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` .

```
Total number of nodes: 307187
```

	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	231044	75.212818
1	[Git, Commit]	11625	3.784340
2	[Java, ByteCode, Parameter]	11202	3.646639
3	[Java, ByteCode, Method, Member]	10666	3.472152
4	[File, Git]	10547	3.433414
5	[Java, ByteCode, Bound, ParameterizedType]	6542	2.129647
6	[Java, ByteCode, Bound]	6316	2.056077
7	[Java, ByteCode, Member, Field]	2847	0.926797
8	[Java, ByteCode, Bound, WildcardType]	2723	0.886431
9	[Java, Value, ByteCode, Annotation]	2225	0.724315
10	[Java, ByteCode, Constructor, Method, Member]	1647	0.536156
11	[Xml, Element]	1534	0.499370
12	[Xml, Text]	1016	0.330743
13	[Java, ByteCode, Bound, TypeVariable]	989	0.321954
14	[Java, ByteCode, Method, Member, Lambda]	769	0.250336
15	[Type, File, Java, ByteCode, Class]	699	0.227549
16	[Type, File, Java, ByteCode, JavaType]	597	0.194344
17	[Type, File, Java, ByteCode, ResolvedDuplicate...	542	0.176440
18	[Java, ByteCode, Method, Member, GenericDeclar...	525	0.170906
19	[Java, Value, ByteCode, Primitive]	467	0.152025
20	[Author, Git, Person]	288	0.093754
21	[Committer, Git, Person]	242	0.078779
22	[Type, File, Java, ByteCode, Class, GenericDec...	215	0.069990
23	[Type, File, Java, ByteCode, ExternalType]	194	0.063154
24	[Type, File, Java, ByteCode, Interface]	167	0.054364
25	[Git, Tag]	165	0.053713
26	[Java, ByteCode, Bound, GenericArrayType]	145	0.047203
27	[Java, Value, ByteCode, Enum]	142	0.046226
28	[Maven, Dependency]	116	0.037762
29	[Value, Property]	109	0.035483

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)

75.21% (231044)

0.54% (1647)
0.72% (2225)
0.89% (2723)
0.93% (2847)
2.06% (6316)
2.13% (6542)
3.19% (9803)
3.43% (10547)
3.47% (10666)
3.65% (11202)
3.78% (11625)

Legend:
- ['Git', 'Change']
- ['Git', 'Commit']
- ['Java', 'ByteCode', 'Parameter']
- ['Java', 'ByteCode', 'Method', 'Member']
- ['File', 'Git']
- others
- ['Java', 'ByteCode', 'Bound', 'ParameterizedType']
- ['Java', 'ByteCode', 'Bound']
- ['Java', 'ByteCode', 'Member', 'Field']
- ['Java', 'ByteCode', 'Bound', 'WildcardType']
- ['Java', 'Value', 'ByteCode', 'Annotation']
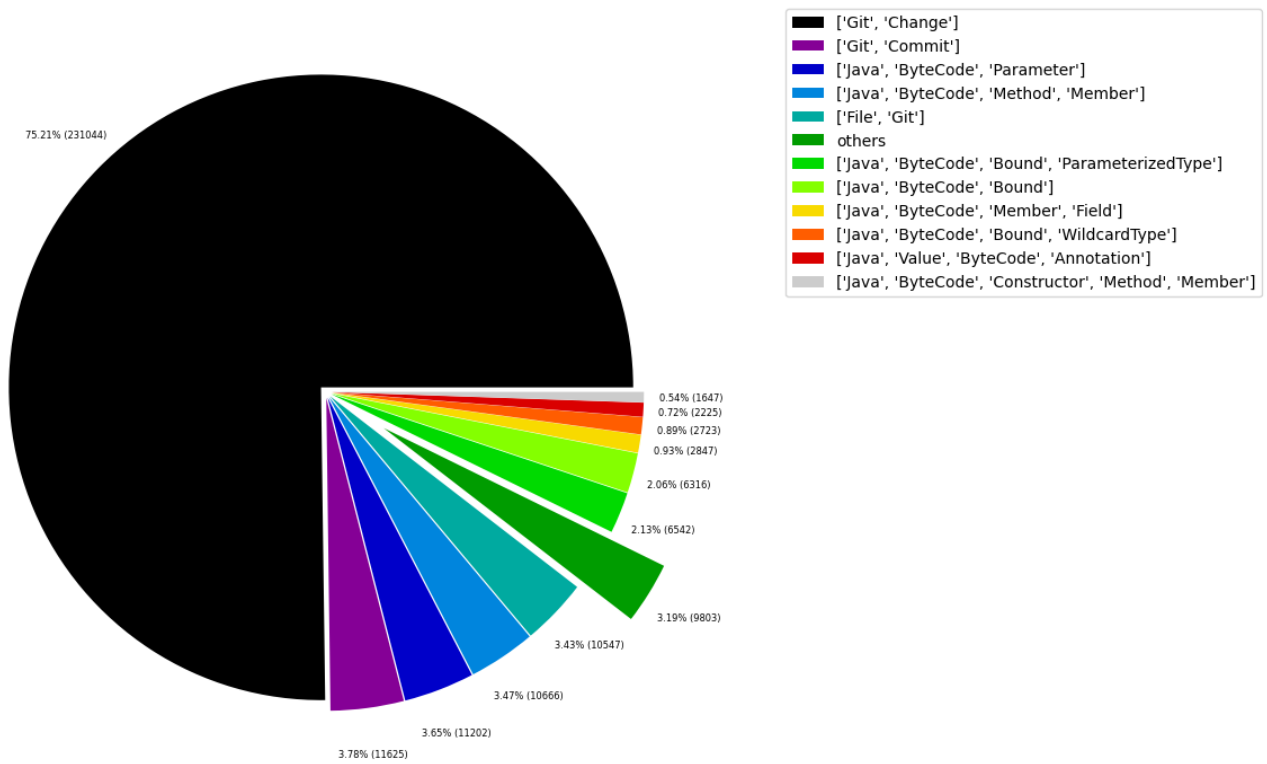- ['Java', 'ByteCode', 'Constructor', 'Method', 'Member']

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	0.000326
1	[Maven, Exclusion]	1	0.000326
2	[Repository, File, Git]	1	0.000326
3	[File, TS, Scan]	1	0.000326
4	[Java, ByteCode, Constructor, Method, Member, ...	3	0.000977
5	[File, Xml, Maven, Pom, Document]	6	0.001953
6	[Type, File, Java, ByteCode, Void]	6	0.001953
7	[Java, ManifestSection]	6	0.001953
8	[File, Java, Manifest]	6	0.001953
9	[Artifact, File, Jar, Archive, Zip, Java]	6	0.001953
10	[File, Java, Properties]	8	0.002604
11	[File, Java, ServiceLoader]	9	0.002930
12	[Xml, Attribute]	12	0.003906
13	[Maven, ExecutionGoal]	14	0.004557
14	[Maven, PluginExecution]	14	0.004557
15	[Maven, Plugin]	16	0.005209
16	[Maven, Configuration]	16	0.005209
17	[jQAssistant, Rule, Concept]	19	0.006185
18	[File]	21	0.006836
19	[Type, File, Java, ByteCode, PrimitiveType]	21	0.006836
20	[Git, Branch]	23	0.007487
21	[Xml, Namespace]	24	0.007813
22	[Type, File, Java, ByteCode, Enum]	25	0.008138
23	[Java, Value, ByteCode, Class]	28	0.009115
24	[File, Directory]	39	0.012696
25	[Type, File, Java, ByteCode, Annotation]	40	0.013021
26	[Type, File, Java, ByteCode, ExternalType, Ext...	65	0.021160
27	[Artifact, Maven]	69	0.022462
28	[Type, File, Java, ByteCode, GenericDeclaratio...	80	0.026043
29	[Value, ManifestEntry]	91	0.029624

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', 'Element']
- ['Xml', 'Text']
- ['Java', 'ByteCode', 'Bound', 'TypeVariable']
- ['Java', 'ByteCode', 'Method', 'Member', 'Lambda']
- ['Type', 'File', 'Java', 'ByteCode', 'Class']
- ['Type', 'File', 'Java', 'ByteCode', 'JavaType']
- ['Type', 'File', 'Java', 'ByteCode', 'ResolvedDuplicateType']
- ['Java', 'ByteCode', 'Method', 'Member', 'GenericDeclaration']
- ['Java', 'Value', 'ByteCode', 'Primitive']
- ['Author', 'Git', 'Person']
- others
- ['Committer', 'Git', 'Person']
- ['Type', 'File', 'Java', 'ByteCode', 'Class', 'GenericDeclaration']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalType']
- ['Type', 'File', 'Java', 'ByteCode', 'Interface']
- ['Git', 'Tag']
- ['Java', 'ByteCode', 'Bound', 'GenericArrayType']
- ['Java', 'Value', 'ByteCode', 'Enum']
- ['Maven', 'Dependency']
- ['Value', 'Property']
- ['Package', 'File', 'Directory', 'Java']
- ['Value', 'Array']
- ['Value', 'ManifestEntry']
- ['Type', 'File', 'Java', 'ByteCode', 'GenericDeclaration', 'Interface']
- ['Artifact', 'Maven']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalType', 'ExternalAnnotation']
- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']
- ['File', 'Directory']

Pie chart values: 15.65% (48069), 10.36% (31837), 10.09% (30991), 7.84% (24097), 7.13% (21904), 6.09% (18708), 5.53% (16984), 5.36% (16451), 4.76% (14634), 2.94% (9025), 2.93% (8993), 2.47% (7583), 2.19% (6737), 1.98% (6079), 1.70% (5233), 1.68% (5170), 1.48% (4544), 1.45% (4450), 1.18% (3635), 1.11% (3416), 1.09% (3353), 1.06% (3259), 0.93% (2852), 0.82% (2507), 0.70% (2162), 0.66% (2037), 0.41% (1253), 0.40% (1222)
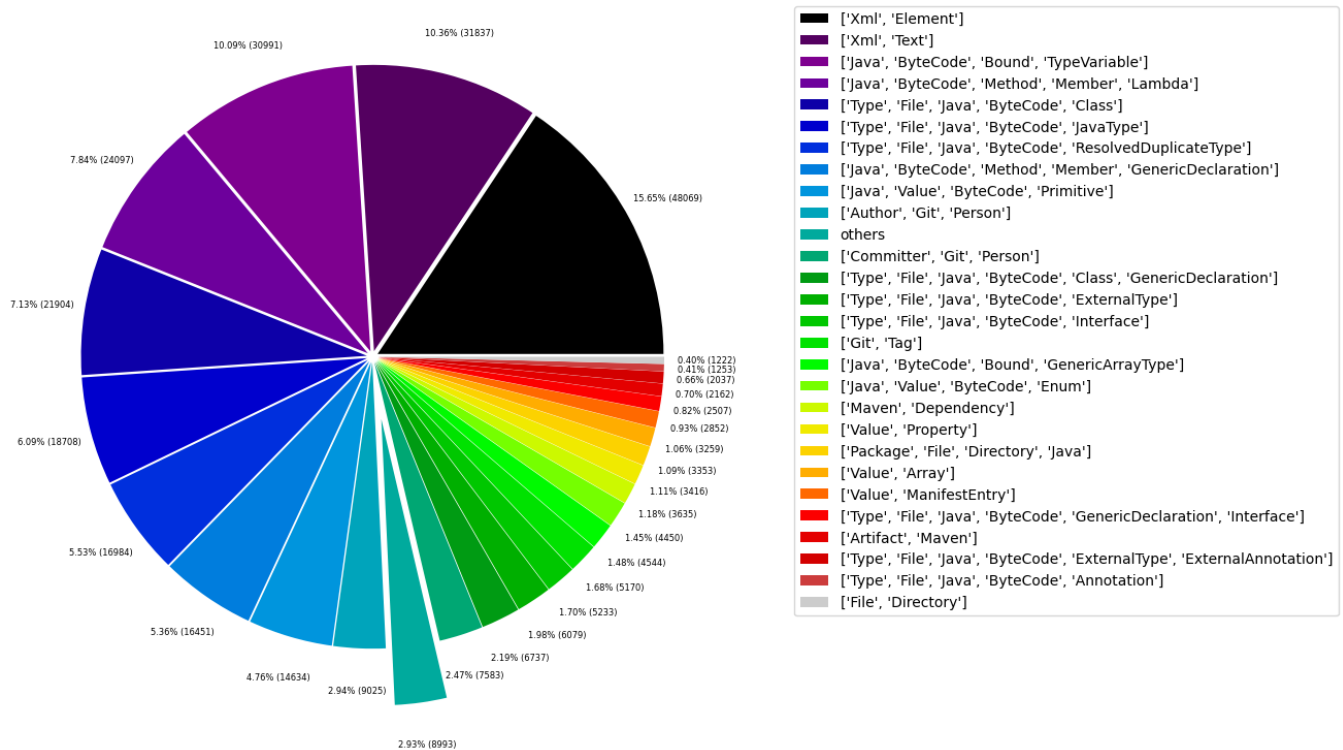
Table 1c - Highest node count by single label

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.
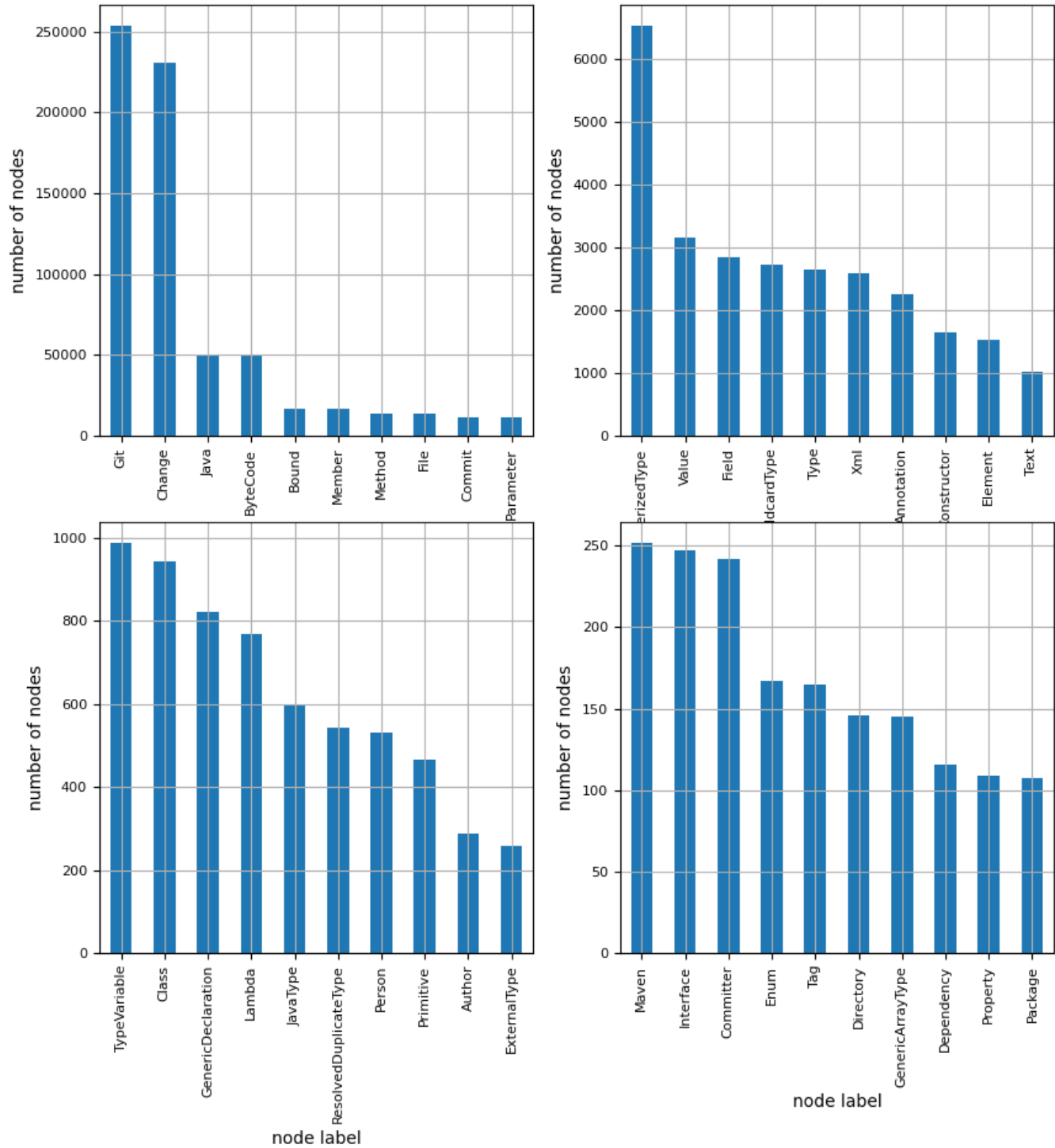
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	253935	82.664631
1	Change	231044	75.212818
2	Java	50029	16.286171
3	ByteCode	49887	16.239945
4	Bound	16715	5.441311
5	Member	16457	5.357323
6	Method	13610	4.430526
7	File	13402	4.362815
8	Commit	11625	3.784340
9	Parameter	11202	3.646639
10	ParameterizedType	6542	2.129647
11	Value	3166	1.030643
12	Field	2847	0.926797
13	WildcardType	2723	0.886431
14	Type	2651	0.862992
15	Xml	2592	0.843786
16	Annotation	2265	0.737336
17	Constructor	1650	0.537132
18	Element	1534	0.499370
19	Text	1016	0.330743
20	TypeVariable	989	0.321954
21	Class	942	0.306654
22	GenericDeclaration	823	0.267915
23	Lambda	769	0.250336
24	JavaType	597	0.194344
25	ResolvedDuplicateType	542	0.176440
26	Person	530	0.172533
27	Primitive	467	0.152025
28	Author	288	0.093754
29	ExternalType	259	0.084313
30	Maven	252	0.082035
31	Interface	247	0.080407
32	Committer	242	0.078779
33	Enum	167	0.054364
34	Tag	165	0.053713
35	Directory	146	0.047528
36	GenericArrayType	145	0.047203
37	Dependency	116	0.037762
38	Property	109	0.035483
39	Package	107	0.034832

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: 953650
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	231044	24.227337
1	MODIFIES	231044	24.227337
2	UPDATES	148020	15.521418
3	CREATES	57541	6.033765
4	DELETES	35338	3.705552
5	INVOKES	29728	3.117286
6	COMMITTED	23250	2.438001
7	OF_TYPE	17636	1.849316
8	DEPENDS_ON	17445	1.829287
9	DECLARES	16819	1.763645
10	OF_RAW_TYPE	15200	1.593876
11	HAS_PARENT	14132	1.481885
12	HAS	11902	1.248047
13	HAS_COMMIT	11625	1.219001
14	RETURNS	10643	1.116028
15	HAS_FILE	10547	1.105961
16	RENAMES	9855	1.033398
17	READS	7695	0.806900
18	HAS_ACTUAL_TYPE_ARGUMENT	7501	0.786557
19	HAS_NEW_NAME	6159	0.645834
20	OF_GENERIC_TYPE	5349	0.560898
21	RESOLVES_TO	3798	0.398259
22	SIMILAR	3419	0.358517
23	RETURNS_GENERIC	3253	0.341110
24	CONTAINS	3084	0.323389
25	WRITES	3066	0.321502
26	ANNOTATED_BY	2213	0.232056
27	HAS_FIRST_CHILD	1534	0.160856
28	HAS_LAST_CHILD	1534	0.160856
29	HAS_ELEMENT	1522	0.159597

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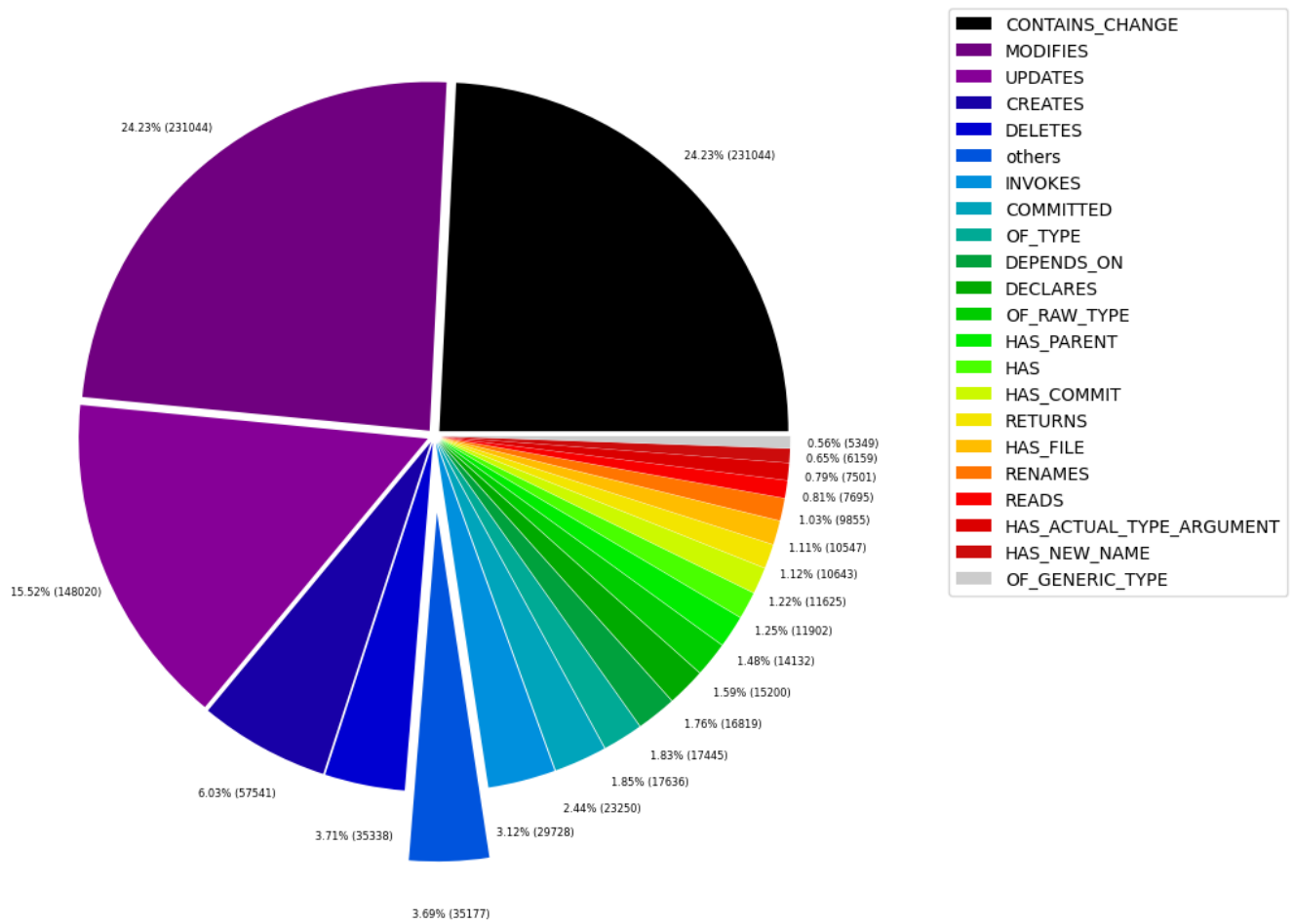
Relationship types (more than 0.5% overall)

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

24.23% (231044)
24.23% (231044)
15.52% (148020)
6.03% (57541)
3.71% (35338)
3.69% (35177)
3.12% (29728)
2.44% (23250)
1.85% (17636)
1.83% (17445)
1.76% (16819)
1.59% (15200)
1.48% (14132)
1.25% (11902)
1.22% (11625)
1.12% (10643)
1.11% (10547)
1.03% (9855)
0.81% (7695)
0.79% (7501)
0.65% (6159)
0.56% (5349)

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.
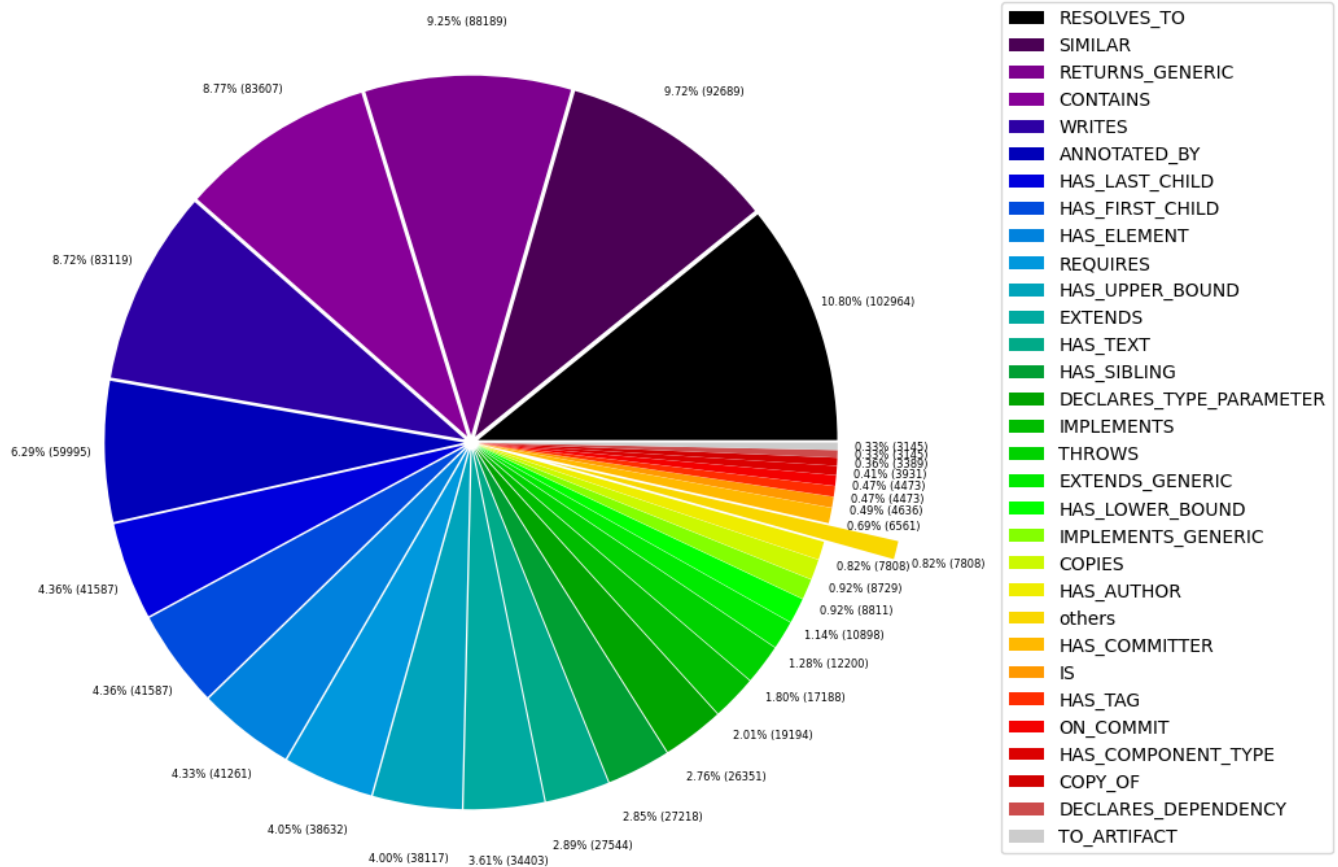This is essentially breaking down the "others" slice from the chart above.

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	EXCLUDES	1	0.000105
1	THROWS_GENERIC	5	0.000524
2	DESCRIBES	6	0.000629
3	HAS_ROOT_ELEMENT	7	0.000734
4	OF_NAMESPACE	12	0.001258
5	HAS_ATTRIBUTE	12	0.001258
6	HAS_GOAL	14	0.001468
7	HAS_EXECUTION	14	0.001468
8	USES_PLUGIN	16	0.001678
9	IS_ARTIFACT	16	0.001678
10	HAS_CONFIGURATION	16	0.001678
11	REQUIRES_TYPE_PARAMETER	17	0.001783
12	INCLUDES_CONCEPT	19	0.001992
13	HAS_BRANCH	23	0.002412
14	HAS_HEAD	24	0.002517
15	DECLARES_NAMESPACE	24	0.002517
16	REQUIRES_CONCEPT	28	0.002936
17	HAS_DEFAULT	34	0.003565
18	DECLARES_DEPENDENCY	116	0.012164
19	TO_ARTIFACT	116	0.012164
20	COPY_OF	125	0.013108
21	HAS_COMPONENT_TYPE	145	0.015205
22	ON_COMMIT	165	0.017302
23	HAS_TAG	165	0.017302
24	IS	171	0.017931
25	HAS_COMMITTER	242	0.025376
26	HAS_AUTHOR	288	0.030200
27	COPIES	322	0.033765
28	IMPLEMENTS_GENERIC	325	0.034080
29	HAS_LOWER_BOUND	402	0.042154

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
- RETURNS_GENERIC
- CONTAINS
- WRITES
- ANNOTATED_BY
- HAS_LAST_CHILD
- HAS_FIRST_CHILD
- HAS_ELEMENT
- REQUIRES
- HAS_UPPER_BOUND
- EXTENDS
- HAS_TEXT
- HAS_SIBLING
- DECLARES_TYPE_PARAMETER
- IMPLEMENTS
- THROWS
- EXTENDS_GENERIC
- HAS_LOWER_BOUND
- IMPLEMENTS_GENERIC
- COPIES
- HAS_AUTHOR
- others
- HAS_COMMITTER
- IS
- HAS_TAG
- ON_COMMIT
- HAS_COMPONENT_TYPE
- COPY_OF
- DECLARES_DEPENDENCY
- TO_ARTIFACT

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]	231044	11625	
1	[Git, Change]	MODIFIES	[File, Git]	231044	231044	
2	[Git, Change]	UPDATES	[File, Git]	148020	231044	
3	[Git, Change]	CREATES	[File, Git]	57541	231044	
4	[Git, Change]	DELETES	[File, Git]	35338	231044	
5	[Java, ByteCode, Method, Member]	INVOKES	[Java, ByteCode, Method, Member]	18206	10579	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14126	11625	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	11625	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	11625	288	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	11625	242	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10547	1	
11	[Git, Change]	RENAMES	[File, Git]	9855	231044	
12	[Java, ByteCode, Method, Member]	HAS	[Java, ByteCode, Parameter]	7306	10579	
13	[Java, ByteCode, Method, Member]	READS	[Java, ByteCode, Member, Field]	6917	10579	
14	[File, Git]	HAS_NEW_NAME	[File, Git]	6159	10547	
15	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	5730	11202	
16	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3218	6316	
17	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	2779	6542	
18	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2723	6542	
19	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2435	11202	
20	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, TypeVariable]	2230	6542	
21	[Java, ByteCode, Constructor, Method, Member]	WRITES	[Java, ByteCode, Member, Field]	1971	1647	
22	[Java, ByteCode, Method, Member]	RETURNS	[Type, File, Java, ByteCode, JavaType]	1892	10579	
23	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalType, Ext...]	1842	2225	
24	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	1819	11202	
25	[Java, ByteCode, Constructor, Method, Member]	INVOKES	[Java, ByteCode, Constructor, Method, Member]	1739	1647	
26	[Java, ByteCode, Constructor, Method, Member]	HAS	[Java, ByteCode, Parameter]	1710	1647	
27	[Java, ByteCode, Method, Member]	INVOKES	[Java, ByteCode, Constructor, Method, Member]	1703	10579	
28	[Java, ByteCode, Parameter]	ANNOTATED_BY	[Java, Value, ByteCode, Annotation]	1594	11202	
29	[Java, ByteCode, Constructor, Method, Member]	INVOKES	[Java, ByteCode, Method, Member]	1539	1647	

Graph Density

```
total_number_of_nodes (vertices): 307187
total_number_of_relationships (edges): 953650
-> total directed graph density: 1.0106127234802471e-05
-> total directed graph density in percent: 0.0010106127234802472
```