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]	240780	73.162950
1	[Java, ByteCode, Member, Method]	13183	4.005761
2	[Java, ByteCode, Parameter]	13000	3.950155
3	[Git, Commit]	11858	3.603149
4	[File, Git]	10614	3.225150
5	[Java, ByteCode, Bound, ParameterizedType]	7166	2.177447
6	[Java, ByteCode, Bound]	7112	2.161039
7	[Java, ByteCode, Member, Field]	3488	1.059857
8	[Java, ByteCode, Bound, WildcardType]	2906	0.883012
9	[Java, Value, ByteCode, Annotation]	2821	0.857184
10	[Xml, Element]	2144	0.651472
11	[Java, ByteCode, Member, Constructor, Method]	2040	0.619870
12	[Xml, Text]	1436	0.436340
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.333028
14	[Java, ByteCode, Member, Method, Lambda]	936	0.284411
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	900	0.273472
16	[Type, File, Java, ByteCode, Class]	861	0.261622
17	[Type, File, Java, ByteCode, JavaType]	746	0.226678
18	[Java, Value, ByteCode, Primitive]	650	0.197508
19	[Java, ByteCode, GenericDeclaration, Member, M...	569	0.172895
20	[Json, Key]	555	0.168641
21	[Value, Json, Scalar]	539	0.163780
22	[Type, File, Java, ByteCode, ExternalType]	396	0.120328
23	[Author, Git, Person]	290	0.088119
24	[Value, Array]	267	0.081130
25	[Committer, Git, Person]	244	0.074141
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.071711
27	[Value, Property]	202	0.061379
28	[Type, File, Java, ByteCode, Interface]	187	0.056821
29	[Java, Value, ByteCode, Class]	172	0.052264

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']
- others
- ['Git', 'Commit']
- ['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', 'Constructor', 'Method']

Pie slice labels:
73.16% (240780)
0.62% (2040)
0.65% (2144)
0.86% (2821)
0.88% (2906)
1.06% (3488)
2.16% (7112)
2.18% (7166)
3.23% (10614)
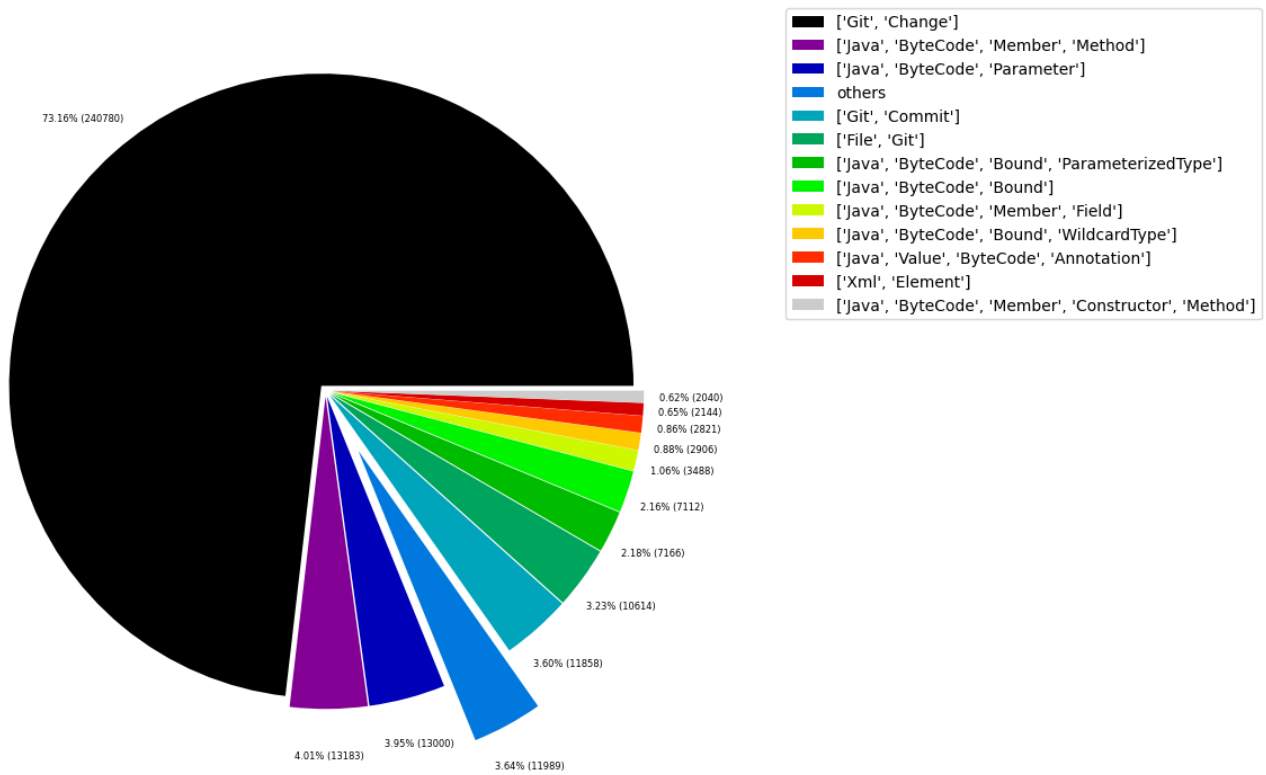3.60% (11858)
3.64% (11989)
3.95% (13000)
4.01% (13183)

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	0.000304
1	[File, TS, Scan]	1	0.000304
2	[File, Json]	2	0.000608
3	[File]	3	0.000912
4	[Java, ByteCode, GenericDeclaration, Member, C...	4	0.001215
5	[Maven, Exclusion]	5	0.001519
6	[Value, Array, Json]	6	0.001823
7	[Type, File, Java, ByteCode, Void]	9	0.002735
8	[File, Maven, Xml, Pom, Document]	9	0.002735
9	[Java, ManifestSection]	9	0.002735
10	[File, Java, Manifest]	9	0.002735
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002735
12	[File, Java, ServiceLoader]	10	0.003039
13	[File, Java, Properties]	12	0.003646
14	[Maven, ExecutionGoal]	16	0.004862
15	[Maven, PluginExecution]	16	0.004862
16	[Xml, Attribute]	18	0.005469
17	[jQAssistant, Rule, Concept]	19	0.005773
18	[Maven, Plugin]	21	0.006381
19	[Maven, Configuration]	21	0.006381
20	[Git, Branch]	28	0.008508
21	[Type, File, Java, ByteCode, Enum]	28	0.008508
22	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008812
23	[Xml, Namespace]	36	0.010939
24	[Type, File, Java, ByteCode, Annotation]	43	0.013066
25	[File, Directory]	44	0.013370
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.025828
27	[Artifact, Maven]	92	0.027955
28	[Type, File, Java, ByteCode, ExternalAnnotatio...	93	0.028259
29	[Value, ManifestEntry]	120	0.036463

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', 'GenericDeclaration', 'Member', 'Method']
- ['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', 'ExternalAnnotation', 'ExternalType']
- ['Artifact', 'Maven']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Interface']
- ['File', 'Directory']
- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']
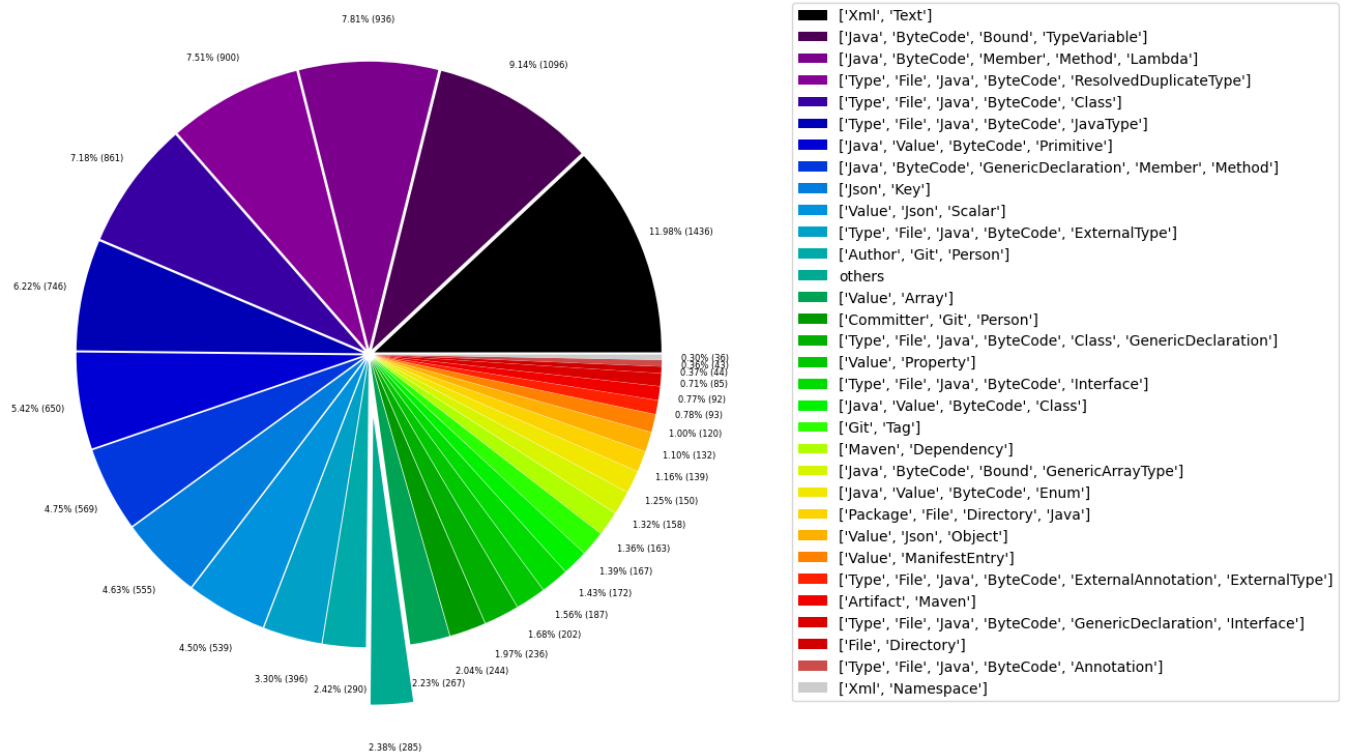- ['Xml', 'Namespace']

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.
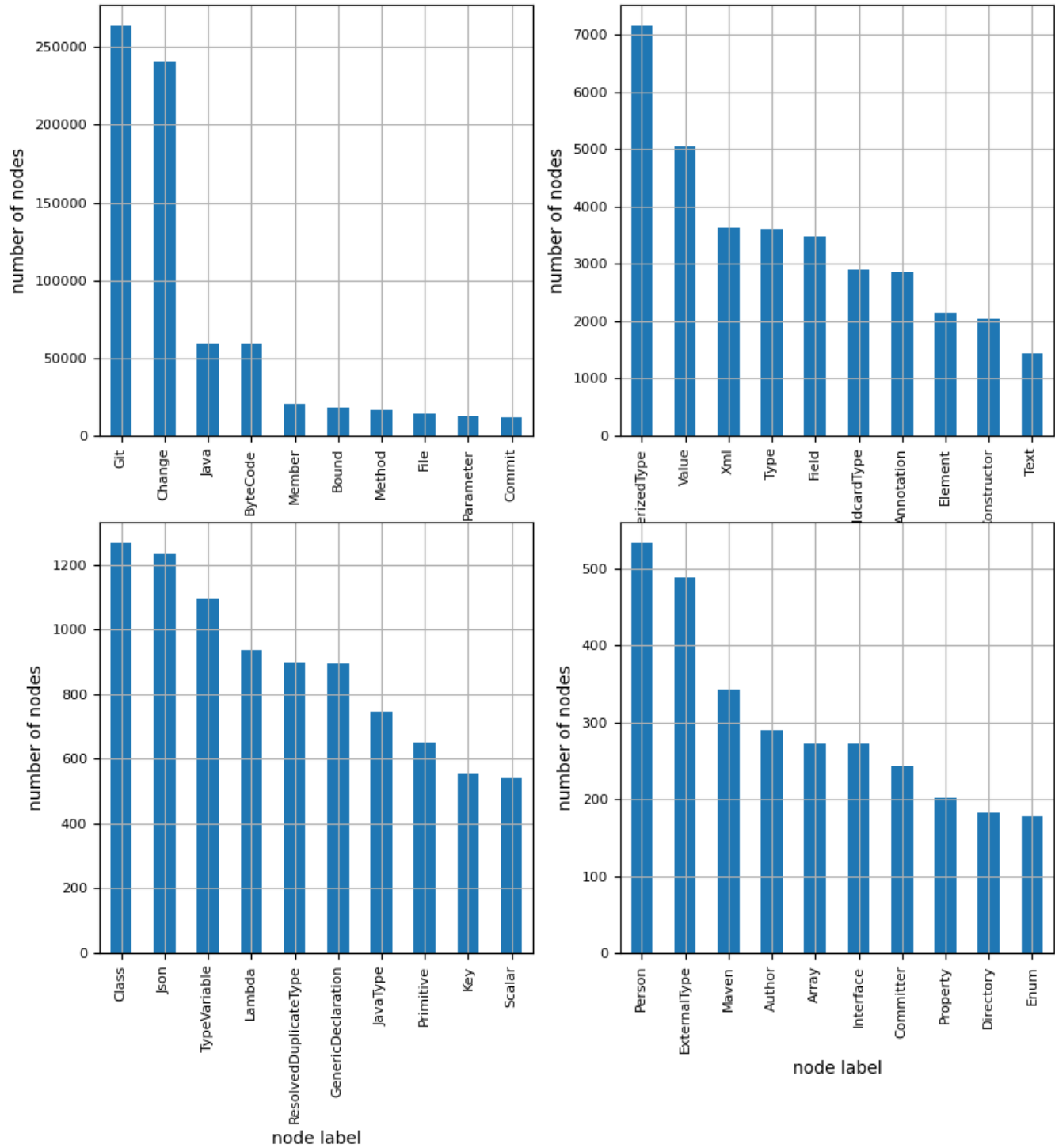
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	263981	80.212761
1	Change	240780	73.162950
2	Java	59252	18.004199
3	ByteCode	59064	17.947074
4	Member	20220	6.144011
5	Bound	18438	5.602535
6	Method	16732	5.084153
7	File	14465	4.395307
8	Parameter	13000	3.950155
9	Commit	11858	3.603149
10	ParameterizedType	7166	2.177447
11	Value	5059	1.537218
12	Xml	3643	1.106955
13	Type	3613	1.097839
14	Field	3488	1.059857
15	WildcardType	2906	0.883012
16	Annotation	2864	0.870250
17	Element	2144	0.651472
18	Constructor	2044	0.621086
19	Text	1436	0.436340
20	Class	1269	0.385596
21	Json	1234	0.374961
22	TypeVariable	1096	0.333028
23	Lambda	936	0.284411
24	ResolvedDuplicateType	900	0.273472
25	GenericDeclaration	894	0.271649
26	JavaType	746	0.226678
27	Primitive	650	0.197508
28	Key	555	0.168641
29	Scalar	539	0.163780
30	Person	534	0.162260
31	ExternalType	489	0.148587
32	Maven	343	0.104223
33	Author	290	0.088119
34	Array	273	0.082953
35	Interface	272	0.082649
36	Committer	244	0.074141
37	Property	202	0.061379
38	Directory	183	0.055606
39	Enum	178	0.054087

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: 997927
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	240780	24.128017
1	MODIFIES	240780	24.128017
2	UPDATES	155401	15.572382
3	CREATES	58833	5.895521
4	DELETES	36677	3.675319
5	INVOKES	35603	3.567696
6	COMMITTED	23716	2.376527
7	DEPENDS_ON	21873	2.191844
8	OF_TYPE	21347	2.139134
9	DECLARES	20661	2.070392
10	OF_RAW_TYPE	17077	1.711247
11	HAS_PARENT	14418	1.444795
12	HAS	14094	1.412328
13	RETURNS	12547	1.257306
14	RENAMES	10131	1.015205
15	READS	9078	0.909686
16	HAS_ACTUAL_TYPE_ARGUMENT	8276	0.829319
17	HAS_NEW_NAME	6178	0.619083
18	OF_GENERIC_TYPE	5902	0.591426
19	RESOLVES_TO	5237	0.524788
20	SIMILAR	3968	0.397624
21	CONTAINS	3805	0.381290
22	WRITES	3803	0.381090
23	RETURNS_GENERIC	3543	0.355036
24	ANNOTATED_BY	2809	0.281484
25	REQUIRES	2173	0.217751
26	HAS_FIRST_CHILD	2144	0.214845
27	HAS_LAST_CHILD	2144	0.214845
28	HAS_ELEMENT	2126	0.213042
29	HAS_UPPER_BOUND	1521	0.152416

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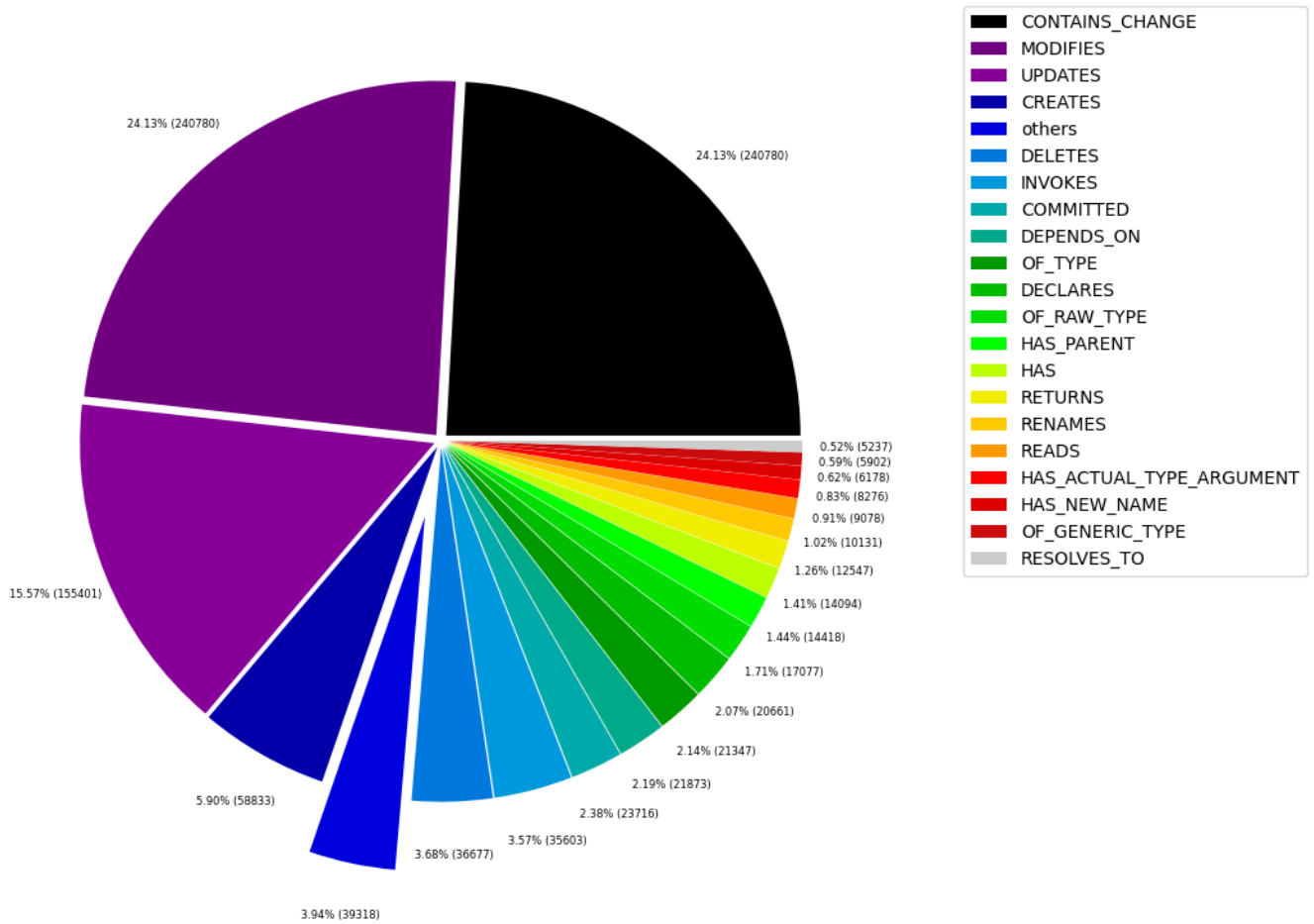
Relationship types (more than 0.5% overall)

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

24.13% (240780)
24.13% (240780)
15.57% (155401)
5.90% (58833)
3.94% (39318)
3.68% (36677)
3.57% (35603)
2.38% (23716)
2.19% (21873)
2.14% (21347)
2.07% (20661)
1.71% (17077)
1.44% (14418)
1.41% (14094)
1.26% (12547)
1.02% (10131)
0.91% (9078)
0.83% (8276)
0.62% (6178)
0.59% (5902)
0.52% (5237)

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.
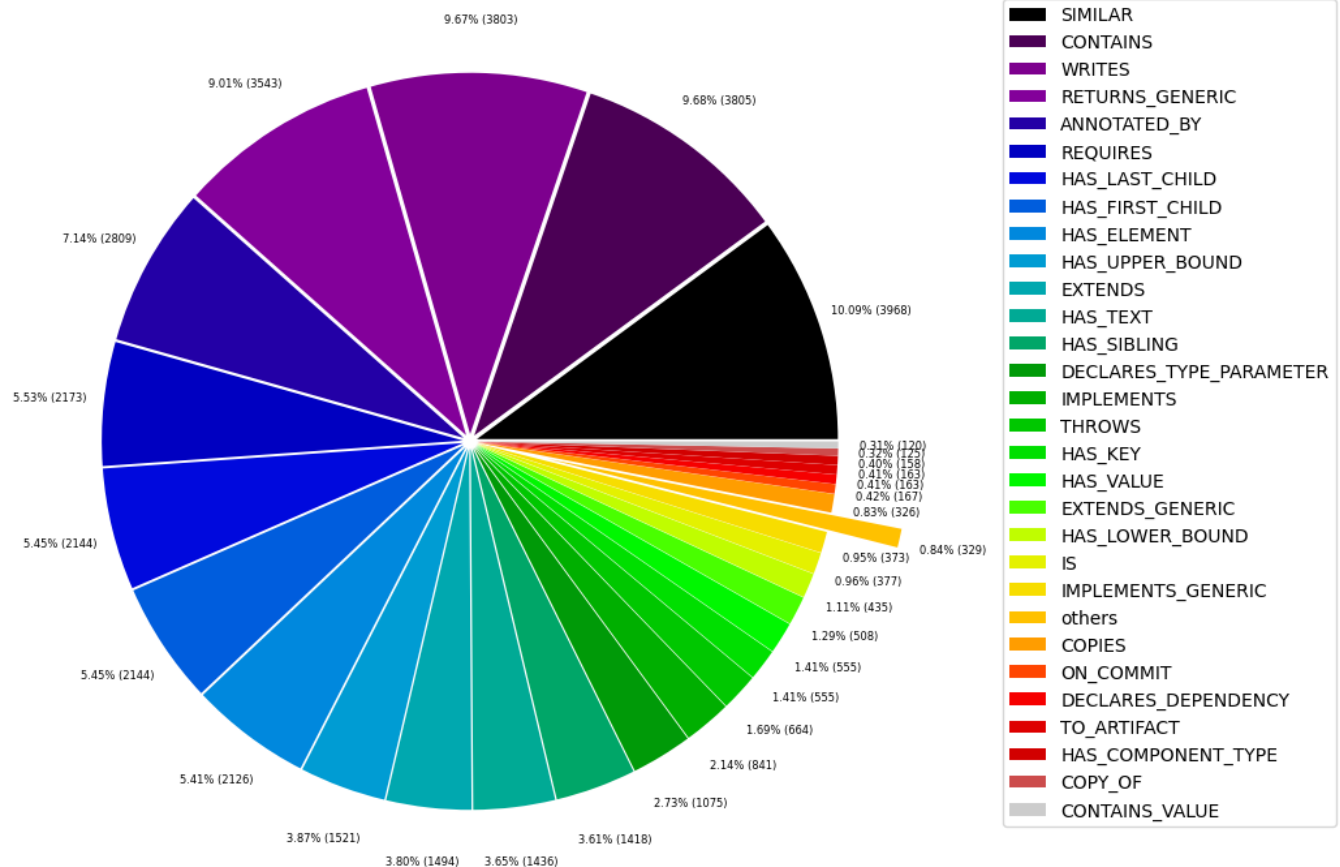
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000100
1	THROWS_GENERIC	5	0.000501
2	EXCLUDES	5	0.000501
3	DESCRIBES	9	0.000902
4	HAS_ROOT_ELEMENT	10	0.001002
5	HAS_GOAL	16	0.001603
6	HAS_EXECUTION	16	0.001603
7	OF_NAMESPACE	18	0.001804
8	HAS_ATTRIBUTE	18	0.001804
9	INCLUDES_CONCEPT	19	0.001904
10	USES_PLUGIN	21	0.002104
11	REQUIRES_TYPE_PARAMETER	21	0.002104
12	IS_ARTIFACT	21	0.002104
13	HAS_CONFIGURATION	21	0.002104
14	REQUIRES_CONCEPT	28	0.002806
15	HAS_HEAD	28	0.002806
16	DECLARES_NAMESPACE	36	0.003607
17	HAS_DEFAULT	36	0.003607
18	CONTAINS_VALUE	120	0.012025
19	COPY_OF	125	0.012526
20	HAS_COMPONENT_TYPE	158	0.015833
21	TO_ARTIFACT	163	0.016334
22	DECLARES_DEPENDENCY	163	0.016334
23	ON_COMMIT	167	0.016735
24	COPIES	326	0.032668
25	IMPLEMENTS_GENERIC	373	0.037377
26	IS	377	0.037778
27	HAS_LOWER_BOUND	435	0.043590
28	EXTENDS_GENERIC	508	0.050906
29	HAS_VALUE	555	0.055615

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>

■	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

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]	240780	11858	
1	[Git, Change]	MODIFIES	[File, Git]	240780	240780	
2	[Git, Change]	UPDATES	[File, Git]	155401	240780	
3	[Git, Change]	CREATES	[File, Git]	58833	240780	
4	[Git, Change]	DELETES	[File, Git]	36677	240780	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21568	13068	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14409	11858	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	11858	290	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	11858	244	
9	[Git, Change]	RENAMES	[File, Git]	10131	240780	
10	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8316	13068	
11	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8129	13068	
12	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6360	13000	
13	[File, Git]	HAS_NEW_NAME	[File, Git]	6178	10614	
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, Member, Constructor, Method]	WRITES	[Java, ByteCode, Member, Field]	2426	2040	
20	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio...	2340	2821	
21	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
22	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2111	2040	
23	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2086	13068	
24	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2030	13068	
25	[Java, ByteCode, Member, Constructor, Method]	HAS	[Java, ByteCode, Parameter]	2023	2040	
26	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2023	13000	
27	[Java, ByteCode, Member, Method, Lambda]	INVOKES	[Java, ByteCode, Member, Method]	1886	936	
28	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound]	1802	7166	
29	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Method]	1797	2040	

Graph Density

```
total_number_of_nodes (vertices): 329101
total_number_of_relationships (edges): 997927
-> total directed graph density: 9.213861692705352e-06
-> total directed graph density in percent: 0.0009213861692705351
```