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: 319229
```

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	231259	72.442980
1	[Java, ByteCode, Method, Member]	13187	4.130890
2	[Java, ByteCode, Parameter]	12968	4.062288
3	[Git, Commit]	11659	3.652237
4	[File, Git]	10572	3.311729
5	[Java, ByteCode, Bound, ParameterizedType]	7154	2.241024
6	[Java, ByteCode, Bound]	7094	2.222229
7	[Java, ByteCode, Member, Field]	3483	1.091066
8	[Java, ByteCode, Bound, WildcardType]	2902	0.909065
9	[Java, Value, ByteCode, Annotation]	2813	0.881186
10	[Xml, Element]	2144	0.671618
11	[Java, ByteCode, Constructor, Method, Member]	2033	0.636847
12	[Xml, Text]	1436	0.449834
13	[Java, ByteCode, Bound, TypeVariable]	1094	0.342701
14	[Java, ByteCode, Method, Member, Lambda]	931	0.291640
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	899	0.281616
16	[Type, File, Java, ByteCode, Class]	858	0.268773
17	[Type, File, Java, ByteCode, JavaType]	746	0.233688
18	[Java, Value, ByteCode, Primitive]	647	0.202676
19	[Java, ByteCode, Method, Member, GenericDeclar...	567	0.177615
20	[Json, Key]	555	0.173856
21	[Value, Json, Scalar]	539	0.168844
22	[Type, File, Java, ByteCode, ExternalType]	396	0.124049
23	[Author, Git, Person]	288	0.090217
24	[Value, Array]	267	0.083639
25	[Committer, Git, Person]	242	0.075808
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.073928
27	[Value, Property]	202	0.063277
28	[Type, File, Java, ByteCode, Interface]	186	0.058265
29	[Java, Value, ByteCode, Class]	170	0.053253

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)

72.44% (231259)

0.64% (2033)
0.67% (2144)
0.88% (2813)
0.91% (2902)
1.09% (3483)
2.22% (7094)
2.24% (7154)
3.31% (10572)
3.65% (11659)
3.75% (11961)
4.06% (12968)
4.13% (13187)

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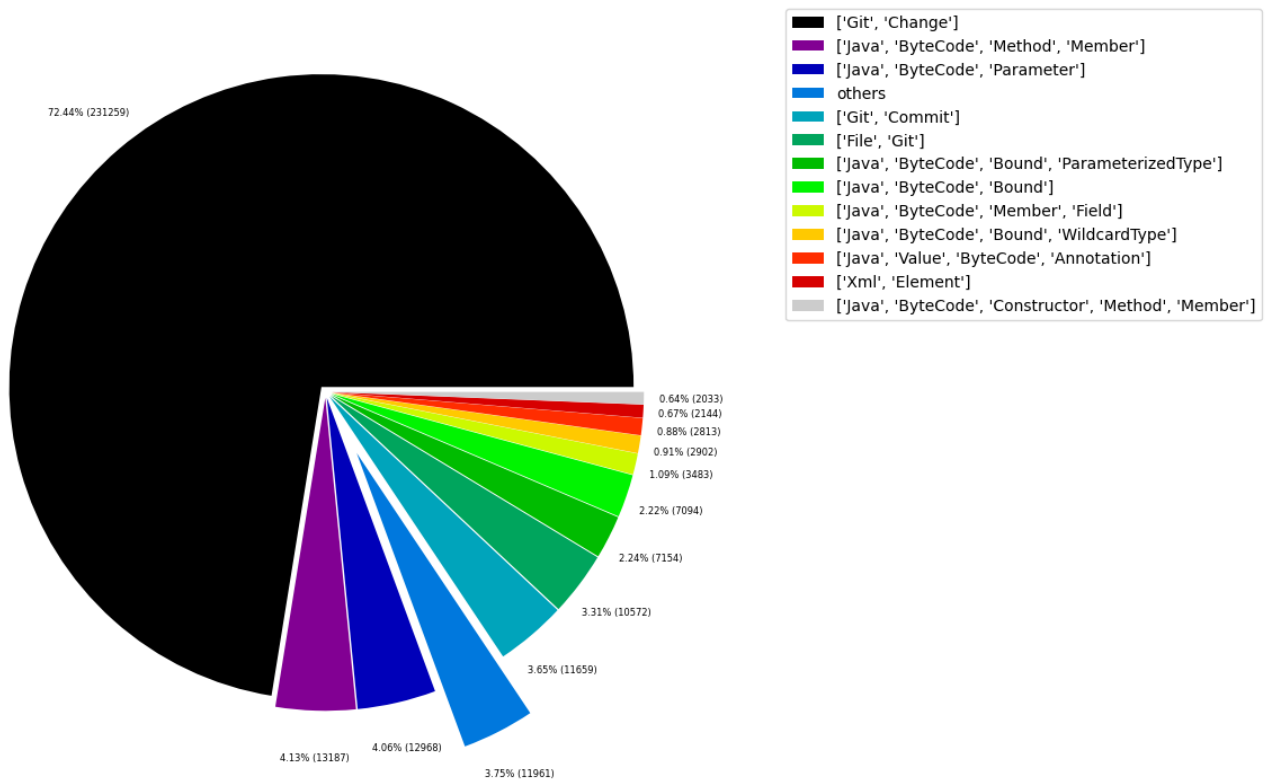- ['Java', 'ByteCode', 'Constructor', 'Method', 'Member']

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	0.000313
1	[File, TS, Scan]	1	0.000313
2	[File, Json]	2	0.000627
3	[File]	3	0.000940
4	[Java, ByteCode, Constructor, Method, Member, ...	4	0.001253
5	[Maven, Exclusion]	5	0.001566
6	[Value, Array, Json]	6	0.001880
7	[File, Maven, Xml, Pom, Document]	9	0.002819
8	[Type, File, Java, ByteCode, Void]	9	0.002819
9	[Java, ManifestSection]	9	0.002819
10	[File, Java, Manifest]	9	0.002819
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002819
12	[File, Java, ServiceLoader]	10	0.003133
13	[File, Java, Properties]	12	0.003759
14	[Maven, ExecutionGoal]	16	0.005012
15	[Maven, PluginExecution]	16	0.005012
16	[Xml, Attribute]	18	0.005639
17	[jQAssistant, Rule, Concept]	19	0.005952
18	[Maven, Plugin]	21	0.006578
19	[Maven, Configuration]	21	0.006578
20	[Git, Branch]	25	0.007831
21	[Type, File, Java, ByteCode, Enum]	28	0.008771
22	[Type, File, Java, ByteCode, PrimitiveType]	29	0.009084
23	[Xml, Namespace]	36	0.011277
24	[Type, File, Java, ByteCode, Annotation]	43	0.013470
25	[File, Directory]	44	0.013783
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.026627
27	[Artifact, Maven]	92	0.028819
28	[Type, File, Java, ByteCode, ExternalType, Ext...	93	0.029133
29	[Value, ManifestEntry]	120	0.037591

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', 'Method', 'Member', 'Lambda']
- ['Type', 'File', 'Java', 'ByteCode', 'ResolvedDuplicateType']
- ['Type', 'File', 'Java', 'ByteCode', 'Class']
- ['Type', 'File', 'Java', 'ByteCode', 'JavaType']
- ['Java', 'Value', 'ByteCode', 'Primitive']
- ['Java', 'ByteCode', 'Method', 'Member', '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', 'ExternalType', 'ExternalAnnotation']
- ['Artifact', 'Maven']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Interface']
- ['File', 'Directory']
- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']
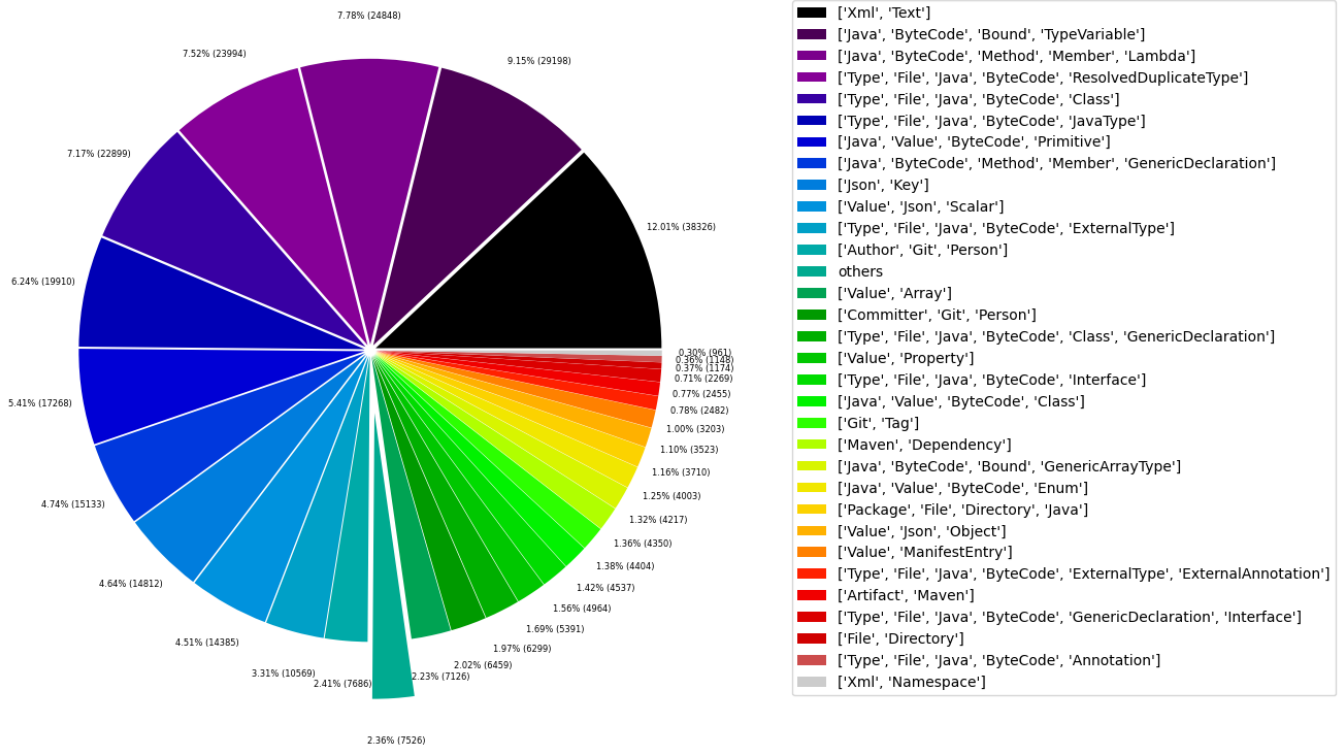- ['Xml', 'Namespace']

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.
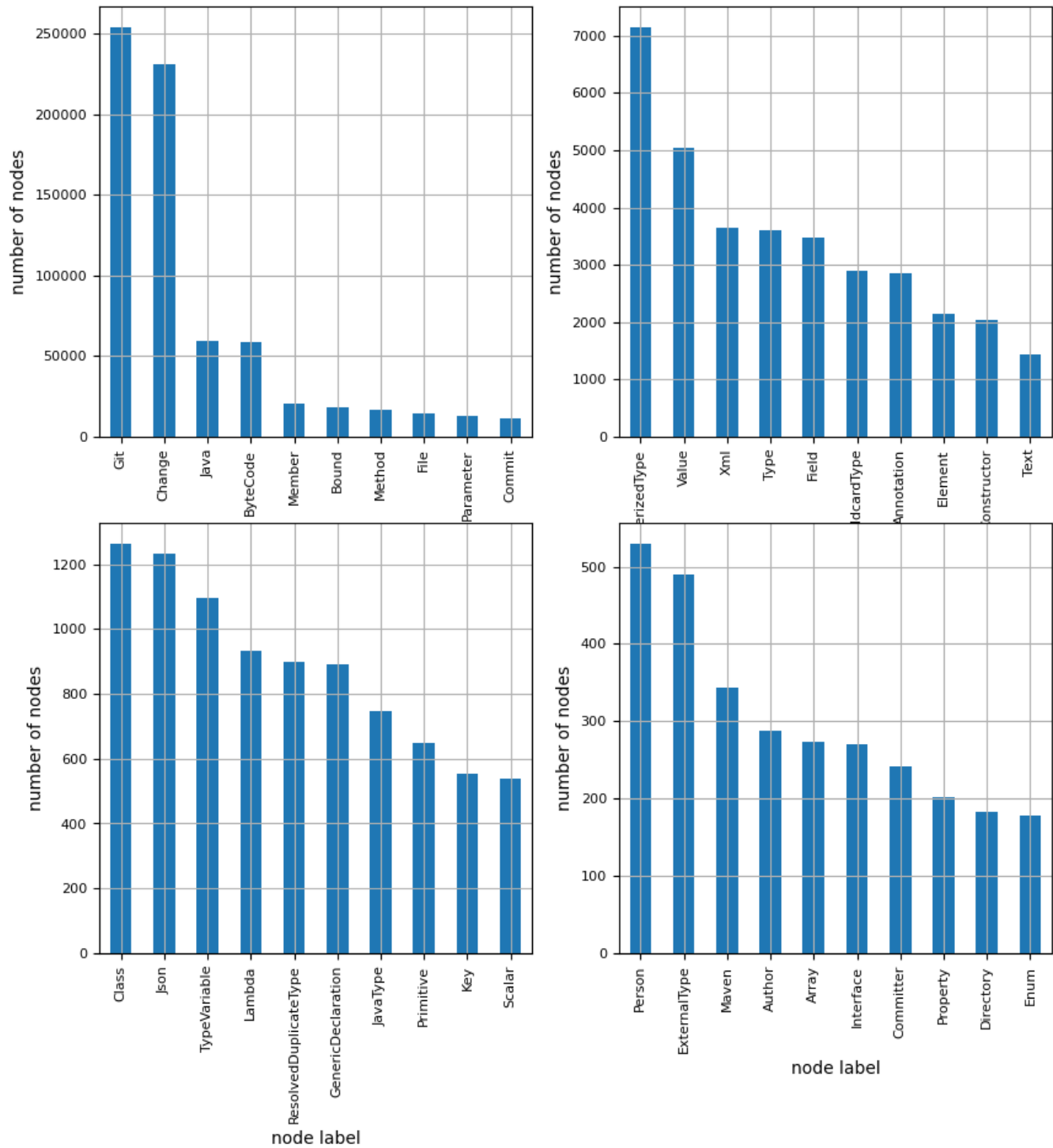
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	254210	79.632490
1	Change	231259	72.442980
2	Java	59151	18.529332
3	ByteCode	58963	18.470440
4	Member	20205	6.329312
5	Bound	18402	5.764514
6	Method	16722	5.238246
7	File	14418	4.516507
8	Parameter	12968	4.062288
9	Commit	11659	3.652237
10	ParameterizedType	7154	2.241024
11	Value	5046	1.580683
12	Xml	3643	1.141187
13	Type	3608	1.130223
14	Field	3483	1.091066
15	WildcardType	2902	0.909065
16	Annotation	2856	0.894656
17	Element	2144	0.671618
18	Constructor	2037	0.638100
19	Text	1436	0.449834
20	Class	1264	0.395954
21	Json	1234	0.386556
22	TypeVariable	1094	0.342701
23	Lambda	931	0.291640
24	ResolvedDuplicateType	899	0.281616
25	GenericDeclaration	892	0.279423
26	JavaType	746	0.233688
27	Primitive	647	0.202676
28	Key	555	0.173856
29	Scalar	539	0.168844
30	Person	530	0.166025
31	ExternalType	489	0.153182
32	Maven	343	0.107446
33	Author	288	0.090217
34	Array	273	0.085519
35	Interface	271	0.084892
36	Committer	242	0.075808
37	Property	202	0.063277
38	Directory	183	0.057326
39	Enum	178	0.055759

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: 967584
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	231259	23.900664
1	MODIFIES	231259	23.900664
2	UPDATES	148166	15.312986
3	CREATES	57605	5.953488
4	INVOKES	35753	3.695080
5	DELETES	35346	3.653016
6	COMMITTED	23318	2.409920
7	DEPENDS_ON	21459	2.217792
8	OF_TYPE	21292	2.200532
9	DECLARES	20644	2.133562
10	OF_RAW_TYPE	17041	1.761191
11	HAS_PARENT	14174	1.464886
12	HAS	14057	1.452794
13	RETURNS	12524	1.294358
14	RENAMES	9858	1.018826
15	READS	9068	0.937180
16	HAS_ACTUAL_TYPE_ARGUMENT	8261	0.853776
17	HAS_NEW_NAME	6161	0.636741
18	OF_GENERIC_TYPE	5892	0.608939
19	RESOLVES_TO	5237	0.541245
20	SIMILAR	3915	0.404616
21	CONTAINS	3797	0.392421
22	WRITES	3796	0.392317
23	RETURNS_GENERIC	3538	0.365653
24	ANNOTATED_BY	2801	0.289484
25	REQUIRES	2172	0.224477
26	HAS_FIRST_CHILD	2144	0.221583
27	HAS_LAST_CHILD	2144	0.221583
28	HAS_ELEMENT	2126	0.219723
29	HAS_UPPER_BOUND	1519	0.156989

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)

- CONTAINS_CHANGE
- MODIFIES
- UPDATES
- CREATES
- others
- INVOKES
- DELETES
- COMMITTED
- DEPENDS_ON
- OF_TYPE
- DECLARES
- OF_RAW_TYPE
- HAS_PARENT
- HAS
- RETURNS
- RENAMES
- READS
- HAS_ACTUAL_TYPE_ARGUMENT
- HAS_NEW_NAME
- OF_GENERIC_TYPE
- RESOLVES_TO

23.90% (231259)
23.90% (231259)
15.31% (148166)
5.95% (57605)
4.05% (39210)
3.70% (35753)
3.65% (35346)
2.41% (23318)
2.22% (21459)
2.20% (21292)
2.13% (20644)
1.76% (17041)
1.46% (14174)
1.45% (14057)
1.29% (12524)
1.02% (9858)
0.94% (9068)
0.85% (8261)
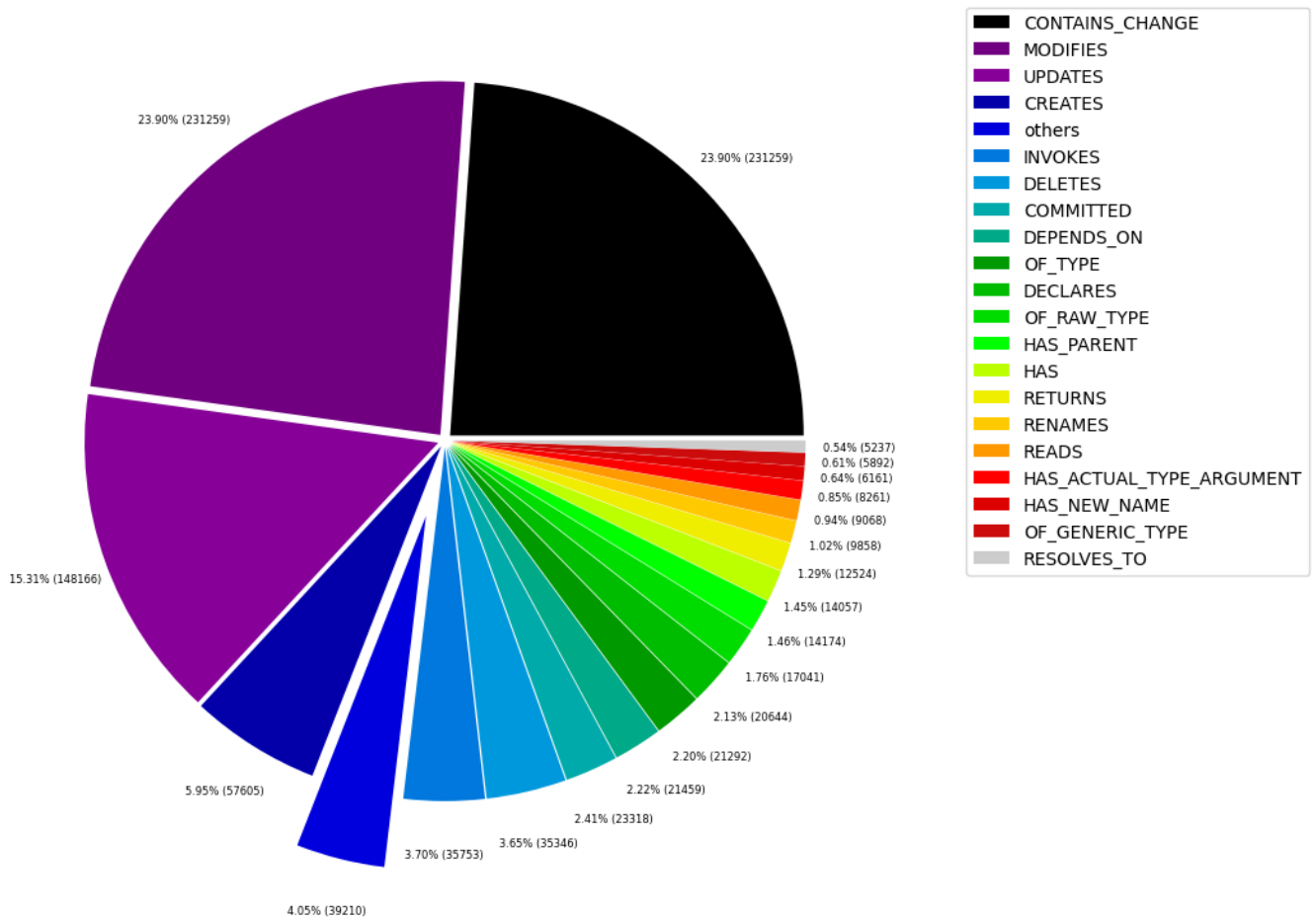0.64% (6161)
0.61% (5892)
0.54% (5237)

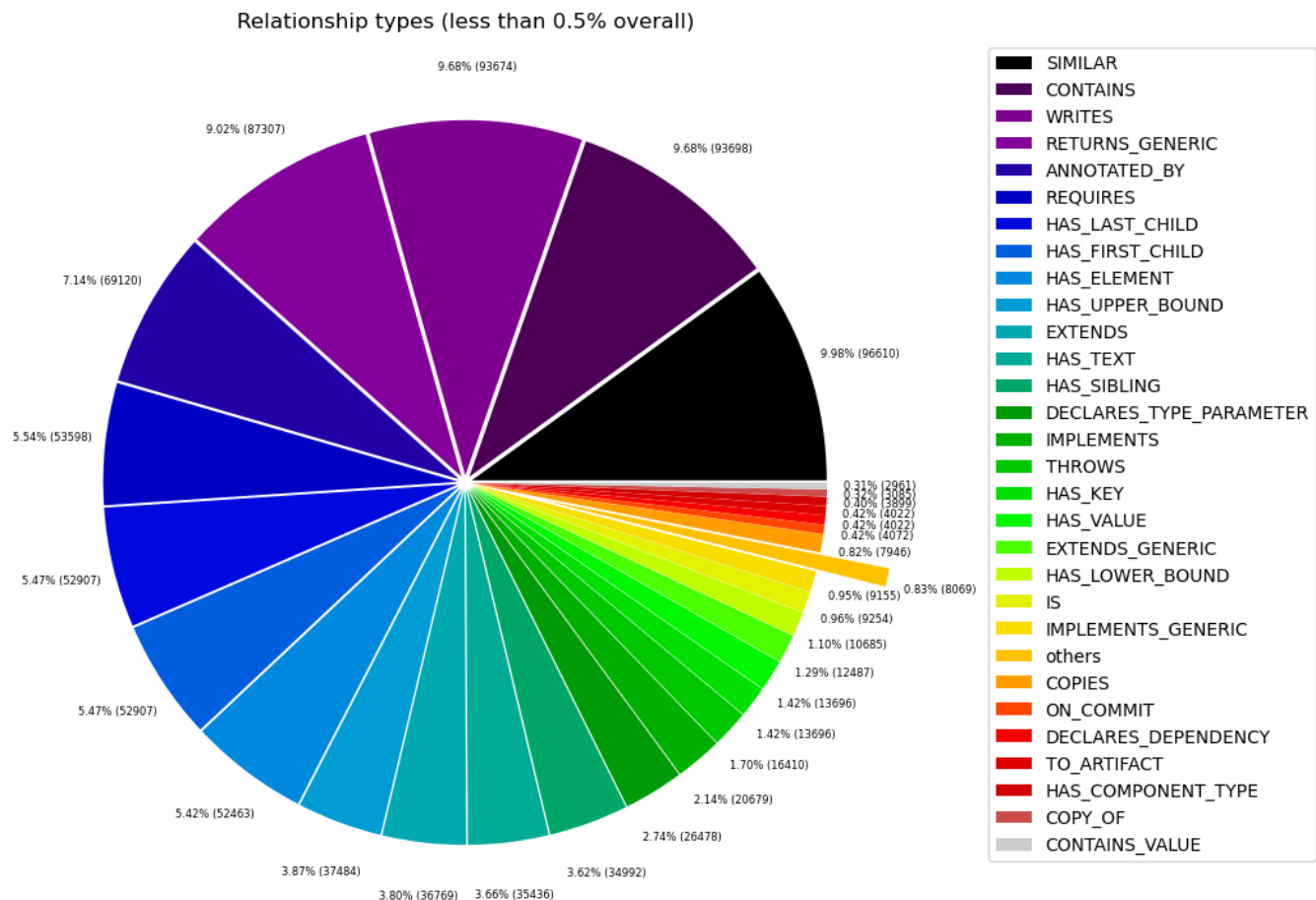## Table 2b - Lowest relationship count by type

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	0.000517
2	EXCLUDES	5	0.000517
3	DESCRIBES	9	0.000930
4	HAS_ROOT_ELEMENT	11	0.001137
5	HAS_GOAL	16	0.001654
6	HAS_EXECUTION	16	0.001654
7	OF_NAMESPACE	18	0.001860
8	HAS_ATTRIBUTE	18	0.001860
9	INCLUDES_CONCEPT	19	0.001964
10	USES_PLUGIN	21	0.002170
11	REQUIRES_TYPE_PARAMETER	21	0.002170
12	IS_ARTIFACT	21	0.002170
13	HAS_CONFIGURATION	21	0.002170
14	HAS_HEAD	25	0.002584
15	REQUIRES_CONCEPT	28	0.002894
16	DECLARES_NAMESPACE	36	0.003721
17	HAS_DEFAULT	36	0.003721
18	CONTAINS_VALUE	120	0.012402
19	COPY_OF	125	0.012919
20	HAS_COMPONENT_TYPE	158	0.016329
21	TO_ARTIFACT	163	0.016846
22	DECLARES_DEPENDENCY	163	0.016846
23	ON_COMMIT	165	0.017053
24	COPIES	322	0.033279
25	IMPLEMENTS_GENERIC	371	0.038343
26	IS	375	0.038756
27	HAS_LOWER_BOUND	433	0.044751
28	EXTENDS_GENERIC	506	0.052295
29	HAS_VALUE	555	0.057359

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)

9.68% (93674)
9.02% (87307)
9.68% (93698)
7.14% (69120)
9.98% (96610)
5.54% (53598)
0.31% (2961)
0.32% (3085)
0.40% (3899)
0.42% (4022)
0.42% (4022)
0.42% (4072)
5.47% (52907)
0.82% (7946)
0.83% (8069)
5.47% (52907)
0.95% (9155)
0.96% (9254)
1.10% (10685)
5.42% (52463)
1.29% (12487)
1.42% (13696)
1.42% (13696)
3.87% (37484)
1.70% (16410)
2.14% (20679)
3.80% (36769)
2.74% (26478)
3.66% (35436)
3.62% (34992)

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]	231259	11659	
1	[Git, Change]	MODIFIES	[File, Git]	231259	231259	
2	[Git, Change]	UPDATES	[File, Git]	148166	231259	
3	[Git, Change]	CREATES	[File, Git]	57605	231259	
4	[Git, Change]	DELETES	[File, Git]	35346	231259	
5	[Java, ByteCode, Method, Member]	INVOKES	[Java, ByteCode, Method, Member]	21885	13071	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14165	11659	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	11659	288	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	11659	242	
9	[Git, Change]	RENAMES	[File, Git]	9858	231259	
10	[Java, ByteCode, Method, Member]	HAS	[Java, ByteCode, Parameter]	8290	13071	
11	[Java, ByteCode, Method, Member]	READS	[Java, ByteCode, Member, Field]	8125	13071	
12	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6335	12968	
13	[File, Git]	HAS_NEW_NAME	[File, Git]	6161	10572	
14	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3554	7094	
15	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3034	7154	
16	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2902	7154	
17	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2650	12968	
18	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2453	7154	
19	[Java, ByteCode, Constructor, Method, Member]	WRITES	[Java, ByteCode, Member, Field]	2423	2033	
20	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalType, Ext...	2336	2813	
21	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
22	[Java, ByteCode, Constructor, Method, Member]	INVOKES	[Java, ByteCode, Constructor, Method, Member]	2100	2033	
23	[Java, ByteCode, Method, Member]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2088	13071	
24	[Java, ByteCode, Method, Member]	INVOKES	[Java, ByteCode, Constructor, Method, Member]	2034	13071	
25	[Java, ByteCode, Constructor, Method, Member]	HAS	[Java, ByteCode, Parameter]	2021	2033	
26	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2020	12968	
27	[Java, ByteCode, Method, Member, Lambda]	INVOKES	[Java, ByteCode, Method, Member]	1866	931	
28	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound]	1792	7154	
29	[Java, ByteCode, Constructor, Method, Member]	INVOKES	[Java, ByteCode, Method, Member]	1790	2033	

Graph Density

```
total_number_of_nodes (vertices): 319229
total_number_of_relationships (edges): 967584
-> total directed graph density: 9.49479000095631e-06
-> total directed graph density in percent: 0.0009494790000956311
```