

# 行政院衛生署委託研究計畫書

年 度： 九十一

計畫名稱：推動國際DICOM標準委員會採用中文碼及國內DICOM標準之應用

研究重點： (請填寫所符合研究重點之編號)

申請機構：中華民國生物醫學工程學會

主 持 人：張恆雄

簽 名：

填報日期：91.5.2.

新增計畫

連續性計畫(非新申請者)

註：本計畫書限用中文書寫

# 目 錄

	頁 碼
封面	
目錄	(2)
壹、綜合資料	(3)
貳、計畫摘要	(4)
參、計畫內容	
一、計畫主旨	(5)
二、背景分析	(7)
三、連續性計畫之執行成果概要	(9)
四、實施方法及進行步驟	(10)
五、參考文獻	(12)
肆、所需經費	(13)
伍、執行期限	(15)
陸、預期成果	(15)
柒、參與人員	(15)
附件一	(16)
	共(19)頁

## 貳、計畫摘要：

關鍵詞：DICOM 欄位，中文碼，工作小組。

本計劃為因應醫療資訊之進步及電子病歷之實施後與 DICOM 標準相關事宜之推動。我國已加入 DICOM 標準委員會成為其會員，並實際參與其會務之運作，於醫療資訊的國際化與本土化，促進相關產業競爭力及提昇民眾醫療服務之要求下，必需就相關事宜研擬方法以配合時代進步之潮流。

本年度將藉由以下之工作來達到預期之目標，包含有：1.積極參加委員會之例行會議，並促使更多我國之專家投入其工作小組研擬各標準草案。2.提出中文碼之規範，使中文碼早日成為標準，以促成國內外儀器廠商在 DICOM 中文化方面之規劃設計。3.與業界及醫界合作，訂定不同醫學影像之 DICOM 欄位，使標準更落實。4.成立台灣 DICOM 協會，舉辦相關之研討會、教育課程，並開放網站，以方便資訊之交流。希望藉由上述的活動及工作，使成果能落實到國際之標準上，並因此提昇醫療資訊業者之競爭力及醫療服務品質。

## 參、計畫內容

### 一、計畫主旨：

本計劃主要係針對如何推動國際 DICOM( Digital Imaging and Communication Medicine ) 之標準在國內落實之相關工作，其主要目標可從國際化、本土化、產業競爭及提昇醫療服務等層面說明，並針對各目標提出可能因應之道。

- 1、標準之國際化：因應目前醫療環境（尤其是有關醫學影像設備之部分）一直在進步當中，PACS 已達實用階段。為了提昇各種不同儀器之相容性，故 DICOM 之標準因應而生，且此標準亦不斷的修正當中，我國在儀器上一向進口居多，故對相關之標準不能忽視。對制訂該項標準之國際組織（DICOM 標準委員會），更需積極參與討論，以做為未來訂定本國標準之準備。
- 2、融入本土化：近年來，因應時代進步，電子病歷逐漸成形。我國亦從今年四月開始實施電子病歷，初期架構較簡單；以後之發展，可能將影像亦列入。因此影像檔中，如何加入一些中文訊息，將是可能發生之問題。目前 DICOM 標準中對於我國所用的中文編碼方式尚無定義，其規範目前正在研議當中，除了參與國際組織會議可要求將我們之需求列入外，其先期之工作有關中文碼如何加入之探討該為一個重要之工作項目。
- 3、提昇產業之競爭性：於研擬中文碼加入標準之同時，若產業界亦能共同參與，則累積此方面之經驗，對國內產業界開發支援中文的 PACS 系統之能力有所提昇，並且亦可協助國外廠商從事此方面之開發工作，提昇國內醫療軟體業者之能力。
- 4、提昇醫療服務：由於標準之加入，使影像之流通性更高，對促進國內相關系統之發展將有所幫助，並且因此促成影像之加入電子病歷，將可使民眾避免一些不必要之檢查，減少舟車勞頓；另一方面，亦可減少國家整體之醫療開支之浪費，並進而提供更多醫療之服務。

綜合上述，本計劃擬由生物醫學工程學會之 DICOM 工作小組及產、官、學、醫界之專家、學者共同進行 DICOM 相關工作之規劃及推展，並積極從事下列幾個工作項目：

- 1、參與國際組織：推派代表參加 DICOM 標準委員會之例行會議，並進而推動專家加入並參與 DICOM 之各個工作小組（Working Group），發揮會員國之影響力。
- 2、協助有關中文碼進入 DICOM 之標準：主要是對現階段各種中文碼作進一步之整理，提供 DICOM 標準委員會參考。DICOM 標準採用中文碼可促成國內外儀器廠商在中文化方面之規劃設計，以利於醫療院所採用符合 DICOM 標準之醫療儀器。
- 3、推動 DICOM 相關之研討教學活動：延請國內外專家學者針對 DICOM 之一些規範進行研討或提供短期課程，使大家對 DICOM 有更進一步的了解。
- 4、實務性推廣工作：各科醫學影像所需的 DICOM 欄位各有不同，本學會將與放射醫學會討論訂定不同影像之 DICOM 格式，今年將對正子電腦斷層影像及傳統 X 光片的 DICOM 欄位、格式的訂定作規範建議，另將對牙根間 X-光片可能用到的 DICOM 格式作分析。正子電腦斷層影像目前已有初步之建議格式，今年度之工作除與放射醫學會作確定此格式外，再加入負責傳遞醫囑及排程資料之 worklist 功

能與醫院資訊系統連接。worklist 內定義許多與造影有關的醫院資訊，像醫生的檢查及造影的時間排程等。

- 5、籌設成立台灣 DICOM 協會，擴大參與人員以進行推廣事宜：生物醫學工程學會 DICOM 工作小組已成立一年多，今年將進行成立台灣 DICOM 協會之工作，依內政部規定申請組織全國性之協會，繼續負責推廣 DICOM 工作，加強對醫界及產業界進行宣導，並促進其對 DICOM 之瞭解。
- 6、成立相關網站，並定期維護：網站之建立有助於資訊之流通及意見之交換，網站之架構已有初步之規劃，今年將加速建置及開放 DICOM 網站，提供醫療院所最新之 DICOM 資訊。

## 二、背景分析：

### 1、政策或法令依據

- (1)、衛生署 89.5.16.衛署企字第 89027608 號  
「推動台灣加入國際醫療資訊標準(HL7 及 DICOM)相關協會事宜」第一次會議記錄
- (2)、衛生署 89.6.26.衛署企字第 89034379 號  
「推動台灣加入國際醫療資訊標準(HL7 及 DICOM)相關協會事宜」第二次會議記錄
- (3)、衛生署 90.6.11.衛署企字第 0900038953 號  
「衛生醫療資訊交換中文碼相關事宜諮商」會議記錄
- (4)、衛生署 90.12.25.衛署資訊字第 0900080397 號  
「制定及推廣醫療資訊標準」專家會議會議記錄
- (5)、衛生署 91.4.15.衛署資訊字第 0910027861 號  
「九十一年度醫療資訊標準 DICOM 推動計畫」會議記錄

### 2、發展需求

隨著醫院資訊電子化的腳步，各式相關儀器之輸出已轉為數位化。數位化的資料若要互通，建立一套共用的標準是必需的。在國際上也因應此趨勢，故分別訂定許多醫療資訊的相關標準，其中又以 DICOM(Digital Imaging and Communication Medicine)及 HL7(Healthcare Level 7)為最重要，此類標準雖已訂立，並已試行中；然而隨著時間的腳步，每年有固定的聚會來進行修正，故有必要派員參與會議。

目前 DICOM 標準中對於我國所用的中文碼的編碼方式並沒有定義。若能將中文的編碼方式加入 DICOM 標準當中，不但國內醫院及廠商可依此標準開發支援中文的 PACS 系統，亦可促進國外儀器廠商在這方面設計規劃，未來國外進口的儀器可以支援中文。要訂立 DICOM 中文編碼標準，需要先期探討現有 DICOM 標準當中對於文字資料編碼方式的相關規定。

PACS 在過去幾年當中，已從制定階段發展到實用階段，目前國內亦有許多醫院積極地在發展及規劃本身的 PACS 系統。為了使 PACS 系統能發揮最大效益，使用者及業者亦應對 DICOM 有初步瞭解，故延請國內外專家學者針對 DICOM 之一些規範進行研討或提供短期課程，應有幫助。

### 3、國內外相關工作之文獻探討

本計畫係對於我國醫療影像儲存、交換、與傳輸標準 DICOM [1-2]部分進行相關之推廣工作。DICOM 主要是訂定各種醫療影像相關資訊之共同交換格式，其發展的主要目的是推動開放式與廠牌無關的醫療數位影像的傳輸與交換，並促使影像儲存與傳輸系統(PACS)的發展與各種醫院資訊系統(HIS, Hospital Information

System)整合,同時可允許所產生的診斷資料庫能廣泛地經由不同地方的設備來存取[3-11]。PACS 在過去幾年當中,已從研發階段發展到實用階段,目前國內亦有許多醫院積極地在發展及規劃本身的 PACS 系統。

目前 DICOM 標準中對於我國所用的中文碼的編碼方式並沒有定義。若能將中文的編碼方式加入 DICOM 標準當中,不但國內醫院及廠商可依此標準開發支援中文的 PACS 系統,亦可促進國外儀器廠商在這方面設計規劃,未來國外進口的儀器可以支援中文。DICOM 標準訂定是一個繁複而嚴謹的過程,參加 DICOM 標準訂立的工作群組可預先獲知在醫學影像相關領域最新的規劃趨勢,這方面的知識對提升國內醫療資訊產業的競爭力非常重要。

要訂立 DICOM 中文編碼標準,需要先期探討現有 DICOM 標準當中對於文字資料編碼方式的相關規定。在現有 DICOM 標準加入中文編碼方式將會牽涉更動到其中哪些章節,以及需要新增哪些內容[12-13]。以 DICOM 當中日文編碼標準為例, DICOM 中日文編碼規定於 DICOM Supplement 9 當中, Supplement 9 Multi-byte Character Set Support 是日本 JIRA (Japanese Industry Radiology Apparatus) 與 DICOM 協會協同訂立。在此當中雖然已規劃了 Multi-byte 的編碼方式,但其主要偏重於日文本身片假名、平假名及日文漢字的編碼相關規定。對於國內常用之 BIG 5 編碼方式並無規定。在 Part 5 Annex H 訂立日文片假名及平假名及日文漢字的字元集。日文片假名及平假名分別有 94 個字元,日文漢字共 96\*96 個字元,每個字元以兩個 byte 表示,字元與編碼數值的對照表分別參照 ISO-IR 13, ISO-IR 14, ISO-IR 87, ISO-IR 159 當中的規定。另外在 DICOM 針對 Part 2, Part 3, Part 4, Part 6 亦將其與此增訂字元集有關的部分做修改。

除了對 DICOM 的文件及涉及中文碼方面的相關工作外,我們亦希望能在計畫期間積極參與 DICOM 組織相關之學術討論,加入 DICOM 與未來中文碼制訂有關的工作群組 (DICOM WG 6:Base Standard),以及瞭解其他有密切相關的工作群組 (DICOM WG 14:Security),(DICOM WG 20:Integration of Imaging and Information Systems)等[14]。透過這些方式,希望瞭解 DICOM 標準制訂的詳細過程,評估及規劃制訂 DICOM 中文編碼標準所需的工作內容,才能初步提出制訂 DICOM 中文碼的文件,並提出制訂 DICOM 中文編碼標準的申請。

#### 4、本計畫與醫療保健之相關性

參加 DICOM 標準訂立的工作群組可預先獲知在醫學影像相關領域最新的規劃趨勢,這方面的知識對提升國內醫療資訊產業的競爭力非常重要。由於 DICOM 標準之加入,使影像之流通性更高。除了對促進國內相關發展將有所幫助,並且可使影像之加入電子病歷有所規劃。如此,將可使民眾避免一些不必要之檢查,減少舟車勞頓;另一方面,亦可減少國家整體之醫療開支之浪費,並進而提供更多醫療之服務。

### 三、連續性計畫之執行成果概要

過去執行成果概要說明如下：

1. 於民國 89 年，學會接受衛生署委託，規劃加入 DICOM 標準委員會之事宜，並於當年 11 月底派出四位代表參加北美放射線醫學年會，並向 DICOM 委員會提出申請，於當年獲通過，成為該委員會會年。
2. 成為會員會後，常常收到該委員會資料徵詢相關意見，並適時參與該委員會之定期會議，去年因 911 事件之故，故分別於京都、柏林、芝加哥舉行三次會議，其中京都之會議因準備不及未參與外，其餘兩次均派員參加，共同討論。
3. DICOM 之網頁已完成初步規劃，尚待更進一步安排維護後，可開放供大家討論。
4. 有關建立 DICOM 認證環境：目前由於中山科學院有參加 IHE 之計畫，故也有了初步 IHE 之架構，本小組成員亦適時加入討論提供意見，未來可依原訂架構完成認證環境之建立。
5. 去年亦參與牙醫學會之有關 DICOM 之討論，並於一些學術研討會提出 DICOM 之報告，提高大家對 DICOM 參與之興趣。
6. 有關中文碼之工作：除蒐集相關資訊外，亦提供給 WG6 之主席 Robert Horn 參考，目前 WG6 已有一些初步之建議 [見附件一]。



#### 四、實施方法及進行步驟：

(一)、以下針對五大工作項目之實施方法及進行步驟作說明:

- 1、參與國際組織: DICOM 標準由 DICOM 標準制定委員會訂立，標準制定委員會的成員包含了生物醫學相關的組織、儀器製造廠商、銷售公司、其他標準訂立組織、政府代表、以及全球對醫學影像標準制訂有興趣的人員。只要加入標準制定委員會就有權力對 DICOM 標準的內容提出增修的提案，對我國在 DICOM 中明定中文碼的需要可以得到適當的討論。我國已成功的加入該標準委員會。加入該委員會後，必須積極參與該委員會各會議，推派代表參加 DICOM 標準委員會之例行會議，並進而推動專家加入並參與 DICOM 之各個工作小組 (Working Group)，發揮會員國之影響力，才能確保會員之地位(不能連續缺席二次)。本年度之會議共有四次，分別為 5/7 在韓國漢城，6/25 在法國巴黎，9/24 在美國華盛頓及 12/5 在美國芝加哥舉行，配合計畫之進展將擇兩次分別派員參加。至於 WG 方面將鼓勵加入 DICOM 與未來中文碼制訂有關的工作群組 (DICOM WG 6:Base Standard)，以及瞭解其他有密切相關的工作群組 (DICOM WG 14: Security)，(DICOM WG 20: Integration of Imaging and Information Systems)。
- 2、協助有關中文碼進入 DICOM 之標準: 主要是對現階段各種中文碼作進一步之整理,提供 DICOM 標準委員會之參考 其中包括 Big-5, Unicode(UTF-8), GB13080, CNS11643, ISO10646 等與現有 DICOM 標準溶入的問題。附件一為 DICOM 委員 Robert Horn 最近提出的有關中文碼在 DICOM 標準中之使用方式，我們認為此須與其作進一步討論以符合我國之需要(參見衛生署 90.6.11.衛署企字第 0900038953 號「衛生醫療資訊交換中文碼相關事宜諮商」會議記錄)。在六月份之 DICOM 會議，邀請中央研究院曾士熊先生共同赴會在 DICOM 委員會中作說明並提出建議。此為一持續性之工作，直到 DICOM 標準中列入中文碼之規範。
- 3、推動 DICOM 相關之研討教學活動: 延請國內外專家學者針對 DICOM 之一些規範進行研討或提供短期課程，使大家對 DICOM 有更進一步的了解。目前由於中山科學院有參加 IHE 之計畫，工研院亦有參與 DICOM 之工作，另外有國外亦有一些負責 DICOM 標準之專家，以上人員今均可擔任講員，此類活動預計辦理二次。本計畫擬於年度內邀請二位國外講員來台針對 DICOM 中文碼及 DICOM 最新發展作研討。
- 4、擴大 DICOM 小組之成員，並進行推廣事宜: 今年內將申請成立台灣 DICOM 協會，加強對醫界及產業界進行宣導，並促進其對 DICOM 之瞭解。主要是利用舉辦 DICOM 相關之研討教學活動使醫界及產業界對 DICOM 有所瞭解並促使他們有意願加入各個工作群組，為未來 DICOM 之推動鋪路。
- 5、成立相關網站，並定期維護: 網站之建立有助於資訊之流通及意見之交換，網站之架構已有初步之規劃，今年將開放此一網站。預定在網頁中放置目前 DICOM 之標準，最近之發展，小組活動，及意見交流等項目。

(二)、預定進度甘梯圖 ( Gantt Chart ):

月次 工作項目	第 1 月	第 2 月	第 3 月	第 4 月	第 5 月	第 6 月	第 7 月	第 8 月			備 註
中文碼相關工作	**	**	**	**	**	**	**	**			
DICOM 實務研習會			**				**				
DICOM 研討會					**						
期中報告					**						
與放射醫學會訂定正子電 腦斷層影像之 DICOM 格 式		**	**	**	**	**	**	**			
與牙醫學會探討牙根間 X- 光片的 DICOM 格式		**	**	**	**	**	**	**			
籌備成立台灣 DICOM 協 會	**	**	**	**	**	**					
台灣 DICOM 協會成立大 會							**				
DICOM 國內網站之建立 及開放	**	**	**	**	**	**	**	**			
撰寫報告								**			
期末報告								**			
預定進度累計百分比	10%	25%	40%	50%	60%	70%	85%	100 %			

## 五、參考文獻

1. Bidgood WD Jr; Horii SC: Introduction to the ACR-NEMA DICOM standard. *Radiographics*. 12(2):345-55,1992.
2. Best DE, Horii SC, Bennett W, Thomson B, Snavely D: Review of the American College of Radiology--National Electrical Manufacturers' Association standards activity. *Computer Methods & Programs in Biomedicine*. 37(4):305-9,1992.
3. Stewart BK, Langer SG: Integration of DICOM images into an electronic medicalrecord using thin viewing clients. *Proceedings / AMIA Annual Symposium*. Pp.:902-906,1998.
4. Behlen F: A DICOM document-oriented approach to PACS infrastructure. *Journal of Digital Imaging*.11(3 Suppl 1):35-38,1998.
5. Oosterwijk H: DICOM versus HL7 for modality interfacing. *Journal of Digital Imaging*.11(3 Suppl 1):39-41,1998.
6. Kimura M, Ohe K, Yoshihara H, Ando Y, Kawamata F, Tsuchiya F, Furukawa H, Horiguchi S, Sakusabe T, Tani S, Akiyama M: MERIT-9: a patient information exchange guideline using MML, HL7 and DICOM. *International Journal of Medical Informatics*. 51(1):59-68,1998.
7. Gerritsen MGJM, van der Putten N, Dijk WA, Dassen WRM, Spruijt HJ, Spruijt EJ, Uijen GJN, Hamers R: General DICOM PACS server for echocardiography images. *Computers in Cardiology*, 26: 431-434, 1999.
8. Prior F: Database access methods for medical imaging: DICOM, SQL and HTML. *Proceedings of the Fourth International Conference on Image Management and Communications*, 1995:288-292.
9. Wahle A, Builtjes JH, Oswald H, Fleck E: DICOM-integration in a heterogeneous environment. 18<sup>th</sup> Annual International Conference for the IEEE Engineering in Medicine and Biology Society, 1996(3): 1228-1229.
10. Bidgood WD Jr: Clinical importance of the DICOM structured reporting standard. *International Journal of Cardiac Imaging*. 14(5):307-315, 1998.
11. Solomon HP: Integration of haemodynamic and electro-cardiographic waveform data with DICOM images. *International Journal of Cardiac Imaging*. 14(5):301-306, 1998.
12. NEMA DICOM 3 Part 5: Data Structures and Encoding.
13. NEMA DICOM 3 Supplement 9 Multi-byte Character Set Support.(修訂中)
14. NEMA DICOM 3 Part 15:Security Profiles

伍、執行期限：民國九十一年五月十五日至民國九十一年十二月十日

#### 陸、預期成果

- 1、參加兩次 DICOM 標準委員會會議，向 DICOM 標準委員會提出在 DICOM 標準中增訂中文碼之申請書。
- 2、定期舉辦之 DICOM 相關論壇、研習會及研討會，可推動 DICOM 標準在國內之使用及醫院規劃相關系統與採購儀器的參考。
- 3、開放 DICOM 國內網站，公佈國內、外各項與 DICOM 相關之最新訊息。
- 4、與放射醫學會訂定正子電腦斷層影像之 DICOM 格式。
- 5、與牙醫學會提出對牙根間 X-光片的 DICOM 格式建議。
- 6、成立台灣 DICOM 協會，配合推動 IHE Taiwan 計畫在國內之工作，使 HL7 與 DICOM 之標準能一起使用。

#### 柒、參與人員

計畫主要參與人員為與美國 DICOM 工作小組接觸過的人員就前次會商之基礎繼續進行，人員包括中央研究院曾士熊先生、陽明大學高材教授及蕭嘉宏博士、中原大學蘇振隆教授、成功大學孫永年、詹寶珠教授、中山科學院羅國璋組長、陳世英先生、高雄榮總楊宗龍醫師及工研院生醫中心林康平組長等，參加在美國、法國舉行的與 DICOM 協會有關的相關會議，如 DICOM Standard Committee meeting, IHE Year 3 Plan test (Chicago, USA), RSNA Annual Conference (Chicago, USA), 及 Annual HIMSS Conference 等。

## 附件一

robert.horn.b@us.agfa.com on 02/18/2002 08:37:14 PM

Please respond to wg6@lists.nema.org

To: "Working Group 6" <wg6@lists.nema.org>  
cc:  
Subject: A radical new proposal for Chinese in DICOM

I have a revised and different technical proposal for handling Chinese text in DICOM messages. I believe it meets the following major goals:

- 1) New objects that contain DICOM compliant Chinese text should not induce errors when received by old DICOM compliant implementations or by new implementations that do not support Chinese. The Chinese text should be processed in accordance with the existing DICOM rules for unknown character sets.
- 2) The UTF-8 character system should be supported to meet the desires of Taiwan and Hong Kong.
- 3) The GB 13080 character system should be supported to meet the desires of mainland China. I have heard both GB 18030 and GBK requested. There are some legal requirements for GB 18030 but it is unclear whether they apply to DICOM. Since GBK is a compatible subset of GB 18030 I think that supporting GB 18030 should meet both needs. Clarification of this is needed. The technical proposal would not be significantly changed if GBK is the proper character set to specify.

The ISO 2022 encoding rules have two features that DICOM has not used before, but which appear simple to add and which meet these goals. DICOM implementations that already support these ISO features and DICOM implementations that only did the minimum necessary for previous character sets should be protected by some of the extra restrictions that are included in this technical approach.

The ISO 2022 features are:

- 1) "DOCS with Return" escape sequences. DOCS stands for Designate Other Coding Set. The escape sequence <ESC> 02/05 04/07 is defined for ISO IR 196: UTF-8 without implementation level. There is a standard generic "return" sequence defined: <ESC> 02/05 04/00 that means return to ISO system.

The effect is that the combination:

<ESC> 02/05 04/07 ....UTF-8 string....<ESC> 02/05 04/00 is an ISO 2022 encapsulation for UTF-8.

ISO also defined IRs for UTF-8 with implementation level and other Unicode encodings. None of these others support the "return" because the escape sequence for return could possibly occur within the text string. So DICOM will need to specify Unicode implementation level separately as part of a DICOM specific name for the character set. This is not a problem.

- 2) The private encodings with return. We can define a private encoding escape sequence: <ESC> 02/03 04/00 to switch to a character set encoding that is not yet assigned an ISO IR. The restriction on this is that the "return" code: <ESC> 02/05 04/00 should not be used within the private system. Neither the GB 18030 nor the GBK utilize these three bytes in any character encoding, so both meet this requirement.

The effect is that the combination:

<ESC> 02/03 04/00 ....GB18030 string .....<ESC> 02/05 04/00 is an ISO 2022 encoding for GB18030 that will be recognized as a "private coding" by a generic ISO 2022 processor, and DICOM can proceed to define its behavior.

Note that the "return" for private encodings is the same as "return" for DOCS.

When ISO defines a "DOCS with return" code for GB 18030, we can make the decision on whether to add that escape sequence as an alternate form. The ISO encoding for GB 18030 as a "DOCS with return" would be the escape sequence <ESC> 02/05 xx/yy. They assign the language codes sequentially, so it would be some byte with a value greater than 04/07. There is no other pattern to assignments so we cannot predict the value.

The private coding system could theoretically conflict with other private coding systems in use by device vendors so we should verify that 02/03 04/00 will be acceptable. I think it is unlikely that any DICOM system is also using a private ISO encoding system within their objects.

Then DICOM needs to add some rules to deal with issues not covered by the ISO IR 196 for UTF-8 and for the private encoding. The key ones will be:

- a) For UTF-8, the implementation level shall be 1. Chinese does not need implementation level 2 or 3. We must also prohibit inclusion of implementation level information within the UTF-8 string because ISO indicates that this might create a conflict with the return sequence. (This means that when we want to implement complex writing systems that require implementation level 2, e.g. Arabic, or level 3, e.g. Devengali, we will be able to use the same escape codes and use the charset attribute to convey this difference.)
- b) For both character sets we make clear that neither encoding is to be used for languages that already have DICOM defined encodings. There will be interoperability problems if people use these for the various European and Asian languages that already have defined DICOM charset encodings. I am aware of the temptation to use UTF-8 for these. Be aware that there is an equally strong temptation to use GB18030 for these purposes, because GB 18030 also provides encodings for all the characters defined in Unicode. They are merely different than the UTF-8 and ISO encodings.
- c) Define charset names, I suggest "DICOM-CN-UTF" and "DICOM-CN-GB". The official ISO name for UTF could imply a requirement for level 3 support. This would also highlight the Chinese-only aspect.
- d) Require that the default charset be the DICOM default IR 100 if either DICOM-CN-UTF or DICOM-CN-GB are present. This ensures that the "return" will always return to a known supported ISO encoding.
- e) Require that the coding system be in IR 100 before switching into UTF-8 or GB 18030. Again, this ensures that the "return" always returns to a known supported ISO encoding.

These, in combination with the existing generic DICOM rules, should handle the embedding of the desired Chinese encodings without inducing errors in old DICOM implementations or new implementations that lack Chinese support. The last two requirements are there so that DICOM implementations that have only the minimum processing needed to handle the current DICOM encodings will not become confused. Those implementations might not recognize these new ISO 2022 escape sequences, but they should recover properly because they must support the DICOM default charset and the rules ensure that string elements always start and stop with it.

R Horn

寄件者: robert.horn.b@us.agfa.com

寄件日期: 2002 年 2 月 27 日星期三 上午 1:13

收件者: eric.pan.ep@us.agfa.com; jos.debaerdemaeker.jd@belgium.agfa.com;

Hewett Andrew; idt dclunie; Lumin Zou; Xiangyun Wang; Herman Oosterwijk;

Dave Harvey; Howard Clark; bmes@indy.be.cycu.edu.tw;

jls@indy.be.cycu.edu.tw; Mike Henderson; Phili Lo

主旨: Re: A radical new proposal for Chinese in DICOM

----- Forwarded by Robert Horn/WIL/AGFA/US/BAYER on 02/26/2002 12:12 PM -----

Tim-P Martin

02/20/2002 04:48 PM

To: Robert Horn/WIL/AGFA/US/BAYER

cc:

Subject: Re: A radical new proposal for Chinese in DICOM

Rob,

I am finally getting around to reviewing your proposal for Chinese characters. I received your WG6 mailing via Jos. (Since I will be working on DICOM tools, can you add me to your WG6 mail list?)

Overall I like the proposal. It seems to meet the requirements well, it is only a small change from the current DICOM methods and should be straight forward to implement.

I prefer the use of the private encoding for GB18030, instead of using IR 58. If you use IR 58 you are mixing standard ISO with private use and that could get confusing. By going to the private encoding, DICOM can define all of the rules.

I couldn't find how you came up with the escape sequence <ESC> 02/03 04/00.

>From my reading of ISO 2022, this looks like an undefined single

>control function. I thought the sequence should be <ESC> 02/05 03/00. This is a DOCS with Return, Reserved for Private Use.

I don't see any need to restrict the use of other character sets if either of these 2 sets are present. They present very little additional requirements on maintaining the state of the invoked character sets. Currently the state of the GL (=G0) and GR (=G1) must be maintained. These character sets will require that an additional "not in ISO" state will need to be maintained. The only benefit of restricting the character set use is that the ISO states would not need to be remembered when going into the non-ISO mode. Remembering the state is not a big deal. I think trying to keep track of the extra restrictions that would need to be put in the standard would be harder. For an implementation that doesn't know anything about these new escape sequences, using a character set other than IR 6, might help them re-synchronize because they will see escape sequences that they know and act accordingly. (By the way, the DICOM default character set is IR 6, not IR 100).

I do think that there should be a requirement for an explicit return to ISO before the end of the line, page, end of value, or end of component or component group. There may be some argument about this since both UTF-8 and GB18030 encode all of the required delimiters in the exact same way as IR 6. The argument will be that the return to ISO could be implicit, as is done now with Korean. I think Korean is different because for Korean, IR 6 will normally be invoked in GL and Korean in GR, so IR 6 is actually active. This requirement will allow implementations to not look for any delimiters between the escape out and escape back. This is particularly important for GB, where the byte values of the "\", "^" and "=" characters can be used as part of a multi-byte character. For an implicit return, all implementations would need to understand how to distinguish single-byte from multi-byte characters.

As a side note on the above point, there is an inaccuracy in the current DICOM standard. PS 3.5-2001, section 6.1.2.5.3 says in the note at the end of the section:

These requirements allow an application to skip lines, values, or components in a textual data element and start the new line with a defined character set without the need to track the character set changes in the text skipped.

The inaccuracy is that values and components cannot be skipped without tracking the character set changes. Since values and components are delimited with printable characters that can have the same value as a byte in a multi-byte character, the only way to know if the byte represents a delimiter is to track the character set changes.

- Tim

----- Forwarded by Tim-P Martin/GLA/AGFA/US/BAYER on 02/20/2002 11:58 AM -----

Jos De\_Baerdemaeker  
02/19/2002 08:08 AM

To: Robert Horn/WIL/AGFA/US/BAYER@AGFA  
cc: Tim-P Martin/GLA/AGFA/US/BAYER@AGFA

Subject: Re: A radical new proposal for Chinese in DICOM (Document link: Tim-P Martin)

Rob,

The Chinese implementation is using the following Defined Terms and escape codes:

[Chinese Guo Biao 2312-80]  
Defined term = ISO 2022 GB2312  
EscapeSequence=1b 24 29 41

[Chinese Big5 ]  
Defined term = ISO 2022 B5  
EscapeSequence=1b 24 29 d7

[Chinese Guo Biao 2312-80, different DT but same escape sequence] Defined term = ISO 2022 GBK  
EscapeSequence=1b 24 29 41

Are you saying I should change them to what?

Jos