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]	246638	73.537892
1	[Java, ByteCode, Member, Method]	13202	3.936325
2	[Java, ByteCode, Parameter]	13016	3.880867
3	[Git, Commit]	12126	3.615503
4	[File, Git]	10692	3.187940
5	[Java, ByteCode, ParameterizedType, Bound]	7172	2.138412
6	[Java, ByteCode, Bound]	7113	2.120821
7	[Java, ByteCode, Member, Field]	3502	1.044161
8	[Java, ByteCode, Bound, WildcardType]	2908	0.867053
9	[Java, Value, ByteCode, Annotation]	2822	0.841411
10	[Xml, Element]	2144	0.639258
11	[Java, ByteCode, Member, Constructor, Method]	2040	0.608249
12	[Xml, Text]	1436	0.428160
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.326785
14	[Java, ByteCode, Member, Method, Lambda]	938	0.279675
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	900	0.268345
16	[Type, File, Java, ByteCode, Class]	861	0.256717
17	[Type, File, Java, ByteCode, JavaType]	748	0.223025
18	[Java, Value, ByteCode, Primitive]	650	0.193805
19	[Java, ByteCode, Member, Method, GenericDeclar...	569	0.169654
20	[Json, Key]	560	0.166970
21	[Value, Json, Scalar]	544	0.162200
22	[Type, File, Java, ByteCode, ExternalType]	396	0.118072
23	[Author, Git, Person]	293	0.087361
24	[Value, Array]	267	0.079609
25	[Committer, Git, Person]	248	0.073944
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.070366
27	[Value, Property]	202	0.060229
28	[Type, File, Java, ByteCode, Interface]	187	0.055756
29	[Java, Value, ByteCode, Class]	172	0.051284

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% (246638)
0.61% (2040)
0.64% (2144)
0.84% (2822)
0.87% (2908)
1.04% (3502)
2.12% (7113)
2.14% (7172)
3.19% (10692)
3.58% (12014)
3.62% (12126)
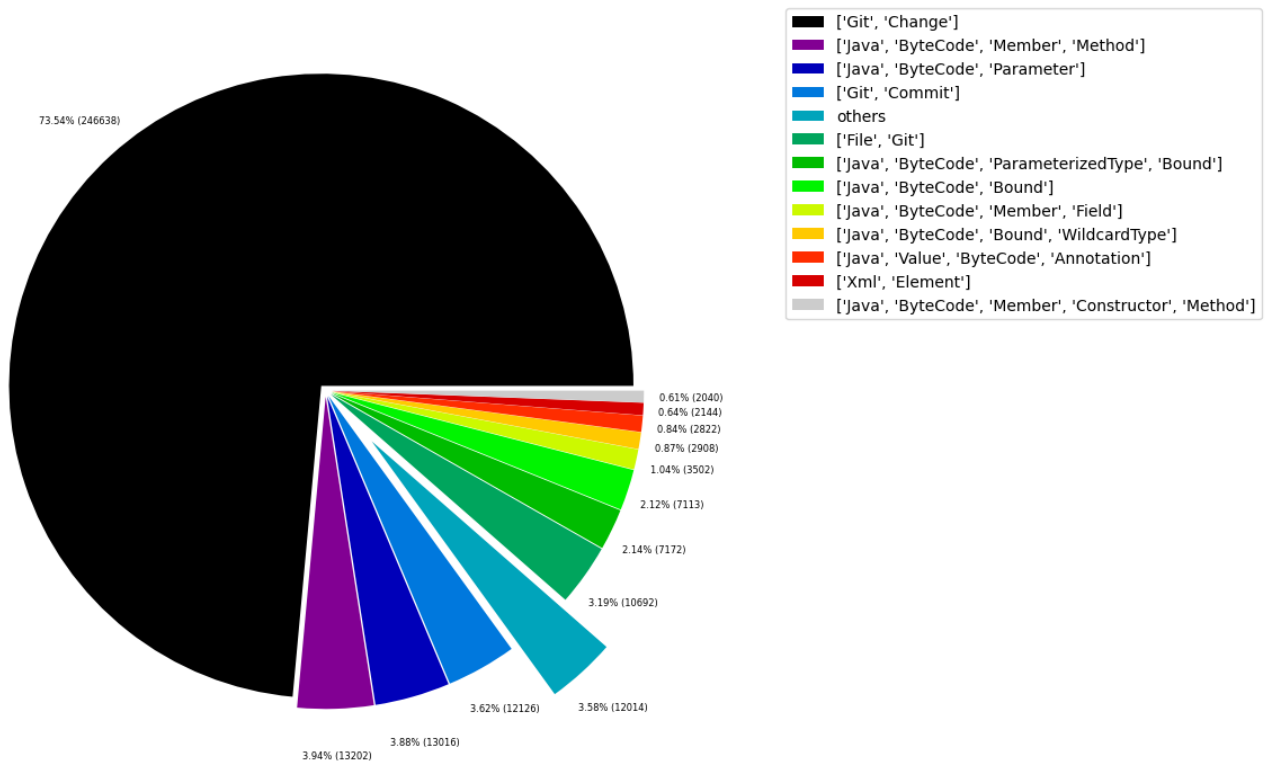3.88% (13016)
3.94% (13202)

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	0.000894
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	0.002683
8	[Type, File, Java, ByteCode, Void]	9	0.002683
9	[Java, ManifestSection]	9	0.002683
10	[File, Java, Manifest]	9	0.002683
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002683
12	[File, Java, ServiceLoader]	10	0.002982
13	[File, Java, Properties]	12	0.003578
14	[Maven, ExecutionGoal]	16	0.004771
15	[Maven, PluginExecution]	16	0.004771
16	[Xml, Attribute]	18	0.005367
17	[jQAssistant, Rule, Concept]	19	0.005665
18	[Maven, Plugin]	21	0.006261
19	[Maven, Configuration]	21	0.006261
20	[Type, File, Java, ByteCode, Enum]	28	0.008349
21	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008647
22	[Git, Branch]	30	0.008945
23	[Xml, Namespace]	36	0.010734
24	[Type, File, Java, ByteCode, Annotation]	43	0.012821
25	[File, Directory]	44	0.013119
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.025344
27	[Artifact, Maven]	92	0.027431
28	[Type, File, Java, ByteCode, ExternalAnnotatio...	93	0.027729
29	[Value, ManifestEntry]	120	0.035779

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:
- 11.95% (1436)
- 9.12% (1096)
- 7.81% (938)
- 7.49% (900)
- 7.17% (861)
- 6.23% (748)
- 5.41% (650)
- 4.74% (569)
- 4.66% (560)
- 4.53% (544)
- 3.30% (396)
- 2.69% (323)
- 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)
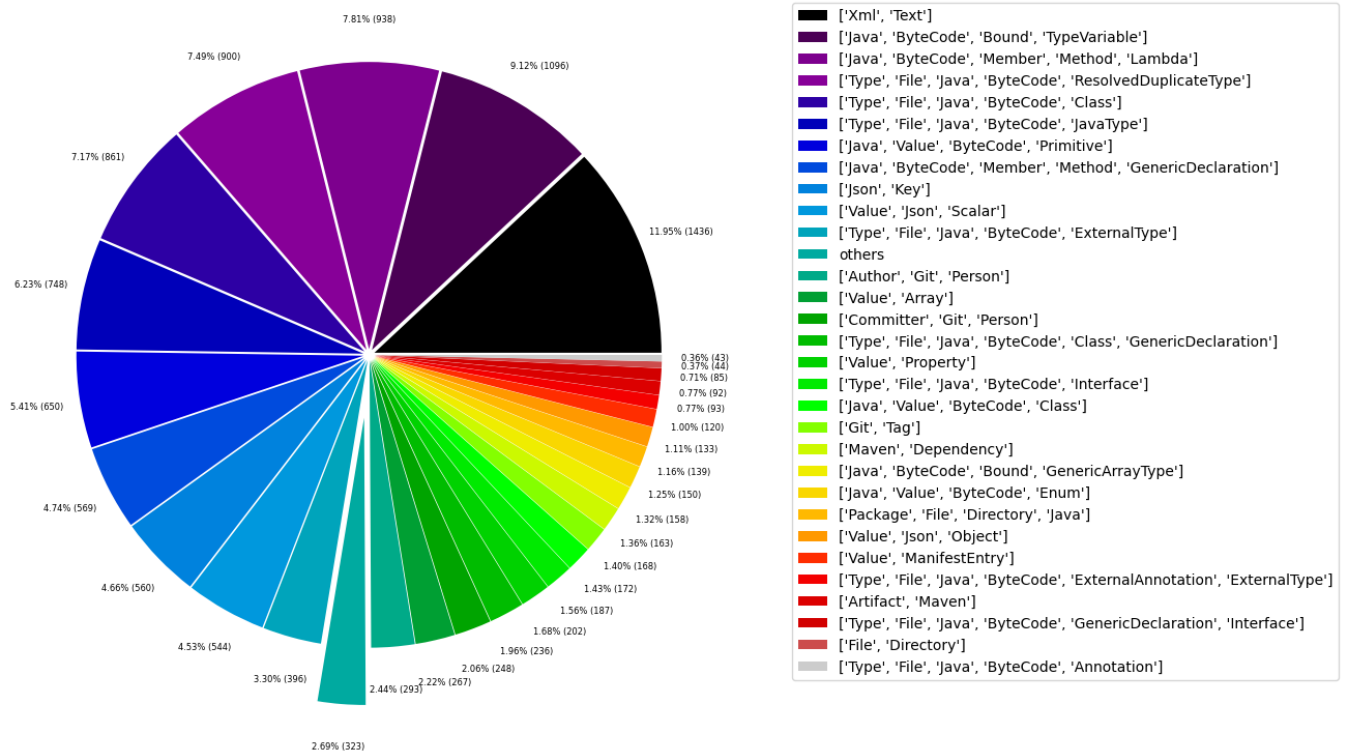- 0.77% (92)
- 0.71% (85)
- 0.37% (44)
- 0.36% (43)

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.
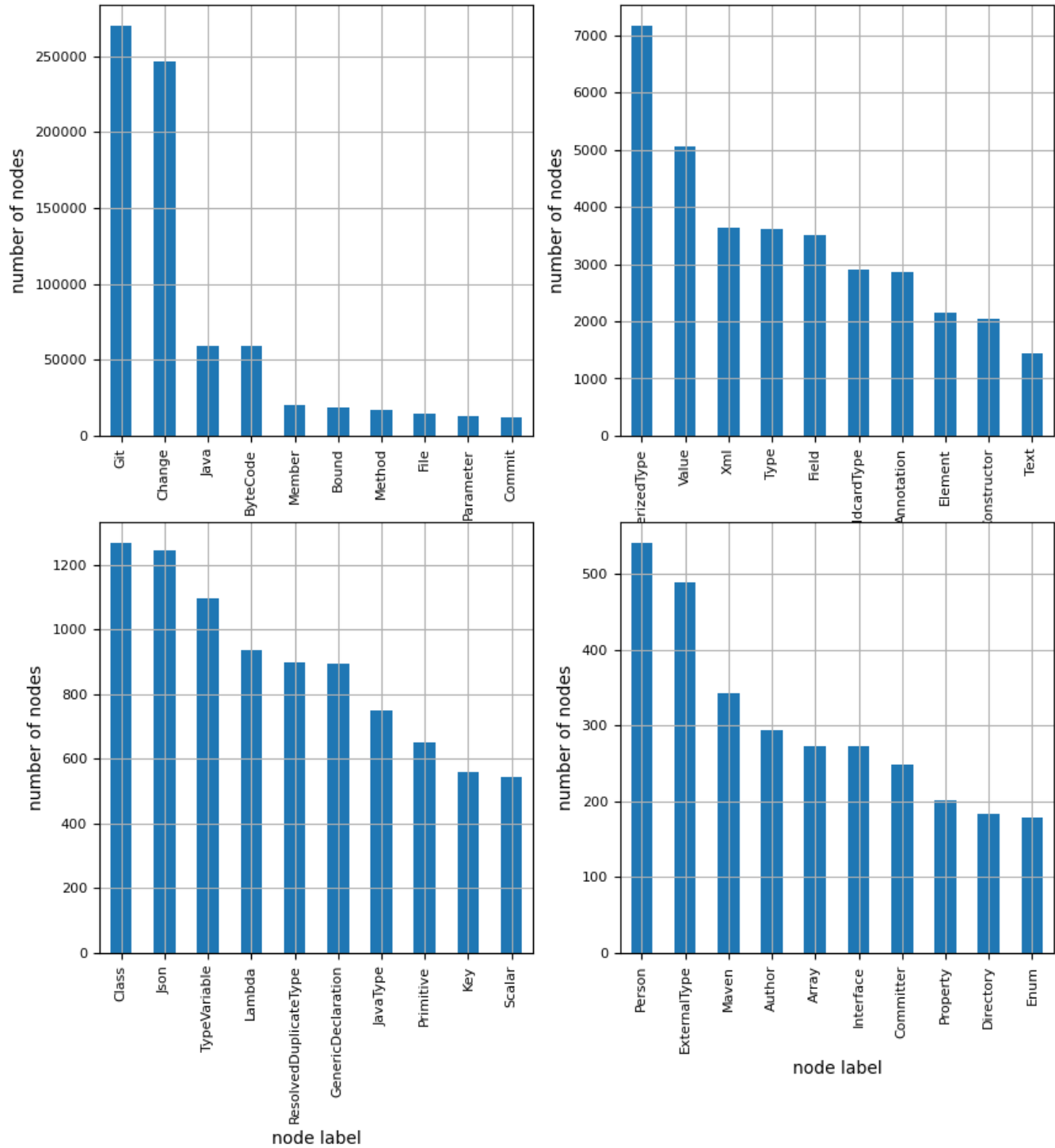
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	270196	80.561974
1	Change	246638	73.537892
2	Java	59315	17.685434
3	ByteCode	59127	17.629380
4	Member	20255	6.039256
5	Bound	18447	5.500180
6	Method	16753	4.995095
7	File	14545	4.336755
8	Parameter	13016	3.880867
9	Commit	12126	3.615503
10	ParameterizedType	7172	2.138412
11	Value	5066	1.510485
12	Xml	3643	1.086201
13	Type	3615	1.077853
14	Field	3502	1.044161
15	WildcardType	2908	0.867053
16	Annotation	2865	0.854232
17	Element	2144	0.639258
18	Constructor	2044	0.609442
19	Text	1436	0.428160
20	Class	1269	0.378367
21	Json	1245	0.371211
22	TypeVariable	1096	0.326785
23	Lambda	938	0.279675
24	ResolvedDuplicateType	900	0.268345
25	GenericDeclaration	894	0.266556
26	JavaType	748	0.223025
27	Primitive	650	0.193805
28	Key	560	0.166970
29	Scalar	544	0.162200
30	Person	541	0.161305
31	ExternalType	489	0.145801
32	Maven	343	0.102269
33	Author	293	0.087361
34	Array	273	0.081398
35	Interface	272	0.081100
36	Committer	248	0.073944
37	Property	202	0.060229
38	Directory	183	0.054564
39	Enum	178	0.053073

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: 1040449
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	246638	23.704958
1	MODIFIES	246638	23.704958
2	UPDATES	159960	15.374132
3	CREATES	59756	5.743290
4	DELETES	37212	3.576533
5	INVOKES	35669	3.428231
6	COMMITTED	24252	2.330917
7	DEPENDS_ON	21887	2.103611
8	OF_TYPE	21373	2.054209
9	DECLARES	20696	1.989141
10	OF_RAW_TYPE	17084	1.641983
11	HAS_PARENT	14734	1.416119
12	HAS	14110	1.356145
13	RETURNS	12562	1.207363
14	HAS_COMMIT	12126	1.165458
15	HAS_FILE	10692	1.027633
16	RENAMES	10290	0.988996
17	READS	9097	0.874334
18	HAS_ACTUAL_TYPE_ARGUMENT	8284	0.796195
19	HAS_NEW_NAME	6199	0.595800
20	OF_GENERIC_TYPE	5903	0.567351
21	RESOLVES_TO	5198	0.499592
22	SIMILAR	3968	0.381374
23	WRITES	3812	0.366380
24	CONTAINS	3806	0.365804
25	RETURNS_GENERIC	3543	0.340526
26	ANNOTATED_BY	2810	0.270076
27	REQUIRES	2175	0.209044
28	HAS_FIRST_CHILD	2144	0.206065
29	HAS_LAST_CHILD	2144	0.206065

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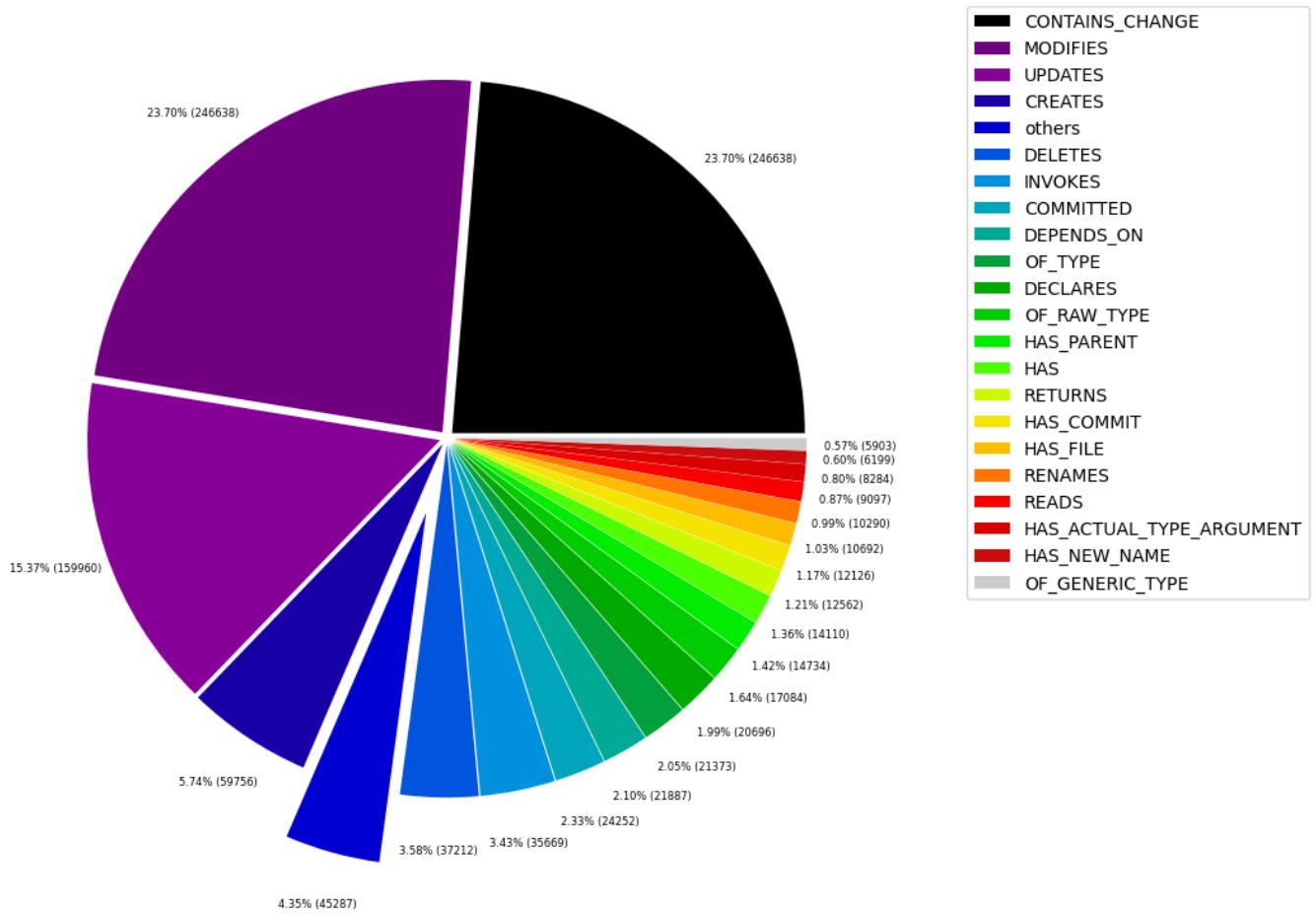
Relationship types (more than 0.5% overall)

CONTAINS_CHANGE
MODIFIES
UPDATES
CREATES
others
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

23.70% (246638)
23.70% (246638)
15.37% (159960)
5.74% (59756)
4.35% (45287)
3.58% (37212)
3.43% (35669)
2.33% (24252)
2.10% (21887)
2.05% (21373)
1.99% (20696)
1.64% (17084)
1.42% (14734)
1.36% (14110)
1.21% (12562)
1.17% (12126)
1.03% (10692)
0.99% (10290)
0.87% (9097)
0.80% (8284)
0.60% (6199)
0.57% (5903)

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.
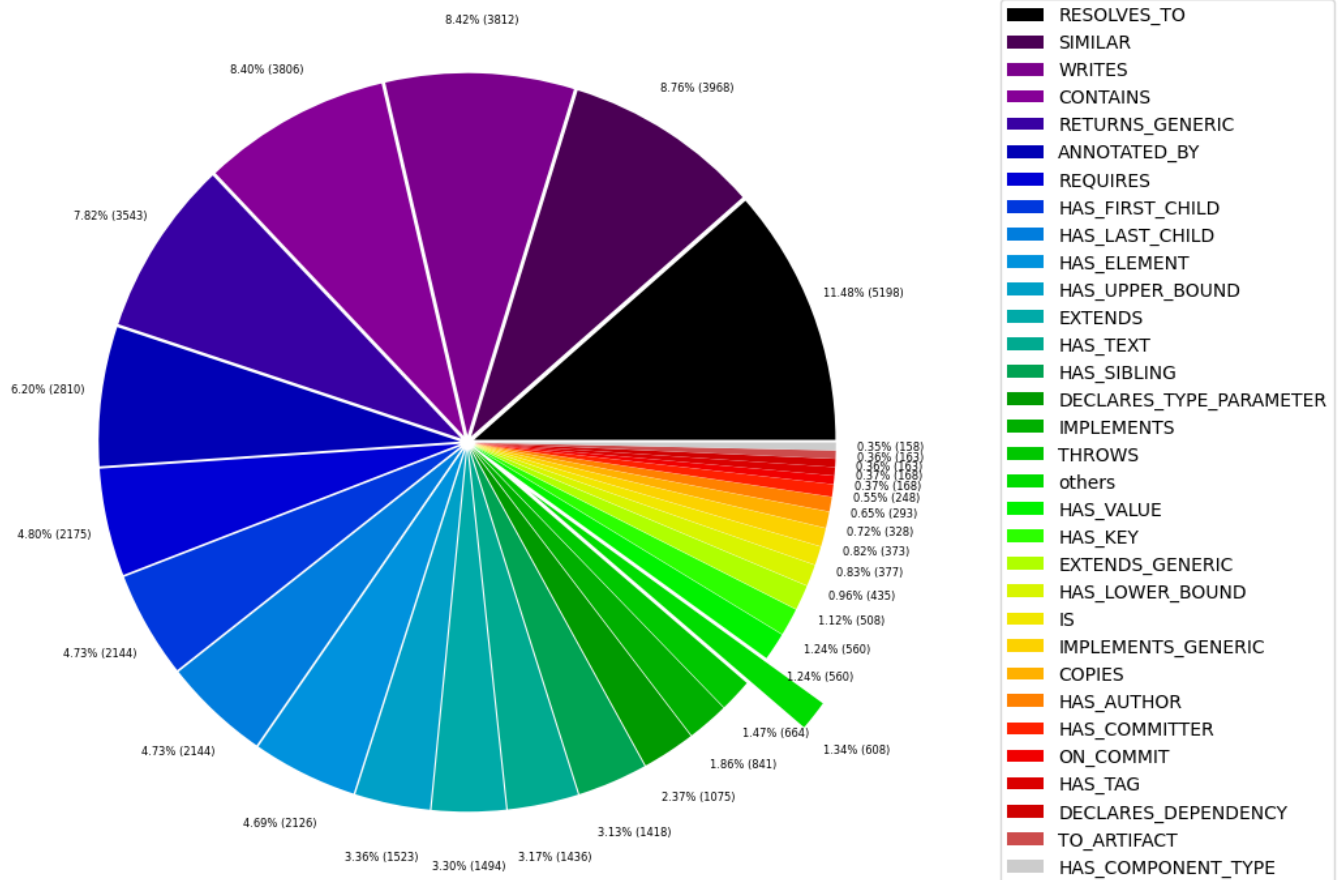
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000096
1	THROWS_GENERIC	5	0.000481
2	EXCLUDES	5	0.000481
3	HAS_ROOT_ELEMENT	9	0.000865
4	DESCRIBES	9	0.000865
5	HAS_GOAL	16	0.001538
6	HAS_EXECUTION	16	0.001538
7	OF_NAMESPACE	18	0.001730
8	HAS_ATTRIBUTE	18	0.001730
9	INCLUDES_CONCEPT	19	0.001826
10	USES_PLUGIN	21	0.002018
11	REQUIRES_TYPE_PARAMETER	21	0.002018
12	IS_ARTIFACT	21	0.002018
13	HAS_CONFIGURATION	21	0.002018
14	REQUIRES_CONCEPT	28	0.002691
15	HAS_BRANCH	30	0.002883
16	HAS_HEAD	31	0.002979
17	DECLARES_NAMESPACE	36	0.003460
18	HAS_DEFAULT	36	0.003460
19	CONTAINS_VALUE	121	0.011630
20	COPY_OF	126	0.012110
21	HAS_COMPONENT_TYPE	158	0.015186
22	TO_ARTIFACT	163	0.015666
23	DECLARES_DEPENDENCY	163	0.015666
24	ON_COMMIT	168	0.016147
25	HAS_TAG	168	0.016147
26	HAS_COMMITTER	248	0.023836
27	HAS_AUTHOR	293	0.028161
28	COPIES	328	0.031525
29	IMPLEMENTS_GENERIC	373	0.035850

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)

Pie chart slice labels:
- 8.76% (3968)
- 8.42% (3812)
- 8.40% (3806)
- 7.82% (3543)
- 11.48% (5198)
- 6.20% (2810)
- 4.80% (2175)
- 4.73% (2144)
- 4.73% (2144)
- 4.69% (2126)
- 3.36% (1523)
- 3.30% (1494)
- 3.17% (1436)
- 3.13% (1418)
- 2.37% (1075)
- 1.86% (841)
- 1.47% (664)
- 1.34% (608)
- 1.24% (560)
- 1.24% (560)
- 1.12% (508)
- 0.96% (435)
- 0.83% (377)
- 0.82% (373)
- 0.72% (328)
- 0.65% (293)
- 0.55% (248)
- 0.37% (168)
- 0.37% (168)
- 0.36% (163)
- 0.36% (163)
- 0.35% (158)

Legend:
- RESOLVES_TO
- 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
- others
- HAS_VALUE
- HAS_KEY
- EXTENDS_GENERIC
- 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, Commit]	CONTAINS_CHANGE	[Git, Change]	246638	12126	
1	[Git, Change]	MODIFIES	[File, Git]	246638	246638	
2	[Git, Change]	UPDATES	[File, Git]	159960	246638	
3	[Git, Change]	CREATES	[File, Git]	59756	246638	
4	[Git, Change]	DELETES	[File, Git]	37212	246638	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21554	13076	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14725	12126	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12126	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12126	293	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12126	248	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10692	1	
11	[Git, Change]	RENAMES	[File, Git]	10290	246638	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8320	13076	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8121	13076	
14	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6368	13016	
15	[File, Git]	HAS_NEW_NAME	[File, Git]	6199	10692	
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	2040	
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]	2109	2040	
25	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2088	13076	
26	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2029	13076	
27	[Java, ByteCode, Member, Constructor, Method]	HAS	[Java, ByteCode, Parameter]	2023	2040	
28	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2023	13016	
29	[Java, ByteCode, Member, Method, Lambda]	INVOKES	[Java, ByteCode, Member, Method]	1878	938	

Graph Density

```
total_number_of_nodes (vertices): 335389
total_number_of_relationships (edges): 1040449
-> total directed graph density: 9.249632175783656e-06
-> total directed graph density in percent: 0.0009249632175783656
```