












# Early Data Assessment Report

**batch name: Custodians - ALL**

**upload date: 11/29/2016 11:23:02 AM**

## Summary

<p>Files</p>  <p><b>361<sup>(1)</sup></b></p>	<p>Expanded Size</p>  <p><b>37.68 MB <sup>(2)</sup></b></p>	<p>Files Requiring OCR</p>  <p><b>0<sup>(3)</sup></b></p>
<p>Email Attachments</p>  <p><b>88</b></p>	<p>Family Groups</p>  <p><b>55</b></p>	<p>Domains</p>  <p><b>246</b></p>
<p>Exceptions</p>  <p><b>0</b></p>	<p>File Types</p>  <p><b>15</b></p>	<p>Conversation threads</p>  <p><b>1</b></p>

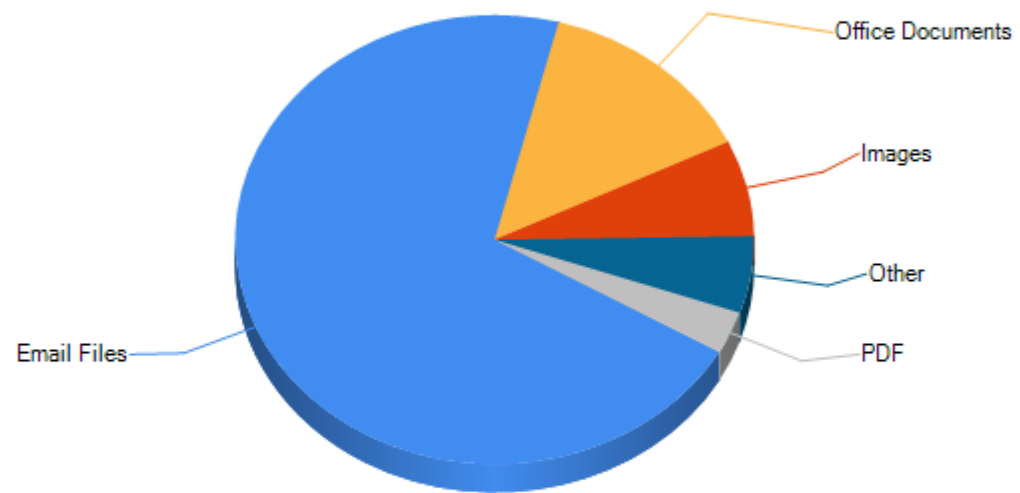
1. does not include PST or compressed containers (ie. .zip, .rar, ect)

2. gigs are based on 1 GB = 1024 MBs

3. number of tiff images (the number of PDFs requiring OCR can not be calculated without further processing)

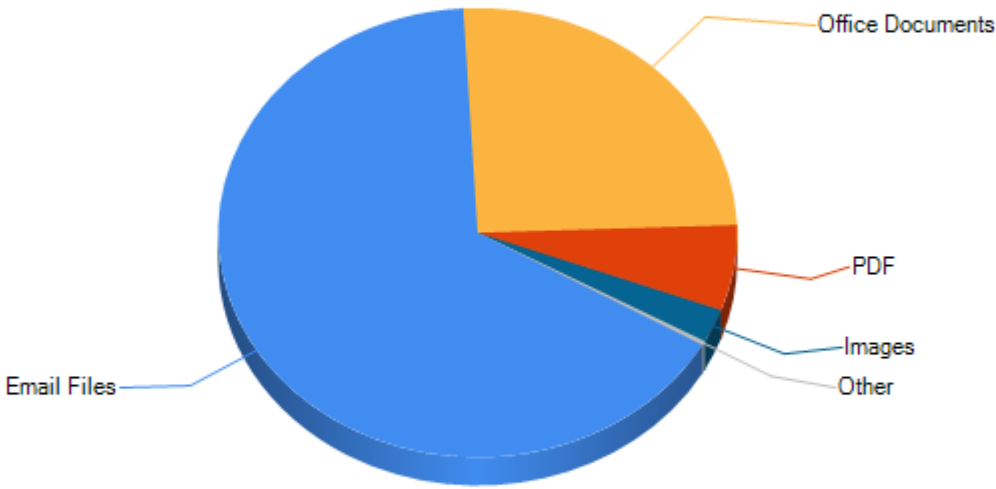
## Top File Categories by Count

File Category	Count	Percentage
Email Files	255	70%
Office Documents	50	13%
Images	25	6%
Other	20	5%
PDF	11	3%



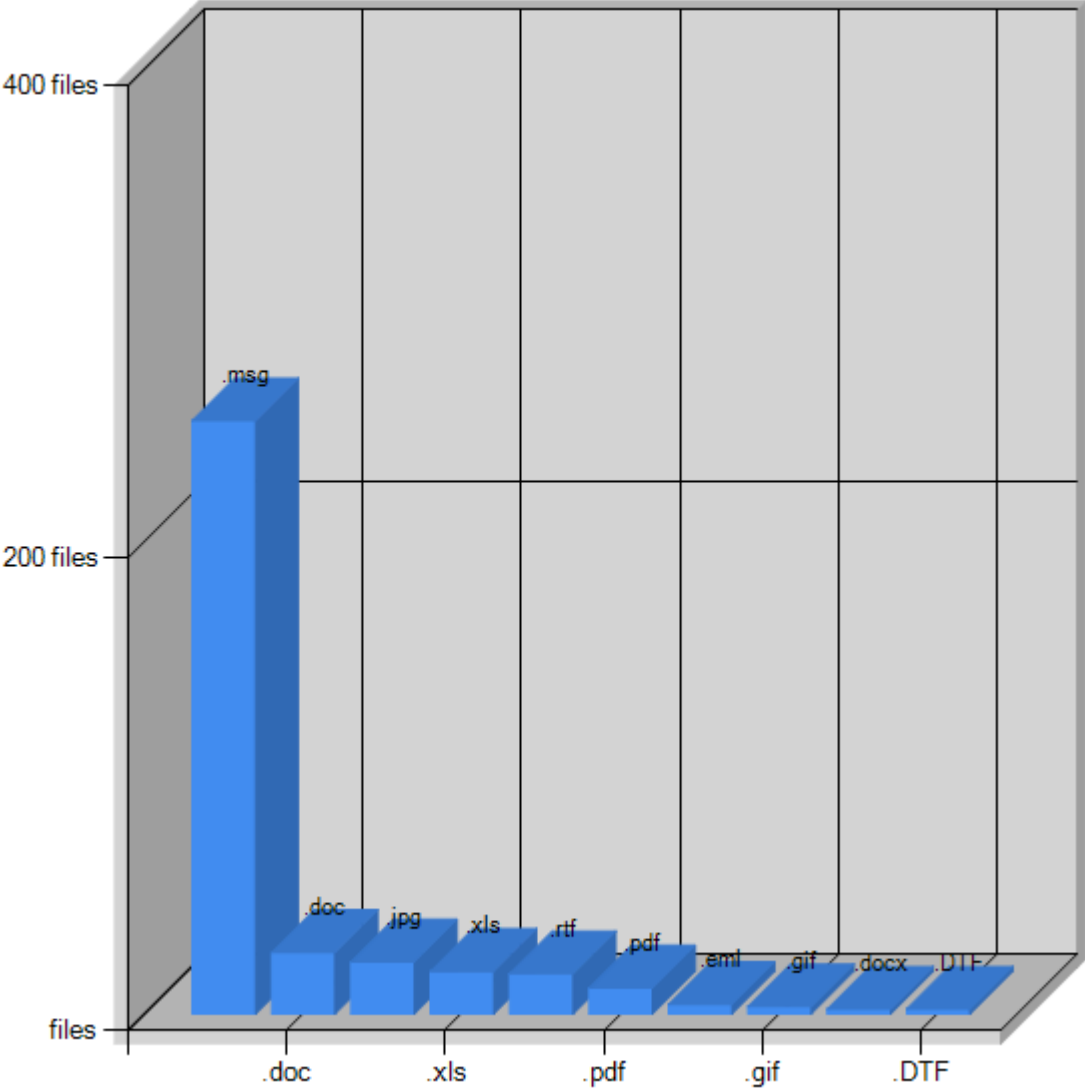
### Top File Categories by Size

File Category	Count	Percentage
Email Files	24.78 MB	65%
Office Documents	9.55 MB	25%
PDF	2.32 MB	6%
Images	933 KB	2%
Other	119 KB	< 1%



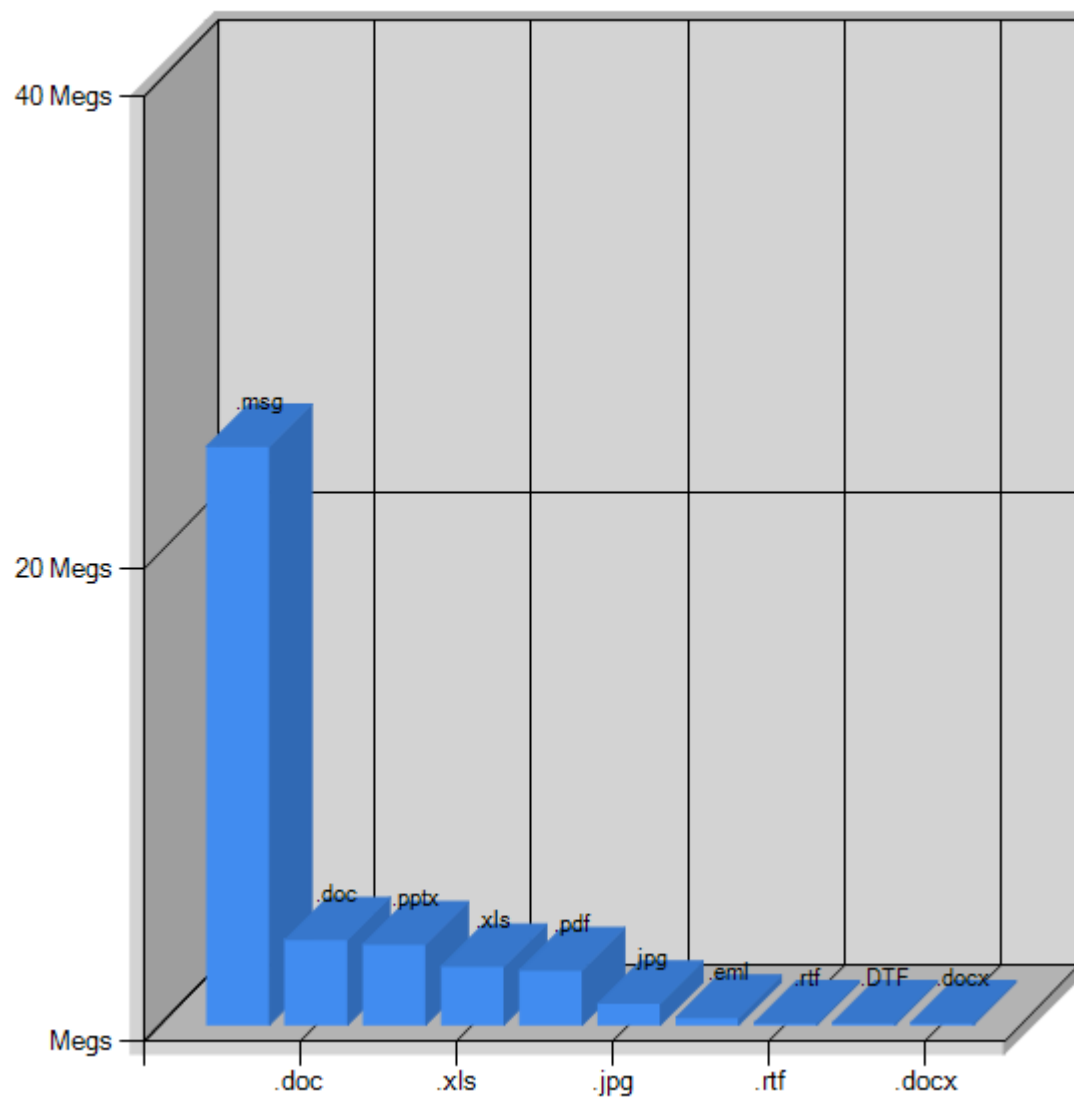
### Top 10 File Types by Count

File Type	Count
.msg	251
.doc	26
.jpg	22
.xls	18
.rtf	17
.pdf	11
.eml	4
.gif	3
.docx	2
.DTF	2



### Top 10 File Types by Size

File Type	Count
.msg	24.48 MB
.doc	3.60 MB
.pptx	3.42 MB
.xls	2.47 MB
.pdf	2.32 MB
.jpg	921 KB
.eml	308 KB
.rtf	67 KB
.DTF	52 KB
.docx	32 KB

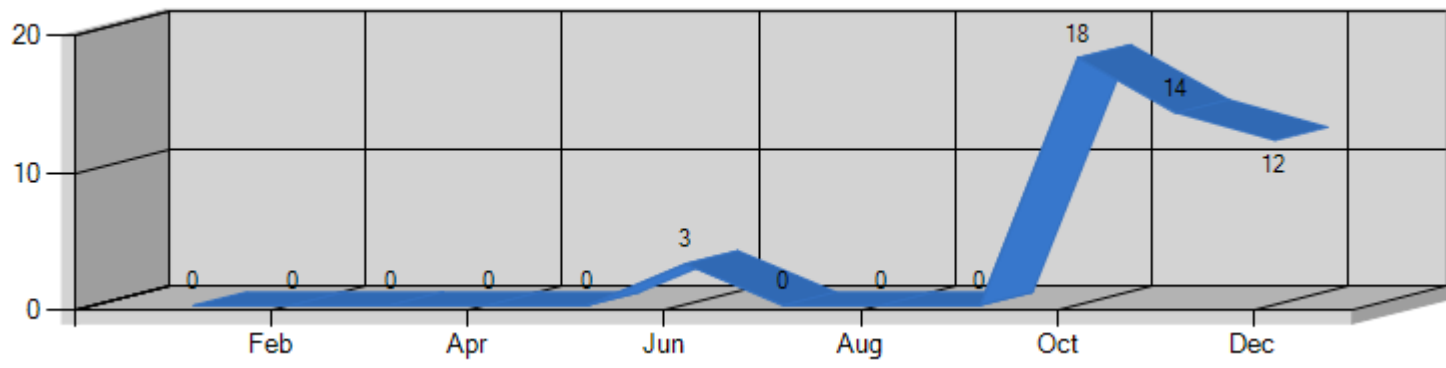


## Top 25 Domains (number of instances)

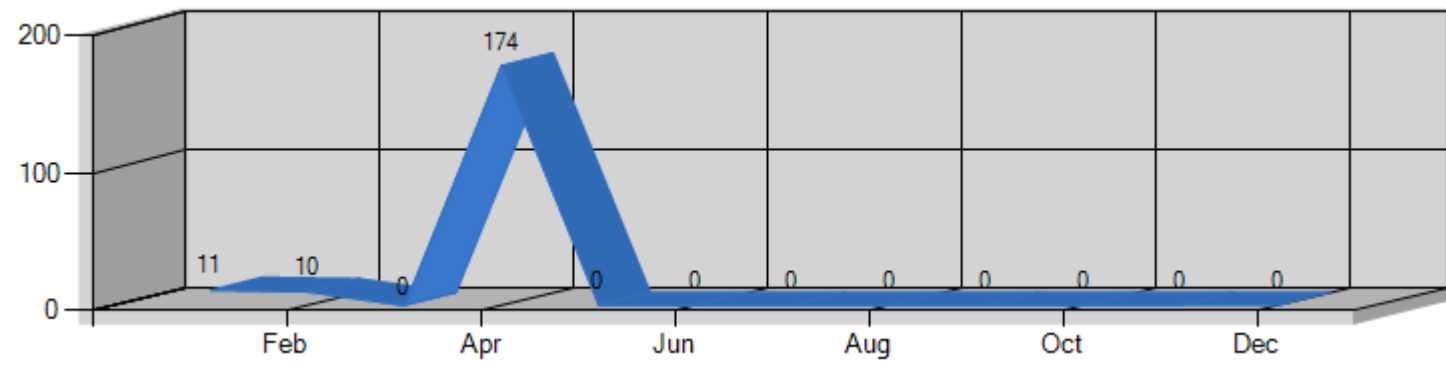
File Type	Count
enron.com	223
bp.com	188
williams.com	188
coral-energy.com	146
utilicorp.com	116
aec.ca	99
br-inc.com	95
engageenergy.com	85
duke-energy.com	80
cngc.com	77
reliantenergy.com	75
avistaenergy.com	74
kimballenergy.com	72
pge.com	71
puget.com	66
avistacorp.com	63
dynegy.com	57
aol.com	55
enserco.com	54
pgn.com	53
epenergy.com	53
wasatchenergy.com	47
westerngas.com	47
neg.pge.com	45
texaco.com	44

# Email Timeline

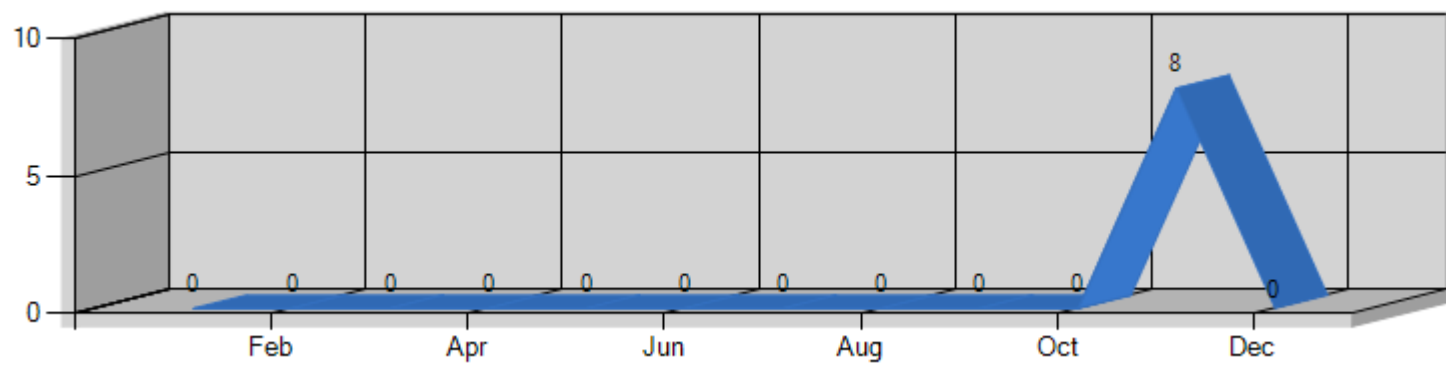
2000



2001



2002



2008

