

《九家舊晉書輯本》舊版造字轉碼說明

中研院資訊所文獻處理實驗室 2007/11/1 陳建安 製作

一、《九家舊晉書輯本》一書(九家舊晉書輯本.xml)使用舊版造字 5 個，字頻 26 次，詳如附件一，但由於全使用重碼部件，故不需轉碼。

二、附件一的造字分析表說明如下：

甲、編號：Big5 造字空間為 6217 個，編號由 1 到 6217。

乙、造字：舊版造字

丙、字頻(txt)：造字在「.txt」文件的出現次數

丁、字頻(xml)：造字在「.xml」文件的出現次數

戊、Big5：造字的 Big5 碼

己、Unicode：造字所對應的 Unicode 碼

庚、WinXP：造字在 Windows XP 的對應字形

辛、備註凡例：

1、重碼部件，不轉碼：同時使用舊漢籍造字及構字式，導致一個字碼代表兩種字形的重碼問題，必須一一比對原書字形，以判斷應取何者之字形，取構字式部件者不轉碼。如編號 4262，舊漢籍造字字形為「麪」，構字式部件為「𠂔」，比對原書發現書中皆取構字式部件字形「𠂔」，故不需轉碼。

三、《九家舊晉書輯本》中未確定字形「❶」共六處，詳細處理內容可參考附件二的未確定字形手動取代表。附件二的手動取代表說明如下：

甲、檔案原文摘錄：摘錄檔案(九家舊晉書輯本.xml)中包含未確定字形「❶」的文句段落，其中紅字底線的部分，為需要手動取代的字形。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有原文。

乙、手動取代結果：變更過後的檔案內容，其中藍字底線的部分，為修改後的結果。文句後方標示（全部取代）者，表示執行全部取代的動作，不一一列舉所有修改後結果。

丙、備註：註明相同文句在檔案中出現的次數，未註明者表示只出現一次。

四、對於在缺字之外發現的錯字，其詳細內容可參考附件三的校出錯

字表。附件三校出錯字表的說明同附件二甲、乙兩項。

附件一、《九家舊晉書輯本》造字分析表

編號	造字	頻次 (txt)	頻次 (xml)	Big5	Unicode	WinXP	備註
4262	麴	1	1	8456	9EAA	麴	與昏重碼，重碼部件，不轉碼。
5302	昏	3	3	8ADA	662C	昏	與羸重碼，重碼部件，不轉碼。
5338	暎	7	16	8AFE	668E	暎	與**重碼，重碼部件，不轉碼。
5355	𡗗	4	2	8B50	20AE4	𡗗	與广重碼，重碼部件，不轉碼。
5583	殮	1	4	8CB9	98F1	殮	與殯重碼，重碼部件，不轉碼。

附件二、《九家舊晉書輯本》未確定字形手動取代表

檔案原文摘錄	手動取代結果	備註
雞之變	雞之變	
訟屠沽不絕	訟屠沽不絕	
血𡗗𡗗莫有心	血𡗗𡗗莫有關心	
與河水	與河水	共 2 處
不問格之罪	不問格之罪	
頗復與人邪	頗復與人邪	

附件三、《九家舊晉書輯本》校出錯字表

檔案原文摘錄	手動取代結果
軍至禾陵	軍至秣陵
挺來往	挺來往