Advanced Batching Tips using Microsoft Excel

*This document assumes a basic understanding of batch functions.

Full Path Batching

Full Path Batching simply means that instead of using the document's simple file name, "document.doc", we can use the entire path that must be traveled to locate the document, "F:\ds_06-14526\To_Merrill\Folder1\document.doc". This allows a single batch to be created for documents that are located within multiple folders.

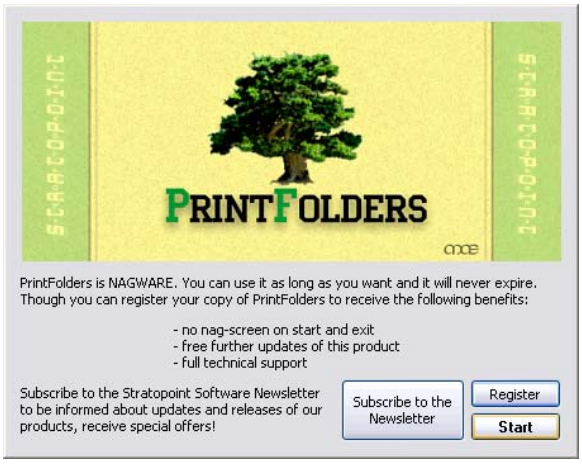
Use full path batching to upload documents from within multiple folders on the FTP site without having to extract documents or merge them into a single folder.

	A	B	C	D
1	**Command**	**Index**	**Full Path File Name**	**Title**
2	e	1.103	F:\06-12563\pub\To_Merrill\Folder1\1.103.doc	Title
3	e	1.7.1	F:\06-12563\pub\To_Merrill\Folder1\1.7.1.pdf	Title1
4	e	5.1	F:\06-12563\pub\To_Merrill\Folder2\5.1.doc	Title2
5	e	2.3.8	F:\06-12563\pub\To_Merrill\Folder2\2.3.8.ppt	Title3
6	e	2.7	F:\06-12563\pub\To_Merrill\Folder2\2.7.pdf	Title4
7	e	3.2.2	F:\06-12563\pub\To_Merrill\Folder3\3.2.2.pdf	Title5
8	e	6.2	F:\06-12563\pub\To_Merrill\Folder3\6.2.doc	Title6
9				

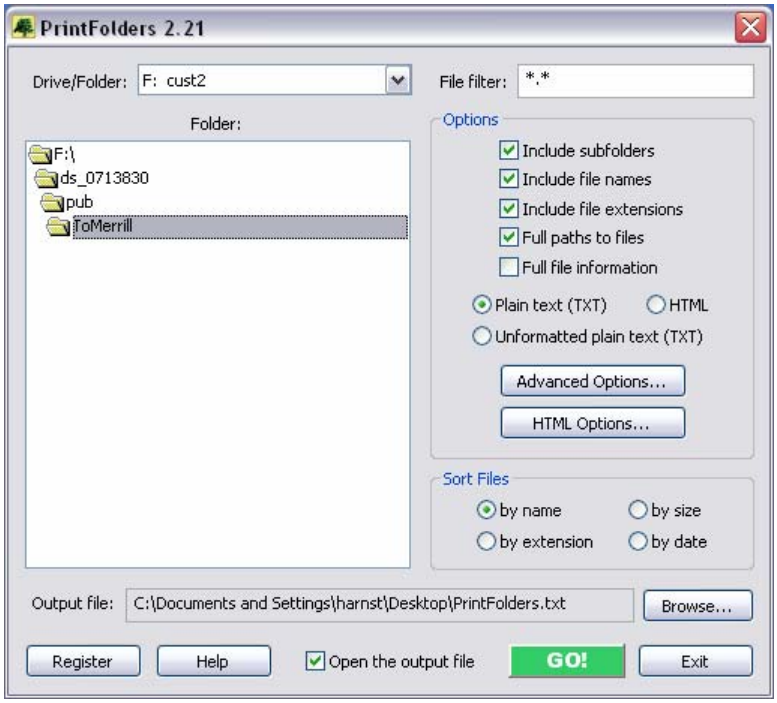
Using Print Folders to capture full path file names

If you are currently using PrintFolders 1.4, it is highly suggested that you upgrade to PrintFolders 2.21. You will find this program located on the T: drive: T:\stpfds02 DataSite CS\PrintFolders.exe. Drag it onto your desktop.

1. Launch PrintFolders.
2. Click **Start**. PrintFolders 2.21 is nag-ware, so it will bring this screen up every time you launch it.



3. Select **Full paths to files** to gather full path file names.
4. Select **Include Subfolders** to pull file names for all sub items.



Using Quotes to "fix" invalid file extension errors

There are several scenarios in which DataSite may incorrectly interpret the full path. When this occurs, DataSite will not be able to locate the specified document and you may receive an error such as "Invalid file extension". Adding quotes to the file path will often correct this error as it prevents DataSite from interpreting the data and simply tells it to locate the exact string of characters.

Insert columns before and after the file name cell. Fill the columns with quotation marks. Copy and paste the batch directly into DataSite.

	A	B	C	D	E	F
1	**Command**	**Index**		**Full Path File Name**		**Title**
2	e	1.103	"	F:\06-12563\pub\To_Merrill\Folder1\1.103.doc	"	Title
3	e	1.7.1	"	F:\06-12563\pub\To_Merrill\Folder1\1.7.1.pdf	"	Title1
4	e	5.1	"	F:\06-12563\pub\To_Merrill\Folder2\5.1.doc	"	Title2
5	e	2.3.8	"	F:\06-12563\pub\To_Merrill\Folder2\2.3.8.ppt	"	Title3
6	e	2.7	"	F:\06-12563\pub\To_Merrill\Folder2\2.7.pdf	"	Title4
7	e	3.2.2	"	F:\06-12563\pub\To_Merrill\Folder3\3.2.2.pdf	"	Title5
8	e	6.2	"	F:\06-12563\pub\To_Merrill\Folder3\6.2.doc	"	Title6
9						

There is no need to concatenate the cells or enter the quotations into the cell with the file name.

Concatenate

The term concatenate refers to connecting or linking items together. You may use concatenation to link information in two different cells. For example, you may concatenate, or merge the information in cell A1 and cell B1 using the following formula: **=A1&B1**

	A	B	C	D
1	1.1.1	.pdf	=A1&B1	
2	1.1.2	.pdf	1.1.2.pdf	
3	1.1.3	.pdf	1.1.3.pdf	
4	1.1.4	.pdf	1.1.4.pdf	
5	1.1.5	.pdf	1.1.5.pdf	
6	1.1.6	.pdf	1.1.6.pdf	
7	1.1.7	.pdf	1.1.7.pdf	
8	1.1.8	.pdf	1.1.8.pdf	
9	1.1.9	.pdf	1.1.9.pdf	

You may also use concatenation to add information to the contents of a cell. For example, when documents are scanned by DataSite production, they are generally named using the Index numbers. The resulting file names are “IndexNumber”.pdf, 1.2.pdf or 13.2.6.pdf.

To create a batch worksheet for production to upload documents to the DataSite, there must be a column for Filename. The file names can be quickly generated using the following formula: =cell&”.pdf”

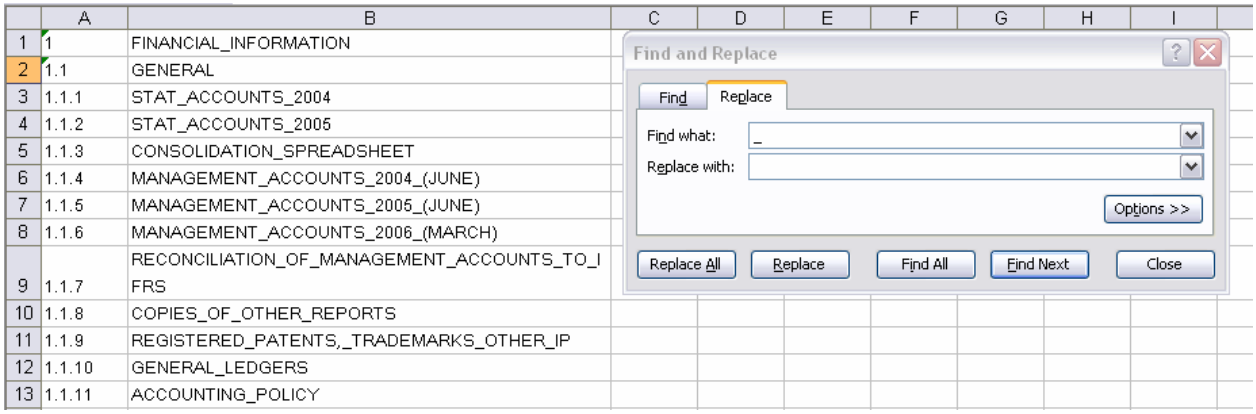
	A	B	C	D
1	Command	Index	Filename	Title
2	e_s	1.1.1	=B2&".pdf"	Title1
3	e_s	1.1.2	1.1.2.pdf	Title2
4	e_s	1.1.3	1.1.3.pdf	Title3
5	e_s	1.1.4	1.1.4.pdf	Title4
6	e_s	1.1.5	1.1.5.pdf	Title5
7	e_s	1.1.6	1.1.6.pdf	Title6
8	e_s	1.1.7	1.1.7.pdf	Title7
9	e_s	1.1.8	1.1.8.pdf	Title8

If you are passing your batch worksheet along for another shift or for production to execute, test your batch to verify that there are not any parsing errors.

Batch Title Rename

Find & Replace unwanted characters

Use the Replace feature in Excel to remove the unwanted underscores in the sample titles. Go to **Edit** | **Replace** or use **Ctrl+H** to launch the **Find and Replace** window.



	A	B
1	1	FINANCIAL_INFORMATION
2	1.1	GENERAL
3	1.1.1	STAT_ACCOUNTS_2004
4	1.1.2	STAT_ACCOUNTS_2005
5	1.1.3	CONSOLIDATION_SPREADSHEET
6	1.1.4	MANAGEMENT_ACCOUNTS_2004_(JUNE)
7	1.1.5	MANAGEMENT_ACCOUNTS_2005_(JUNE)
8	1.1.6	MANAGEMENT_ACCOUNTS_2006_(MARCH)
9	1.1.7	RECONCILIATION_OF_MANAGEMENT_ACCOUNTS_TO_IFRS
10	1.1.8	COPIES_OF_OTHER_REPORTS
11	1.1.9	REGISTERED_PATENTS,_TRADEMARKS_OTHER_IP
12	1.1.10	GENERAL_LEDGERS
13	1.1.11	ACCOUNTING_POLICY

Change Casing

Casing may also be changed within Excel. Use the PROPER formula to change to Title Case in which the first letter of each word is capitalized: **=PROPER(cell)**

	A	B	C	D	E	F
1	1	FINANCIAL INFORMATION	=PROPER(B1)			
2	1.1	GENERAL	G PROPER(text)			
3	1.1.1	STAT ACCOUNTS 2004	Stat Accounts 2004			
4	1.1.2	STAT ACCOUNTS 2005	Stat Accounts 2005			
5	1.1.3	CONSOLIDATION SPREADSHEET	Consolidation Spreadsheet			
6	1.1.4	MANAGEMENT ACCOUNTS 2004 (JUNE)	Management Accounts 2004 (June)			
7	1.1.5	MANAGEMENT ACCOUNTS 2005 (JUNE)	Management Accounts 2005 (June)			
8	1.1.6	MANAGEMENT ACCOUNTS 2006 (MARCH)	Management Accounts 2006 (March)			
9	1.1.7	RECONCILIATION OF MANAGEMENT ACCOUNTS TO IFRS	Reconciliation Of Management Accounts To Ifrs			
10	1.1.8	COPIES OF OTHER REPORTS	Copies Of Other Reports			
11	1.1.9	REGISTERED PATENTS, TRADEMARKS OTHER IP	Registered Patents, Trademarks Other Ip			
12	1.1.10	GENERAL LEDGERS	General Ledgers			
13	1.1.11	ACCOUNTING POLICY	Accounting Policy			

Be cautious of words that should be all caps. For example, acronyms such as IPO will be converted to Ipo.

Use **=LOWER(cell)** to change text to all lowercase or **=UPPER(cell)** to change text to all uppercase.

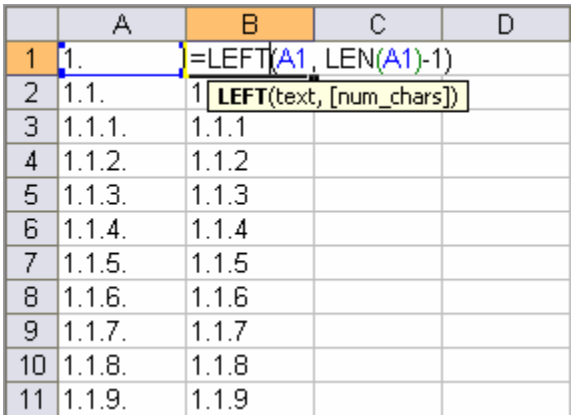
Eliminate training period in Excel

Trailing periods can be removed in a number of ways. Excel has a formula that will allow you to remove a single character from the end of a string of characters. This formula can be used to remove trailing periods from the index numbers: =LEFT(cell,Len(cell)-x)

Formula Breakdown:

- **LEFT** tells the formal to read the cell from left to right. If you wanted to remove characters from the beginning, you could use RIGHT to tell the formula to read from right to left.
- "**cell**" is the cell from which we are removing characters.
- **LEN** returns the number of characters in the string.
- "**x**" is the number of characters that need to be removed from the end of the character string.

Hence, =LEFT(A1,LEN(A1)-1) tells Excel to remove a single character from the end of the character string in cell A1.



This formula will remove the last character in the cell, regardless of what it is. If there is not a trailing period on a particular index number, it will cut the last digit of the index number.

If the above formula makes your brain hurt, there are other options for removing trailing periods. You can use the combined power of Concatenation and Find & Replace to assist with this task.

If you use Find & Replace to remove periods, it would remove all periods and not just the trailing ones. If you create a unique series of characters and then use Find & Replace, it will allow you to only remove the unwanted trailing periods.

Use your power of Concatenation to add "xx" to the existing index numbers. This will create an ".xx" combination that can be removed using Find & Replace. For any stray index numbers that did not have trailing periods, you can use the Find feature to search for any strangling occurrences of "xx".

Vlookup to match document title

Scenario: A client sends us an index for their DataSite. The client then sends us documents with file names corresponding to the index ("Standard Service Agreement.pdf"). The client requests that the documents be placed on the DataSite according to the provided index. You now have a list of document titles and we need to determine their index numbers.

Vlookup will allow us to take a list of titles and compare them to the titles in the index. For matching items, it will then return the appropriate index number. **=vlookup(cell,A:B,2,False)**

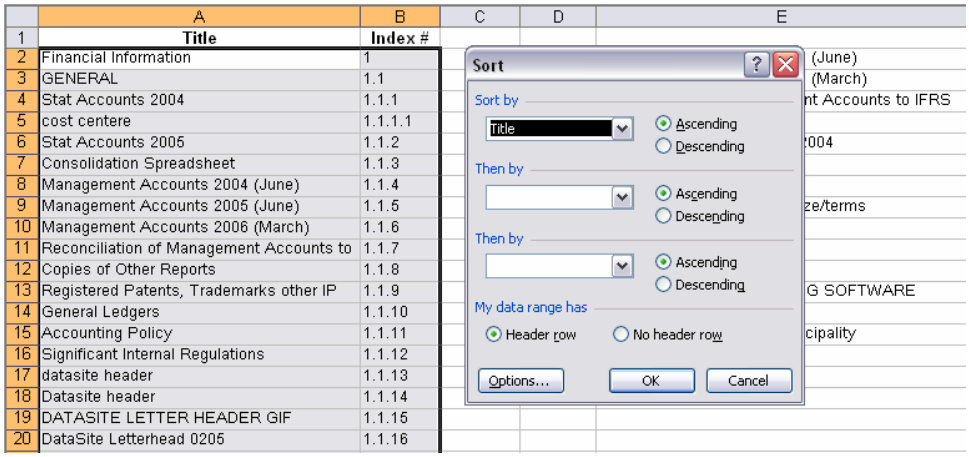
1. Gather list of document titles (use PrintFolders to pull this information from the documents)
2. Paste the document titles into the Index spreadsheet.

	A	B	C	D	E
1	**Index #**	**Title**			
2	1	Financial Information			Management Accounts 2005 (June)
3	1.1	GENERAL			Management Accounts 2006 (March)
4	1.1.1	Stat Accounts 2004			Reconciliation of Management Accounts to IFRS
5	1.1.1.1	cost centere			Top 10 clients - by revenue
6	1.1.2	Stat Accounts 2005			Top 10 clients - contracts - 2004
7	1.1.3	Consolidation Spreadsheet			Standard service contract
8	1.1.4	Management Accounts 2004 (June)			DIRECT COSTS
9	1.1.5	Management Accounts 2005 (June)			Property details - address/size/terms
10	1.1.6	Management Accounts 2006 (March)			Property Leases
11	1.1.7	Reconciliation of Management Accounts to			Insurance polices
12	1.1.8	Copies of Other Reports			Board Minutes
13	1.1.9	Registered Patents, Trademarks other IP			IT SYSTEMS - ACCOUNTING SOFTWARE
14	1.1.10	General Ledgers			Local business tax returns
15	1.1.11	Accounting Policy			Tax accounts - state & municipality
16	1.1.12	Significant Internal Regulations			
17	1.1.13	datasite header			
18	1.1.14	Datasite header			
19	1.1.15	DATASITE LETTER HEADER GIF			
20	1.1.16	DataSite Letterhead 0205			

3. Reorder the Index columns such that the Title field is first, followed by the index number.

	A	B	C	D	E
1	**Title**	**Index #**			
2	Financial Information	1			Management Accounts 2005 (June)
3	GENERAL	1.1			Management Accounts 2006 (March)
4	Stat Accounts 2004	1.1.1			Reconciliation of Management Accounts to IFRS
5	cost centere	1.1.1.1			Top 10 clients - by revenue
6	Stat Accounts 2005	1.1.2			Top 10 clients - contracts - 2004
7	Consolidation Spreadsheet	1.1.3			Standard service contract
8	Management Accounts 2004 (June)	1.1.4			DIRECT COSTS
9	Management Accounts 2005 (June)	1.1.5			Property details - address/size/terms
10	Management Accounts 2006 (March)	1.1.6			Property Leases
11	Reconciliation of Management Accounts to	1.1.7			Insurance polices
12	Copies of Other Reports	1.1.8			Board Minutes
13	Registered Patents, Trademarks other IP	1.1.9			IT SYSTEMS - ACCOUNTING SOFTWARE
14	General Ledgers	1.1.10			Local business tax returns
15	Accounting Policy	1.1.11			Tax accounts - state & municipality
16	Significant Internal Regulations	1.1.12			
17	datasite header	1.1.13			
18	Datasite header	1.1.14			
19	DATASITE LETTER HEADER GIF	1.1.15			
20	DataSite Letterhead 0205	1.1.16			

4. Sort columns A and B in ascending order. This formula is going to look for the first matching instance. If multiple documents have the same title, this will group them together in the list and report on the first occurrence.



	A	B	C	D	E
1	**Title**	**Index #**			
2	Financial Information	1			
3	GENERAL	1.1			
4	Stat Accounts 2004	1.1.1			
5	cost centere	1.1.1.1			
6	Stat Accounts 2005	1.1.2			
7	Consolidation Spreadsheet	1.1.3			
8	Management Accounts 2004 (June)	1.1.4			
9	Management Accounts 2005 (June)	1.1.5			
10	Management Accounts 2006 (March)	1.1.6			
11	Reconciliation of Management Accounts to	1.1.7			
12	Copies of Other Reports	1.1.8			
13	Registered Patents, Trademarks other IP	1.1.9			
14	General Ledgers	1.1.10			
15	Accounting Policy	1.1.11			
16	Significant Internal Regulations	1.1.12			
17	datasite header	1.1.13			
18	Datasite header	1.1.14			
19	DATASITE LETTER HEADER GIF	1.1.15			
20	DataSite Letterhead 0205	1.1.16			

5. Enter the Vloopup forumula **=vlookup(cell,A:B,2,FALSE)**. In the example worksheet, enter the Vlookup formula in cell D2.

Formula Breakdown:

- **Cell** is the cell that we are trying to match – the document title provided by the client. In this example, cell=E2.
- **A:B** is the range in which we are looking for information. We are looking in columns A through B to find our title and index number.
- **2** is the column that will produce our answer. In the example, we are looking at columns A through B. **2** is telling the formula to provide the data from column 2 as the answer.
- **FALSE** is going to report items that are exact matches. Yes, **FALSE** is going to provide exact match results. If **FALSE** does not produce any results, then there is not an exact title match. You can change the value to **TRUE** and locate approximate matches.

	A	B	C	D	E
1	**Title**	**Index #**			
2	3 sample employment contracts	1.12.14		=VLOOKUP(E2,A:B,2,FALSE)	
3	Accounting Policy	1.1.11		1	VLOOKUP(lookup_value, table_array, col_index_num, **[range_lookup]**)
4	ASSET BASE	1.9		1.1.7	Reconciliation of Management Accounts to IFRS
5	Board Minutes	1.10.2		1.2.4	Top 10 clients - by revenue
6	Breakdown of retainer fees and success	1.2.3		1.2.5	Top 10 clients - contracts - 2004
7	Budgets	1.8.1		1.2.13	Standard service contract
8	Budgets FY2006	1.8.1.3		1.3	DIRECT COSTS
9	BUSINESS PLAN	1.8		1.5.2	Property details - address/size/terms
10	CASHFLOW & FINANCING	1.7		1.5.3	Property Leases
11	Consolidation Spreadsheet	1.1.3		1.10.1	Insurance polices
12	CONTINGENCIES	1.10		1.10.2	Board Minutes
13	Copies of Other Reports	1.1.8		1.11	IT SYSTEMS - ACCOUNTING SOFTWARE
14	Copy of LON DataSite Operations Emp List	1.3.4		1.12.17	Local business tax returns
15	cost centere	1.1.1.1		1.12.20	Tax accounts - state & municipality
16	cost centere	1.3.5			
17	Creditor Analysis	1.9.3			
18	datasite header	1.1.13			
19	Datasite header	1.1.14			
20	DATASITE LETTER HEADER GIF	1.1.15			