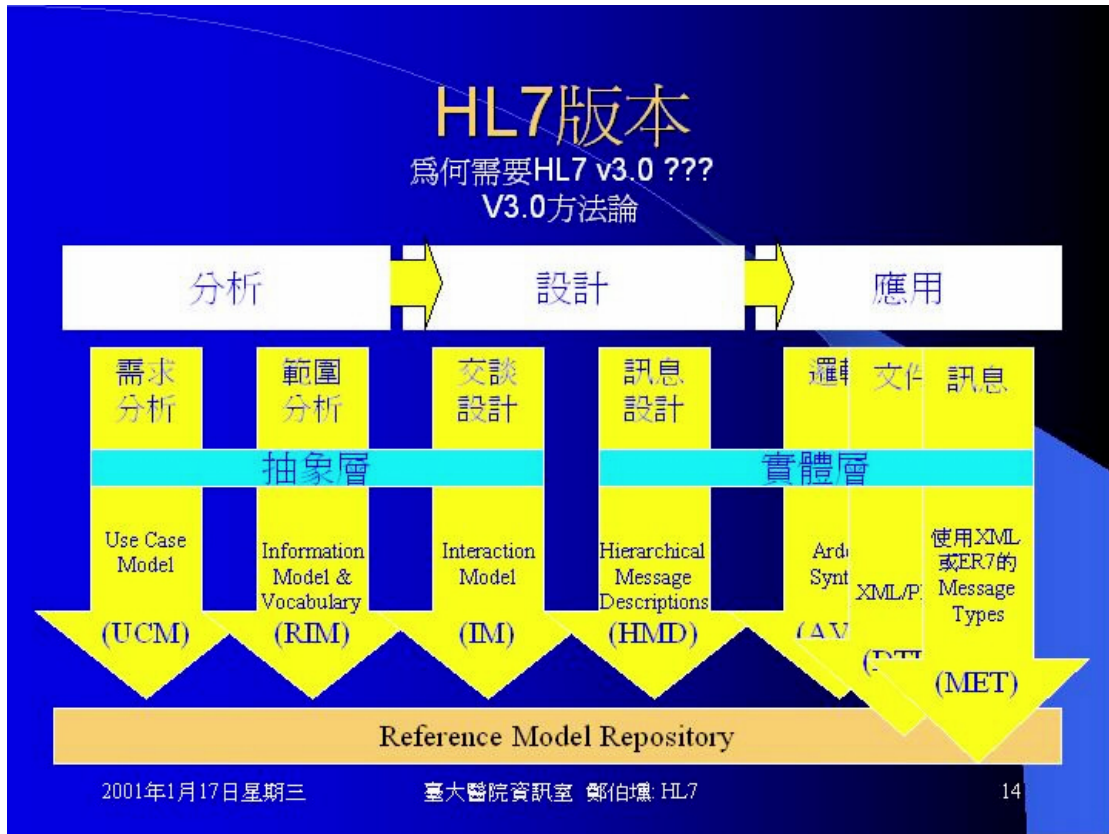
5.1.6.4. V3.0 vs. V2.x

- V2.x
 - ✓ 將建議的資訊內容片段化(Segmentation)。
 - ✓ 事件(Event)將指明應用的行為。
 - ✓ 使用例子說明標準之內容(但是從未正式的文件化)。
- V3.0
 - ✓ 尋求正式的應用物件分析方法，將以應用模型到 HL7 作業的方式來達成。
 - ✓ 改善標準內部的一致性。
 - ✓ 提供語意的定義。
 - ✓ 使未來標準的架構能夠順利套用。
 - ✓ 是一種進化(Evolution)，不是革命(Revolution)。

5.1.6.5. V3.0 方法論

- 引用模型與共用的倉庫(Repository)，使標準之內容具有更好的控制與一致性。
- 在 Information Model 中使用 State Diagram 定義 Trigger Events，內含更清楚，更仔細，更細微的規格說明。
- 提供更多訊息以簡化選擇項目。
- 內含標準字彙，專用術語，代碼集(Code Sets)。

(出國類別：研究)



由上圖可知 V3.0 方法論之大略步驟如下所示：

- 針對醫療資料定義一個適合的 RIM。
- 收集專用術語(Terminology)及資料型別(Data Type)，以描述 RIM 中的屬性。
- 將 RIM，字彙(Vocabulary)，型別(Type)應用到訊息(Message)，病友資料錄(Patient Record)之 DTD，醫療邏輯模組(Medical Logic Module)，元件規格(Component Specification)上。
- 對每一個應用程式，畫出 RIM 以建立出抽象訊息結構(Abstract Message Structure)，也就是所謂的 HMD。
- 為每一個建置技術(Implementation Technology)，HL7 將定義一個 ITS 以便將 HMD 對應到技術上。
- 當訊息(Message)被傳送後，HMD 會被用來排列(Marshal)資料，而 ITS 會被用來格式化通訊用的資料。

5.1.6.6. V3.0 模型：抽象層

- Use Case Model (UCM)：說明工作或程序的階層關係。
- Reference Information Model (RIM)：說明 Classes, Relationships, States, Life Cycles。
- Vocabulary：說明所有的 HL7 內建(Coded)屬性，將 HL7 結構化表格代碼化(Enumerated Codes)Interaction Model (IM)：說明 Trigger Events, 抽象訊息, Application Profiles。

5.1.6.7. V3.0 模型：實體層

- Message Design Model：使用 Hierarchical Message Description (HMD)來建立模型。
- Implementation Specification：使用 Message Element Type (MET) 語言來建立, Implementation Technology Specification (ITS)。

5.1.6.8. V3.0 的挑戰

- 語意的一致性：從 1998 開始，RIM 已經開發 3 年，且已經被文件化與公開化，使用目前世界上最好的專用術語所制定，想要將 HL7 產品依附到單一的模型與範圍內。
- 最低的不明確性：每個觀念從清楚的定義開始說明，方法論裡已經清楚說明規則與環境，方法論需要一致性的應用程式來宣告執行時的訊息與欄位，方法論便於開發訊息的其他變形。
- 製鞋匠小孩的鞋子：方法論是被預期處在一個單一且共享的倉庫 (Repository)上，同時擁有所有 HL7 模型的定義，而倉庫主要是管理 RIM 的版本，字彙的一致性，公佈模型與範圍，被使用在技術委員會裡的訊

(出國類別：研究)

息設計工具，定義且票決被提名的標準內容的一種基礎。

- 引用其他技術的能力：例如，HL7 要適用於 XML 且從中取得到優點，優點包括提供 OTS (Off-the-shelf) Processors 免費軟體與商用 Parsers：
OTS Parsers 可以確認 Message/Document 的內容。
 - ✓ 轉換與瀏覽工具：XSL Processor + Web Browser = Intelligent，
Message/Document，XSLT 可以當作為 Interface Engine 規則語言，
提供個人訓練課程。
 - ✓ 以範圍而言：HL7 語法 << XML 語法。
 - ✓ 底限是最便宜，可快速發展，易維護。
- 面對技術改變的能力：HL7 標準的方法論之抽象層可以允許以新的技術重新建置(Re-implementation)標準，而不必重新定義(Re-define)HL7 標準內容，此外以工具(Tools)與倉庫(Repository)為基礎的標準允許 HL7 引用最新的技術。
- 面對市場改變的能力：RIM 開發工作，應用到 RIM 的字彙範圍，因為有強大的語意基礎，所以可以成為 Ad hoc 之標準開發者。

5.1.6.9. HL7 v3.0 時程

- 1996 年
 - ✓ 1st TSC (Technical Steering Committee 退回 v3.0 methodology。
 - ✓ 2nd TSC 退回 v3.0 methodology。
 - ✓ 介紹模型給各 TCs。
 - ✓ 開始開發 RIM。
 - ✓ 1st 對會員教授 v3.0。
 - ✓ 建立 SIG 字彙。
 - ✓ HIPAA 成為正式法律。
- 1997 年
 - ✓ RIM v0。8。
 - ✓ 1st 介紹 MDF (Message Development Framework)。

- ✓ TSC 審查 v3.0 的一致性。
- ✓ 通過 v2.3 (最後的 2.x???) 。
- ✓ 1st RIM 協調會議。
- ✓ 修訂 HMD 的方法論。
- 1998 年
 - ✓ 發表 MDF v3.0.1 。
 - ✓ 應用 Rational Rose 模型。
 - ✓ 開始 HIPAA Claims Attachment 。
 - ✓ 1st 提出 v3.0 Data Type 。
 - ✓ 開始加入 v30 XML 建置技術。
 - ✓ 應用 RoseTree 。
 - ✓ 開始討論 USAM 。
- 1999 年
 - ✓ 通過 v2.3.1 (最後的 2.x???) 。
 - ✓ 審查 v3.0 Data Type 。
 - ✓ 1st v3.0 在 HIMSS 展示。
 - ✓ 通過 HIPAA Implementation Guide 。
 - ✓ MDF 加入 R-MIN 符號。
 - ✓ MDF 加入字彙。
 - ✓ 開始 v3.0 進階訓練，RoseTree 建立完整的 HMD 。
 - ✓ RIM 加入 USAM 。
 - ✓ 發表 MDF-99 。
 - ✓ 1st ISO TSC215 MDF 草案。
- 2000 年
 - ✓ 開始 LHSR 專案。
 - ✓ 票決 v3.0 Data Type 。
 - ✓ 1st 字彙協調會。
 - ✓ 開始 v3.0 Acceleration 專案。
 - ✓ RIM 加入 LHSR 內容。

(出國類別：研究)

- ✓ V3.0 開始世界化。
- ✓ 完成 v2。4 票決 (最後的 2.x???)。
- 2001 年預計工作項目
 - ✓ Balloted package。
 - ✓ Support documentation。
 - ✓ Education prepared。
 - ✓ Membership ballot package。
 - ✓ Support documentation ready for final package。
 - ✓ Final ballot on Data Type。
 - ✓ Final ballot on XML ITS。
- 2002 年預計工作項目
 - ✓ Message Design ready for committee ballot。
 - ✓ Documentation support to committee vote。
 - ✓ Committee ballot on XML ITS and Data Type。
 - ✓ Interaction design。
 - ✓ Preliminary message design : patterns & common types。
 - ✓ CMET。
 - ✓ XML ITS and Data Types ready for committee ballot。

5.1.7. HL7 v2.3.1 標準心得

5.1.7.1. 觀念部分

- Trigger Event vs. Unsolicited Update
 - ✓ Trigger Event : The real world event that initiates an exchange of message。
 - ✓ Query + Response。

- ✓ Unsolicited Update + Acknowledgement。

HL7 Ch2: Query & Control

Trigger Event vs. Unsolicited Update

類型	Trigger Event	傳送端 (Sender/Placer)	接收端 (Responder/Filler)
Unsolicited Update	病友住院	住院通知	Acknowledgement (accept/application)
	緊急檢查報告	顯示緊急檢查報告	
Query	醫師查詢檢驗結果	顯示查詢結果	回應並顯示報告或加以列印
		以筆數回應查詢結果	以筆數為基礎回應到相對應的工作站

2001年1月17日星期三 臺大醫院資訊室 鄭伯璦 HL7 5

- Acknowledgement
 - ✓ Original Mode。
 - ✓ Enhancement Mode：包括 Accept Acknowledgement，Application Acknowledgement。
- Query：共有 5 種
 - ✓ QRD + QRF，Original Mode Query Definition/Filter。
 - ✓ EQL，Embedded Query Language。
 - ✓ VTQ，Virtual Table Query。
 - ✓ SPQ，Stored Procedure Query。
 - ✓ ERQ，Event Replay Query。
- Communication Environment：2+ 應用程式並未緊密的整合在同一機器上，共有以下四種類型
 - ✓ 具基本傳輸功能。
 - ✓ RS-232C --> HL7/LLP(Low Level Protocol)。
 - ✓ 有 Robust 網路，但未達高標準：如 TCP/IP，SNA，DECNet。

(出國類別：研究)

- ✓ ISO 與專屬網路：如 SNA/LU6.2, NFS。
- HL7 假設的網路環境為
 - ✓ Error Free Transmission。
 - ✓ Character Conversion。
 - ✓ Unlimited Message Length。
- Message：The atomic unit of data transferred b'n systems。
- Message Type：Define message purpose。
- Message Type vs. Trigger Event，其關係圖如下，

HL7 Ch2: Query & Control

Message Type vs. Trigger Event

- 一對多

```

graph TD
    MT[Message Type] ---|1| TE[Trigger Event]
            
```

- Z打頭的Message Type與Trigger Event是保留給Locally-defined Message使用的

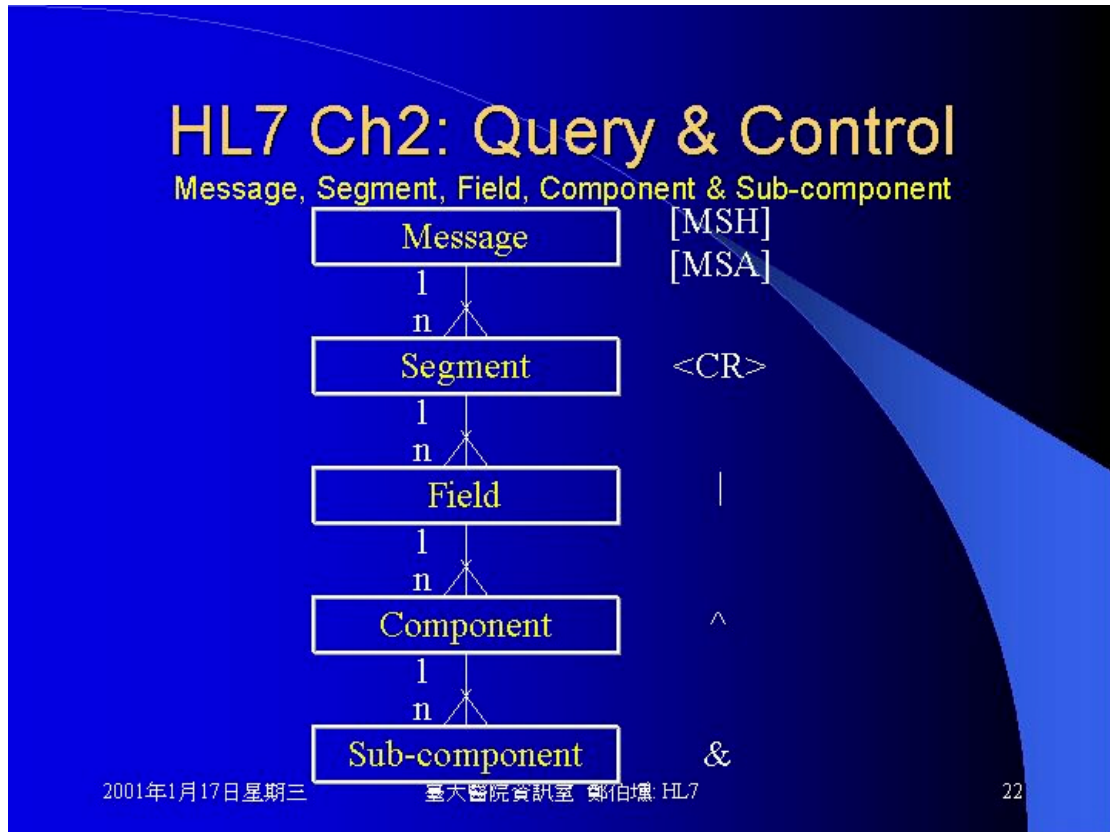
2001年1月17日星期三 臺大醫院資訊室 鄭伯堦 HL7 12

- Segment：A logical grouping of data fields。
- Segment ID：Each segment is identified by a unique 3-char code，Z 打頭的 Segment ID Code 是保留給 Locally-defined Message 使用的。
- Field：A string of characters，其欄位定義如下，
 - ✓ SEQ：Position，sequence w/in the segment。
 - ✓ LEN：Maximum length。
 - ✓ DT：Data type。
 - ✓ OPT：Optionality (R = Required，O = Optional，C = Conditional on

trigger event / other fields , X = not used w/ this trigger event , B = backward compatible) 。

- ✓ RP/# : Repetition (N = No repetition , Y = 無限次或 Site defined , Integer = 次數) 。
- ✓ TBL#
 - ◆ IS = coded value for user-defined table 。
 - ◆ ID = coded value for HL7 table , 如 Event Type 。
 - ◆ CE = encode by reference to other standard documents 。
 - ◆ User-defined Table , but not official standard 。
 - ◆ #9000 以上保留給 HL7 Externally-defined table 用 。
- ✓ ITEM# : Small integer that uniquely identifies the data item throughout the standard 。
- ✓ ELEMENT NAME : Descriptive name of the data item 。
- Message Delimiters : 定義於 MSH 之 Encoding Character 中 , 4 bytes , 順序為
 - ✓ Component , 預設值為^
 - ✓ Repetition , 預設值為~
 - ✓ Escape , 預設值為\
 - ✓ Sub-component , 預設值為&

(出國類別：研究)



✓ Message , Segment , Field , Component & Sub-component 之關係圖如下，

HL7 Ch2: Query & Control

Message Delimiters

Delimiter Name	Character	Hexadecimal	Decimal
Segment Terminator	<CR>	0x0d	10
Field Separator		0x7c	124
Component Separator	^	0x5e	94
Subcomponent Separator	&	0x26	38
Repetition Separator	~	0x7e	126
Escape Character	\	0x5c	92

2001年1月17日星期三 臺大醫院資訊室 鄭伯堉 HL7 20

5.1.7.2. 資料類型

- CNE (composite w/ no exceptions)
 - ✓ 被用於必須且強制性的編碼欄位上。
 - ✓ <id (ST)> ^<text (ST)> ^<name of coding system (ST)> ^<alternate id (ST)> ^<alternate text (ST)> ^<alternate name of coding system (ST)> ^<coding system version id (ST)> ^<alternate coding system version id (ST)> ^<original text (ST)> 。
 - ✓ 如果 CNE-1：identifier 已經被定義成 CNE 之一個欄位，或者該 CNE-1：identifier 並不存在於 Value Set 中，則無法產生一個有效的 OBX Instance。
 - ✓ 舉例來說，if lab sys。had a identifier="Decimal Range"，又因 no corresponding identifier available in HL7 table，所以 no valid OBX instance could be created。
 - ✓ 在 all valid instance of CNE fields，the identifier field must have a valid value from the specified table。
 - ✓ CNE-3：name of coding system 與 CNE-6：alternated name of coding system 必須於 coding system table 存在才可。
 - ✓ If the code scheme is anything other than HL7 table identifier，a version #(CNE-7orCNE-8) must be present。
- CP (composite price)
 - ✓ 不同時段對應價格不同。
 - ✓ Price type
 - ◆ AP = administrative price or handling fee，管理費。
 - ◆ PF = professional fee，專業加給。
 - ◆ UP = unit price，單價。
 - ◆ TF = technology fee，技術費。
 - ◆ DC = direct cost，直接費用。
 - ◆ IC = indirect cost，間接費用。

(出國類別：研究)

- ◆ TP = total price，總價。
- CQ (composite quantity w/ units)
 - ✓ 已經不再使用，因為可以用兩個獨立的 NM 與 CE 資料型別來分別傳送，所以不必存在此複合式的資料型別。
 - ✓ <quantity (NM)> ^ <units (CE)>
 - ◆ CE = <id (ST)> ^ <text (ST)> ^ <name of coding system (ST)>
 - ✓ |123 ° 7^kg|
 - ◆ <id> = kg
 - ◆ <text> = null
 - ◆ <name of coding system> = null
 - ✓ |150^lb&&ANSI+|
 - ◆ <id> = lb
 - ◆ <text> = null
 - ◆ <name of coding system> = ANSI+
- CWE (coded w/ exceptions)
 - ✓ <id (ST)> ^ <text (ST)> ^ <name of coding system (ST)> ^ <alternate id (ST)> ^ <alternate text (ST)> ^ <alternate name of coding system (ST)> ^ <coding system version id (ST)> ^ <alternate coding system version id (ST)> ^ <original text (ST)>
 - ✓ 兩組 coding systems，一主一副。
 - ◆ 主：CWE-1，CWE-2，CWE-3。
 - ◆ 副：CWE-4，CWE-5，CWE-6。
 - ✓ 為了與舊版本相容而存在的，新版中可以不必存在與使用的 components。
 - ✓ CWE-7：coding system version id。
 - ✓ CWE-8：alternate coding system version id。
 - ✓ 使用時機
 - ◆ 被用於當 coding system 被用到時，卻沒有相對應的 code 或 site agreement。
 - ◆ 被用於 Optional code field 或傳送那些尚未被允許的 value set
 - ◆ CWE-1~3 & CWE-7 之三種使用狀況，

- i. 狀況一：coded
 - If coding sys. = null, then 視為 HL7 coding sys.
 - If coding sys. <> HL7, then version id = 值。
 - If coding sys. = HL7, then version id = null / 值。
 - If version id = null, then 會被視為與 MSH 之 version id 相同。
 - ii. 狀況二：un-coded
 - If version id = null and text = 值, then coding sys. 與 version id 同狀況一(coded) 之處理方式。
 - 一般而言, id 會從 value set 中取值。
 - If id = 值, 但該值不是 value set 之一部份, then id = null 且 text = 值。
 - iii. 狀況三：data missing
 - If coding sys. = "HL7 CE Status" and version id = null / 值, then 會被視為與 MSH 之 version id 相同, 也就是與狀況一(coded)之型三之處理方式相似。
 - If id = 值, then 使用上述處理方式, 且 CWE Status 為
 - U unknown
 - UASK ask but unknown
 - NAV not available
 - NA not applicable
 - NASK not asked
- CWE vs. CNE
 - ✓ CWE = Coded w/ exceptions
 - ✓ CNE = Coded w/ no exceptions
 - ✓ 同樣是 <id (ST)> ^<text (ST)> ^<name of coding system (ST)> ^<alternate id (ST)> ^<alternate text (ST)> ^<alternate name of coding system (ST)> ^<coding system version id (ST)> ^<alternate coding system version id (ST)> ^<original text (ST)> 。

(出國類別：研究)

- CX (extended composite ID w/ check digit)
 - ✓ <ID (ST)>^<check digit (ST)> ^<code identifying the check digit scheme employed (ID)> ^<assigning authority (HD)> ^<identifier type code (IS)> ^<assigning facility (HD)> 。
 - ✓ 前半部很像 CK 。
 - ✓ CK = <ID (NM)>^<check digit (NM)> ^<code identifying the check digit scheme employed (ID)> ^<assigning authority (HD)>
 - ✓ <ID>與<check digit>之 data type 。
 - ✓ CX 是 ST，CK 是 NM 。
 - ✓ 如果 ID 是字元的話，check digit = NULL 。
 - ✓ <Identifier authority> vs. <assigning facility>
 - ◆ 可以視為信用卡 vs. 信用卡發卡銀行或是病歷號 vs. 病歷號給號醫院 。
 - ◆ 範例
 - |1234567^4^M11^ADT01^MR^University Hospital|
 - |3326090^2^M10^ADT01^MR^NTUH|
- DLN (driver's license #)
 - ✓ <license # (ST)>^<issuing state, province, country (IS)>^<expiration date (DT)> 。
 - ✓ 好像可以用 CX 來代替，但是 CX 沒有<expiration date> 。
 - ✓ 範例
 - ◆ |F123456789^TWN^20001217|
- DR (date / time range)
 - ✓ <range start date/time (TS)>^<range end date/time (TS)> 。
 - ✓ 範例
 - ◆ |20001124^20001127|，準備課程內容時間 。
 - ◆ |200011271400^200011271700|，上課時間 。
 - ◆ |200011271400+0800^200011271700+0800|，加上台北時區 。
- DT (date)
 - ✓ YYYY[MM[DD]]
 - ✓ 範例

- ◆ |2000|
- ◆ |200011|
- ◆ |20001127|
- ED (encapsulated data)
 - ✓ 直接攜帶編碼後的資料，似 RP (Reference Pointer)，RP 是用指標指到另一系統之資料上。
 - ✓ $\langle \text{source app. (HD)} \rangle \wedge \langle \text{type of data (ID)} \rangle \wedge \langle \text{data subtype (ID)} \rangle \wedge \langle \text{encoding (ID)} \rangle \wedge \langle \text{data (ST)} \rangle$ 。
 - ◆ $\langle \text{type of data} \rangle$ ：資料類型。
 - SI (scanned image)，NS (non-scanned image)，SD (scanned document)，TX (machine readable text)，FT (formatted text)：v2.3.1 以前使用。
 - TEXT (machine readable text)，IM (image data)，AU (audio data)，AP (other application data)，Image，Audio，Application：v2.3 以後使用。
 - ◆ $\langle \text{data subtype} \rangle$ ：資料格式。
 - TIFF，JPEG，GIF，FAX，...
 - ✓ MIME = Multi-purpose Internet Mail Extensions。
 - ◆ RFC 1521。
 - ◆ Base 64：uses 65 US-ASCII characters，lower/upper 英文字母，0~9，+，/，=。
 - ◆ 以 6 bits 來 encode 資料， $6 \text{ bits} \times 4 = 24 \text{ bits} = 3 \text{ 個 octets (8 bits} \times 3)$ ，也就是用 3 個 bytes 來放 4 個 ASCII 碼。
 - ◆ If $< 24 \text{ input bits}$ ，then 從右補零為 24 bits。
 - ◆ If output 不須顯示出來，then 補成"="。
 - ◆ Output 一定是 8 的倍數。
 - If final input = $24N \text{ bit}$ ，then output 是 $4N \text{ 個 char} + 0 \text{ 個 "="}$ 。
 - If final input = 8 bit ，then output 是 $2 \text{ 個 char} + 2 \text{ 個 "="}$ 。
 - If final input = 16 bit ，then output 是 $3 \text{ 個 char} + 1 \text{ 個 "="}$ 。

(出國類別：研究)

- ✓ If <encoding> = "A" 。
- ✓ If <encoding> = "Hex" 或 "Base64" 。
- EI (entity identifier)
 - ✓ <entity identifier (ST)>^<namespace (IS)> ^<universal ID (ST)> ^<universal ID type (ID)> 。
 - ✓ 定義一連串 ID 中之某一個 ID 。
 - ✓ 似 HD (Hierarchic Designator) 。
- FC (financial class)
 - ✓ <financial class (IS)>^<effective date (TS)> 。
 - ◆ <financial class>：針對某一人之財物分類。
- FT (formatted text)
 - ✓ FT <= 64K 。
 - ✓ 範例
 - ◆ \|。sp\((skip one vertical line)|
- HD (hierarchic designator)
 - ✓ <namespace ID (IS)> ^ <universal ID (ST)> ^ <universal ID type (ID)> 。
 - ✓ 內容為(HD-1) or (HD-2 and HD-3) 。
 - ✓ <namespace ID>:如果 HD 被其他資料型別所引用,此 namespace ID table 必須被 re-defined 。
 - ✓ <universal ID>：是依據 UID 之條件所格式化之字串,所以必須依循 UID type scheme 所定之規則,但此規格並非由 HL7 所制定的。
 - ✓ HD 主要指明某一(管理者/系統/應用程式/組織)所(管理/指定)的一組 ID(Placer/Filler #/PID/Provider ID) 。
 - ✓ HD 代替掉 v2.1,v2.2 之 app°ID 功能,也就是說,app°ID = namespace ID 。
 - ✓ 似 EI (Entity Identifier) 。
 - ✓ HD 定義
 - ◆ An assigning authority (應用程式或組織) 。
 - 應用程式：Registration sys。--> assign PID 。

- 組織：License authority --> assign driver license # 或 assign professional id。
- ◆ An assigning facility (設備) 。
 - 被指定 ID 的設備。
- ◆ Assigning authority vs. coded element-defined
 - Assigning authority：真實世界的事物。
 - Coded element-defined：一種觀念上的代碼。
 - 兩者都在指明一組離散的 ID。
 - 2 種 identifiers
 - Local identifier --> HD-1：name space ID。
 - Universal identifier --> HD-2：UID & HD-3：UID type，又叫做 Public-assigned ID。
 - Universal ID (UID) Type
 - DNS = Internet dot name。
 - GUID = UUID。
 - HCD = CEN Healthcare Coding System Designator。
 - HL7。
 - ISO = International standard organization object identifier。
 - L，M，N = reserved for local definition coding scheme。
 - Random = Base 64。
 - UUID = DCE Universal Unique Identifier。
 - X400 = X。400 MHS format identifier。
 - X500 = X。500 directory name。
- ID (coded value for HL7 defined tables)
 - ✓ HL7 所定義的 ID 代碼，可以從 ID table 中查詢得到。
 - ✓ 使用 ID 的欄位範例。
 - ✓ MSH-12：version id。
 - ✓ OBR-25：result status。

(出國類別：研究)

- IS (coded value for user-defined tables)
 - ✓ 非 HL7 所定義的 ID 代碼，無法從 ID table 中查詢得到，是使用者自行定義的代碼。
 - ✓ 使用 IS 的欄位範例。
 - ✓ Event reason code
 - ◆ 01 = patient request。
 - ◆ 02 = physician order。
 - ◆ 03 = census management。
- JCC (job code / class)
 - ✓ <job code (IS)>^<job class (IS)>。
 - ✓ <job class>：職業分類。
- MA (multiplexed array)
 - ✓ <sample 1 from channel 1 (NM)> ^ <sample 1 from channel 2 (NM)> ^ <sample 1 from channel 3 (NM)> ...~ <sample 2 from channel 1 (NM)> ^ <sample 2 from channel 2 (NM)> ^ <sample 2 from channel 3> ...~ ...。
 - ✓ 主要是用來表示多工頻道之波形資料。
- MO (money)
 - ✓ <quantity (NM)> ^ <denomination (ID)>。
 - ✓ <denomination>：貨幣單位，參 ISO-4217。
 - ✓ 範例
 - ◆ |99。50^USD|
 - ◆ |1000^TWD|
- NA (numeric array)
 - ✓ 用來表示一連串的數值資料。
- PL (person location)
 - ✓ 用來表示病友在醫院內的位置。
 - ✓ <point of care (IS)> ^ <room (IS)> ^ <bed (IS)> ^ <facility (HD)> ^ <location status (IS)> ^ <person location type (IS)> ^ <building (IS)> ^ <floor (IS)> ^ <location description (ST)>。
 - ◆ <point of care>：就是 Ward #(護理站，部門，門診，...)。

- ◆ <location status>：什麼類型的床等(頭等，二等，加護，...)。
- ◆ <person location type>：什麼類型的位置(內科，外科，...，護理站)。
- ◆ <building>：哪一棟(A，B，C，D)。
- ◆ <floor>：哪一樓層。
- PN (person name)
 - ✓ <family name (ST)> & <last name prefix (ST)> ^ <given name (ST)> ^ <middle or initial name (ST)> ^ <suffix (舉例來說，JR or III (ST)> ^ <prefix (舉例來說，DR) (ST)> ^ <degree (舉例來說，MD) (IS)>。
 - ✓ 最大 48 字元，含 delimiter。
 - ✓ Receiver 會全部轉為大寫字母。
 - ✓ 參考並比較 CN、XCN、XPN。
 - ✓ 可以重複設定不同的字元集，但最後必須回到預設的字元集上。例如，日本可以依次叫用 4 個不同的字元集。
 - ✓ 與 DICOM 共舞。
 - ◆ Character Repertoires (字元演奏)。
 - ◆ 參 DICOM part 5 supplement 9 之 62E1 小節說明。
 - ◆ 參 ISO-8859。
- PPN (performing person time stamp)
 - ✓ <ID # (ST)> ^ <family name (ST)> ^ <last name prefix (ST)> ^ <given name (ST)> ^ <middle or initial name (ST)> ^ <suffix (舉例來說，JR or III) (ST)> ^ <prefix (舉例來說，DR) (ST)> ^ <degree (舉例來說，MD) (ST)> ^ <source table (IS)> ^ <assigning authority (HD)> ^ <name type code (ID)> ^ <identifier check digit (ST)> ^ <code identifying the check digit scheme employed (ID)> ^ <identifier type code (IS)> ^ <assigning facility (HD)> ^ <date/time action performed (TS)> ^ <name representation code (ID)>。
 - ◆ <ID #>：If <ID #> = null，then <source table> = 值或<assigning authority> = 值。
 - ◆ <assigning facility>：第一次 assign 病友之機構。

(出國類別：研究)

- ✓ 似 XCN + TS 。
- PT (processing type)
 - ✓ <processing ID (ID)> ^ <processing mode (ID)>
 - ✓ <processing ID>
 - ◆ D = debugging 。
 - ◆ P = production 。
 - ◆ T = training 。
 - ✓ <processing mode>
 - ◆ A = archive 。
 - ◆ R = restore from archive 。
 - ◆ I = initial load 。
 - ◆ T = current processing , X'mitted at intervals (schedule/on-demand) 。
 - ◆ Not present = 目前的 processing(預設值) 。
- QIP (query input parameter list)
 - ✓ <segment field name (ST)> ^ <value1 (ST) & value2 (ST) & value3 (ST) & ...> 。
 - ✓ 範例
 - ◆ [|@PID.5.1^鄭|](#)
 - ◆ [|@PID.5.1^鄭&龐|](#)
- QSC (query selection criteria)
 - ✓ <segment field name (ST)> ^ <relational operator (ID)> ^ <value (ST)> ^ <relational conjunction (ID)> 。
 - ✓ 範例
 - ◆ [|@PID.5.1^EQ^鄭|](#)
- RCD (row column definition)
 - ✓ 定義欄位 。
 - ✓ <segment field name (ST)> ^ <HL7 data type (ST)> ^ <max 。
 - width (NM)> 。
 - ✓ 範例

◆ [@PID.5.1^ST^4]

- RI (repeat interval)
 - ✓ 定義排程之時間間隔。
 - ✓ <repeat pattern (IS)> ^ <explicit time interval (ST)> 。
- RP (reference pointer)
 - ✓ 傳送存放在其他系統中之資料的資訊。
 - ✓ <pointer (ST)> ^ <application ID (HD)> ^ <type of data (ID)> ^ <subtype (ID)> 。
 - ✓ 範例
 - ◆ |1234A321634BC^EFC^SD|
 - ✓ 似 ED。
 - ✓ HL7 只使用 DICOM 之 image 資料部份。
 - ✓ DICOM 含
 - ◆ File Meta Information Header 。
 - ◆ Data Set 。
- SCV (scheduling class value pair)
 - ✓ 問 receiver 可以接受所安排的排程嗎?
 - ✓ <parameter class (IS)> ^ <parameter value (ST)>
 - ◆ <parameter class>：排程編碼。
 - ◆ <parameter value>：排程值。
- SI (sequence ID)
 - ✓ 資料表中的某一非負整數。
- SN (structured numeric)
 - ✓ <comparator (ST)> ^ <num1 (NM)> ^ <separator/suffix (ST)> ^ <num2 (NM)> 。
 - ◆ <comparator>：">"，">="，"<"，"<="，"="，"<>"。
 - ◆ <separator/suffix>："-","+", "/", "。", ":"。
 - ✓ 範例
 - ◆ |>^100| >100
 - ◆ |^100^-^200| 100~200
 - ◆ |^1^：^228| 1：228

(出國類別：研究)

- ◆ |²⁺| 2+
- ST (string data)
 - ✓ 適用於較短的文字串上，<=200。
 - ✓ 較長的文字串最好用 TX 或 FT。
- TM (time)
 - ✓ HH[MM[SS[。S[S[S[S]]]]]] [+/-ZZZZ]。
 - ✓ 範例
 - ◆ |235959+1100|
 - ◆ |0800|
 - ◆ |093544。2312|
 - ◆ |13|
- TN (telephone #)
 - ✓ [NN] [(999)]999-9999[X99999][B99999][C any text]。
 - ◆ [NN]：country code。
 - ✓ 範例
 - ◆ |(415)236-8618X1082|
 - ◆ |236-8618CWEEKENDS|
- TQ (timing quantity)
- TS (time stamp)
 - ✓ YYYY[MM[DD[HHMM[SS[。S[S[S[S]]]]]]]] [+/-ZZZZ] ^ <degree of precision>。
 - ✓ 範例
 - ◆ |197612170159-0500|
 - ◆ |19641217|
- TX (text data)
 - ✓ 大字串。
 - ✓ 最大 64KB，含 delimiter，似
 - ◆ ST (String Text)，小字串。
 - ◆ FT (Format Text)。
- VH (visiting hours)
 - ✓ 診間開放時間。

- ✓ <start day range (ID)> ^ <end day range (ID)> ^ <start hour range (TM)> ^ <end hour range (TM)> 。
- ✓ 範例
 - ◆ |MON^FRI|
- VID (version identifier)
 - ✓ HL7 version identifier 。
 - ✓ <version ID (ID)> ^ <internationalization code (CE)> ^ <international version ID (CE)> 。
 - ◆ <version ID>：HL7 版本。
 - ◆ <internationalization code>：國家碼，參 ISO 3166-1：1997。
 - ◆ <international version ID>：該國家 HL7 副版本。
 - ✓ 範例
 - ◆ |2。4^TWN^2。4-TW|
- XAD (extended address)
 - ✓ <street address (ST)> ^ <other designation (ST)> ^ <city (ST)> ^ <state or province (ST)> ^ <zip or postal code (ST)> ^ <country (ID)> ^ <address type (ID)> ^ <other geographic designation (ST)> ^ <country/parish code (IS)> ^ <census tract (IS)> ^ <address representation code (ID)> 。
 - ◆ <country>：參 ISO-3166。
 - ◆ <country/parish code>：牧師管區。
 - ◆ <census tract>：普查土地。
 - ✓ 範例
 - ◆ |7 Chung-san S。Rd。^^TAIPEI^TWN^106^TWN^M|
- XAD (extended address)
 - ✓ If XAD-9 <> null，then XAD-8 = null，且不可重覆，此外最好以 XAD-9 取代 XAD-8，因為 XAD-8 在 v2.3 版以後不再使用。
- XCN (extended composite ID # and name for person)
 - ✓ <ID # (ST)> ^ <family name (ST)> & <last name prefix (ST)> ^ <given name (ST)> ^ <middle or initial name (ST)> ^ <suffix (舉例來說，JR，III) (ST)> ^ <prefix (舉例來說，DR) (ST)> ^ <degree (舉例

(出國類別：研究)

來說，MD)> ^ <source table (IS)> ^ <assigning authority (HD)> ^ <name type code (ID)> ^ <identifier check digit (ST)> ^ <code identifying the check digit scheme employed (ID)> ^ <identifier type code (IS)> ^ <assigning facility (HD)> ^ <name representation code (ID)> 。

✓ 範例

◆ |12372^GA^JOHN^J^III^DR^PHD^ADT1^^L^4^M11^MR|

◆ |APA02^LAI^JS^^^DR^MD^^&NTUH-KK 。

HIS&L^L^9^M10^DN^ &NTUH 。

◆ |3326090^CHENG^PH^^^MS^^&NTUH-KK 。

HIS&L^L^3^M10^MR^ &NTUH 。

● XON (extended composite name & identifier # for organization)

✓ 加強版複合式組織名稱與代號。

✓ <organization name (ST)> ^ <organization name type code (IS)> ^ <ID # (NM)> ^ <check digit (NM)> ^ <code identifying the check digit scheme employed (IS)> ^ <assigning authority (HD)> ^ <identifier type code (IS)> ^ <assigning facility ID (HD)> ^ <name representation code (ID)> 。

✓ 範例

◆ |NTUH^L^000716^9^M10^&NTUH 。

HIS&L^XX^&HIN 。

HINS&L^A|

◆ |NTUH^L^000716^9^M10^&NTUH 。

HIS&L^XX^&HIN 。

HINS&L^A|

● XPN (extended person name)

✓ 加強版病友姓名。

✓ <family name (ST)> & <last name prefix (ST)> ^ <given name (ST)> ^ <middle or initial name (ST)> ^ <suffix (舉例來說，JR，III) (ST)> ^ <prefix (舉例來說，DR) (ST)> ^ <degree (舉例來說，MD)> ^ <name type code (ID)> ^ <name representation code (ID)> 。

✓ 範例

◆ |CHENG^PH^^^MS^L|

- XTN (extended telecommunication #)
 - ✓ 加強版通訊號碼。
 - ✓ [NNN][(999)999-9999[X99999][B99999][C any text] ^
 - <telecommunication use code (ID)> ^ <telecommunication equipment type (ID)> ^ <email address (ST)> ^ <country code (NM)> ^
 - <area/city code (NM)> ^ <phone # (NM)> ^ <extension (NM)> ^ <any text (ST)> 。
 - ✓ 範例
 - ◆ |886(022)321-0037^ORN^FAX^cph@ms5。hinet。net|

5.1.8. 控制(Control)

5.1.8.1. Escape

- \H\ = highlight text 。
- \N\ = normal text 。
- \F\ = field separator 。
- \S\ = component separator 。
- \T\ = sub-component separator 。
- \R\ = repetition separator 。
- \E\ = escape separator 。
- \Xddd...\ = hexadecimal data 。
- \Zddd...\ = locally defined escape sequence 。
- \Cxyy\ = single-byte character set escape sequence 。
- \Mxyyzz\ = multi-byte character set escape sequence where zz is optional 。

Escape sequence 不准巢狀(Nested)使用，supports multiple character set for some data types，includes PN、XPN、XON、XCN、XAD。

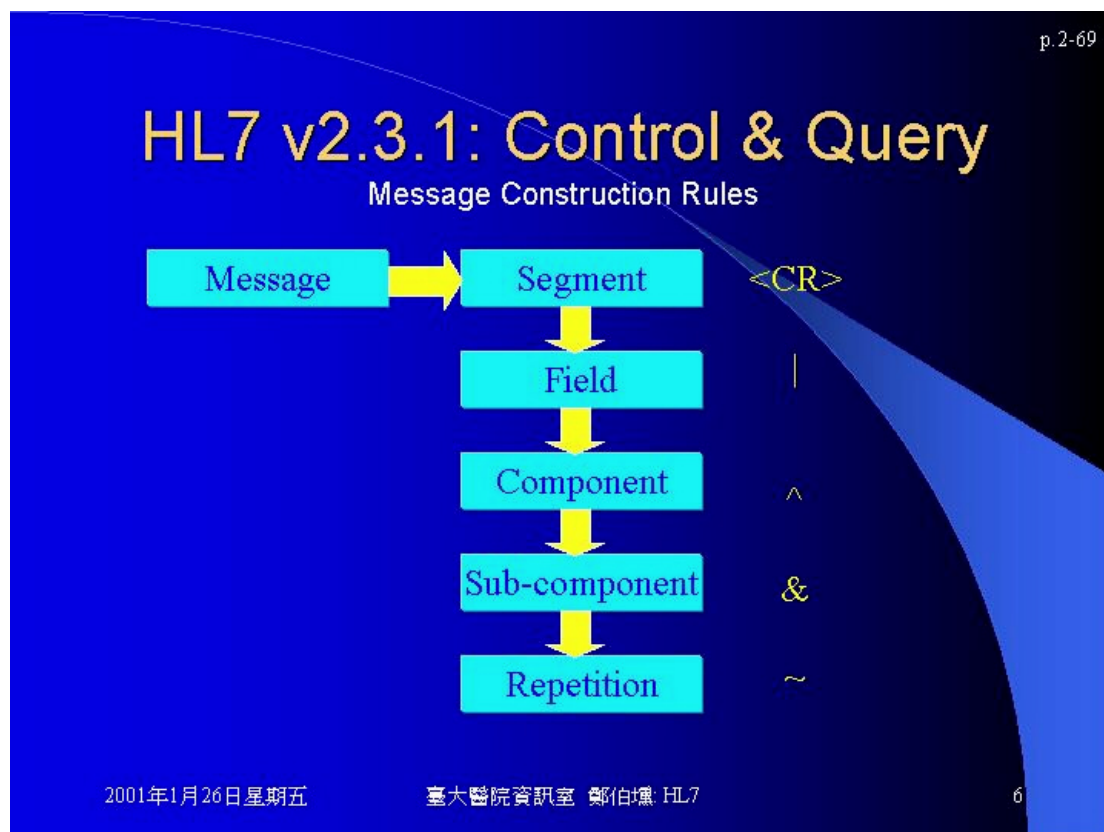
(出國類別：研究)

- \C2842\ = ISO IR6 G0 (ISO 646 : ASCII) 。
- \C2D41\ = ISO IR 100 (ISO 8859 : Latin Alphabet 1) 。
- \C2D42\ = ISO IR 101 (ISO 8859 : Latin Alphabet 2) 。
- \C2D43\ = ISO IR 109 (ISO 8859 : Latin Alphabet 3) 。
- \C2D44\ = ISO IR 110 (ISO 8859 : Latin Alphabet 4) 。
- \C2D4C\ = ISO IR 144 (ISO 8859 : Cyrillic) 。
- \C2D47\ = ISO IR 127 (ISO 8859 : Arabic) 。
- \C2D46\ = ISO IR 126 (ISO 8859 : Greek) 。
- \C2D48\ = ISO IR 138 (ISO 8859 : Hebrew) 。
- \C2D4D\ = ISO IR 148 (ISO 8859 : Latin Alphabet 5) 。
- \C284A\ = ISO IR 14 (JIS X0201-1976 : Romaji) 。
- \C2849\ = ISO IR 13 (JIS X0201 : Katakana) 。
- \M2442\ = ISO IR 87 (JIS X0208 : Kanji , Hiragana , Katakana) 。
- \M242844\ = ISO IR 159 (JIS X0212 : Supplementary Kanji) 。

Escape sequence 在 FT (Formatted Text) 中 。

- .sp <number> 換行且垂直向下跳數行 。
- .br 換行 。
- .fi word warp mode (fill mode) 。
- .nf no-warp mode 。
- .in <number> 跳數個空格，不能放在第一列列印字元之後 。
- .ti <number> 向左/右跳數個空格，不能放在第一列列印字元之後 。
- .sk <number> 向跳右數個空格 。
- .ce 結束本行，下行靠中對齊 。

5.1.8.2. 訊息建置規則



5.1.8.3. 符號(Notations)

- [...] = 表示其內之 segments is optional。
- {...} = 表示其內之 segments is repetition。
- [...] = 表示其內之 segments is repetition 且 optional。
- [...] = {...}。

5.1.8.4. Acknowledgement Mode

- Version Compatibility, 為了版本相容起見, V2.3 以前的版本叫做 Original Acknowledgement Mode, V2.3 以後的版本叫做 Enhanced

(出國類別：研究)

Acknowledgement Mode ◦

- Original Acknowledgement Mode

p.2-72

HL7 v2.3.1: Control & Query

Original Acknowledgement Mode

#	Sender / Placer		Receiver / Filler
1	Message	→	
2			Validate message
3			Pass to application
4		←	Response message Error message Reject message
5	Pass to application		

2001年1月26日星期五
臺大醫院資訊室 鄭伯璦 HL7
10

p.2-72

HL7 v2.3.1: Control & Query

Original Acknowledgement Mode

- S = Sender / Placer
- R = Receiver / Filler
- SPS = Sender Protocol Software
- RPS = Receiver Protocol Software

```

graph LR
    S((S)) <--> SPS((SPS))
    SPS <--> RPS((RPS))
    RPS <--> R((R))
    
```

2001年1月26日星期五
臺大醫院資訊室 鄭伯璦 HL7
11

HL7 v2.3.1: Control & Query

Original Acknowledgement Mode



1		Validate message MSH-9: message type MSH-11: process id MSH-12: version id	→	
2		If fail, then reject ACK^MSA-1 = AR	←	
3			→	if ok then pass msg to app.
4		If ok, ACK^MSA-1=AA/AE/AR	←	
5	Pass message MSA-2 = MSH-10 MSA-4 = MSH-13 + 1		←	

HL7 v2.3.1: Control & Query

Original Acknowledgement Mode

- MSH-3: sending app.
- MSH-4: sending facility
- MSH-5: receiving app.
- MSH-6: receiving facility
- MSH-7: D/T of message
- MSH-9: message type
- MSH-10: message control id
- MSH-11: process id
- MSH-12: version id
- MSH-13: sequence #
- MSH-14: continuation pointer
- MSA-1: ack. code (AA/AE/AR)
- MSA-2: message control id
- MSA-3: text message
- MSA-4: expected sequence #
- MSA-5: delay ack. type(D/F/Null)

(出國類別：研究)

5.1.8.5. Enhanced Acknowledgement Mode

p.2-72

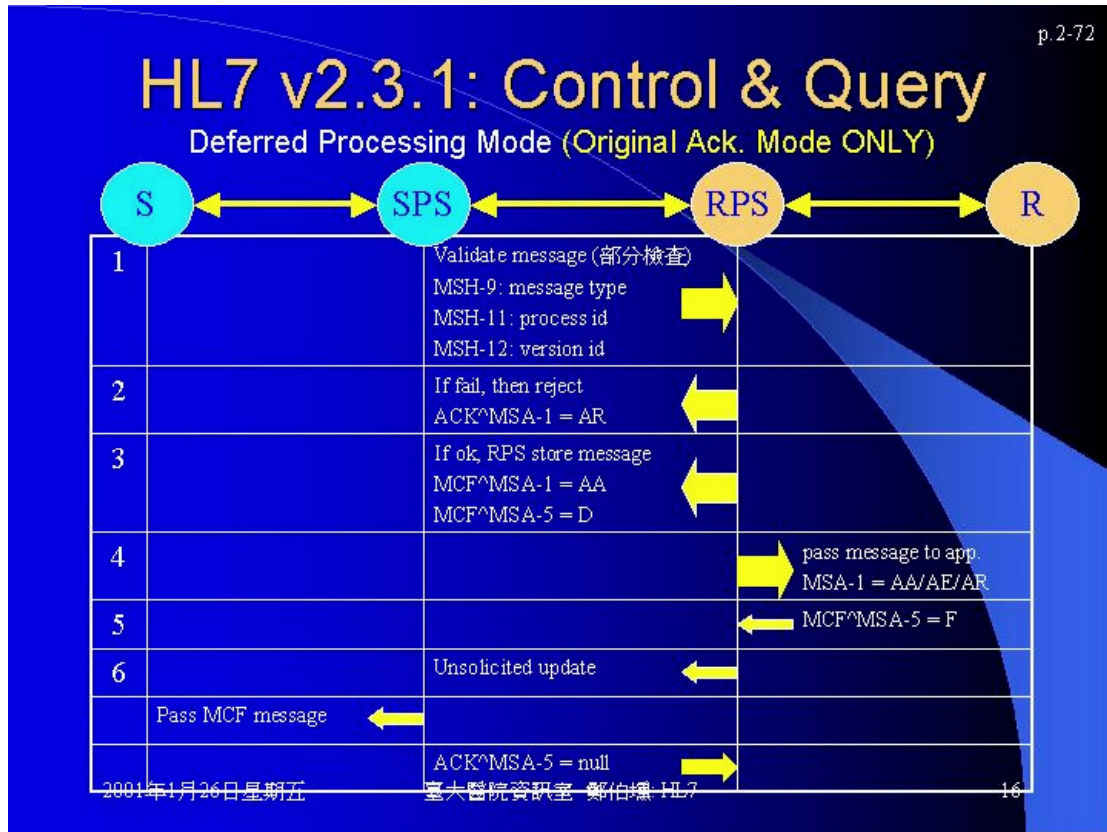
HL7 v2.3.1: Control & Query

Enhanced Acknowledgement Mode

#	Sender / Placer		Receiver / Filler
1	Message		
2		→	Commit & save message
3			Check message header, 是否要回傳accept acknowledgement message
4		←	Accept acknowledgement message, 如果 Filler 想要回送 Application-dependent 資訊給 Placer, 可以隨後回傳 application acknowledgement message
5		←	Application acknowledgement message
6	Receiver / Filler		Sender / Placer

此時Placer / Filler的角色開始互換了

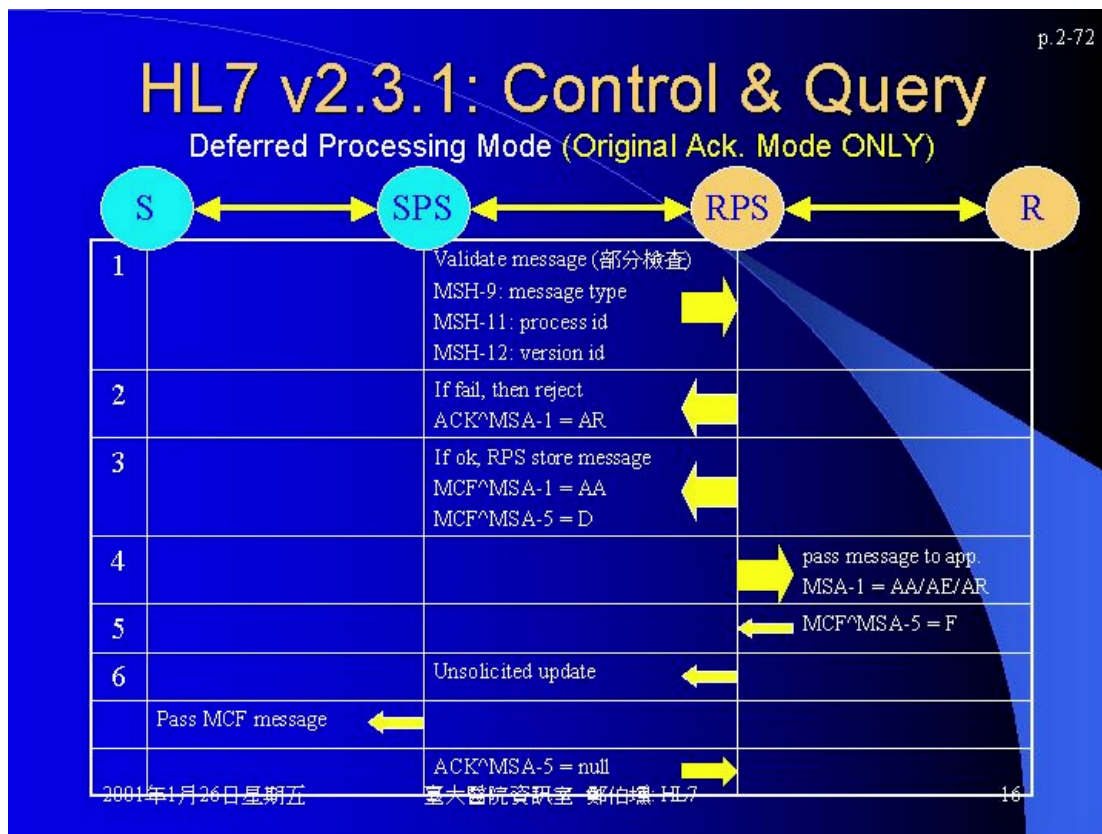
2001年1月26日星期五 臺大醫院資訊室 鄭伯璦 HL7 14



5.1.8.6. Deferred Processing Mode

Original Acknowledge Mode ONLY。

(出國類別：研究)



5.1.8.7. ACK, QAK vs. MCF

- ACK = general Acknowledge，用於
 - ✓ Application 沒有定義特定的 application-level Acknowledge message 時。
 - ✓ Application 有 error 時。
 - ✓ Accept-level Acknowledge。
- MCF = delayed Acknowledge，也就是說，deferred Acknowledge，使用於非同步 application Acknowledge。
- QAK 用於
 - ✓ Original mode：於 error segment 之後，QAK-1：query tag 不使用。
 - ✓ Enhanced mode：任一 query 之後，QAK-2：query response status。
 - ◆ OK = data found，no error (default)
 - ◆ NF = no data found，no error

- ◆ AE = application error
- ◆ AR = application reject

5.1.8.8. Message Data Format

- Display-oriented data
 - ✓ Display only, 不會 update receiver DB, 也就是說, unsolicited display
 - ✓ 方法: Unsolicited display message。
- Record-oriented data, 型式
 - ✓ Tabular (詳參 Ch2-15)。
 - ✓ Application message segment-oriented (詳參 Ch3~Ch12)。
 - ✓ 需要 functionally-specific capabilities。

5.1.8.9. Message : ADD 使用時機

p.2-89

HL7 v2.3.1: Control & Query

Message: ADD使用時機

- 當一個segment未完成時, 須加上空的 ADD, 再加上DSC

...
 ANY
 ADD
 DSC
- 延續query時, receiver順序一定是

...
 QRD
 QRF
 ADD
 ANY
- Unsolicited update時, sender再次查詢

MSH
 ADD
 ANY

2001年1月26日星期五
臺大醫院資訊室 鄭伯堉 HL7
20

(出國類別：研究)

5.1.9. 查詢(Query)

5.1.9.1. 型式

- Display ◦
- Record-oriented query ◦
 - ✓ Original mode ◦
 - ✓ Embedded query language ◦
 - ✓ Virtual table ◦
 - ✓ Stored procedure query ◦
- Event replay request ◦

5.1.9.2. Query : Response

Query Response 包括

- Display response (O/E) ◦
- Record-oriented response (O) ◦
 - ✓ Original mode (O) ◦
 - ✓ Tabular response (E) ◦
 - ✓ Event replay response (E) ◦

Query Response Mode 包括

- Original mode (O) ◦
 - ✓ Display ◦
 - ✓ Record-oriented ◦
- Enhanced mode (E)
 - ✓ Display ◦

- ✓ Tabular ◦
- ✓ Event replay ◦

5.1.9.3. Original Mode Query

HL7 v2.3.1: Control & Query

Query
Original Mode

Query / Response Name	Sender / Placer Msg^Evt(Msg type, Seg. ID)	Receiver / Filler Msg^Evt(Msg type, Seg. ID, ...)
Display Query	QRY^Q01 (QRD, QRF)	→
Immediate Response	←	DSR^Q01 (QAK, QRD, QRF, DSP)
Deferred Query	QRY^Q02 (QRD, QRF)	→
	←	QCK^Q02 (QAK)
Deferred Response	←	DSR^Q03 (QAK, QRD, QRF, DSP)
	ACK^Q03	→

- QRY
 - display-oriented
- QRY, DSR, ACK
 - record-oriented
 - specific to a functional area

2001年1月26日星期五 臺大醫院資訊室 鄭伯璦 HL7 4

(出國類別：研究)

5.1.9.4. Enhanced Mode Query

HL7 v2.3.1: Control & Query

Query Enhanced Mode

Query / Response Name	Sender / Placer Msg^Evt(Msg type, Seg. Id)	Receiver / Filler Msg^Evt(Msg type, Seg. Id, ...)
Embedded Query	EQQ^Q04 (EQL)	
Display Response		EDR^R07 (QAK, SDP)
Virtual Table Query	VQQ^Q07 (VTQ, RDF [o])	
Tabular Response		TBR^R08 (QAK, RDF [m], RDT)
Stored Procedure Query	SPQ^Q08 (SPR, RDF [o])	
Tabular Response		TBR^R08 (QAK, RDF [m], RDT)
Event Replay Query	RQQ^Q09 (ERQ)	
Event Replay Response		ERP^R09 (QAK, ERQ)

2001年1月26日星期五 臺大醫院資訊室 鄭伯堉 HL7 5

5.1.9.5. Original Mode Query Segments

- Original
 - ✓ QRD = original mode query definition segment。
 - ✓ QRF = original mode query filter segment。
- Unsolicited
 - ✓ URD = unsolicited result/update definition seg。
 - ✓ URS = unsolicited selection segment。
- Tabular
 - ✓ RDF = table row definition segment。
 - ◆ 可用於 SPQ/VQQ message 中。
 - ◆ 必須用於 TBR message 中。

- ✓ RDT = table data segment 。
 - ◆ 必須用於 TBR message 中。
- QRD , URD vs. RDF

HL7 v2.3.1: Control & Query

Original Mode Query Segments QRD, URD vs. RDF

QRD	URD	RDF
<ul style="list-style-type: none"> ● Query DT ● Query Format Code (D/R/T) ● Query Priority (D/I) ● Query ID ● Deferred Response Type (B/L) ● Deferred Response DT ● Quantity Limited Request ● Who Subject Filter ● What Subject Filter ● What Department Data Code ● What Data Code Value Quality ● Query Results Level (O/R/S/T) 	<ul style="list-style-type: none"> ● R/U DT ● Query Priority (D/I) ● R/U Who Subject Filter, 查詢誰 ● R/U What Subject Filter, 查什麼 ● R/U What Dept. Data Code, 屬於哪一部門的資料 ● R/U Display / Print Locations ● R/U Query Results Level (O/R/S/T) 	<ul style="list-style-type: none"> ● # Of Column Per Row ● Column Description (RCD)

2001年1月26日星期五
臺大醫院資訊室 鄭伯璦 HL7
7

(出國類別：研究)

HL7 v2.3.1: Control & Query

Original Mode Query Segments QRF, URS vs. RDT

QRF	URS	RDT
<ul style="list-style-type: none"> • Where Subject Filter • When Data Start DT • When Data End DT • What User Qualifier • Other QRY Subject Filter • Which DT Qualifier • Which DT Status Qualifier • DT Selection Qualifier • When Quantity / Timing Qualifier 	<ul style="list-style-type: none"> • R/U Where Subject Filter • R/U When Data Start DT • R/U When Data End DT • R/U What User Qualifier • R/U Other QRY Subject Filter • R/U Which DT Qualifier • R/U Which DT Status Qualifier • R/U DT Selection Qualifier • R/U When Quantity / Timing Qualifier 	<ul style="list-style-type: none"> • Column Value

2001年1月26日星期五
臺大醫院資訊室 鄭伯璦 HL7
8

● QRF

- ✓ Where Subject Filter，哪裡在要求查詢。
- ✓ When Data Start DT。
- ✓ When Data End DT。
- ✓ What User Qualifier，進一部資料說明。
- ✓ Other QRY Subject Filter，區域過濾，用於兩系統之間。
- ✓ Which DT Qualifier，介於起迄時間內之何種資料，
ANY/COL/ORD/RCT/REP/SCHED。
- ✓ Which DT Status Qualifier，介於起迄時間內資料之報告狀態，
ANY/CFN/COR/FIN/PRE/REP。
- ✓ DT Selection Qualifier，查詢結果(排列順序)依時間而異，
1ST/ALL/LST/REV。
- ✓ When Quantity / Timing Qualifier，多重回傳查詢結果之時間間隔。

5.1.9.6. Enhanced Mode Query Segments

- EQL , VTQ , SPR vs. ERQ
 - ✓ EQL = embedded query language segment ◦
 - ✓ VTQ = virtual table query request segment ◦
 - ✓ SPR = stored procedure request definition segment ◦
 - ✓ ERQ = event replay query segment ◦

HL7 v2.3.1: Control & Query

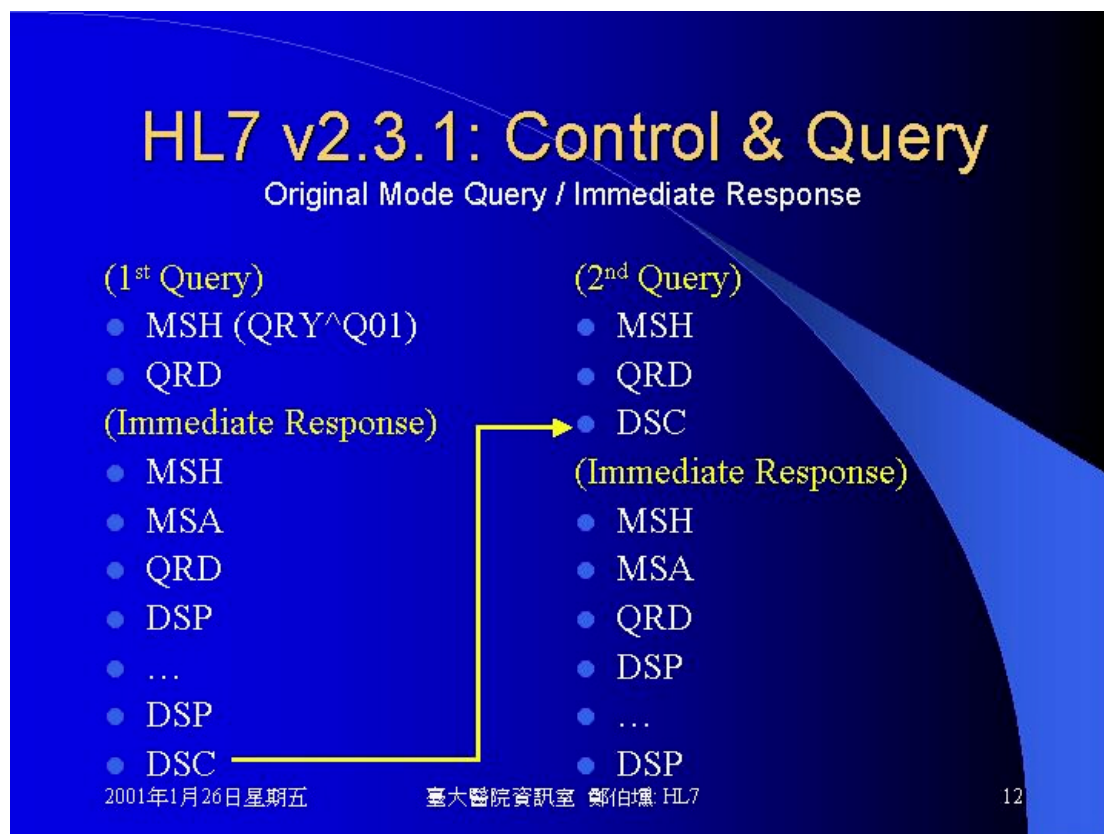
Enhanced Mode Query Segments EQL, VTQ, SPR vs. ERQ

EQL	VTQ	SPQ	ERQ
<ul style="list-style-type: none"> ● Query tag ● Query/response format code ● EQL query name ● EQL query statement (ST) 	<ul style="list-style-type: none"> ● Query tag ● Query/response format code ● VT query name ● VT name ● Selection criteria (QSC) 	<ul style="list-style-type: none"> ● Query tag ● Query/response format code ● SP name ● Input parameter list (QIP) 	<ul style="list-style-type: none"> ● Query tag ● Event identifier ● Input parameter list (QIP)
<div style="text-align: center;">↕</div>	<div style="text-align: center;">↕</div>	<div style="text-align: center;">↕</div>	<div style="text-align: center;">↕</div>
Select @PID.1,@PID.2 From PID Where @PID.5=3326090	@PID.5^EQ3326090	@PID.5^3326090	@PID.5^ST^7^3326090

2001年1月26日星期五
臺大醫院資訊室 鄭伯璦 HL7
11

(出國類別：研究)

5.1.9.7. Original Mode Query / Immediate Response



5.1.9.8. Query Error

當 query 發生 error 時，

- DSR, TBR, ERP 之 MSA segment 中含 AE/AR。
- MSA-6: error condition。
- DSC not sent 或 DSC-1=null。
- QAK-2: query response status = AE/AR。

5.1.10. 訊息(Message)

5.1.10.1. Message：MSH-20

- MSH-20：alternate char set handling (ID)。
 - ✓ ISO 2022-1994：可以使用 ASCII "esc" 字元作為 escape 字元。
 - ✓ v2.3：v2.3 escape 方法，但不使用 ASCII "esc" 字元。
 - ✓ Null：No char set switching，預設值。

5.1.10.2. Representation Name Type

- I = Ideographic，意表文字的。
- A = Alphabetic，音節的。
- P = Phonetic，語音的。

5.1.10.3. Z Locally-Defined

- HL7 message。
- HL7 segment。
- RI 之 data type。
- TX 之 \Zdddd...\，local escape sequence。

5.1.10.4. Data Type Length

- TS <= 26。
- CE <= 60。

(出國類別：研究)

- PN <= 48。
- TX <= 64K。
- FT <= 64K。
- ST 最好 < 200。

5.1.10.5. Message : CD (Channel Definition)

- Channel identifier : (channel # [NM<=4]) & (channel name [ST<=17]) 。
- Waveform source : 可以定義一個或兩個 source name , 以&隔開。
- Channel sensitivity & units : (sensitivity [ST<=20]) & (unit 1) & ...& (unit 6) , 也就是說 , 一個 sensitivity 與六個 units 。
- Channel calibration parameters : (channel sensitivity [NM<=20]) & (baseline [NM<=20]) & (time skew [NM<=20]) 。
- Channel sampling frequency : NM<=20 , Hertz 。
- Min ° & max ° data value : (min ° data value) & (max ° data value) , NM<=20。

p.7-109

HL7 v2.3.1: Control & Query

Message: CD (Channel Definition)
Min. & max. data value: Example

- 例如, $n=3$ bits $\rightarrow 2^n=2^3=8$
- 所以含0, i.e., 0~7共有8個值變

$H \rightarrow 2^{n-1}-1$ $B \rightarrow 0$ $L \rightarrow -2^{n-1}$	➔	$H \rightarrow 3$ $B \rightarrow 0$ $L \rightarrow -4$
$H \rightarrow 2^n-1$ $B \rightarrow 2^{n-1}$ $L \rightarrow 0$	➔	$H \rightarrow 7$ $B \rightarrow 4$ $L \rightarrow 0$

2001年1月26日星期五
臺大醫院資訊室 鄭伯璦 HL7
8

- Actual Single Amplitude
 - ✓ Actual Single Amplitude (A) : $A = S * C * (D - B)$ 。
 - ◆ S = 定義上的 normal sensitivity 。
 - ◆ C = actual sensitivity correction factor 。
 - ◆ D = value b'n (H , L) 。
 - ◆ B = actual baseline value 。
- Time Skew
 - ✓ 是指第 m 個 sample 於第 n 個 channel 之時間點 。
 - ✓ = (normal sampling time) - (actual sampling time) 。
 - ✓ = $R + (m-1)/f + (n-1)*t$ 。
 - ◆ R = 起始之時間點 。
 - ◆ $(m-1)/f$ = 第 m 個 sample 之時間點 。
 - ◆ $(n-1)*t$ = 第 n 個 channel 之時間點 。

5.1.10.6. RI (Repeat Interval)

- = (repeat pattern [IS]) ^ (explicit time interval [ST]) 。
- 定義排程時間間隔 。

5.1.10.7. TQ (Timing / Quantity)

Timing / Quantity (TQ)的 components 包括

- Quantity (CQ) = (quantity [NM]) & (units [CE]) 。
- Interval (CM)
 - ✓ = RI (Repeat Interval) , 定義排程時間間隔 。
 - ✓ = (repeat pattern) & (explicit time interval) 。
 - ◆ QnS 每隔 n 秒 。
 - ◆ QnM 每隔 n 分 。
 - ◆ QnH 每隔 n 時 。

(出國類別：研究)

- ◆ QnD 每隔 n 日。
- ◆ QnW 每隔 n 週。
- ◆ QnL 每隔 n 月。
- ◆ QnJm 每隔 n 週的週 m。
- Duration (ST) 。
- Start DT (TS) 。
- End DT (TS) 。
- Priority (ST) 。
- Condition (ST)，注意事項(服藥，打針...) 。
- Text (TX)，醫令。
- Conjunction (ST) 。
- Order sequencing (CM) 。
- Occurrence duration (CE) 。
- Total occurrence (NM) 。

5.1.11. HL7 之會議與認證

5.1.11.1. HL7 國際會議

- 2000/06/01 以個人名義加入 HL7 組織。
- 2000/09/11-15 參加年會與工作群組會議。
- 會議內容適合醫療資訊相關人員參與。
- 建議先加入會員再參加會議比較划算。
- 以機構名義加入可以彈性調派與會人員。
- 會議內容
 - ✓ Educational Sessions 。
 - ✓ Introductory Track 。

- ◆ Introductory Tutorial 。
- ◆ Implementation Class for Project Manager 。
- ◆ Implementation Class for Interface Engineer 。
- ◆ Overview of Financial Management 。
- ◆ Overview of Order Entry Tutorial 。
- ◆ 適合出入門者 。
- ✓ Message and More Track 。
- ◆ CCOW Tutorial 。
- ◆ Clinical Decision Support：System Architecture，Knowledge Representation and Arden Syntax 。
- ◆ Introductory Vocabulary Tutorial 。
- ◆ Advanced Vocabulary Tutorial 。
- ◆ Conformance Based Queries Tutorial 。
- ◆ Scheduling and Logistics Tutorial 。
- ◆ Security Tutorial 。
- ◆ 適合出進階者 。
- ✓ HL7 Toward the Future Track 。
- ◆ Version 3 Intermediate Tutorial 。
- ◆ Version 3 Advanced Tutorial 。
- ◆ XML and the PRA 。
- ◆ XML and the PRA in Depth 。
- ◆ Introduction to W3C XML Schema 。
- ◆ Reference Information Model (RIM) Tutorial 。
- ◆ 適合出進階者 。
- ✓ HL7 and Other Standard Track 。
- ◆ European Standards for Health Communication 。
- ◆ HL7 International Tutorial 。
- ◆ DICOM Tutorial 。
- ◆ X12 Tutorial 。

(出國類別：研究)

- ◆ HIPAA - What's Happening?
- ◆ HL7 and HIPAA , The Claims Attachment Recommendation 。
- ◆ 適合出進階者。

5.1.11.2. HL7 認證經驗

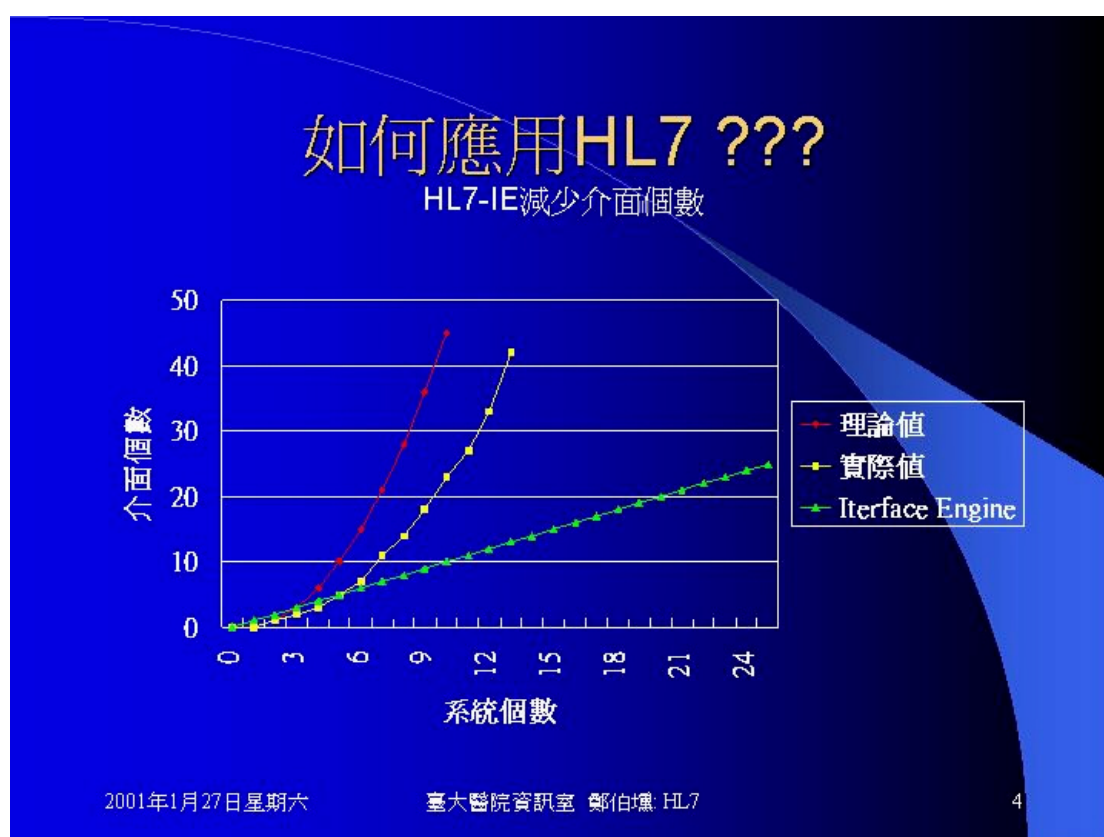
- 資訊來源：<http://www.hl7.org/>。
- 考期：每年有四次國際會議可以報考，每次會議有兩次考試時段，只能選一天考試，兩次考試期間通常會有提供認證課程。
- 也可以申請 On-Site 認證考試。
- 課程：提供 On-Site 認證課程 。
- 考題
 - ✓ 網路上有考題範本。
 - ✓ 考題為五選一，答錯不會倒扣。
 - ✓ 共七十題，兩小時內答對四十九題者過關。
- 考試內容
 - ✓ 第二章：Control 。
 - ✓ 第五章：Query 。
- 考試內容概說：一定要熟讀，甚至記下 Data Type 與 Message 。

5.1.12. HL7 之 Interface Engine

5.1.12.1. 如何應用 HL7 ???

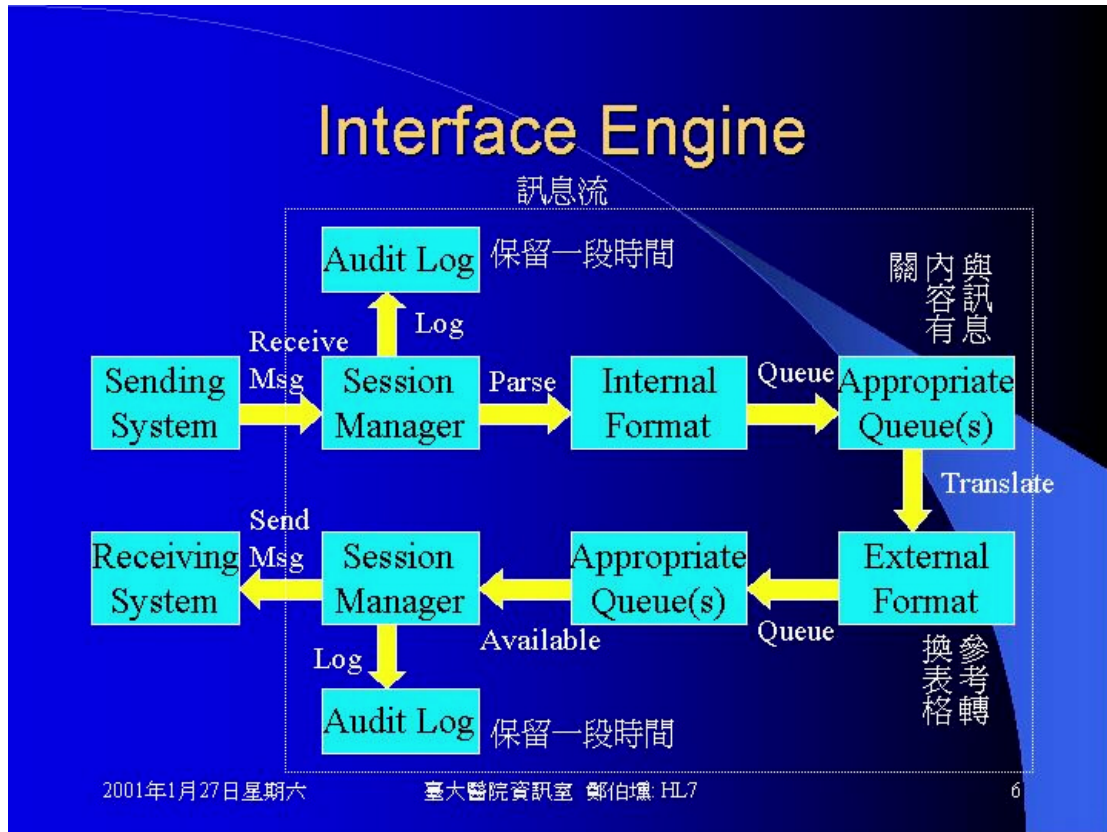
- Hospital Interfaces
 - ✓ 4 個子系統可以只有 3 個資料交換介面。
 - ✓ 15 個子系統需要???個資料交換介面。

- 問題：你有???個子系統，有???個資料交換介面。
- 方法：將所有子系統的資料交換全部透過 HL7 的標準介面來轉換，HL7 稱呼此介面為 Interface Engine (IE)。
- 結果：N 個子系統只有 N 個資料交換介面，所以 HL7-IE 會減少介面個數。

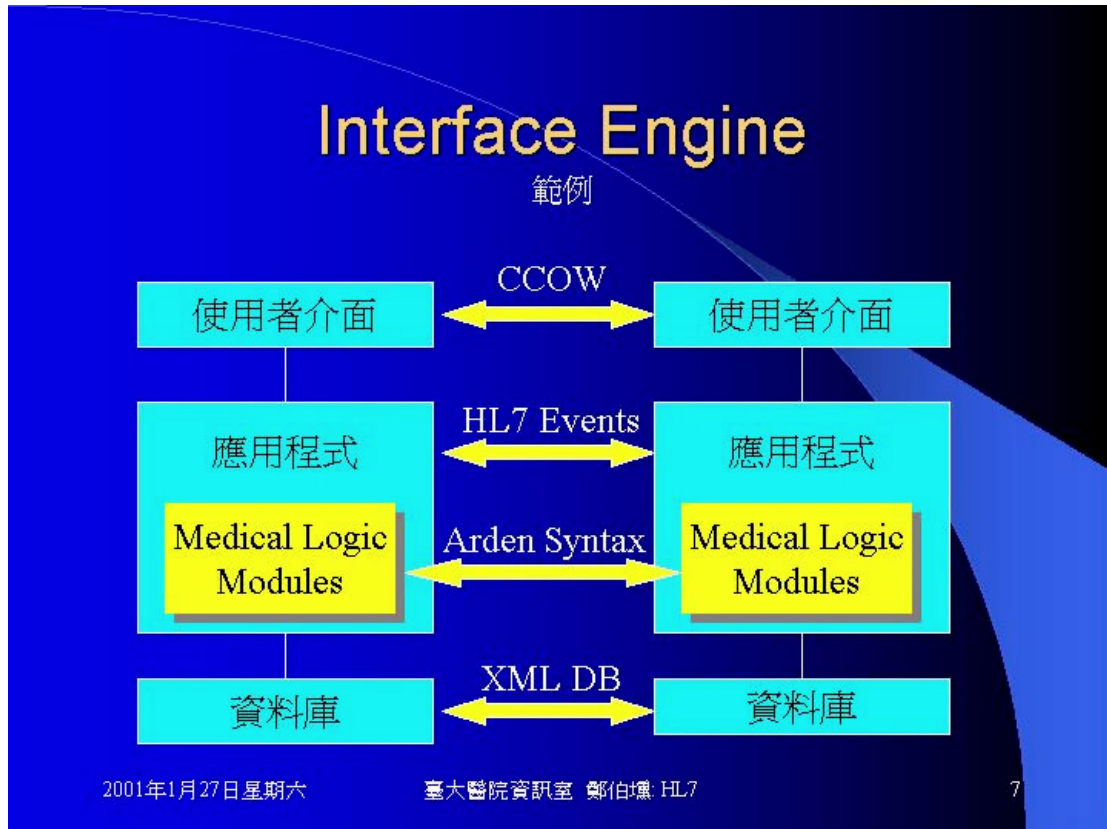


- 選擇指南
 - ✓ 先選擇重要系統加以整合，如醫令。
 - ✓ 容易上手，增加產能。
 - ✓ 提供強大的 Mission-Critical Message 功能。
 - ✓ 系統易於擴充。
 - ✓ 系統往前相容。
- IE 訊息流

(出國類別：研究)



● IE 範例



5.1.12.2. 深入了解 IE

- Interface Engine 功能
 - ✓ Support development , test and production 。
 - ✓ Internal "loop-back" 。
 - ✓ New interfaces should not affect production 。
 - ✓ Different modes : real-time , batch , unsolicited , query & response 。
 - ✓ Store and forward 。
 - ✓ Non-messaging interfaces : FTP , Kermit , X/Y/Z Modem , ODBC , etc 。
- Interface Engine 選擇準則
 - ✓ 醫療資訊系統裡的三大抉擇，包括
 - ◆ Interface Engine
 - ◆ Data Repository
 - ◆ User Interface
 - ✓ 真正支援最新 HL7 標準。
 - ✓ 容易上手使用。
 - ✓ 有詳細的規格。
 - ✓ 容易擴充。
 - ✓ 支援 High Availability 。
 - ✓ 提供自行設定表格與對應碼。
 - ✓ 容易組態且具有彈性。
 - ✓ 不是第一個試用者。
 - ✓ 廠商是 HL7 的組織成員。
 - ✓ 提供錯誤訊息檔案功能。
 - ✓ 容易啟動與停止介面功能。
 - ✓ 提供介面當機自動回復功能。
 - ✓ 可以限制資料的所有者。
 - ✓ 具有非常強壯的系統技術架構。

(出國類別：研究)

- ✓ 具有智慧型的錯誤處理機制。
- ✓ 從既有廠商中選擇加入 HL7 轉換模組。

深入了解 IE

Interface Engine 選擇準則

- 從既有廠商中選擇加入 HL7 轉換模組

```

    graph LR
        PharmacySystem[Pharmacy System] <--> PharmacyHL7[Pharmacy HL7 Translator]
        LabSystem[Lab System] <--> LabHL7[Lab HL7 Translator]
        PharmacyHL7 <--> IE[Interface Engine]
        LabHL7 <--> IE
    
```

2001年1月27日星期六 臺大醫院資訊室 鄭伯璦 HL7 12

- Interface Engine 廠商與產品

- ✓ STC : DataGate , <http://www.stc.com/> 。
- ✓ HIE : Cloverleaf , <http://www.hie.com/> 。
- ✓ SMS : OPENlink , <http://www.smed.com/> 。
- ✓ eWebIT Solutions : eLink , eView , eSign , ePI ,
<http://www.ewebit.com/> 。
- ✓ LINK Medical Computing : Interface Development Kit ,
<http://www.linkmed.com> 。
- ✓ CAI : Transaction Distribution Manager (TDM) ,
<http://www.cainc.com/> 。
- ✓ Cerner Corporation : Open Engine 。
- ✓ Neon , <http://www.neon.com/> 。
- ✓ Healthcare.com , <http://www.healthcare.com/> , Cloverleaf 。
- ✓ Eclipsys , <http://www.Eclipsys.com/> 。

- Interface Engine 相關問題
 - ✓ 在台有無分支機構?
 - ✓ 系統是否充分支援 Multi-byte?
 - ✓ 如何教育訓練?
 - ✓ 如何解決線上問題?
 - ✓ 從何下手?

5.1.13. HL7 系統觀摩

內容包括以下美國醫療機構，

- PAMF, Palo Alto Medical Foundation (2000/05/08)。
- Stanford Hospital (2000/05/12)。
- UCSF Hospital (2000/05/12)。
- Univ. Hospital of Columbia & Cornell (2000/09/13)。
- U. of Virginia Hospital (2000/09/15)。
- UCLA Mattel Child Hospital (2000/09/17)。
 - ✓ Medical Information Services Department。
 - ◆ Richard, CIO。
 - ◆ Timothy Carison, Programmer Analyst IV。
 - ◆ Spencer Sun, System Administrator。
 - ✓ STC: DataGate。

5.1.13.1. Palo Alto Medical Foundation

- 組織
 - ✓ PAMF, <http://www.pamf.org>
 - ◆ Brown & Toland Medical Group。
 - ◆ CHRC, Community Health Resource Center。

(出國類別：研究)

- ◆ CPMC , California Pacific Medical Center 。
- ◆ UCSF Medical Center 。
- ✓ 人員
 - ◆ Steven Lan , MD , Medical Director , Medical Information Department 。
 - ◆ Paul Tang , MD 。
- 工作內容
 - ✓ Health Risk Approval 。
 - ✓ Email Guideline --> Patients --> Clinician 。
 - ✓ Population Management 。
 - ◆ Disease Management 。
 - ◆ Resource Management 。
 - ◆ 資料來源：Patient Responses , Evidences 。
- Internet
 - ✓ Goal
 - ◆ Administration Cost：減少 10% (US\$3M) 。
 - ◆ Doctor Office：減少 US\$30M 。
 - ✓ 整體成本約減少 30~50% , 包括 Health Plan , Doctor , Hospital , Vendor 。
 - ✓ 5 大成本(RACER)
 - ◆ Referral (轉診) 。
 - ◆ Authorization (醫師授權) 。
 - ◆ Claims (處方) 。
 - ◆ Eligibility (醫師證書) 。
 - ◆ Report (報告) 。
- 系統
 - ✓ Administration System：IDX 公司 。
 - ✓ On-Line Systems：EpicCare 公司 , MS-Access , Oracle 。
 - ✓ Data Warehousing：SQL Server --> Report 。

- ✓ 電子病歷(EMR)：Sutter Health 公司。
 - ◆ 1999/06：Schedule，Transcription，abstraction。
 - ◆ 1999/09：Prescribing，Message，HM alerts。
 - ◆ 1999/12：Charting Tools，Decision Support。
 - ◆ 2000/05：Order，Billing，Result。

5.1.13.2. Stanford Hospital

- 系統
 - ✓ 1999/06 開始與 UCSF 醫院之資訊系統分家。
 - ✓ 使用 IDX 系統。
 - ✓ 5 大部分
 - ◆ Core Vendor。
 - ◆ System Integration。
 - ◆ People & Processes。
 - ◆ Data Warehousing。
 - ◆ Infrastructure。

(出國類別：研究)

HL7系統觀摩

Stanford Hospital

住院床數	1200床
醫師人數	300人
Technical Mapping	1.3M美金/年
Project Management	25.3M美金/年

2001年1月27日星期六
臺大醫院資訊室 鄭伯璦 HL7
10

- EMR Ready Characteristics
 - ✓ Standard Compliant ◦
 - ✓ Scalable ◦
 - ✓ Upgradeable ◦
 - ✓ Security ◦
 - ✓ Backup Strategic ◦
- 未來展望
 - ✓ Web-based Technology ◦
 - ✓ ASP (Application Service Provider) ◦
 - ✓ Mobile Computing ◦
 - ✓ Consumer Demands ◦

HL7系統觀摩

Stanford Hospital

- **Rex Healthcare系統**
 – Raleigh, NC, <http://www.rexhealth.com>

2001年1月27日星期六 臺大醫院資訊室 鄭伯璦 HL7 13

5.1.13.3. UCSF Hospital

- 部門：Medical Knowledge Information Services。
- 人員：曹式孟，Sherman Tsao，Operations Director，Stokes & Co.，L.L.C.，stsao@attglobal.net。
- 系統：IDX，1999/06 開始與 UCSF 醫院之資訊系統分家。

以下是 2000/05/12 當天參觀 UCSF 電子處方系統的描述：其電子處方系統是最近才更新且使用 Microsoft Windows NT 環境，該系統直接與美國之 First Data Bank 連線，而藥師可以直接使用三種不同的功能。

- 第一個功能畫面叫做 Allergen Master，允許藥師檢視病友對於使用任一種藥物的過敏性反應。
- 第二個功能畫面叫做 Drug Master，允許藥師檢視所建議的安全劑量及任一種特殊藥物的使用頻率，以便使得病友不會服用過多或過少的藥物。
- 第三個功能畫面叫做 Interaction Master，允許藥師檢視使用在病友上的

(出國類別：研究)

任何一種藥物對其他藥物的交互作用反應。

- 第四個功能畫面叫做 Diagnostic Master，並未被完整的介紹到且該院藥師指出尚未用到該項功能。

基於上述三種功能，藥師可以檢視較緊急的資訊，例如，藥物敏感性試驗，是否醫師開立了過量的藥物，或者任一種藥物交互作用於某一種特定的藥物上。如果要是發現某種藥物的使用有上述任一種衝突的資訊存在，藥師必須與醫師進一步的確認用藥，如果醫師堅持其所開立的處方內容，則藥師可以隨後 Override 該系統並且分配該藥物，而同時此資訊將會被系統監控且紀錄在系統中。很意外的是在該樓層工作的醫師並無法直接使用該系統，反而只能夠間接使用 Micromedix 來檢視藥物的相關資訊。

當藥師將醫師所開立的處方簽內容輸入系統後，藥物會自動被送至 Auto Dispensing System，此系統狀似一個具有七至九各抽屜的緊急用推車，抽屜中有個特定護理站所需使用的所有藥物，此緊急用推車也被放置在醫院的每一個護理站。所以當某一病友需要特定藥物時，護士人員必須登入系統中並從緊急用推車上取得所需的藥物，此系統的優點是護理站的護士已經有所有所需的藥物且不必等候特定的處方藥物。順帶一提的是所有的麻醉藥物都會被精細的分配與手工紀錄下來。結論是此舉將可以更有效率的將病友的藥物在用藥時間內快速且精準的送達，此外，對任一個病友的每一個用藥內容將會被精確的計價而且當並有出院時不會有藥物會被退回藥局。總之此系統有許多的優缺點，而此系統不可能毫不修改的直接使用在現在的台灣各醫療機構中。

5.1.13.4. University Hospital of Columbia & Cornell

- 部門：New York Presbyterian Health Network。
- 人員：Virginia Lorenzi，vlorenzi@fcg.com。

HL7系統觀摩

University Hospital of Columbia & Cornell

醫院數目	Hospitals	32家
15家贊助醫院床數	Sponsored Hospitals	5,772床
13家結盟醫院床數	Affiliated Hospitals	3,592床
專業機構	Specialty Institutions	3個
遠距看護設備	Long-term Care Facilities	8套
家庭醫學代理商	Home Health Agencies	11個
衛星重點看護中心	Satellite Primary Care Center	97個
醫師群組	Physician Groups	12個
管理看護實體	Managed Care Entities	4個

2001年1月27日星期六
臺大醫院資訊室 鄭伯璦 HL7
16

- 不是在整合介面，因為我們
 - ✓ 想要取得系統間最好且有意義的交換資料(Best of Breed)。
 - ✓ 不可能找到一個可以達成所有事情的介面。
- HL7
 - ✓ 讓介面工作更簡單。
 - ✓ 並不會解決所有的問題。
 - ✓ 介面的專案性工作仍舊需要。

5.1.13.5. University of Virginia Hospital

- 部門：DT7 Software。
- 人員：David John Marotta，marotta@stanfordalumni.org。

(出國類別：研究)

HL7系統觀摩

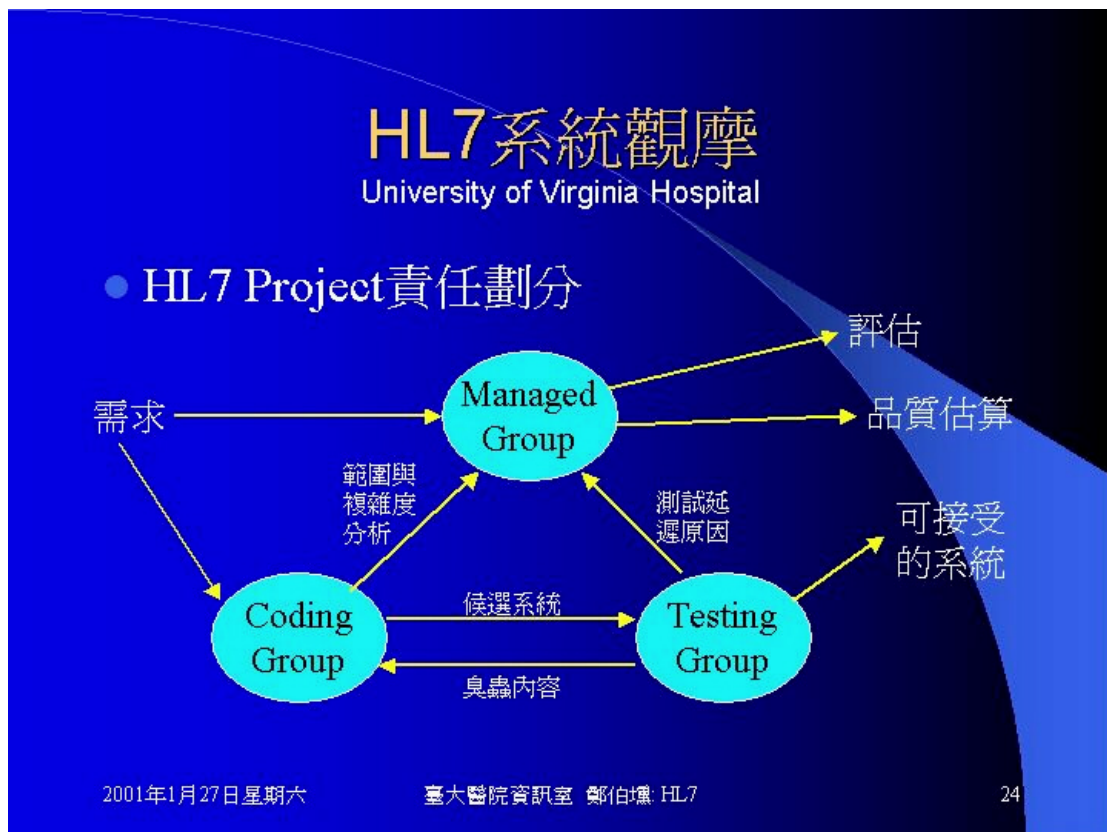
University of Virginia Hospital

住院床數	600床
住院病友數目	27,000人/年
住院病友平均住院日	6.2天/人
門診病友數目	420,000人/年
急診病友數目	60,000人/年
有效檢驗量	1,500,000件/年
作業成本	400,000,000美金/年

2001年1月27日星期六
臺大醫院資訊室 鄭伯璦 HL7
19

- 5 P's
 - ✓ Paper how it's supposed to work?
 - ✓ Process what is supposed to make it work?
 - ✓ Politics who can make it fail?
 - ✓ Practice what really makes it works?
 - ✓ Problems why it does not work?
- 決定 HL7 Interface Engine 應用範圍。
 - ✓ 很能估計真實的應用範圍。
 - ✓ 對未來介面的快速成長要有心理準備。
 - ✓ IE 會吸引更多的介面。
- 招募與培訓 HL7 介面專家。
 - ✓ 技術面
 - ◆ 知道需要用的工具與技術。
 - ◆ 知道軟硬體的技术瓶頸。
 - ◆ 要精於系統分析，尤其是細部規格。
 - ◆ 最好是 HL7 專家。

- ✓ 功能面
 - ◆ 醫療經驗。
 - ◆ 知道介面(能夠/不能夠)提供的功能。
 - ◆ 可以決定使用者的真正需求。
- ✓ 溝通技巧：介面的開發是一種社交經驗與行為。
- HL7 Project Planning
 - ✓ Scope。
 - ✓ Analysis。
 - ✓ Communications。
 - ✓ Sending System。
 - ✓ Receiving System。
 - ✓ Interface Engine。
 - ✓ Testing。



(出國類別：研究)

5.1.13.6. UCLA Mattel Child Hospital

- 醫院與資訊室不在同一地點。
- 該資訊室連接兩家醫院
 - ✓ UCLA Mattel Child Hospital。
 - ✓ Santa Monica SPD Hospital。
- 病床約有 500 床。
- 使用 HL7 v2.1。
- HL7 使用 STC 之 DataGate 軟體，在 Sun SPARC 機器上執行，沒有備援方案。
- 部門：Medical Information Services Department。
- 人員
 - ✓ Robert Konishi，CIO，children@mednet.ucla.edu。
 - ✓ Timothy Carison，Programmer Analyst IV。
 - ✓ Spencer Sun，System Administrator。
- 系統：STC 之 DataGate。
- 子系統至少包括
 - ✓ 醫令：IBM Mainframe。
 - ✓ 藥局：IBM RS/6000，BDM Pharmacy。
 - ✓ 檢驗：DEC Alpha，Meditech Lab。
 - ✓ 報告瀏覽：Sequent，Clinicomp。
 - ✓ 病理：IBM RS/6000，Comed Pathology。
- STC 公司產品
 - ✓ Interface Engine：DataGate，<http://www.stc.com/>。
 - ✓ 未開發亞洲市場，故 Multi-byte 問題尚未釐清。
 - ✓ DataGate Interface Engine
 - ◆ 可以視為一個 Gateway。
 - ◆ 所有系統間的資料交換全部要經過它。
 - ◆ 至少有兩種作業方式。

以下是申請參觀該院的信件內容，

Dear Prof.,

I act as a visiting scholar in US and am a system team leader at Medical Information Systems Department, National Taiwan University Hospital (NTUH). I have been interested in the field of hospital Laboratory Information System (LIS) for 3.5+ years and am working and leading a four-team-member for developing and maintaining our LIS. Meanwhile, we focus on the HL7 standard implementation for Electrical Patient Records (EPR) and try to integrate Health Level 7 (HL7) related software into our legacy systems.

I am an HL7 individual membership and seeking an experience (NOT FOR CREDIT) for observing, even learning, how to implement or integrate HL7 related tools in your hospital legacy systems for 1+ weeks. I also know that both of our hospitals (universities) maintain an alliance relationship of international medical student exchange program right now.

Actually, my government will fully support me the expenses for this international practice. Is there a related program available where I might be able to follow around learning new HL7 techniques?

Currently, I act as a visiting scholar at Computer Science Department, University of California, Riverside (UCR) and live on-campus now. I will stay in US until 30th Oct., 2000 and return to Taiwan. I am learning and studying the Data Warehousing/Mining at UCR with Prof. Dimitris Gunopulos. I will attend the HL7 2000 annual meeting and workshop meeting at St. Louis, MO, from 10th to 15th Sept, 2000. If possible, I hope you could offer me one learning opportunity about HL7 real-world implementation or integration at your site after 16th Sept.

My undergraduate major is Information and Computer Engineering in Taiwan. I spent 1.5 years and earned an Electrical Engineering master degree at University of Southern California (USC) in 1992 spring. I would like to learn some of the HL7 professional skills in US. In addition, I can learn more about the field of HL7 from an international perspective. I believe my experiences in your hospital will increase my understanding of HL7 and improve Taiwan healthcare systems as well as myself.

I am excited by the possibility of working in conjunction with the your hospital. I can be contacted through E-mail at pcheng@cs.ucr.edu and cph@ha.mc.ntu.edu.tw.

(出國類別：研究)

Thank you for your attention and I look forward to hearing from you.

Sincerely yours,

Po-Hsun Cheng
 Application Engineer (Level 3)
 Medical Information Systems Department
 National Taiwan University Hospital
 Taipei, Taiwan, R.O.C.

以下是整個參觀過程的描述：今早(2000/09/17)開了 4 小時的車(因為塞車)到 UCLA 兒童醫院資訊單位參觀，總共見到了兩位工程師，

- Timothy Carison，Programmer Analyst IV。
- Spencer Sun，System Administrator。

從上午 10：30 開始至下午 1：30，總共討論了約三個小時，席間並實際參觀該醫院醫療系統 STC 控制平台，該平台是在 UNIX 上以 X Window 方式操作的。同時也向對方要了一些系統架構圖，各子系統在 STC 控制平台下的部分 Log 檔案，當然短短的三個小時無法詳細的知道該控制平台的內容，據說該平台的學習曲線很高，但是下一個版本聽說使用者介面會好一些。好玩的是她們的 STC 控制平台竟然沒有備援方案，不過據說上線好幾年了還沒有死當過ㄟ。

此外該資訊室與醫院有一段空間上的距離，好像公館分院與台大本院的模式，而該院約 500 多床，也擁有 IBM 主機，更與臺大醫院一樣使用 PCS 程式，當然也有一些開放系統，分別是 lab，pharmacy，radiology，pathology 等系統。

而 STC 的控制平台叫做 DataGate，所有的系統全部都會連接到該平台上，也就是說不管是否是開放系統或是 IBM 主機系統，資料的傳送與接收都會透過 DataGate 做 Gateway。而每一個醫療子系統都需具有 Translator 與 Interface 兩個模組，Translator 可以買也可以自己寫程式，Interface 是與 DataGate 之間的通訊模組，以下舉個小例子說明：

例如從 IBM 傳送資料給 Lab



- T 表示 Translator。
- I 表示 Interface。
- 每個醫療子系統在傳送資料時要先轉換資料格式。
 - ✓ IBM 可以是 2k 以上的長字串。
 - ✓ 開放系統可以是自訂的資料格式或是使用 HL7 標準所定義的資料格式。
- 轉換資料格式後再透過 Interface 與 DataGate。
- DataGate 會自動 Routing 資料封包到 Lab 系統。
- Lab 系統透過 Interface 接收資料之後在轉換成她自己所認得的資料格式。
- 當然子系統可以自行設計程式也可以購買。

Timothy 分析師建議我們從 ADT 開始下手，因為這是最複雜的一環，ADT 也就是我們所謂的醫令系統，而工作的流程是先 mapping 資料流格式(一個人月)，接下來就和資訊工程的上線流程相似。目前她們也不知道哪些廠商支援 Multi-byte 轉碼功能，此外目前她們是使用 HL7 v2.1 的標準，而目前 HL7 最新的標準版本是 v. 2.3.1，預計明年 v2.4 會通過 ANSI。

目前她們與 Radiology 子系統連接並不傳送與接收影像的資料，而只是傳收單純的文字資料，而 HL7 v2.3.1 可以傳收影像資料，這一點是她們未來要做的。廠商除了 STC 之外，她們還建議了一些其他 Interface Engine 的相關廠商，包括 Neon，Healthcare.com，Eclipsys 等。

此外，她們將 STC 與 UC-San Francisco 醫院所用的 IDX 產品做了一個初略的比較，總之 IDX 的產品將 Interface 與 Translator 綁的比較緊，也就是說比較不像 STC 的產品 Independent。所以站在分析師的角度來看，STC 的產品比較合乎 HL7 未來 Plug-and-Play 的功能性需求定義。

5.1.13.7. 有關 HL7 的建議

- HL7 標準資訊交換應用。
 - ✓ 去除院際間病歷流通阻礙。

(出國類別：研究)

- ✓ 降低院內病歷資料交換的困難度。
- ✓ 降低醫療資訊系統的複雜度。
- ✓ 加速建置其他醫療系統。
- ✓ 也可改善醫療資源重複浪費之缺失。
- ✓ 將進一步活化電子病歷。
- HL7 是電子病歷的基礎工程，障礙包括既有系統轉換能力，醫師接受程度，院際間的門戶之見等。
- 應由主管單位(衛生署)主導，(如日本)。
 - ✓ 讓各醫院齊步走入 HL7 領域。
 - ✓ 團結才有力量。
 - ✓ 各醫院將互信互利。
 - ✓ 節省未來不必要的醫療資訊系統投資與浪費。
 - ✓ 加速本國醫療系統國際化的腳步。
- 宜慎選跨國性且充分支援 HL7 最新標準之廠商，甚至應鼓勵台灣本土 IT 廠商與美國相關廠商結盟互惠。

5.1.13.8. 我對 HL7 的思考方向

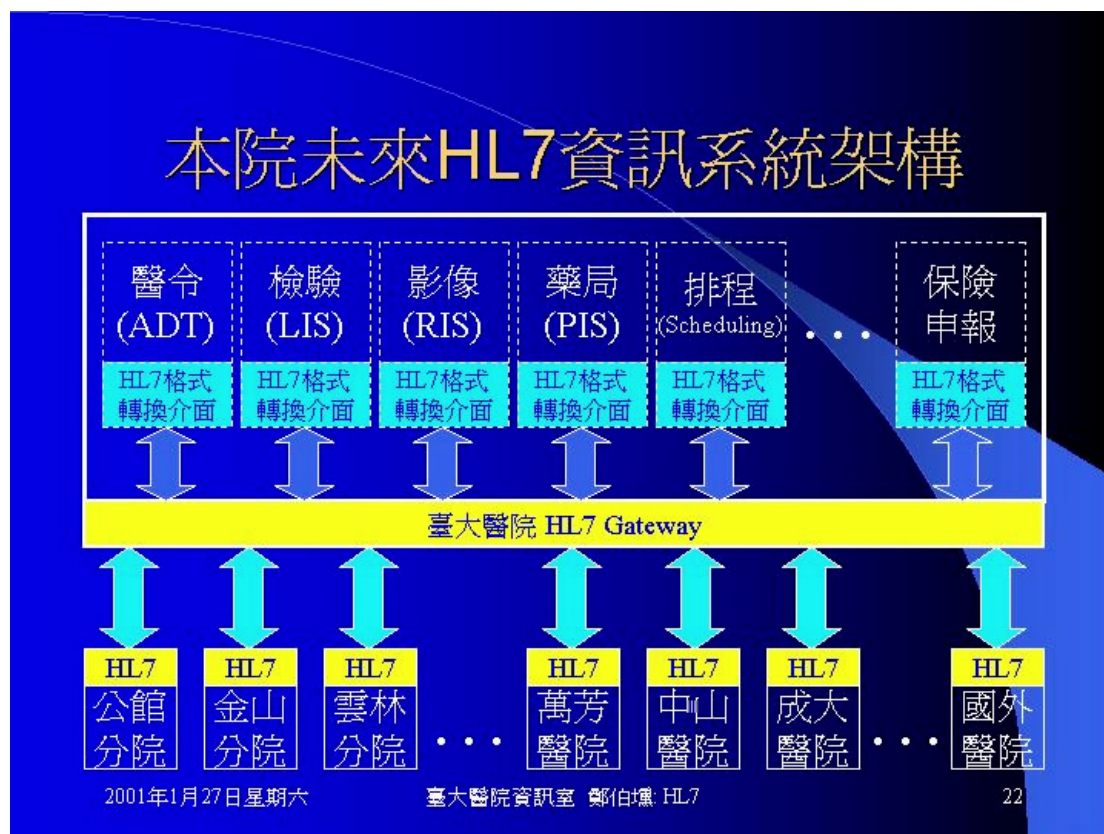
- 目前台灣與美國醫療資訊系統之建構過程迥異，如何快速且有效的移植 HL7 標準，還有賴 IT 人員三思?
- 如何取得充分的 HL7 成功建構資訊?失敗案例?(如，Per-se 日本經驗)
- 如何評估 HL7 建構所帶來的影響與效能?
- 如何透過 HL7 的建置，快速且有效的提昇進而鞏固本院的醫療資訊地位?

5.1.13.9. 資訊室對 HL7 的因應計劃

- HL7-Taiwan

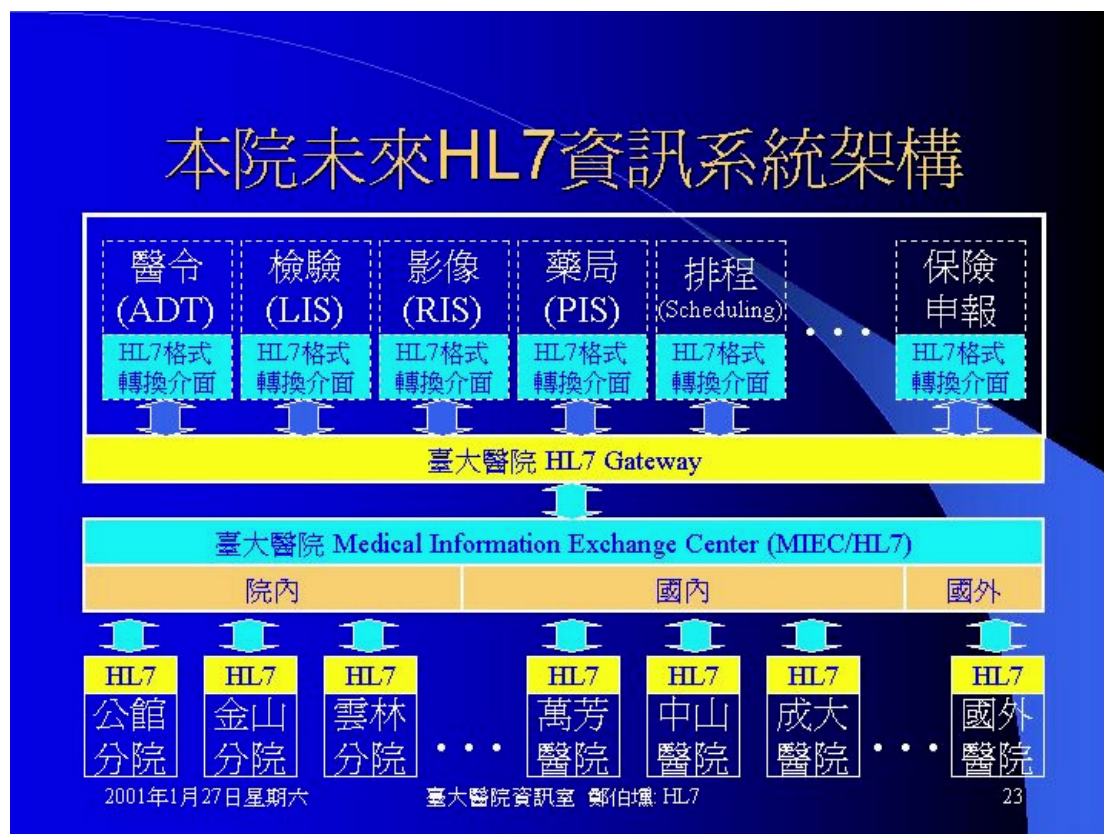
- ✓ 本室賴主任 2000/11 成為電子病歷 工作小組(EPR Working Group) 召集人。
- ✓ 2000/11/24 針對本院召開電子病歷研發專案協調會議。
- 加速推動本院電子病歷建置計劃。
 - ✓ 配合建置全院電子病歷系統，並成立特攻小組。
 - ✓ 資訊室內部組織再造，增加本室 HL7 認證人數。
 - ✓ 成為本院／台灣／亞洲醫療資訊交換中心(MIEC，Medical Information Exchange Center)。
- 持續推動本院電子病歷，相關推動單位至少包括
 - ✓ 台大醫學資訊社。
 - ✓ 台大醫學院醫療資訊組。
- 台大總院
 - ✓ 建置 HL7 標準化測試系統。
 - ✓ 建置 HL7 標準化正式系統。
 - ◆ 短程目標：與公館分院 HL7 連線。
 - ◆ 中程目標：與國內友院 HL7 連線(萬芳，中山，成大等)。
 - ◆ 長程目標：與國外友院 HL7 連線(美國，澳洲等)。
 - ◆ 公館分院：配合 HL7 標準化系統上線。
- 本院未來 HL7 資訊系統架構

(出國類別：研究)



● 小結論

- ✓ 醫療資訊交換中心(Medical Information Exchange Center, MIEC)
 - ◆ 類似金融資料交換中心，各銀行都透過其交換資料(轉帳，跨行提款等)，並酌付過路費。
 - ◆ 建議成立基金會加以發展與管理。
 - ◆ 不要總是紙上談兵，要積極的建置與應用，否則新科技與新標準的價值性將會一去不返。



- ✓ 如果政府短期內沒有政策與目標，國內誰先成為 MIEC，誰就是未來國內醫療資訊的龍頭。
- ✓ 路上的石頭：中文碼標準？
- ✓ HL7 不是萬靈丹。

(出國類別：研究)

5.2. 資料倉儲與資料採礦心得

5.2.1. 進修資訊

- 學校：University of California，Riverside。
- 科系：Computer Science。
- 指導老師
 - ✓ Prof. Dimitris Gunopulos (Data Mining)。
 - ✓ Prof. Suntruis (Data Warehousing)。
- 進修期間：2000/05/04~11/03。
- 進修主題：Non-Visualized Data Mining (Clustering，Decision Tree，Neural Network)。
- 指導老師研究方向
 - ✓ Prof. Dimitris Gunopulos (Data Mining)
 - ◆ 以前
 - Regression Analysis。
 - Time Series Similarity Models。
 - Finding Associations in the spatial domain。
 - Finding Associations in the temporal domain。
 - ◆ 目前
 - Finding frequent spatiotemporal evolution patterns。
 - ◆ 應用範圍
 - Epidemiological Data。
 - Earth Science Data。

5.2.2. 資料倉儲(Data Warehousing)

主要內容：定義 DW 開發與設計的方法與主要特性，了解不同 DW 架構模組間的異同點，DW 人員實際的建置與維護工作。

5.2.2.1. DW 定義

- DW = Data Warehousing = 資料倉儲。
- DW 鼻祖：Bill Inmon。
- "DW 是一業務導向，整合性，隨時間序列變動且唯讀之大量歷史資料庫，目的在支援管理決策之用"。
- 收集並轉換組織內各種資料至單一資料庫(DW)。
- 定義資料於業務應用上的內在意義。
- 提供快速且多樣化的資料存取方法。
- 提昇而非取代現有系統效能。
- 一種資料流程整合與再利用的工作。
- 需經由一連串的建置活動始可完成，而非隨插即用的產品。

(出國類別：研究)

DW定義

Operational	資料較細微, 每天收集資料, 目前的資料, 被高度存取, 應用程式導向	某人目前的信用額度為何?
Atomic / DW	資料更細微, 資料依時間而異, 整合性的資料, 科目導向, 部分屬於結論資料	某人以往的信用狀況如何?
Departmental	部門性的資料, 部分資料是推導來的	我們是否吸引更多的使用者?
Individual	暫時性的資料, 嚐試錯誤導向, 非重複性的資料,	我們在分析哪一類的客戶?

2001年1月26日星期五
臺大醫院資訊室鄭伯堉 資料倉儲
8

5.2.2.2. 為何需要 DW ???

- 使用者需求的改變
 - ✓ 使用者越來越聰明了。
 - ✓ 都想要 "更快, 更好, 更多, 最好不要成本"。
 - ✓ 需要即時得知 Free Query 結果。
 - ✓ 需要知道客戶的需求在哪裡。
 - ✓ 策略聯盟。
 - ✓ 需要馬上知道市場的變化。
 - ✓ 競爭對手變化速度比自己快速。
 - ✓ 產品附加價值觀點：從產品變成每次交易的行為。
- 改善既有組織的痛處
 - ✓ 因為各自獨立的應用系統會有以下狀況。
 - ◆ 存取資料困難。

- ◆ 不同的作業平台。
- ◆ 不同的資料格式。
- ◆ 大量的重複性資料。
- ◆ 不良的資料模型與定義
- ◆ 大量的資料轉換與擷取程式。
- ◆ 多種異質性資料庫並存。
- ◆ 資料內容 / 格式不一致 --> 不同報表不同結果。
- ✓ 強大的競爭壓力。
- ✓ 專注在使用者的需求上。
- ✓ 組織重整與改造。
 - ◆ 分散式決策需求。
 - ◆ 需要不同部門的資料，需要剪貼別處的資料。
- ✓ 市場區隔。
- ✓ 絕大多數的資訊系統是報表導向的。
- ✓ 針對新的需求常需要重新分析。
- ✓ 傳統 RDBMS 可行之方法幾乎用盡，包括
 - ◆ Pre-Tune query。
 - ◆ Create more indexes。
 - ◆ Create more summary tables。
 - ◆ Create new database schema designs。
 - ◆ Parallel processing。
 - ◆ Clustering。
 - ◆ Increase hard disk。
 - ◆ Increase memory。
 - ◆ Increase processors speed。
 - ◆ Increase number of processors。
- OS 的工作
 - ✓ OS 可以做到
 - ◆ 處理訂單。

(出國類別：研究)

- ◆ 找出年度最佳銷售員。
- ◆ 計算流水帳。
- ✓ OS 無法做到
 - ◆ 誰是我最大的客戶？
 - ◆ 為何他是年度最佳銷售員？
 - ◆ 如果銷價競爭會賺更多錢嗎？
- 投資報酬率 (ROI)
 - ✓ 馬上回收
 - ◆ 資料一致性的定義。
 - ◆ 單一資料。
 - ◆ 改善準確度。
 - ◆ 改善資料處理時間。
 - ◆ 提供細微的資料。
 - ◆ 提供歷史資料。

為何需要DW ???

投資報酬率 (ROI)

- 1996 IDC ROI Study

	Data Mart	Enterprise
成本回收	1.1年	-3年
投資報酬率	400%	301%
平均成本	US\$1.3M	US\$2.2M
風險評估	Low	High

2001年1月26日星期五
臺大醫院資訊室鄭伯壘資料倉儲
16

- 效益
 - ✓ 將不同資訊系統之資料加以整合。

- ✓ 對公司資訊提供一個一致性的版本。
- ✓ 將資料轉換成有價值的資訊。
- ✓ 提供 Knowledge Management 技術平台，讓需要資訊的人，適時掌握所需的資訊。
- DW 的存在價值
 - ✓ 太多的資料，太少的資訊。
 - ◆ 資料太多無法快速回應。
 - ◆ 對以往未預期到的問題需要較複雜的分析。
 - ◆ 資訊系統一直在變更中。
 - ◆ 主管需要線上即時決策。
 - ✓ 成長趨勢
 - ◆ 85%：美國五百大企業。
 - ◆ 1999：US\$10B。
 - ◆ Data Mart 方法不斷擴充。
 - ◆ DW 產業一直在整合擴張。
 - ◆ 改善取得資料的過程，因 OLAP 會干擾組織內線上的作業系統，影響 OLAP 的功能與使用意願。
 - ◆ 簡化資料收集的難度。
 - ◆ 擴大資料分析的功能，傳統決策通常來自 Summary 資料，決策品質無法精緻化。
 - ◆ 減少 Turn Around Time，因傳統決策支援方式無法滿足決策部門資訊需求之倍數成長與即時提供的要求。
 - ◆ 提供正確的資料，傳統決策通常來自 Summary 資料，決策品質無法精緻化，Randy Mott，CIO，Wal-Mart，Summary Date --> Average Information --> Average Decisions --> Average Performance --> Average Customer Feedback Acknowledge。

(出國類別：研究)

5.2.2.3. 如何成功建置 DW ???

- 投資者的決心 (Executive Sponsorship)
 - ✓ 專案負責人 (Project Champion(s)) 。
 - ✓ 管理委員會 (Advisory Board) 。
 - ✓ 期望什麼 (Expectations) 。
 - ✓ 有效的溝通 (Communications) 。
- 要解決的問題 (Business Case)
 - ✓ 為何要建置 DW ???
 - ✓ 對組織的價值何在 ???
 - ✓ 投資報酬率 ???
 - ✓ 未來是不是一顆炸彈 ???
- 充分定義需求 (Well-defined Requirements)
 - ✓ Detail data 。
 - ✓ Summary data 。
 - ✓ History 。
 - ✓ Refresh 。
 - ✓ Representative queries 。
 - ✓ Canned reports 。
 - ✓ Archive 。
 - ✓ Performance 。
 - ✓ Technical support 。
- 問題的大小範圍 (Subject Area(s))
 - ✓ 一次只解決一個問題，不要貪心 。
 - ✓ 第一個 DW 最好是小的問題 。
 - ✓ 一次 DW 專案叫做一個 Roll 。
- 計劃的角色扮演與責任歸屬 (Project Roles and Responsibilities)

如何成功建置DW ???

計劃的角色扮演與責任歸屬

組織與使用者	資訊人員
說明需要的資料	建立Data Models
定義Data Archiving的需求	提供Data Archiving
提供Data Sources	指明Data Sources
指明與解決Data Integrity問題	指明與解決Data Integrity問題
定義可接受的系統回應時間	系統調整至可接受的時間範圍
將查詢最佳化	建議最佳化查詢的技巧
定義與開發專家系統應用程式	協助與開發專家系統應用程式
將資料定義與文件化	將資料文件化
訓練使用者	訓練使用者

2001年1月26日星期五
臺大醫院資訊室鄭伯堦 資料倉儲
26

- ✓ IT 資訊成員
 - ◆ IT Executive
 - Provides resources , monitors project progress , sets priorities , resolves conflicts , integrates project w/in IT .
 - ◆ Data Guardian(s)
 - Owners of the physical implementation of the organizations's data .
 - Resolves data integration issues .
 - Implements data security .
 - Documents data definition , calculations , summarizations .
 - Maintains and updates business rules .
 - ◆ DW Project Manager
 - Managing the project schedule and effort .
 - Tracking and reporting progress and changes .
 - Resolving conflicts .
 - ◆ Project Quality Manager

(出國類別：研究)

- Verifying the processes have been defined for ensuring quality of deliverables ◦
- Performing project quality reviews ◦
- ◆ DW Architect
 - Leading the DW architecture efforts ◦
 - Synthesizing business , data , and architecture requirements into an architecture design ◦
 - Developing project plans for the architecture analysis , design , and implementation activities ◦
- ◆ Data Architect
 - Translates business requirements into data models and data management requirements ◦
 - Creates and maintains the subject area diagrams , logical and DW data models , and datamart data models ◦
 - Assists in data integration problem resolution ◦
 - Defines metadata and maintains the metadata repository ◦
- ◆ Application Architect : Translates business requirements into DW application architecture requirements and designs ◦
- ◆ Technology Architect : Develops technology designs to meet data and application requirements ◦
- ◆ Support Architect : Develops support definitions for practices , procedures , staff , and technologies to manage the DW environment ◦
- ◆ IT Infrastructure Manager
 - IT resource w/ a background in H/W , systems software , and other infrastructure components in use at the organization ◦
 - Works w/ the architecture team , focusing on requirements for and design of the technology and support architecture components ◦
- ◆ Database Administrator

- Designs the physical DB environment , including any distributed DB components ◦
- Creates physical DB schema , manages performance , performs backups and archives data ◦
- ◆ Application Developer(s)
 - Develops the data extraction and transformation processes from the legacy systems ◦
 - Develops applications used to access the DW ◦
- ◆ Hardware Administrators : Implements the H/W configuration required to support the DW ◦
- ◆ Network Administrators : Implements the networks required to support the DW ◦
- ◆ Production Operations : Documenting and managing all batch processes required w/in the DW ◦
- ◆ Security Administrators : Establishing and administering security controls for the DW ◦
- ◆ Configuration Management : Managing and maintaining all programs and processes used to build , use , and manage the DW ◦
- ◆ Technical Writer : Writes and maintains all technical and user documents ◦
- ◆ Help Desk : Helps the users in accessing the metadata and DW data ◦
- ◆ Trainer : Create the training curriculum and trains the users ◦
- ✓ 組織與使用者
 - ◆ Business Executive
 - Provides budget and resources , monitors project progress , sets priorities , resolves conflicts , approves scope changes ◦
 - Guides the project team ◦
 - Breaks roadblocks ◦
 - ◆ Advisory Team

(出國類別：研究)

- Reviews project deliverables ◦
- Maintains a broad , enterprise perspective of the project ◦
- Sets priorities , resolves conflicts , facilitates , approves scope changes ◦
- ◆ Data Owner(s)
 - Business owner of the data ◦
 - Determines data security ◦
 - Resolves data integration issues ◦
 - Defines data and business rules ◦
 - Defines data use , calculations , and summarizations ◦
- ◆ Business Client(s)
 - Provides budget and resources , sets priorities , resolve conflicts , monitors project progress ◦
 - Reviews works in progress ◦
 - Approves deliverables and change requests in a timely manner ◦
 - Works w/ IT project manager to design and implement the DW environment ◦
- ◆ Subject Matter Experts
 - Knowledgeable in the needs and requirements of users , business processes , and existing source systems ◦
 - Validates data models , source of data , and user access ◦
 - Provides insight into business issues specific to the organization ◦
 - Provides knowledge of existing legacy systems where information resides and how it is used ◦
 - Works w/ the architecture team ◦
 - Focuses on the requirements for and design of the data architecture and application architecture components ◦
- 資料來源 (Data Sources)

- ✓ 內部資料
 - ◆ Relational。
 - ◆ Non-Relational。
 - ◆ Sequential。
- ✓ 外部資料：組織外部資料，可以是任何形式的資料，包括書面文字資料。
- 資訊架構與設計 (Architecture Design)
 - ✓ 如何選擇適合組織既有系統架構的設計。
 - ✓ RDBMS vs. MDDDBMS
 - ✓ Multi-Datamart vs. Centralized DW
 - ✓ 不適合委外作業，因為工作牽涉組織內部各相關應用系統，工作內容隨時間異動，但需要顧問型式的外援，以減少不必要的投資與浪費。
 - ✓ 如何減輕建置時的使用者抱怨。
- 分成幾次的建置計劃 (Multi-Generation Implementation Plan)
 - ✓ 每一次 DW 建置計劃是各自獨立的。
 - ✓ 不要將每一個 Roll 複雜化。
 - ✓ 當開始執行計劃後，將當初為納入計劃的部分與問題留待到下一個 Roll 解決。

5.2.2.4. DW 系統架構

- 系統架構分成
 - ✓ Data Architecture：The organization of sources and stores business informations，包括
 - ◆ Data stewards and guardians。
 - ◆ Business rules。
 - ◆ Data definitions。
 - ◆ Corporate logical data model。

(出國類別：研究)

- ◆ Source operational systems ◦
- ◆ Data integrity ◦
- ◆ Data distribution and synchronization ◦
- ◆ Data segmentation ◦
- ◆ Physical database design ◦
- ✓ Application Architecture : A S/W framework for building DW components that guides the overall implementation of business functionality , 包括
 - ◆ Assemble data from source systems ◦
 - ◆ Transform data into a consistent format ◦
 - ◆ Distribute data to locations close to the users ◦
 - ◆ Access the data for analysis and problem solving ◦
 - ◆ Metadata guides the processing , while instructing users about warehouse content and context ◦
 - ◆ Warehouse process management keeps the data moving ◦
- ✓ Technical Architecture : A computing infrastructure that enables the data and application architectures , 包括
 - ◆ Platforms : Server , Client ◦
 - ◆ Database : MPP , VLDB , DSS ◦
 - ◆ Networks : LAN , WAN ◦
 - ◆ Operating environments ◦
 - ◆ Data movement ◦
 - ◆ Performance monitoring ◦
 - ◆ Fault tolerance ◦
 - ◆ Remote or mobile connectivity ◦
- ✓ * Security Support Architecture : The S/W components and organizational functions requires to effectively manage the technology investment , 包括
 - ◆ Managing data movement ◦

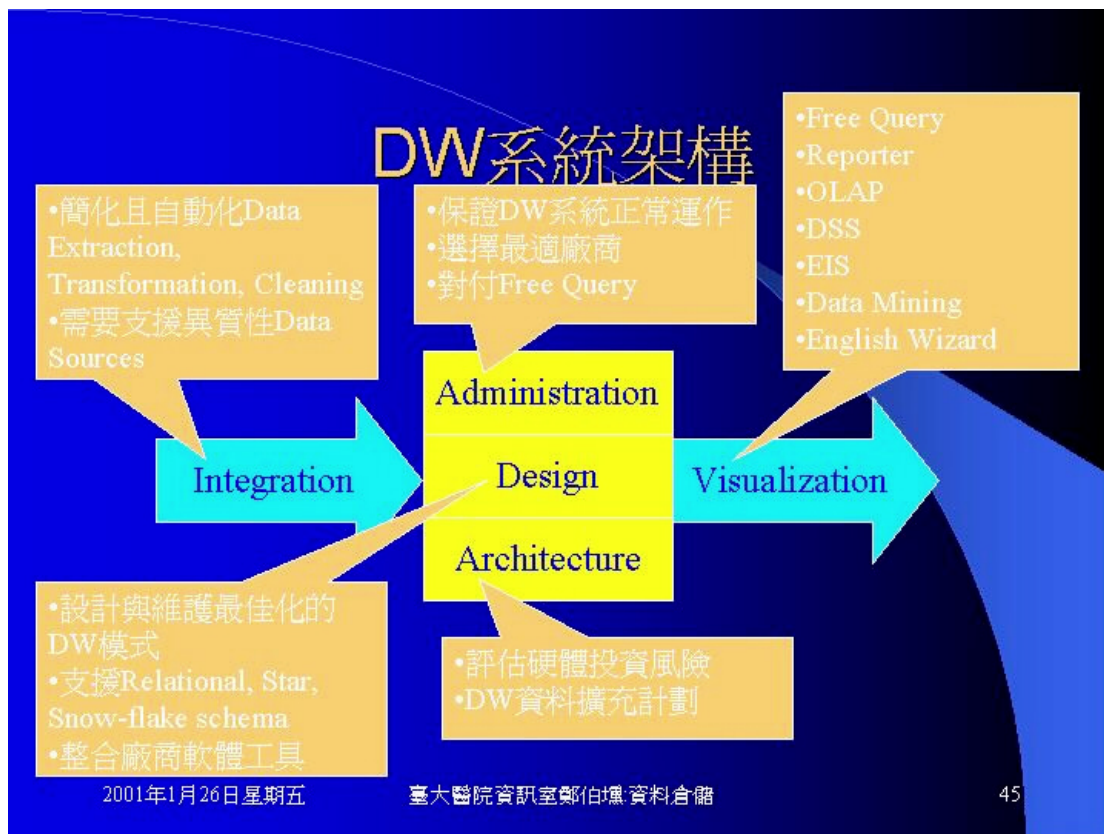
- ◆ Performance management ◦
 - ◆ Distributed system management ◦
 - ◆ Reliability requirements ◦
 - ◆ Service levels ◦
 - ◆ Support roles , staffing , training ◦
 - ◆ Help desk management ◦
 - ◆ Change management ◦
 - ◆ Source system changes ◦
- Architecture Model

DW系統架構 Architecture Model

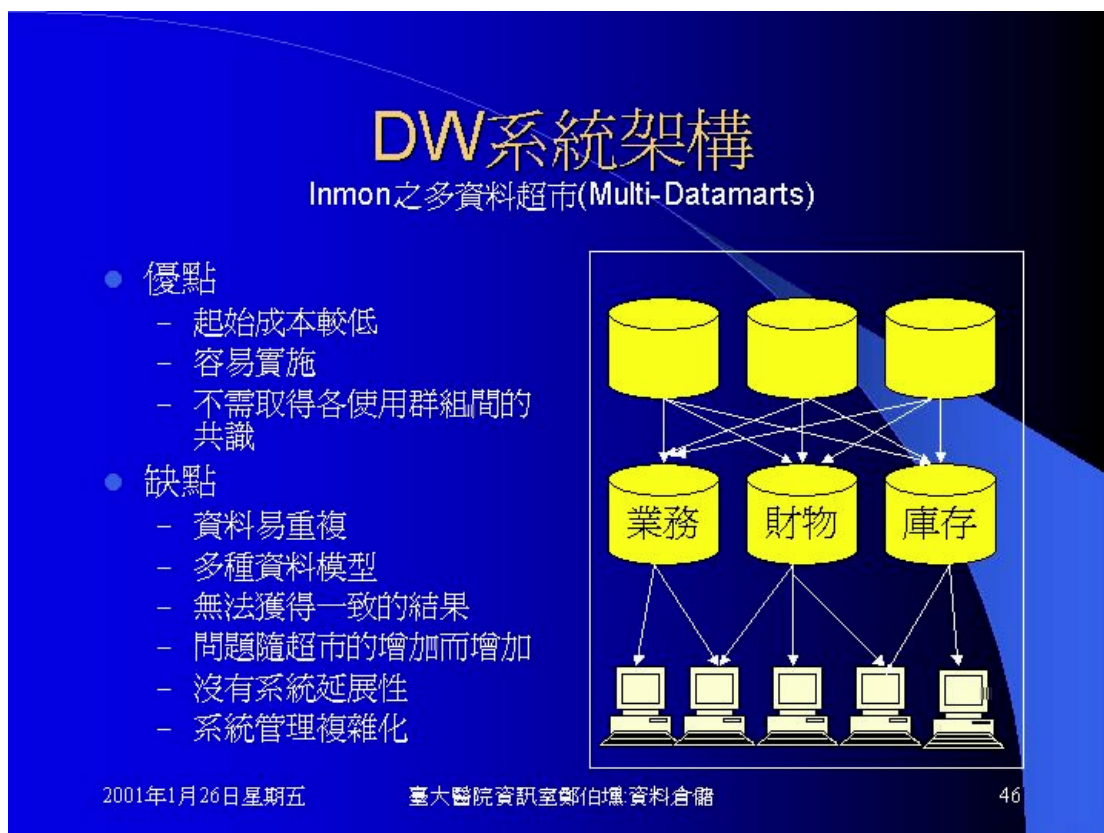
Application	Operational Business	Discovery Analysis	Target Manager	Source Logs	Admin. Data
Application Middleware	Metadata	Data Assembly	Data Transform	Data Access	Framework Service
Middleware	System Manager	Workflow	Data Distributed	Directory Service	C/S
Hardware Enabling	RDBMS	Network Services	OS	Fault Tolerance	Backup
Hardware	PCs	Servers	Mainframe	Data Storage	Networks

2001年1月26日星期五
臺大醫院資訊室鄭伯璦資料倉儲
44

(出國類別：研究)



● Inmon 之多資料超市(Multi-Datamarts)



● 集中式資料倉儲(Centralized Data Warehouse)

DW系統架構

集中式資料倉儲(Centralized Data Warehouse)

- 優點
 - 資料一致性
 - 一種資料模型
 - 問題單純化
 - 系統容易延展性
 - 系統管理單純化
- 缺點
 - 起始成本較高
 - 不易實施
 - 需取得各使用群組間的共識

2001年1月26日星期五
臺大醫院資訊室鄭伯堉 資料倉儲
47

- 集中式資料倉儲含分散式資料超市(Centralized Data Warehouse w/ Distributed Datamarts)，又稱為 Centributed 架構。

DW系統架構

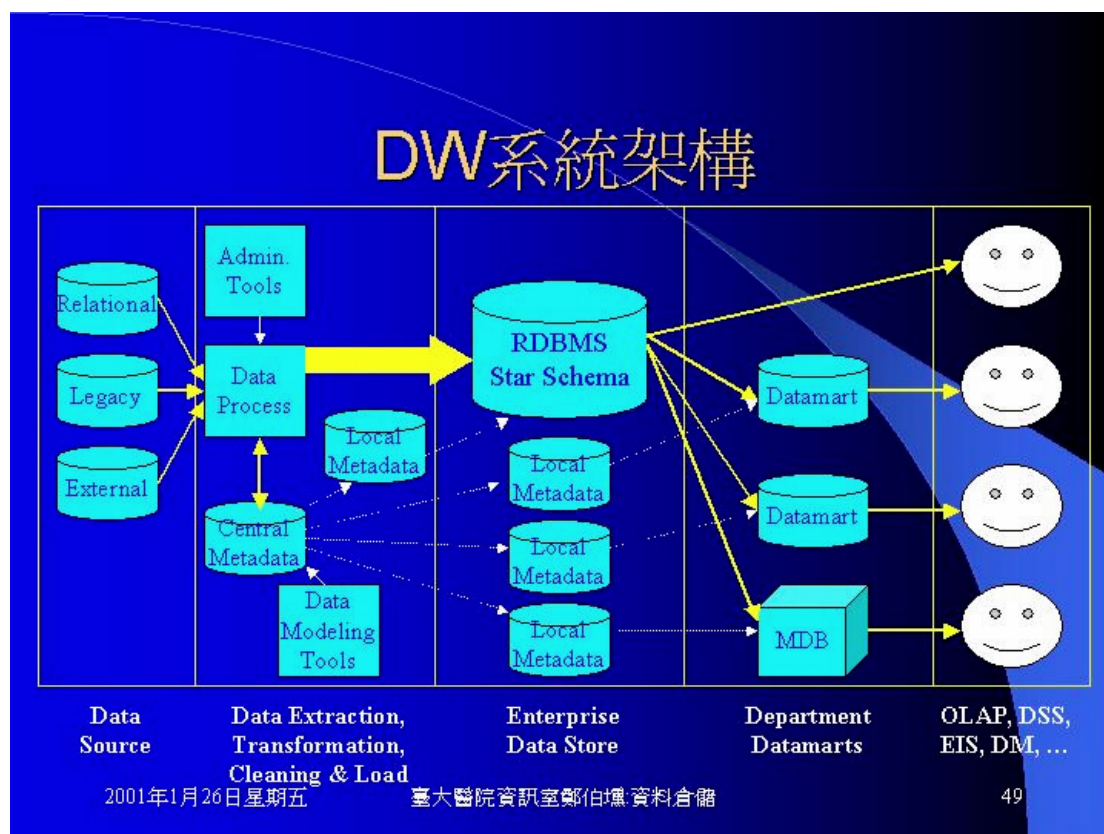
集中式資料倉儲含分散式資料超市 (Centralized Data Warehouse w/ Distributed Datamarts)

- 又稱為Centributed架構
- 優點
 - 資料一致性
 - 一種資料模型
 - 問題單純化
 - 系統容易延展性
 - 系統管理單純化
- 缺點
 - 起始成本太高
 - 不易實施
 - 需取得各使用群組間的共識

2001年1月26日星期五
臺大醫院資訊室鄭伯堉 資料倉儲
48

(出國類別：研究)

- DW 系統架構



5.2.2.5. Data Movement Models

- Bulk Data Movement
 - ✓ Data collected in batches on scheduled basis ◦
 - ✓ Data refreshed and delta changes ◦
- Store-and-forward Data Movement
 - ✓ Capture selected changes from source system ◦
 - ✓ Must provide full transaction integrity ◦
- Real-time Synchronous Data Movement ◦

5.2.2.6. 架構設計考量

- Normalization vs. De-normalization ◦

- Parallel Processing ◦
- Database Management ◦

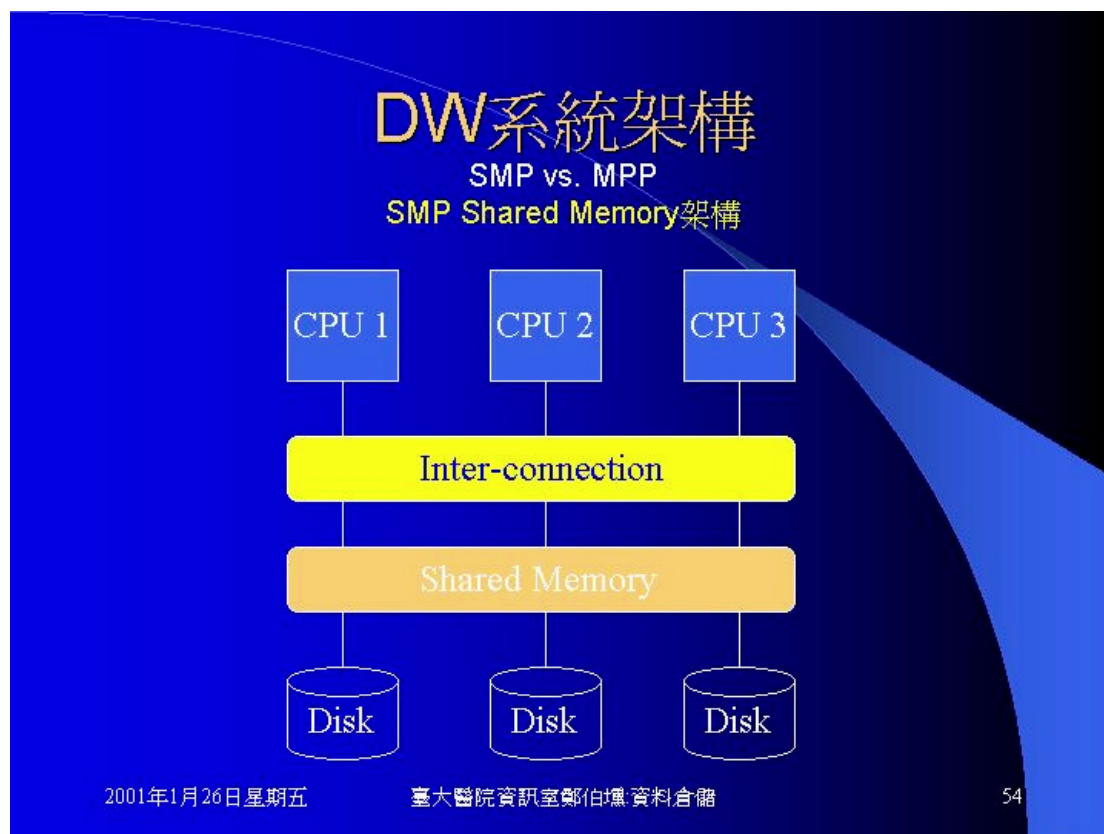
5.2.2.7. 建置過程

- Business Requirements --> Architecture Requirements ◦
- Architecture Requirements --> Architecture Design ◦
- Dimensional Data Modeling ◦
- Sizing the DW ◦
- Accessing the DW ◦

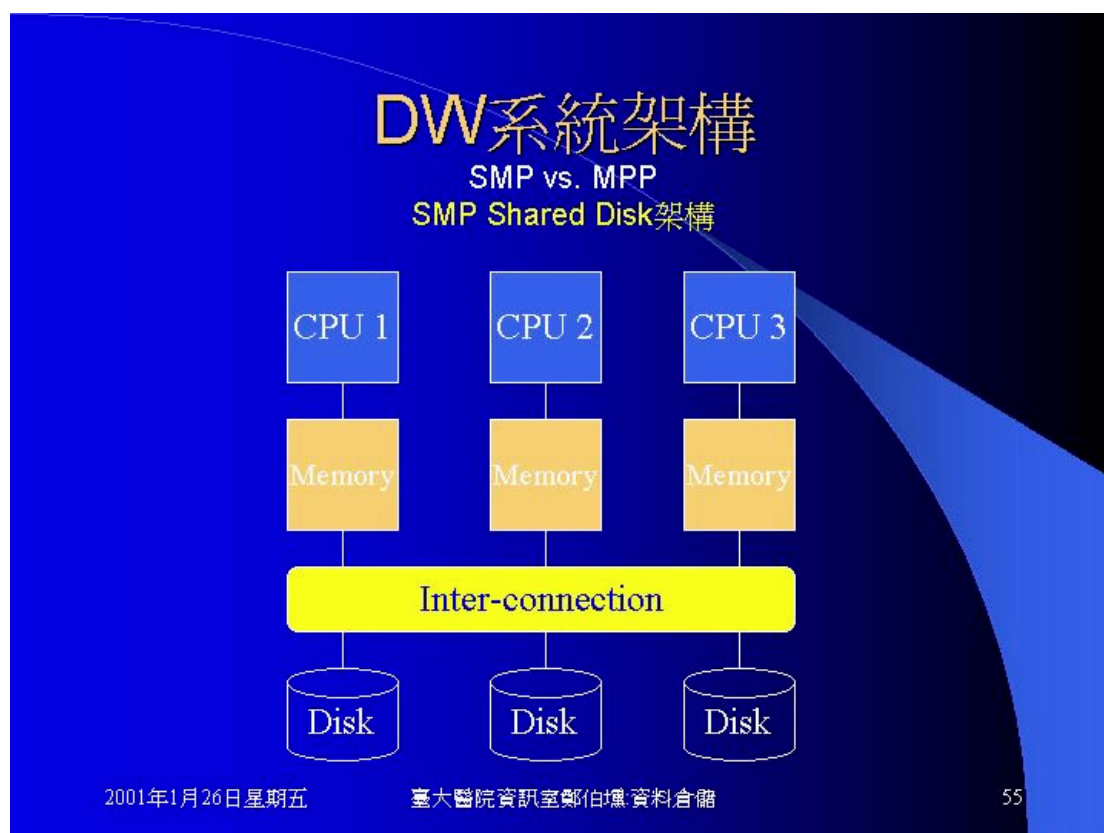
5.2.2.8. SMP vs. MPP

- SMP (Symmetric Multiprocessing)
 - ✓ 所謂的 Shared-Everything Processors ◦
 - ✓ 2~128 Processors, Shared Memory/Disk, Scaleable Shared Memory ◦
 - ✓ 使用於中大型 Data Marts 與中型 Data Repositories ◦
 - ✓ 廠商：IBM, HP, DEC, Sun, Digital, NCR, Sequent, Pyramid ◦
- MPP (Massively Parallel Processing)
 - ✓ 所謂的 Shared-Nothing Processors ◦
 - ✓ ~1024 Processors, 每一個 Processor 有自己的 Memory ◦
 - ✓ 系統功能與作業平台通常不獨立 ◦
 - ✓ 使用於超大型資料庫(VLDB, Very Large Database) ◦
 - ✓ 廠商：IBM/SP2, HP, NCR ◦
- SMP Shared Memory 架構

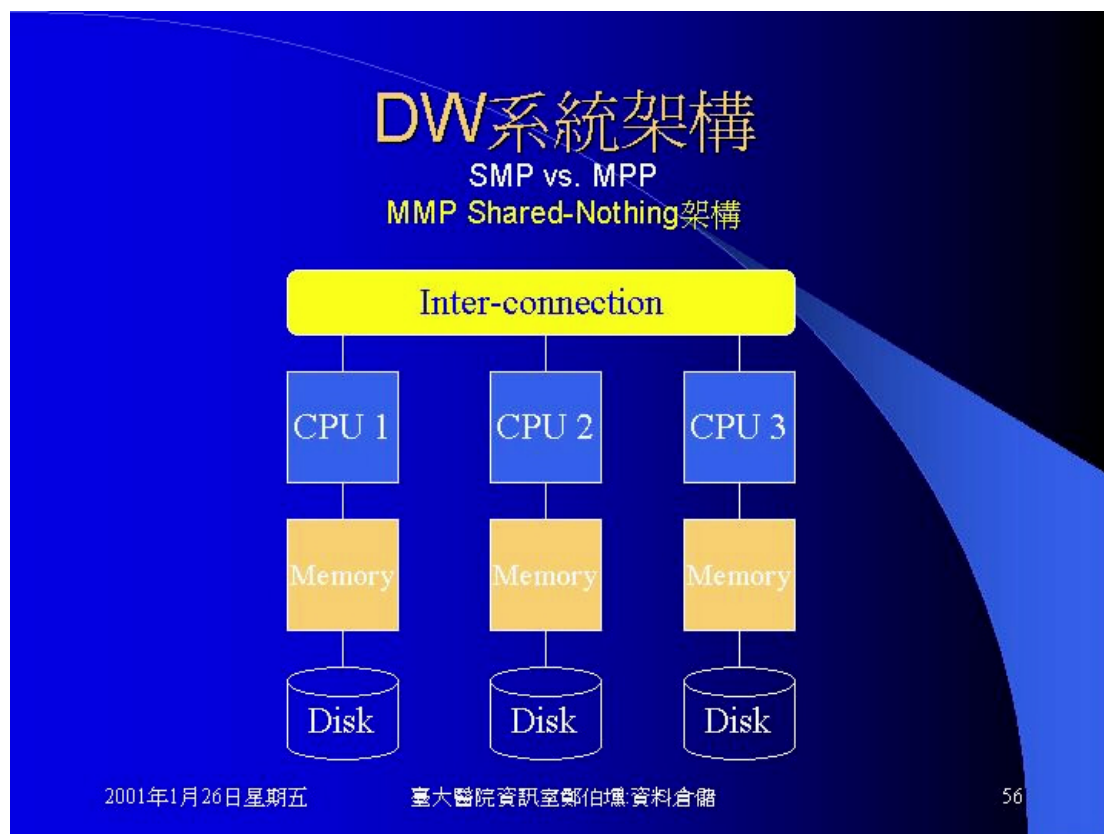
(出國類別：研究)



- SMP Shared Disk 架構



- MMP Shared-Nothing 架構



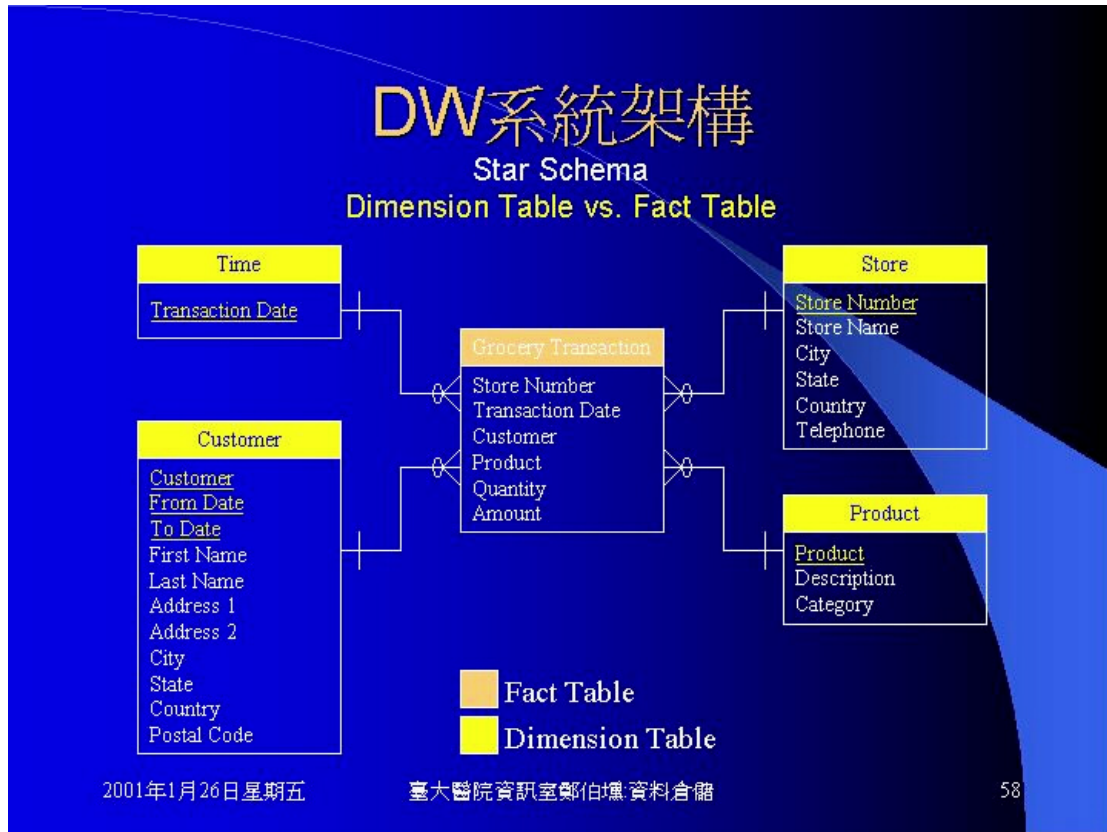
5.2.2.9. Data Modeling

- Star Schema：將已知的 Dimension 利用 Aggregate 方法，將資料以 Multi-dimensional 方式儲存在 2D 的 RDBMS 中。
- Snowflake Schema：在關聯環境中使用階層方式對應到 Star Schema 中的 Dimension，可以變形為 Multiple Parent Schema。
- Constellation (星座) Schema：也就是 Multiple Fact Table Schema。
- Snowstorm Schema：也就是混合 Multiple Parent Schema & Constellation Schema。
- Aggregate：有 2D Aggregate Table，3D Aggregate Table。

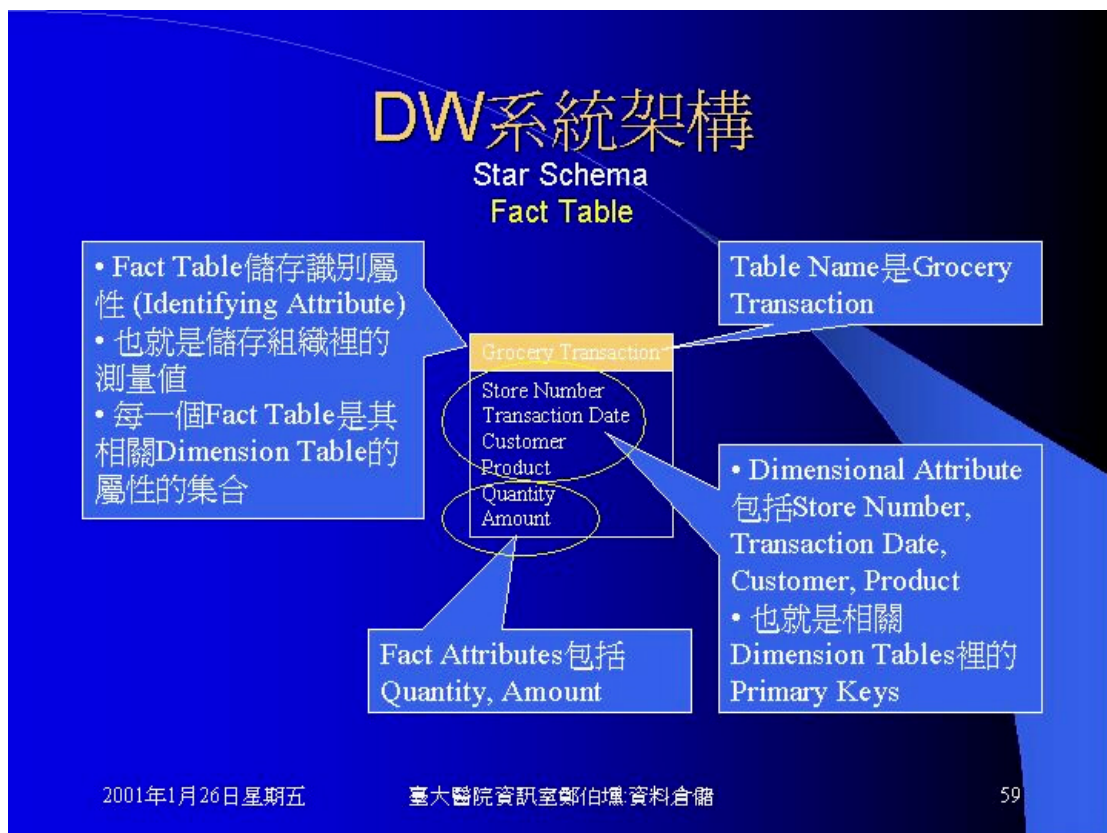
5.2.2.10. Star Schema

- Dimension Table vs. Fact Table

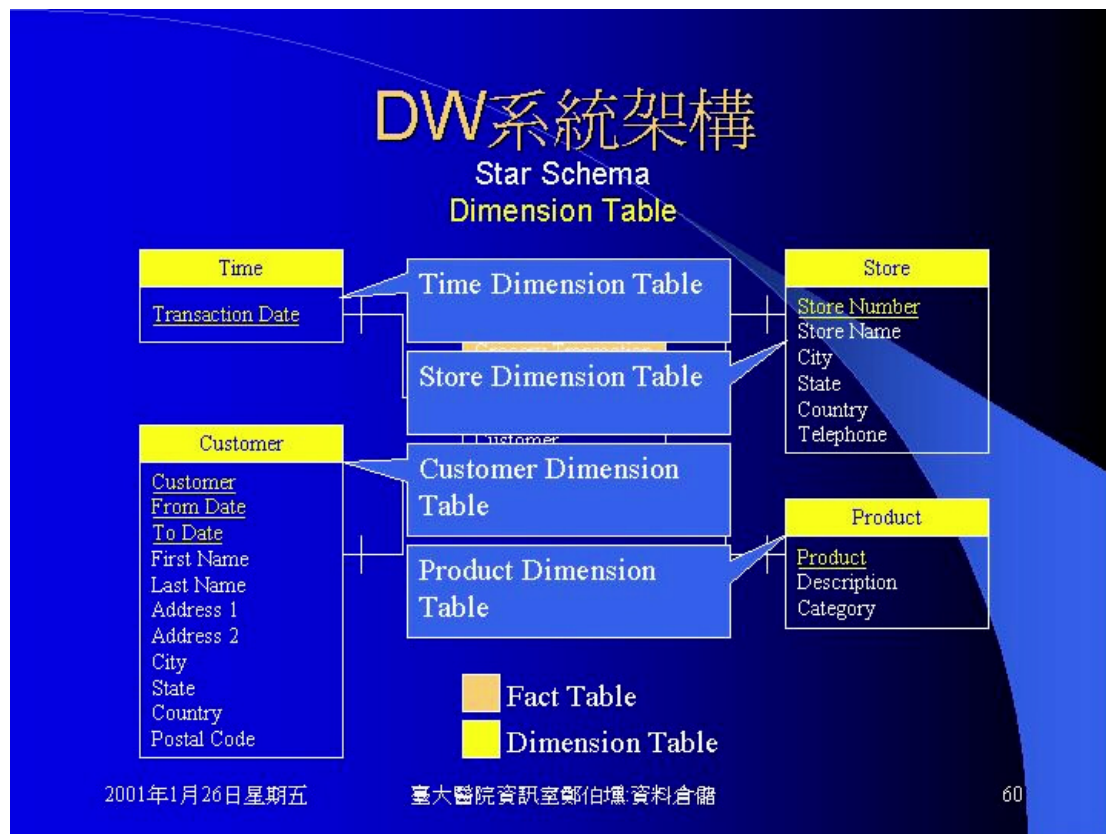
(出國類別：研究)



● Fact Table



● Dimension Table



(出國類別：研究)

5.2.2.11. Star Schema vs. Fully Normalized Relational

DW系統架構

Star Schema

Star Schema vs. Fully Normalized Relational

Star Schema

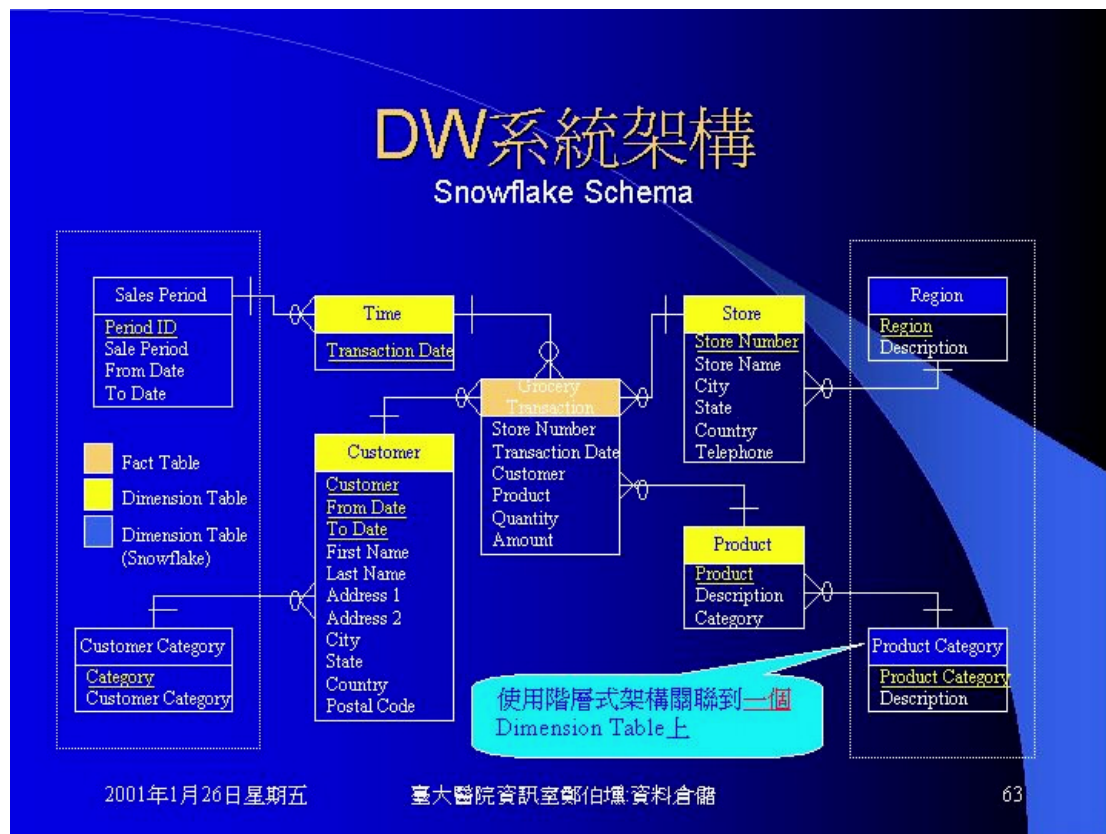
- 高度反正規化 (De-normalization) 以求達到最佳效能
- 交易資料儲存在 Fact Table
- 參考資料儲存在 Dimension Table

Fully Normalized Relational

- 正規化 (Normalization) 過程在企圖消除關聯式資料庫裡重複的資料
- 交易資料與參考資料可以儲存在任一個資料表中

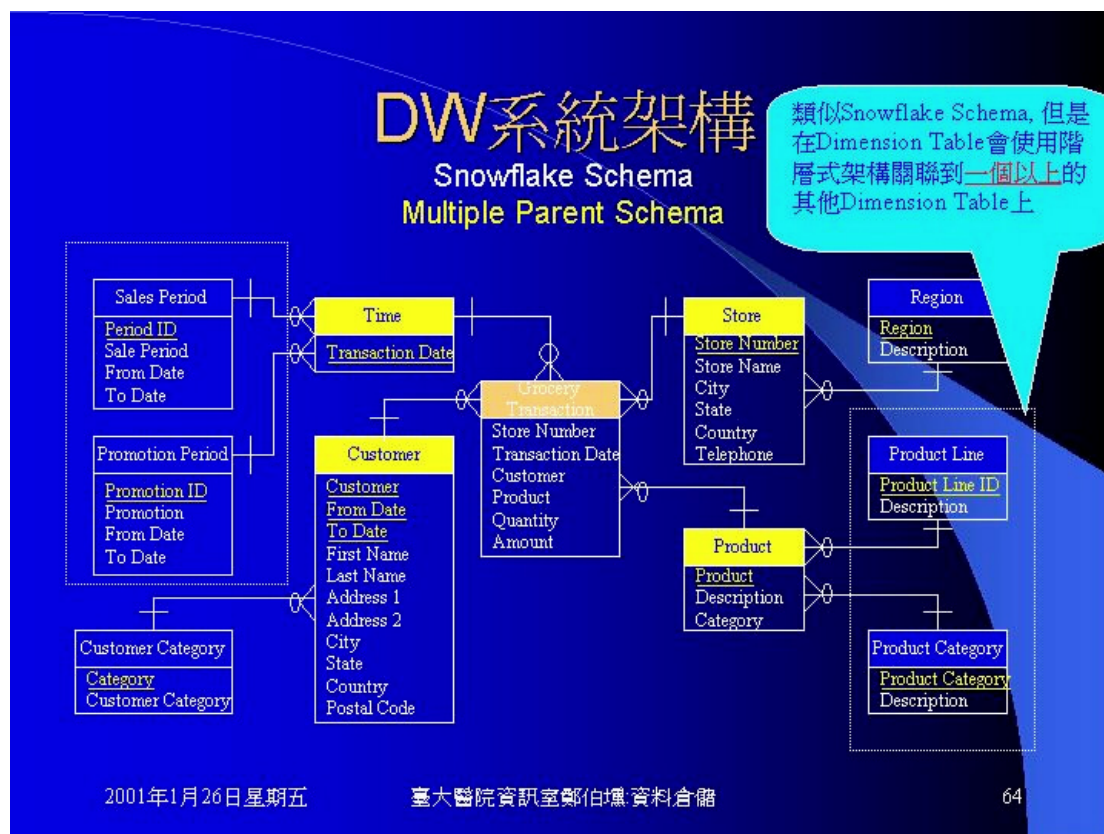
2001年1月26日星期五 臺大醫院資訊室鄭伯堉 資料倉儲 62

5.2.2.12. Snowflake Schema

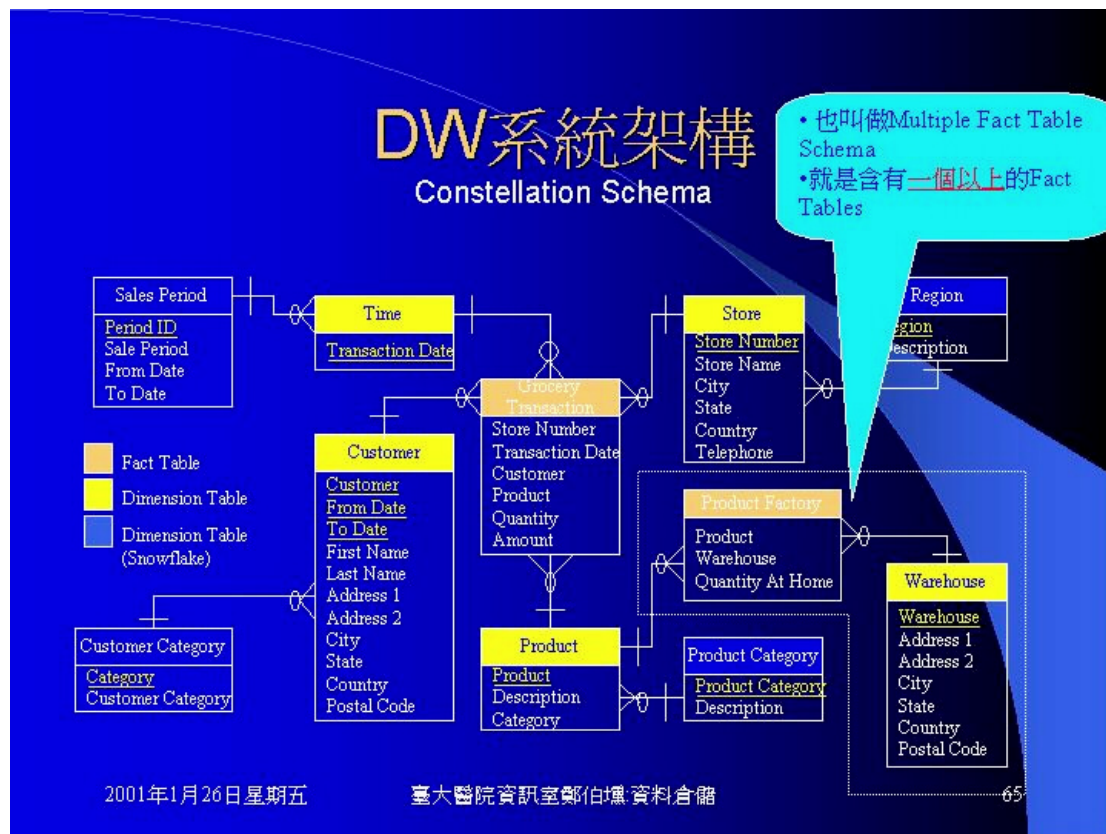


- Multiple Parent Schema

(出國類別：研究)

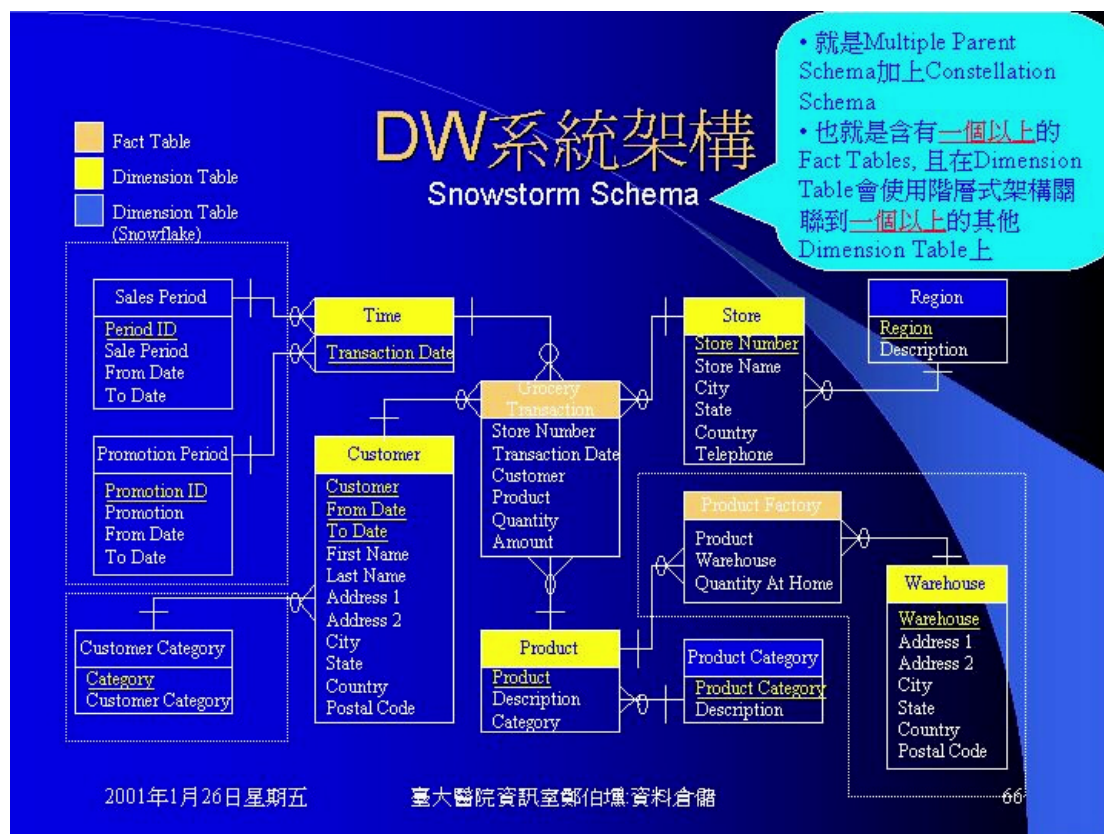


5.2.2.13. Constellation Schema



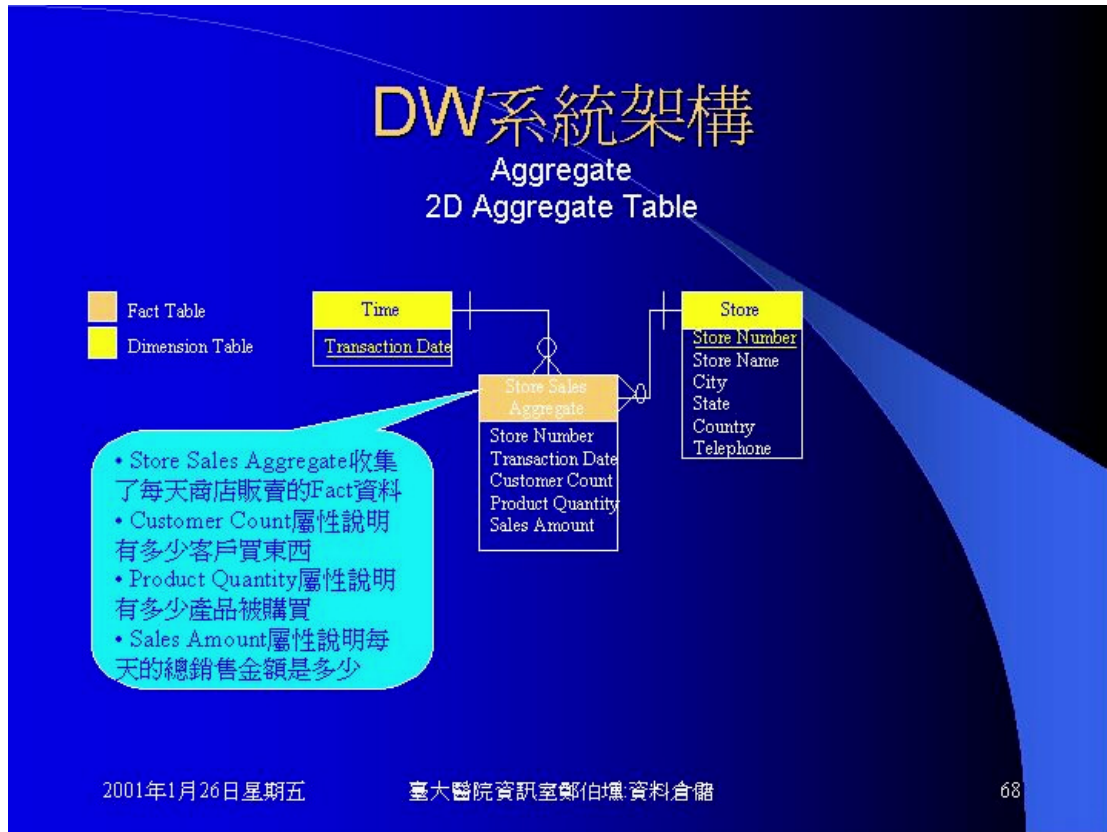
(出國類別：研究)

5.2.2.14. Snowstorm Schema

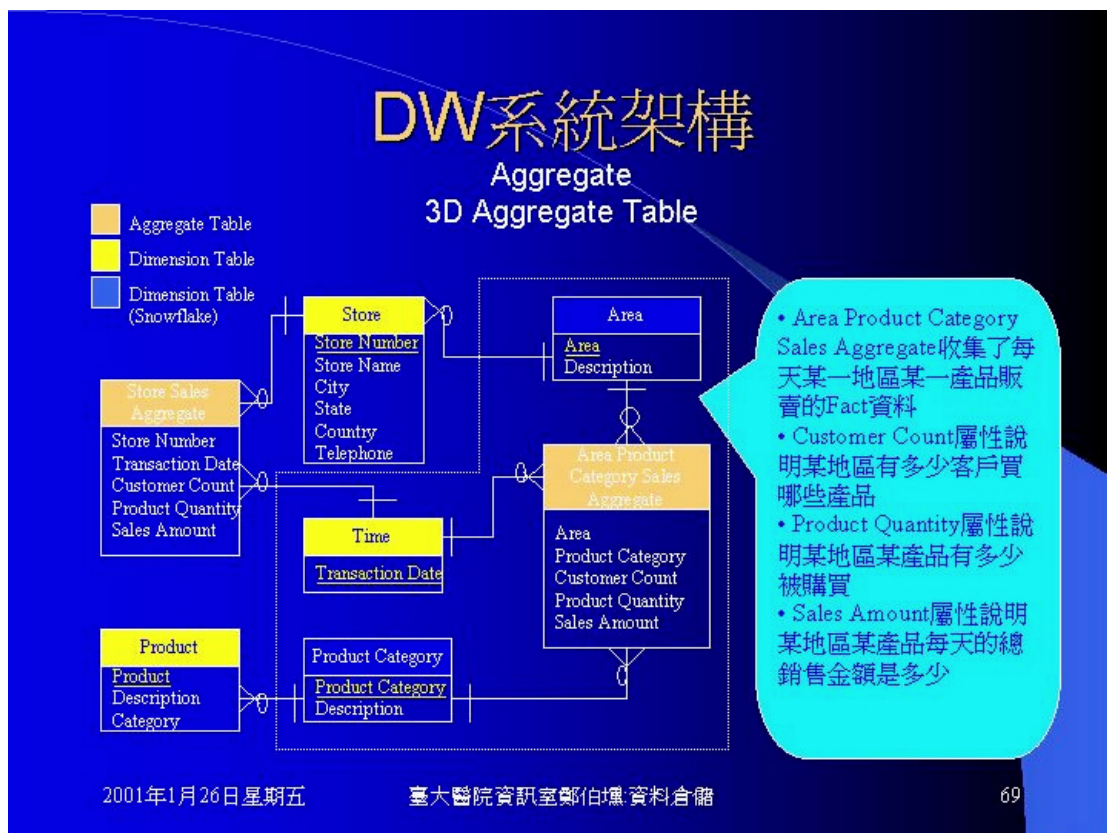


5.2.2.15. Aggregate

- 提供多階層的 Fact 資訊細節。
- 事先計算好查詢的結果以便增加查詢效能。
- 可以收集不同群組的資訊，通常收集自許多不同的 Dimension 或許多 Dimension 的組合。
- 2D Aggregate Table。



● 3D Aggregate Table



(出國類別：研究)

5.2.2.16. Dimensional Data Modeling (DDM)

- 廣泛使用 De-normalization
 - ✓ Dimensional Data Model 廣泛使用 De-normalization，以避免 Table Join De-normalization
 - ◆ 強調效能，所以資料易重複，會犧牲儲存空間。
 - ◆ 很適用於 OLTP，尤其是事先計劃好的作業，但是對於未事先計劃的作業則不適用。
 - ◆ 當重複的資料異動時，需要額外作許多工作。
- 使用 Multiple Fact Table 查詢方式
 - ✓ Fact Table 經常非常的大，也就是比一般正規化後的資料表大非常多。
 - ✓ 因為 Fact Table 非常大，所以 Dimensional Data Modeling 避免了 Table Join，讓查詢速度加快。
 - ✓ 而與 Dimension Table Join 則增加了使用者對資料的了解程度。
- 善用 Aggregates
 - ✓ Aggregate 必須被事先計劃。
 - ✓ 80/20 理論。
 - ✓ 指數遞增風險。
 - ◆ Aggregate Table Explosion：N!
 - ◆ 例如
 - 一年內的資料。
 - 100 家商店，25 種產品分類，50 種客戶類型。
- 盡量不要 Snowflake
 - ✓ Ralph Kimball 的 "The Data Warehouse Toolkit" 書中說到 "Do not snowflake your dimensions, even if they are large. If you do so, be prepared to live with poor query performance."

5.2.2.17. Logical Data Model vs. DW Data Model

DW系統架構

Logical Data Model vs. DW Data Model

Logical Data Model	DW Data Model
Normalized	De-normalized
Detailed data	Detailed and summarized data
Corporate-wide perspective	Strategic decision-making perspective
May contain time w/ key	Must contain an element of time
No derived data	Derived, strategic data
Repeating groups shown only once	Can contain data arrays
Organized around business rules	Organized around usage and stability

2001年1月26日星期五
臺大醫院資訊室鄭伯璦 資料倉儲
75

- Bill Inmon 的 Logical Data Model --> DW Data Model 轉換步驟
 - ✓ Removal of purely non-strategic data。
 - ✓ Addition of an element of time to the key structure。
 - ✓ Addition of derived data。
 - ✓ Accommodation of different levels of granularity。
 - ✓ Summarization schemes。
 - ✓ Merging like data from different tables。
 - ✓ Creation of arrays of data。
 - ✓ Separation of data attributes based on their stability。
- Logical Data Model --> DW Data Model 轉換注意事項
 - ✓ 加入時間。
 - ✓ 每次將資料加入 DW 時，先問是否值得加入。

(出國類別：研究)

5.2.2.18. Data Granularity vs. Transaction Granularity

- Data Granularity
 - ✓ 儲存資料了細微程度。
 - ✓ 與 DW 的建置類型有關。
 - ✓ 有些情況下，作業系統與 DW 的資料細微程度會有明顯的差異。
 - ✓ 當資料細微程度改變時，要反映在 DW 資料模型上。
 - ✓ 如果 DW 只儲存 Summarize 資料，將無法 Drill-down。
- Transaction Granularity：交易資料的細微程度。

5.2.2.19. Multi-dimensional

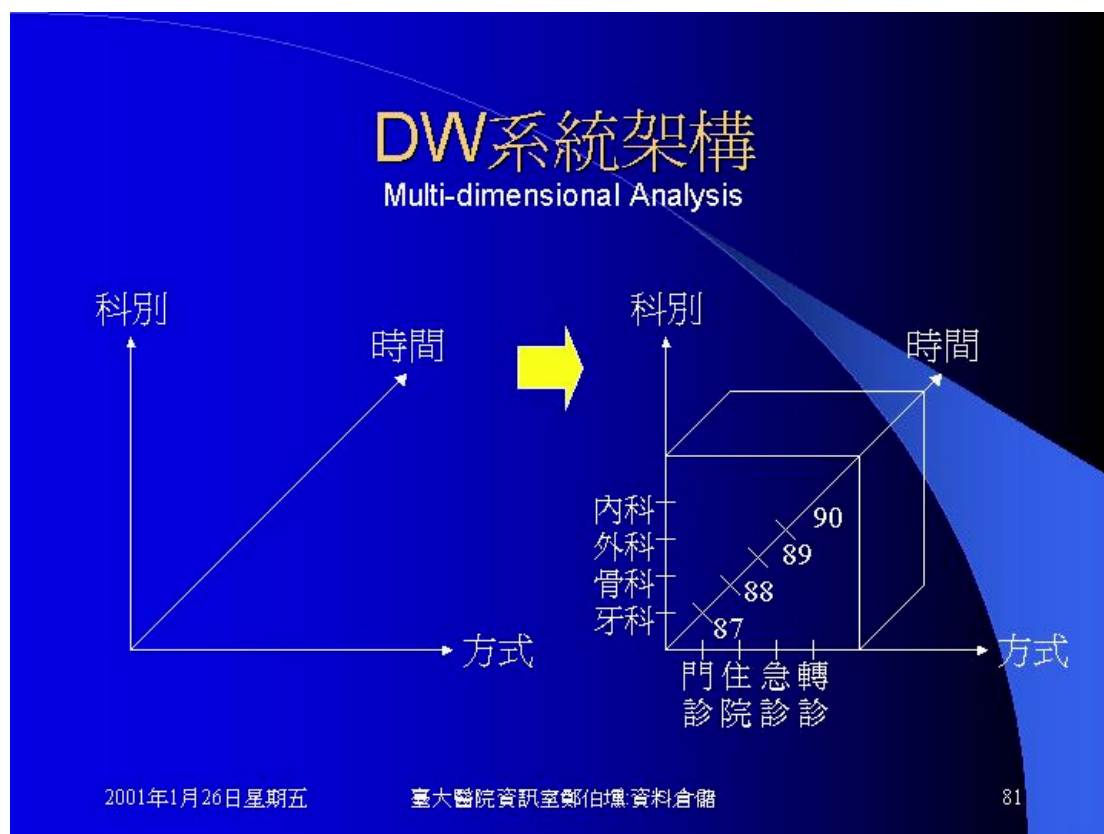
- Hypercube
 - ✓ 也就是 Multi-dimensional Data。
 - ✓ 一種視覺化的表示方法。
- Multi-dimensional Analysis
 - ✓ 又叫做 Drill-Down，Drill-Across 或 Slice-and-Dice。
 - ✓ 經由許多相關的 Dimension 處理資料以方便分析與了解資料。
- Multi-dimensional Database (MDDDB)：非關聯式資料庫管理工具，可以儲存與管理 Multi-dimension 資料。
- OLAP：一組功能以方便 Multi-dimension Analysis。

5.2.2.20. Multi-dimensional vs. Multi-relational

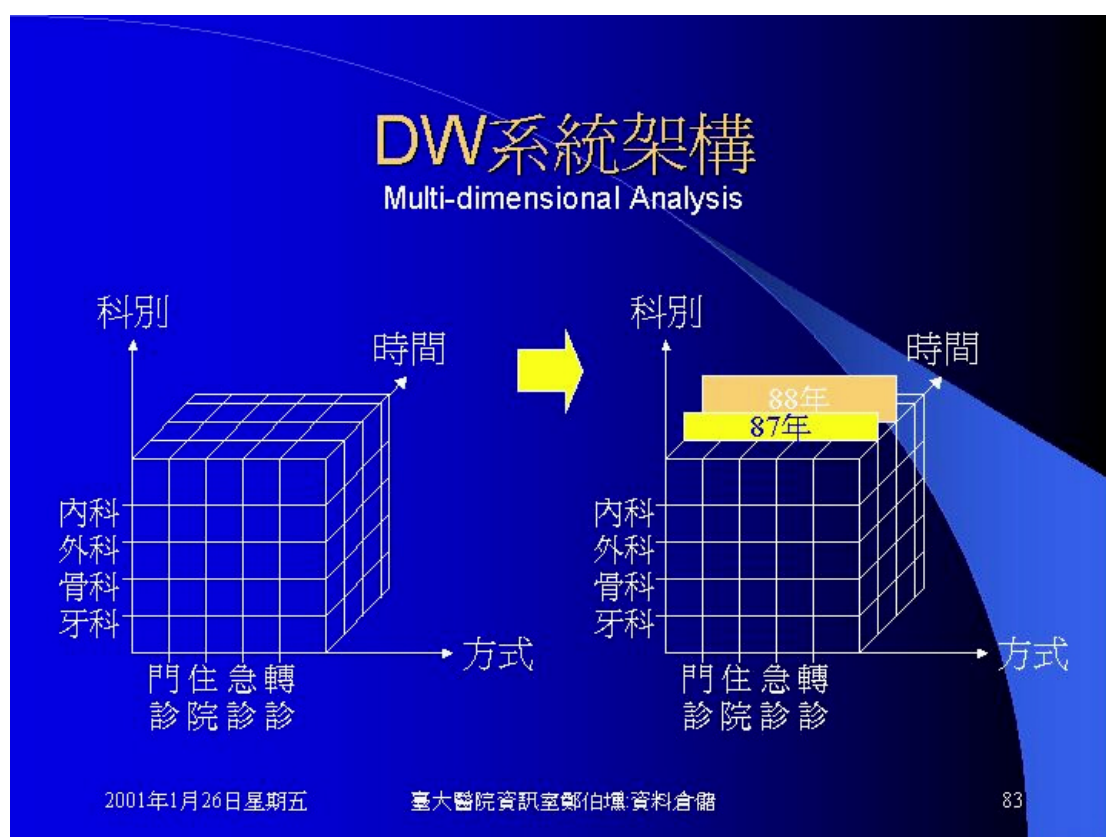
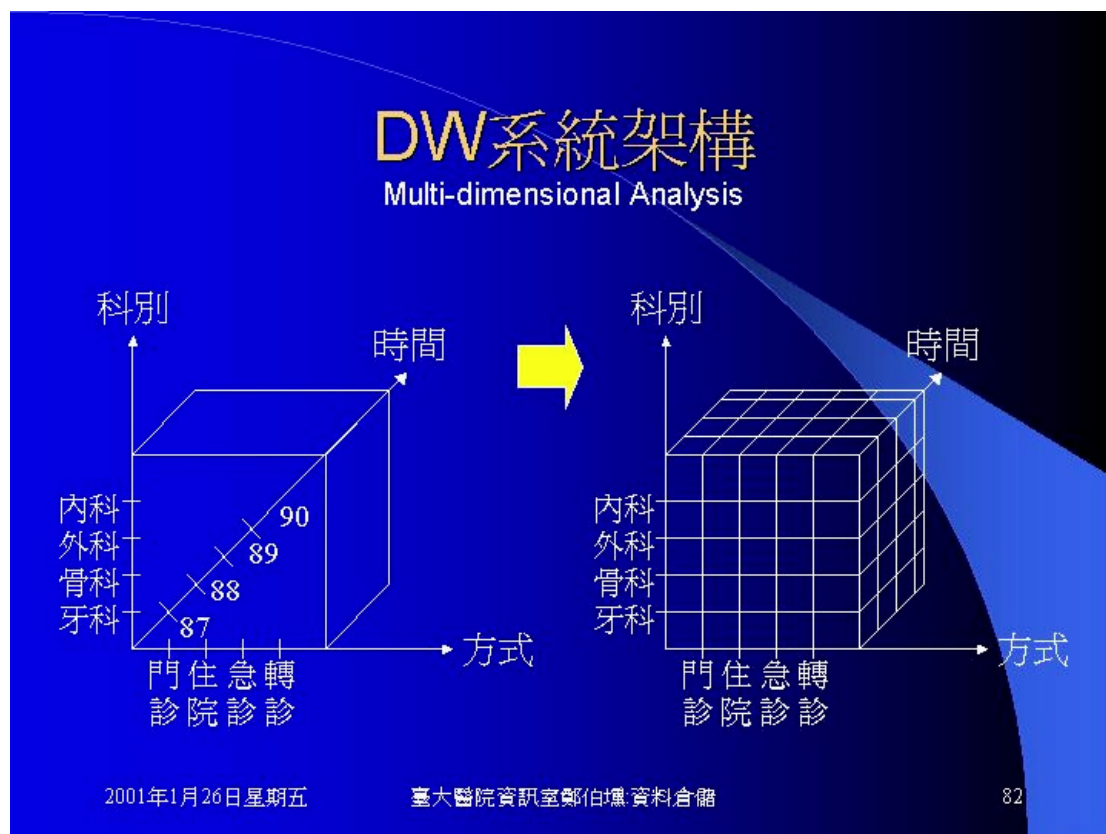
- Multi-dimensional：一種環境通常針對一定範圍內的許多可能組合的結果利用 Bit-mapped 方法作索引，事先計算與收集好數值。經由許多相關的 Dimension 處理資料以方便分析與了解資料。
- * Multi-relational：一種環境能經由 Cache 或線上 RDBMS 資訊，很快的計算與收集數值，通常是以 MDDDB 為前台，RDBMS 為後台。

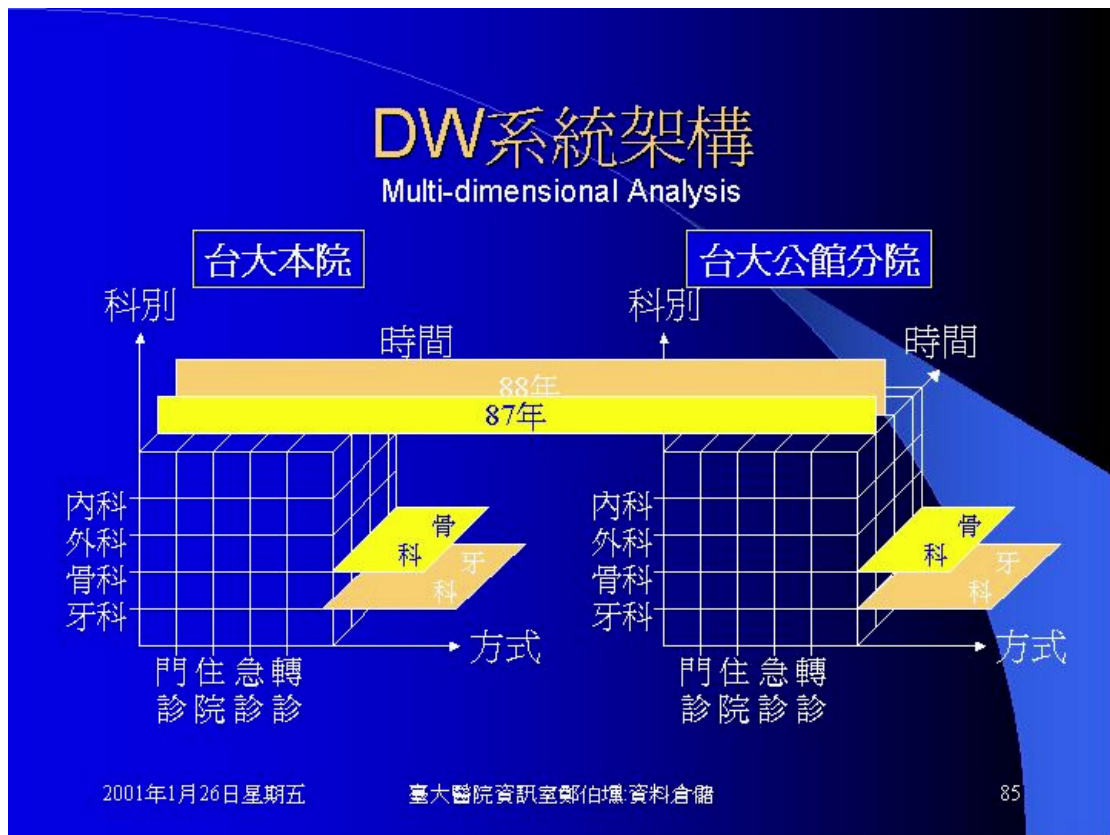
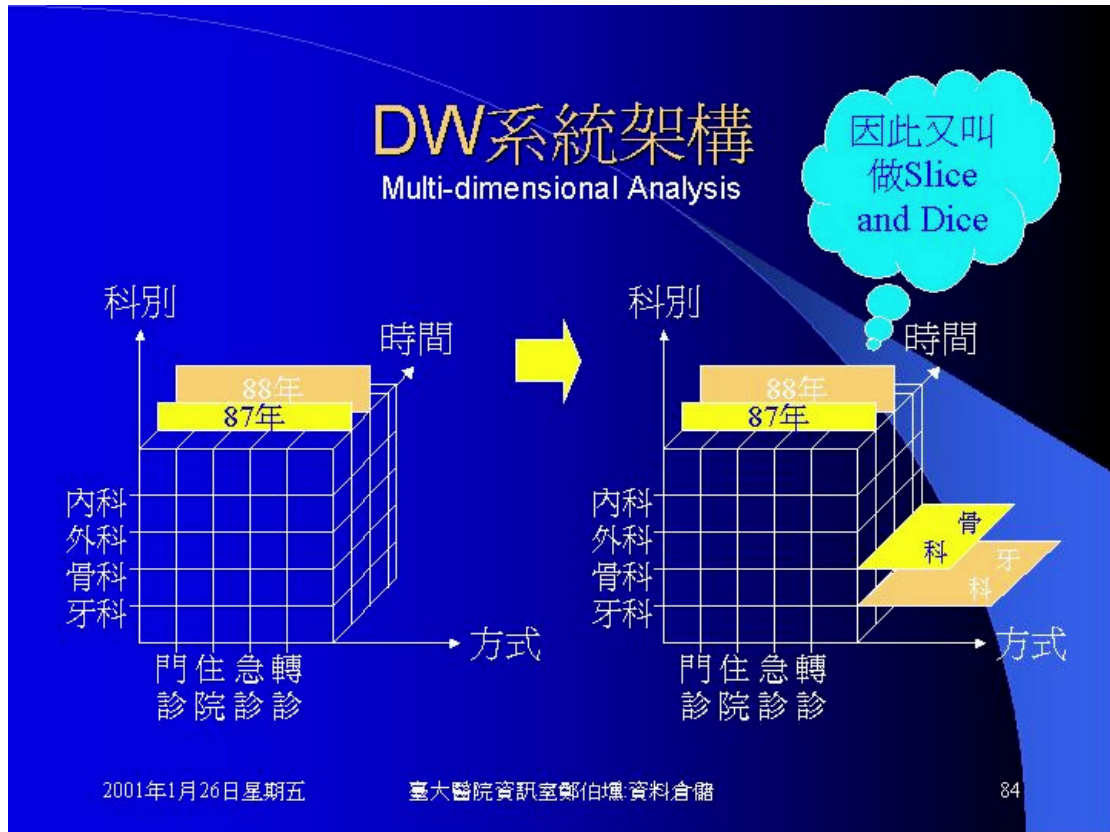
- Ad-hoc Query：一種環境能方便使用者點選資料來產生 SQL 語法，進而從 RDBMS 中取得資訊。

5.2.2.21. Multi-dimensional Analysis



(出國類別：研究)





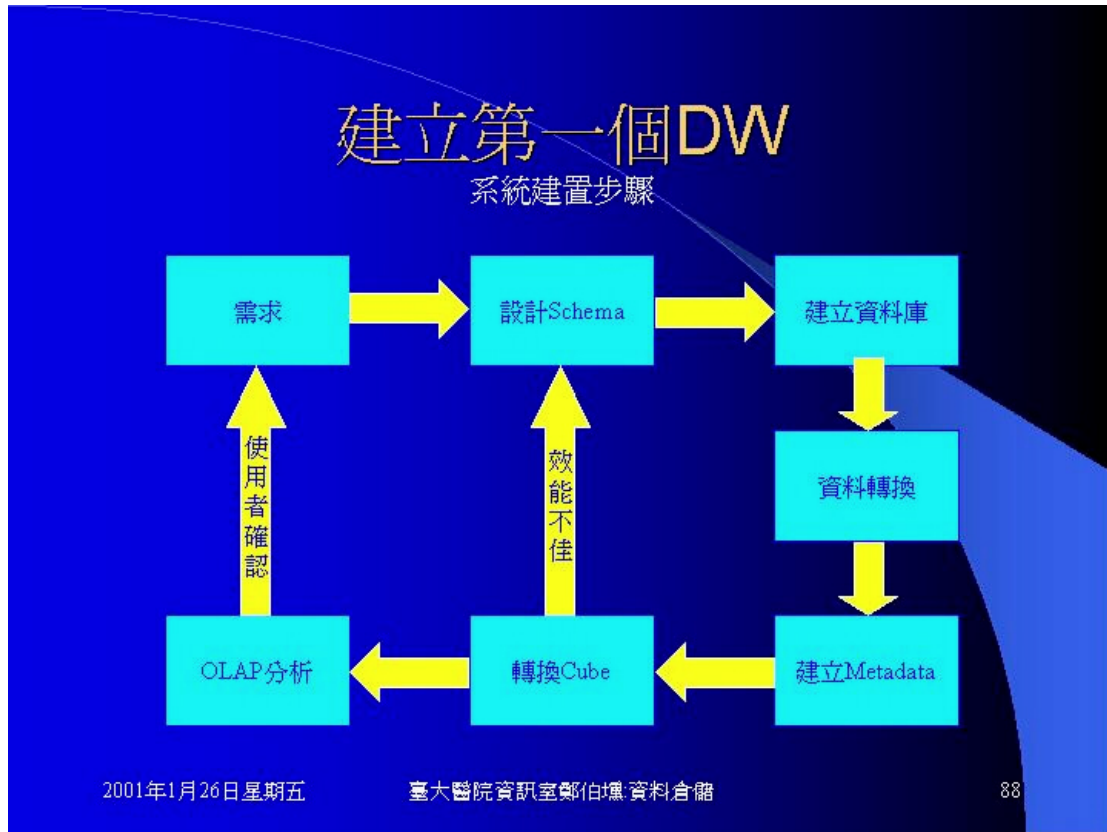
(出國類別：研究)

5.2.2.22. 資料容量

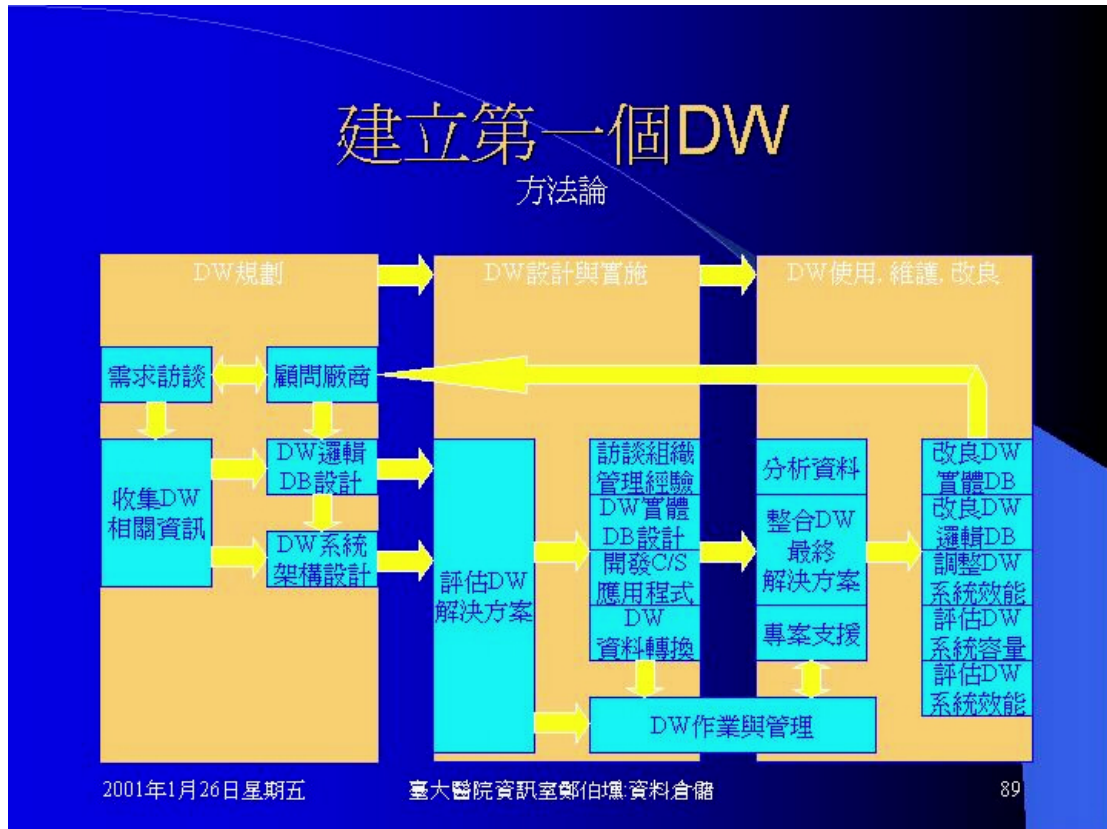
- DW 總大小= Raw Data (30%)+ Indexes (40%)+ Aggregates (20%)+ Workspace(10%)。
- Data Mart 總大小= Fact Table 總大小+ Indexes+ Aggregates+ DBMS Overhead。
- Fact Table Row Count = # of Base Fact Records。
 - ✓ # of Key Columns' Unique Values。
 - ✓ Sparse Factor。

5.2.2.23. 建立第一個 DW

- 重點確認
 - ✓ 確認此 DW 的大小範圍。
 - ✓ 開發 DW Data Model。
 - ✓ 取得資料。
 - ✓ 75%以上的投資會花在資料轉換上。
 - ✓ 多使用現成廠商之工具。
- Hard Coding：一定會成功，但是費時，不夠彈性，且不易維護。
- 系統建置步驟



● 方法論



● DW 上線

(出國類別：研究)

✓ 重點確認

- ◆ 調整出最佳的 DW 效能。
- ◆ 使用者分群
 - 系統維運者 (Administrators)：系統管理者，網路管理者，資料庫管理者，機房維運者。
 - 應用系統開發者 (Developers)：DW 研發者，Data Mart 研發者，Data Extraction 研發者。
 - 資深使用者 (Power Users)：熟悉業務領域與分析工具之使用者。
 - 一般使用者 (End Users)：依照所屬單位職務與工作性質分類授權之一般使用者。

● DW 維護

✓ 重點確認

- ◆ 與使用者確認服務的範圍。
- ◆ 作業流程管理。
- ◆ 使用 Metadata 來管理 DW。
- ◆ 資料備援與備份。
- ◆ 確認資料庫架構的一致性。
- ◆ 資料取得與還原。
- ◆ 與下一次 DW 計劃異動的管理。

5.2.2.24. DW 工具

- Data Access / Query Tools。
- * Report Writer。
- * MDBMS (Multidimensional DBMS)。
- * DSS (Decision Support System)。
- * EIS (Executive Information System)。

5.2.2.25. DW 資料轉換

- 工作
 - ✓ Data Model Design。
 - ✓ Data Extraction。
 - ✓ Data Transformation。
 - ✓ Data Cleaning。
 - ✓ Data Load。
 - ✓ Monitor / Manager。
- 工作困難點
 - ✓ 需佔整體建置人力與時間的 80%。
 - ✓ 程式撰寫不易。
 - ✓ 無法完全自動化作業。
 - ✓ 無法即時整合不同資料庫內的資料。
 - ✓ 資料庫架構異動時需要更改既有 Data Extraction 與 Data Load 程式。

5.2.2.26. Homonyms vs. Synonyms

- Homonyms
 - ✓ Same name, different object。
 - ✓ 例如：序號(SeqNo)：可以是檢驗單序號或工作清單序號。
- Synonyms
 - ✓ Same object, different name。
 - ✓ 例如：帳號：AccountNo, CaseNo。

(出國類別：研究)

5.2.2.27. DW 的延伸應用

- DM (Data Mining) 。
- OLAP (On-Line Analytical Processing) 。
- DSS (Decision Support System) 。
- EIS (Executive Information System) 。
- Statistics Analysis 。
- Access Reporting 。
- Free Query 。
- English Wizard 。
- Internet 。

5.2.2.28. OLTP vs. DSS

- OLTP，On-Line Transaction Processing(線上交易) 。
- ✓ 查詢與處理方式是事先制定好的。
- ✓ 簡單的查詢，為已知問題之查詢。
- ✓ 資料量小且處理時段較寬鬆，反應時間快速。
- ✓ 例如，某病友掛號骨科門診。
- ✓ On-Line Transaction Processing，線上交易。
- ✓ 查詢與處理方式是事先制定好的。
- ✓ 例如，某病友掛號骨科門診。
- DSS，Decision Support Systems (決策支援系統) 。
- ✓ 複雜且多表格查詢，為未知問題之查詢。
- ✓ 資料量大且密集處理，反應時間與資料庫大小成反比。
- ✓ 例如，哪種病友容易預先掛號後爽約。
- ✓ Decision Support Systems，決策支援系統 。
- ✓ 例如，哪種病友容易預先掛號後爽約 。

5.2.2.29. OLAP vs. DM

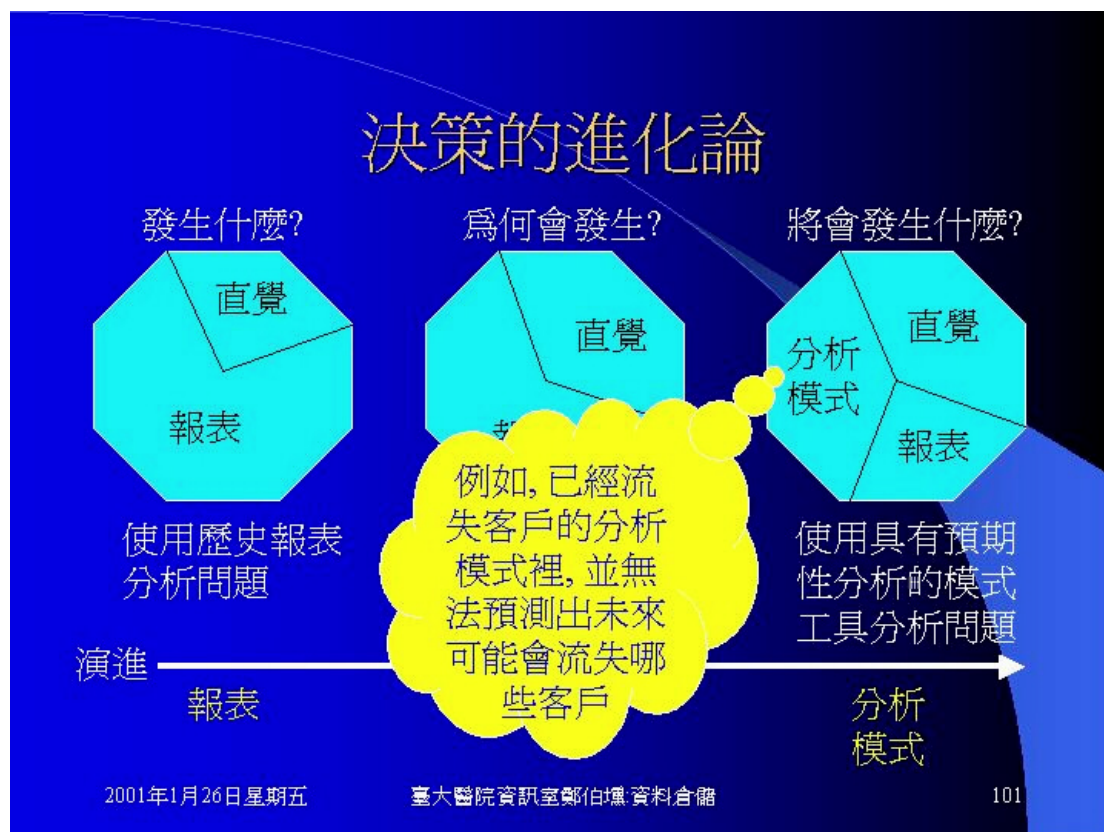
- OLAP
 - ✓ On-Line Analysis Processing，線上分析處理。
 - ✓ 使用者主動。
 - ✓ 使用者產生假說，再用 OLAP 工具查證假說的正確性。
 - ✓ 使用者自行探求(Explore)結果。
- DM
 - ✓ Data Mining，資料採礦。
 - ✓ 使用者被動。
 - ✓ 幫使用者建立假說。
 - ✓ 使用工具探求結果。

5.2.2.30. DW 的未來

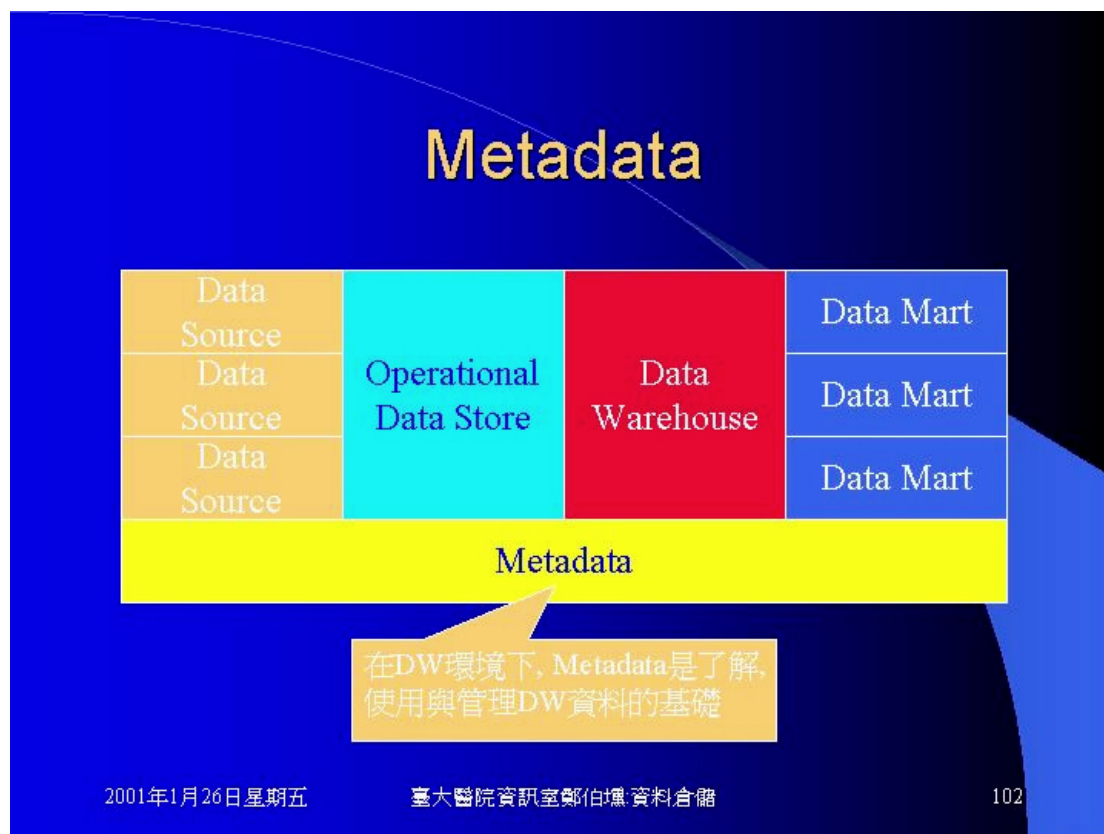
- Metadata 國際規格尚未制定：OMG vs. MDC：2000/10 以 OMG 為主流。
- Cube 國際規格尚未制定：Cube 資料由 OLAP 與 Database 產生。
- 如何有效縮短 Data Cleaning 的工作時間。
- OLAP 將直接透過 SQL 讀取資料，不經由 DW。
- HL7 已納入 DW 工作群組，未來該標準會對 DW 多所著墨。
- 高延展性，集中化 DW，分散式 Datamarts。

(出國類別：研究)

5.2.2.31. 決策的進化論



5.2.2.32. Metadata



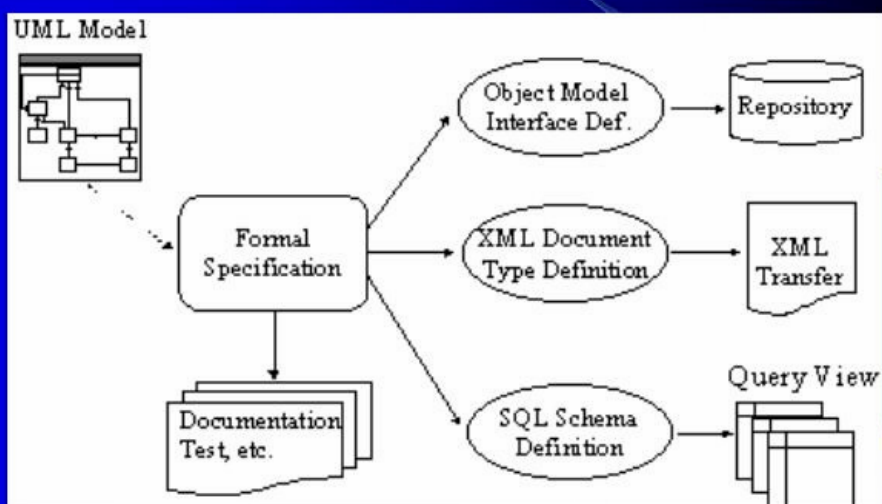
- **OMG (Object Management Group)**
 - ✓ <http://www.omg.org/technology/cwm/>，由 Oracle 等廠商所提倡，2000/09：OMG(Object Management Group)合併 MDC 制定 Metadata 標準。
 - ✓ **CWM (Common Warehouse Matamodel)** 。
 - ◆ 2000/01 公佈。
 - ◆ 說明在 data warehousing，business intelligence，knowledge management，portal technologies 系統間 meta data 資料交換的規格。
 - ✓ **MOF (Meta-Object Facility)** 。
 - ◆ 2000/04 公佈。
 - ◆ 提供一個共用的 Meta Model 來連接兩個不同 Meta Models
 - ◆ 如果兩個不同 Meta Models 都是與 MOF 相容，則可以並存於

(出國類別：研究)

同一個儲存媒體上。

- MDC (Meta Data Coalition) 。
 - ✓ <http://www.mdcinfo.com/>，由 Microsoft 等廠商所提倡 1996/10：成立，1999/08：OIM v1.0(正式)，1999/09：OIM v1.1(草案)，2000/10：合併至 OMG 組織下制定 Metadata。
 - ✓ OIM (Open Information Model) 。
 - ◆ 一組 Meta Data 規格，在使程式開發與資料倉儲可以容易共享與重複使用。
 - ◆ 使用 UML (Unified Modeling Language)來描述規格內容。
 - ◆ UML 可以用來將軟體之抽象層面的知識：視覺化，文件化，指明與建置。
 - ◆ 使用 XML (Extensible Markup Language) 來交換 Meta Data。
 - ◆ 使用 SQL (Structured Query Language)來查詢 Meta Data。
 - ◆ 含 MDIS (Meta Data Interchange Specification) 規格。
 - ◆ 包括
 - Application Metamodel，針對應用程式用 Table 來存放 Metadata 的 Schemas。
 - Meta Data Metamodel，用來說明 Data Dictionary，Data Extraction，Replication，User Query，Database Server 等工具的資訊。

Metadata: MDC-OIM



2001年1月26日星期五

臺大醫院資訊室鄭伯璦資料倉儲

106

- ✓ MDC-OIM v1.0 涵蓋範圍，包括
 - ◆ Analysis and Design Models。
 - ◆ Objects and Components Models。
 - ◆ Database and Data Warehousing Models。
 - ◆ Knowledge Management Models。
 - ◆ Business Engineering Models。
- ✓ MDC-OIM v1.1 增加
 - ◆ Business Rule models。
 - ◆ Knowledge Description Models。

5.2.2.33. 相關參考資源

- 書籍
 - ✓ "Data Warehousing Architecture and Implementation"，Mark Humprise，Michael W.Hawkins，Michelle C.Dy，1999，Prentice

(出國類別：研究)

- Hall , ISBN : 0-13-080902-0 。
- ✓ "Building the Data Warehouse" , William H.Inmon , 1992 。
- ✓ "Building A Data Warehouse For Decision Support" , Vidette Poe , 1996 。
- ✓ "Official Sybase Data Warehousing on the Internet" , Sybase , 1999 。
- ✓ "The Data Warehouse Toolkit" , Ralph Kimball , 1996 , John Wiley & Sons , ISBN : 0471153370 。
- ✓ "Data Warehouse : From Architecture to Implementation" , Barry Devlin , 1997 , Addison Wesley , ISBN : 0201964252 。
- ✓ "OLAP Solutions : Building Multidimensional Information Systems" , Erik Thomsen , 1997 , John Wiley & Sons; ISBN : 0471149314 。
- 研討會，簡報
 - ✓ Microsoft SQL Server's BOL (Books On-Line) 。
 - ✓ "From Data Glut To Decision Intelligence : Better Decisions From Your Systems Without Conversion" , John W.Boyer M.D. , Decision Intelligence Inc. , 2000/05/10 , TEPR Meeting 。
 - ✓ "資訊倉儲結構與應用" , 中華電信資訊處 , 2000/02/23 。
 - ✓ "資料倉儲系統建議架構簡報" , 艾群科技 , 2000/01/25 。
 - ✓ "淺談資料倉儲暨離形系統建置展示" , 國防資訊中心尹寒柏少校 , 2000/01/14 。
 - ✓ "Architecting the Data Warehouse" , Harnam Thandi , 1997 , Sybase Asian Users Group & Powersoft Developers Conference 。
- Metadata
 - ✓ Object Management Group , <http://www.omg.org/technology/cwm/> 。
 - ✓ Meta Data Coalition , <http://www.Mdcinfo.com/> 。
- * 免費技術評估報告
 - ✓ <http://www.dwinfocenter.org/> 。
 - ✓ <http://www.olapreport.com/> 。
- 商業雜誌

- ✓ Intelligent Enterprise , <http://www.intelligententerprise.com/> 。
- 新聞群組
 - ✓ news : //comp.database.olap/ 。
 - ✓ news : //microsoft.public.sqlserver.olap/ 。
 - ✓ news : //microsoft.public.sqlserver.datawarehouse/ 。
- 免費資訊郵件
 - ✓ Dwlist : 第一行標明"subscribe" , 將電子郵件寄至 dwlist-request@datawarehousing.com 。

5.2.2.34. 有關 DW 的建議

- DW 是組織內資料一致性的必要機制，醫療單位可利用 DW 避免資料不一致的情形。
- DW 是 DM 的所不可或缺的基础建設。
- DW 需要龐大的硬體資料儲存設備，如何避免當機發生，應事先想好防治機制。
- DW 會保留所有資料，其他系統各自儲存一定時間內的所屬資料。
- 因為資料來自各單位，整理彙總後儲存於 DW 中，如何定義資料的擁有權需要事先考量。

5.2.3. 資料採礦(Data Mining)

5.2.3.1. DM 定義

- DM = Data Mining = 資料採礦。
- 可以說是
 - ✓ 統計的延伸，在資訊領域裡利用電腦快速運算的特性，將以往的

(出國類別：研究)

統計方法以另一種方式呈現出來。

- ✓ 人工智慧(AI) 與資料庫(DB) 的結合。
- ✓ 資料倉儲(DW) 的一種延伸。

5.2.3.2. DM 應用範圍

- 市場研究。
- 國防。
- 犯罪防治。
- 銀行。
- 地震。
- 天然火災。
- 百貨倉儲。
- 醫療。

5.2.3.3. DM 之工作

- Tasks 包括
 - ✓ Classification (屬性分類)
 - ◆ 依照非連續性數值(如是或否，男或女)分門別類加以定義，建立類組(Class)。
 - ◆ 例如
 - 將信用卡申請者的風險屬性分成高度風險，中度風險，低度風險。
 - 區分哪一之電話號碼同時也是傳真機號碼。
 - 找出詐領保險費的申請案件。
 - ◆ 方法
 - Decision Trees。

- Memory-Based Reasoning。
 - Link Analysis (部分適用)。
- ✓ Estimation (異質推估)
 - ◆ 依照既有連續性數值之相關屬性資料，推估某一屬性可能的未知數值。
 - ◆ 例如
 - 依照信用卡申請者的教育程度，性別等屬性，可以推估其未來的信用卡消費量。
 - 推估某家庭的小孩數目。
 - 推估某家庭的總收入。
 - ◆ 方法
 - Statistics Analysis，Statistics Regression Analysis。
 - Neural Networks。
- ✓ Prediction (同質預測)
 - ◆ 根據過去觀察資料與未來的預測值，預測可能的未知數直，要知道結果的精確度只有等著瞧。
 - ◆ 例如
 - 依信用卡消費紀錄，預測未來的信用卡消費量。
 - 預測哪些客戶會在六個月內流失。
 - ◆ 方法
 - Market Basket Analysis。
 - Memory-Based Reasoning。
 - Decision Trees。
 - Statistics Regression Analysis。
 - Time Series Analysis。
 - Neural Networks。
- ✓ Affinity Grouping (關聯分組)
 - ◆ Affinity Grouping = Market Basket Analysis。
 - ◆ 根據物件屬性的關聯性，決定哪些物件應該放在一起，目的在

(出國類別：研究)

確認 Cross-Selling 的機率，以設計出吸引人的產品銷售群組。

◆ 例如

- 超級市場中依照物件屬性，可以將牛奶，飲料，麵包等食物放在一起。

◆ 方法

- Market Basket Analysis。
- Association Rules。

✓ Clustering (同質分組)

◆ Clustering = Segmentation。

◆ 在異質性的資料中，區分出較具同質性的群組 (Cluster)。

◆ 區分前並不對區隔的方式加以定義，是自然產生區隔的。

◆ 例如

- 透過病歷資料可以區分出 A，B，C，D，E 等肝炎分組。

◆ 方法

- Cluster Detection。
- K-means。
- Agglomeration。

✓ Description (描述)

◆ 對某一複雜資料庫的發展簡單說明以增加了解程度。

◆ 例如

- 美國總統選舉，女性比男性更加支持民主黨。

◆ 方法

- Market Basket Analysis。
- Neural Networks。

- 目前沒有任何工具或技術可以對等解決上面所有的問題。

5.2.3.4. DM 方法論(Methodology)

- DM 基本作業流程

- ✓ Identifying the problem。
- ✓ Analyzing the data。
- ✓ Taking action。
- ✓ Measuring the outcome。
- 類型
 - ✓ Hypothesis Testing
 - ◆ 屬於 Top-Down 方式，企圖證明或否決所謂的先見之明 (Preconceived Idea)。
 - ◆ 步驟
 - 產生一些假說。
 - 決定哪些資料允許上述假說的測試。
 - 取得資料：Data Warehouse 或外部資料。
 - 準備分析資料：Data Transformation，Data Cleaning 等。
 - 建立電腦模型。
 - 統計軟體：SAS，SPSS，S-Plus。
 - 其他 DM 工具，或 Decision Support Database Query。
 - 評估電腦模型：證明或否決假說。
 - 針對結果採取行動，並且評估行動的效果。
 - 使用行動後所產生的有利新資料，重複以上工作。
 - ✓ Knowledge Discovery
 - ◆ 屬於 Bottom-Up 方式，從資料下手找出一些原本未知的事實或原理。
 - ◆ 分成
 - Directed Knowledge Discovery。
 - 會選定特定的目標欄位(Target Field)，如收入，年齡。
 - 解釋(Explain)所發現的屬性關係。
 - 可能處理的問題，例如
 - 什麼產品會與豆腐一起賣出。
 - 當牛奶降價時，什麼產品售價會提高。

(出國類別：研究)

- 誰有可能會買人壽保險。

- 步驟

- 指明已經事先分類的資料。
- 準備分析的資料。
- 將資料分成 Training, Test, Evaluation Datasets。
- 利用 Training Dataset 建立電腦模型，並加以訓練。
- 將電腦模型應用到 Test Dataset 上。
- 評估電腦模型的精準度。
- 針對結果採取行動，並且評估行動的效果。
- 使用行動後所產生的有利新資料，重複以上工作。

- Undirected Knowledge Discovery

- 不指定任何目標欄位。

- 分辨(Recognize)屬性間的關係。

- 可能處理的問題，例如

- 什麼產品會一起賣出。
- 為何某些產品會一起賣出 (Why)?
- 誰會買特別組合的產品 (Who)?
- 何時會賣出特別組合的產品 (When)?

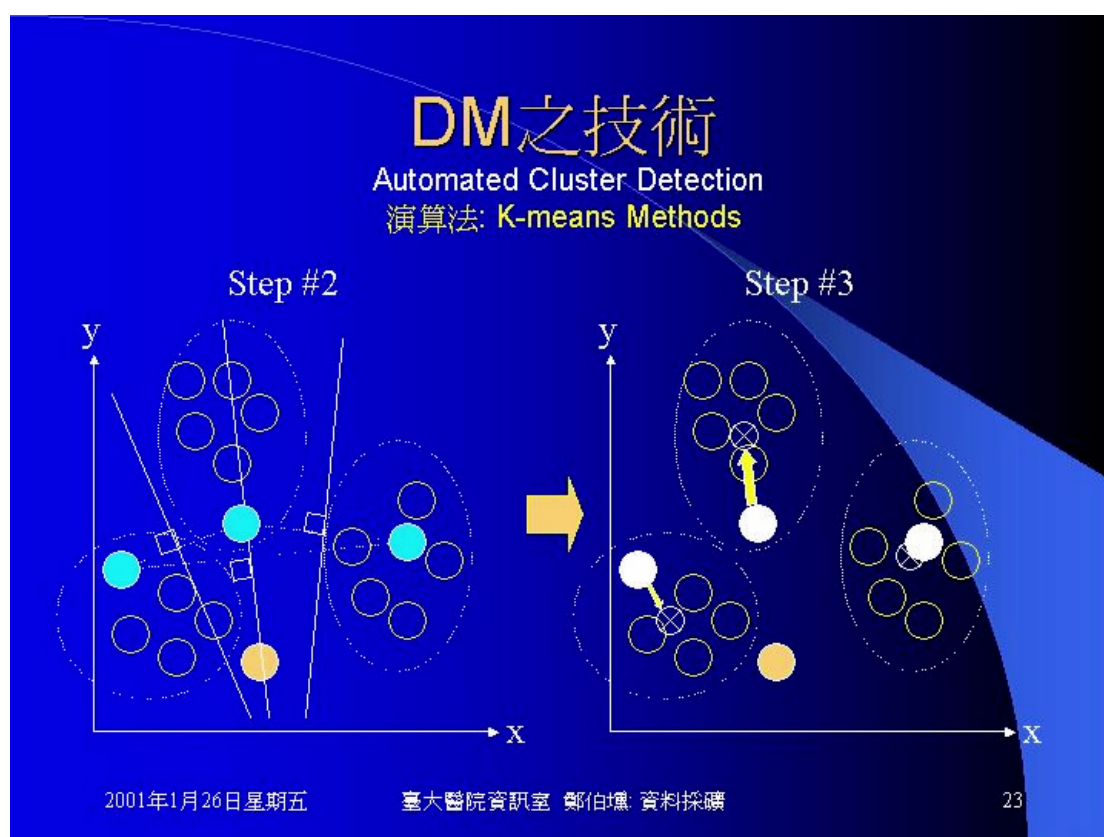
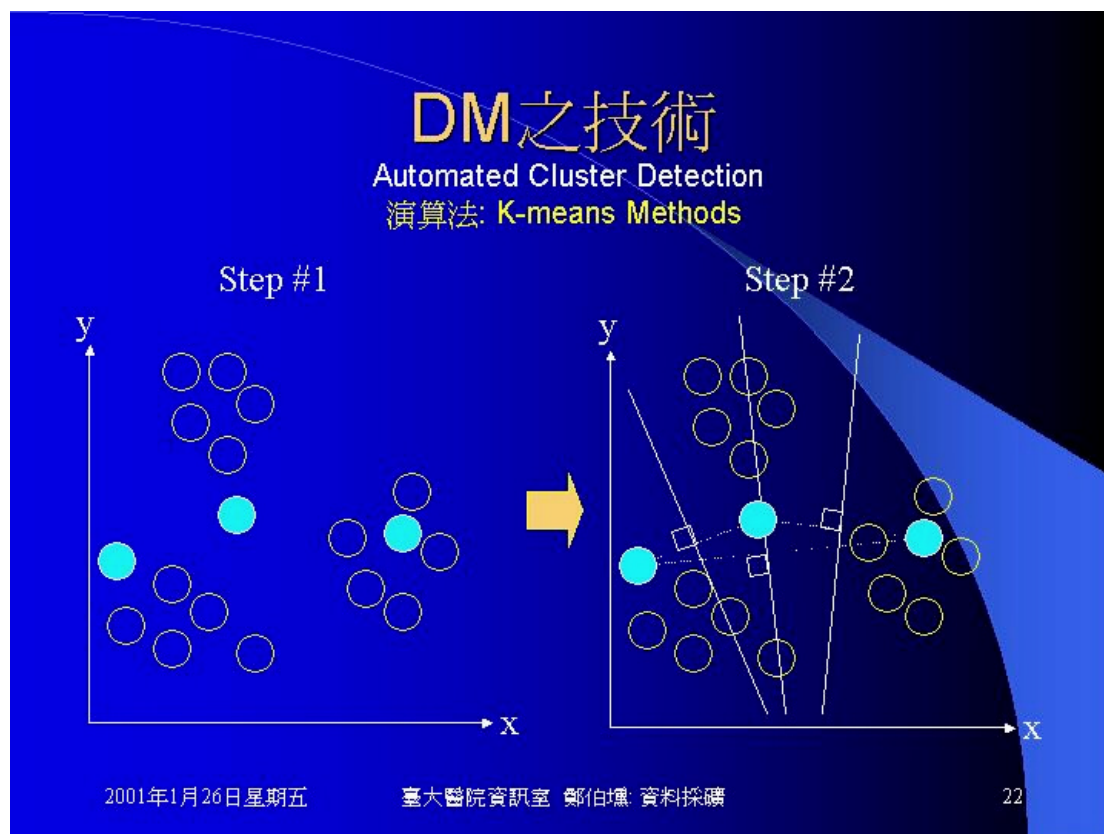
- 步驟

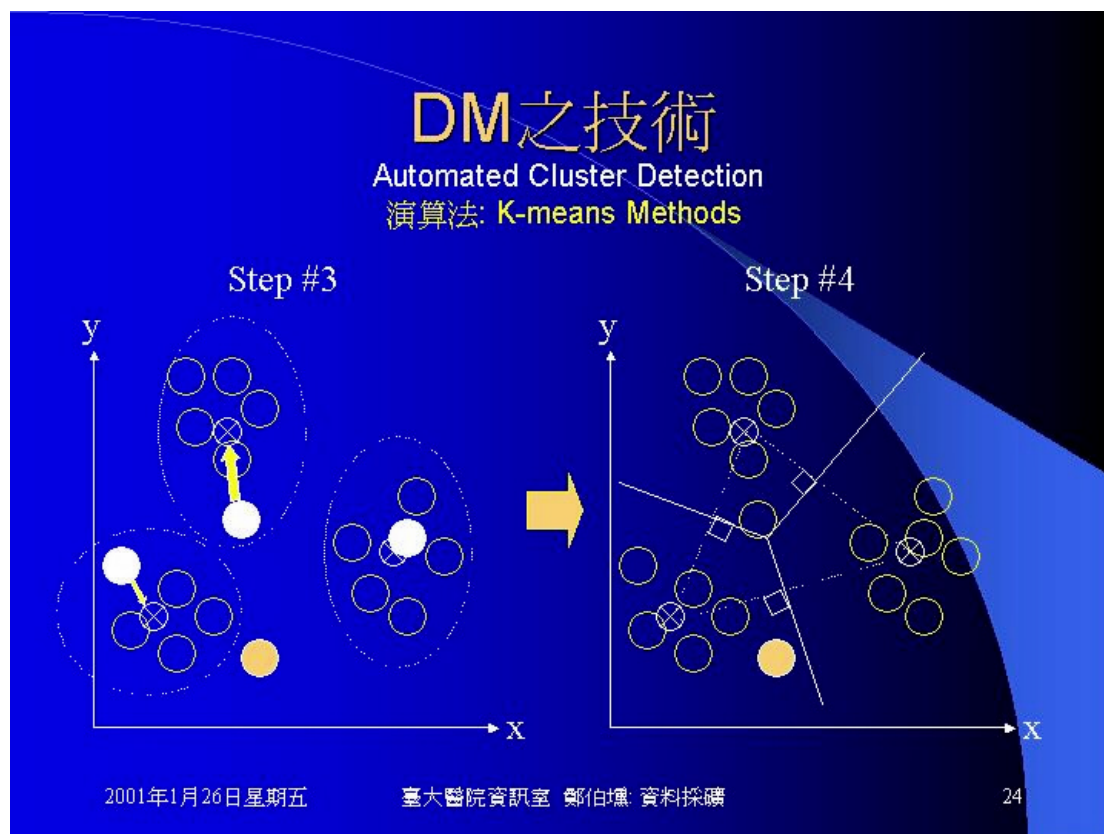
- 指明資料。
- 準備分析的資料。
- 建立電腦模型，並加以訓練。
- 評估電腦模型。
- 將電腦模型應用到新的資料上。
- 指明 Directed Knowledge Discovery 的潛在目標。
- 產生新的測試假說。

5.2.3.5. DM 之技術

- Visualization 技術。
- Non-Visualization 技術。
 - ✓ Market Basket Analysis。
 - ✓ Memory-Based Reasoning。
 - ✓ Automated Cluster Detection。
 - ✓ Link Analysis。
 - ✓ Decision Trees。
 - ✓ Artificial Neural Networks。
 - ✓ Genetic Algorithm。
 - ✓ On-Line Analytic Processing (OLAP)。
- Automated Cluster Detection
 - ✓ 演算法
 - ◆ K-Means Methods：
 - J。B。MacQueen，1967 所提出。
 - 2 維演算法。
 - 取前 K 筆資料作為種子(Seed)。
 - 在種子間取的中分線，將所有資料指定到最接近之中分線的種子上，形成分組。
 - 針對原本各種子所形成之分組資料求取中心點，並將中心點其設定為新的種子。
 - 重複前面步驟。

(出國類別：研究)





- 可以擴充成 N 維演算法。
- Natural Association：被分在同一個 Cluster 中的資料的關係。
- Similarity：被分在同一個 Cluster 中的資料的關係是非常相似的。
- $K = ???$
 - 根據 Michael Anderberg 書中所述之橋牌理論，K 最好是 1，2，3，4，8，12，26，52。
- 各種資料的變數設定方式。
 - Categories：直覺分類，如水果類，蔬菜類，海鮮類等。
 - Ranks： $X > Y > Z$ ，但不一定 $(X - Y) > (Y - Z)$ 。
 - Intervals：兩點間的距離，通常使用 Geometric Distance Metrics 方法。
 - True Measures：對應原點(0,0) 的值，通常可以使用兩向量間的角度求得。

(出國類別：研究)

- Scaling vs. Weighting。
 - Scaling
 - 處理不同變數，且變數是以不同的單位來測量的。
 - 例如：溫度，溼度，壓力之間無法比較，可以設定適當的 Scale 加以處理。
 - Weighting
 - 針對某些變數特別在意，所以加上權重以示區別。
 - 例如：汽車保險費，單身會比已婚的保險費來的貴，所以可能 $\text{Weight 單身} > \text{Weight 已婚}$ 。
- K-Mean 的缺點
 - 無法有效處理重疊的分組。
 - 分組的中心點很容易被其他點所拉開。
 - 每一點都是在分組內部或外部，沒有辦法說明某些分組內的點比分組內其他點更相似一些。
- K-means 方法的變形包括
 - 如何選擇不同的起始種子(Initial Seeds)。
 - 如何計算下一次的分組中心點(Centroid)。
 - 使用 Probability Density 取代 Distance 來分組。
- ◆ Gaussian Mixture Models
 - 是 K-means 方法的一種變形，使用 Gaussian Distribution (Probability Density)取代原有的 Distance 來分組。
 - 通常應用在高維度(High Dimensional)問題。
 - Mixture Model: 因為每一點的機率是混合了許多高斯分佈的總和(Sum)，故稱。
 - Soft Clustering : 因此法最會把每一點依照機率高低，分配到不同的分組，故稱。
 - 演算法
 - 取 K 筆資料作為種子，並視種子為高斯分佈的

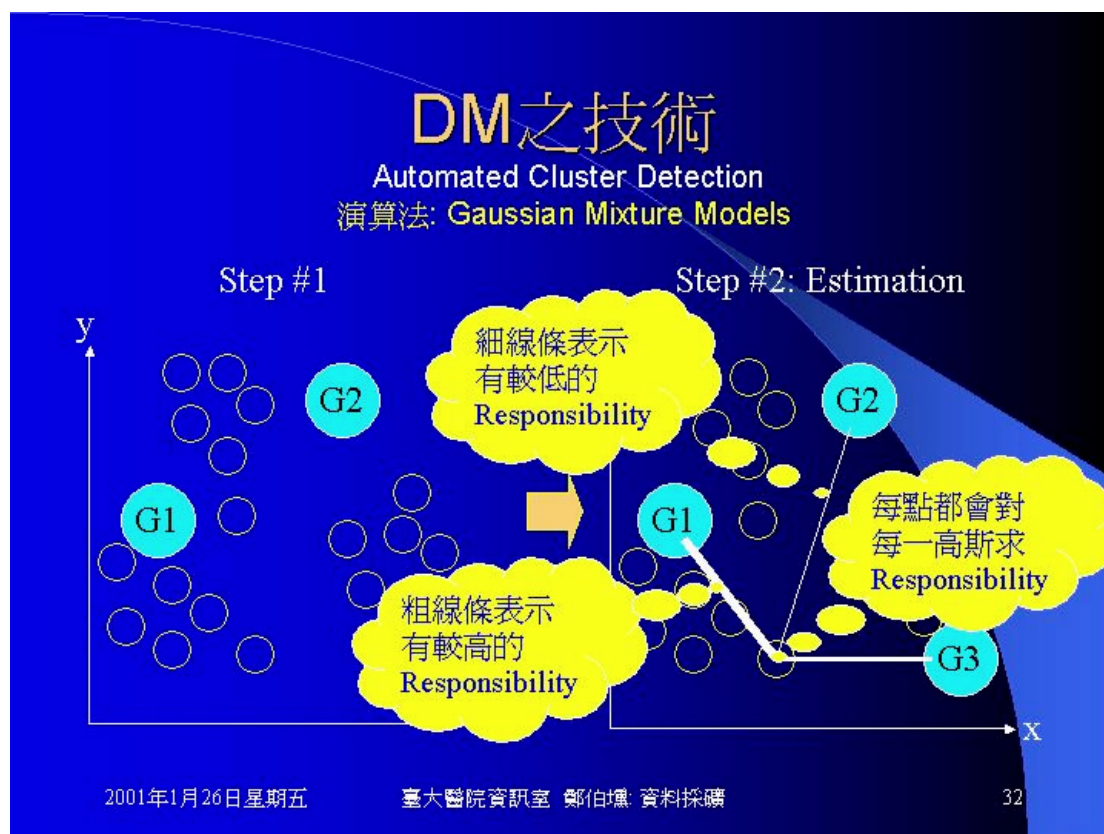
Means。

■ Estimation Step

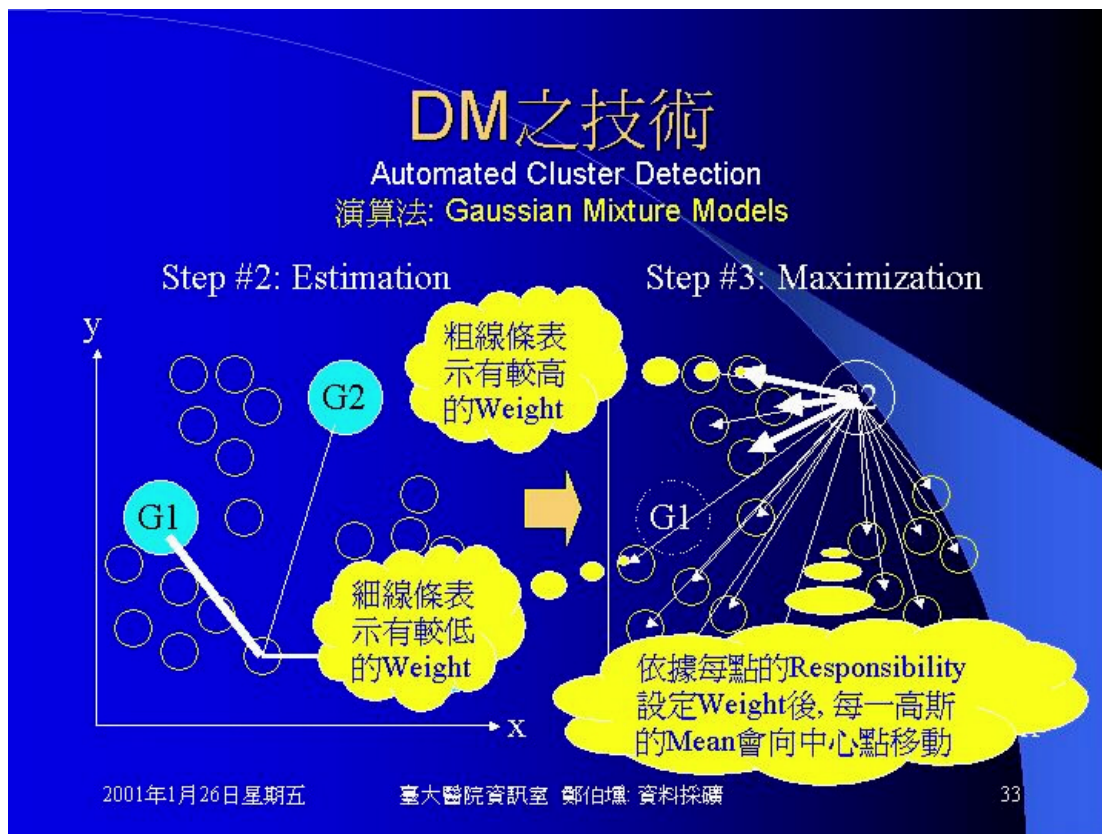
- 計算每一高斯的 Responsibility，並將其作為 Weights。
- 每一高斯具有強的 Responsibility 的點會較接近，反之具有弱的 Responsibility 的點會較遠離些。

■ Maximization Step

- 每一個高斯的 Mean 會向所有資料的中心點移動。
- 重複此步驟直到高斯不動為止。



(出國類別：研究)



◆ Agglomeration Methods

□ K-means vs. Agglomeration (黏聚)

■ K-means

- 使用固定的點開始分組。
- 將所有點分組到各固定的點的分組內。

■ Agglomeration

- 每一點開始時為一分組。
- 隨後將黏聚相近的點直到形成大分組為止。

□ 建立 Similarity Matrix

◆ Divisive Methods

◆ Self-Organizing Maps

5.2.3.6. DM vs. OLAP

● OLAP

- ✓ 使用者主動。

- ✓ 使用者產生假說，再使用 OLAP 工具查證假說的正確性。
- ✓ 使用者自行探求(Explore)結果。
- DM
 - ✓ 使用者被動。
 - ✓ 幫使用者建立假說。
 - ✓ 使用工具探求結果。

5.2.3.7. DM 的未來

- 整體應用上
 - ✓ 會將大多數 Non-Visualization 方法轉化並以 Visualization 方法呈現。
 - ✓ 如何在既有的 DM 軟體裡快速引用最新最快的 DM 演算法。
 - ✓ 如何確認演算法所得到的結果。
 - ✓ 如何 DM 網站資源，導入 Web 與 XML 環境成為 WM (Web Mining) 。
 - ✓ 如何使用 DM 來提高 CRM (Customer Relationship Management) 。
- 醫療應用上
 - ✓ 需要強化時間維度的驗算法 。
 - ✓ 影像部分的 DM 應用 (UCLA) 。
 - ✓ 醫療用專家系統的建置 。
- * Kdnuggets 2000/10 DM 人員年薪調查，有效樣本 310 份，單位千美元

✓ 40 以下	55	18%
✓ 40~60	46	15%
✓ 60~80	70	23%
✓ 80~100	57	18%
✓ 100~120	53	17%
✓ 120 以上	29	9%

(出國類別：研究)

5.2.3.8. DM 相關資源

- 組織
 - ✓ DMI (Data Mining Institute) , <http://www.cs.wisc.edu/dmi/> 。
 - ✓ KDNuggets , <http://www.kdnuggets.com/> 。
- 軟體廠商
 - ✓ IBM (DB2 Intelligent Miner for Data v6.1)
 - ◆ 又叫做 IM for Data v6.1 。
 - ◆ Enables knowledge workers to identify and extract high value business intelligence and communicate results to business users for further analysis and decision making 。
 - ◆ 主從架構 。
 - ◆ 直接連接 DB2 或透過 DB2 Joiner 連接其他非 DB2 資料庫 。
 - ◆ 功能面
 - Analyze business information more effectively
 - A new associations visualizer 。
 - Extended data exploration 。
 - DB2 Intelligent Miner for Relationship Marketing v6.1 。
 - Improve productivity through
 - Visual Warehouse Integration 。
 - Export to spreadsheet 。
 - Copy to clipboard 。
 - Print visualizations 。
 - Control cost w/ new scalability options
 - DB2 UDB (Universal Database) v6.1 support 。
 - SMP (Symmetric Multiprocessor) for AIX , NT , Solaris 。
 - Parallel processing for Value Prediction 。

- Take advantage of powerful resources through enhanced mining solutions
 - SAP Business Warehouse integration。
 - SPSS v10.0 integration。
 - IBM GBIS (Global Business Intelligence Solutions) consultants。
- ◆ 演算法
 - Classification：財物公司透過客戶信用卡使用 Pattern 的 Deviation，可以用來偵測出冒用信用卡。
 - Predictive Modeling：大盤交易商透過客戶購買方式(In-Store 與 Internet)之 Pattern，可以預測未來變化。
 - Association Discovery：超級市場可以找出哪些產品會伴隨其他產品一起賣出。
 - Sequential Pattern Discovery：醫師可以找出導致特別疾病之相同症狀的 Pattern。
 - Clustering：保險公司可以建立市場目標敵人或保險公司對已存在的客戶進行 Cross-Selling。
- ◆ 演算法來源
 - IBM
 - ISV (Independent Software Vendors)。
 - SI (System Integrators)。
 - VAR (Value-Added Resellers)。
- ◆ OLAP Server，內含
 - Essbase OLAP Engine。
 - Hyperion APIs。
 - DB2 RDBMS。
- ◆ 系統架構

(出國類別：研究)

DM相關資源

軟體廠商: IBM (DB2 Intelligent Miner for Data v6.1)
系統架構

DB2 Intelligent Miner for Data v6.1			95/98/NT
UDB Enterprise Ed. Ext. V5.0			AIX
UDB Enterprise Ed. V5.0			OS/2
HPC Cluster Tools v3.0	Parallel Environ v2.2	Genias PaTENT MPI v4.0	
HPC Software Foundation Package V2.0	Parallel System Support Program v2.2		
	DB2 Parallel v1.2		
DB2 V2.1.1			
Solaris V2.6	AIX V4.2.1	NT V4.0	

2001年1月26日星期五 臺大醫院資訊室 鄭伯璦 資料採購 55

DM相關資源

軟體廠商: IBM (DB2 Intelligent Miner for Data v6.1)
系統架構: Client

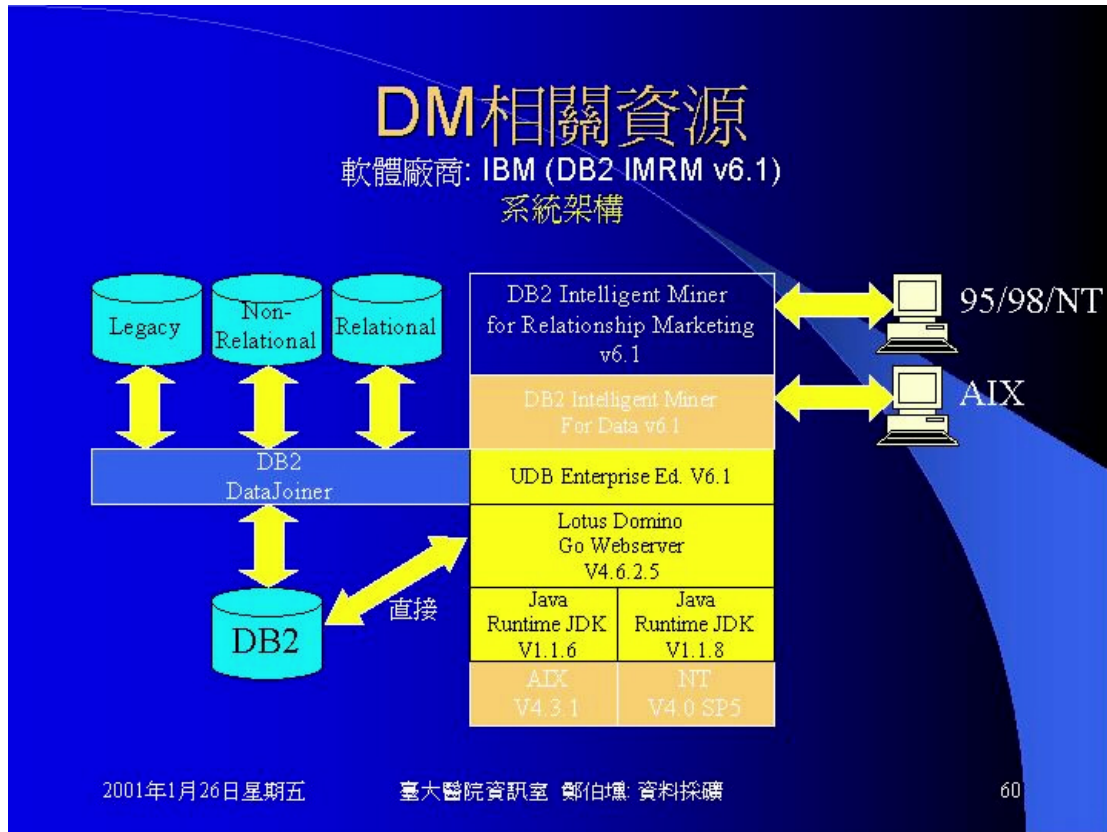
Netscape Navigator V2.0.2	Netscape Navigator V3.0	Netscape Navigator V3.0.1	MS Internet Explorer V3.0
Java Runtime Environment From Java Development Kit (JDK) v1.1.8	Java Runtime Environment From Java Development Kit (JDK) v1.1.6		
OS/2 V4.0	AIX V4.2.1	95/98/NT V4.0	

2001年1月26日星期五 臺大醫院資訊室 鄭伯璦 資料採購 56

- ✓ IBM (DB2 IMRM v6.1)
- ◆ IMRM = Intelligent Miner for Relationship Marketing.

- ◆ A suite of industry- and task-specific customer relationship marketing applications which employ DM in a structured process。
- ◆ 功能面
 - Parallel support on SP and SMP。
 - Logging enhancement。
 - Database access optimization。
 - DB2 Intelligent Miner for Data v6.1 support。
 - DB2 UDB v6.1 support。
 - Euro Currency support。
- ◆ 功能模組
 - Data Preparation Module：指明特定的變數，然後準備 Predictive Modeling 作業。
 - Model Build Module：在測試資料中使用 Tree Induction 或 RBF (Radial Basis Functions)技術，建立與評估不同的 Predictive 模式。
 - Customer Focus Module：使用 Model Build Module 之評分資料作業，輸出目標客戶清單。
 - Business Insights Module：使用 Clustering 技術開發出目標客戶的認知資訊。
 - Data Manager：管理工具。
- ◆ 系統架構

(出國類別：研究)



✓ 參考書籍

- ◆ "Data Mining Techniques : For Marketing , Sales , and Customer

Support" , Michael J.A.Berry , Gordon Linoff , 2000 。

- ◆ "Data Mining Solutions" , Christopher Westphal , Teresa Blaxton , 2000 。
- ◆ "Educated Guessing : How to cope in an uncertain world" , Samuel Kotz , Donna F . Stroup , 1991 。
- ◆ "Developing Analytical Database Applications" , 1999 。
- ◆ "Cluster Analysis for Applications" , Michael Anderberg , 1973 。
- ◆ "Modeling Data for Expert Systems" , IBM Red Books 。
- ◆ "Intelligent Data Analysis : An Introduction" , Michael Berthold , David J . Hand , 2000 , <http://www.springer.de> 。

5.2.3.9. 我的 DM 研究題目

- Clustering for Acute Viral Hepatitis with Numerical Clinical Data 。
- 題目由來
 - ✓ 因肝炎為東方人的常見疾病。
 - ✓ 猛暴性肝炎(Acute Viral Hepatitis)又是另醫師較為困擾的一種肝炎。
 - ✓ 猛暴性肝炎所牽涉到的醫學常識較少，比較適合無醫學背景的我作為研究題目。
 - ✓ 因所需的醫學常識較少，故較適合在 DM 上有更多更深入的探討。
 - ✓ 為了簡化題目的複雜度，目前只探討 DM 其中之一種方法：Clustering 。
 - ✓ 因為醫療資料的範圍至少包括數字與文字，為了簡化 DM 的方法複雜度，目前只考慮數字類型的資料。
- 目標
 - ✓ 肝炎方面的研究有非常多的論文是用統計的方法，但是以 DM 方式的研究尚未有之。
 - ◆ DM 除了 Clustering 方法外，還有許多其他方法可以應用到醫

(出國類別：研究)

療資訊上，至少包括 Decision Trees，Neural Networks 等，而每一種方法又有許多種變形方法，至於哪一種方法最適用於醫療資訊上，目前查無相關文獻。

- ◆ 可以互相對照結果 (統計方法 vs. DM 方法)。
- ◆ 可以發現新的疾病分組 (如果幸運的話)。
- ◆ 可以進一步建立疾病分類專家支援系統 (目前美國已有類似系統，可以在足夠的資料下立即線上告知臨床醫師，由高至低列出可能疾病的百分比，但是以統計方法為基礎所建立的系統)。
- ✓ 希望成為日後非醫學背景人員從事醫療資訊相關研究的一個低門檻研究題目範本與參考，以精省複雜的醫學常識學習過程，進而在資訊方面有較深入的探討與應用。
- 原本擬定的研究思考方向
 - ✓ 方向一
 - ◆ 使用所有的資料用 Clustering 方法套用各種參數找出猛暴性肝炎或所有肝炎的 Clusters，可能的話，可以找出某個小 Cluster 是一般臨床醫師所容易忽略或尚未發現的。
 - ◆ 使用手中既有的 Data Mining 軟體，先行試做一下。
 - ◆ 可以加入時間參數，因為目前沒有現成軟體可以試做，將需要程式設計納入自己的演算法。
 - ◆ 使用所有的資料用 Clustering 方法找出其他疾病的 Clusters，可能的話，可以找出某個小 Cluster 是醫師所容易忽略或尚未發現的。
 - ◆ <1>+<4>可以進一步建立疾病分類專家系統，當系統接收到臨床醫師所給定的症狀與相關檢驗資料後，系統可以使用百分比方式由高到低列出病患最有可能罹患的疾病，提供醫師做進一步參考。
 - ✓ 方向二：找出使用統計方法的猛暴性肝炎的相關論文，再擇一與自行設計的演算法(Clustering，Decision Tree，Neural Network 等)做

一比較，兩者之間哪一個的可信度比較高。

- ✓ 方向三：使用所有的資料找出猛暴性肝炎的 Pattern，與現行醫療臨床診斷教科書中所列出的症狀做比較，找出差異性。

- 研究題目所需的資料：猛暴性肝炎所需資料包括

- ✓ ID。
- ✓ Sex。
- ✓ Age。
- ✓ Diagnosis Code。
- ✓ Lab Data。
- ✓ Clinic Data (理學檢查，主述)。
- ✓ Pathology Data。

- 研究摘要

- ✓ Given
 - ◆ Data from a set of medical patients，where for each patient one has a various number of time-series observations，each time-series of different lengths and different intervals (舉例來說，measured several times a day for a various number of days，depending on the patient)。

(出國類別：研究)

我的DM研究題目

Clustering for Acute Viral Hepatitis with Numerical Clinical Data

研究摘要

		ICD Code	GOT	GPT	X	Y	Z
情況一	0309 10.00						
	0400 10.00			N/A		N/A	
	0400 14.00			N/A		N/A	
	0400 20.00			N/A		N/A	
	0401 10.00			N/A		N/A	
	0401 14.00			N/A		N/A	
	0401 20.00			N/A		N/A	
情況二	0301 10.00						
	0503 10.00	N/A				N/A	N/A
	0504 10.00						N/A
	0400 10.00						N/A
	0400 10.00	N/A				N/A	N/A
	0700 10.00	N/A				N/A	N/A
	0805 10.00						N/A
	0805 14.00						N/A
	0805 20.00	N/A					N/A
	0804 10.00					N/A	N/A
	0804 14.00					N/A	N/A
	0804 20.00	N/A				N/A	N/A

2001年1月26日星期五
臺大醫院資訊室 鄭伯璿 資料採購
69

- Objective
 - ✓ Group patients who show similar patterns on the measurements ◦ The data also includes medical codes describing known diseases each patient has ◦ By clustering we can also check if patients that belong to the same cluster have similar patterns of codes (也就是說，they have the same diseases) ◦
- Complication
 - ✓ Each patient might have observations that are measured on a prior visit to the hospital，long before he or she was diagnosed of the disease ◦ For accuracy，those data should be omitted because they don't represent the measurement of a person who has the disease ◦ The problem is，how do we distinguish the data taken from a prior visit and a current visit，especially if 2 visits are close in terms of time ◦ Note that a visit here can range from days to months ◦ How do we know which set of data is useful and which set is not? It is important to take out un-useful data because they might affect the accuracy of the result ◦

- 研究步驟
 - ✓ Formulate the Problem
 - ◆ Given a model for data。
 - ◆ Define a distance/difference model。
 - ◆ Cluster data using the distance model。
 - ✓ Get some real clinical data eventually
- 參考資料：醫學相關資料
 - ✓ "Acute Viral Hepatitis：Clinical And Laboratory Features，Symptoms And Signs"，DOI 10.1036/1096-7133.ch295，McGraw-Hill，1999/02/10。
 - ✓ "Acute Viral Hepatitis：Differential Diagnosis"，DOI 10.1036/1096-7133.ch295，McGraw-Hill，1999/02/10。
 - ✓ "Progression Criteria for Cancer Antigen 15。3 and Carcinoembryonic Antigen in Meta-static Breast Cancer Compared by Computer Simulation of Marker Data"，Gyorgy Soletormos，Per Hyltoft Petersen，Per Dombernowsky，Clinical Chemistry 46：7，p939-949，2000。
 - ✓ "Description of a Computer Program to Assess Cancer Antigen 15。3，Carcinoembryonic Antigen，and Tissue Polypeptide Antigen Information during Monitoring of Meta-static Breast Cancer"，Gyorgy Soletormos，Vibeke Schioler，Clinical Chemistry 46：8，p1106-1113，2000。
 - ✓ "猛爆型肝炎，醫師也抓狂"，蕭旭峰，2000/06/28，台南縣佳里綜合醫院家醫科醫師。
 - ✓ "General Probabilistic Framework for Clustering Individual"，New Yoik University，2000/03，<http://www.ics.uci.edu/~datalab/papers-by-date.html>。
 - ✓ "Time series"，<http://www.cs.ucr.edu/~dg/kddtut-dg-das.ps>。

(出國類別：研究)

5.2.3.10. 有關 DM 的建議

- 建立 DM 之前，一定要先建置好 DW，而 DW 資料的正確性，將決定 DM 是否成功。
- DM 在醫療的應用需要以醫師作為主角提供相關醫療知識，資訊人員提供適當軟體與程式搭配應用。
- 資訊人員需要熟悉大多數 DM 演算法與相關軟體功能，而選擇適合的 DM 演算法，是一種經驗值。
- 資訊人員最好具有高等統計運算背景。
- 因 DM 價格昂貴，建議應由主管單位主導，提供各醫療院所相關 DM 軟體應用資訊，節省未來不必要的醫療資訊系統投資與浪費。

6. 建議

6.1. 資訊部門的挑戰

現代企業的毒藥視資料太多資訊太少，到處都是電腦，累積了數十億位元的龐大資料，但不知可用的資訊在哪裡隨著企業的成长與規模的擴大，一般公司內部每天要處理的資料量與日俱增身為管理人員，是否曾經為了產生一份報表花了整整一個禮拜的時間來收集，分析與處理各方面的資料，最後再辛辛苦苦的將所蒐集到的原是資料轉化成有用的資訊呢？

是否你也曾經感嘆為何沒有一個系統能夠在輸入查詢後，就能自動將不同來源的資料整理萃取出有用且完整的訊息，而不用花費大批人力，時間，物力去蒐集，甚至能更進一步地自動產生容易讀取，易於了解的報表，協助決策的制定或是問題的解決呢？

是否您也會覺得維護醫療資訊相關系統間的資料交換時須要反覆仔細的對照與保持系統間自訂的資料流介面之完整性？甚至在抽換某一系統時，須要努力保持抽換之新舊系統間的自訂的資料流介面之一致性？

其實上述的問題都可以應用本次的修習內容加以化解，也就是說，可以使用資料倉儲將所辛苦收集的資料完整且正確的保存成為可用與易用的資料超市，然後應用各種資料倉儲的後製作工具將資料轉換化成對組織有用的資訊，也可以進一步應用資料採礦的各種演算法或工具從資料超市中的資料找出一些被組織所忽略但是非常重要的資訊，進一步改變組織的策略與行為模式。而 HL7 國際醫療資訊交換標準則可以更進一步提供醫療機構間資料交換的媒介，讓資料可以應用統一的格式在網際網路上暢行無阻，使醫療資訊的延展性更強。