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]	245365	73.478635
1	[Java, ByteCode, Method, Member]	13204	3.954158
2	[Java, ByteCode, Parameter]	13000	3.893066
3	[Git, Commit]	12005	3.595097
4	[File, Git]	10675	3.196806
5	[Java, ByteCode, Bound, ParameterizedType]	7166	2.145978
6	[Java, ByteCode, Bound]	7112	2.129807
7	[Java, ByteCode, Member, Field]	3494	1.046336
8	[Java, ByteCode, Bound, WildcardType]	2906	0.870250
9	[Java, Value, ByteCode, Annotation]	2821	0.844795
10	[Xml, Element]	2144	0.642056
11	[Java, ByteCode, Constructor, Method, Member]	2040	0.610912
12	[Xml, Text]	1436	0.430034
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.328215
14	[Java, ByteCode, Method, Member, Lambda]	936	0.280301
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	900	0.269520
16	[Type, File, Java, ByteCode, Class]	861	0.257841
17	[Type, File, Java, ByteCode, JavaType]	746	0.223402
18	[Java, Value, ByteCode, Primitive]	650	0.194653
19	[Java, ByteCode, Method, Member, GenericDeclar...	569	0.170397
20	[Json, Key]	555	0.166204
21	[Value, Json, Scalar]	539	0.161413
22	[Type, File, Java, ByteCode, ExternalType]	396	0.118589
23	[Author, Git, Person]	292	0.087444
24	[Value, Array]	267	0.079958
25	[Committer, Git, Person]	247	0.073968
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.070674
27	[Value, Property]	202	0.060492
28	[Type, File, Java, ByteCode, Interface]	187	0.056000
29	[Java, Value, ByteCode, Class]	172	0.051508

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', 'Method', 'Member']
- ['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', 'Constructor', 'Method', 'Member']

Pie slice labels:
- 73.48% (245365)
- 0.61% (2040)
- 0.64% (2144)
- 0.84% (2821)
- 0.87% (2906)
- 1.05% (3494)
- 2.13% (7112)
- 2.15% (7166)
- 3.20% (10675)
- 3.59% (11995)
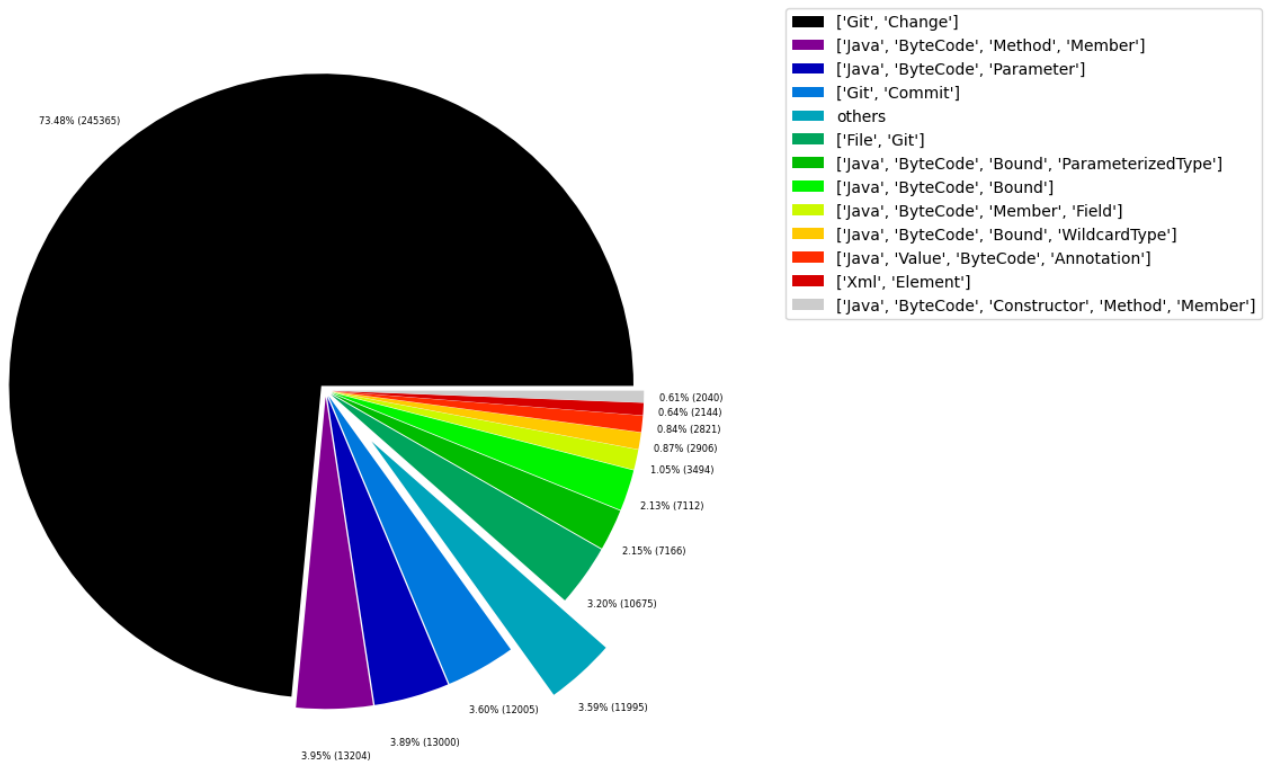- 3.60% (12005)
- 3.89% (13000)
- 3.95% (13204)

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	0.000299
1	[File, Json]	2	0.000599
2	[File]	3	0.000898
3	[Java, ByteCode, Constructor, Method, Member, ...	4	0.001198
4	[Maven, Exclusion]	5	0.001497
5	[Value, Array, Json]	6	0.001797
6	[File, Xml, Maven, Pom, Document]	9	0.002695
7	[Type, File, Java, ByteCode, Void]	9	0.002695
8	[Java, ManifestSection]	9	0.002695
9	[File, Java, Manifest]	9	0.002695
10	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002695
11	[File, Java, ServiceLoader]	10	0.002995
12	[File, Java, Properties]	12	0.003594
13	[Maven, ExecutionGoal]	16	0.004791
14	[Maven, PluginExecution]	16	0.004791
15	[Xml, Attribute]	18	0.005390
16	[jQAssistant, Rule, Concept]	19	0.005690
17	[Maven, Plugin]	21	0.006289
18	[Maven, Configuration]	21	0.006289
19	[Type, File, Java, ByteCode, Enum]	28	0.008385
20	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008685
21	[Git, Branch]	30	0.008984
22	[Xml, Namespace]	36	0.010781
23	[Type, File, Java, ByteCode, Annotation]	43	0.012877
24	[File, Directory]	44	0.013177
25	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.025455
26	[Artifact, Maven]	92	0.027551
27	[Type, File, Java, ByteCode, ExternalType, Ext...	93	0.027850
28	[Value, ManifestEntry]	120	0.035936
29	[Value, Json, Object]	132	0.039530

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)

7.80% (936)
7.50% (900)
9.14% (1096)
7.18% (861)
11.97% (1436)
6.22% (746)
5.42% (650)
0.30% (36)
0.31% (37)
0.37% (44)
0.71% (85)
0.77% (92)
0.78% (93)
1.00% (120)
1.10% (132)
1.16% (139)
1.25% (150)
1.32% (158)
1.36% (163)
1.39% (167)
1.43% (172)
1.56% (187)
1.68% (202)
1.97% (236)
2.06% (247)
2.23% (267)
4.74% (569)
4.63% (555)
4.49% (539)
3.30% (396)
2.43% (292)
2.38% (286)

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']
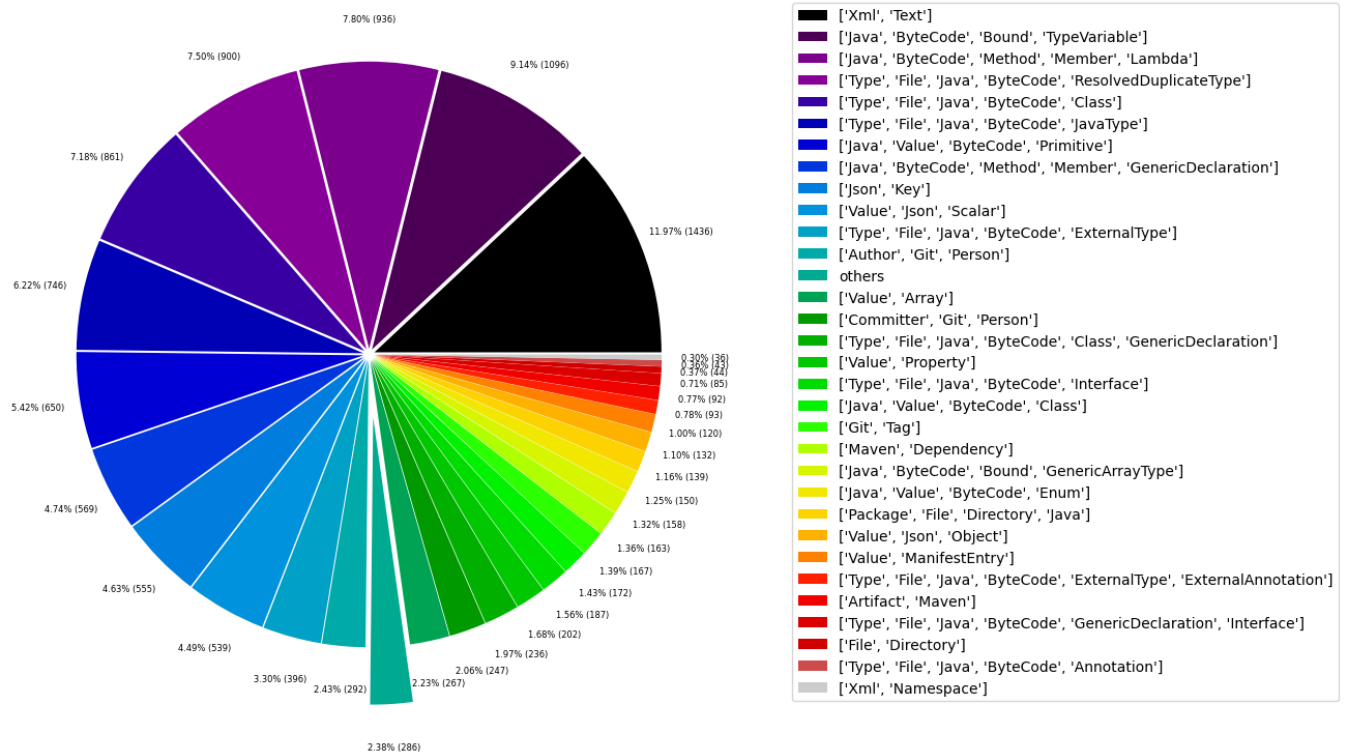- ['Xml', 'Namespace']

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.
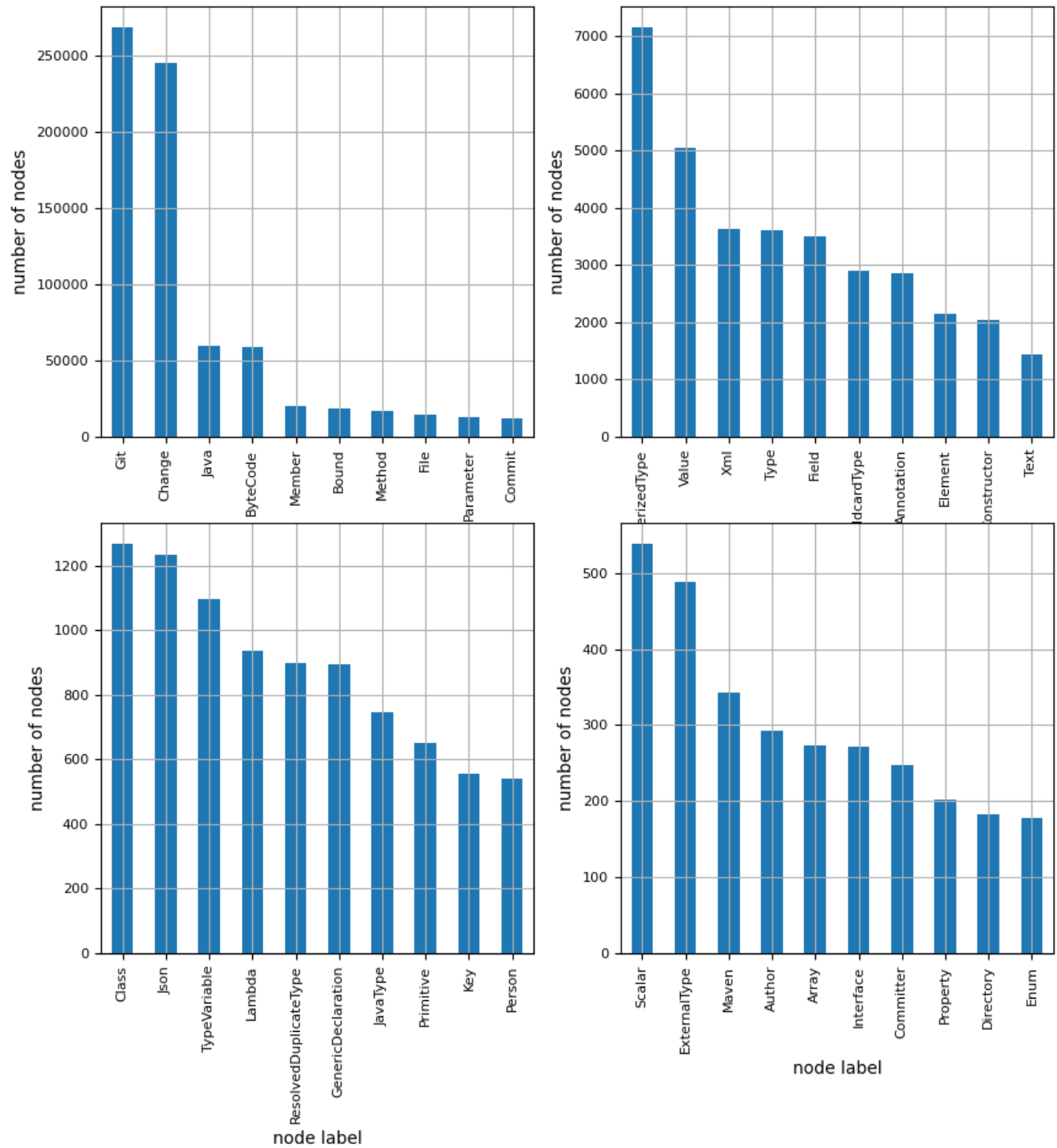
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	268781	80.490946
1	Change	245365	73.478635
2	Java	59279	17.752084
3	ByteCode	59091	17.695784
4	Member	20247	6.063301
5	Bound	18438	5.521566
6	Method	16753	5.016965
7	File	14525	4.349753
8	Parameter	13000	3.893066
9	Commit	12005	3.595097
10	ParameterizedType	7166	2.145978
11	Value	5059	1.515002
12	Xml	3643	1.090957
13	Type	3613	1.081973
14	Field	3494	1.046336
15	WildcardType	2906	0.870250
16	Annotation	2864	0.857672
17	Element	2144	0.642056
18	Constructor	2044	0.612110
19	Text	1436	0.430034
20	Class	1269	0.380023
21	Json	1234	0.369542
22	TypeVariable	1096	0.328215
23	Lambda	936	0.280301
24	ResolvedDuplicateType	900	0.269520
25	GenericDeclaration	894	0.267723
26	JavaType	746	0.223402
27	Primitive	650	0.194653
28	Key	555	0.166204
29	Person	539	0.161413
30	Scalar	539	0.161413
31	ExternalType	489	0.146439
32	Maven	343	0.102717
33	Author	292	0.087444
34	Array	273	0.081754
35	Interface	272	0.081455
36	Committer	247	0.073968
37	Property	202	0.060492
38	Directory	183	0.054802
39	Enum	178	0.053305

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: 1012480
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	245365	24.234059
1	MODIFIES	245365	24.234059
2	UPDATES	158745	15.678828
3	CREATES	59716	5.897993
4	DELETES	37193	3.673455
5	INVOKES	35574	3.513551
6	COMMITTED	24010	2.371405
7	DEPENDS_ON	21873	2.160339
8	OF_TYPE	21347	2.108387
9	DECLARES	20688	2.043300
10	OF_RAW_TYPE	17077	1.686651
11	HAS_PARENT	14598	1.441806
12	HAS	14094	1.392027
13	RETURNS	12547	1.239234
14	RENAMES	10289	1.016218
15	READS	9076	0.896413
16	HAS_ACTUAL_TYPE_ARGUMENT	8276	0.817399
17	HAS_NEW_NAME	6198	0.612160
18	OF_GENERIC_TYPE	5902	0.582925
19	RESOLVES_TO	5225	0.516060
20	SIMILAR	3968	0.391909
21	CONTAINS	3805	0.375810
22	WRITES	3803	0.375612
23	RETURNS_GENERIC	3543	0.349933
24	ANNOTATED_BY	2809	0.277438
25	REQUIRES	2173	0.214622
26	HAS_FIRST_CHILD	2144	0.211757
27	HAS_LAST_CHILD	2144	0.211757
28	HAS_ELEMENT	2126	0.209979
29	HAS_UPPER_BOUND	1521	0.150225

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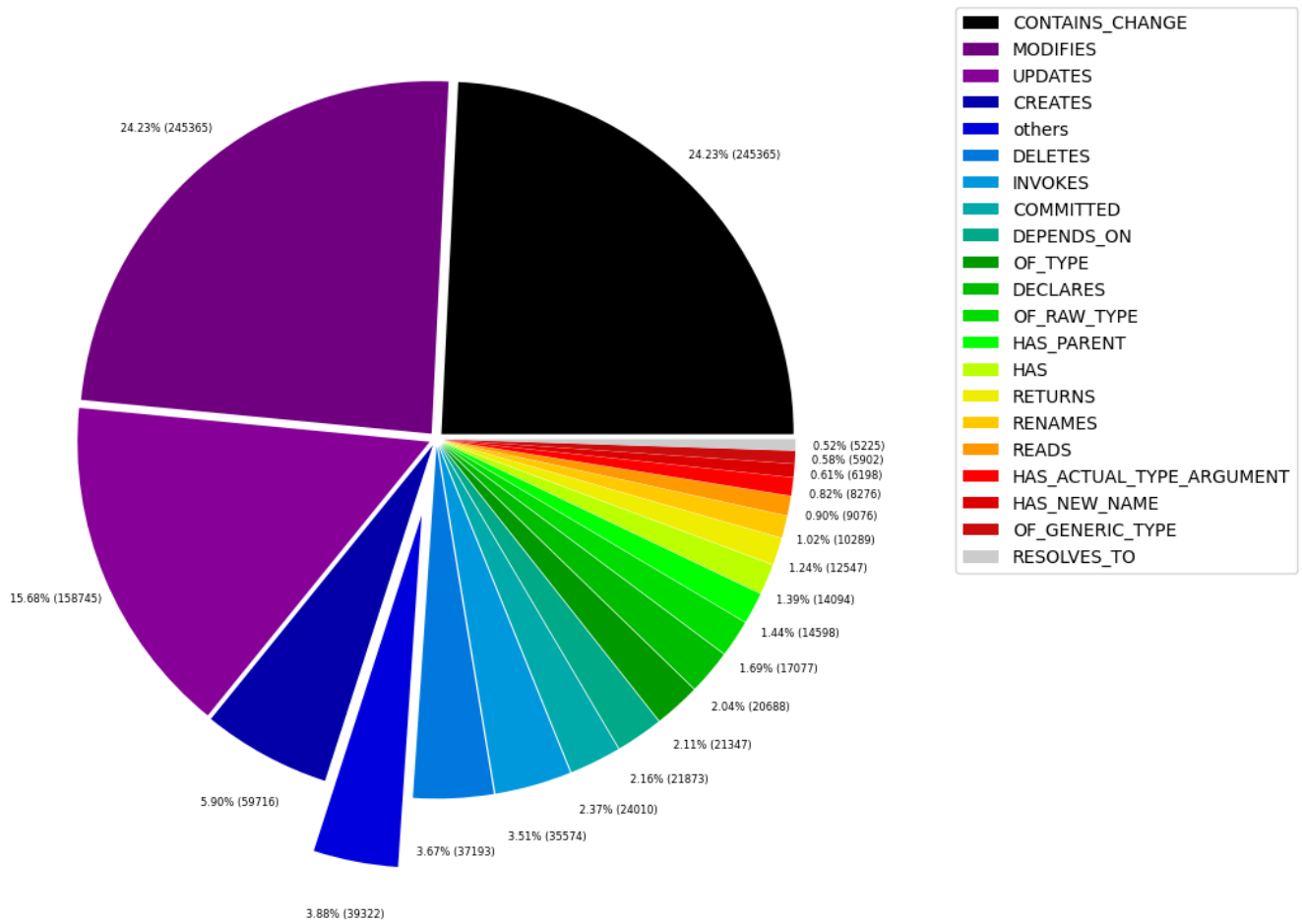
Relationship types (more than 0.5% overall)

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

Data labels:
- 24.23% (245365)
- 24.23% (245365)
- 15.68% (158745)
- 5.90% (59716)
- 3.88% (39322)
- 3.67% (37193)
- 3.51% (35574)
- 2.37% (24010)
- 2.16% (21873)
- 2.11% (21347)
- 2.04% (20688)
- 1.69% (17077)
- 1.44% (14598)
- 1.39% (14094)
- 1.24% (12547)
- 1.02% (10289)
- 0.90% (9076)
- 0.82% (8276)
- 0.61% (6198)
- 0.58% (5902)
- 0.52% (5225)

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.
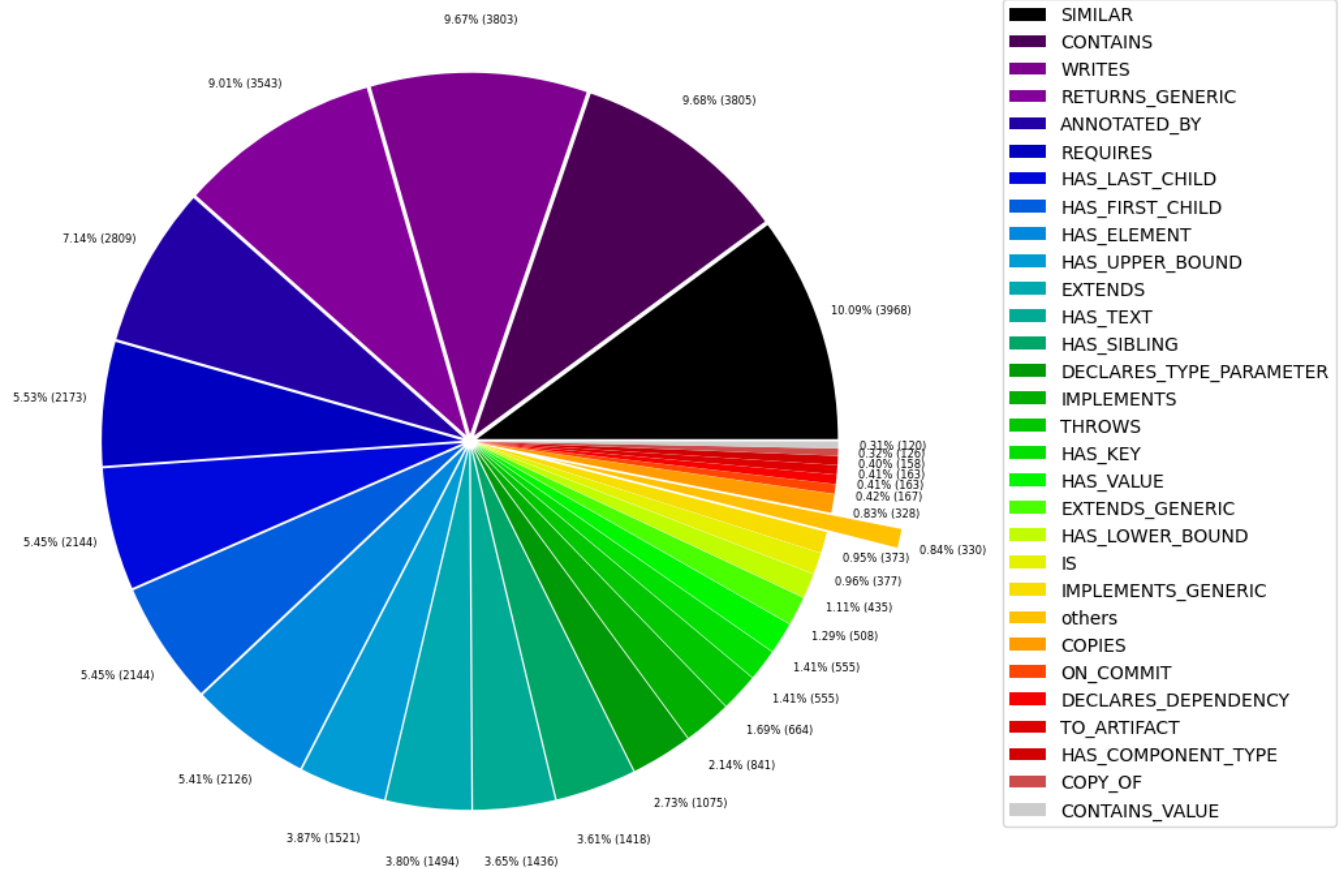
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000099
1	THROWS_GENERIC	5	0.000494
2	EXCLUDES	5	0.000494
3	HAS_ROOT_ELEMENT	9	0.000889
4	DESCRIBES	9	0.000889
5	HAS_GOAL	16	0.001580
6	HAS_EXECUTION	16	0.001580
7	OF_NAMESPACE	18	0.001778
8	HAS_ATTRIBUTE	18	0.001778
9	INCLUDES_CONCEPT	19	0.001877
10	USES_PLUGIN	21	0.002074
11	REQUIRES_TYPE_PARAMETER	21	0.002074
12	IS_ARTIFACT	21	0.002074
13	HAS_CONFIGURATION	21	0.002074
14	REQUIRES_CONCEPT	28	0.002765
15	HAS_HEAD	30	0.002963
16	DECLARES_NAMESPACE	36	0.003556
17	HAS_DEFAULT	36	0.003556
18	CONTAINS_VALUE	120	0.011852
19	COPY_OF	126	0.012445
20	HAS_COMPONENT_TYPE	158	0.015605
21	TO_ARTIFACT	163	0.016099
22	DECLARES_DEPENDENCY	163	0.016099
23	ON_COMMIT	167	0.016494
24	COPIES	328	0.032396
25	IMPLEMENTS_GENERIC	373	0.036840
26	IS	377	0.037235
27	HAS_LOWER_BOUND	435	0.042964
28	EXTENDS_GENERIC	508	0.050174
29	HAS_VALUE	555	0.054816

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:
- 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

Pie chart values:
- 9.67% (3803)
- 9.68% (3805)
- 9.01% (3543)
- 10.09% (3968)
- 7.14% (2809)
- 5.53% (2173)
- 0.31% (120)
- 0.32% (126)
- 0.40% (158)
- 0.41% (163)
- 0.41% (163)
- 0.42% (167)
- 0.83% (328)
- 0.84% (330)
- 5.45% (2144)
- 0.95% (373)
- 0.96% (377)
- 1.11% (435)
- 1.29% (508)
- 1.41% (555)
- 1.41% (555)
- 1.69% (664)
- 2.14% (841)
- 5.45% (2144)
- 2.73% (1075)
- 5.41% (2126)
- 3.61% (1418)
- 3.87% (1521)
- 3.80% (1494)
- 3.65% (1436)

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]	245365	12005	
1	[Git, Change]	MODIFIES	[File, Git]	245365	245365	
2	[Git, Change]	UPDATES	[File, Git]	158745	245365	
3	[Git, Change]	CREATES	[File, Git]	59716	245365	
4	[Git, Change]	DELETES	[File, Git]	37193	245365	
5	[Java, ByteCode, Method, Member]	INVOKES	[Java, ByteCode, Method, Member]	21521	13080	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14589	12005	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	12005	292	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12005	247	
9	[Git, Change]	RENAMES	[File, Git]	10289	245365	
10	[Java, ByteCode, Method, Member]	HAS	[Java, ByteCode, Parameter]	8297	13080	
11	[Java, ByteCode, Method, Member]	READS	[Java, ByteCode, Member, Field]	8100	13080	
12	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6360	13000	
13	[File, Git]	HAS_NEW_NAME	[File, Git]	6198	10675	
14	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3569	7112	
15	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3043	7166	
16	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2906	7166	
17	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2655	13000	
18	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2454	7166	
19	[Java, ByteCode, Constructor, Method, Member]	WRITES	[Java, ByteCode, Member, Field]	2426	2040	
20	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalType, Ext...	2340	2821	
21	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
22	[Java, ByteCode, Constructor, Method, Member]	INVOKES	[Java, ByteCode, Constructor, Method, Member]	2111	2040	
23	[Java, ByteCode, Method, Member]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2088	13080	
24	[Java, ByteCode, Constructor, Method, Member]	HAS	[Java, ByteCode, Parameter]	2023	2040	
25	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2023	13000	
26	[Java, ByteCode, Method, Member]	INVOKES	[Java, ByteCode, Constructor, Method, Member]	2007	13080	
27	[Java, ByteCode, Method, Member, Lambda]	INVOKES	[Java, ByteCode, Method, Member]	1883	936	
28	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound]	1802	7166	
29	[Java, ByteCode, Constructor, Method, Member]	INVOKES	[Java, ByteCode, Method, Member]	1796	2040	

Graph Density

```
total_number_of_nodes (vertices): 333927
total_number_of_relationships (edges): 1012480
-> total directed graph density: 9.079975613975641e-06
-> total directed graph density in percent: 0.0009079975613975641
```