

## Re: Large text file to a CSV file...

**Source:**

<http://www.tech-archive.net/Archive/Access/microsoft.public.access.externaldata/2004-05/0708.html>

---

**From:** Ted Allen (*anonymous\_at\_discussions.microsoft.com*)

**Date:** 05/26/04

Date: Wed, 26 May 2004 09:13:17 -0700

Hi Louie,

Well, you may be able to get by without having to write code to process the file if you can experiment with using search and replace in a text editor that can handle the file. As John mentioned, just replacing the []'s around the dates with ""'s may be enough. The key is that any fields that include spaces in the field values need to be qualified with ""'s around them so that Access will not interpret the spaces as field delimiters. When you are importing the file, it is also important to specify that the ""'s are the text qualifiers.

One other thing you could try is to import to Excel first and save as an Excel file. I have noticed that Excel is sometimes better (or at least different) in the way it handles csv files. But, this may not be an option for you if your file has more records than Excel can handle.

Other than that I think your only options that I know of would be to try to find some other program that can import the file, or write some VBA code to make whatever changes are necessary in the text file prior to importing. This really isn't much work for those familiar with VBA, but there would obviously be a learning curve if you aren't familiar. There are many in this user group though that are happy to help with code if you can define what you want it to do.

Hope that helps.

-Ted Allen

>-----Original Message-----

>There's the rub... I would have to write a preprocessor

>routine because a 3,094,112KB file would be difficult

(to

>say the least) to edit. WSP does imbed the - when there

>is no value. I may be finding that this is more work  
than  
>I need to go through.  
>  
>  
>>-----Original Message-----  
>>Hi Louie,  
>>  
>>Access's text import routine has problems with files  
like  
>this, because  
>>the space is not just being used as a delimiter, it  
also  
>appears in the  
>>data without being "qualified" by enclosing the entire  
>field in quote  
>>marks.  
>>  
>>It looks as if the problematic field is the datestamp.  
>Using "-" for an  
>>unknown username shouldn't be a problem because it  
>doesn't introduce any  
>>extra spaces, though if WebSite Pro uses " - " it  
would.  
>>  
>>One approach would be to pre-process the file by simply  
>replacing the [  
>>and ] around the datestamp with " and ". That should  
>convert it into a  
>>regular space-delimited file, with fields that may  
>contain a space  
>>qualified with " ". It would then be possible to  
import  
>it to a  
>>temporary Access table, where you could run a query to  
>parse the  
>>datestamp field and convert it to an Access date time  
>value as you  
>>append it to your "permanent" table.  
>>  
>>The pre-processing could be done in a text editor, or  
if  
>it's a regular  
>>task you could work up a VBA routine using Line Input #  
>to read the file  
>>line by line, the Replace() function to change the  
line,  
>and Print # to  
>>write it to a new file. Then import the new file.  
>>  
>>Another approach would be to use a regular expression

>engine to parse  
>>the input line into fields and append these directly to  
>your table.  
>>  
>>  
>>On Fri, 21 May 2004 09:50:16 -0700, "Louie Warren"  
><lwarren@ticom.com>  
>>wrote:  
>>  
>>>I have a WebsitePro Access Log Combined (W3C/NCSA)  
>Format  
>>>file; and yes, I know there is an ODBC routine out  
there  
>>>to do this, but my company doesn't want to spend the  
>money  
>>>to get it because we are abandoning WebSite Pro in the  
>>>next few months. The data is on a classified system  
so  
>I  
>>>will attempt to recreate the problem with generic  
data.  
>I  
>>>already see problems with the blank delimiters...  
>>>  
>>>An example from the book (string these together with a  
>>>space between fields.  
>>>  
>>>172.18.96.157 (IP Address)  
>>>  
>>>gryphonaire.com (Hostname of server  
>>>receiving request)  
>>>  
>>>Admin (Authentication user  
>name)  
>>>  
>>>[02/Jan/2001:09:44:56 - 0800] (Date timestamp UTC  
>offset)  
>>>  
>>>"GET /members.html HTTP/1.1" (Complete HTTP request)  
>>>  
>>>200 (HTTP Response Code)  
>>>  
>>>178 (Number of bytes  
>transferred)  
>>>  
>>>"<http://gryphonaire.com/index.html>" (Complete  
referring  
>>>URL)  
>>>  
>>>"Mozilla/4.0 (compatible; MSIE 4.01; Windows NT)" (ID  
>of

>>>browser)  
>>>  
>>>"Member=WebSiteRules" (Cookie)  
>>>  
>>>The problem in the Access import is if the  
>Authentication  
>>>username isn't available. WSP puts in a – This is  
>>>common for most of the requests. When an actual ID is  
>>>present, it doesn't fit the pattern with the others.  
>The  
>>>data is consistantly like the above example (" [, and  
>>>all).  
>>>  
>>>Hopefully that explains it.  
>>>  
>>>L  
>>  
>>--  
>>John Nurick [Microsoft Access MVP]  
>>  
>>Please respond in the newgroup and not by email.  
>>.br/>>>  
>.br/>>