

資訊的分享

謝清俊

前言—資訊分享的背景

合理的分享資訊是大家的願望，也是國家社會的福祉。在民主社會裡，資訊分享是必須的，在民主社會裡，資訊分享是必須的，否則就無法公平地執行民主政治：如果溝通不良而影響到資訊分享的實行，就會造成特權和許多不平之事，成為民主政治的隱憂和大患。

電腦和通信系統的結合，給現代社會提供了分享資訊的大道，也就理想中能充分資訊的社會帶來了實現的希望。事實上，自從電腦發明以來，這個「分享資訊」的理想就從未被遺忘過：數不清的創作與努力都是朝著這個目標，一次又一次地改良和加強，為的是提供更好的資訊分享環境。

從訂定交換碼（ interchange code ）標準，高階語言的標準，資料庫標準，資料登錄標準， 介面標準...等等， 都對資訊分享的基礎。之後， 有建立程式庫和推動可攜程式（ portable software ）的風潮，以期達到最佳的程式分享；有精簡指令電腦（ R I S C ）的推動，以期建立機器語言的相容性； 有標準索引典（ thesaurus ）和各種專業符號集的訂定，以期減少分享資訊時的語意障礙；有各種通訊標準之訂定，以期構成相容的通信網；...這種種努力，都是為了能讓大家更容易取用到資訊。

最近，關於建立圖形介面（ graphic user's interface ）標準的競爭，建立標準作業系統（如 Open System Foundation, OSF ）的努力，以及開放系統介面（ Open System Inte-face, OSI ）的一連串標準的推出，更在使用者介面，操作系統，電腦網路以及各種應用環境中，如文書處理（ text processing ），資料處理（ data processing ），辦公室文件系統（ office document system ），甚至到銀行、圖書館、...等特定領域中，建立了分享資訊的基石。

最近的這些為分享資訊而訂定的標準，不只給電腦界帶來極大的衝擊，更給整個人類社會帶來了實現真正能充份分享資訊的契機—資訊社會的腳步近

了。給電腦界帶來的衝擊是：這種努力把以往被各大廠商視為禁●的系統軟體，如操作系統及各種系統內的伺服程式（utilities and servers）予以公開化、標準化。這種情形，使得各電腦製造商的基本競爭基礎起了變化，也給我們（發展中國家）帶來千載難逢的機會：這正是迎頭趕上電腦軟硬體和開發軟硬體工業的絕佳時機。我今天想和大家說的，就是在這樣的背景之下，在資訊分享中，關於資料部份，目前的狀況和相關的標準，以及它給我們帶來的影響。

資料分享的模式

資料的分享，一直是個煩人的問題。譬如說，同樣的一篇文章，對於文書軟體（word processor）、編輯軟體（Editor）、排版軟體（typesetters），印刷軟體（press software）、傳輸軟體（communication software）等等都需要更改其格式，才能被這些軟體處理；試想看，如果要出一本書，或是把一篇文章收集到資料庫裡，豈不是要把資料格式一改、再改、三改，煩死人了嗎？能不能以一種資料格式不論對任何電腦中的任何程式而都能通用到底呢？這正是資料分享追求的理想。

現在，就讓我們談談資料分享這個問題。從傳統上來說，兩個系統之間要分享資料，首先要建立個共識，譬如，是什麼資料，有多長，是用什麼碼表示的，有多少欄位，...等等，當這些共識建立之後，雙方才能溝通無阻。因此，要分享資料，就要有系統地建立起這個共識—資料結構的描述；然後依協定好的方法和程序，如行外交禮儀（protocol）一般地，行禮如儀進行資料的交換。所以，在 ISO 訂定關次資料分享的標準中，都是上述的模式進行的。以下，就讓我們對此模式加以說明。

首先，我們要用一種標準語言（或亦稱 protocol）將資料的抽象的資料結構描述和資料本身，就構成了交換的主體。其次，再依一定的規則將此主體以一串標字碼表示之，以便適合循序傳輸的狀態，此稱為序化

（serialization）。經過這樣整理的資料，便可隨時置於網路上和別人交換了。接收的人則反其道而行，只要用同樣的協定，就可以將接收到序化的資料變成資料主體，再以主體中的抽象資料結構描述，去詮釋資料本身是些什麼。這樣的做法，事先無需和接收資料的人做任何接觸；換言之，分享資料的人你不需

認識，甚至是生於不同時代的人；只要大家能用同樣的標準；將資料的抽象結構描述出來，再用同樣的序化協定。這個模定，用圖表示的話，如圖一。（圖表省略）

資料分享的重要標準

在OSI裡，有三種非常重要的標準，是針對應用程式開發者或是使用者而設立的標準，它們是：

1．基本數據表達標準：

ISO 8824 Specification of Abstract Syntax Notation 1(1987 10/ 簡稱

ASN.1 ISO 8825 Specification of Basic Encoding Rules for ASN.1 (1987 11/ 15)，簡稱 BER。

2．格式化資料表達標準

ISO 9735 EDIFACT Syntax Rules(1987 7)簡稱 EDI。

3．文件資料表達標準

ISO 8879 Standard Generalized Markup Language (1986 10/15) 簡稱

SGML ISO 9069 SGML/ Interchange Format (1988 9/15) 簡稱 SDIF。

這些標準，訂定之後陸續有所延伸〔ISO 8824/DADI 1988 12/9，ISO 8825/dadi 1988 12/9 ISO 8879/AMI 1988 7/1〕。第一類標準，也就是ISO 8824，ISO 8825，是屬於OSI中不層次均使用的（layer independent standard），也就是說，在OSI第七層次均使用的（layer independent standard），所有的基本資料型態（basic datatypes）的制式表達方法，都沿用此標準。而第二類和第三類標準，則是屬於OSI應用層（application layer）的標準。

ISO 8824描述對象是基本資料型態；包括 詞語（words），含大小寫字母、數字和連接用的橫線（hyphens） 數目字 字串，含文字串、十六進位字串、二進位字串等 標準符號。在資料種類方面，目前已訂定的列在表一中，請參表。表一的資料類別是以編注（tag）的方式呈現於次

料描述之中。譬如：Tag 值為 2 3，則表示資料為 Coordinated Universal Time 型式；年月日均為兩位數，時間則精確到一秒的解析度。UTC 900320143117 則表示 1 9 9 0 年 3 月 2 0 日 1 4 點 3 1 分 1 7 秒。

Universal Tag	ASN.1 Type
1	BOOLEAN
2	INTEGER
3	BIT STRING
4	OCTET STRING
5	NULL
6	OBJECT IDENTIFIER
7	Object Descriptor
8	EXTERNAL
9	ENUMERATED
2-15	Reserverd for addenda
16	SEQUENCE, SEQUENCE OF
17	SET, SET OF
18	Numeric String
19	Pringable String
20	Teletex String
21	Videotex String
22	IA5 String
23	UTC Time
24	Generalized Type
25	Graphics String
26	Visible String
27	General String
28	Character String
29-..	Reserverd for addenda

表一 ASN.1 Universal Tags

I S O 9 7 3 5 即目前甚為風行的電子資料交換標準（E D I）。此標準是針對傳統資料處理（Data processing, DP）而設立的。處理的對象是欄位化

以格式呈現的資料。譬如：人事、會計、物料、庫存、訂單、銀行作業、圖書館作業...均屬此類之運作。E D I 的每個欄位基本資料型態的標準，仍然基於 A S N · 1。而 E D I 主要訂定的是每筆記錄（**t-transaction record**）的標準。

I S O 8 8 7 9 則是為了非欄位式的資料而設；亦即一般的文件、書本、信件...等。這個標準把電腦處理資料的能力，由欄位化的型式延申到一般性的資料，亦即時下流行所稱的全文資料。這個延申，義意非凡，大大地擴充了電腦對文獻處理的能力，使電腦對文化、社會、語言上產生無比的影響力。I S O 8 8 7 9 名為標準標誌語言，所謂標誌（**markup**）語言也者，即描述文件的抽象結構所用之標準語言也。其實此語言就是排印時在稿子上所作的排印符號或命令。亦類似於文字處理（**word processor**）軟體中用的指令集。所不同的是：在 I S O 8 8 7 9 中所定的是這種標誌語言的語法（**Syntax**）及語意（**semantics**）部分，而指令所用的控制詞彙則可依設計而抽換。這樣的做法，則可適用於不同的商品，各商品間只要以一個表來表示控制詞彙的轉換，就可以互通訊息，達到資料分享的目的。

上述的各種標準，如果和前敘的資料分享模式對照來看，I S O 8 8 2 4，I S O 9 7 3 5，和 I S O 8 8 7 9 是表示資料抽象結構的語言或是協定（**protocol**）；而 I S O 8 8 2 5 和 I S O 9 0 6 9 則是序化的標準。這三種標準涵蓋了主要的資料形式，而建立了分享資料的基本共識和工具。

結語

要建立資訊分享的社會，有四個要件，那就是：

- 四通八達的通信網路—硬體建設
- 分享資訊的協定和標準—分享的共識基礎
- 分享資訊的秩序、以及相關的規範和法令
- 使用資訊設施及詮釋資訊的能力。

我們今天談的，集中在第二項，即分享資訊的協定和標準。這是硬體建設上第一層重要的軟體建設，如果把硬體比作道路，那麼這些協定和標準就如同公路局的售票和乘搭制度；它是達到交通 Communication 或溝通的必要架構。

除了上述的三種標準以外，當然，其他在 O S I 應用層的標準無不是為分享資訊而設；譬如：I S O 7 9 4 2 Graphical Kernd System Functional Description (GKS)是針對圖形 (graphics) 資料而設，I S O 1 0 1 8 0 Standard Page description language (S P D L) 是針對印刷而設，等等，都對資訊之分享有不可缺少的貢獻，並有其特定的價值。然而，今天我們介紹的，對使用人來說，是最為基礎者。如果能熟知這些標準，遵照這些標準開發軟體，則其效果將遠甚於以往傳統的作法。而真正能達到無論遠近，不管今昔，無分國界種族，均可分享資訊的目的。

我們國家要引進這些標準，當然還要大家共同努力；中文化是必需的，我們責無旁貸，從另一方面說，這也正是我們建立軟體工業和建立資訊社會的契機；真的是千載難逢，因為到目前為止，並沒有一個真正 O S I 的系統建立起來，雖然大家、全世界都在往這個方面努力。像 O S I 或是 O S F 的做法：先訂下技術規格和標準並公開發行，這種機緣堪稱空前，以後再要有這種機會亦將少之又少。如果我們不能趁此良機有所作為，搭上這個不花錢的列車把我們軟體能力建立起來，那就怨不得別人，以後只有自己怨自己的份兒。

日本、歐洲、美國等先進國家，對此都重視無比，都有國家型的計劃，推動資訊分享的政策。例如，美國軍方的 C A L S 計劃 (Computer Assisted Acquisition and Legislat-ion and Legislative System)，就是以 S G M L 標準為骨幹，包括 E D I，S P D L，G K S 等等；它包含軍方全部所有的資訊處理設施，最近美國政府亦部份採用，並宣稱到 1 9 9 2 年 (明年)，若任何軟體或資料，無法符合 C A L S 標準的話，軍方將禁止採購。又如，新加坡、大陸、日本所推行的海關通關系統，銀行連線作業系統則均是以 E D I 為基礎所發展的。

反觀國內，雖有些動靜，但是對此大勢仍然若睡夢未醒，一副漠然表情；這就是我選這個題目和大家討論的主要動機；若各位能從今日談話中，獲得一些心得而有所作為，則正是本人深深期則者。