

中文界面諮詢委員會

ISO 10646 國際編碼標準與 GB 18030 編碼字符集

目的

1. 本文件旨在向各委員簡述 ISO 10646 國際編碼標準與 GB 18030 編碼字符集的關係。

背景

2. 目前，不同地區會用不同的編碼標準來處理電子中文資訊，例如中國內地採用國標碼，香港及台灣則主要使用大五碼。不同的編碼標準有機會導致在使用中文進行電子通訊及資料交換時，產生亂碼和資料失誤的問題。
3. 為提供一個共通的技术基礎以處理不同語文的電子信息，國際標準化組織 (International Organization for Standardization, 簡稱為 ISO) 發展了一套名為 ISO 10646 的國際編碼標準。該標準為世界各種主要語文的字符包括繁體及簡體的中文字，編訂統一的電腦內碼。
4. 香港特區政府在「數碼 21」資訊科技策略下，為香港特別行政區制訂一套共通的電子中文界面，以方便市民能更有效地使用中文進行電子通訊。共通電子中文界面的基礎包括採用 ISO 10646 國際編碼標準及制訂《香港增補字符集》。

ISO 10646 國際編碼標準

5. 國際標準化組織於一九九三年發表 ISO 10646 國際編碼標準的首個版本，全名是「ISO/IEC 10646-1:1993」。ISO 10646 國際編碼標準的最新版本是在二零零零年十月發表的「ISO/IEC 10646-1:2000」。這個版本收納了上一個版本內的 20,902 個表意文字，及新增收的 6,582 個表意文字。表意文字是指一些字形與字義有關的文字，包括中、日、韓及亞洲各地使用的漢字。
6. ISO 10646 國際編碼標準的下一個版本估計會在二零零一年年底前發表。這個新的版本所增收的表意文字將會超過 40,000 個，其中包括所

有收錄在《康熙字典》、《漢語大字典》和《漢語大詞典》內的漢字。

7. 現時市場上已有多種軟件透過採用統一碼(Unicode)支援 ISO 10646 國際編碼標準，包括操作系統、辦公室自動化軟件、網頁瀏覽器、中文字型產品、手寫輸入裝置、數據庫軟件等。

GB 18030 編碼字符集

8. 中國內地採用國家強制性的國標碼作為漢字信息交換的編碼標準。最新版本是中國國家質量技術監督局於二零零零年三月公布的 GB 18030 - 2000 《資訊技術 資訊交換用漢字編碼字元集 基本集的擴充》。
9. GB 18030 編碼字符集是 GB 2312-80¹的擴展版，總編碼空間超過 150 萬個碼位，共收錄了 27,000 多個表意文字。GB 18030 編碼字符集所收納的字符與 ISO/IEC 10646-1:2000 版本的字符是相互對應的。就兩個編碼標準的比較，請參閱附件。
10. 現時已有產品可同時兼容 ISO 10646 國際編碼標準與 GB 18030 編碼字符集。這些產品包括操作系統、辦公室自動化軟件及數據庫軟件等。隨著電腦產品製造商開發更多可同時兼容這兩個編碼標準的產品，當用戶逐漸同時採用這些產品時，可更方便及更準確地透過這兩個編碼標準進行電子通訊及資料交換。
11. 我們會密切留意 GB 18030 編碼字符集在中國內地的發展，及電腦產品市場上採納 GB 18030 編碼字符集的情況。

資訊科技署

二零零一年十一月

¹ GB2312-80 全稱《信息交換用漢字編碼字符集 基本集》，由中華人民共和國國家標準總局發布，於 1981 年 5 月 1 日起實施。GB2312-80 包含了 6,763 個漢字和 682 個符號。

ISO 10646 國際編碼標準與 GB 18030 編碼字符集的比較

	ISO 10646 國際編碼標準	國標碼 GB 18030
支援字符	涵蓋了各種主要語文的字符,包括繁體及簡體的中文字	所收納的字符緊貼 ISO 10646 國際編碼標準,但內碼編排與 ISO 10646 國際編碼標準不同
軟件支援	操作系統、辦公室自動化軟件、網頁瀏覽器、中文字型產品、手寫輸入裝置、數據庫軟件等	操作系統、辦公室自動化軟件、網頁瀏覽器、中文字型產品、手寫輸入裝置、數據庫軟件等
通用地區	世界各地	中國
制訂機構	國際標準化組織	中國國家質量技術監督局
最新版本	ISO/IEC 10646-1:2000	GB 18030-2000
公布日期	二零零零年十月	二零零零年三月
字符數目	27,484 個表意文字	約 27,000 個表意文字