



EDRM DATA CALCULATOR INSTRUCTIONS

ABOUT

The EDRM Data Calculator is an Excel spreadsheet file that helps you better estimate how much data you may have in matters involving eDiscovery. The estimates should help you and your organization prepare budgets, manage workflows, and measure and improve your eDiscovery processes.

The EDRM Data Calculator consists of a Core Data Calculator and a Supplemental Data Calculator. As the names suggest, you may use only the Core Data Calculator or you may choose to use the Supplemental Data Calculator as well.

The Core Data Calculator uses information you enter or select to prepare two sets of estimates. First, it calculates how much your data may increase in size because of steps taken to expand the data – steps such as unpacking compressed files. Second, the Core Data Calculator estimates how much the size of your data set will decrease as a result of processing steps such as de-Nisting, de-duplication, search terms, and CAR (computer assisted review).

The Supplemental Data Calculator uses additional information you enter to arrive at two additional sets of estimates about your data after it has been expanded and then reduced. First, it delivers four sets of estimates about three major data types: email files, structured data, and unstructured data. The sets are percentage expected (e.g., 15% of your data will be email files), estimated GBs (e.g., you will have 12 GB of email), estimated files/GB (e.g., you will have 500 files per GB), and total files (e.g., you will have 14,000 files). Second, the Supplemental Data Calculator delivers the same four sets of estimates for six subcategories of unstructured data: word processing files, spreadsheet files, presentation files, image files, PDF files, and other unstructured data.

USE

In the Entry portion of the spreadsheet (columns D through K), you enter numbers and select options. For example, you need to enter how many gigabytes of data you plan to start with and select which data reductions steps, if any, you intend to use.

From the data you enter, the core part of the spreadsheet calculates the amount of data you may have after expansion and the reductions in data volumes you may get during processing. The supplemental part of the spreadsheet calculates data volumes by major data types – email files, structured data, and unstructured data – as well as by subcategories for unstructured

data. These estimates appear both in the Entry side of the spreadsheet (columns D through K) and the Report side of the spreadsheet (columns N through R).

INSTRUCTIONS

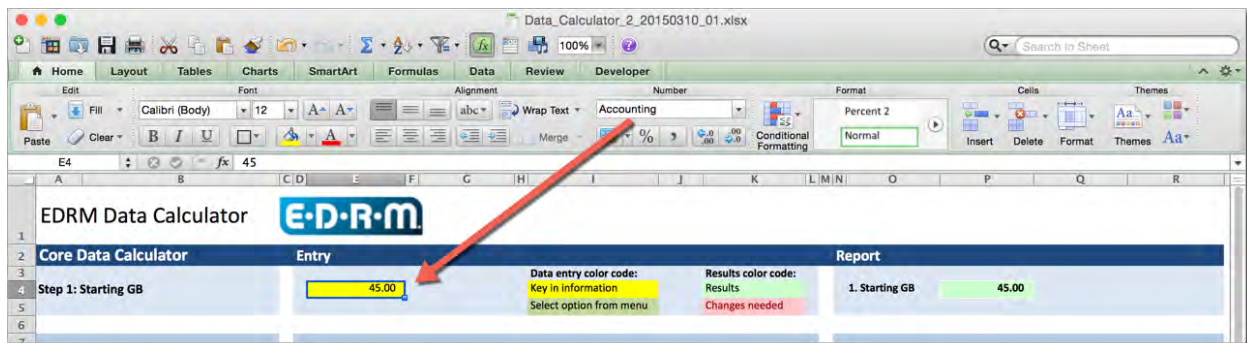
CORE DATA CALCULATOR

Step 1: Starting GB

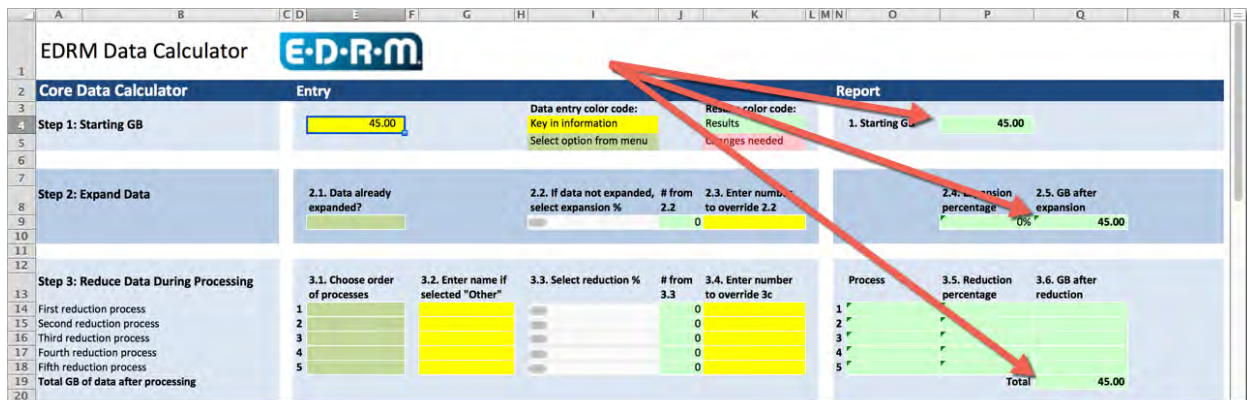
To start using the EDRM Data Calculator, enter the amount of data you intend to process.

In cell E4, enter the number of GBs of data you intend to start with. You may use whole numbers (e.g., “45” for 45 gigabytes) and decimals (e.g., “45.5” for 45 and ½ gigabytes). Once you enter a number, you should see it displayed with two decimal digits (e.g., “45” will appear as “45.00”).

For the value you enter in cell E4, use a known number or an estimate.



When you enter a value in cell E4, that number will appear in cells P4, Q9, and Q19.

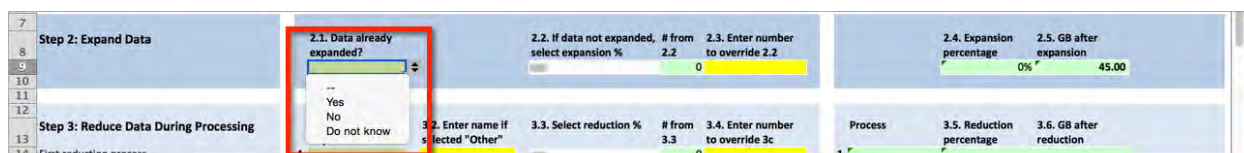


Step 2: Expand Data

Processing electronic data often causes the data's volume to expand. This expansion can make predicting costs a challenge. Use step 2 to help gain clearer insight into the amount of data you can expect for processing purposes.

Step 2.1: Has Your Data Been Expanded?

In cell E9, select the option that fits whether your data has been expanded. Cell E9 contains a drop-down menu from which you can select one of three options: "Yes", meaning your data already has been expanded; "No", meaning your data has not been expanded; or "Do not know", meaning you do not know whether your data has been expanded.



If you select "Yes", go to Step 3 where you will enter information about reducing data during processing. The values in cells Q9 and Q19 will reflect the starting GB value you entered in cell E4.

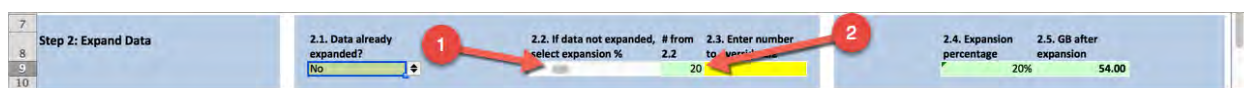
If you select "No," continue to Step 2.2 or 2.3 where you will enter an expansion percentage.

If you select "Do not know," either continue to Step 2.2 or 2.3 where you will enter an expansion percentage **or go to Step 3** where you will enter information about reducing data during processing.

Step 2.2: Data Expansion Value

If you selected "No" or "Do not know" in Step 2.1 (cell E9), you can use the slider in cell I9 to set an expansion percentage for your data. To use the slider, click on the bar in cell I9 (1, below) and move it to the right or the left until you see the percentage you want displayed in cell J9 (2, below).

For the percentage you set, use a known number or an estimate.



When you set a percentage, the number will appear as a percentage on the Report side of the spreadsheet, in cell P9 (1, below).

The spreadsheet will use that percentage and the Starting GB number to estimate how many GBs of data you will have after expansion. That number will appear in cell Q9 (2, below) as well as in cell Q19.

7	Step 2: Expand Data	2.1. Data already expanded?	2.2. If data not expanded, # from select expansion %	2.3. Enter number to override 2.2	2.4. Expansion percentage	2.5. GB after expansion
8		<input type="text" value="No"/>	<input type="text" value="20"/>	<input type="text" value="20"/>	20%	54.00
9						
10						

Step 2.3: Data Expansion Value Override

If it is more convenient for you to enter the expansion percentage manually, type a number in cell K9 instead of using the slider. Do not enter a percentage sign (%). You may use whole numbers (e.g., “20” for 20%) and decimals (e.g., “20.5” for 20 and ½ percentage). Once you enter a number, you should see it displayed with two decimal digits followed by a percentage sign (e.g., “20” will appear as “20.00%”).

Entering a value in cell K9 will override any value selected using the slider in cell I9.

7	Step 2: Expand Data	2.1. Data already expanded?	2.2. If data not expanded, # from select expansion %	2.3. Enter number to override 2.2	2.4. Expansion percentage	2.5. GB after expansion
8		<input type="text" value="No"/>	<input type="text" value="0"/>	<input type="text" value="20.00%"/>	20.00%	54.00
9						
10						

The number you enter will appear in cells K9 and P9. The spreadsheet will use that percentage and the Starting GB number to estimate how many GBs of data you will have after expansion. That number will appear in cell Q9 (2, below) as well as in cell Q19.

Step 3: Reduce Data During Processing

You may want to use one or more methods to reduce the amount of data you use. You can reduce data by various methods, including through de-Nesting, de-duplication, search terms, and CAR (computer assisted review). Step 3 offers a means of factoring in data reduction processes as you estimate how much data you will have after the data has been processed.

Step 3.1: Select Data Reduction Processes

Using the five dropdowns in cells E14 through E18, select the data reduction processes you intend to use with your data set. You may choose up to five processes. You may choose any of the four pre-set processes (De-nesting, De-duplication, Search Terms, and CAR). If you prefer, you may add your own processes (see Step 3.2).

To choose a process, start with the first row (cell E14), click in the cell, click on the up/down arrows to the right of the cell, and select the desired option. You can follow the same approach with cells E15, E16, E17, and E18.

Step 3: Reduce Data During Processing					Process	3.5. Reduction percentage	3.6. GB after reduction	
3.1. Choose order of processes	3.2. Enter name if selected "Other"	3.3. Select reduction %	# from 3.3	3.4. Enter number to override 3c				
1			0		1			
2			0		2			
3	De-nesting		0		3			
4	De-duplication		0		4			
5	Search terms		0		5			
							Total	54.00

Once you choose a process, it will appear in the corresponding cells in columns E and O.

Step 3: Reduce Data During Processing					Process	3.5. Reduction percentage	3.6. GB after reduction
3.1. Choose order of processes	3.2. Enter name if selected "Other"	3.3. Select reduction %	# from 3.3	3.4. Enter number to override 3c			
1	De-nesting		0		1	0%	54.00
2			0		2		
3			0		3		

Step 3.2: "Other" Data Reduction Processes

If you want to use a data reduction process that is not specifically listed in the drop-down options, select "Other" and then manually enter the process in the corresponding row in column G. For example, if in cell E14 you select "Other" (1, below) then enter the specific process in Step cell G14 (2, below). The text you enter will appear in cells G14 (2, below) and O14 (3, below).

Step 3: Reduce Data During Processing					Process	3.5. Reduction percentage	3.6. GB after reduction	
3.1. Choose order of processes	3.2. Enter name if selected "Other"	3.3. Select reduction %	# from 3.3	3.4. Enter number to override 3c				
1	Other	Date range	0		1	0%	54.00	
2			0		2			
3			0		3			
4			0		4			
5			0		5			
							Total	54.00

Step 3.3: Reduction Percentage

After you chose reduction methods (cells E14-E18), you can use the corresponding sliders to set the reduction percentages for your data. To use one of the sliders, click on the bar in the cell (I14, for example (1, below)) and move it to the right or the left until you see the percentage you want displayed in the corresponding cell (J14, in this example (2, below)).

Step 3: Reduce Data During Processing					Process	3.5. Reduction percentage	3.6. GB after reduction
3.1. Choose order of processes	3.2. Enter name if selected "Other"	3.3. Select reduction %	# from 3.3	3.4. Enter number to override 3c			
1	De-nesting		15		1	15.00%	45.90
2			0		2		
3			0		3		

When you set a percentage, the number will appear as a percentage on the Report side of the spreadsheet in the corresponding cell (P14 in this example (1, below)).

For the first row, the spreadsheet will use that percentage and either the starting GB number if the data already has been expanded (cell P4) or the post-expansion number (cell Q9) to estimate how many GBs of data you will have after the first reduction process. That number will appear in cell Q14 (2, below).

12	Step 3: Reduce Data During Processing					1	Process	3.5. Reduction percentage	3.6. GB after reduction
13		3.1. Choose order of processes	3.2. Enter name if selected "Other"	3.3. Select reduction %	# from 3.3	3.4. Enter number to override 3c			
14	1	De-nesting			15		15.00%	45.90	
15	2	De-duplication			0		15.00%	39.02	
16	3	Search terms			0		30.00%	27.31	

Repeat as desired for rows 15 through 18.

Once you have entered all values and reduction percentages you want in rows 15-18, cell Q19 will reflect the total GB remaining after factoring in all reduction processes and amounts.

12	Step 3: Reduce Data During Processing					Process	3.5. Reduction percentage	3.6. GB after reduction
13		3.1. Choose order of processes	3.2. Enter name if selected "Other"	3.3. Select reduction %	# from 3.3	3.4. Enter number to override 3c		
14	1	De-nesting			15		15.00%	45.90
15	2	De-duplication			15		15.00%	39.02
16	3	Search terms			30		30.00%	27.31
17	4				0			
18	5				0			
19	Total GB of data after processing						Total	27.31

Step 3.4: Data Reduction Value Override

If it is more convenient for you to enter the reduction percentage manually, type a number in the corresponding cell (K14 in this example) instead of using the slider. Do not enter a percentage sign (%). You may use whole numbers (e.g., "20" for 20%) and decimals (e.g., "20.5" for 20 and ½ percentage). Once you enter a number, you should see it displayed with two decimal digits followed by a percentage sign (e.g., "20" will appear as "20.00%").

Entering a value in cells K14-K18 will override any values selected using the sliders in cells I14-I18.

12	Step 3: Reduce Data During Processing					Process	3.5. Reduction percentage	3.6. GB after reduction
13		3.1. Choose order of processes	3.2. Enter name if selected "Other"	3.3. Select reduction %	# from 3.3	3.4. Enter number to override 3c		
14	1	De-nesting			6	15.00%	15.00%	45.90
15	2	De-duplication			15		15.00%	39.02
16	3	Search terms			30		30.00%	27.31

SUPPLEMENTAL DATA CALCULATOR

Step 4: Estimate By Data Type

The Supplemental Data Calculator is optional. You may stop at Step 3.4, if you like. You may continue to Step 4.1 and stop there. You may continue to Step 4.2 as well.

With the Supplemental Data Calculator you can further refine your calculations based on the type of email, structured, and unstructured data files being handled. Step 4.1 focuses on email, structured data, and unstructured data. Step 4.2 goes into subcategories of unstructured data.

If you do not have enough information about unstructured data subcategories or are not interested in that level of detail, you can stop at Step 4.1.

If you have sufficiently reliable information and are interested in the level of detail found in Step 4.2, complete that step as well. The results displayed in Step 4.2 will be solely for the

unstructured data whose information you entered in row 28. Those results also will be shown in row 30.

Step 4.1: Major Data Types

For email files, structured data, and unstructured data, enter your estimates for the percentage of your total data set each of those data types account for. Do not enter a percentage sign (%). For example, you might enter “60” in cell E26 for email (1, below), “10” for in cell E27 for structured data (2, below), and “30” in cell E30 for unstructured data (3, below). You may use whole numbers (e.g., “20” for 20%) and decimals (e.g., “20.5” for 20 and ½ percentage). Once you enter a number, you should see it displayed with two decimal digits followed by a percentage sign (e.g., “20” will appear as “20.00%”).

A running total will appear in cells G26, G27, and G30. Once you have entered percentages for email, structured data, and unstructured data, the total should be 100.00%. If it is not, adjust your percentages until you reach 100%.

Supplemental Data Calculator - Entry						Report			
Step 4: Estimate by Data Type						4.1.5. Percentage expected	4.1.6. Estimated GB	4.1.7. Estimated files/GB	4.1.8. Total files
4.1. Major Data Types	4.1.1. Expected % of total population	4.1.2. Running total	4.1.3. Select estimated number of files per GB	# from 4.1.3	4.1.4. Enter number to override				
Email files	60.00%	60.00%	<input type="text"/>	0		60.00%	16.39	-	-
Structured data	10.00%	70.00%	<input type="text"/>	0		10.00%	2.73	-	-
Unstructured data	30.00%	100.00%	<input type="text"/>	0		30.00%	8.19	-	-
Totals and average: Email, structured data, and unstructured data WITHOUT subcategories	Total too low by	100.00%	You are not using subcategories. The numbers from 4.1.3 or 4.1.4 will be used.			100.00%	27.31	-	-
						Total	Total	Average	Total

After you enter data type percentages, you can use the corresponding sliders to enter your estimates for the number of files per GB you expect to have for each data type. To use one of the sliders, click on the bar in the corresponding row and move it to the right or the left until you see the percentage you want displayed in the corresponding cell.

Supplemental Data Calculator - Entry						Report			
Step 4: Estimate by Data Type						4.1.5. Percentage expected	4.1.6. Estimated GB	4.1.7. Estimated files/GB	4.1.8. Total files
4.1. Major Data Types	4.1.1. Expected % of total population	4.1.2. Running total	4.1.3. Select estimated number of files per GB	# from 4.1.3	4.1.4. Enter number to override				
Email files	60.00%	60.00%	<input type="text"/>	345		60.00%	16.39	345	5,653
Structured data	10.00%	70.00%	<input type="text"/>	2500		10.00%	2.73	2,500	6,828
Unstructured data	30.00%	100.00%	<input type="text"/>	1026		30.00%	8.19	1,026	8,406
Totals and average: Email, structured data, and unstructured data WITHOUT subcategories	Total too low by	100.00%	You are not using subcategories. The numbers from 4.1.3 or 4.1.4 will be used.			100.00%	27.31	765	20,887
						Total	Total	Average	Total

If it is more convenient for you to enter the data type percentages manually, type the number in the corresponding cell instead of using the slider. Do not enter a percentage sign (%). You may use whole numbers (e.g., “20” for 20%) and decimals (e.g., “20.5” for 20 and ½ percentage). Once you enter a number, you should see it displayed with two decimal digits followed by a percentage sign (e.g., “20” will appear as “20.00%”).

Entering a value in cells K26, K27, or K30 will override any values selected using the sliders in cells I26, I17, and I30.

Supplemental Data Calculator		Entry				Report				
22	Step 4: Estimate by Data Type	4.1.1. Expected % of total population	4.1.2. Running total	4.1.3. Select estimated number of files per GB	# from 4.1.3	4.1.4. Enter number to override	4.1.5. Percentage expected	4.1.6. Estimated GB	4.1.7. Estimated files/GB	4.1.8. Total files
24	4.1. Major Data Types									
25	Email files	60.00%	60.00%		345	500	60.00%	16.39	500	8,193
26	Structured data	10.00%	70.00%		2500	1500	10.00%	2.73	1,500	4,097
27	Unstructured data									
28	Use subcategories? If "No", then complete 4.1.1 AND (4.1.3 OR 4.1.4)	30.00%	100.00%		1026	750	30.00%	8.19	750	6,145
29										
30										
31										

In cell E29, select “No” if you do not wish to use subcategories for estimating unstructured data OR select “Yes” if you do wish to use subcategories for estimating.

If you selected “No”, the results of the spreadsheet’s calculations will be displayed in the “Report” section.

Supplemental Data Calculator		Entry				Report				
22	Step 4: Estimate by Data Type	4.1.1. Expected % of total population	4.1.2. Running total	4.1.3. Select estimated number of files per GB	# from 4.1.3	4.1.4. Enter number to override	4.1.5. Percentage expected	4.1.6. Estimated GB	4.1.7. Estimated files/GB	4.1.8. Total files
24	4.1. Major Data Types									
25	Email files	60.00%	60.00%		345	500	60.00%	16.39	500	8,193
26	Structured data	10.00%	70.00%		2500	1500	10.00%	2.73	1,500	4,097
27	Unstructured data									
28	Use subcategories? If "No", then complete 4.1.1 AND (4.1.3 OR 4.1.4)	No	100.00%		1026	750	30.00%	8.19	750	6,145
29										
30										
31	Totals and average: Email, structured data, and unstructured data WITHOUT subcategories		100.00%				100.00%	27.31	675	18,435
32							Total	Total	Average	Total
33										
34										

If you selected “Yes”, continue to Step 4.2.

Step 4.2: Unstructured Data By Subcategories

If you selected “Yes” in Step 4.1 above, enter your estimates for unstructured data types in this section.

For each subcategory, enter the expected percentage of unstructured data. Do not enter a percentage sign (%). For example, you might enter “60” in cell E39 for word processing files, “10” for in cell E40 for spreadsheet files, and “30” in cell E43 for PDF files. You may use whole numbers (e.g., “20” for 20%) and decimals (e.g., “20.5” for 20 and ½ percentage). Once you enter a number, you should see it displayed with two decimal digits followed by a percentage sign (e.g., “20” will appear as “20.00%”).

Fill in any combination of subcategories that is appropriate in your situation. You do not need to fill in all the rows.

A running total will appear in cells G39-G44. You want to make sure that your total in cell G46 is 100.00%. If it is not, adjust your percentages until you reach 100%.

4.2. Unstructured Data by Subcategories		4.2.1. Expected % of unstructured	4.2.2. Running total	4.2.3. Select estimated number of files per GB	# from 4.2.3	4.2.4. Enter number to override	4.2.5. % of unstructured	4.2.6. Estimated GB	4.2.7. Estimated files/GB	4.2.8. Total files
36	Word processing files	60.00%	60.00%		0		60.00%	4.92	-	-
37	Spreadsheet files	10.00%	70.00%		0		10.00%	0.82	-	-
38	Presentation files		70.00%		0		0.00%	-	-	-
39	Image files (JPEG, TIFF, etc)		70.00%		0		0.00%	-	-	-
40	PDF files	30.00%	100.00%		0		30.00%	2.46	-	-
41	Other unstructured data		100.00%		0		0.00%	-	-	-
42	Unstructured data: totals and average		100.00%				100.00%	8.19	-	-
43							Total	Total	Average	Total
44										
45										
46										
47										
48										

After you enter subcategory percentages, you can use the corresponding sliders to enter your estimates for the number of files per GB you expect to have for each subcategory. To use one of the sliders, click on the bar in the corresponding row and move it to the right or the left until you see the percentage you want displayed in the corresponding cell.

4.2. Unstructured Data by Subcategories	4.2.1. Expected % of unstructured	4.2.2. Running total	4.2.3. Select estimated number of files per GB	# from 4.2.3	4.2.4. Enter number to override	4.2.5. % of unstructured	4.2.6. Estimated GB	4.2.7. Estimated files/GB	4.2.8. Total files
Word processing files	60.00%	60.00%	<input type="range"/>	1754		60.00%	4.92	1,754	8,622
Spreadsheet files	10.00%	70.00%	<input type="range"/>	3130		10.00%	0.82	3,130	2,564
Presentation files		70.00%	<input type="range"/>	0		0.00%			
Image files (JPEG, TIFF, etc)		70.00%	<input type="range"/>	0		0.00%			
PDF files	30.00%	100.00%	<input type="range"/>	1379		30.00%	2.46	1,379	3,390
Other unstructured data		100.00%	<input type="range"/>	0		0.00%			

If it is more convenient for you to enter the subcategory percentages manually, type the number in the corresponding cell instead of using the slider. Do not enter a percentage sign (%). You may use whole numbers (e.g., “20” for 20%) and decimals (e.g., “20.5” for 20 and ½ percentage). Once you enter a number, you should see it displayed with two decimal digits followed by a percentage sign (e.g., “20” will appear as “20.00%”).

Entering a value in cells K39-K44 will override any values selected using the sliders in cells 39-144.

4.2. Unstructured Data by Subcategories	4.2.1. Expected % of unstructured	4.2.2. Running total	4.2.3. Select estimated number of files per GB	# from 4.2.3	4.2.4. Enter number to override	4.2.5. % of unstructured	4.2.6. Estimated GB	4.2.7. Estimated files/GB	4.2.8. Total files
Word processing files	60.00%	60.00%	<input type="range"/>	1754	500	60.00%	4.92	500	2,458
Spreadsheet files	10.00%	70.00%	<input type="range"/>	3130	1000	10.00%	0.82	1,000	819
Presentation files		70.00%	<input type="range"/>	0		0.00%			
Image files (JPEG, TIFF, etc)		70.00%	<input type="range"/>	0		0.00%			
PDF files	30.00%	100.00%	<input type="range"/>	1379	750	30.00%	2.46	750	1,843
Other unstructured data		100.00%	<input type="range"/>	0		0.00%			

The results of the spreadsheet’s calculations will be displayed in the “Report” section for both Step 4.1 and Step 4.2.

Supplemental Data Calculator					Report			
Entry					Report			
Step 4: Estimate by Data Type					4.1.5. Percentage expected	4.1.6. Estimated GB	4.1.7. Estimated files/GB	4.1.8. Total files
4.1. Major Data Types					60.00%	16.39	500	8,193
Email files					10.00%	2.73	1,500	4,097
Structured data								
Unstructured data								
Use subcategories? If 'Yes', then complete 4.1.1 (right) and 4.2 (below)					30.00%	8.19	625	5,121
Totals and average: Email, structured data, and unstructured data WITH subcategories					100.00%	27.31	638	17,410
Total too low by					Total	Total	Average	Total
4.2. Unstructured Data by Subcategories					4.2.5. % of unstructured	4.2.6. Estimated GB	4.2.7. Estimated files/GB	4.2.8. Total files
Word processing files					60.00%	4.92	500	2,458
Spreadsheet files					10.00%	0.82	1,000	819
Presentation files					0.00%			
Image files (JPEG, TIFF, etc)					0.00%			
PDF files					30.00%	2.46	750	1,843
Other unstructured data					0.00%			
Unstructured data: totals and average					100.00%	8.19	625	5,121
Total unstructured					Total	Total	Average	Total

PREPARED BY

Neda Shakoori, Sheri Towne and George Socha