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]	256171	74.178930
1	[Java, ByteCode, Member, Method]	13201	3.822587
2	[Java, ByteCode, Parameter]	13029	3.772782
3	[Git, Commit]	12400	3.590643
4	[File, Git]	10780	3.121543
5	[Java, ByteCode, Bound, ParameterizedType]	7176	2.077940
6	[Java, ByteCode, Bound]	7118	2.061145
7	[Java, ByteCode, Member, Field]	3505	1.014936
8	[Java, ByteCode, Bound, WildcardType]	2908	0.842064
9	[Java, Value, ByteCode, Annotation]	2822	0.817161
10	[Xml, Element]	2144	0.620834
11	[Java, ByteCode, Member, Method, Constructor]	2049	0.593325
12	[Xml, Text]	1436	0.415820
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.317367
14	[Java, ByteCode, Member, Method, Lambda]	942	0.272773
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	899	0.260322
16	[Type, File, Java, ByteCode, Class]	864	0.250187
17	[Type, File, Java, ByteCode, JavaType]	748	0.216597
18	[Java, Value, ByteCode, Primitive]	652	0.188798
19	[Java, ByteCode, GenericDeclaration, Member, M...	569	0.164764
20	[Json, Key]	560	0.162158
21	[Value, Json, Scalar]	544	0.157525
22	[Type, File, Java, ByteCode, ExternalType]	396	0.114669
23	[Author, Git, Person]	298	0.086291
24	[Value, Array]	267	0.077315
25	[Committer, Git, Person]	252	0.072971
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.068338
27	[Value, Property]	202	0.058493
28	[Type, File, Java, ByteCode, Interface]	188	0.054439
29	[Java, Value, ByteCode, Class]	172	0.049806

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']

74.18% (256171)
0.59% (2049)
0.62% (2144)
0.82% (2822)
0.84% (2908)
1.01% (3505)
2.06% (7118)
2.08% (7176)
3.12% (10780)
3.49% (12039)
3.59% (12400)
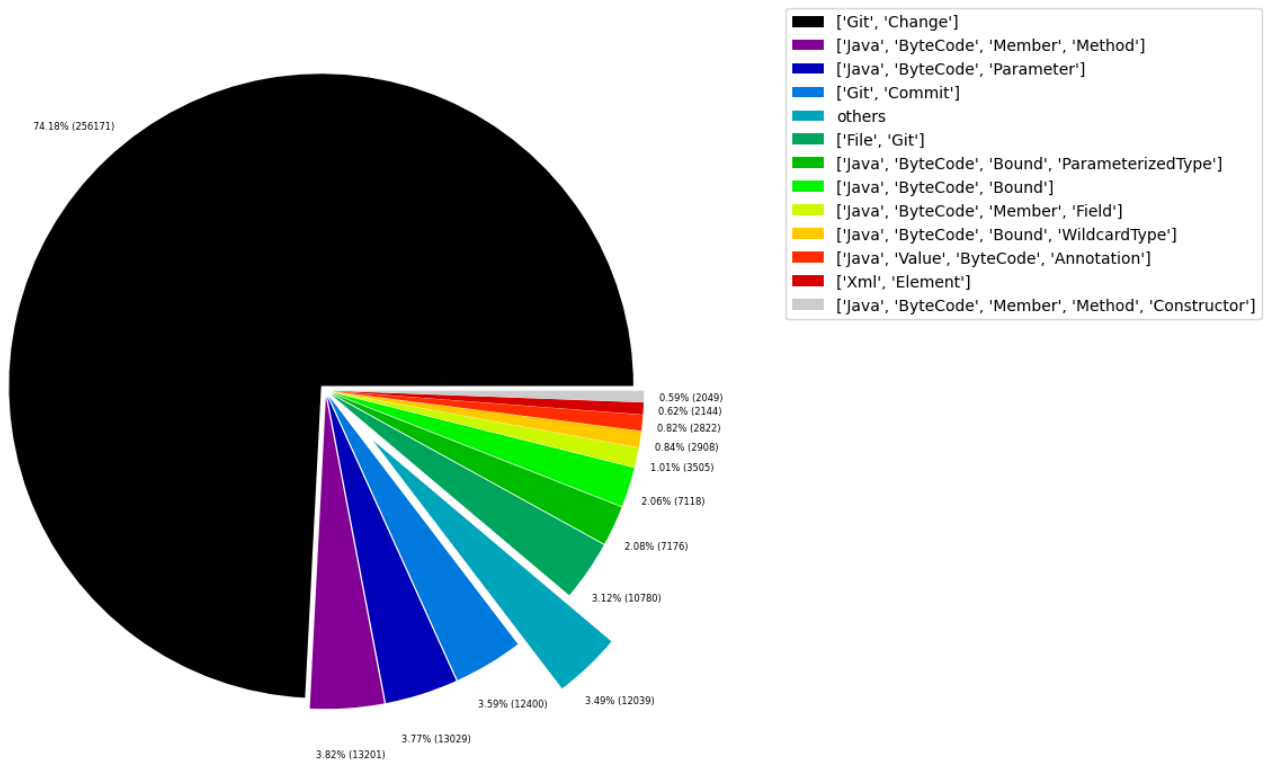3.77% (13029)
3.82% (13201)

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	0.000290
1	[Repository, File, Git]	1	0.000290
2	[File, Json]	2	0.000579
3	[File]	3	0.000869
4	[Java, ByteCode, GenericDeclaration, Member, M...	4	0.001158
5	[Maven, Exclusion]	5	0.001448
6	[Value, Array, Json]	6	0.001737
7	[Type, File, Java, ByteCode, Void]	9	0.002606
8	[File, Maven, Xml, Pom, Document]	9	0.002606
9	[Java, ManifestSection]	9	0.002606
10	[File, Java, Manifest]	9	0.002606
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002606
12	[File, Java, ServiceLoader]	10	0.002896
13	[File, Java, Properties]	12	0.003475
14	[Maven, ExecutionGoal]	16	0.004633
15	[Maven, PluginExecution]	16	0.004633
16	[Xml, Attribute]	18	0.005212
17	[jQAssistant, Rule, Concept]	19	0.005502
18	[Maven, Plugin]	21	0.006081
19	[Maven, Configuration]	21	0.006081
20	[Type, File, Java, ByteCode, Enum]	28	0.008108
21	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008397
22	[Git, Branch]	36	0.010424
23	[Xml, Namespace]	36	0.010424
24	[Type, File, Java, ByteCode, Annotation]	43	0.012451
25	[File, Directory]	44	0.012741
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.024613
27	[Artifact, Maven]	92	0.026640
28	[Type, File, Java, ByteCode, ExternalAnnotatio...	93	0.026930
29	[Value, ManifestEntry]	120	0.034748

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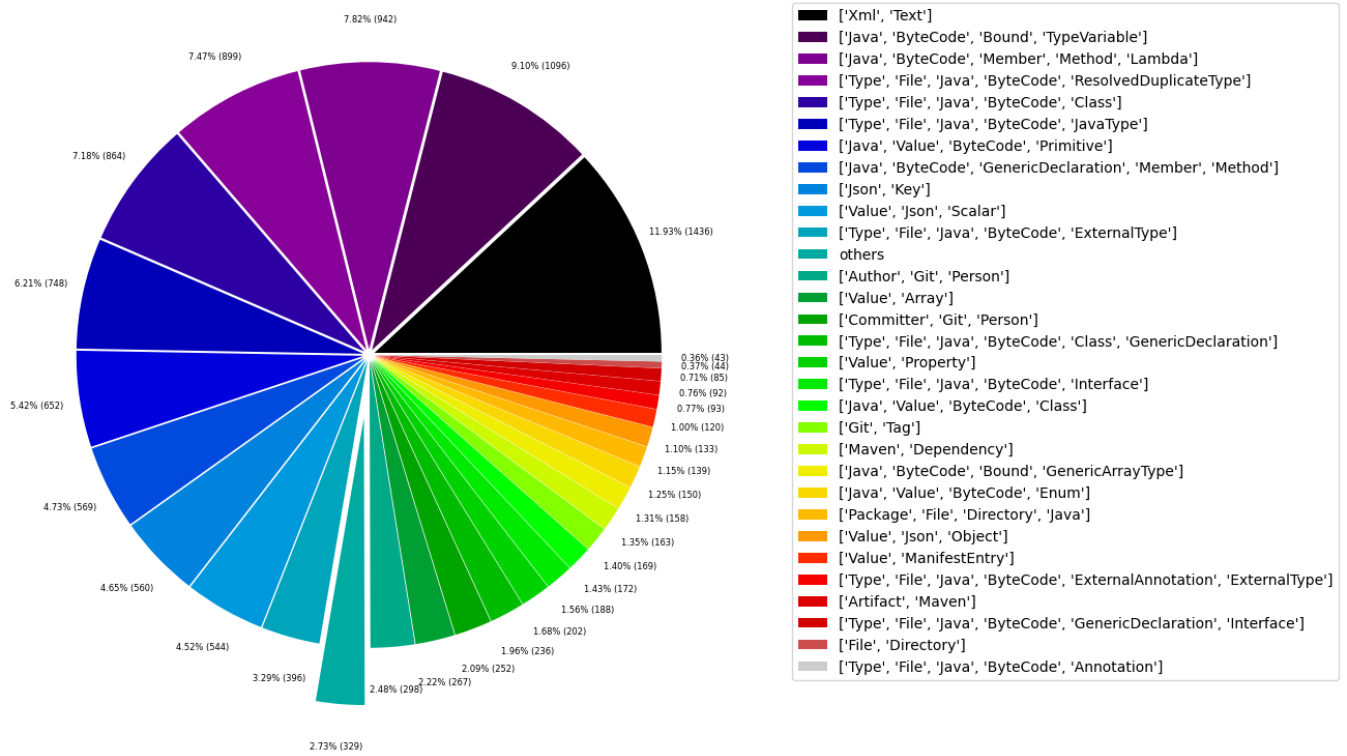- ['Type', 'File', 'Java', 'ByteCode', 'Annotation']

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.
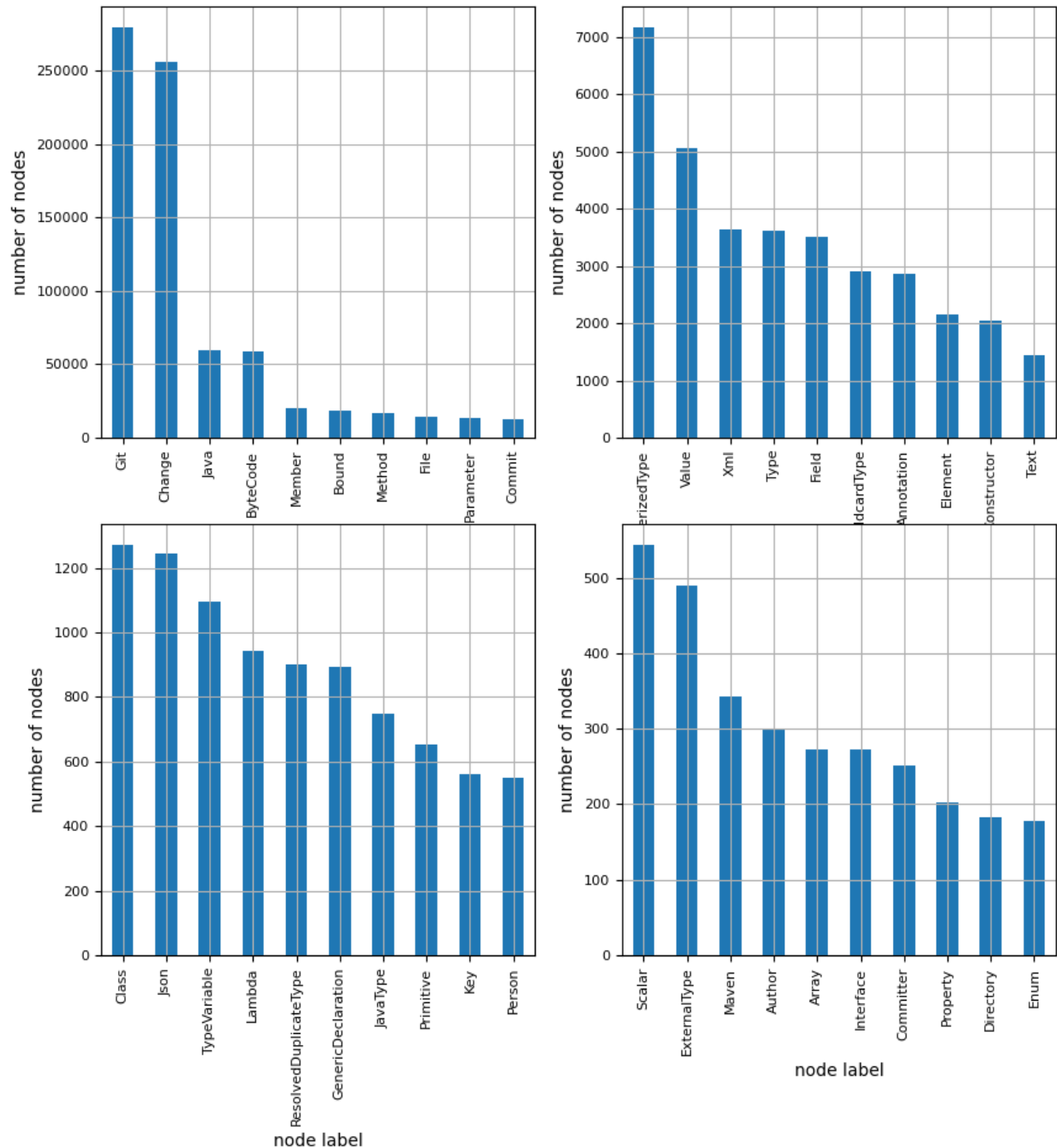
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	280107	81.110030
1	Change	256171	74.178930
2	Java	59357	17.187889
3	ByteCode	59169	17.133450
4	Member	20270	5.869544
5	Bound	18456	5.344267
6	Method	16765	4.854608
7	File	14636	4.238118
8	Parameter	13029	3.772782
9	Commit	12400	3.590643
10	ParameterizedType	7176	2.077940
11	Value	5068	1.467531
12	Xml	3643	1.054896
13	Type	3618	1.047657
14	Field	3505	1.014936
15	WildcardType	2908	0.842064
16	Annotation	2865	0.829612
17	Element	2144	0.620834
18	Constructor	2053	0.594483
19	Text	1436	0.415820
20	Class	1272	0.368331
21	Json	1245	0.360512
22	TypeVariable	1096	0.317367
23	Lambda	942	0.272773
24	ResolvedDuplicateType	899	0.260322
25	GenericDeclaration	894	0.258874
26	JavaType	748	0.216597
27	Primitive	652	0.188798
28	Key	560	0.162158
29	Person	550	0.159262
30	Scalar	544	0.157525
31	ExternalType	489	0.141599
32	Maven	343	0.099322
33	Author	298	0.086291
34	Array	273	0.079052
35	Interface	273	0.079052
36	Committer	252	0.072971
37	Property	202	0.058493
38	Directory	183	0.052991
39	Enum	178	0.051543

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: 1071642
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	256171	23.904532
1	MODIFIES	256171	23.904532
2	UPDATES	167870	15.664746
3	CREATES	60933	5.685947
4	DELETES	37874	3.534203
5	INVOKES	35641	3.325831
6	COMMITTED	24800	2.314206
7	DEPENDS_ON	21931	2.046486
8	OF_TYPE	21394	1.996376
9	DECLARES	20715	1.933015
10	OF_RAW_TYPE	17093	1.595029
11	HAS_PARENT	15066	1.405880
12	HAS	14125	1.318071
13	RETURNS	12578	1.173713
14	HAS_COMMIT	12400	1.157103
15	HAS_FILE	10780	1.005933
16	RENAMES	10506	0.980365
17	READS	9164	0.855136
18	HAS_ACTUAL_TYPE_ARGUMENT	8288	0.773393
19	HAS_NEW_NAME	6220	0.580418
20	OF_GENERIC_TYPE	5906	0.551117
21	RESOLVES_TO	5176	0.482997
22	SIMILAR	3977	0.371113
23	WRITES	3818	0.356276
24	CONTAINS	3814	0.355902
25	RETURNS_GENERIC	3545	0.330801
26	ANNOTATED_BY	2810	0.262214
27	REQUIRES	2174	0.202866
28	HAS_FIRST_CHILD	2144	0.200067
29	HAS_LAST_CHILD	2144	0.200067

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>
```

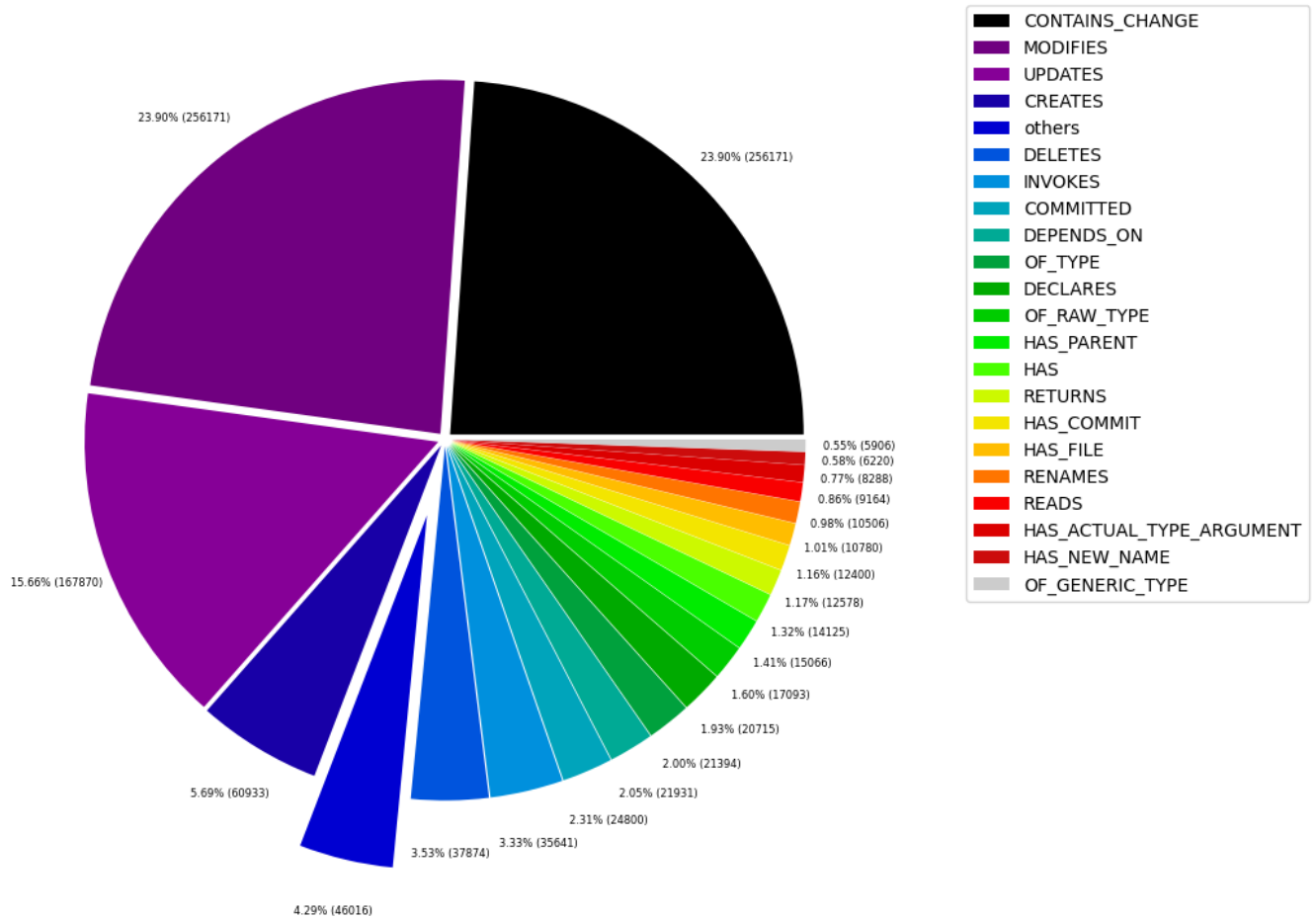Relationship types (more than 0.5% overall)

23.90% (256171)

23.90% (256171)

0.55% (5906)
0.58% (6220)
0.77% (8288)
0.86% (9164)
0.98% (10506)
1.01% (10780)
1.16% (12400)
1.17% (12578)
1.32% (14125)
1.41% (15066)
1.60% (17093)
1.93% (20715)
2.00% (21394)
2.05% (21931)
2.31% (24800)

15.66% (167870)

5.69% (60933)

3.53% (37874)
3.33% (35641)

4.29% (46016)

Legend:
- 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

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.
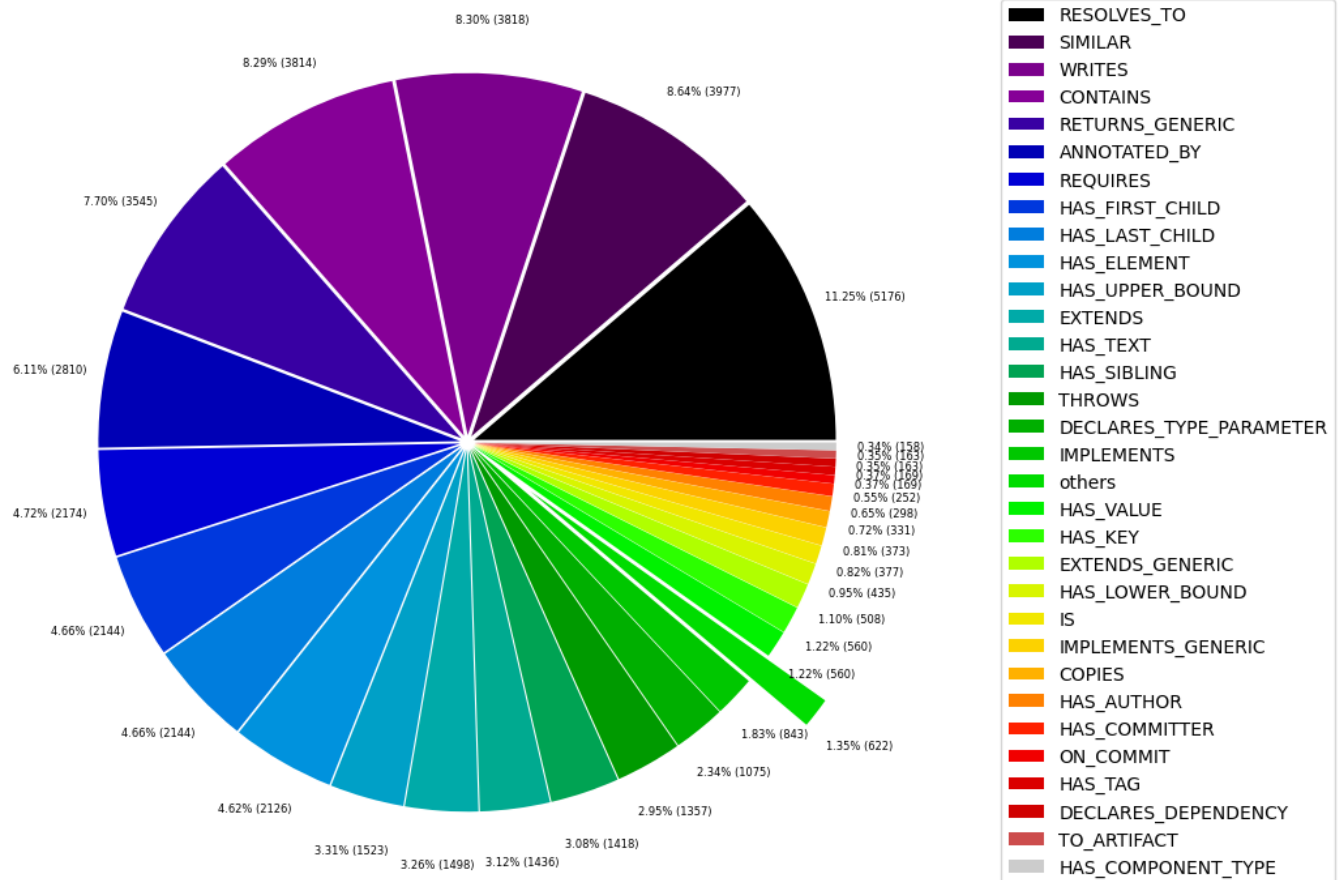
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000093
1	THROWS_GENERIC	5	0.000467
2	EXCLUDES	5	0.000467
3	DESCRIBES	9	0.000840
4	HAS_ROOT_ELEMENT	11	0.001026
5	HAS_GOAL	16	0.001493
6	HAS_EXECUTION	16	0.001493
7	OF_NAMESPACE	18	0.001680
8	HAS_ATTRIBUTE	18	0.001680
9	INCLUDES_CONCEPT	19	0.001773
10	USES_PLUGIN	21	0.001960
11	REQUIRES_TYPE_PARAMETER	21	0.001960
12	IS_ARTIFACT	21	0.001960
13	HAS_CONFIGURATION	21	0.001960
14	REQUIRES_CONCEPT	28	0.002613
15	HAS_DEFAULT	36	0.003359
16	HAS_BRANCH	36	0.003359
17	DECLARES_NAMESPACE	36	0.003359
18	HAS_HEAD	37	0.003453
19	CONTAINS_VALUE	121	0.011291
20	COPY_OF	126	0.011758
21	HAS_COMPONENT_TYPE	158	0.014744
22	TO_ARTIFACT	163	0.015210
23	DECLARES_DEPENDENCY	163	0.015210
24	ON_COMMIT	169	0.015770
25	HAS_TAG	169	0.015770
26	HAS_COMMITTER	252	0.023515
27	HAS_AUTHOR	298	0.027808
28	COPIES	331	0.030887
29	IMPLEMENTS_GENERIC	373	0.034806

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
- WRITES
- CONTAINS
- RETURNS_GENERIC
- ANNOTATED_BY
- REQUIRES
- HAS_FIRST_CHILD
- HAS_LAST_CHILD
- HAS_ELEMENT
- HAS_UPPER_BOUND
- EXTENDS
- HAS_TEXT
- HAS_SIBLING
- THROWS
- DECLARES_TYPE_PARAMETER
- IMPLEMENTS
- 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

Pie chart values: 11.25% (5176), 8.64% (3977), 8.30% (3818), 8.29% (3814), 7.70% (3545), 6.11% (2810), 4.72% (2174), 4.66% (2144), 4.66% (2144), 4.62% (2126), 3.31% (1523), 3.26% (1498), 3.12% (1436), 3.08% (1418), 2.95% (1357), 2.34% (1075), 1.83% (843), 1.35% (622), 1.22% (560), 1.22% (560), 1.10% (508), 0.95% (435), 0.82% (377), 0.81% (373), 0.72% (331), 0.65% (298), 0.55% (252), 0.37% (169), 0.37% (169), 0.35% (163), 0.35% (163), 0.34% (158)

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]	256171	12400	
1	[Git, Change]	MODIFIES	[File, Git]	256171	256171	
2	[Git, Change]	UPDATES	[File, Git]	167870	256171	
3	[Git, Change]	CREATES	[File, Git]	60933	256171	
4	[Git, Change]	DELETES	[File, Git]	37874	256171	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21705	13107	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15057	12400	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12400	1	
8	[Author, Git, Person]	COMMITTED	[Git, Commit]	12400	298	
9	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12400	252	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10780	1	
11	[Git, Change]	RENAMES	[File, Git]	10506	256171	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8360	13107	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8218	13107	
14	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6371	13029	
15	[File, Git]	HAS_NEW_NAME	[File, Git]	6220	10780	
16	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Method]	5103	814	
17	[Type, File, Java, ByteCode, Class]	DEPENDS_ON	[Type, File, Java, ByteCode, JavaType]	4480	814	
18	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3570	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	2049	
24	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalAnnotatio...]	2341	2822	
25	[Type, File, Java, ByteCode, Class]	DECLARES	[Java, ByteCode, Member, Field]	2203	814	
26	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2126	2144	
27	[Java, ByteCode, Member, Method, Constructor]	INVOKES	[Java, ByteCode, Member, Method, Constructor]	2117	2049	
28	[Type, File, Java, ByteCode, Class, GenericDec...]	DECLARES	[Java, ByteCode, Member, Method]	2115	231	
29	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2091	13107	

Graph Density

```
total_number_of_nodes (vertices): 345342
total_number_of_relationships (edges): 1071642
-> total directed graph density: 8.985705568891573e-06
-> total directed graph density in percent: 0.0008985705568891573
```