Joomla! 1.0.x 與 1.5.x 均有內建功能讓網站訪客產生所閱讀文章的 PDF 檔案。這項功能對於西方語言沒有問題,但對 亞洲語言像是中文及日文卻無法產生令人滿意的 PDF 檔案。對於 Joomla! 1.0.x,已經存在了修補方式來讓它對亞洲語 言作業正常。例如 台灣 Joomla! 釋出本地化的正體中文 Joomla! 1.0.x,裏面包含了可以正常作業的 PDF 產生子系統。

因為 Joomla! 1.5.x 使用 UTF-8 編碼並允許對不同語言指定各自的字體來產生 PDF,它基本上在 PDF 功能支援廣泛的 語言。只要字體資料、腳本程式與字體本身安裝無誤, Joomla! 1.5.x 的確可以正確地產生亞洲語言 PDF 檔案。然而, 它對於亞洲語言的支援仍然有些微問題。

目前 Joomla! (1.5.7) 使用 TCPDF v2.5.000。這個版本的 TCPDF 永遠在產生 PDF 檔案時將完整的字體嵌入,除非該字 體是一個核心字體。這對亞洲字體會有兩個問體,而這兩個問題都由通常很大的字體檔案造成。

第一個問題是所產生 PDF 檔案的大小。因為它包含了整個字體,PDF 的大小會至少是字體的大小。另外一個問題是 伺服器需要分配大塊的記憶區來使用,因為 TCPDF 在製作 PDF 時把整個字體載入記憶體。一個 15 MB 正體中文字 體大約可用到至少 40 MB 記憶體來處理。

在這裡我們將提供程序如何對正體中文文章產生不嵌入整個字體的 PDF,使產生的檔案大小是在數百 KB 而不是幾個 MB。經過些許修改,類似的程序應該也可用於其他亞洲語言。

下載 TCPDF

之前我們提及 Joomla! 1.5.7 提供的 TCPDF v2.5.000 並沒有選項不嵌入完整的字體檔案。幸運的是新版的 TCPDF 有 這個選項。因此我們的計畫基本上是將 Joomla! 內的 TCPDF 用最新版來取代。

請前往官方 TCPDF 網站來<u>下載最新版本</u>。將它解壓縮到你的電腦硬碟裏。你會看到一個目錄名稱為 tcpdf 被建立, 裏面含有 TCPDF 的程式碼。

建造字體腳本程式

雖然我們不要 TCPDF 把整個字體檔案放入 PDF 內,我們仍然必須讓它包含一些字體的資訊。你需要執行在 TCPDF 軟體包裏的兩件工具以產生一個含有字體重要資料的 PHP 腳本程式。

為了要使中文 PDF 能被 Adobe Acrobat 正常顯示且不用實際嵌入字體檔案,所用的字體必須已經是 Acrobat 安裝一部 份。如果你檢查 Acrobat 安裝目錄下的 Resource\CIDFont 檔案夾,你會發現幾個給亞洲語言使用的字體。對於正體 中文 Adobe Acrobat 提供了 AdobeMingStd-Light.otf,那是我們要使用的字體。(我們假設你自身使用的機器是一部視 窗電腦,因為其中一個你必須使用的指令是 DOS 的程式。)

D:\tcpdf\fonts\utils> ttf2ufm -a -F adobemingstd-light.otf

這指令會花一點時間來執行。等它完成後,你應該會在同樣的目錄下看到多出三個檔案,它們的延伸檔名會分別是.afm、.t1a 與.ufm。

接下來你必須執行一個 PHP 腳本程式 makefont.php。(你需要有 PHP 安裝在你的視窗機器上,或者你也可以將那三個 剛建造的檔案複製到你的伺服器上,然後在伺服器上執行下面的指令。)

D:\tcpdf\fonts\utils> php -q makefont.php adobemingstd-light.otf adobemingstd-light.ufm false

這會製作出一個有.z 延伸檔名的壓縮檔及一個 PHP 腳本程式叫做 adobemingstd-light.php, 這是 TCPDF 產生 PDF 時

「準啦!」1.5.x

需要用到的檔案。但在這個 PHP 腳本程式能被使用之前,還需經過一些修改。請用你慣用的文字編輯器開啟 adobemingstd-light.php,做以下的改變:

- 改變字體型式成為: \$type='cidfont0';
- 增加這行以設定預設的字型寬度: \$dw=1000;
- 移除 \$enc、\$file 和 \$ctg 變數的定義
- 於檔案的最後加入下面這段文字:

```
// Chinese Traditional
$enc='UniCNS-UTF16-H';
$cidinfo=array('Registry'=>'Adobe', 'Ordering'=>'CNS1','Supplement'=>0);
include(dirname(__FILE__).'/uni2cid_ac15.php');
```

要知道如何針對其他亞洲語言做修改,請參考<u>TCPDF 官方網站上的文件</u>。

複製檔案到 Joomla! 的安裝內

全部檔案準備好之後,我們現在可以把它們複製到伺服器上的適當位置裡。請使用 FTP 或 SFTP 將你的 tcpdf 檔案夾內的所有檔案以及 cache、config 與 images 三個目錄複製到你的 Joomla! 安裝內的 libraries/tcpdf 目錄裡面。(為了安全 你最好把原來的 libraries/tcpdf 目錄改名,再建立一個新的 libraries/tcpdf 目錄。)

然後請把之前製作的 adobemingstd-light.php 複製到你的 Joomla! 安裝下的 language/pdf_fonts。同樣地將你電腦 tcpdf/fonts 檔案夾裡的 helvetica.php 和 uni2cid_ac15.php 複製過去。

編輯語言敘述 XML 檔

產生 PDF 時使用的字體是在個別語言目錄裡的語言敘述 XML 檔內指定。對正體中文來說,那個檔案是 language/zh-TW/zh-TW.xml。我們需在 和 標籤之間加入下面這行:

adobemingstd-light

修改 Joomla! PDF 腳本程式

在 Joomla! PDF 腳本程式裡似乎有個臭蟲。它並沒有正確地為文章內容設定字體。請用你愛用的文字編輯器開啟檔案 libraries/joomla/document/pdf/pdf.php。搜尋 function render(\$cache = false, \$params = array())。將下面這行插入到 \$pdf->WriteHTML(\$this->getBuffer(), true); 之前。

\$pdf->setFont(\$font, '', 10);

額外步驟

如果你能容忍 PDF 檔案中文章內容醜陋的編排,這其實是選擇性的步驟。這裡的問題是 TCPDF 使用空白字元來排版,但中文字通常是連在一起,並沒有空白字元夾在中間。

要修正這個問題,你必須修改 TCPDF 核心程式。在你的伺服器上用你愛用的編輯器開啟檔案 libraries/tcpdf/tcpdf.php,找到 public function Write(\$h, \$txt, \$link='', \$fill=0, \$align='', \$ln=false, \$stretch=0, \$firstline=false) 函式的定義。在該函式內,搜尋以下這行:

if (preg_match("/[\s]/u", \$this->unichr(\$c))) {

以這行文字取代

if (preg_match("/[\s\p{Lo}]/u", \$this->unichr(\$c))) {

現在從你的 Joomla! 網站前台,試著產生一個中文 PDF 檔案,看看編排是否有改進。如果沒有改進,表示你的伺服器 上之 PCRE 函式庫沒有把 Unicode 性能支援編譯進去。你可以在伺服器上用這個指令來檢查:

\$ pcretest -C

你會看見類似下面的輸出

PCRE version 6.6 06-Feb-2006 Compiled with UTF-8 support Unicode properties support Newline character is LF Internal link size = 2 POSIX malloc threshold = 10 Default match limit = 10000000 Default recursion depth limit = 10000000 Match recursion uses stack

如果它顯示 Unicode 性能支援沒有啟動,這步驟裡的修改便無效。在這狀況下,你可以下載 PCRE 函式庫的程式 碼,自行編譯以啟用 Unicode 性能支援。一旦新編譯的 PCRE 函式庫安裝好了,所產生之中文 PDF 的排版應該變得 能被接受。

如果覺得懶惰.....

要是看到這些步驟讓你覺得氣餒,這裡告訴你一個捷徑。我把所有的檔案(除了額外步驟提及的檔案外,因為需要特別的 PCRE 函式庫)打包成<u>一個套件</u>,放在 Joomla! Taiwan。你可以將它下載,放進你的伺服器上。