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]	259908	74.382904
1	[Java, ByteCode, Member, Method]	13197	3.776841
2	[Java, ByteCode, Parameter]	13029	3.728761
3	[Git, Commit]	12722	3.640901
4	[File, Git]	10810	3.093707
5	[Java, ByteCode, Bound, ParameterizedType]	7176	2.053695
6	[Java, ByteCode, Bound]	7118	2.037096
7	[Java, ByteCode, Member, Field]	3502	1.002235
8	[Java, ByteCode, Bound, WildcardType]	2908	0.832239
9	[Java, Value, ByteCode, Annotation]	2822	0.807626
10	[Xml, Element]	2144	0.613590
11	[Java, ByteCode, Member, Method, Constructor]	2039	0.583540
12	[Xml, Text]	1436	0.410968
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.313664
14	[Java, ByteCode, Member, Method, Lambda]	942	0.269590
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	877	0.250988
16	[Type, File, Java, ByteCode, Class]	810	0.231813
17	[Java, Value, ByteCode, Primitive]	652	0.186595
18	[Type, File, Java, ByteCode, JavaType]	644	0.184306
19	[Java, ByteCode, Member, Method, GenericDeclar...	569	0.162842
20	[Json, Key]	560	0.160266
21	[Value, Json, Scalar]	544	0.155687
22	[Type, File, Java, ByteCode, ExternalType]	380	0.108752
23	[Author, Git, Person]	298	0.085284
24	[Value, Array]	267	0.076413
25	[Committer, Git, Person]	252	0.072120
26	[Type, File, Java, ByteCode, GenericDeclaratio...	236	0.067541
27	[Value, Property]	202	0.057810
28	[Type, File, Java, ByteCode, Interface]	188	0.053804
29	[Java, Value, ByteCode, Class]	172	0.049225

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)

Legend:
- ['Git', 'Change']
- ['Java', 'ByteCode', 'Member', 'Method']
- ['Java', 'ByteCode', 'Parameter']
- ['Git', 'Commit']
- others
- ['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', 'Member', 'Method', 'Constructor']

Values shown on chart:
- 74.38% (259908)
- 0.58% (2039)
- 0.61% (2144)
- 0.81% (2822)
- 0.83% (2908)
- 1.00% (3502)
- 2.04% (7118)
- 2.05% (7176)
- 3.09% (10810)
- 3.45% (12044)
- 3.64% (12722)
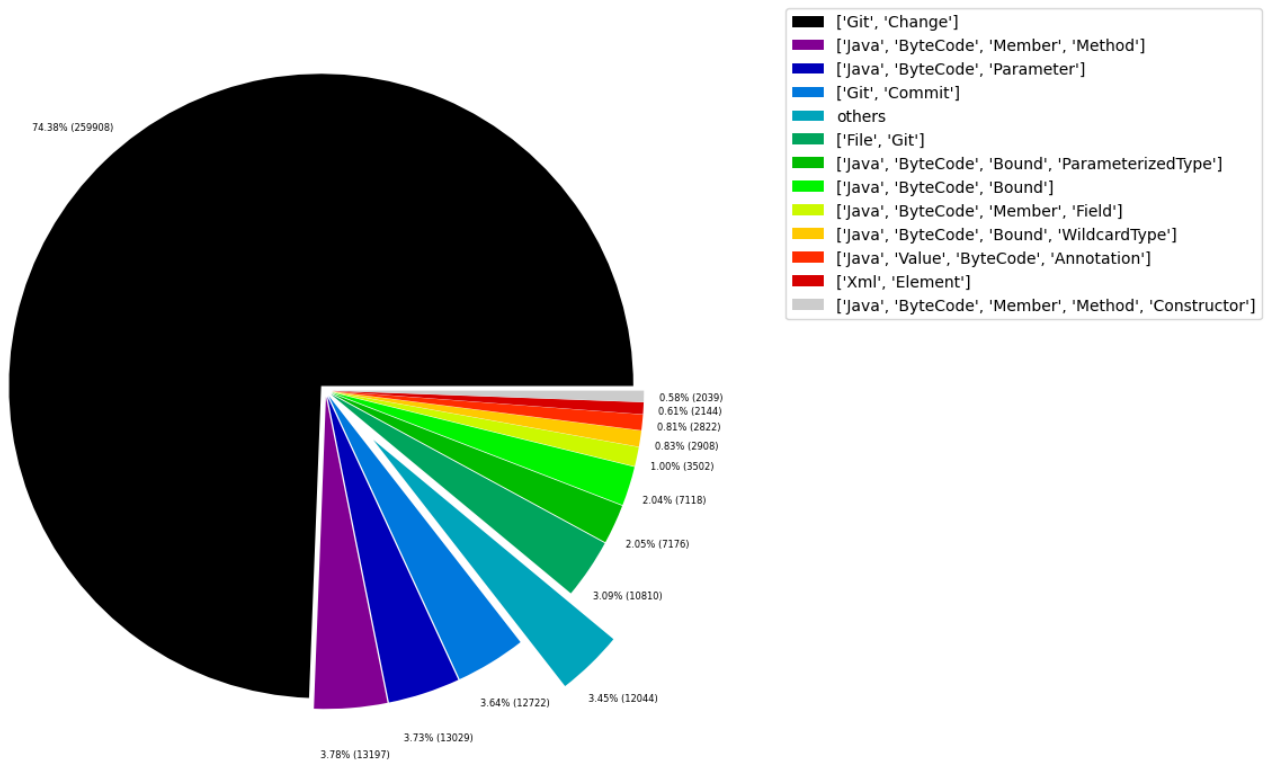- 3.73% (13029)
- 3.78% (13197)

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	0.000286
1	[Repository, File, Git]	1	0.000286
2	[File, Json]	2	0.000572
3	[File]	3	0.000859
4	[Java, ByteCode, Member, Method, GenericDeclar...	4	0.001145
5	[Maven, Exclusion]	5	0.001431
6	[Value, Array, Json]	6	0.001717
7	[Type, File, Java, ByteCode, Void]	9	0.002576
8	[File, Maven, Xml, Pom, Document]	9	0.002576
9	[Java, ManifestSection]	9	0.002576
10	[File, Java, Manifest]	9	0.002576
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002576
12	[File, Java, ServiceLoader]	10	0.002862
13	[File, Java, Properties]	12	0.003434
14	[Type, File, Java, ByteCode, Throwable, Extern...	16	0.004579
15	[Maven, ExecutionGoal]	16	0.004579
16	[Maven, PluginExecution]	16	0.004579
17	[Xml, Attribute]	18	0.005151
18	[jQAssistant, Rule, Concept]	19	0.005438
19	[Maven, Plugin]	21	0.006010
20	[Maven, Configuration]	21	0.006010
21	[Type, File, Java, ByteCode, Throwable, Resolv...	22	0.006296
22	[Type, File, Java, ByteCode, Enum]	28	0.008013
23	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008299
24	[Xml, Namespace]	36	0.010303
25	[Git, Branch]	41	0.011734
26	[Type, File, Java, ByteCode, Annotation]	43	0.012306
27	[File, Directory]	44	0.012592
28	[Type, File, Java, ByteCode, Class, Throwable]	54	0.015454
29	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.024326

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']
- ['Java', 'Value', 'ByteCode', 'Primitive']
- ['Type', 'File', 'Java', 'ByteCode', 'JavaType']
- ['Java', 'ByteCode', 'Member', 'Method', 'GenericDeclaration']
- ['Json', 'Key']
- ['Value', 'Json', 'Scalar']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalType']
- others
- ['Author', 'Git', 'Person']
- ['Value', 'Array']
- ['Committer', 'Git', 'Person']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Class']
- ['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', 'Throwable', 'JavaType']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalType', 'ExternalAnnotation']
- ['Artifact', 'Maven']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Interface']
- ['Type', 'File', 'Java', 'ByteCode', 'Class', 'Throwable']
- ['File', 'Directory']
- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']
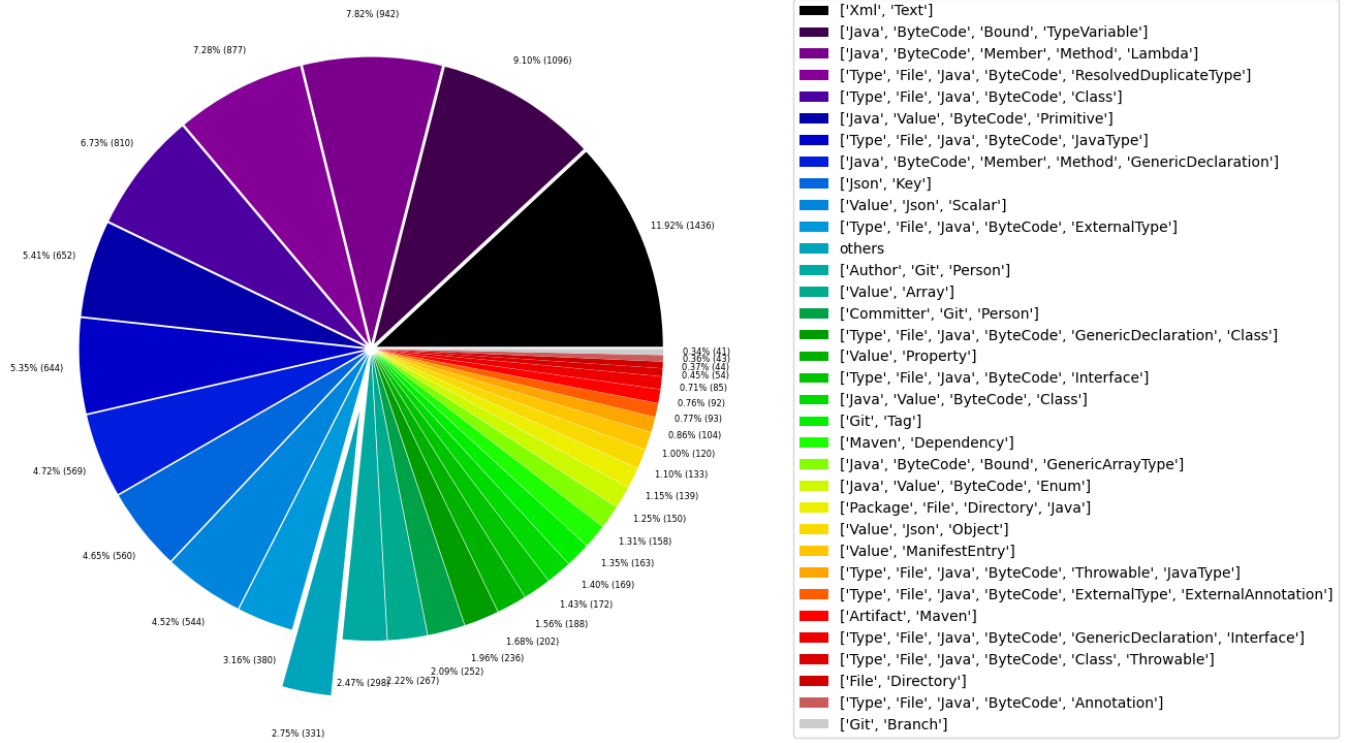- ['Git', 'Branch']

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.
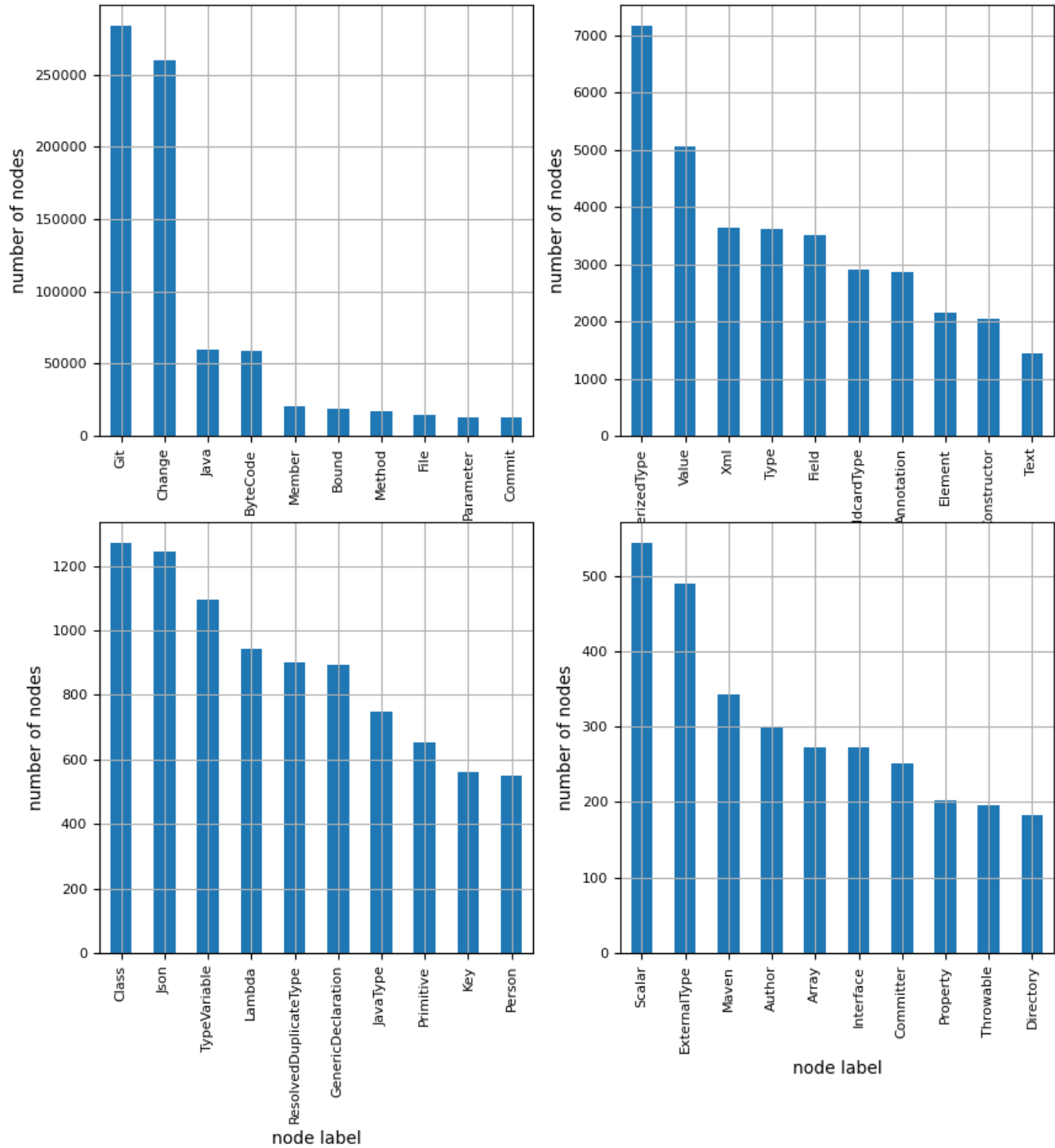
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	284201	81.335302
1	Change	259908	74.382904
2	Java	59340	16.982477
3	ByteCode	59152	16.928673
4	Member	20253	5.796193
5	Bound	18456	5.281911
6	Method	16751	4.793958
7	File	14666	4.197253
8	Parameter	13029	3.728761
9	Commit	12722	3.640901
10	ParameterizedType	7176	2.053695
11	Value	5068	1.450408
12	Xml	3643	1.042588
13	Type	3618	1.035433
14	Field	3502	1.002235
15	WildcardType	2908	0.832239
16	Annotation	2865	0.819933
17	Element	2144	0.613590
18	Constructor	2043	0.584685
19	Text	1436	0.410968
20	Class	1272	0.364033
21	Json	1245	0.356306
22	TypeVariable	1096	0.313664
23	Lambda	942	0.269590
24	ResolvedDuplicateType	899	0.257284
25	GenericDeclaration	894	0.255853
26	JavaType	748	0.214070
27	Primitive	652	0.186595
28	Key	560	0.160266
29	Person	550	0.157404
30	Scalar	544	0.155687
31	ExternalType	489	0.139947
32	Maven	343	0.098163
33	Author	298	0.085284
34	Array	273	0.078130
35	Interface	273	0.078130
36	Committer	252	0.072120
37	Property	202	0.057810
38	Throwable	196	0.056093
39	Directory	183	0.052373

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: 1085117
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	259908	23.952072
1	MODIFIES	259908	23.952072
2	UPDATES	171251	15.781800
3	CREATES	61217	5.641511
4	DELETES	37993	3.501281
5	INVOKES	35947	3.312730
6	COMMITTED	25444	2.344816
7	DEPENDS_ON	21931	2.021072
8	OF_TYPE	21394	1.971585
9	DECLARES	20698	1.907444
10	OF_RAW_TYPE	17093	1.575222
11	HAS_PARENT	15406	1.419755
12	HAS	14125	1.301703
13	HAS_COMMIT	12722	1.172408
14	RETURNS	12578	1.159138
15	HAS_FILE	10810	0.996206
16	RENAMES	10553	0.972522
17	READS	9161	0.844241
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.763789
19	HAS_NEW_NAME	6237	0.574777
20	OF_GENERIC_TYPE	5906	0.544273
21	RESOLVES_TO	5215	0.480593
22	SIMILAR	3977	0.366504
23	WRITES	3818	0.351851
24	CONTAINS	3814	0.351483
25	RETURNS_GENERIC	3545	0.326693
26	ANNOTATED_BY	2810	0.258958
27	REQUIRES	2174	0.200347
28	HAS_FIRST_CHILD	2144	0.197582
29	HAS_LAST_CHILD	2144	0.197582

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)

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

Pie chart values:
- 23.95% (259908)
- 23.95% (259908)
- 15.78% (171251)
- 5.64% (61217)
- 4.29% (46547)
- 3.50% (37993)
- 3.31% (35947)
- 2.34% (25444)
- 2.02% (21931)
- 1.97% (21394)
- 1.91% (20698)
- 1.58% (17093)
- 1.42% (15406)
- 1.30% (14125)
- 1.17% (12722)
- 1.16% (12578)
- 1.00% (10810)
- 0.97% (10553)
- 0.84% (9161)
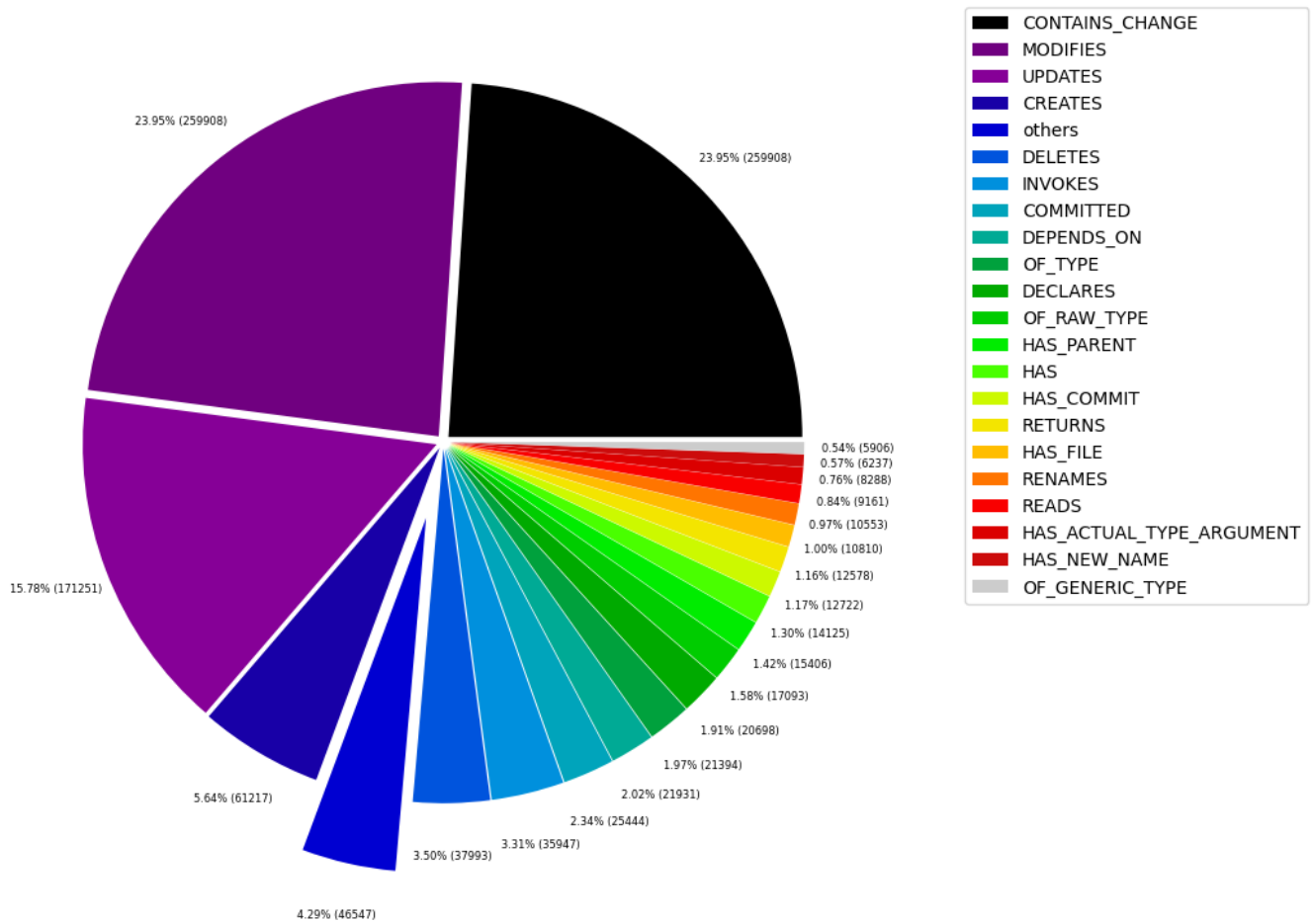- 0.76% (8288)
- 0.57% (6237)
- 0.54% (5906)

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.
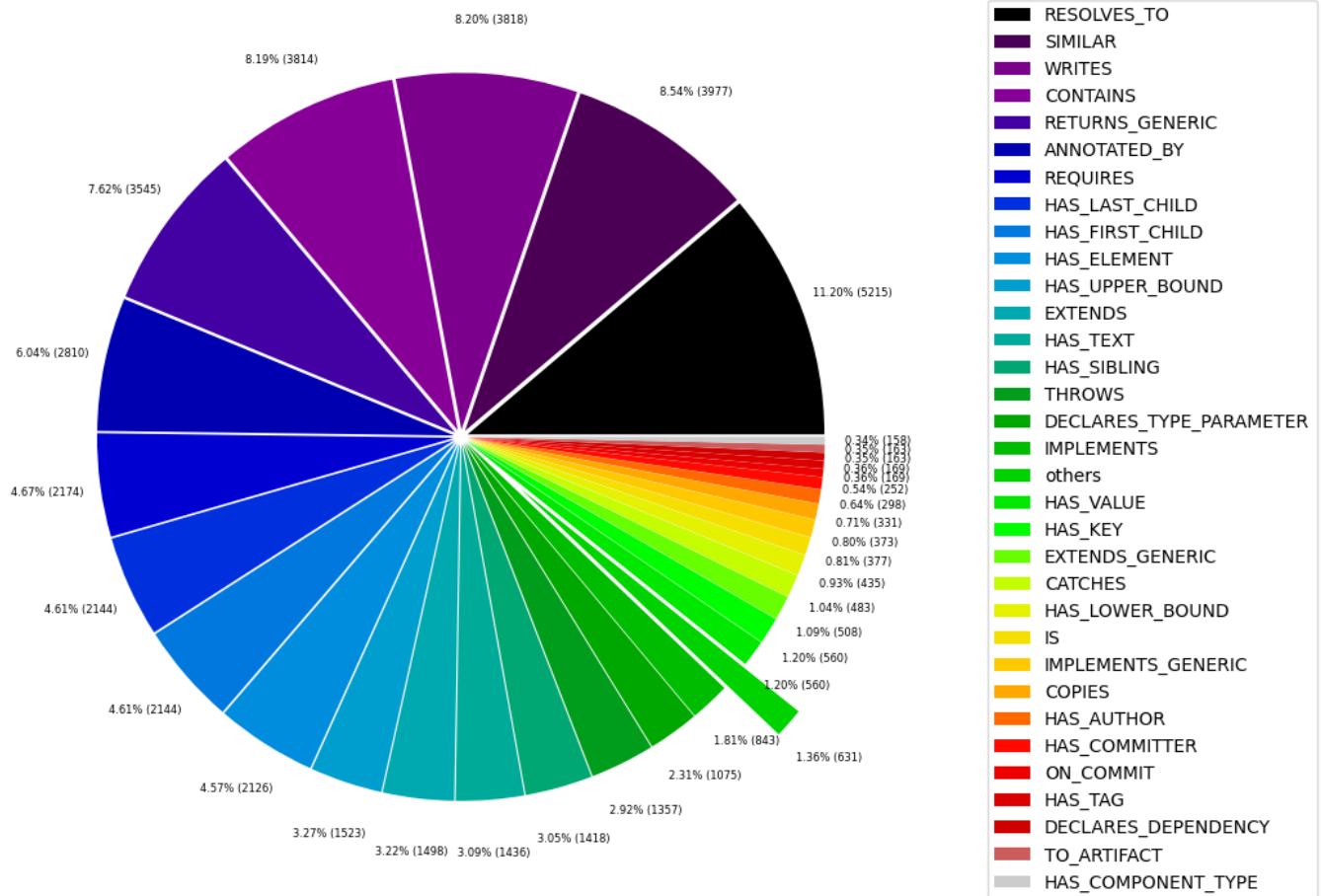
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000092
1	THROWS_GENERIC	5	0.000461
2	EXCLUDES	5	0.000461
3	DESCRIBES	9	0.000829
4	HAS_ROOT_ELEMENT	10	0.000922
5	HAS_GOAL	16	0.001474
6	HAS_EXECUTION	16	0.001474
7	OF_NAMESPACE	18	0.001659
8	HAS_ATTRIBUTE	18	0.001659
9	INCLUDES_CONCEPT	19	0.001751
10	USES_PLUGIN	21	0.001935
11	IS_ARTIFACT	21	0.001935
12	HAS_CONFIGURATION	21	0.001935
13	REQUIRES_TYPE_PARAMETER	21	0.001935
14	REQUIRES_CONCEPT	28	0.002580
15	HAS_DEFAULT	36	0.003318
16	DECLARES_NAMESPACE	36	0.003318
17	HAS_BRANCH	41	0.003778
18	HAS_HEAD	42	0.003871
19	CONTAINS_VALUE	121	0.011151
20	COPY_OF	126	0.011612
21	HAS_COMPONENT_TYPE	158	0.014561
22	DECLARES_DEPENDENCY	163	0.015021
23	TO_ARTIFACT	163	0.015021
24	ON_COMMIT	169	0.015574
25	HAS_TAG	169	0.015574
26	HAS_COMMITTER	252	0.023223
27	HAS_AUTHOR	298	0.027462
28	COPIES	331	0.030504
29	IMPLEMENTS_GENERIC	373	0.034374

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)

■	RESOLVES_TO
■	SIMILAR
■	WRITES
■	CONTAINS
■	RETURNS_GENERIC
■	ANNOTATED_BY
■	REQUIRES
■	HAS_LAST_CHILD
■	HAS_FIRST_CHILD
■	HAS_ELEMENT
■	HAS_UPPER_BOUND
■	EXTENDS
■	HAS_TEXT
■	HAS_SIBLING
■	THROWS
■	DECLARES_TYPE_PARAMETER
■	IMPLEMENTS
■	others
■	HAS_VALUE
■	HAS_KEY
■	EXTENDS_GENERIC
■	CATCHES
■	HAS_LOWER_BOUND
■	IS
■	IMPLEMENTS_GENERIC
■	COPIES
■	HAS_AUTHOR
■	HAS_COMMITTER
■	ON_COMMIT
■	HAS_TAG
■	DECLARES_DEPENDENCY
■	TO_ARTIFACT
■	HAS_COMPONENT_TYPE

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, Change]	MODIFIES	[File, Git]	259908	259908	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	259908	12722	
2	[Git, Change]	UPDATES	[File, Git]	171251	259908	
3	[Git, Change]	CREATES	[File, Git]	61217	259908	
4	[Git, Change]	DELETES	[File, Git]	37993	259908	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21999	13102	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15397	12722	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12722	1	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12722	252	
9	[Author, Git, Person]	COMMITTED	[Git, Commit]	12722	298	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10810	1	
11	[Git, Change]	RENAMES	[File, Git]	10553	259908	
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]	6237	10810	
15	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6079	13029	
16	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Method]	5090	770	
17	[Type, File, Java, ByteCode, Class]	DEPENDS_ON	[Type, File, Java, ByteCode, JavaType]	4017	770	
18	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3427	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	2039	
24	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalType, Ext...]	2341	2822	
25	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Field]	2142	770	
26	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
27	[Java, ByteCode, Member, Method, Constructor]	INVOKES	[Java, ByteCode, Member, Method, Constructor]	2119	2039	
28	[Type, File, Java, ByteCode, GenericDeclaratio...]	DECLARES	[Java, ByteCode, Member, Method]	2106	230	
29	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2080	13102	

Graph Density

total_number_of_nodes (vertices): 349419
total_number_of_relationships (edges): 1085117
-> total directed graph density: 8.887605670329396e-06
-> total directed graph density in percent: 0.0008887605670329395