



拡張子「.html」と「.htm」どっちがいい？

text 牧野武文

さまざまなサイトを閲覧している時、ふとWebブラウザのアドレスバーに目を移すと、HTMLページであっても、ファイルの拡張子が「.html」と「.htm」という2種類が混在していることに気づく。傾向としては「.html」の方が圧倒的に多く、混在していても問題もなく閲覧できてしまう。とはいえ、どうして混在しているのか、ハッキリさせておきたいところだ。

なぜ「html」が一般的なのに「htm」という拡張子があるのかといえば、単にMS-DOS時代からの慣例でしかない。今やWindowsもMacもUNIXも4文字以上の拡張子に対応しているが、MS-DOSの時代は拡張子が3文字限定だったため、

MS-DOSのテキストエディタなどでHTML文書を作成した場合などは、「htm」までしか記述できなかったのである。

それでも問題なくネットが閲覧できている理由は、Webサーバ側の「MIME設定」にある。この設定は、WebサーバからWebブラウザへとファイルを送信する際に、「このファイルは〇〇ですよ」という情報を付け加えて送信するという設定のことだ。極端な話をすれば、Webサーバがファイルの種類を正しく認識してさえすれば、拡張子が一般的なものでなくてもWebブラウザは正しく認識してくれるのである。

しかし、Webサーバも万能ではない。ファイル内

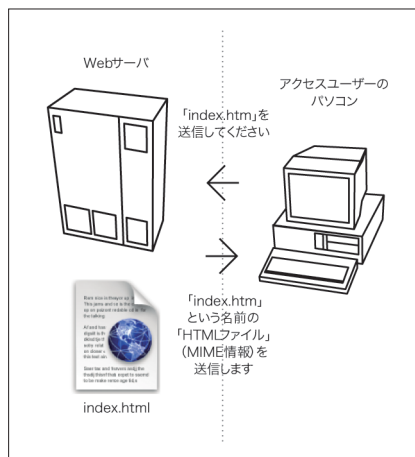
のデータ構造などによっては、ファイル種別の認識を誤ることもある。そんな時でも、ファイルの拡張子が正しくつけられていれば、Webブラウザはファイル種別に合った正しい動作を行ってくれるのだ。

インターネット上には、HTMLやCSS、画像にswf、PDFなどさまざまな種類のファイルが公開されている。Webブラウザがそれぞれのファイル種別に適した動作を行ったり、プラグインを呼び出せたりしているのも、すべてはこの拡張子とWebサーバによる二重チェックのおかげだったのである。

さて、肝心の「.html」と「.htm」という話だが、どちらもHTMLとして認識するのでご安心を。

■Webブラウザで表示できる代表的なファイルの拡張子

拡張子	
.avi	古くからあるWindows用の動画ファイル形式
.css	CSSファイル
.cgi	Perlなどの実行ファイル
.chm	HTMLヘルプファイル
.class	Javaの実行ファイル
.exe	Windowsなどの実行ファイル(プログラム本体)
.flv	Flash Videoファイル
.gif	GIF形式の画像ファイル
.htm / .html	HTMLまたはXHTML
.ico	Windows用のアイコンファイル。Favicon用にも使われる
.ini	PHPなどが使う初期設定ファイル
.jpg / .jpeg	JPEG形式の画像ファイル
.js	JavaScriptファイル
.log	Webサーバなどが記録するログファイル
.mp3	MP3形式の音声ファイル
.mp4	MP4形式の動画(または音声)ファイル
.mov	QuickTime形式の動画ファイル
.mpg / .mpeg	MPEG形式の動画ファイル
.pl	Perlのソースファイル
.png	PNG形式の画像ファイル
.ram	Real Player用ファイル
.swf	Flash Player用ファイル
.txt	テキストファイル
.wav	Windowsの標準音声ファイル。音楽CDのデータ記録にも使われている



WebブラウザからWebサーバへとファイルの送信要求を行った際、Webサーバはファイルの送信と同時に、そのファイルがどんな種類(HTMLや画像など)なのかという情報(MIME)を送信している

なぜMS-DOSの拡張子は3文字になったのか？

MS-DOS系列の拡張子は、ファイル名8文字+拡張子3文字といった制約があった。ここで拡張子の文字数を3文字に決めた理由は、欧米の文化が「省略は3文字」という考え方があったためだと言われている。Jan、Feb、Marといった月名や、Sun、Mon、Tueなどの曜日を省略する時も3文字。IBM、CIAなど、社名や機関名も3文字に省略することが多い。

そういった理由に加えて、当時のパソコンはメモリ容量などのスペックが低く、1文字でもデータを節約することが求められていた。かと言って、拡張子を2文字にしてしまうと将来的に拡張子が不足してしまう可能性もあるため、結果的に3文字に決められたのだと思う。