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	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	255467	74.138016
1	[Java, ByteCode, Member, Method]	13198	3.830137
2	[Java, ByteCode, Parameter]	13029	3.781092
3	[Git, Commit]	12364	3.588105
4	[File, Git]	10767	3.124646
5	[Java, ByteCode, Bound, ParameterizedType]	7176	2.082517
6	[Java, ByteCode, Bound]	7118	2.065685
7	[Java, ByteCode, Member, Field]	3505	1.017171
8	[Java, ByteCode, Bound, WildcardType]	2908	0.843919
9	[Java, Value, ByteCode, Annotation]	2822	0.818961
10	[Xml, Element]	2144	0.622201
11	[Java, ByteCode, Member, Method, Constructor]	2049	0.594632
12	[Xml, Text]	1436	0.416736
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.318066
14	[Java, ByteCode, Member, Method, Lambda]	942	0.273374
15	[Type, File, Java, ByteCode, ResolvedDuplicate	899	0.260895
16	[Type, File, Java, ByteCode, Class]	864	0.250738
17	[Type, File, Java, ByteCode, JavaType]	748	0.217074
18	[Java, Value, ByteCode, Primitive]	652	0.189214
19	$[{\tt Java,ByteCode,GenericDeclaration,Member,M}$	569	0.165127
20	[Json, Key]	560	0.162515
21	[Value, Json, Scalar]	544	0.157872
22	[Type, File, Java, ByteCode, ExternalType]	396	0.114922
23	[Author, Git, Person]	298	0.086481
24	[Value, Array]	267	0.077485
25	[Committer, Git, Person]	252	0.073132
26	[Type, File, Java, ByteCode, Class, GenericDec	236	0.068489
27	[Value, Property]	202	0.058622
28	[Type, File, Java, ByteCode, Interface]	188	0.054559
29	[Java, Value, ByteCode, Class]	172	0.049915

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.

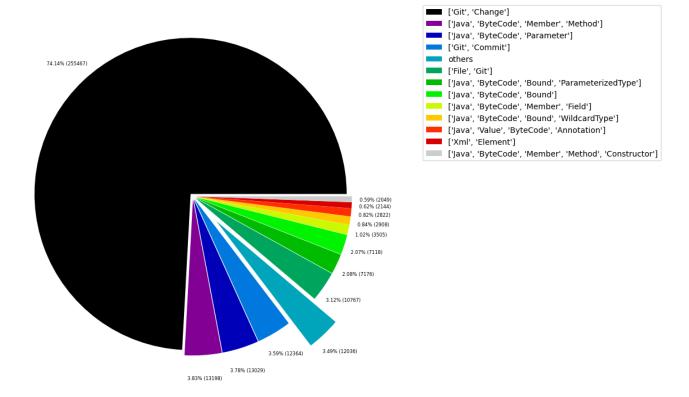


Table 1b - Lowest node count by label combination

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	0.000290
1	[Repository, File, Git]	1	0.000290
2	[File, Json]	2	0.000580
3	[File]	3	0.000871
4	[Java,ByteCode,GenericDeclaration,Member,M	4	0.001161
5	[Maven, Exclusion]	5	0.001451
6	[Value, Array, Json]	6	0.001741
7	[Type, File, Java, ByteCode, Void]	9	0.002612
8	[File, Maven, Xml, Pom, Document]	9	0.002612
9	[Java, ManifestSection]	9	0.002612
10	[File, Java, Manifest]	9	0.002612
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002612
12	[File, Java, ServiceLoader]	10	0.002902
13	[File, Java, Properties]	12	0.003482
14	[Maven, ExecutionGoal]	16	0.004643
15	[Maven, PluginExecution]	16	0.004643
16	[Xml, Attribute]	18	0.005224
17	[jQAssistant, Rule, Concept]	19	0.005514
18	[Maven, Plugin]	21	0.006094
19	[Maven, Configuration]	21	0.006094
20	[Type, File, Java, ByteCode, Enum]	28	0.008126
21	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008416
22	[Git, Branch]	33	0.009577
23	[Xml, Namespace]	36	0.010447
24	[Type, File, Java, ByteCode, Annotation]	43	0.012479
25	[File, Directory]	44	0.012769
26	[Type, File, Java, ByteCode, GenericDeclaratio	85	0.024667
27	[Artifact, Maven]	92	0.026699
28	[Type, File, Java, ByteCode, ExternalAnnotatio	93	0.026989
29	[Value, ManifestEntry]	120	0.034825

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.

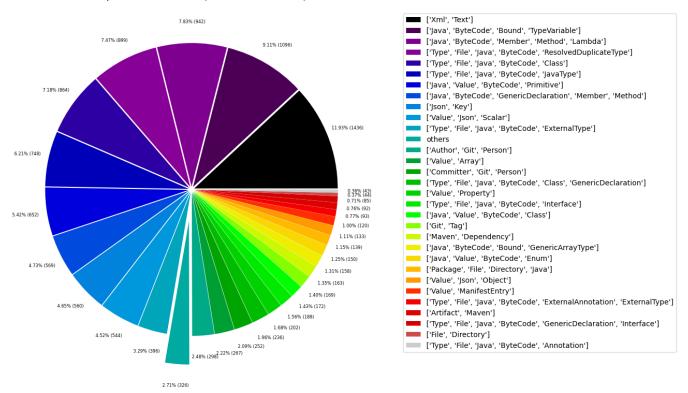


Table 1c - Highest node count by single label

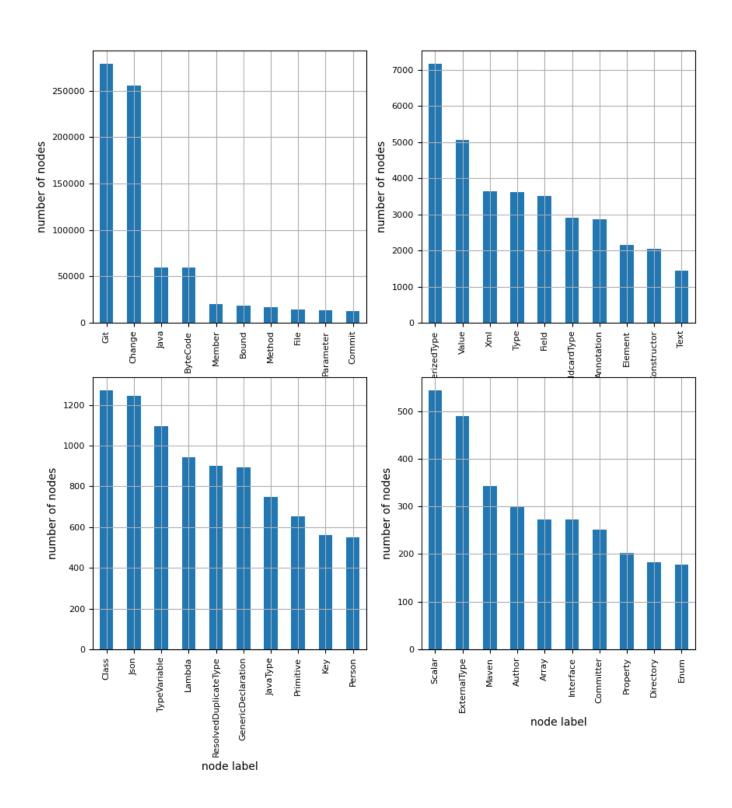
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	nodes With That Label	nodesWithThatLabelPercent	
0	Git	279351	81.069292	
1	Change	255467	74.138016	
2	Java	59354	17.224878	
3	ByteCode	59166	17.170319	
4	Member	20267	5.881602	
5	Bound	18456	5.356039	
6	Method	16762	4.864430	
7	File	14623	4.243680	
8	Parameter	13029	3.781092	
9	Commit	12364	3.588105	
10	ParameterizedType	7176	2.082517	
11	Value	5068	1.470763	
12	Xml	3643	1.057220	
13	Туре	3618	1.049965	
14	Field	3505	1.017171	
15	WildcardType	2908	0.843919	
16	Annotation	2865	0.831440	
17	Element	2144	0.622201	
18	Constructor	2053	0.595793	
19	Text	1436	0.416736	
20	Class	1272	0.369142	
21	Json	1245	0.361306	
22	TypeVariable	1096	0.318066	
23	Lambda	942	0.273374	
24	ResolvedDuplicateType	899	0.260895	
25	GenericDeclaration	894	0.259444	
26	JavaType	748	0.217074	
27	Primitive	652	0.189214	
28	Key	560	0.162515	
29	Person	550	0.159613	
30	Scalar	544	0.157872	
31	ExternalType	489	0.141911	
32	Maven	343	0.099541	
33	Author	298	0.086481	
34	Array	273	0.079226	
35	Interface	273	0.079226	
36	Committer	252	0.073132	
37	Property	202	0.058622	
38	Directory	183	0.053108	
39	Enum	178	0.051657	

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: 1069381

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	255467	23.889241
1	MODIFIES	255467	23.889241
2	UPDATES	167307	15.645219
3	CREATES	60811	5.686561
4	DELETES	37841	3.538589
5	INVOKES	35684	3.336884
6	COMMITTED	24728	2.312366
7	DEPENDS_ON	21931	2.050813
8	OF_TYPE	21394	2.000597
9	DECLARES	20712	1.936821
10	OF_RAW_TYPE	17093	1.598401
11	HAS_PARENT	15025	1.405018
12	HAS	14125	1.320858
13	RETURNS	12578	1.176194
14	HAS_COMMIT	12364	1.156183
15	HAS_FILE	10767	1.006844
16	RENAMES	10492	0.981128
17	READS	9164	0.856944
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.775028
19	HAS_NEW_NAME	6219	0.581551
20	OF_GENERIC_TYPE	5906	0.552282
21	RESOLVES_TO	5184	0.484766
22	SIMILAR	3977	0.371897
23	WRITES	3818	0.357029
24	CONTAINS	3814	0.356655
25	RETURNS_GENERIC	3545	0.331500
26	ANNOTATED_BY	2810	0.262769
27	REQUIRES	2174	0.203295
28	HAS_FIRST_CHILD	2144	0.200490
29	HAS_LAST_CHILD	2144	0.200490

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.

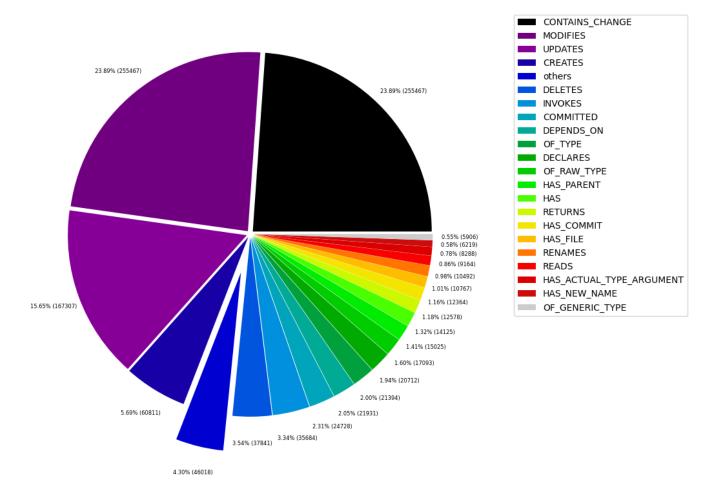


Table 2b - Lowest relationship count by type

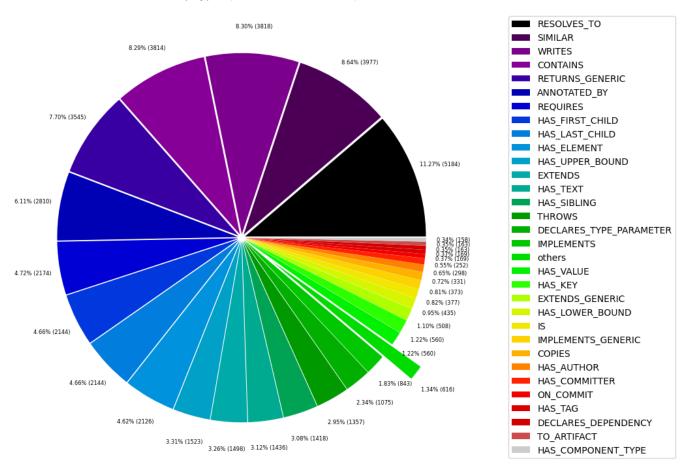
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodes With That Relationship Type	nodesWithThatRelationshipTypePercent	
0	HAS_PROPERTY	1	0.000094	
1	THROWS_GENERIC	5	0.000468	
2	EXCLUDES	5	0.000468	
3	DESCRIBES	9	0.000842	
4	HAS_ROOT_ELEMENT	11	0.001029	
5	HAS_GOAL	16	0.001496	
6	HAS_EXECUTION	16	0.001496	
7	OF_NAMESPACE	18	0.001683	
8	HAS_ATTRIBUTE	18	0.001683	
9	INCLUDES_CONCEPT	19	0.001777	
10	USES_PLUGIN	21	0.001964	
11	REQUIRES_TYPE_PARAMETER	21	0.001964	
12	IS_ARTIFACT	21	0.001964	
13	HAS_CONFIGURATION	21	0.001964	
14	REQUIRES_CONCEPT	28	0.002618	
15	HAS_BRANCH	33	0.003086	
16	HAS_HEAD	34	0.003179	
17	DECLARES_NAMESPACE	36	0.003366	
18	HAS_DEFAULT	36	0.003366	
19	CONTAINS_VALUE	121	0.011315	
20	COPY_OF	126	0.011783	
21	HAS_COMPONENT_TYPE	158	0.014775	
22	TO_ARTIFACT	163	0.015242	
23	DECLARES_DEPENDENCY	163	0.015242	
24	ON_COMMIT	169	0.015804	
25	HAS_TAG	169	0.015804	
26	HAS_COMMITTER	252	0.023565	
27	HAS_AUTHOR	298	0.027867	
28	COPIES	331	0.030952	
29	IMPLEMENTS_GENERIC	373	0.034880	

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	numberOfRelationships	number Of Nodes With Same Labels As Source	numberOfNodesWithS
0	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	255467	12364	
1	[Git, Change]	MODIFIES	[File, Git]	255467	255467	
2	[Git, Change]	UPDATES	[File, Git]	167307	255467	
3	[Git, Change]	CREATES	[File, Git]	60811	255467	
4	[Git, Change]	DELETES	[File, Git]	37841	255467	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21758	13104	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15016	12364	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12364	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12364	298	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12364	252	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10767	1	
11	[Git, Change]	RENAMES	[File, Git]	10492	255467	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8361	13104	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8218	13104	
14	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6371	13029	
15	[File, Git]	HAS_NEW_NAME	[File, Git]	6219	10767	
16	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Method]	5103	814	
17	[Type, File, Java, ByteCode, Class]	DEPENDS_ON	[Type, File, Java, ByteCode, JavaType]	4480	814	
18	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3570	7118	
19	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3053	7176	
20	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2908	7176	
21	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2657	13029	
22	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7176	
23	[Java, ByteCode, Member, Method, Constructor]	WRITES	[Java, ByteCode, Member, Field]	2438	2049	
24	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio	2341	2822	
25	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Field]	2203	814	
26	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
27	[Java, ByteCode, Member, Method, Constructor]	INVOKES	[Java, ByteCode, Member, Method, Constructor]	2117	2049	
28	[Type, File, Java, ByteCode, Class, GenericDec	DECLARES	[Java, ByteCode, Member, Method]	2114	231	
29	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2092	13104	

Graph Density

total_number_of_nodes (vertices): 344583

total_number_of_relationships (edges): 1069381

- -> total directed graph density: 9.006292104576975e-06
- -> total directed graph density in percent: 0.0009006292104576975