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`.

```
Total number of nodes: 321599
```

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	233497	72.605014
1	[Java, ByteCode, Member, Method]	13217	4.109776
2	[Java, ByteCode, Parameter]	12968	4.032351
3	[Git, Commit]	11729	3.647088
4	[File, Git]	10590	3.292921
5	[Java, ByteCode, ParameterizedType, Bound]	7154	2.224509
6	[Java, ByteCode, Bound]	7094	2.205853
7	[Java, ByteCode, Member, Field]	3485	1.083648
8	[Java, ByteCode, Bound, WildcardType]	2902	0.902366
9	[Java, Value, ByteCode, Annotation]	2813	0.874692
10	[Xml, Element]	2144	0.666669
11	[Java, ByteCode, Member, Constructor, Method]	2043	0.635263
12	[Xml, Text]	1436	0.446519
13	[Java, ByteCode, Bound, TypeVariable]	1094	0.340175
14	[Java, ByteCode, Member, Method, Lambda]	931	0.289491
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	899	0.279541
16	[Type, File, Java, ByteCode, Class]	858	0.266792
17	[Type, File, Java, ByteCode, JavaType]	746	0.231966
18	[Java, Value, ByteCode, Primitive]	647	0.201182
19	[Java, ByteCode, Member, Method, GenericDeclar...	567	0.176307
20	[Json, Key]	555	0.172575
21	[Value, Json, Scalar]	539	0.167600
22	[Type, File, Java, ByteCode, ExternalType]	396	0.123135
23	[Author, Git, Person]	289	0.089863
24	[Value, Array]	267	0.083023
25	[Committer, Git, Person]	243	0.075560
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.073383
27	[Value, Property]	202	0.062811
28	[Type, File, Java, ByteCode, Interface]	186	0.057836
29	[Java, Value, ByteCode, Class]	170	0.052861

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.

<Figure size 640x480 with 0 Axes>

Nodes per label combination (more than 0.5% overall)

72.61% (233497)

0.64% (2043)
0.67% (2144)
0.87% (2813)
0.90% (2902)
1.08% (3485)
2.21% (7094)
2.22% (7154)
3.29% (10590)
3.65% (11729)
3.72% (11963)
4.03% (12968)
4.11% (13217)

Legend:
- ['Git', 'Change']
- ['Java', 'ByteCode', 'Member', 'Method']
- ['Java', 'ByteCode', 'Parameter']
- others
- ['Git', 'Commit']
- ['File', 'Git']
- ['Java', 'ByteCode', 'ParameterizedType', 'Bound']
- ['Java', 'ByteCode', 'Bound']
- ['Java', 'ByteCode', 'Member', 'Field']
- ['Java', 'ByteCode', 'Bound', 'WildcardType']
- ['Java', 'Value', 'ByteCode', 'Annotation']
- ['Xml', 'Element']
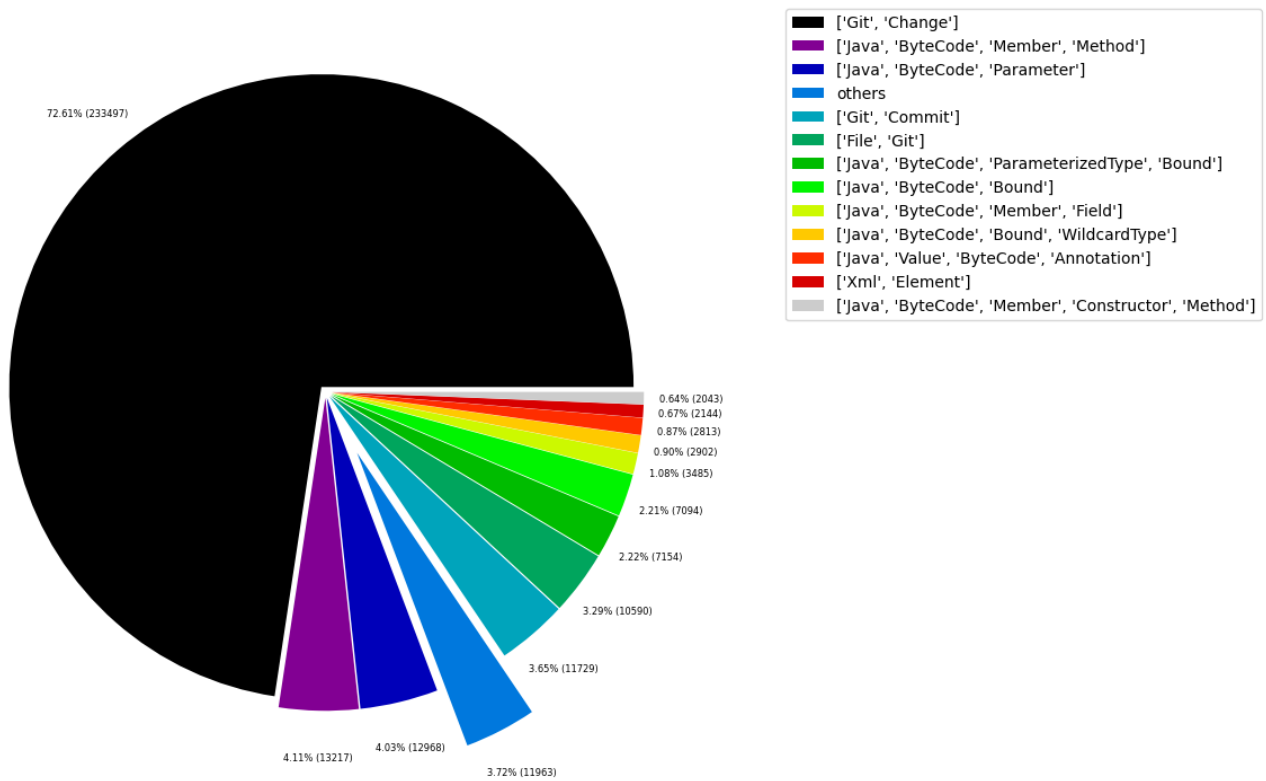- ['Java', 'ByteCode', 'Member', 'Constructor', 'Method']

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	0.000311
1	[File, TS, Scan]	1	0.000311
2	[File, Json]	2	0.000622
3	[File]	3	0.000933
4	[Java, ByteCode, Member, Constructor, Method, ...	4	0.001244
5	[Maven, Exclusion]	5	0.001555
6	[Value, Json, Array]	6	0.001866
7	[File, Maven, Xml, Pom, Document]	9	0.002799
8	[Type, File, Java, ByteCode, Void]	9	0.002799
9	[Java, ManifestSection]	9	0.002799
10	[File, Java, Manifest]	9	0.002799
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002799
12	[File, Java, ServiceLoader]	10	0.003109
13	[File, Java, Properties]	12	0.003731
14	[Maven, ExecutionGoal]	16	0.004975
15	[Maven, PluginExecution]	16	0.004975
16	[Xml, Attribute]	18	0.005597
17	[jQAssistant, Rule, Concept]	19	0.005908
18	[Maven, Plugin]	21	0.006530
19	[Maven, Configuration]	21	0.006530
20	[Git, Branch]	25	0.007774
21	[Type, File, Java, ByteCode, Enum]	28	0.008706
22	[Type, File, Java, ByteCode, PrimitiveType]	29	0.009017
23	[Xml, Namespace]	36	0.011194
24	[Type, File, Java, ByteCode, Annotation]	43	0.013371
25	[File, Directory]	44	0.013682
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.026430
27	[Artifact, Maven]	92	0.028607
28	[Type, File, Java, ByteCode, ExternalAnnotatio...	93	0.028918
29	[Value, ManifestEntry]	120	0.037314

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.

```
<Figure size 640x480 with 0 Axes>
```

Nodes per label combination (less than 0.5% overall)

Legend:
- ['Xml', 'Text']
- ['Java', 'ByteCode', 'Bound', 'TypeVariable']
- ['Java', 'ByteCode', 'Member', 'Method', 'Lambda']
- ['Type', 'File', 'Java', 'ByteCode', 'ResolvedDuplicateType']
- ['Type', 'File', 'Java', 'ByteCode', 'Class']
- ['Type', 'File', 'Java', 'ByteCode', 'JavaType']
- ['Java', 'Value', 'ByteCode', 'Primitive']
- ['Java', 'ByteCode', 'Member', 'Method', 'GenericDeclaration']
- ['Json', 'Key']
- ['Value', 'Json', 'Scalar']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalType']
- ['Author', 'Git', 'Person']
- others
- ['Value', 'Array']
- ['Committer', 'Git', 'Person']
- ['Type', 'File', 'Java', 'ByteCode', 'Class', 'GenericDeclaration']
- ['Value', 'Property']
- ['Type', 'File', 'Java', 'ByteCode', 'Interface']
- ['Java', 'Value', 'ByteCode', 'Class']
- ['Git', 'Tag']
- ['Maven', 'Dependency']
- ['Java', 'ByteCode', 'Bound', 'GenericArrayType']
- ['Java', 'Value', 'ByteCode', 'Enum']
- ['Package', 'File', 'Directory', 'Java']
- ['Value', 'Json', 'Object']
- ['Value', 'ManifestEntry']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalAnnotation', 'ExternalType']
- ['Artifact', 'Maven']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Interface']
- ['File', 'Directory']
- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']
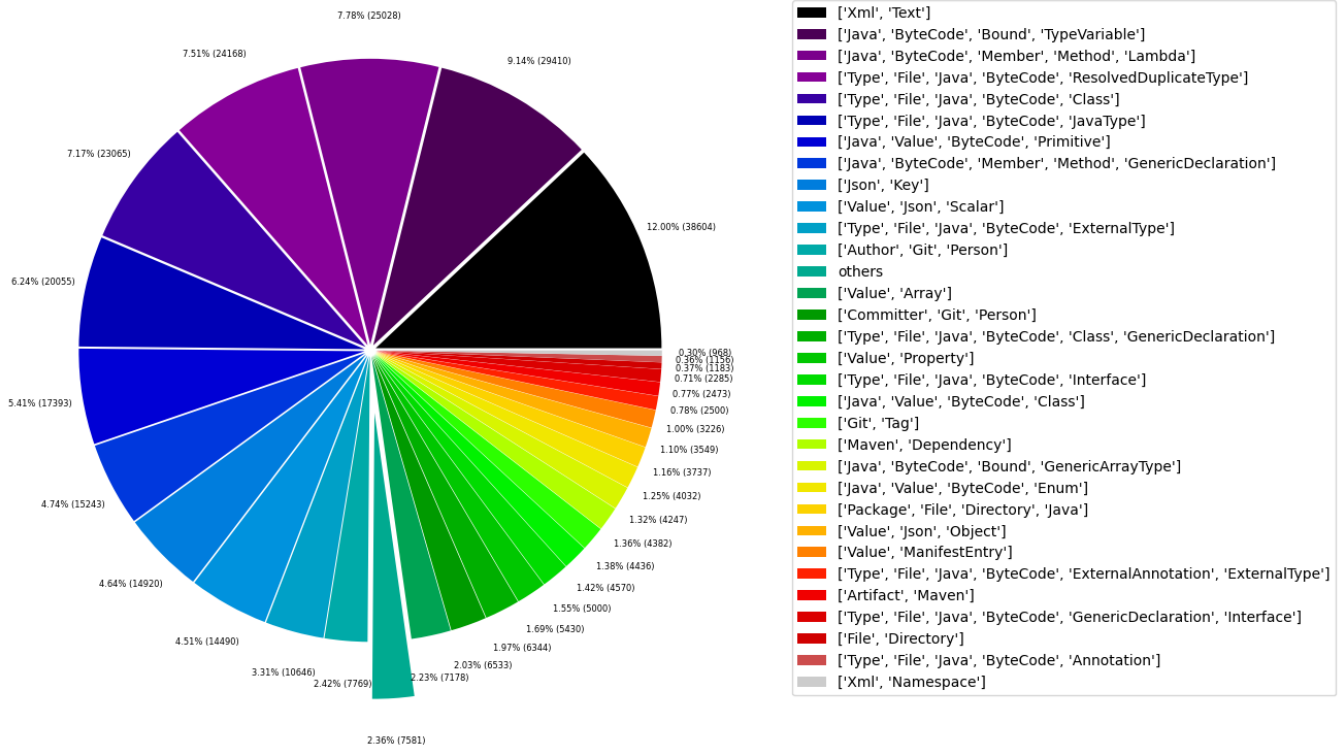- ['Xml', 'Namespace']

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.
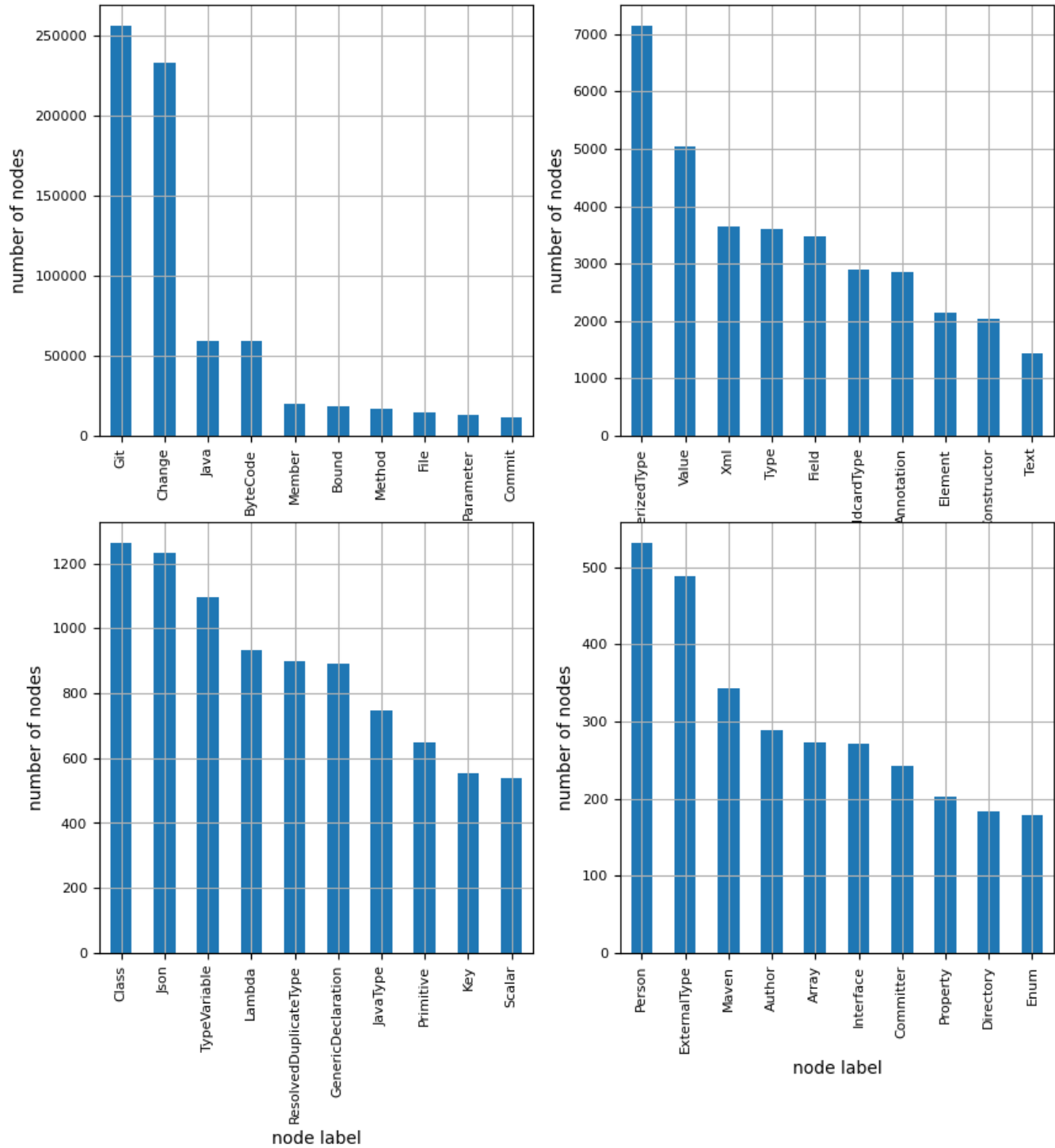
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	256538	79.769527
1	Change	233497	72.605014
2	Java	59193	18.405841
3	ByteCode	59005	18.347383
4	Member	20247	6.295729
5	Bound	18402	5.722033
6	Method	16762	5.212081
7	File	14436	4.488820
8	Parameter	12968	4.032351
9	Commit	11729	3.647088
10	ParameterizedType	7154	2.224509
11	Value	5046	1.569035
12	Xml	3643	1.132777
13	Type	3608	1.121894
14	Field	3485	1.083648
15	WildcardType	2902	0.902366
16	Annotation	2856	0.888062
17	Element	2144	0.666669
18	Constructor	2047	0.636507
19	Text	1436	0.446519
20	Class	1264	0.393036
21	Json	1234	0.383708
22	TypeVariable	1094	0.340175
23	Lambda	931	0.289491
24	ResolvedDuplicateType	899	0.279541
25	GenericDeclaration	892	0.277364
26	JavaType	746	0.231966
27	Primitive	647	0.201182
28	Key	555	0.172575
29	Scalar	539	0.167600
30	Person	532	0.165423
31	ExternalType	489	0.152053
32	Maven	343	0.106655
33	Author	289	0.089863
34	Array	273	0.084888
35	Interface	271	0.084266
36	Committer	243	0.075560
37	Property	202	0.062811
38	Directory	183	0.056903
39	Enum	178	0.055348

Chart 1c - Highest node count by label

Shows the 40 labels with the highest number of nodes.

<Figure size 640x480 with 0 Axes>

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: 974947
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	233497	23.949712
1	MODIFIES	233497	23.949712
2	UPDATES	149837	15.368733
3	CREATES	57905	5.939297
4	DELETES	35670	3.658660
5	INVOKES	35596	3.651070
6	COMMITTED	23458	2.406080
7	DEPENDS_ON	21820	2.238070
8	OF_TYPE	21292	2.183914
9	DECLARES	20686	2.121756
10	OF_RAW_TYPE	17041	1.747890
11	HAS_PARENT	14255	1.462131
12	HAS	14057	1.441822
13	RETURNS	12524	1.284583
14	RENAMES	9915	1.016978
15	READS	9069	0.930204
16	HAS_ACTUAL_TYPE_ARGUMENT	8261	0.847328
17	HAS_NEW_NAME	6164	0.632239
18	OF_GENERIC_TYPE	5892	0.604341
19	RESOLVES_TO	5259	0.539414
20	SIMILAR	3955	0.405663
21	CONTAINS	3797	0.389457
22	WRITES	3796	0.389354
23	RETURNS_GENERIC	3538	0.362892
24	ANNOTATED_BY	2801	0.287298
25	REQUIRES	2172	0.222781
26	HAS_FIRST_CHILD	2144	0.219909
27	HAS_LAST_CHILD	2144	0.219909
28	HAS_ELEMENT	2126	0.218063
29	HAS_UPPER_BOUND	1519	0.155803

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.

```
<Figure size 640x480 with 0 Axes>
```

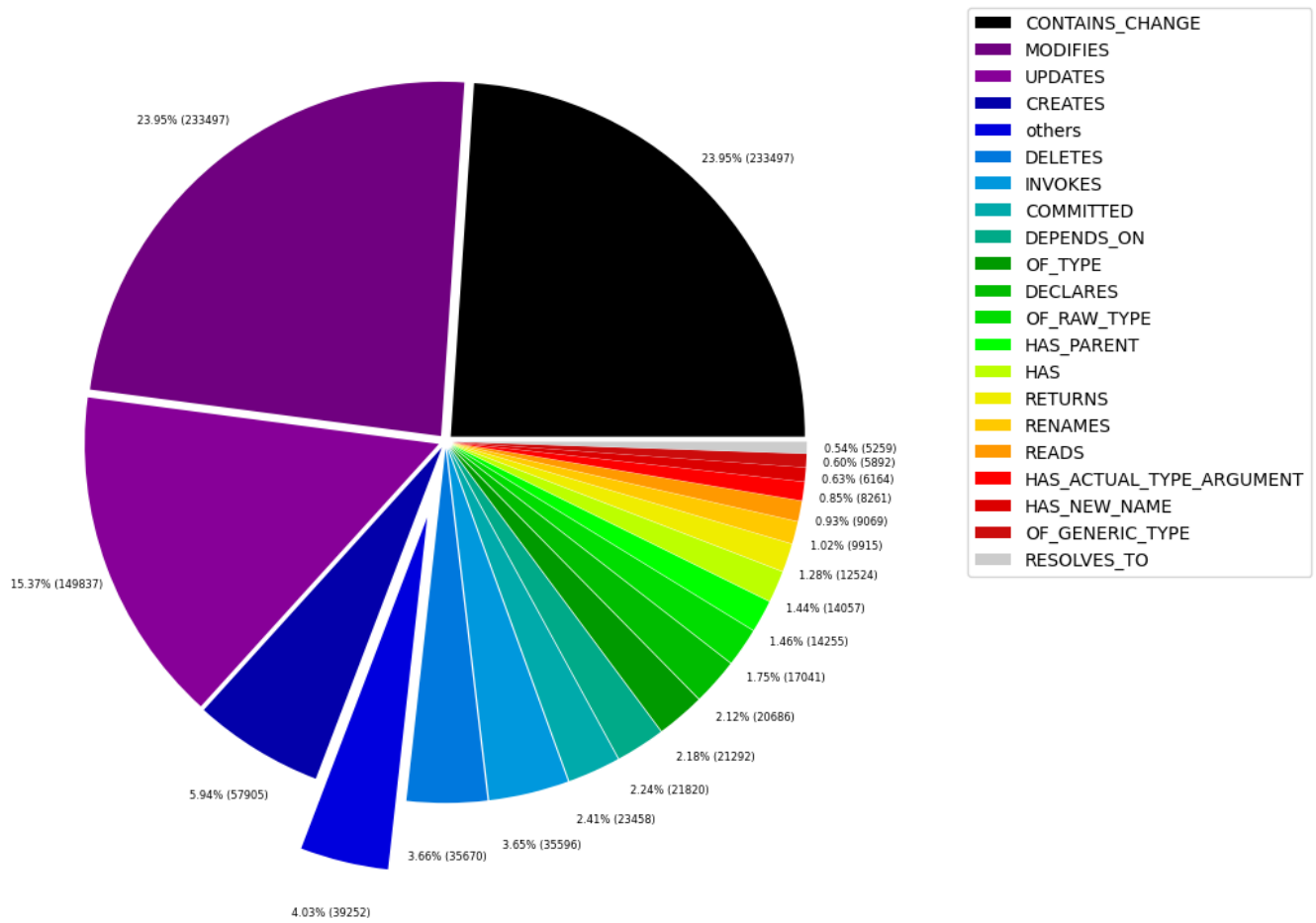Relationship types (more than 0.5% overall)

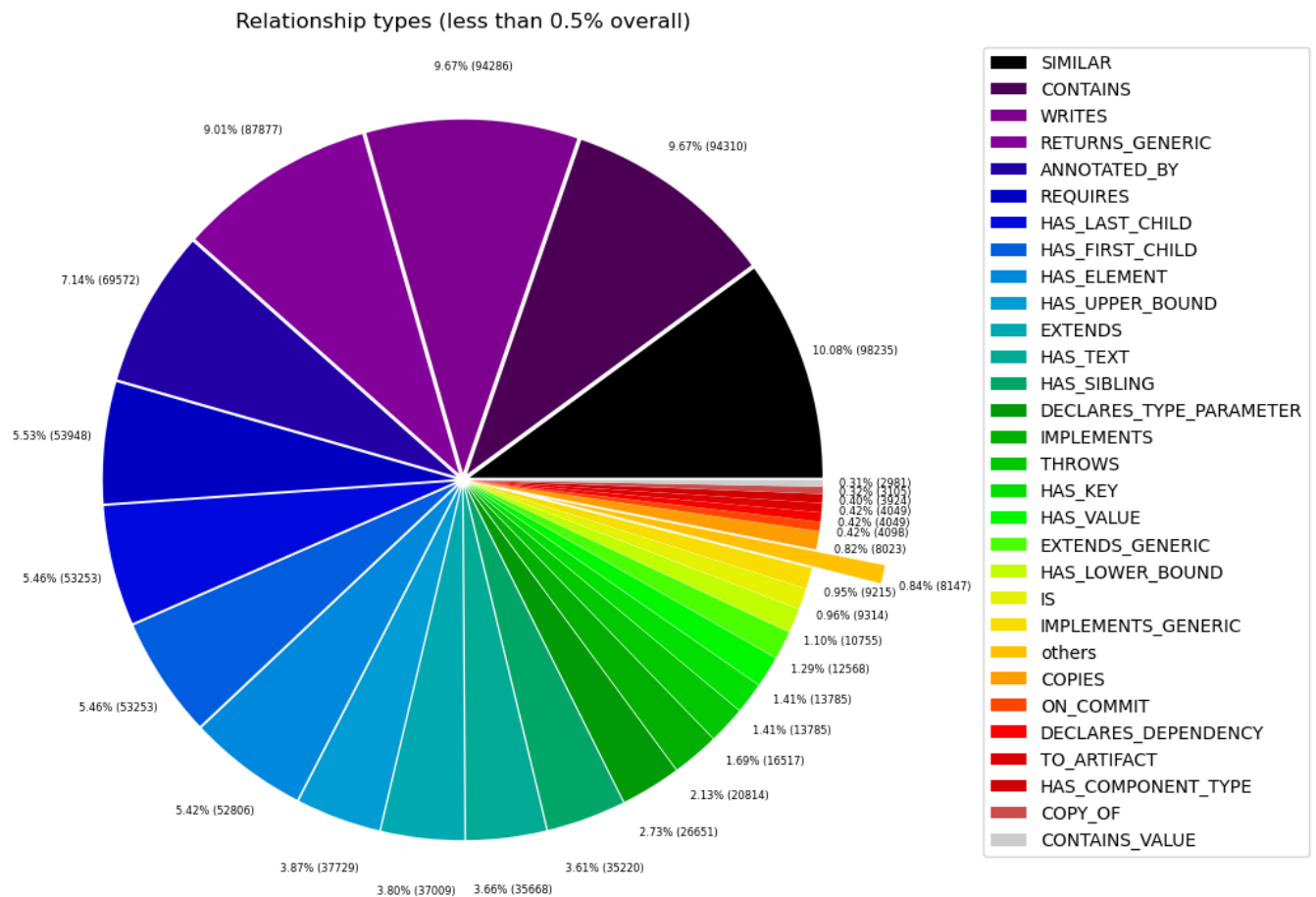## Table 2b - Lowest relationship count by type

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	HAS_PROPERTY	1	0.000103
1	THROWS_GENERIC	5	0.000513
2	EXCLUDES	5	0.000513
3	DESCRIBES	9	0.000923
4	HAS_ROOT_ELEMENT	12	0.001231
5	HAS_GOAL	16	0.001641
6	HAS_EXECUTION	16	0.001641
7	OF_NAMESPACE	18	0.001846
8	HAS_ATTRIBUTE	18	0.001846
9	INCLUDES_CONCEPT	19	0.001949
10	USES_PLUGIN	21	0.002154
11	REQUIRES_TYPE_PARAMETER	21	0.002154
12	IS_ARTIFACT	21	0.002154
13	HAS_CONFIGURATION	21	0.002154
14	HAS_HEAD	25	0.002564
15	REQUIRES_CONCEPT	28	0.002872
16	DECLARES_NAMESPACE	36	0.003693
17	HAS_DEFAULT	36	0.003693
18	CONTAINS_VALUE	120	0.012308
19	COPY_OF	125	0.012821
20	HAS_COMPONENT_TYPE	158	0.016206
21	TO_ARTIFACT	163	0.016719
22	DECLARES_DEPENDENCY	163	0.016719
23	ON_COMMIT	165	0.016924
24	COPIES	323	0.033130
25	IMPLEMENTS_GENERIC	371	0.038053
26	IS	375	0.038464
27	HAS_LOWER_BOUND	433	0.044413
28	EXTENDS_GENERIC	506	0.051900
29	HAS_VALUE	555	0.056926

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.

```
<Figure size 640x480 with 0 Axes>
```

Relationship types (less than 0.5% overall)

Legend:
- SIMILAR
- CONTAINS
- WRITES
- RETURNS_GENERIC
- ANNOTATED_BY
- REQUIRES
- HAS_LAST_CHILD
- HAS_FIRST_CHILD
- HAS_ELEMENT
- HAS_UPPER_BOUND
- EXTENDS
- HAS_TEXT
- HAS_SIBLING
- DECLARES_TYPE_PARAMETER
- IMPLEMENTS
- THROWS
- HAS_KEY
- HAS_VALUE
- EXTENDS_GENERIC
- HAS_LOWER_BOUND
- IS
- IMPLEMENTS_GENERIC
- others
- COPIES
- ON_COMMIT
- DECLARES_DEPENDENCY
- TO_ARTIFACT
- HAS_COMPONENT_TYPE
- COPY_OF
- CONTAINS_VALUE

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	numberOfNodesWithSameLabelsAsSource	numberOfNodesWithS
0	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	233497	11729	
1	[Git, Change]	MODIFIES	[File, Git]	233497	233497	
2	[Git, Change]	UPDATES	[File, Git]	149837	233497	
3	[Git, Change]	CREATES	[File, Git]	57905	233497	
4	[Git, Change]	DELETES	[File, Git]	35670	233497	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21654	13098	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14246	11729	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	11729	289	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	11729	243	
9	[Git, Change]	RENAMES	[File, Git]	9915	233497	
10	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8295	13098	
11	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8122	13098	
12	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6335	12968	
13	[File, Git]	HAS_NEW_NAME	[File, Git]	6164	10590	
14	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3554	7094	
15	[Java, ByteCode, ParameterizedType, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3034	7154	
16	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2902	7154	
17	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, ParameterizedType, Bound]	2650	12968	
18	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2453	7154	
19	[Java, ByteCode, Member, Constructor, Method]	WRITES	[Java, ByteCode, Member, Field]	2423	2043	
20	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio...]	2336	2813	
21	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
22	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2107	2043	
23	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2084	13098	
24	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2037	13098	
25	[Java, ByteCode, Member, Constructor, Method]	HAS	[Java, ByteCode, Parameter]	2021	2043	
26	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2020	12968	
27	[Java, ByteCode, Member, Method, Lambda]	INVOKES	[Java, ByteCode, Member, Method]	1880	931	
28	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Method]	1800	2043	
29	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound]	1792	7154	

Graph Density

```
total_number_of_nodes (vertices): 321599
total_number_of_relationships (edges): 974947
-> total directed graph density: 9.426554404602033e-06
-> total directed graph density in percent: 0.0009426554404602033
```