

IMLS MUSEUM UNIVERSE DATA FILE



Data Catalog

~24 TB WARC

9.9 GB CDX (compressed)

138,905,796 URLs

Based on Museum Universe Data File FY 2016 Q1
 This dataset provides a list of known museums and related organizations in the United States

MID	COMMONNAME	WEBURL
1	8,400,400 SAWTOOTH MOUNTAINS HISTORICAL SOCIETY	HTTP://DISCOVERSAWTOOTH.ORG/

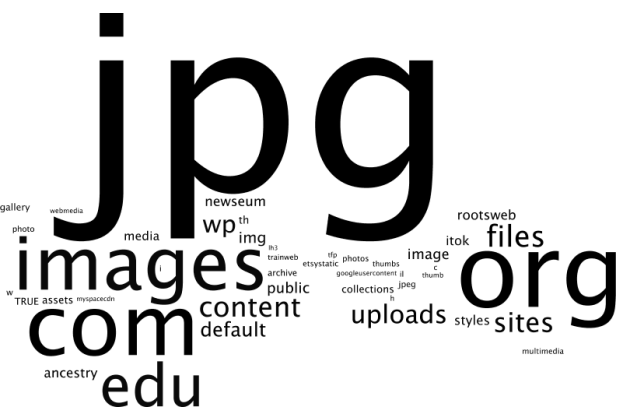
		Name	URL	Income	Pages	Size	404	500	Robots	size_gb
2	8,400,800 MESA VE									
3	8,401,200 SPADY M									
4	8,401,300 LANIER M	VALLEY INSTITUTE VISUAL ART	www.vivagallery.org	22.0	2153	123098457	8	0	False	0.114644
5	8,402,300 ACADIA M									
6	8,402,500 ROBERT									
7	8,402,500 GEORGE	6 COLLEGE OF EASTERN UTAH	www.ceu.edu	336.0	22676	1589230787	138	0	True	1.480087
8	8,402,500 BABSON									
9	8,403,800 FORT BU									
10	8,404,000 PAOLI OK	4 SHELBY COUNTY HISTORICAL SOCIETY INC	shelby.mogenweb.org	7156.0	10980	22380537929	94	0	False	20.843500
11	8,404,200 SUSQUEH									
12	8,405,000 SHOREH									
13	8,405,100 WARREN									
14	8,405,300 MUSEUM	TEMPE HISTORICAL SOCIETY	www.tempe.gov	6053.0	119784	5044019549	467	78	True	4.697609
15	8,409,400 WAQUON									
16	8,403,800 MOUSE R									
17	8,409,501 BENT'S O	ALLIED GARDENS PARK & RECREATION COUNCIL	www.sandiego.gov	113766.0	53357	74895087335	3093	0	True	69.751486
18	8,403,800 BUFFALO									
19	8,404,800 FORT DA									
20	8,409,400 1761 POV									
21	8,403,700 RYLAND I	9 ALLEN COUNTY COURTHOUSE PRESERVATION TRUST INC	in.gov	91289.0	152042	50989630130	11144	0	True	47.487794

DOMAIN TREE MAP



- 72,929,853 .org
- 27,779,573 .com
- 23,703,090 .edu
- 4,225,733 .us
- 2,387,102 .gov
- 1,730,192 .net

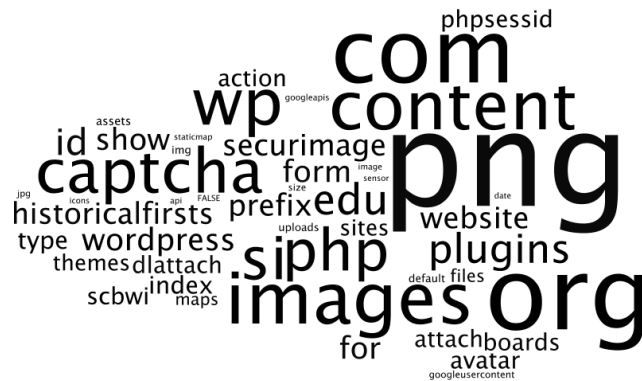
URL TEXT ANALYSIS by MIMETYPE



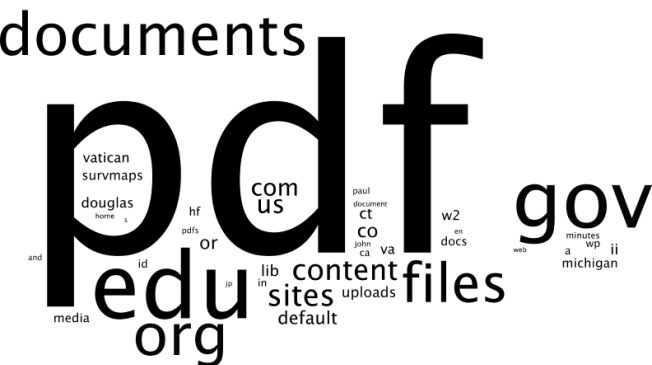
14,101,088 JPEGs



2,415,025 GIFs



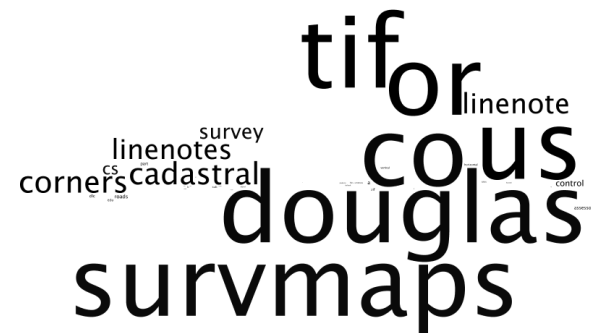
1,840,497 PNGs



2,093,525 PDFs



1,718,110
Unknown type



69,843 TIFFs