

Data Set Name	Document Name	Size (MB)	Tags	Number of Nodes	Depth	Data Ratio
EXI	EXI-Telecomp.xml	0.65	39	651398	7	0.48
	EXI-Weblog.xml	2.60	12	178419	3	0.31
	EXI-Invoice.xml	0.93	52	78377	7	0.57
	EXI-Array.xml	22.18	47	1168115	10	0.68
	EXI-Factbook.xml	4.12	199	104117	5	0.53
	EXI-Geographic Coordinates.xml	16.20	17	55	3	1
XMark	XMark1.xml (Small)	11.40	74	520546	12	0.74
	XMark2.xml (Medium)	113.80	74	5167121	12	0.74
	XMark3.xml (Large)	571.75	74	25900899	12	0.74
XBench	DCSD-Small.xml	10.60	50	6190628	8	0.45
	DCSD-Normal.xml	105.60	50	6190628	8	0.45
	TCSD-Small.xml	10.95	24	831393	8	0.78
	TCSD-Normal.xml	106.25	24	8085816	8	0.78
Wikipedia	EnWikiNews.xml	71.09	20	2013778	5	0.91
	EnWikiQuote.xml	127.25	20	2672870	5	0.97
	EnWikiSource.xml	1036.66	20	13423014	5	0.98
	EnWikiVersity.xml	83.35	20	3333622	5	0.91
	EnWikTionary.xml	570.00	20	28656178	5	0.77
DBLP	DBLP.xml	130.72	32	4718588	5	0.58
U.S House	USHouse.xml	0.52	43	16963	16	0.77
SwissProt	SwissProt.xml	112.13	85	13917441	5	0.60
NASA	NASA.xml	24.45	61	2278447	8	0.66
Shakespeare	Shakespeare.xml	7.47	22	574156	7	0.64
Lineitem	Lineitem.xml	31.48	18	2045953	3	0.19
Mondial	Mondial.xml	1.75	23	147207	5	0.77
BaseBall	BaseBall.xml	0.65	46	57812	6	0.11
Treebank	Treebank.xml	84.06	250	10795711	36	0.70
Random	Random-R1.xml	14.20	100	1249997	28	0
	Random-R2.xml	53.90	200	3750002	34	0
	Random-R3.xml	97.85	300	7500017	30	0