Overview in General

This file contains a general overview of the data in the graph including node labels and relationships types.

References

- jqassistant
- Neo4j Python Driver

Node Labels

Table 1a - Highest node count by label combination

Lists the 30 label combinations with the highest number of nodes. The labels with the lowest node count are listed in table 1b. The total list would sum up to the total number of labels (100%).

The whole table can be found in the CSV report Node_label_combination_count.

	nodeLabels	nodes With That Labels	nodes With That Labels Percent
0	[Git, Change]	81170	78.309359
1	[Git, Commit]	10546	10.174332
2	[File, Git]	5454	5.261787
3	[Git, Tag]	1308	1.261903
4	[Author, Git, Person]	1211	1.168321
5	[Json, Key]	668	0.644458
6	[Json, Value, Scalar]	603	0.581749
7	[Committer, Git, Person]	371	0.357925
8	[NPM, Dependency]	338	0.326088
9	[Type, TS, Primitive]	291	0.280744
10	[Type, TS, Declared]	276	0.266273
11	[TS, ExternalDeclaration]	216	0.208388
12	[Type, TS, Literal]	136	0.131207
13	[Json, Value, Object]	133	0.128313
14	[Type, TS, Union]	119	0.114806
15	[Type, TS, ObjectMember]	101	0.097440
16	[NPM, Script]	91	0.087793
17	[TS, Property]	65	0.062709
18	[TS, Function]	47	0.045344
19	[Type, TS, FunctionParameter]	40	0.038590
20	[Type, Object, TS]	39	0.037626
21	[File, Directory]	34	0.032802
22	[Type, TS, Function]	34	0.032802
23	[TS, Parameter]	33	0.031837
24	[Git, Branch]	32	0.030872
25	[Package, File, Json, NPM]	29	0.027978
26	[TS, ExternalModule]	25	0.024119
27	[TS, Variable]	24	0.023154
28	[Value, TS, Literal]	20	0.019295
29	[jQAssistant, Rule, Concept]	19	0.018330

Chart 1a - Highest node count by label combination

Values under 0.5% will be grouped into "others" to get a cleaner plot. The group "others" is then broken down in Chart 1b.

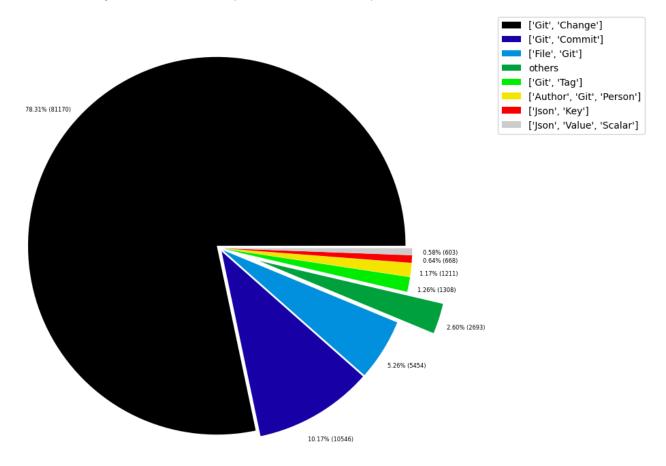


Table 1b - Lowest node count by label combination

Lists the 30 label combinations with the lowest number of nodes until they reach 0.5% of the total node count, which are shown above.

	nodeLabels	nodesWithThatLabels	nodes With That Labels Percent
0	[Analyze, Task, jQAssistant]	1	0.000965
1	[File, TS, Scan]	1	0.000965
2	[TS, Method]	1	0.000965
3	[Repository, File, Git]	1	0.000965
4	[TS, Constructor]	1	0.000965
5	[Value, TS, ObjectMember]	1	0.000965
6	[TS, Class]	1	0.000965
7	[TS, Enum]	2	0.001930
8	[Value, Object, TS]	3	0.002894
9	[Type, TS, Tuple]	3	0.002894
10	[Value, TS, Function]	4	0.003859
11	[TS, TypeParameter]	4	0.003859
12	[Value, TS, Complex]	5	0.004824
13	[NPM, Engine]	6	0.005789
14	[Project, TS]	6	0.005789
15	[File, Local]	6	0.005789
16	[Value, TS, Call]	6	0.005789
17	[Value, TS, Member]	6	0.005789
18	[File, TS, Local, Module]	6	0.005789
19	[Type, TS, TypeParameterReference]	6	0.005789
20	[TS, EnumMember]	8	0.007718
21	[Type, TS, NotIdentified]	11	0.010612
22	[Json, Value, Array]	12	0.011577
23	[Value, TS, Declared]	13	0.012542
24	[TS, TypeAlias]	16	0.015436
25	[File, Directory, Local]	16	0.015436
26	[TS, Interface]	17	0.016401
27	[Type, TS, Intersection]	17	0.016401
28	[jQAssistant, Rule, Concept]	19	0.018330
29	[Value, TS, Literal]	20	0.019295

Chart 1b - Lowest node count by label combination

Shows the lowest (less than 0.5% overall) node count label combinations. Therefore, this plot breaks down the "others" slice of the pie chart above. Values under 0.01% will be grouped into "others" to get a cleaner plot.

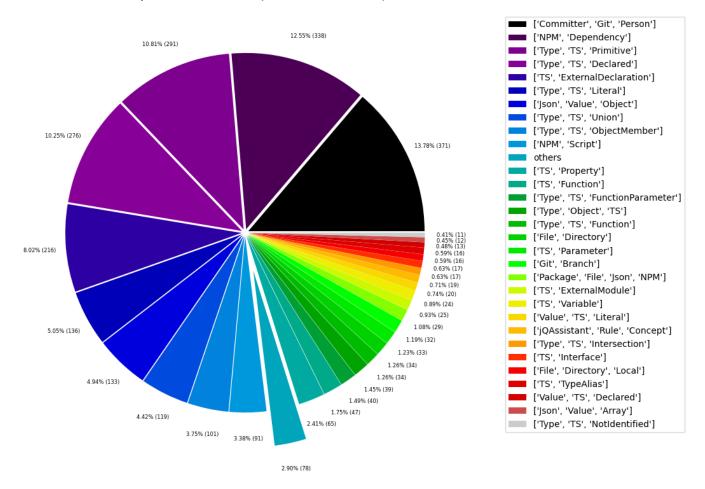


Table 1c - Highest node count by single label

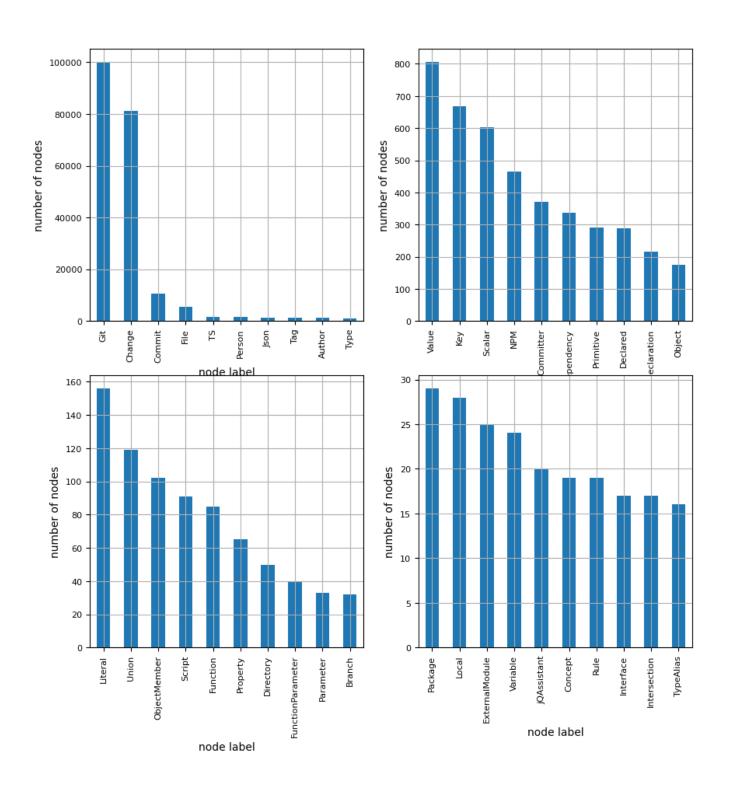
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent	
0	Git	100093	96.565464	
1	Change	81170	78.309359	
2	Commit	10546	10.174332	
3	File	5547	5.351509	
4	TS	1604	1.547471	
5	Person	1582	1.526246	
6	Json	1445	1.394074	
7	Tag	1308	1.261903	
8	Author	1211	1.168321	
9	Туре	1073	1.035185	
10	Value	806	0.777594	
11	Key	668	0.644458	
12	Scalar	603	0.581749	
13	NPM	464	0.447647	
14	Committer	371	0.357925	
15	Dependency	338	0.326088	
16	Primitive	291	0.280744	
17	Declared	289	0.278815	
18	ExternalDeclaration	216	0.208388	
19	Object	175	0.168833	
20	Literal	156	0.150502	
21	Union	119	0.114806	
22	ObjectMember	102	0.098405	
23	Script	91	0.087793	
24	Function	85	0.082004	
25	Property	65	0.062709	
26	Directory	50	0.048238	
27	FunctionParameter	40	0.038590	
28	Parameter	33	0.031837	
29	Branch	32	0.030872	
30	Package	29	0.027978	
31	Local	28	0.027013	
32	ExternalModule	25	0.024119	
33	Variable	24	0.023154	
34	jQAssistant	20	0.019295	
35	Concept	19	0.018330	
36	Rule	19	0.018330	
37	Interface	17	0.016401	
38	Intersection	17	0.016401	
39	TypeAlias	16	0.015436	

Chart 1c - Highest node count by label

Shows the 40 labels with the highest number of nodes.

Node count by label



Relationship Types

Table 2a - Highest relationship count by type

Lists the 30 relationship types with the highest number of occurrences. The whole table can be found in the CSV report Relationship_type_count .

Total number of relationships: 309577

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS CHANGE	81170	26.219648
1	MODIFIES	81170	26.219648
2	UPDATES	53075	17.144361
3	COMMITTED	21092	6.813168
4	CREATES	19585	6.326374
5	DELETES	11683	3.773859
6	HAS_PARENT	11590	3.743818
7	HAS COMMIT	10546	3.406584
8	HAS_FILE	5454	1.761759
9	RENAMES	3173	1.024947
10	HAS NEW NAME	1718	0.554951
11	HAS TAG	1308	0.422512
12	ON COMMIT	1308	0.422512
13	HAS AUTHOR	1211	0.391179
14	DEPENDS_ON	971	0.313654
15	HAS_KEY	668	0.215778
16	HAS_VALUE	668	0.215778
17	CONTAINS	594	0.191875
18	HAS_COMMITTER	371	0.119841
19	OF_TYPE	337	0.108858
20	EXPORTS	277	0.089477
21	REFERENCES	197	0.063635
22	DECLARES	186	0.060082
23	DECLARES_DEV_DEPENDENCY	169	0.054591
24	DECLARES_DEPENDENCY	161	0.052006
25	HAS_MEMBER	102	0.032948
26	HAS_TYPE_ARGUMENT	94	0.030364
27	DECLARES_SCRIPT	91	0.029395
28	RETURNS	82	0.026488
29	HAS_PARAMETER	73	0.023581

Chart 2a - Highest relationship count by type

Values under 0.5% will be grouped into "others" to get a cleaner plot. The group "others" is then broken down in the second chart.

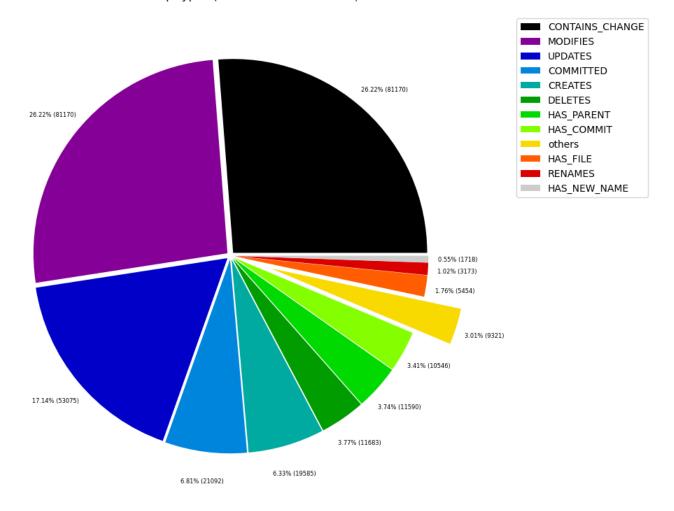


Table 2b - Lowest relationship count by type

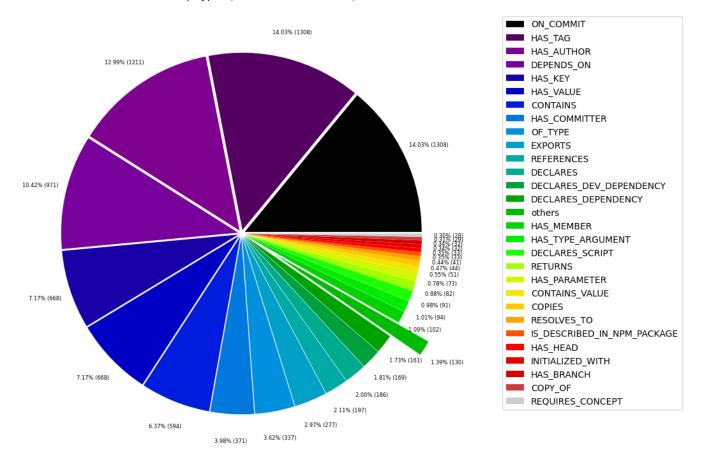
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent	
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000323	
1	IS_IMPLEMENTED_IN	2	0.000646	
2	CONSTRAINED_BY	4	0.001292	
3	REFERENCED_PROJECTS	REFERENCED_PROJECTS 5	0.001615	
4	CONTAINS_PROJECT	6	0.001938	
5	DECLARES_ENGINE	6	0.001938	
6	EXTENDS	6	0.001938	
7	HAS_ARGUMENT	6	0.001938	
8	CALLS	6	0.001938	
9	HAS_NPM_PACKAGE	6	0.001938	
10	HAS_ROOT	6	0.001938	
11	MEMBER	6	0.001938	
12	PARENT	6	0.001938	
13	HAS_CONFIG	6	0.001938	
14	SIMILAR	6	0.001938	
15	DECLARES_PEER_DEPENDENCY	8	0.002584	
16	INCLUDES_CONCEPT	19	0.006137	
17	USES	25	0.008076	
18	REQUIRES_CONCEPT	28	0.009045	
19	COPY_OF	29	0.009368	
20	INITIALIZED_WITH	32	0.010337	
21	HAS_BRANCH	32	0.010337	
22	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010660	
23	HAS_HEAD	33	0.010660	
24	RESOLVES_TO	41	0.013244	
25	COPIES	44	0.014213	
26	CONTAINS_VALUE	51	0.016474	
27	HAS_PARAMETER	73	0.023581	
28	RETURNS	82	0.026488	
29	DECLARES_SCRIPT	91	0.029395	

Chart 2b - Lowest relationship count by type

Shows the lowest (less than 0.5% overall) relationship types. This plot breaks down the "others" slice of the pie chart above. Values under 0.01% will be grouped into "others" to get a cleaner plot.

Relationship types (less than 0.5% overall)



Node labels with their relationships

Table 3a - Highest relationship count by node labels and relationship type

Lists the 30 node labels and their relationship types with the highest number of occurrences.

	sourceLabels	relationType	targetLabels	number Of Relationships	number Of Nodes With Same Labels As Source	numberOfNodes
0	[Git, Change]	MODIFIES	[File, Git]	81170	81170	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	81170	10546	
2	[Git, Change]	UPDATES	[File, Git]	53075	81170	
3	[Git, Change]	CREATES	[File, Git]	19585	81170	
4	[Git, Change]	DELETES	[File, Git]	11683	81170	
5	[Git, Commit]	HAS_PARENT	[Git, Commit]	11590	10546	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10546	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10546	1211	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10546	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5454	1	
10	[Git, Change]	RENAMES	[File, Git]	3173	81170	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1718	5454	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1308	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1308	1308	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1211	1	
15	[Json, Value, Object]	HAS_KEY	[Json, Key]	668	133	
16	[Json, Key]	HAS_VALUE	[Json, Value, Scalar]	552	668	
17	[Repository, File, Git]	HAS_COMMITTER	[Committer, Git, Person]	371	1	
18	[TS, Function]	DEPENDS_ON	[TS, ExternalDeclaration]	292	47	
19	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	216	25	
20	[File, TS, Local, Module, Mark4ModuleWeaklyCon	DEPENDS_ON	[TS, ExternalDeclaration]	193	2	
21	[Package, File, Json, NPM]	DECLARES_DEV_DEPENDENCY	[NPM, Dependency]	169	29	
22	[Package, File, Json, NPM]	DECLARES_DEPENDENCY	[NPM, Dependency]	161	29	
23	[Type, TS, Union]	CONTAINS	[Type, TS, Primitive]	147	119	
24	[Type, TS, Declared]	REFERENCES	[TS, ExternalDeclaration]	142	276	
25	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	135	47	
26	[Type, TS, Union]	CONTAINS	[Type, TS, Literal]	119	119	
27	[Json, Key]	HAS_VALUE	[Json, Value, Object]	104	668	
28	[Type, Object, TS]	HAS_MEMBER	[Type, TS, ObjectMember]	101	39	
29	[Package, File, Json, NPM]	DECLARES_SCRIPT	[NPM, Script]	91	29	

Graph Density

total_number_of_nodes (vertices): 103653
total_number_of_relationships (edges): 309577

-> total directed graph density: 2.881436974269316e-05

-> total directed graph density in percent: 0.002881436974269316