

3.5 Memory Limits

The Sci2 Tool's database functionality greatly improves the volume of data which can be loaded into and analyzed by the tool. Whereas most scientometric tools available as of March 2010 require powerful computing resources to perform large scale analyses each time a network needs to be extracted, the pre-loaded database runs network extraction algorithms quickly and allows the users to format their own custom queries. This functionality has some front-heavy memory requirements in order to initially load the data into the database, the upper limits of which can be seen in Tables 3.1 and 3.2.

Entries	Load	Merge People	Merge Journals	Match	Extract Authors	Extract Documents	Extract Co-Authors	Extract Document Citation (with outer)	Extract Author Citation	Extract Document Co-Citation	Extract Document Co-Citation (core)	Author Co-Citation
50	5	1	1	1	1	1	1	1	1	10	4	2
500	10	4	4	1	1	1	1	1	1	20	4	5
5000	100	48	180	3	1	1	3	18	2	3300	20	35
10000	210	120	480	5	2	1	10	30	7		60	100
20000	400	300	1200	15	5	1	22	70	20		164	250

Table 3.1: The number of seconds to perform each action on a dataset of the given size, on a computer with 12GB of memory

Entries	Load	Merge People	Merge Journals	Match	Extract Authors	Extract Documents	Extract Co-Authors	Extract Document Citation (with outer)	Extract Author Citation	Extract Document Co-Citation	Extract Document Co-Citation (core)	Author Co-Citation
50	50	55	9	4	3	1	1	1	1	1	40	13
500	500	420	52	13	3	1	1	1	2	1	75	13
5000	5000	3600	1080	480	25	1	1	12	90	2		180

Table 3.2: The number of seconds to perform each action on a dataset of the given size, on a computer with 1.2GB of memory