

中文界面諮詢委員會
中文資訊科技工作小組
工作進展報告
(2013年12月4日)

1. 主要工作

自中文界面諮詢委員會（中諮會）第 20 次會議後，工作小組曾召開一次會議。該次會議為工作小組第 29 次會議，於 2013 年 10 月 25 日舉行，主要議題如下：

- 匯報表意文字小組(IRG)第 40 次會議的討論事項
- 匯報香港中文異體字字形規範的工作進展
- 為不獲納入《香港增補字符集》的字符提供檢索工具

2. IRG 第 40 次會議討論事項

秘書處匯報香港特別行政區政府跟進 IRG 編輯工作的進展情況，並簡介 IRG 第 40 次會議的決議要點如下：

- IRG 通過把中日韓表意文字區擴展區 F (CJK F) 工作集分拆成 F1 和 F2 兩部分，以加快工作進度，並訂出審議 CJK_F1 的工作時間表。
- IRG 商議精簡《原則和程序》(Principles and Procedures) 的某些項目，包括確認未來擴展版本不再採用“原字集分別編碼原則(Source Code Separation Rule)”，以及每個工作集只限容納 4,000 個字符。
- Unicode Consortium 要求把 19 個表意文字列為“迫切需要編碼字符”(Urgently Needed Characters)處理。

3. 香港中文異體字字形規範

背景

在中諮會第20次會議上，陸勤教授向與會者介紹表意文字異體字數據庫(Ideographic Variation Database)(IVD)及其用途。IVD為需要處理異體字的電腦系統，提供一套符合統一碼(Unicode)標準而又可互用(interoperable)的方案。

創新及科技基金資助項目“香港中文異體字的編碼及字形規範”已在2013年6月開展，為期18個月。

推展上述項目的原因，是考慮到在電腦應用上，同一文檔可能有多個異體字並存；這些異體字必須在字符層面上分開編碼和定義。該項目會利用由Unicode Consortium開發的嶄新技術——表意文字異體字序列(Ideographic Variation Sequence)(IVS)，為異體字編碼，以補充香港字形規範，使其更為可用。在該項目向創新科技基金申請前，得到工作小組的支持。

工作建議

工作小組會就項目的進行情況做出討論，並認可由中文字工作小組審議、制訂有關規範。與會者有提出可否請更多的香港廠家參與制訂該規範的工作，並有公平參與的機會。陸勤教授認同此建議，並承諾會將項目中所有審議和工作字形的修改意見全部放在網上公開。

4. 為不獲納入《香港增補字符集》的字符提供檢索工具

背景

根據《香港增補字符集》字符增收原則，由2008年3月31日起，凡在ISO/IEC 10646國際編碼標準字集及《香港增補字符集》內已有可視為認同之字符者，概不增收。由該日起至2013年9月30日止，工作小組曾處理超過640宗字符增收申請，除少數字

符(7 個)缺乏充分佐證資料外，其他字符均已收錄於 ISO/IEC 10646 國際編碼標準字集，因此無須增收。

工作小組處理的增收申請中，大部分字符並未收錄於國際表意文字子集 (International Ideographs Core)(IICore)或《香港增補字符集》內，因此一般中文輸入法未必適用，而電腦作業系統提供的字體亦可能有別於使用者要求的字形。使用者無法輸入所需字符，便以為字符尚未編碼，因此提交增收申請。這種情況不但對電腦使用者造成不便，亦未能有效運用審核字符的資源。

工作建議

工作小組建議定期整理不獲納入《香港增補字符集》的字符，並提供網上檢索工具，以便市民和政府部門在提交字符增收申請前查核。除了現有的倉頡碼和部首外，還可考慮加入筆劃總數、粵語拼音和漢語拼音等資料，供有需要使用這些字符的人士、中文輸入法或字體開發者參考。這種安排既可減省申請人整理和提交申請的時間，亦能更有效運用審核字符的資源。工作小組認為中諮會秘書處為最合適的發放部門。