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]	81536	78.300618
1	[Git, Commit]	10605	10.184189
2	[File, Git]	5495	5.276956
3	[Git, Tag]	1314	1.261860
4	[Author, Git, Person]	1219	1.170630
5	[Json, Key]	668	0.641493
6	[Json, Value, Scalar]	603	0.579073
7	[Committer, Git, Person]	371	0.356279
8	[NPM, Dependency]	338	0.324588
9	[Type, TS, Primitive]	291	0.279453
10	[Type, TS, Declared]	276	0.265048
11	[TS, ExternalDeclaration]	214	0.205508
12	[Type, TS, Literal]	136	0.130603
13	[Json, Value, Object]	133	0.127723
14	[Type, TS, Union]	119	0.114278
15	[Type, TS, ObjectMember]	101	0.096992
16	[NPM, Script]	91	0.087389
17	[TS, Property]	65	0.062421
18	[TS, Function]	47	0.045135
19	[Type, TS, FunctionParameter]	40	0.038413
20	[Type, Object, TS]	39	0.037452
21	[Git, Branch]	35	0.033611
22	[File, Directory]	34	0.032651
23	[Type, TS, Function]	34	0.032651
24	[TS, Parameter]	33	0.031691
25	[Package, File, Json, NPM]	29	0.027849
26	[TS, Variable]	24	0.023048
27	[TS, ExternalModule]	23	0.022087
28	[Value, TS, Literal]	20	0.019206
29	[jQAssistant, Rule, Concept]	19	0.018246

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.

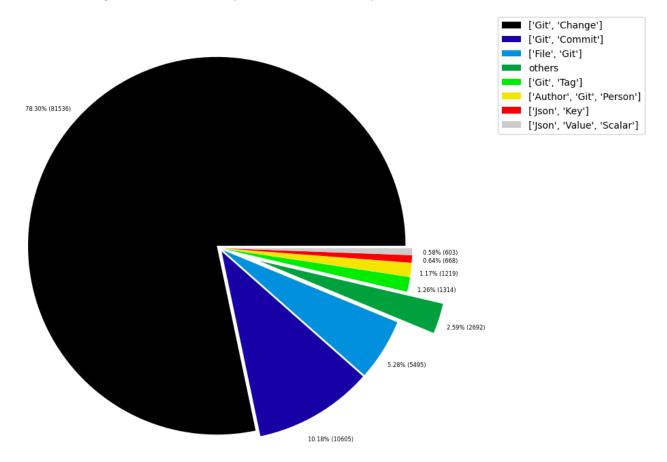


Table 1b - Lowest node count by label combination

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	nodesWithThatLabelsPercent
0	[Analyze, Task, jQAssistant]	1	0.000960
1	[File, TS, Scan]	1	0.000960
2	[TS, Method]	1	0.000960
3	[Repository, File, Git]	1	0.000960
4	[TS, Constructor]	1	0.000960
5	[Value, TS, ObjectMember]	1	0.000960
6	[TS, Class]	1	0.000960
7	[TS, Enum]	2	0.001921
8	[Value, Object, TS]	3	0.002881
9	[Type, TS, Tuple]	3	0.002881
10	[Value, TS, Function]	4	0.003841
11	[TS, TypeParameter]	4	0.003841
12	[Value, TS, Complex]	5	0.004802
13	[NPM, Engine]	6	0.005762
14	[Project, TS]	6	0.005762
15	[File, Local]	6	0.005762
16	[Value, TS, Call]	6	0.005762
17	[Value, TS, Member]	6	0.005762
18	[File, TS, Local, Module]	6	0.005762
19	[Type, TS, TypeParameterReference]	6	0.005762
20	[TS, EnumMember]	8	0.007683
21	[Type, TS, NotIdentified]	11	0.010564
22	[Json, Value, Array]	12	0.011524
23	[Value, TS, Declared]	13	0.012484
24	[TS, TypeAlias]	16	0.015365
25	[File, Directory, Local]	16	0.015365
26	[TS, Interface]	17	0.016325
27	[Type, TS, Intersection]	17	0.016325
28	[jQAssistant, Rule, Concept]	19	0.018246
29	[Value, TS, Literal]	20	0.019206

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.

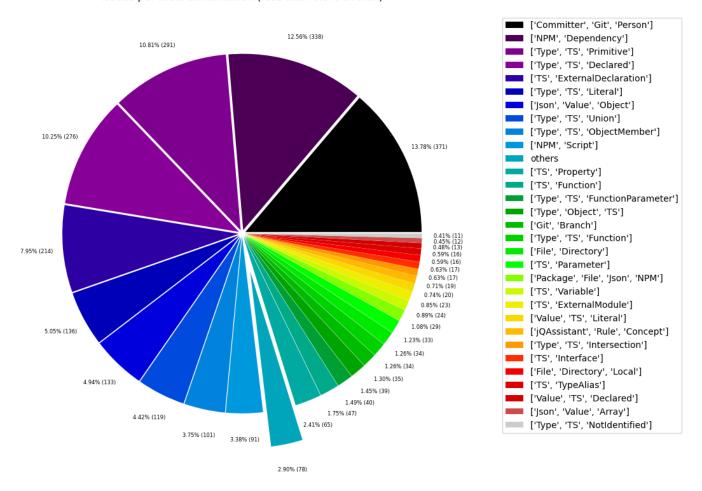


Table 1c - Highest node count by single label

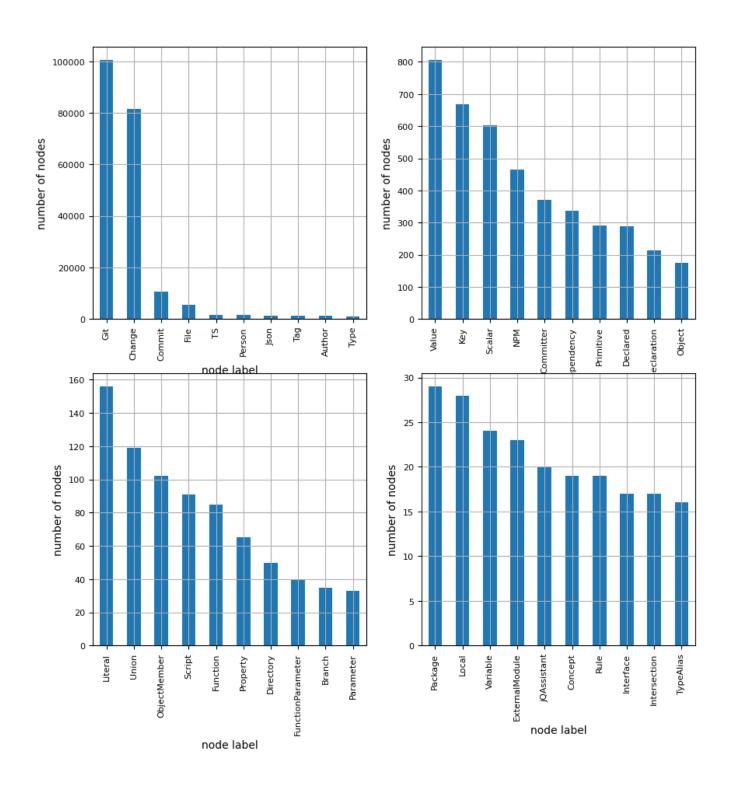
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodes With That Label	nodesWithThatLabelPercent	
0	Git	100576	96.585104	
1	Change	81536	78.300618	
2	Commit	10605	10.184189	
3	File	5588	5.366266	
4	TS	1600	1.536511	
5	Person	1590	1.526908	
6	Json	1445	1.387662	
7	Tag	1314	1.261860	
8	Author	1219	1.170630	
9	Туре	1073	1.030423	
10	Value	806	0.774018	
11	Key	668	0.641493	
12	Scalar	603	0.579073	
13	NPM	464	0.445588	
14	Committer	371	0.356279	
15	Dependency	338	0.324588	
16	Primitive	291	0.279453	
17	Declared	289	0.277532	
18	ExternalDeclaration	214	0.205508	
19	Object	175	0.168056	
20	Literal	156	0.149810	
21	Union	119	0.114278	
22	ObjectMember	102	0.097953	
23	Script	91	0.087389	
24	Function	85	0.081627	
25	Property	65	0.062421	
26	Directory	50	0.048016	
27	FunctionParameter	40	0.038413	
28	Branch	35	0.033611	
29	Parameter	33	0.031691	
30	Package	29	0.027849	
31	Local	28	0.026889	
32	Variable	24	0.023048	
33	ExternalModule	23	0.022087	
34	jQAssistant	20	0.019206	
35	Concept	19	0.018246	
36	Rule	19	0.018246	
37	Interface	17	0.016325	
38	Intersection	17	0.016325	
39	TypeAlias	16	0.015365	

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: 311060

		•	
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	81536	26.212306
1	MODIFIES	81536	26.212306
2	UPDATES	53348	17.150389
3	COMMITTED	21210	6.818620
4	CREATES	19662	6.320967
5	DELETES	11723	3.768726
6	HAS_PARENT	11651	3.745580
7	HAS_COMMIT	10605	3.409310
8	HAS_FILE	5495	1.766540
9	RENAMES	3197	1.027776
10	HAS_NEW_NAME	1726	0.554877
11	HAS_TAG	1314	0.422427
12	ON_COMMIT	1314	0.422427
13	HAS_AUTHOR	1219	0.391886
14	DEPENDS_ON	961	0.308944
15	HAS_KEY	668	0.214750
16	HAS_VALUE	668	0.214750
17	CONTAINS	594	0.190960
18	HAS_COMMITTER	371	0.119270
19	OF_TYPE	337	0.108339
20	EXPORTS	283	0.090979
21	REFERENCES	197	0.063332
22	DECLARES	186	0.059796
23	DECLARES_DEV_DEPENDENCY	169	0.054330
24	DECLARES_DEPENDENCY	161	0.051759
25	HAS_MEMBER	102	0.032791
26	HAS_TYPE_ARGUMENT	94	0.030219
27	DECLARES_SCRIPT	91	0.029255
28	RETURNS	82	0.026361
29	HAS_PARAMETER	73	0.023468

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.

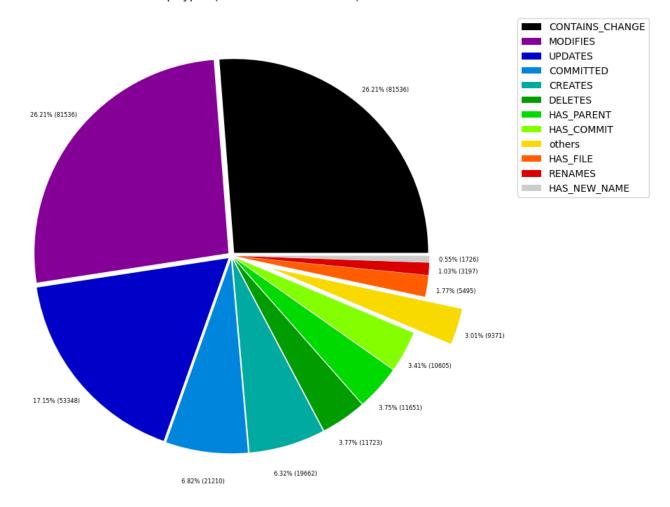


Table 2b - Lowest relationship count by type

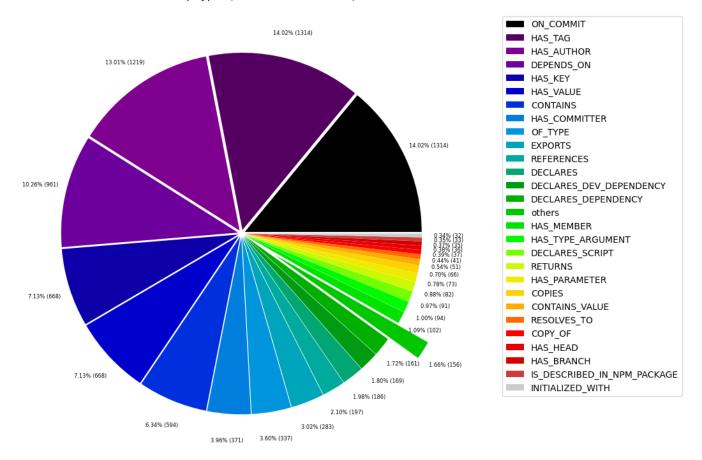
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent	
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000321	
1	IS_IMPLEMENTED_IN	2	0.000643	
2	CONSTRAINED_BY	4	0.001286	
3	REFERENCED_PROJECTS	5	0.001607	
4	CONTAINS_PROJECT	6	0.001929	
5	DECLARES_ENGINE	6	0.001929	
6	EXTENDS	6	0.001929	
7	HAS_ARGUMENT	6	0.001929	
8	CALLS	6	0.001929	
9	HAS_NPM_PACKAGE	6	0.001929	
10	HAS_ROOT	6	0.001929	
11	MEMBER	6	0.001929	
12	PARENT	6	0.001929	
13	HAS_CONFIG	6	0.001929	
14	SIMILAR	6	0.001929	
15	DECLARES_PEER_DEPENDENCY	8	0.002572	
16	INCLUDES_CONCEPT	19	0.006108	
17	USES	23	0.007394	
18	REQUIRES_CONCEPT	28	0.009001	
19	INITIALIZED_WITH	32	0.010287	
20	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010609	
21	HAS_BRANCH	35	0.011252	
22	HAS_HEAD	36	0.011573	
23	COPY_OF	37	0.011895	
24	RESOLVES_TO	41	0.013181	
25	CONTAINS_VALUE	51	0.016396	
26	COPIES	66	0.021218	
27	HAS_PARAMETER	73	0.023468	
28	RETURNS	82	0.026361	
29	DECLARES_SCRIPT	91	0.029255	

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]	81536	81536	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	81536	10605	
2	[Git, Change]	UPDATES	[File, Git]	53348	81536	
3	[Git, Change]	CREATES	[File, Git]	19662	81536	
4	[Git, Change]	DELETES	[File, Git]	11723	81536	
5	[Git, Commit]	HAS_PARENT	[Git, Commit]	11651	10605	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10605	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10605	1219	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10605	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5495	1	
10	[Git, Change]	RENAMES	[File, Git]	3197	81536	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1726	5495	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1314	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1314	1314	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1219	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]	289	47	
19	[File, TS, Local, Module, Mark4ModuleWeaklyCon	DEPENDS_ON	[TS, ExternalDeclaration]	232	4	
20	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	214	23	
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]	132	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): 104132
total_number_of_relationships (edges): 311060

-> total directed graph density: 2.8686655560969712e-05

-> total directed graph density in percent: 0.0028686655560969713