### 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]	261796	74.462288
1	[Java, ByteCode, Member, Method]	13197	3.753605
2	[Java, ByteCode, Parameter]	13029	3.705821
3	[Git, Commit]	12886	3.665148
4	[File, Git]	10887	3.096575
5	[Java, ByteCode, Bound, ParameterizedType]	7176	2.041060
6	[Java, ByteCode, Bound]	7118	2.024563
7	[Java, ByteCode, Member, Field]	3502	0.996069
8	[Java, ByteCode, Bound, WildcardType]	2908	0.827119
9	[Java, Value, ByteCode, Annotation]	2822	0.802658
10	[Xml, Element]	2144	0.609815
11	[Java, ByteCode, Member, Method, Constructor]	2039	0.579950
12	[Xml, Text]	1436	0.408440
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.311734
14	[Java, ByteCode, Member, Method, Lambda]	942	0.267932
15	[Type, File, Java, ByteCode, ResolvedDuplicate	877	0.249444
16	[Type, File, Java, ByteCode, Class]	810	0.230387
17	[Java, Value, ByteCode, Primitive]	652	0.185447
18	[Type, File, Java, ByteCode, JavaType]	644	0.183172
19	$[{\sf Java,ByteCode,Member,Method,GenericDeclar}$	569	0.161840
20	[Json, Key]	568	0.161555
21	[Value, Json, Scalar]	551	0.156720
22	[Type, File, Java, ByteCode, ExternalType]	380	0.108083
23	[Author, Git, Person]	298	0.084760
24	[Value, Array]	267	0.075942
25	[Committer, Git, Person]	252	0.071676
26	[Type, File, Java, ByteCode, GenericDeclaratio	236	0.067125
27	[Value, Property]	202	0.057455
28	[Type, File, Java, ByteCode, Interface]	188	0.053473
29	[Java, Value, ByteCode, Class]	172	0.048922

# 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.

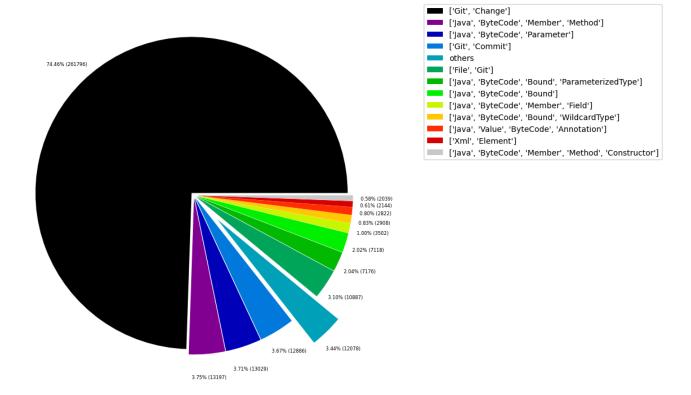


Table 1b - Lowest node count by label combination

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	0.000284
1	[Package, File, Json, NPM]	1	0.000284
2	[Repository, File, Git]	1	0.000284
3	[File, TS, Scan]	1	0.000284
4	[File, Json]	2	0.000569
5	[File]	3	0.000853
6	[Java, ByteCode, Member, Method, GenericDeclar	4	0.001138
7	[Maven, Exclusion]	5	0.001422
8	[Value, Array, Json]	6	0.001707
9	[Dependency, NPM]	7	0.001991
10	[Type, File, Java, ByteCode, Void]	9	0.002560
11	[File, Maven, Xml, Pom, Document]	9	0.002560
12	[Java, ManifestSection]	9	0.002560
13	[File, Java, Manifest]	9	0.002560
14	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002560
15	[File, Java, ServiceLoader]	10	0.002844
16	[File, Java, Properties]	12	0.003413
17	[Maven, PluginExecution]	16	0.004551
18	[Maven, ExecutionGoal]	16	0.004551
19	[Type, File, Java, ByteCode, Throwable, Extern	16	0.004551
20	[Xml, Attribute]	18	0.005120
21	[jQAssistant, Rule, Concept]	19	0.005404
22	[Maven, Configuration]	21	0.005973
23	[Maven, Plugin]	21	0.005973
24	[Type, File, Java, ByteCode, Throwable, Resolv	22	0.006257
25	[Type, File, Java, ByteCode, Enum]	28	0.007964
26	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008248
27	[Xml, Namespace]	36	0.010239
28	[Type, File, Java, ByteCode, Annotation]	43	0.012230
29	[Git, Branch]	44	0.012515

#### 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.

nodel abels nodesWithThatLabels nodesWithThatLabelsPercent

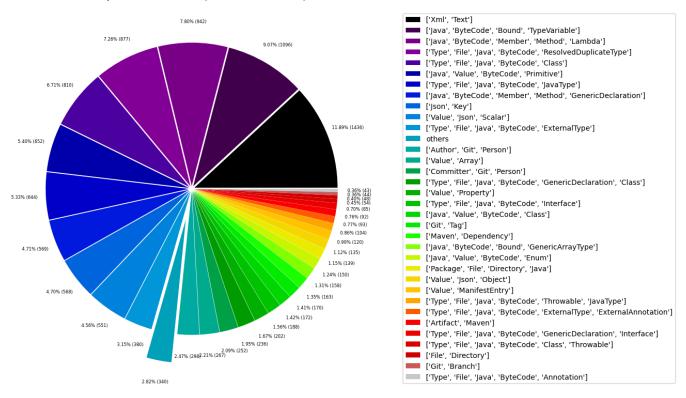


Table 1c - Highest node count by single label

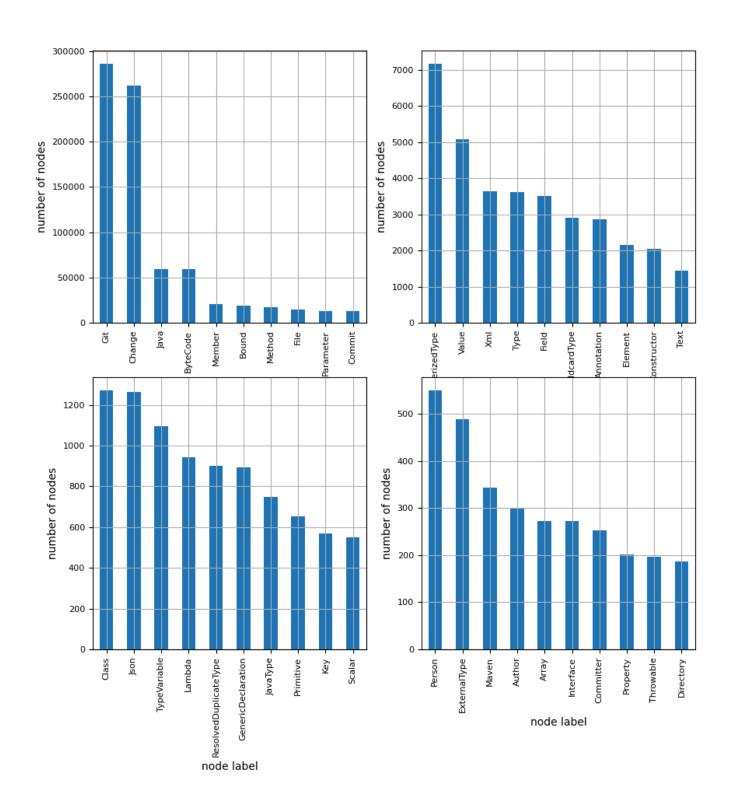
Lists the 40 labels with the highest number of nodes. Doesn't sum up to the total number of nodes or 100% because one node can have multiple labels. Helps to identify commonly used labels.

	nodeLabel	${\bf nodesWithThatLabel}$	${\bf nodes With That Label Percent}$
0	Git	286334	81.441598
1	Change	261796	74.462288
2	Java	59340	16.877997
3	ByteCode	59152	16.824525
4	Member	20253	5.760534
5	Bound	18456	5.249415
6	Method	16751	4.764465
7	File	14749	4.195038
8	Parameter	13029	3.705821
9	Commit	12886	3.665148
10	ParameterizedType	7176	2.041060
11	Value	5077	1.444044
12	Xml	3643	1.036174
13	Туре	3618	1.029063
14	Field	3502	0.996069
15	WildcardType	2908	0.827119
16	Annotation	2865	0.814888
17	Element	2144	0.609815
18	Constructor	2043	0.581088
19	Text	1436	0.408440
20	Class	1272	0.361793
21	Json	1263	0.359233
22	TypeVariable	1096	0.311734
23	Lambda	942	0.267932
24	ResolvedDuplicateType	899	0.255701
25	GenericDeclaration	894	0.254279
26	JavaType	748	0.212753
27	Primitive	652	0.185447
28	Key	568	0.161555
29	Scalar	551	0.156720
30	Person	550	0.156436
31	ExternalType	489	0.139086
32	Maven	343	0.097559
33	Author	298	0.084760
34	Array	273	0.077649
35	Interface	273	0.077649
36	Committer	252	0.071676
37	Property	202	0.057455
38	Throwable	196	0.055748
39	Directory	187	0.053188

# 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: 1091605

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	261796	23.982668
1	MODIFIES	261796	23.982668
2	UPDATES	172614	15.812863
3	CREATES	61715	5.653602
4	DELETES	38029	3.483769
5	INVOKES	35947	3.293041
6	COMMITTED	25772	2.360927
7	DEPENDS_ON	21931	2.009060
8	OF_TYPE	21394	1.959866
9	DECLARES	20698	1.896107
10	OF_RAW_TYPE	17093	1.565859
11	HAS_PARENT	15598	1.428905
12	HAS	14125	1.293966
13	HAS_COMMIT	12886	1.180464
14	RETURNS	12578	1.152248
15	HAS_FILE	10887	0.997339
16	RENAMES	10562	0.967566
17	READS	9161	0.839223
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.759249
19	HAS_NEW_NAME	6240	0.571635
20	OF_GENERIC_TYPE	5906	0.541038
21	RESOLVES_TO	5216	0.477829
22	SIMILAR	3977	0.364326
23	CONTAINS	3824	0.350310
24	WRITES	3818	0.349760
25	RETURNS_GENERIC	3545	0.324751
26	ANNOTATED_BY	2810	0.257419
27	REQUIRES	2174	0.199156
28	HAS_FIRST_CHILD	2144	0.196408
29	HAS_LAST_CHILD	2144	0.196408

#### 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.

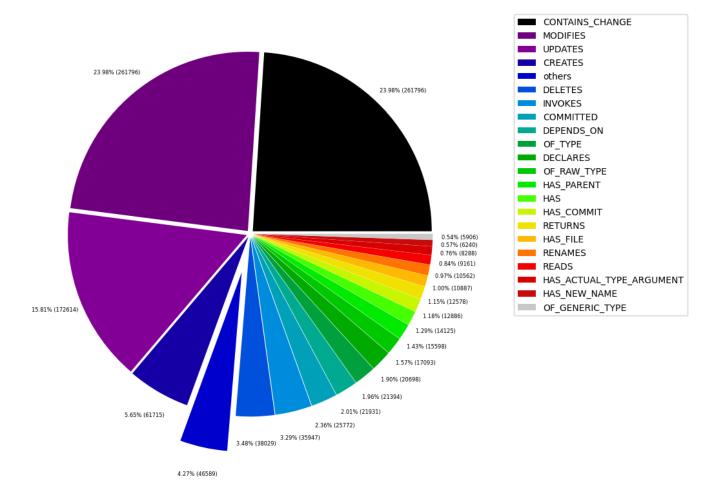


Table 2b - Lowest relationship count by type

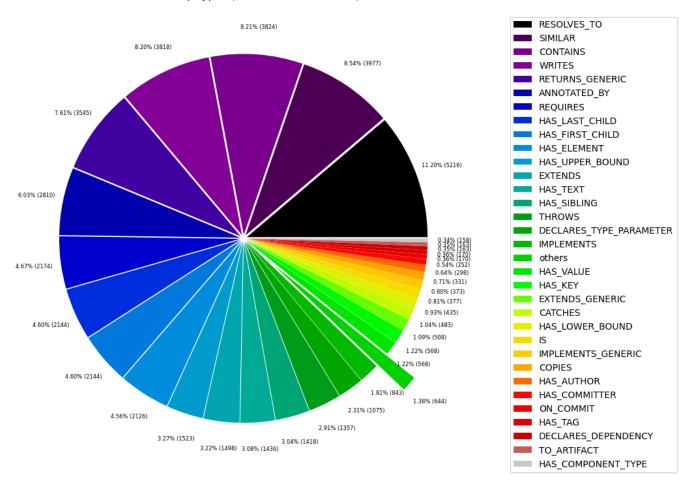
Lists the 30 relationships type with the lowest number of occurrences up to 0.5% of the total node count. This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodes With That Relationship Type	nodes With That Relationship Type Percent
0	HAS_PROPERTY	1	0.000092
1	THROWS_GENERIC	5	0.000458
2	EXCLUDES	5	0.000458
3	DECLARES_DEV_DEPENDENCY	7	0.000641
4	DESCRIBES	9	0.000824
5	HAS_ROOT_ELEMENT	10	0.000916
6	HAS_GOAL	16	0.001466
7	HAS_EXECUTION	16	0.001466
8	OF_NAMESPACE	18	0.001649
9	HAS_ATTRIBUTE	18	0.001649
10	INCLUDES_CONCEPT	19	0.001741
11	REQUIRES_TYPE_PARAMETER	21	0.001924
12	IS_ARTIFACT	21	0.001924
13	HAS_CONFIGURATION	21	0.001924
14	USES_PLUGIN	21	0.001924
15	REQUIRES_CONCEPT	28	0.002565
16	HAS_DEFAULT	36	0.003298
17	DECLARES_NAMESPACE	36	0.003298
18	HAS_BRANCH	44	0.004031
19	HAS_HEAD	45	0.004122
20	CONTAINS_VALUE	121	0.011085
21	COPY_OF	126	0.011543
22	HAS_COMPONENT_TYPE	158	0.014474
23	TO_ARTIFACT	163	0.014932
24	DECLARES_DEPENDENCY	163	0.014932
25	ON_COMMIT	170	0.015573
26	HAS_TAG	170	0.015573
27	HAS_COMMITTER	252	0.023085
28	HAS_AUTHOR	298	0.027299
29	COPIES	331	0.030322

### 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	numbe
0	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	261796	12886	
1	[Git, Change]	MODIFIES	[File, Git]	261796	261796	
2	[Git, Change]	UPDATES	[File, Git]	172614	261796	
3	[Git, Change]	CREATES	[File, Git]	61715	261796	
4	[Git, Change]	DELETES	[File, Git]	38029	261796	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21999	13102	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15589	12886	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12886	1	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12886	252	
9	[Author, Git, Person]	COMMITTED	[Git, Commit]	12886	298	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10887	1	
11	[Git, Change]	RENAMES	[File, Git]	10562	261796	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8363	13102	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8221	13102	
14	[File, Git]	HAS_NEW_NAME	[File, Git]	6240	10887	
15	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6079	13029	
16	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3427	7118	
17	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3053	7176	
18	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2908	7176	
19	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2657	13029	
20	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7176	
21	[Java, ByteCode, Member, Method, Constructor]	WRITES	[Java, ByteCode, Member, Field]	2438	2039	
22	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalType, Ext	2341	2822	
23	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
24	[Java, ByteCode, Member, Method, Constructor]	INVOKES	[Java, ByteCode, Member, Method, Constructor]	2119	2039	
25	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2080	13102	
26	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method, Constructor]	2028	13102	
27	[Java, ByteCode, Member, Method, Constructor]	HAS	[Java, ByteCode, Parameter]	2027	2039	
28	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2022	13029	
29	[Java, ByteCode, Member, Method, Lambda]	INVOKES	[Java, ByteCode, Member, Method]	1892	942	

# Graph Density

total\_number\_of\_nodes (vertices): 351582

- total\_number\_of\_relationships (edges): 1091605
  -> total directed graph density: 8.831073245674903e-06
- -> total directed graph density in percent: 0.0008831073245674903