

Should talk about finished datasets: There are websites that host content like malware that you can search and download, but I will not consider this a dataset.

Describing the possible values for the columns (starting with column 1) in table ?

1. I stand for Item and is the sequence number (1,2,3,...,end of list)
2. Acc stand for Access and can be {Pub=public, Req=By request}
3. DT stands for Data type and can be {S=synthetic,R=Real, H=Hybrid, + file types extension?}
4. Cat stands for category and can be {MAL=Malware, Net=Network traffic, ..., }
5. Size=size of dataset in Giga Bytes (GB), or maby in samples (have to include this as not all is given any other size information), compressed/uncompressed, unknown=?
6. N/A for NOT available

I	Acc	DT	Cat	Size	Description
1	P	R	MAL	0.05	
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					