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]	259914	74.375256
1	[Java, ByteCode, Member, Method]	13197	3.776365
2	[Java, ByteCode, Parameter]	13029	3.728292
3	[Git, Commit]	12727	3.641873
4	[File, Git]	10810	3.093317
5	[Java, ByteCode, Bound, ParameterizedType]	7176	2.053436
6	[Java, ByteCode, Bound]	7118	2.036839
7	[Java, ByteCode, Member, Field]	3502	1.002109
8	[Java, ByteCode, Bound, WildcardType]	2908	0.832134
9	[Java, Value, ByteCode, Annotation]	2822	0.807525
10	[Xml, Element]	2144	0.613513
11	[Java, ByteCode, Member, Method, Constructor]	2039	0.583467
12	[Xml, Text]	1436	0.410916
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.313624
14	[Java, ByteCode, Member, Method, Lambda]	942	0.269556
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	877	0.250956
16	[Type, File, Java, ByteCode, Class]	810	0.231784
17	[Java, Value, ByteCode, Primitive]	652	0.186572
18	[Type, File, Java, ByteCode, JavaType]	644	0.184283
19	[Java, ByteCode, Member, Method, GenericDeclar...	569	0.162821
20	[Json, Key]	568	0.162535
21	[Value, Json, Scalar]	551	0.157670
22	[Type, File, Java, ByteCode, ExternalType]	380	0.108738
23	[Author, Git, Person]	298	0.085274
24	[Value, Array]	267	0.076403
25	[Committer, Git, Person]	252	0.072111
26	[Type, File, Java, ByteCode, GenericDeclaratio...	236	0.067532
27	[Value, Property]	202	0.057803
28	[Type, File, Java, ByteCode, Interface]	188	0.053797
29	[Java, Value, ByteCode, Class]	172	0.049218

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']

74.38% (259914)

0.58% (2039)
0.61% (2144)
0.81% (2822)
0.83% (2908)
1.00% (3502)
2.04% (7118)
2.05% (7176)
3.09% (10810)
3.46% (12077)
3.64% (12727)
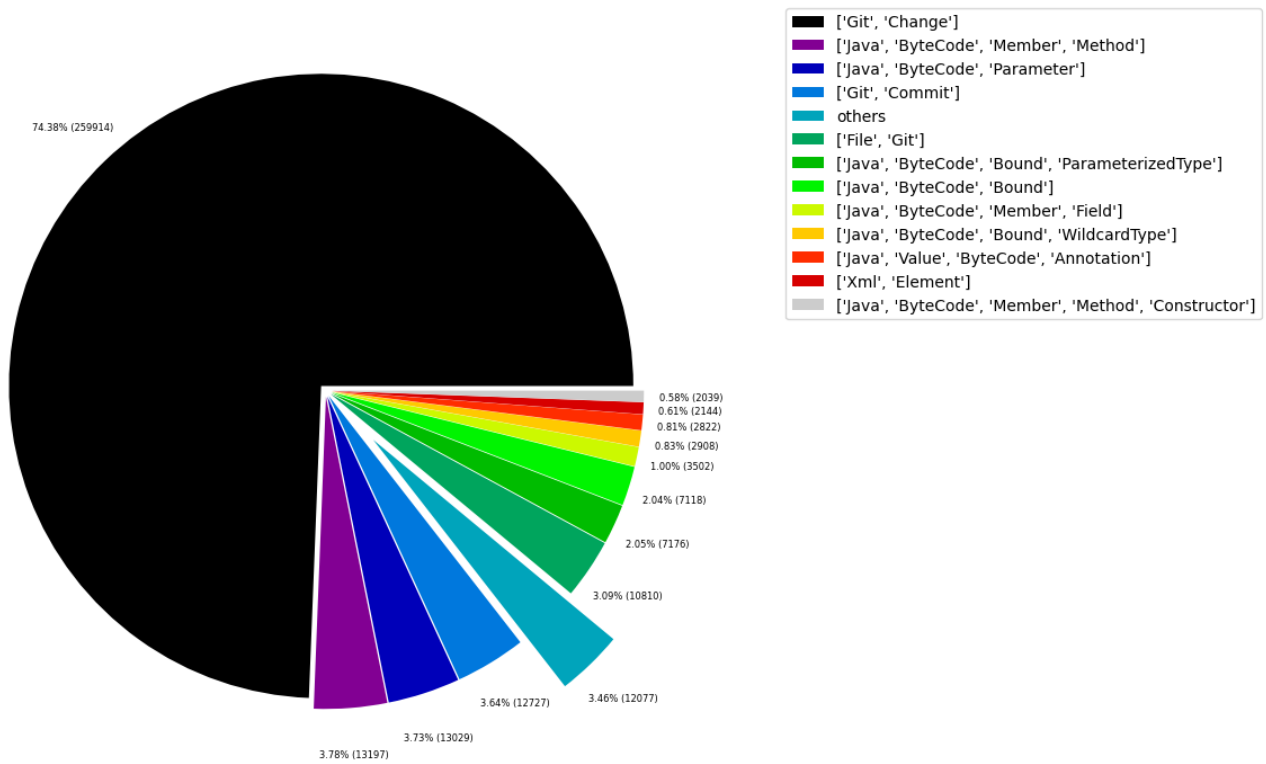3.73% (13029)
3.78% (13197)

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	[Package, File, Json, NPM]	1	0.000286
2	[Repository, File, Git]	1	0.000286
3	[File, TS, Scan]	1	0.000286
4	[File, Json]	2	0.000572
5	[File]	3	0.000858
6	[Java, ByteCode, Member, Method, GenericDeclar...	4	0.001145
7	[Maven, Exclusion]	5	0.001431
8	[Value, Array, Json]	6	0.001717
9	[Dependency, NPM]	7	0.002003
10	[Type, File, Java, ByteCode, Void]	9	0.002575
11	[File, Maven, Xml, Pom, Document]	9	0.002575
12	[Java, ManifestSection]	9	0.002575
13	[File, Java, Manifest]	9	0.002575
14	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002575
15	[File, Java, ServiceLoader]	10	0.002862
16	[File, Java, Properties]	12	0.003434
17	[Maven, PluginExecution]	16	0.004578
18	[Maven, ExecutionGoal]	16	0.004578
19	[Type, File, Java, ByteCode, Throwable, Extern...	16	0.004578
20	[Xml, Attribute]	18	0.005151
21	[jQAssistant, Rule, Concept]	19	0.005437
22	[Maven, Configuration]	21	0.006009
23	[Maven, Plugin]	21	0.006009
24	[Type, File, Java, ByteCode, Throwable, Resolv...	22	0.006295
25	[Type, File, Java, ByteCode, Enum]	28	0.008012
26	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008298
27	[Xml, Namespace]	36	0.010302
28	[Type, File, Java, ByteCode, Annotation]	43	0.012305
29	[Git, Branch]	44	0.012591

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']
- ['Git', 'Branch']
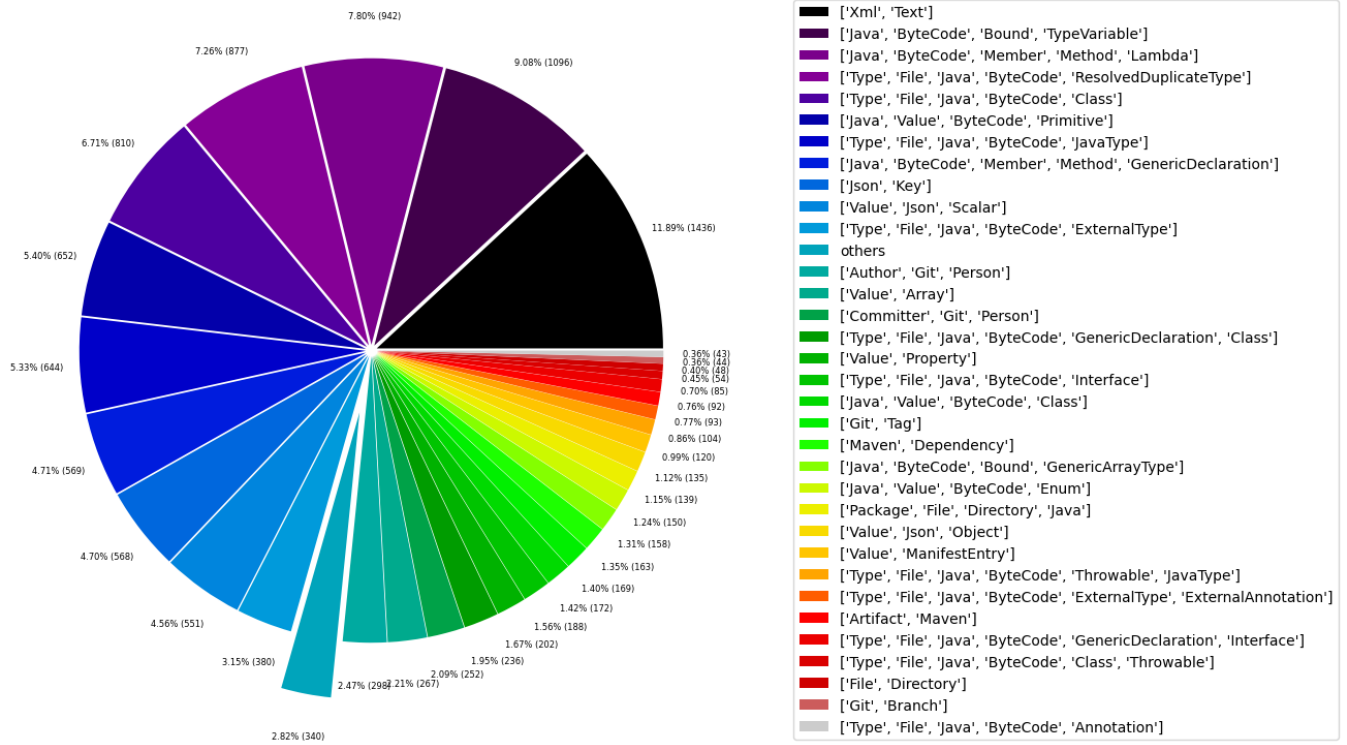- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']

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.
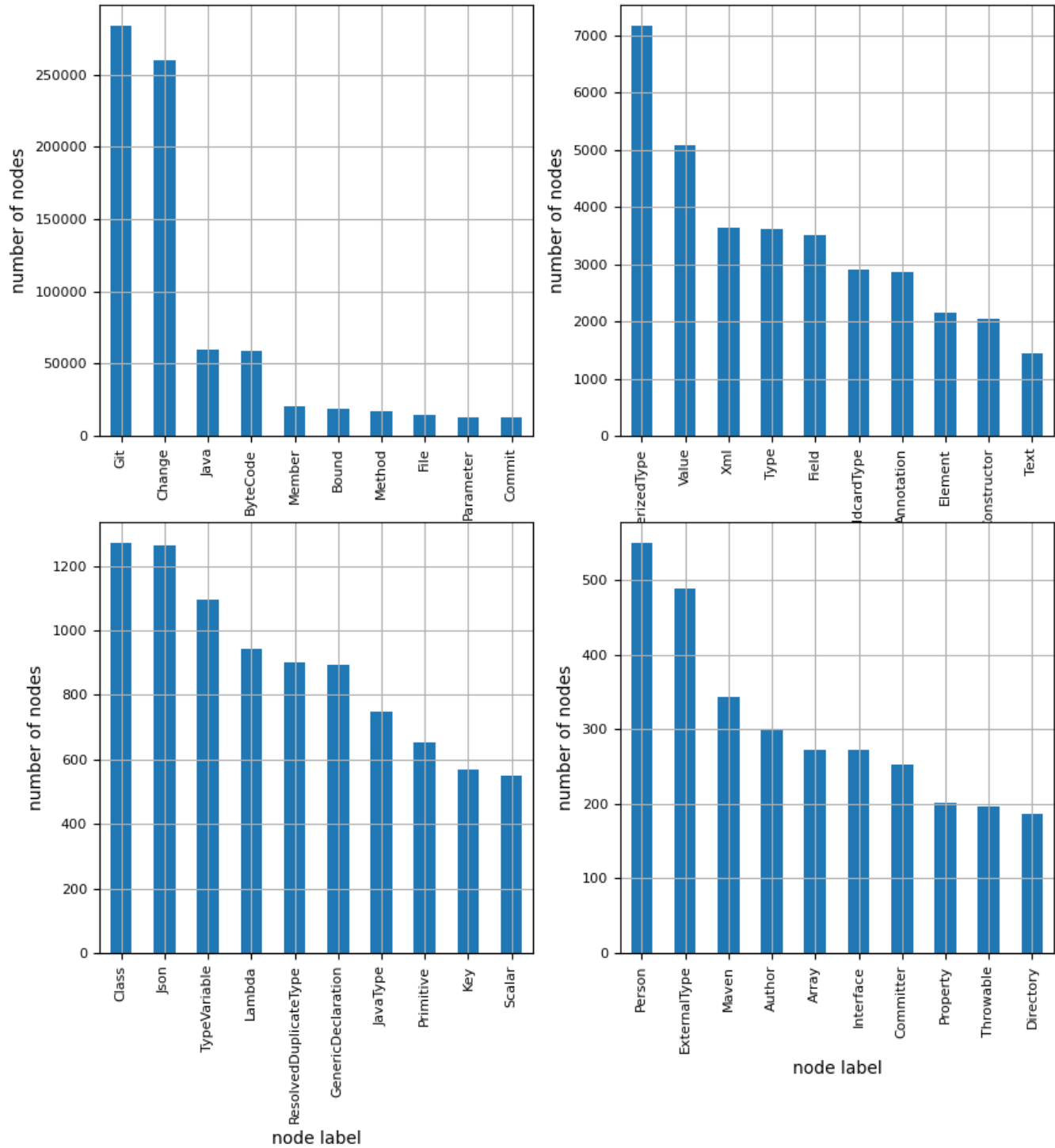
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	284215	81.329068
1	Change	259914	74.375256
2	Java	59340	16.980338
3	ByteCode	59152	16.926542
4	Member	20253	5.795463
5	Bound	18456	5.281246
6	Method	16751	4.793354
7	File	14672	4.198442
8	Parameter	13029	3.728292
9	Commit	12727	3.641873
10	ParameterizedType	7176	2.053436
11	Value	5077	1.452800
12	Xml	3643	1.042457
13	Type	3618	1.035303
14	Field	3502	1.002109
15	WildcardType	2908	0.832134
16	Annotation	2865	0.819829
17	Element	2144	0.613513
18	Constructor	2043	0.584611
19	Text	1436	0.410916
20	Class	1272	0.363987
21	Json	1263	0.361412
22	TypeVariable	1096	0.313624
23	Lambda	942	0.269556
24	ResolvedDuplicateType	899	0.257252
25	GenericDeclaration	894	0.255821
26	JavaType	748	0.214043
27	Primitive	652	0.186572
28	Key	568	0.162535
29	Scalar	551	0.157670
30	Person	550	0.157384
31	ExternalType	489	0.139929
32	Maven	343	0.098151
33	Author	298	0.085274
34	Array	273	0.078120
35	Interface	273	0.078120
36	Committer	252	0.072111
37	Property	202	0.057803
38	Throwable	196	0.056086
39	Directory	187	0.053511

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: 1085195
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	259914	23.950903
1	MODIFIES	259914	23.950903
2	UPDATES	171257	15.781219
3	CREATES	61217	5.641106
4	DELETES	37993	3.501030
5	INVOKES	35947	3.312492
6	COMMITTED	25454	2.345569
7	DEPENDS_ON	21931	2.020927
8	OF_TYPE	21394	1.971443
9	DECLARES	20698	1.907307
10	OF_RAW_TYPE	17093	1.575109
11	HAS_PARENT	15411	1.420113
12	HAS	14125	1.301609
13	HAS_COMMIT	12727	1.172785
14	RETURNS	12578	1.159054
15	HAS_FILE	10810	0.996134
16	RENAMES	10553	0.972452
17	READS	9161	0.844180
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.763734
19	HAS_NEW_NAME	6237	0.574735
20	OF_GENERIC_TYPE	5906	0.544234
21	RESOLVES_TO	5216	0.480651
22	SIMILAR	3977	0.366478
23	CONTAINS	3824	0.352379
24	WRITES	3818	0.351826
25	RETURNS_GENERIC	3545	0.326669
26	ANNOTATED_BY	2810	0.258940
27	REQUIRES	2174	0.200333
28	HAS_FIRST_CHILD	2144	0.197568
29	HAS_LAST_CHILD	2144	0.197568

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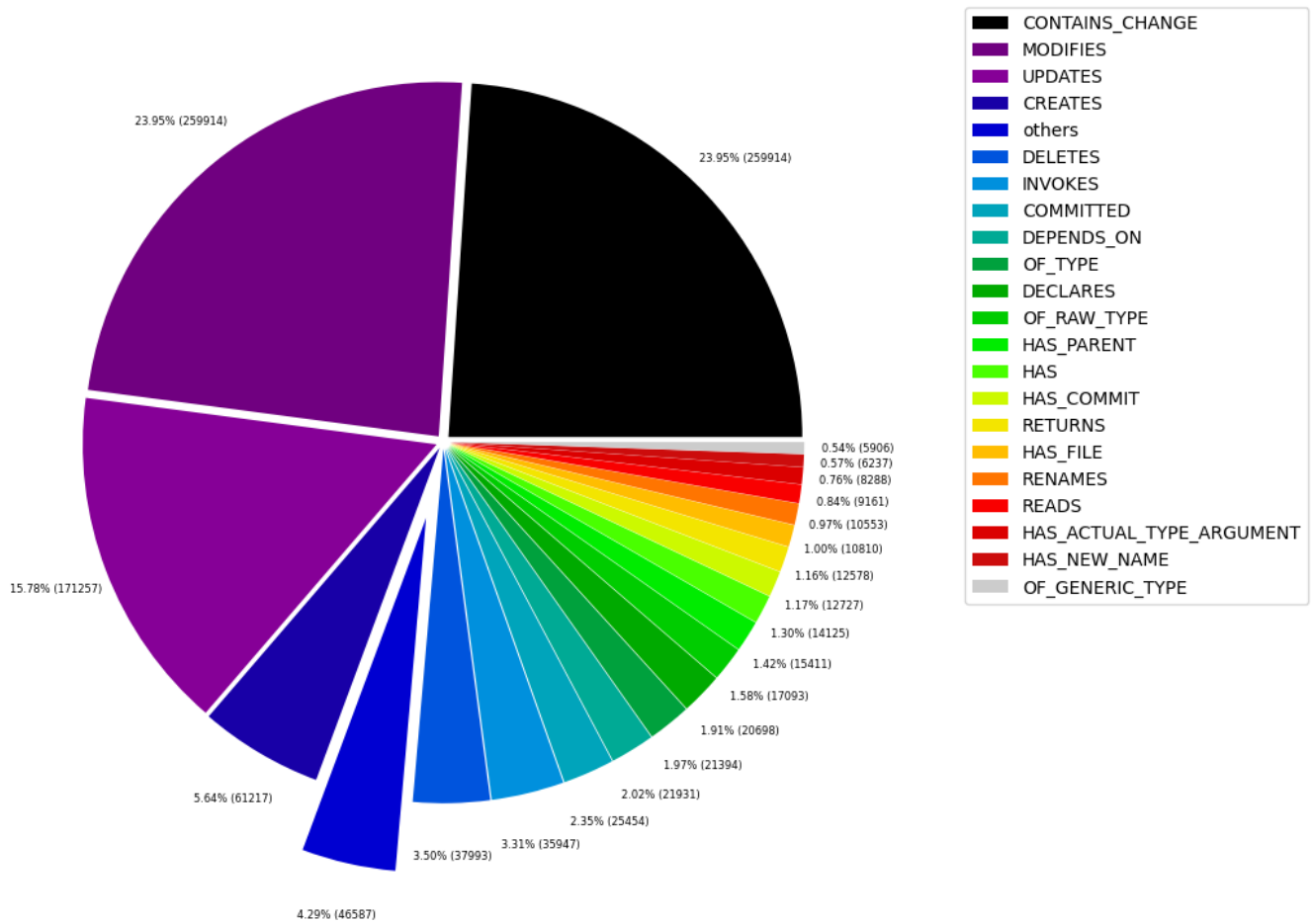
Relationship types (more than 0.5% overall)

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

23.95% (259914)
23.95% (259914)
15.78% (171257)
5.64% (61217)
4.29% (46587)
3.50% (37993)
3.31% (35947)
2.35% (25454)
2.02% (21931)
1.97% (21394)
1.91% (20698)
1.58% (17093)
1.42% (15411)
1.30% (14125)
1.17% (12727)
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.
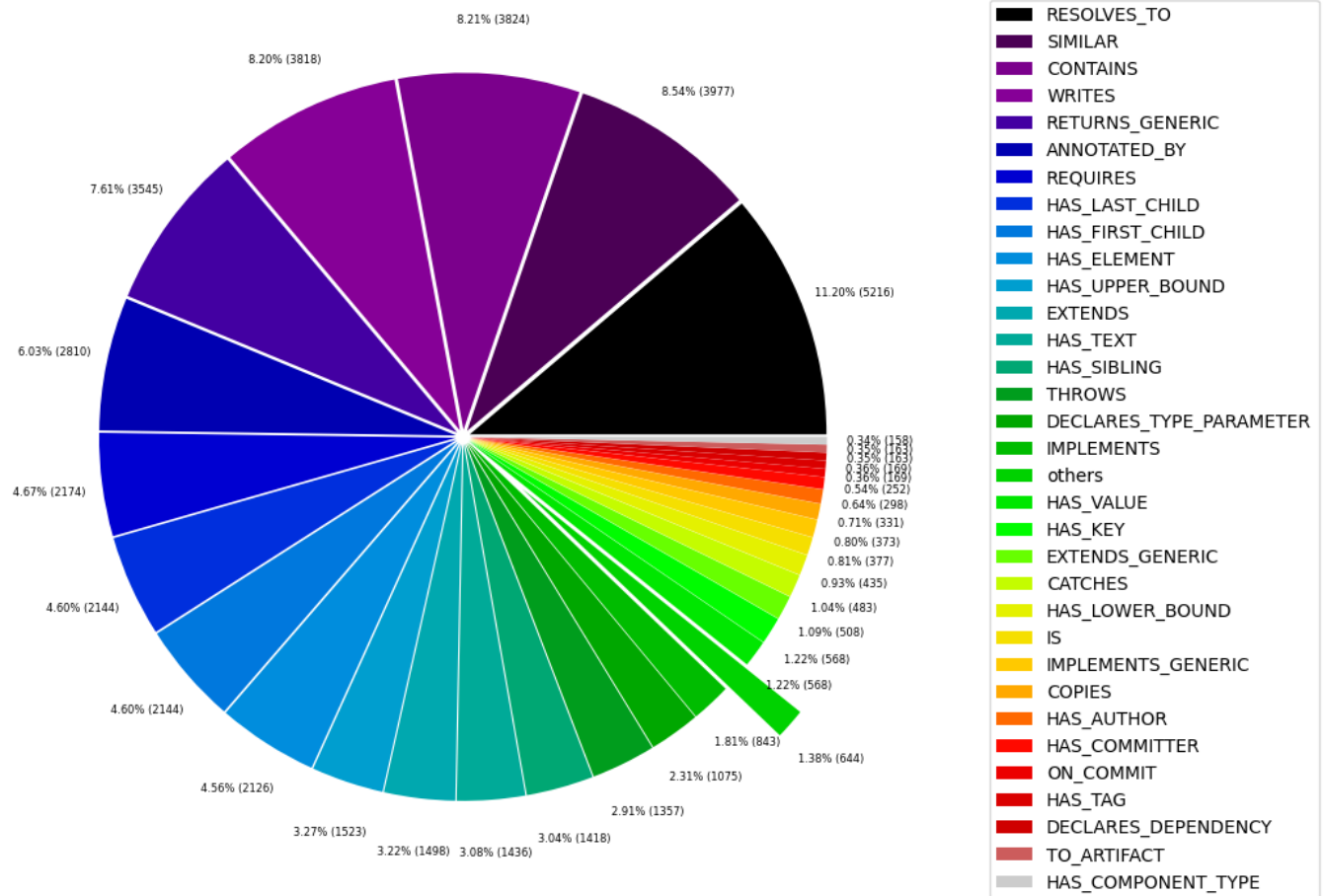
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000092
1	THROWS_GENERIC	5	0.000461
2	EXCLUDES	5	0.000461
3	DECLARES_DEV_DEPENDENCY	7	0.000645
4	DESCRIBES	9	0.000829
5	HAS_ROOT_ELEMENT	10	0.000921
6	HAS_GOAL	16	0.001474
7	HAS_EXECUTION	16	0.001474
8	OF_NAMESPACE	18	0.001659
9	HAS_ATTRIBUTE	18	0.001659
10	INCLUDES_CONCEPT	19	0.001751
11	REQUIRES_TYPE_PARAMETER	21	0.001935
12	IS_ARTIFACT	21	0.001935
13	HAS_CONFIGURATION	21	0.001935
14	USES_PLUGIN	21	0.001935
15	REQUIRES_CONCEPT	28	0.002580
16	HAS_DEFAULT	36	0.003317
17	DECLARES_NAMESPACE	36	0.003317
18	HAS_BRANCH	44	0.004055
19	HAS_HEAD	45	0.004147
20	CONTAINS_VALUE	121	0.011150
21	COPY_OF	126	0.011611
22	HAS_COMPONENT_TYPE	158	0.014560
23	TO_ARTIFACT	163	0.015020
24	DECLARES_DEPENDENCY	163	0.015020
25	ON_COMMIT	169	0.015573
26	HAS_TAG	169	0.015573
27	HAS_COMMITTER	252	0.023222
28	HAS_AUTHOR	298	0.027461
29	COPIES	331	0.030501

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:
- RESOLVES_TO
- SIMILAR
- CONTAINS
- WRITES
- 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]	259914	259914	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	259914	12727	
2	[Git, Change]	UPDATES	[File, Git]	171257	259914	
3	[Git, Change]	CREATES	[File, Git]	61217	259914	
4	[Git, Change]	DELETES	[File, Git]	37993	259914	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21999	13102	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15402	12727	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12727	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12727	298	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12727	252	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10810	1	
11	[Git, Change]	RENAMES	[File, Git]	10553	259914	
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	[Type, File, Java, ByteCode, GenericDeclaratio...]	DECLARES	[Java, ByteCode, Member, Method]	2137	230	
27	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
28	[Java, ByteCode, Member, Method, Constructor]	INVOKES	[Java, ByteCode, Member, Method, Constructor]	2119	2039	
29	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2080	13102	

Graph Density

```
total_number_of_nodes (vertices): 349463
total_number_of_relationships (edges): 1085195
-> total directed graph density: 8.886006471136158e-06
-> total directed graph density in percent: 0.0008886006471136159
```