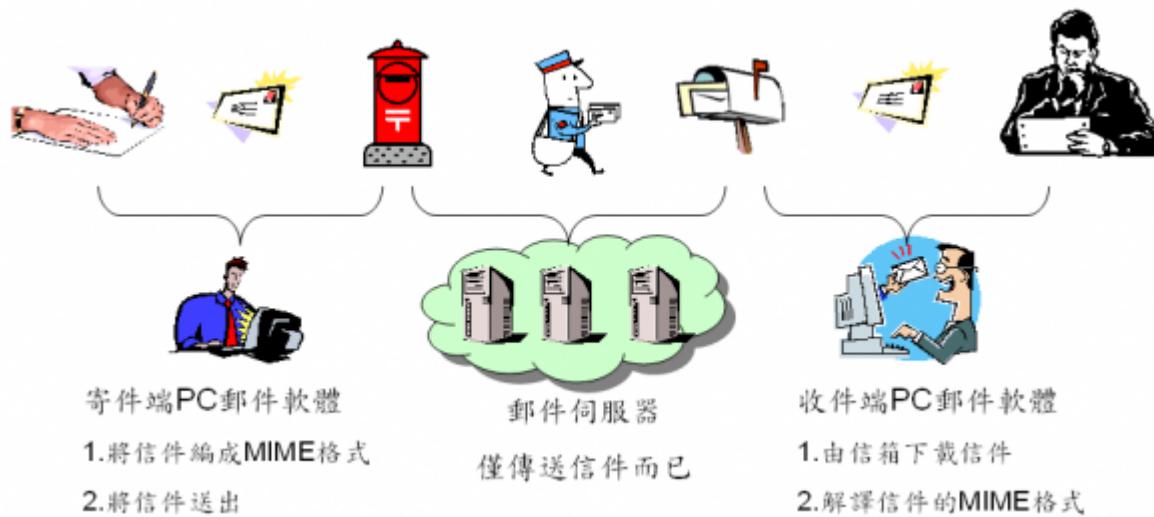


# 為何我收到的信會是亂碼？

- 要能順利閱讀信件，除了寄件者使用正確的編碼方式，收件者也要能正確解碼，如此才能避免因亂碼而無法閱讀的問題。



信件中的 **MIME** 格式與編碼有關的部份，分別是「**字元集 (charset)**」與「**編碼方式 (Content-Transfer-Encoding)**」。前者有我們常見的 **ISO-8859-1**、**Big5**、**UTF-8** 等；後者則有 **7bit**、**8bit**、**quoted-printable**、**base64**、**binary** 多種編碼方式。

多數形況下，郵件軟體都能正確編碼、解碼，但實作上仍有些差異和缺陷，如未能正確指定字元集與編碼方式、無法妥善的解碼或缺少特定字元集的轉換能力。除此之外，使用者也可能手動選擇不適當的字元集或編碼方式，造成字元轉換錯誤，以致於信件無法正確呈現。

## 案例一：收件端郵件軟體使用錯誤字元集解碼

- 文字「這是一封信」，以下用不同的編碼/解碼的方式。

錯誤：原文字以 <b>BIG5</b> 編碼，再以 <b>UTF-8</b> 解碼，可能看到情形。
b00@□
正確：原文字以 <b>BIG5</b> 編碼，再以 <b>BIG5</b> 解碼。
這是一封中文信
錯誤：原文字以 <b>UTF-8</b> 編碼，再以 <b>BIG5</b> 解碼，可能看到情形。
? 銚 撻 葉? 繼
正確：原文字以 <b>UTF-8</b> 編碼，再以 <b>UTF-8</b> 解碼。
這是一封中文信

## 案例二：寄件端郵件軟體提供錯誤字元集資訊

- 信件顯示編碼資訊如下，但卻無法正確解碼。  
 Content-Type: text/plain; charset="GB2312"  
 Content-Transfer-Encoding: 8bit

**使用 GB2312 解碼**

From: 峴弁 <[REDACTED]@my.nthu.edu.tw>  
 Date: Wed, 25 Mar 2009 09:38:53 +0800  
 Subject: ? 狗EMAIL跪 兖乙 韓 峴弁

-----

蓄 淬 98爛翩艙 脹? 妝 麻詔礀呀?  
 珑 1 掛莓9 8爛3棆12 跋菴0980500154  
 煙燐諒?  
 翩艙 脹? 汲〔

**使用 GB18030解碼**

From: [REDACTED] <[REDACTED]@my.nthu.edu.tw>  
 Date: Wed, 25 Mar 2009 09:38:53 +0800  
 Subject: 請協助EMAIL各單位，並掛於[REDACTED]

-----

國立[REDACTED]年[REDACTED]報名注意事項  
 一、本校[REDACTED]第098050015[REDACTED]  
 [REDACTED]

**案例補充說明**

下面範例包含了「正體字」、「簡體字」，並使用了 [WGB2312](#) 字元集，但信件內的部份字元卻不在 GB2312 字元集範圍內，如：

**华 堑 編**

- 华 - 在 UTF-8\GB2312 範圍，不在 Big5 範圍內
- 堑 - 在 UTF-8 範圍，不在 Big5\GB2312 範圍內
- 編 - 在 UTF-8\Big5 範圍，不在 GB2312 範圍內 (在 GBK 範圍內)

上述範例中編輯者指定 GB2312 進行編碼，將造成閱讀者解碼時的錯誤。如需要同時使用「正體字」與「簡體字」，編輯時可以改用範圍較大的 [UTF-8](#)

**常見問題發生原因**

目前多數輸入法 ([微軟新注音輸入法](#)、[新酷音輸入法](#)) 支援輸入「正體字」與「簡體字」，使用者在選字時往往因為字體較小或字型差異，不自覺選了相似的字，如「編○○○編」，其中前者為「正體字」；後者則為「簡體字」。此時同時混合使用「正體字」與「簡體字」，但卻使用了字元範圍較小的 GB2312 字元集，將會造成部份「正體字」編碼與解碼的錯誤。(另一例和語系無關但容易選錯字，則是國字的「一」與注音的「ㄧ」)

— —

國字      注音

編      編

正體字      簡體字

### 建議改善方式：

- 如果是選錯字的問題，建議調整字體大小或改用容易辨讀的字型，以避免無法正確辨別差異，選擇了相似的字。
- 如果僅需要輸入 Big5 字元集，而無使用「簡體字」的需求，此時可以限定輸入法只提供 Big5 候選字，避免誤輸入。
- 如有必要同時使用「正體字」、「簡體字」或是其他字元(如日文、韓文等)，可依據字元內容使用範圍較大的 UTF-8 或其他字元集。

### 相關補充資料

- 使用 `wiconv` 或 `piconv` 進行字元編碼的轉換，如轉換 Big5 或 GBK 至 UTF-8

```
# iconv -f big5/gbk -t utf-8 輸入檔案  
# piconv -f big5/gbk -t utf-8 輸入檔案
```

- 列出 `iconv` 或 `piconv` 目前支援的字元集

```
# iconv -l  
# piconv -l
```

### 相關網站

- [CNS11643 中文全字庫](#)
- [中文碼介紹](#)
- [Outlook 2007 主旨會變亂碼](#)
- [Configuring Encoding Options in Outlook 2003](#)

From:  
<https://net.nthu.edu.tw/netsys/> - 網路系統組

Permanent link:  
[https://net.nthu.edu.tw/netsys/faq:mail\\_encode\\_charset](https://net.nthu.edu.tw/netsys/faq:mail_encode_charset)

Last update: 2009/04/06 09:36