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]	246598	73.539999
1	[Java, ByteCode, Member, Method]	13201	3.936778
2	[Java, ByteCode, Parameter]	13016	3.881607
3	[Git, Commit]	12102	3.609036
4	[File, Git]	10690	3.187952
5	[Java, ByteCode, ParameterizedType, Bound]	7172	2.138821
6	[Java, ByteCode, Bound]	7113	2.121226
7	[Java, ByteCode, Member, Field]	3503	1.044658
8	[Java, ByteCode, Bound, WildcardType]	2908	0.867218
9	[Java, Value, ByteCode, Annotation]	2822	0.841572
10	[Xml, Element]	2144	0.639380
11	[Java, ByteCode, Member, Constructor, Method]	2043	0.609260
12	[Xml, Text]	1436	0.428241
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.326847
14	[Java, ByteCode, Member, Method, Lambda]	938	0.279729
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	900	0.268396
16	[Type, File, Java, ByteCode, Class]	861	0.256766
17	[Type, File, Java, ByteCode, JavaType]	748	0.223067
18	[Java, Value, ByteCode, Primitive]	650	0.193842
19	[Java, ByteCode, Member, Method, GenericDeclar...	569	0.169686
20	[Json, Key]	560	0.167002
21	[Value, Json, Scalar]	544	0.162231
22	[Type, File, Java, ByteCode, ExternalType]	396	0.118094
23	[Author, Git, Person]	293	0.087378
24	[Value, Array]	267	0.079624
25	[Committer, Git, Person]	248	0.073958
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.070379
27	[Value, Property]	202	0.060240
28	[Type, File, Java, ByteCode, Interface]	187	0.055767
29	[Java, Value, ByteCode, Class]	172	0.051294

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', 'ParameterizedType', 'Bound']
- ['Java', 'ByteCode', 'Bound']
- ['Java', 'ByteCode', 'Member', 'Field']
- ['Java', 'ByteCode', 'Bound', 'WildcardType']
- ['Java', 'Value', 'ByteCode', 'Annotation']
- ['Xml', 'Element']
- ['Java', 'ByteCode', 'Member', 'Constructor', 'Method']

73.54% (246598)
0.61% (2043)
0.64% (2144)
0.84% (2822)
0.87% (2908)
1.04% (3503)
2.12% (7113)
2.14% (7172)
3.19% (10690)
3.58% (12013)
3.61% (12102)
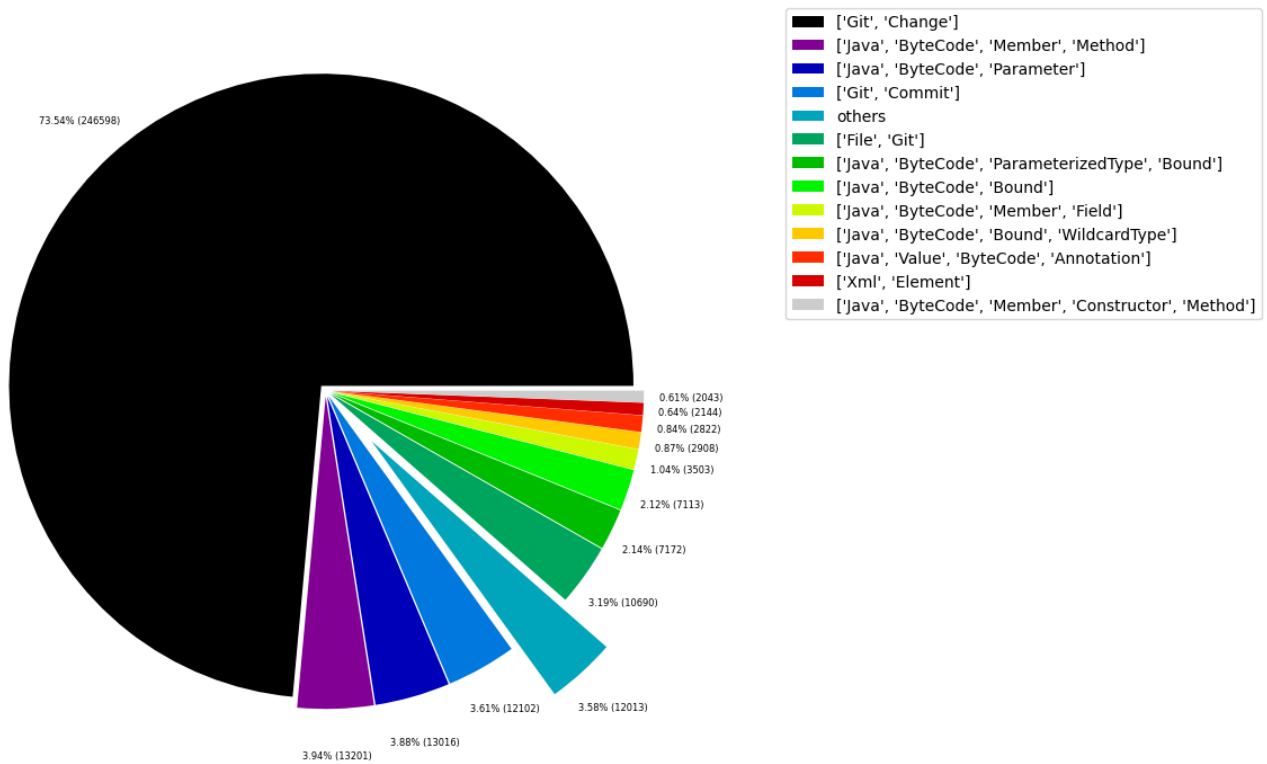3.88% (13016)
3.94% (13201)

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	0.000298
1	[Repository, File, Git]	1	0.000298
2	[File, Json]	2	0.000596
3	[File]	3	0.000895
4	[Java, ByteCode, Member, Constructor, Method, ...	4	0.001193
5	[Maven, Exclusion]	5	0.001491
6	[Value, Json, Array]	6	0.001789
7	[File, Maven, Xml, Pom, Document]	9	0.002684
8	[Type, File, Java, ByteCode, Void]	9	0.002684
9	[Java, ManifestSection]	9	0.002684
10	[File, Java, Manifest]	9	0.002684
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002684
12	[File, Java, ServiceLoader]	10	0.002982
13	[File, Java, Properties]	12	0.003579
14	[Maven, ExecutionGoal]	16	0.004771
15	[Maven, PluginExecution]	16	0.004771
16	[Xml, Attribute]	18	0.005368
17	[jQAssistant, Rule, Concept]	19	0.005666
18	[Maven, Plugin]	21	0.006263
19	[Maven, Configuration]	21	0.006263
20	[Type, File, Java, ByteCode, Enum]	28	0.008350
21	[Git, Branch]	29	0.008648
22	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008648
23	[Xml, Namespace]	36	0.010736
24	[Type, File, Java, ByteCode, Annotation]	43	0.012823
25	[File, Directory]	44	0.013122
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.025349
27	[Artifact, Maven]	92	0.027436
28	[Type, File, Java, ByteCode, ExternalAnnotatio...	93	0.027734
29	[Value, ManifestEntry]	120	0.035786

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']
- others
- ['Author', 'Git', 'Person']
- ['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']

Pie chart values: 7.81% (938), 9.12% (1096), 7.49% (900), 11.95% (1436), 7.17% (861), 6.23% (748), 5.41% (650), 4.74% (569), 4.66% (560), 4.53% (544), 3.30% (396), 2.68% (322), 2.44% (293), 2.22% (267), 2.06% (248), 1.96% (236), 1.68% (202), 1.56% (187), 1.43% (172), 1.40% (168), 1.36% (163), 1.32% (158), 1.25% (150), 1.16% (139), 1.11% (133), 1.00% (120), 0.77% (93), 0.77% (92), 0.71% (85), 0.37% (44), 0.36% (43)
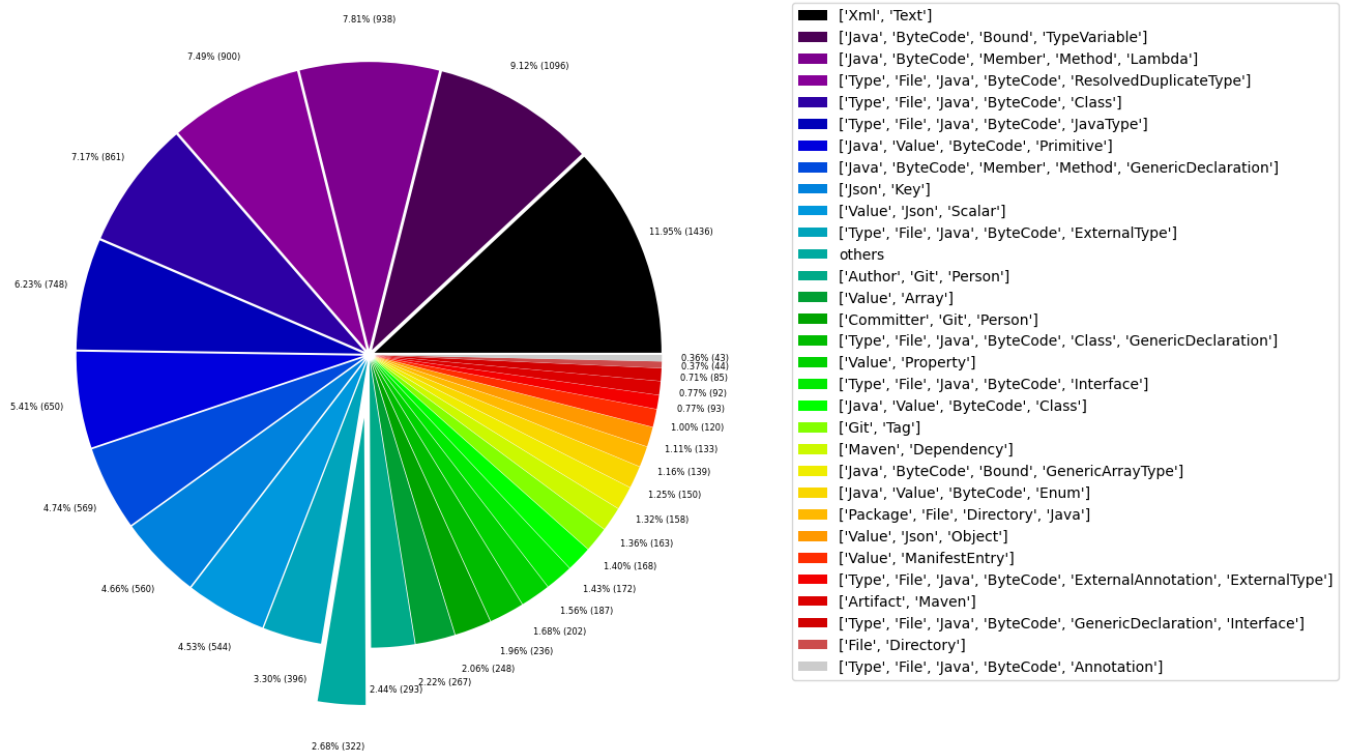
Table 1c - Highest node count by single label

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.
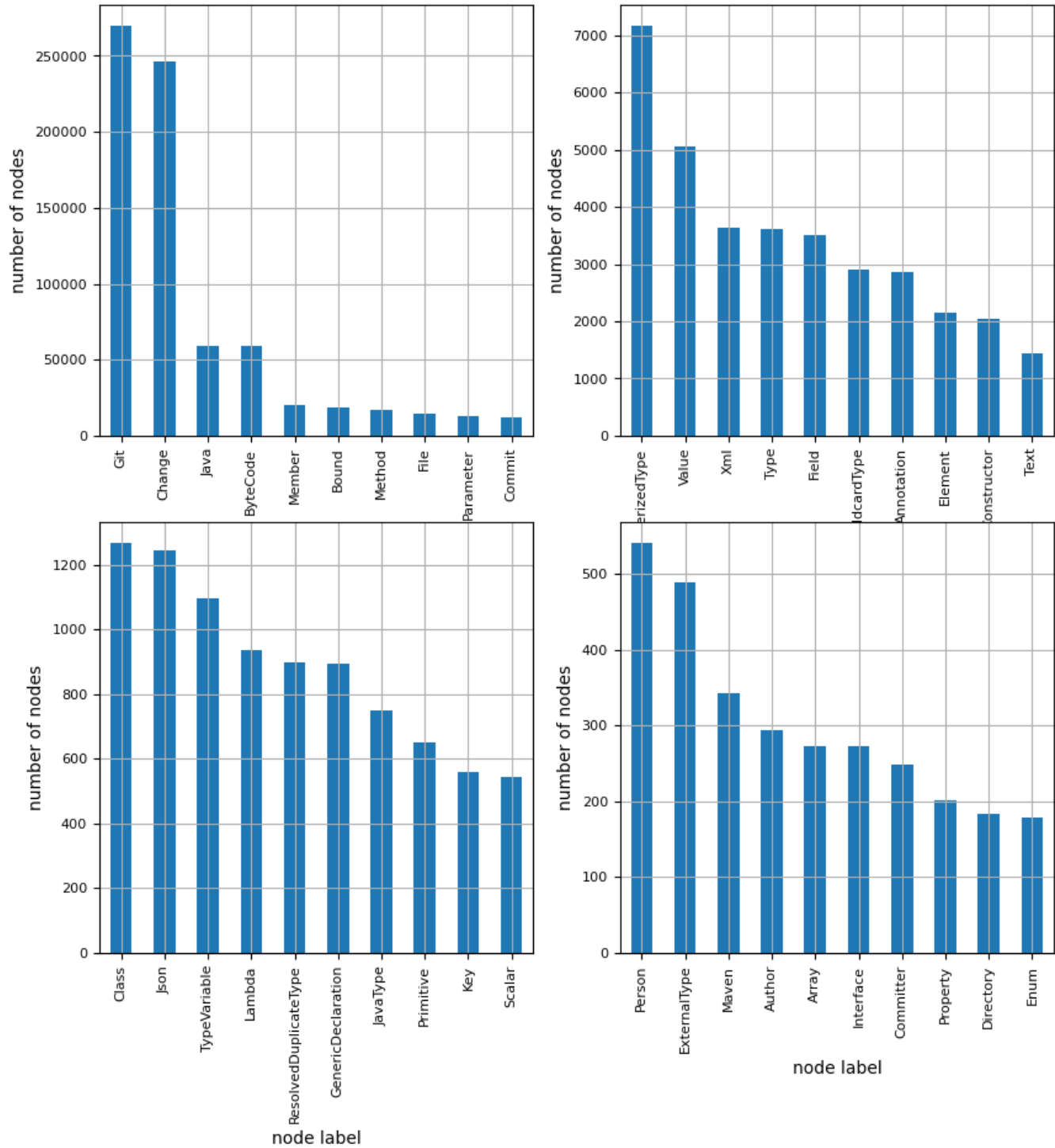
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	270129	80.557370
1	Change	246598	73.539999
2	Java	59318	17.689704
3	ByteCode	59130	17.633639
4	Member	20258	6.041303
5	Bound	18447	5.501230
6	Method	16755	4.996645
7	File	14543	4.336987
8	Parameter	13016	3.881607
9	Commit	12102	3.609036
10	ParameterizedType	7172	2.138821
11	Value	5066	1.510773
12	Xml	3643	1.086409
13	Type	3615	1.078059
14	Field	3503	1.044658
15	WildcardType	2908	0.867218
16	Annotation	2865	0.854395
17	Element	2144	0.639380
18	Constructor	2047	0.610453
19	Text	1436	0.428241
20	Class	1269	0.378439
21	Json	1245	0.371282
22	TypeVariable	1096	0.326847
23	Lambda	938	0.279729
24	ResolvedDuplicateType	900	0.268396
25	GenericDeclaration	894	0.266607
26	JavaType	748	0.223067
27	Primitive	650	0.193842
28	Key	560	0.167002
29	Scalar	544	0.162231
30	Person	541	0.161336
31	ExternalType	489	0.145829
32	Maven	343	0.102289
33	Author	293	0.087378
34	Array	273	0.081414
35	Interface	272	0.081115
36	Committer	248	0.073958
37	Property	202	0.060240
38	Directory	183	0.054574
39	Enum	178	0.053083

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: 1040249
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	246598	23.705670
1	MODIFIES	246598	23.705670
2	UPDATES	159927	15.373915
3	CREATES	59754	5.744202
4	DELETES	37207	3.576740
5	INVOKES	35679	3.429852
6	COMMITTED	24204	2.326751
7	DEPENDS_ON	21887	2.104015
8	OF_TYPE	21373	2.054604
9	DECLARES	20699	1.989812
10	OF_RAW_TYPE	17084	1.642299
11	HAS_PARENT	14710	1.414085
12	HAS	14110	1.356406
13	RETURNS	12562	1.207595
14	HAS_COMMIT	12102	1.163375
15	HAS_FILE	10690	1.027639
16	RENAMES	10290	0.989186
17	READS	9097	0.874502
18	HAS_ACTUAL_TYPE_ARGUMENT	8284	0.796348
19	HAS_NEW_NAME	6199	0.595915
20	OF_GENERIC_TYPE	5903	0.567460
21	RESOLVES_TO	5205	0.500361
22	SIMILAR	3968	0.381447
23	WRITES	3812	0.366451
24	CONTAINS	3806	0.365874
25	RETURNS_GENERIC	3543	0.340592
26	ANNOTATED_BY	2810	0.270128
27	REQUIRES	2175	0.209085
28	HAS_FIRST_CHILD	2144	0.206104
29	HAS_LAST_CHILD	2144	0.206104

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>
```
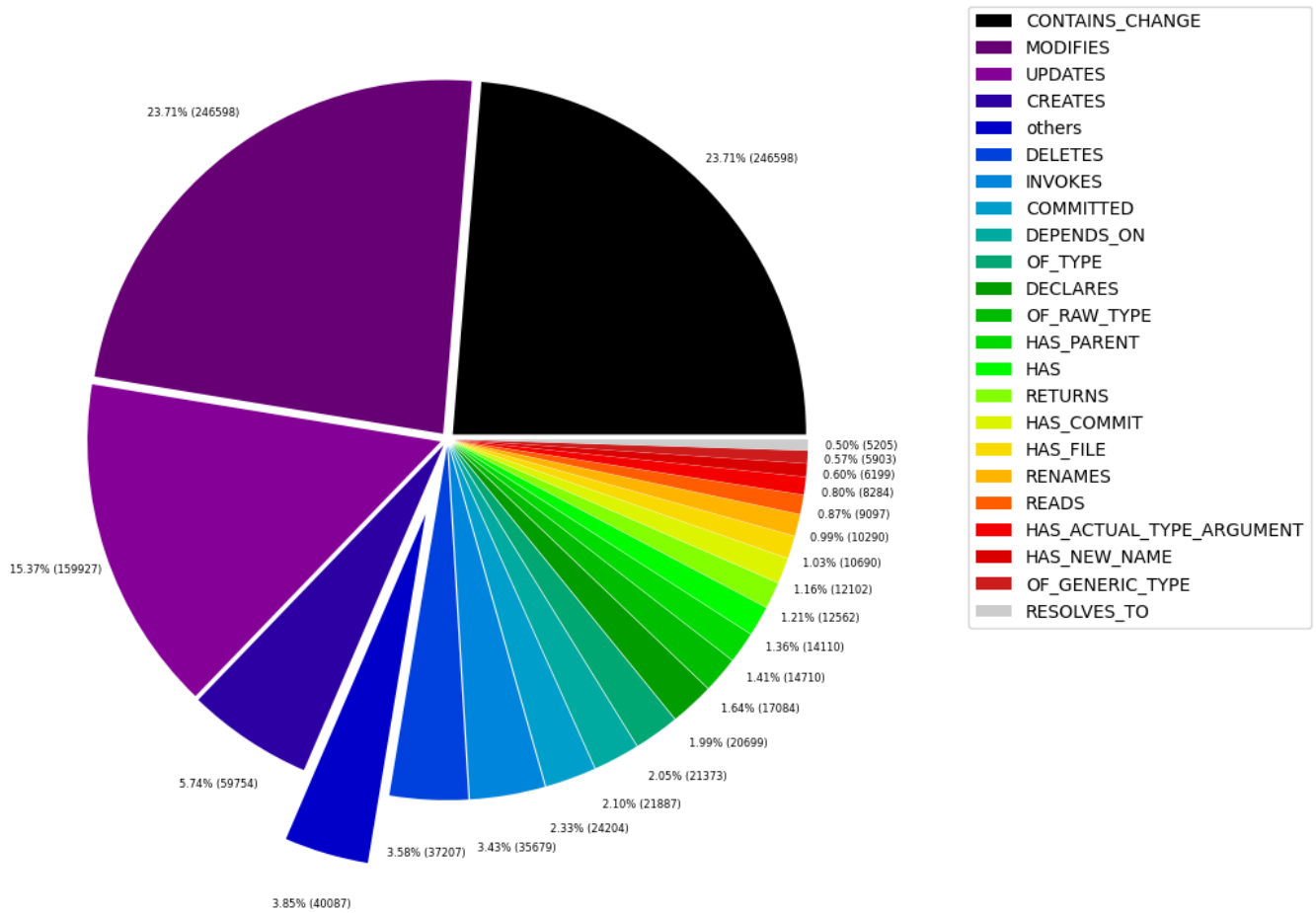
Relationship types (more than 0.5% overall)

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

Slice labels:
- 23.71% (246598)
- 23.71% (246598)
- 15.37% (159927)
- 5.74% (59754)
- 3.85% (40087)
- 3.58% (37207)
- 3.43% (35679)
- 2.33% (24204)
- 2.10% (21887)
- 2.05% (21373)
- 1.99% (20699)
- 1.64% (17084)
- 1.41% (14710)
- 1.36% (14110)
- 1.21% (12562)
- 1.16% (12102)
- 1.03% (10690)
- 0.99% (10290)
- 0.87% (9097)
- 0.80% (8284)
- 0.60% (6199)
- 0.57% (5903)
- 0.50% (5205)

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.
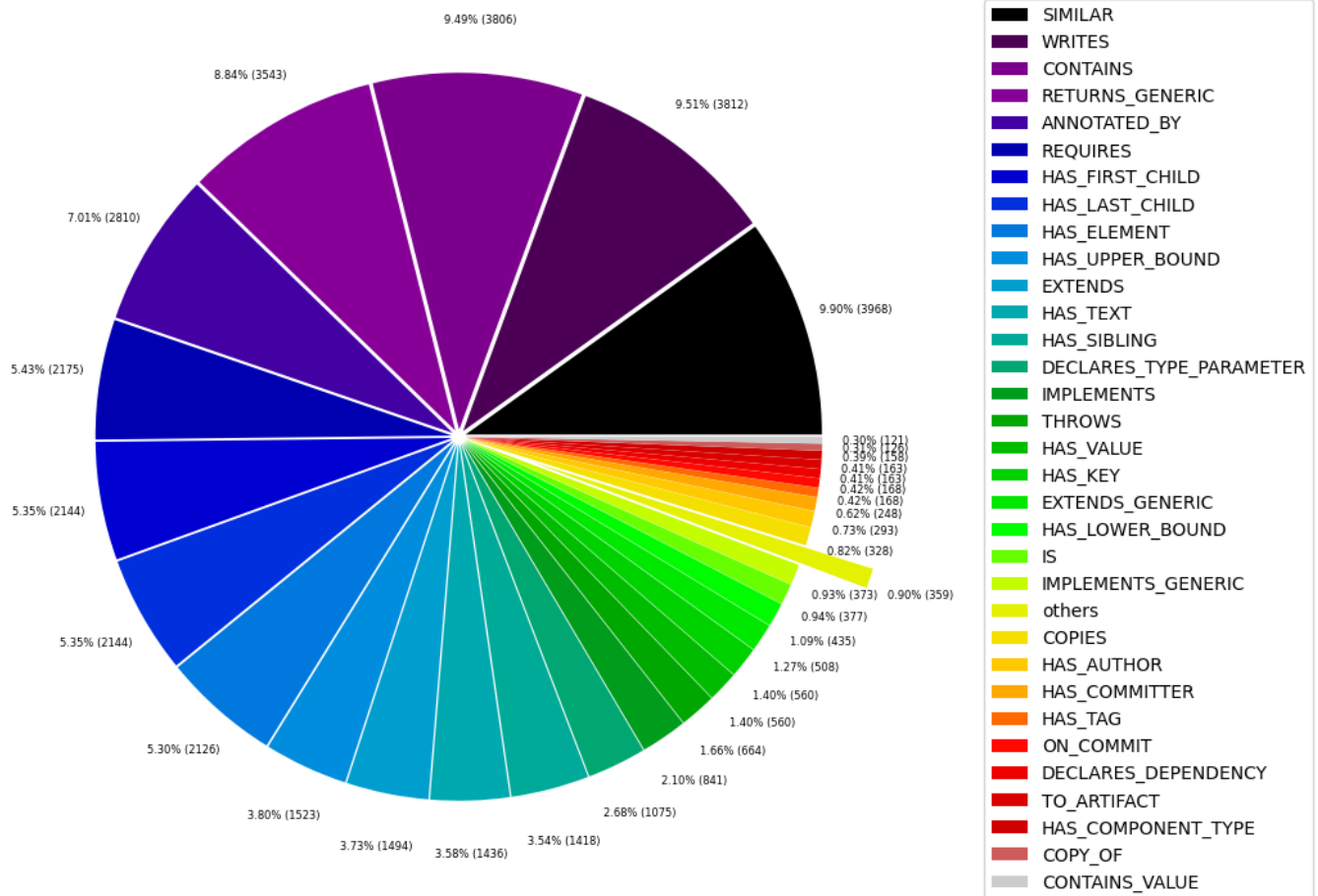
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000096
1	THROWS_GENERIC	5	0.000481
2	EXCLUDES	5	0.000481
3	HAS_ROOT_ELEMENT	9	0.000865
4	DESCRIBES	9	0.000865
5	HAS_GOAL	16	0.001538
6	HAS_EXECUTION	16	0.001538
7	OF_NAMESPACE	18	0.001730
8	HAS_ATTRIBUTE	18	0.001730
9	INCLUDES_CONCEPT	19	0.001826
10	USES_PLUGIN	21	0.002019
11	REQUIRES_TYPE_PARAMETER	21	0.002019
12	IS_ARTIFACT	21	0.002019
13	HAS_CONFIGURATION	21	0.002019
14	REQUIRES_CONCEPT	28	0.002692
15	HAS_BRANCH	29	0.002788
16	HAS_HEAD	30	0.002884
17	DECLARES_NAMESPACE	36	0.003461
18	HAS_DEFAULT	36	0.003461
19	CONTAINS_VALUE	121	0.011632
20	COPY_OF	126	0.012112
21	HAS_COMPONENT_TYPE	158	0.015189
22	TO_ARTIFACT	163	0.015669
23	DECLARES_DEPENDENCY	163	0.015669
24	ON_COMMIT	168	0.016150
25	HAS_TAG	168	0.016150
26	HAS_COMMITTER	248	0.023840
27	HAS_AUTHOR	293	0.028166
28	COPIES	328	0.031531
29	IMPLEMENTS_GENERIC	373	0.035857

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)

9.49% (3806)
8.84% (3543)
9.51% (3812)
7.01% (2810)
9.90% (3968)
5.43% (2175)
5.35% (2144)
5.35% (2144)
5.30% (2126)
3.80% (1523)
3.73% (1494)
3.58% (1436)
3.54% (1418)
2.68% (1075)
2.10% (841)
1.66% (664)
1.40% (560)
1.40% (560)
1.27% (508)
1.09% (435)
0.94% (377)
0.93% (373)
0.90% (359)
0.82% (328)
0.73% (293)
0.62% (248)
0.42% (168)
0.42% (168)
0.41% (163)
0.41% (163)
0.39% (158)
0.30% (126)
0.30% (121)

Legend:
- SIMILAR
- WRITES
- CONTAINS
- RETURNS_GENERIC
- ANNOTATED_BY
- REQUIRES
- HAS_FIRST_CHILD
- HAS_LAST_CHILD
- HAS_ELEMENT
- HAS_UPPER_BOUND
- EXTENDS
- HAS_TEXT
- HAS_SIBLING
- DECLARES_TYPE_PARAMETER
- IMPLEMENTS
- THROWS
- HAS_VALUE
- HAS_KEY
- EXTENDS_GENERIC
- HAS_LOWER_BOUND
- IS
- IMPLEMENTS_GENERIC
- others
- COPIES
- HAS_AUTHOR
- HAS_COMMITTER
- HAS_TAG
- 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]	246598	12102	
1	[Git, Change]	MODIFIES	[File, Git]	246598	246598	
2	[Git, Change]	UPDATES	[File, Git]	159927	246598	
3	[Git, Change]	CREATES	[File, Git]	59754	246598	
4	[Git, Change]	DELETES	[File, Git]	37207	246598	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21566	13076	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14701	12102	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12102	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12102	293	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12102	248	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10690	1	
11	[Git, Change]	RENAMES	[File, Git]	10290	246598	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8320	13076	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8124	13076	
14	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6368	13016	
15	[File, Git]	HAS_NEW_NAME	[File, Git]	6199	10690	
16	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3570	7113	
17	[Java, ByteCode, ParameterizedType, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3049	7172	
18	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2908	7172	
19	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, ParameterizedType, Bound]	2655	13016	
20	[Java, ByteCode, ParameterizedType, Bound]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7172	
21	[Java, ByteCode, Member, Constructor, Method]	WRITES	[Java, ByteCode, Member, Field]	2432	2043	
22	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio...]	2341	2822	
23	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
24	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2113	2043	
25	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2089	13076	
26	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2030	13076	
27	[Java, ByteCode, Member, Constructor, Method]	HAS	[Java, ByteCode, Parameter]	2023	2043	
28	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2023	13016	
29	[Java, ByteCode, Member, Method, Lambda]	INVOKES	[Java, ByteCode, Member, Method]	1879	938	

Graph Density

```
total_number_of_nodes (vertices): 335325
total_number_of_relationships (edges): 1040249
-> total directed graph density: 9.251384593875916e-06
-> total directed graph density in percent: 0.0009251384593875916
```