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`.
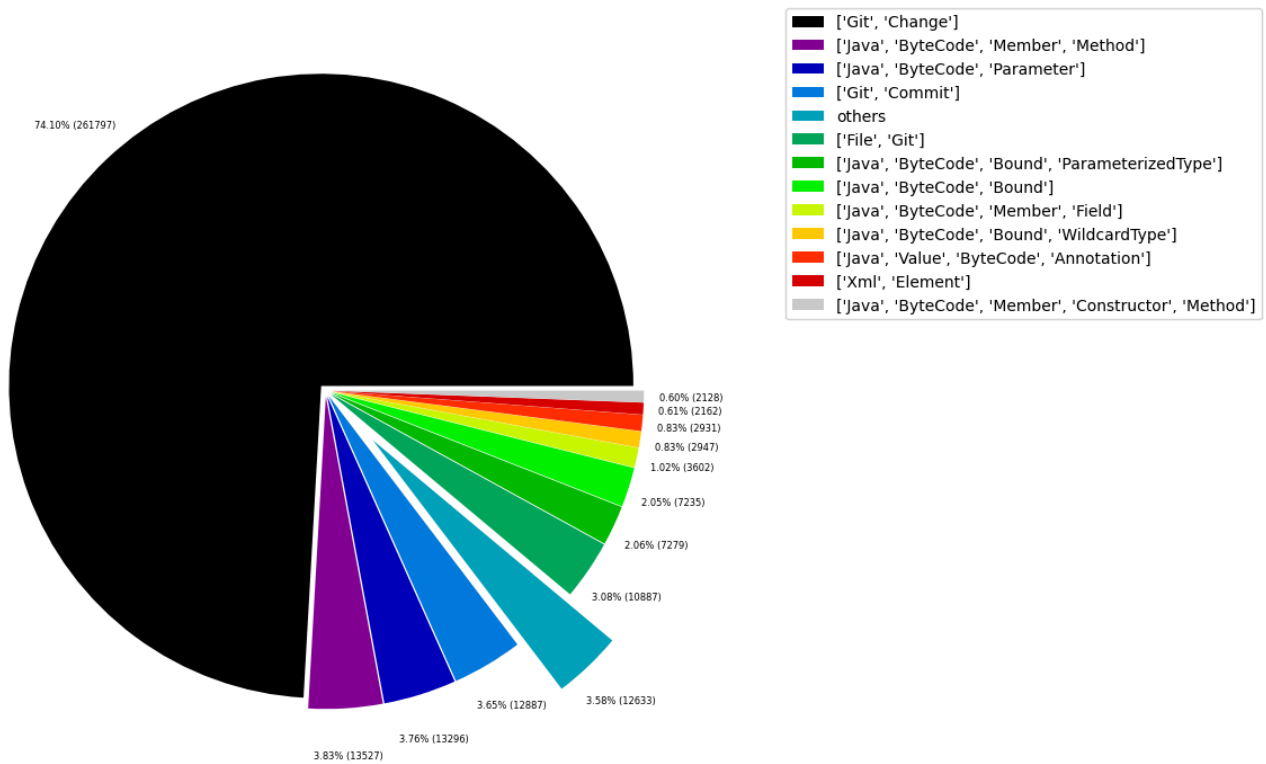
	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Git, Change]	261797	74.098174
1	[Java, ByteCode, Member, Method]	13527	3.828638
2	[Java, ByteCode, Parameter]	13296	3.763257
3	[Git, Commit]	12887	3.647495
4	[File, Git]	10887	3.081421
5	[Java, ByteCode, Bound, ParameterizedType]	7279	2.060225
6	[Java, ByteCode, Bound]	7235	2.047771
7	[Java, ByteCode, Member, Field]	3602	1.019498
8	[Java, ByteCode, Bound, WildcardType]	2947	0.834109
9	[Java, Value, ByteCode, Annotation]	2931	0.829581
10	[Xml, Element]	2162	0.611925
11	[Java, ByteCode, Member, Constructor, Method]	2128	0.602302
12	[Xml, Text]	1450	0.410403
13	[Java, ByteCode, TypeVariable, Bound]	1111	0.314454
14	[Java, ByteCode, Member, Method, Lambda]	972	0.275112
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	888	0.251337
16	[Type, File, Java, Class, ByteCode]	846	0.239449
17	[Json, Key]	702	0.198692
18	[Value, Json, Scalar]	685	0.193880
19	[Java, Value, ByteCode, Primitive]	672	0.190201
20	[Type, File, Java, ByteCode, JavaType]	649	0.183691
21	[Java, ByteCode, Member, Method, GenericDeclar...	578	0.163595
22	[Type, File, Java, ByteCode, ExternalType]	411	0.116328
23	[Author, Git, Person]	298	0.084345
24	[Value, Array]	287	0.081232
25	[Committer, Git, Person]	252	0.071325
26	[Type, File, Java, Class, ByteCode, GenericDec...	237	0.067080
27	[Value, Property]	211	0.059721
28	[Type, File, Java, ByteCode, Interface]	190	0.053777
29	[Java, Value, Class, ByteCode]	189	0.053494

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', 'Constructor', 'Method']

74.10% (261797)
0.60% (2128)
0.61% (2162)
0.83% (2931)
0.83% (2947)
1.02% (3602)
2.05% (7235)
2.06% (7279)
3.08% (10887)
3.58% (12633)
3.65% (12887)
3.76% (13296)
3.83% (13527)

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.
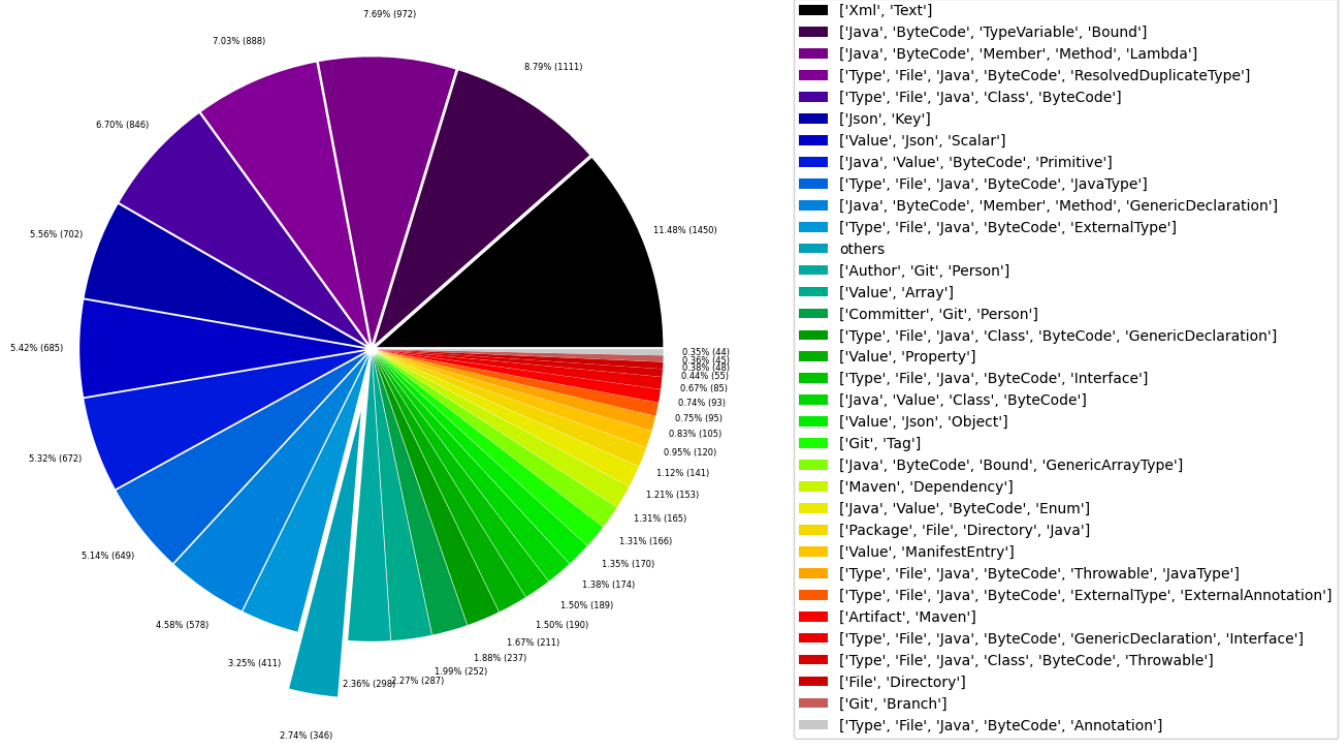
	nodeLabels	nodesWithThatLabels	nodesWithThatLabelsPercent
0	[Analyze, Task, jQAssistant]	1	0.000283
1	[Package, File, Json, NPM]	1	0.000283
2	[Repository, File, Git]	1	0.000283
3	[File, TS, Scan]	1	0.000283
4	[File, Json]	2	0.000566
5	[File]	3	0.000849
6	[Java, ByteCode, Member, Constructor, Method, ...	4	0.001132
7	[Maven, Exclusion]	5	0.001415
8	[Value, Array, Json]	6	0.001698
9	[Dependency, NPM]	7	0.001981
10	[Type, File, Java, ByteCode, Void]	9	0.002547
11	[File, Maven, Xml, Pom, Document]	9	0.002547
12	[Java, ManifestSection]	9	0.002547
13	[File, Java, Manifest]	9	0.002547
14	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002547
15	[File, Java, ServiceLoader]	10	0.002830
16	[File, Java, Properties]	12	0.003396
17	[Maven, PluginExecution]	16	0.004529
18	[Maven, ExecutionGoal]	16	0.004529
19	[Xml, Attribute]	18	0.005095
20	[jQAssistant, Rule, Concept]	19	0.005378
21	[Maven, Configuration]	21	0.005944
22	[Maven, Plugin]	21	0.005944
23	[Type, File, Java, ByteCode, Throwable, Extern...	21	0.005944
24	[Type, File, Java, ByteCode, Throwable, Resolv...	22	0.006227
25	[Type, File, Java, ByteCode, Enum]	28	0.007925
26	[Type, File, Java, ByteCode, PrimitiveType]	30	0.008491
27	[Xml, Namespace]	36	0.010189
28	[Type, File, Java, ByteCode, Annotation]	44	0.012454
29	[Git, Branch]	45	0.012737

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', 'TypeVariable', 'Bound']
- ['Java', 'ByteCode', 'Member', 'Method', 'Lambda']
- ['Type', 'File', 'Java', 'ByteCode', 'ResolvedDuplicateType']
- ['Type', 'File', 'Java', 'Class', 'ByteCode']
- ['Json', 'Key']
- ['Value', 'Json', 'Scalar']
- ['Java', 'Value', 'ByteCode', 'Primitive']
- ['Type', 'File', 'Java', 'ByteCode', 'JavaType']
- ['Java', 'ByteCode', 'Member', 'Method', 'GenericDeclaration']
- ['Type', 'File', 'Java', 'ByteCode', 'ExternalType']
- others
- ['Author', 'Git', 'Person']
- ['Value', 'Array']
- ['Committer', 'Git', 'Person']
- ['Type', 'File', 'Java', 'Class', 'ByteCode', 'GenericDeclaration']
- ['Value', 'Property']
- ['Type', 'File', 'Java', 'ByteCode', 'Interface']
- ['Java', 'Value', 'Class', 'ByteCode']
- ['Value', 'Json', 'Object']
- ['Git', 'Tag']
- ['Java', 'ByteCode', 'Bound', 'GenericArrayType']
- ['Maven', 'Dependency']
- ['Java', 'Value', 'ByteCode', 'Enum']
- ['Package', 'File', 'Directory', 'Java']
- ['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', 'Class', 'ByteCode', '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.
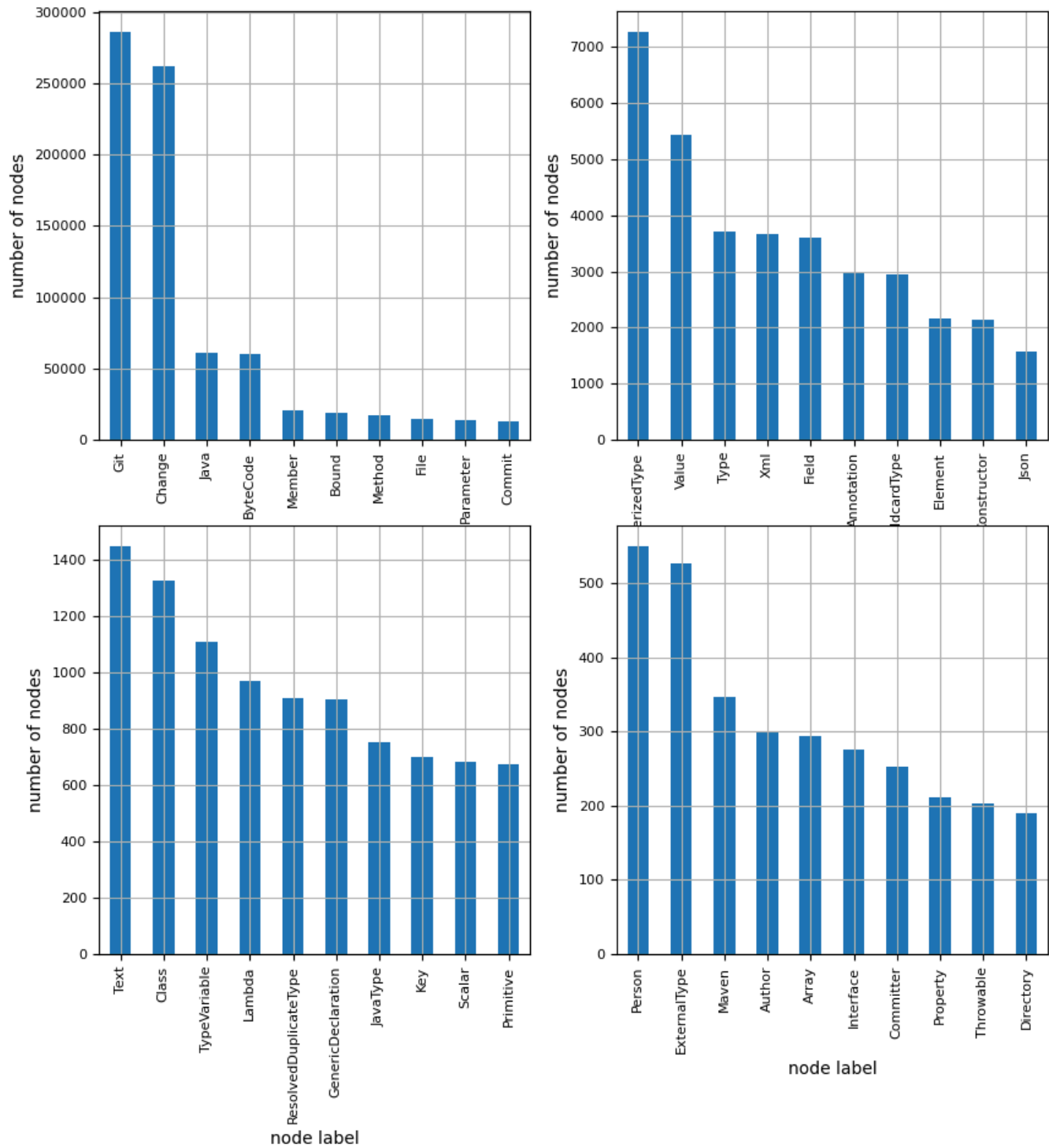
	nodeLabel	nodesWithThatLabel	nodesWithThatLabelPercent
0	Git	286337	81.043896
1	Change	261797	74.098174
2	Java	60695	17.178916
3	ByteCode	60505	17.125139
4	Member	20811	5.890278
5	Bound	18738	5.303543
6	Method	17209	4.870780
7	File	14848	4.202530
8	Parameter	13296	3.763257
9	Commit	12887	3.647495
10	ParameterizedType	7279	2.060225
11	Value	5428	1.536324
12	Type	3715	1.051482
13	Xml	3675	1.040160
14	Field	3602	1.019498
15	Annotation	2975	0.842034
16	WildcardType	2947	0.834109
17	Element	2162	0.611925
18	Constructor	2132	0.603434
19	Json	1570	0.444368
20	Text	1450	0.410403
21	Class	1327	0.375590
22	TypeVariable	1111	0.314454
23	Lambda	972	0.275112
24	ResolvedDuplicateType	910	0.257563
25	GenericDeclaration	904	0.255865
26	JavaType	754	0.213410
27	Key	702	0.198692
28	Scalar	685	0.193880
29	Primitive	672	0.190201
30	Person	550	0.155670
31	ExternalType	527	0.149160
32	Maven	346	0.097931
33	Author	298	0.084345
34	Array	293	0.082930
35	Interface	275	0.077835
36	Committer	252	0.071325
37	Property	211	0.059721
38	Throwable	203	0.057456
39	Directory	189	0.053494

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: 1096433
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	261797	23.877154
1	MODIFIES	261797	23.877154
2	UPDATES	172615	15.743324
3	CREATES	61715	5.628707
4	DELETES	38029	3.468429
5	INVOKES	36628	3.340651
6	COMMITTED	25774	2.350714
7	DEPENDS_ON	22488	2.051015
8	OF_TYPE	21922	1.999393
9	DECLARES	21272	1.940109
10	OF_RAW_TYPE	17348	1.582222
11	HAS_PARENT	15599	1.422704
12	HAS	14432	1.316268
13	HAS_COMMIT	12887	1.175357
14	RETURNS	12853	1.172256
15	HAS_FILE	10887	0.992947
16	RENAMES	10562	0.963306
17	READS	9378	0.855319
18	HAS_ACTUAL_TYPE_ARGUMENT	8407	0.766759
19	HAS_NEW_NAME	6240	0.569118
20	OF_GENERIC_TYPE	5991	0.546408
21	RESOLVES_TO	5277	0.481288
22	SIMILAR	4078	0.371933
23	WRITES	3944	0.359712
24	CONTAINS	3936	0.358982
25	RETURNS_GENERIC	3593	0.327699
26	ANNOTATED_BY	2919	0.266227
27	REQUIRES	2230	0.203387
28	HAS_FIRST_CHILD	2162	0.197185
29	HAS_LAST_CHILD	2162	0.197185

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
- HAS_COMMIT
- RETURNS
- HAS_FILE
- RENAMES
- READS
- HAS_ACTUAL_TYPE_ARGUMENT
- HAS_NEW_NAME
- OF_GENERIC_TYPE

Slice labels:
- 23.88% (261797)
- 23.88% (261797)
- 15.74% (172615)
- 5.63% (61715)
- 4.36% (47812)
- 3.47% (38029)
- 3.34% (36628)
- 2.35% (25774)
- 2.05% (22488)
- 2.00% (21922)
- 1.94% (21272)
- 1.58% (17348)
- 1.42% (15599)
- 1.32% (14432)
- 1.18% (12887)
- 1.17% (12853)
- 0.99% (10887)
- 0.96% (10562)
- 0.86% (9378)
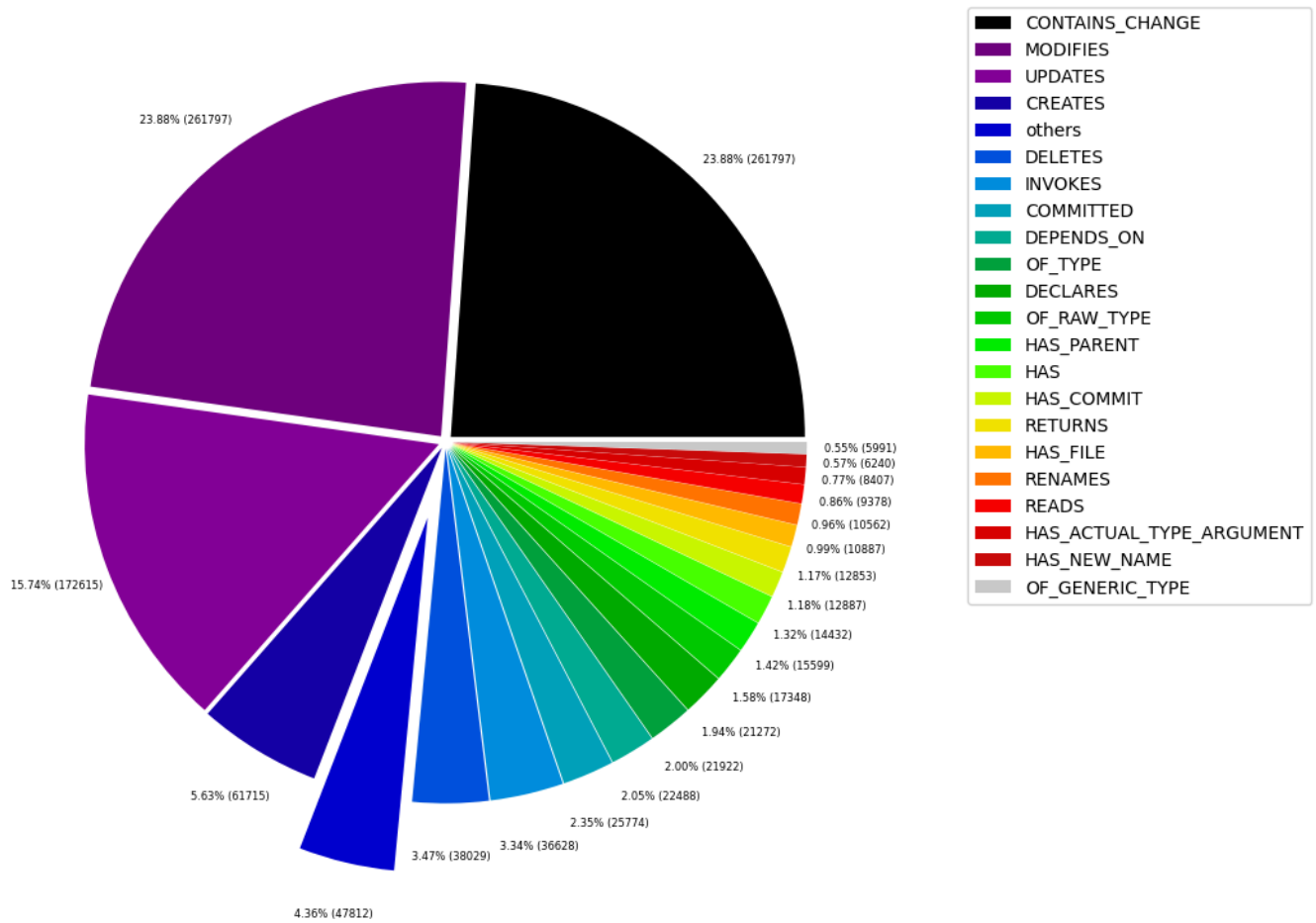- 0.77% (8407)
- 0.57% (6240)
- 0.55% (5991)

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.
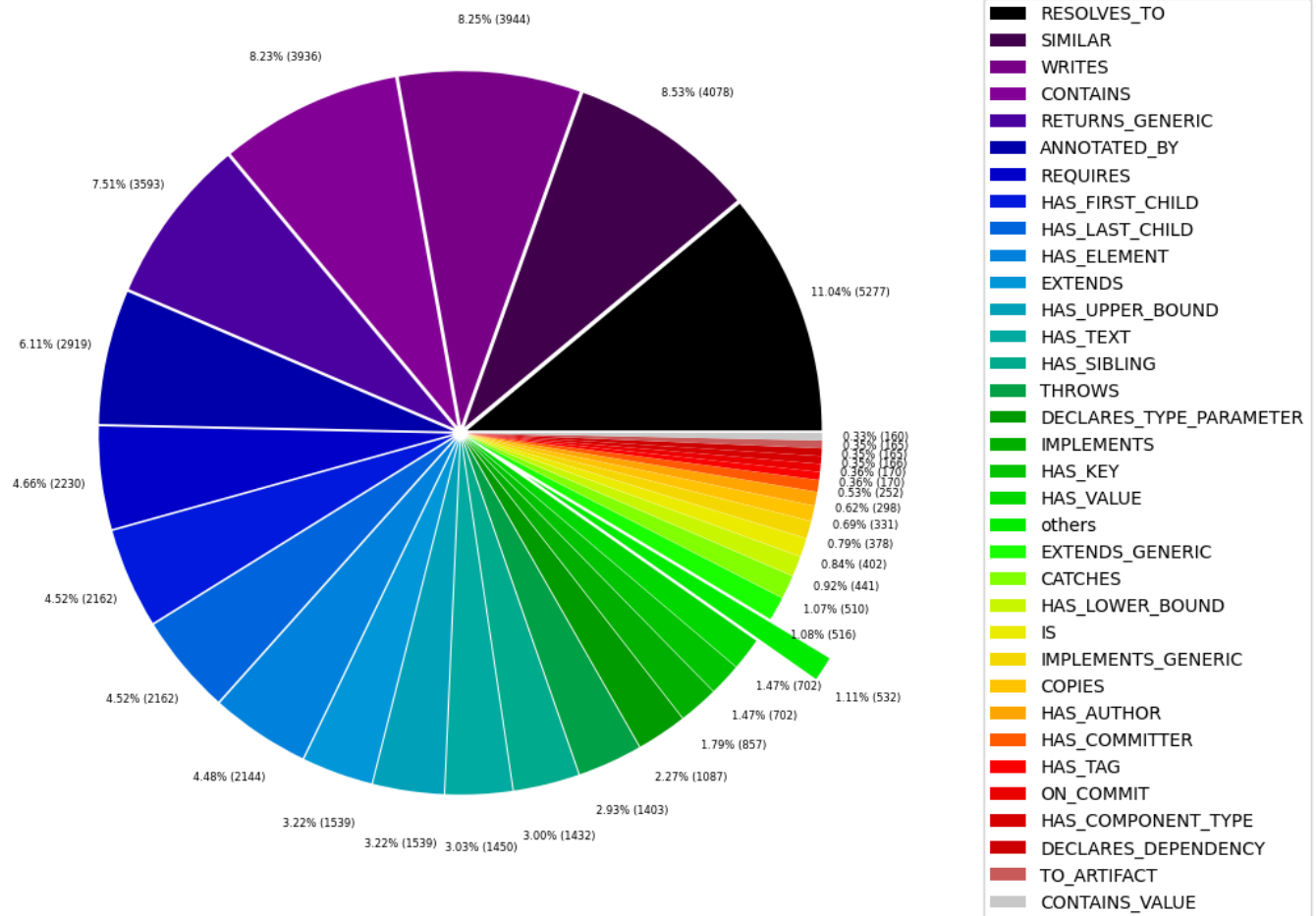
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	HAS_PROPERTY	1	0.000091
1	THROWS_GENERIC	5	0.000456
2	EXCLUDES	5	0.000456
3	DECLARES_DEV_DEPENDENCY	7	0.000638
4	DESCRIBES	9	0.000821
5	HAS_ROOT_ELEMENT	11	0.001003
6	HAS_GOAL	16	0.001459
7	HAS_EXECUTION	16	0.001459
8	OF_NAMESPACE	18	0.001642
9	HAS_ATTRIBUTE	18	0.001642
10	INCLUDES_CONCEPT	19	0.001733
11	IS_ARTIFACT	21	0.001915
12	HAS_CONFIGURATION	21	0.001915
13	USES_PLUGIN	21	0.001915
14	REQUIRES_TYPE_PARAMETER	24	0.002189
15	REQUIRES_CONCEPT	28	0.002554
16	DECLARES_NAMESPACE	36	0.003283
17	HAS_DEFAULT	39	0.003557
18	HAS_BRANCH	45	0.004104
19	HAS_HEAD	46	0.004195
20	COPY_OF	126	0.011492
21	CONTAINS_VALUE	160	0.014593
22	DECLARES_DEPENDENCY	165	0.015049
23	TO_ARTIFACT	165	0.015049
24	HAS_COMPONENT_TYPE	166	0.015140
25	ON_COMMIT	170	0.015505
26	HAS_TAG	170	0.015505
27	HAS_COMMITTER	252	0.022984
28	HAS_AUTHOR	298	0.027179
29	COPIES	331	0.030189

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
- EXTENDS
- HAS_UPPER_BOUND
- HAS_TEXT
- HAS_SIBLING
- THROWS
- DECLARES_TYPE_PARAMETER
- IMPLEMENTS
- HAS_KEY
- HAS_VALUE
- others
- EXTENDS_GENERIC
- CATCHES
- HAS_LOWER_BOUND
- IS
- IMPLEMENTS_GENERIC
- COPIES
- HAS_AUTHOR
- HAS_COMMITTER
- HAS_TAG
- ON_COMMIT
- HAS_COMPONENT_TYPE
- DECLARES_DEPENDENCY
- TO_ARTIFACT
- CONTAINS_VALUE

Pie slice labels: 8.25% (3944), 8.23% (3936), 8.53% (4078), 7.51% (3593), 11.04% (5277), 6.11% (2919), 4.66% (2230), 4.52% (2162), 4.52% (2162), 4.48% (2144), 3.22% (1539), 3.22% (1539), 3.03% (1450), 3.00% (1432), 2.93% (1403), 2.27% (1087), 1.79% (857), 1.47% (702), 1.47% (702), 1.11% (532), 1.08% (516), 1.07% (510), 0.92% (441), 0.84% (402), 0.79% (378), 0.69% (331), 0.62% (298), 0.53% (252), 0.36% (170), 0.36% (170), 0.35% (165), 0.35% (165), 0.33% (160)

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	numbe
0	[Git, Change]	MODIFIES	[File, Git]	261797	261797	
1	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	261797	12887	
2	[Git, Change]	UPDATES	[File, Git]	172615	261797	
3	[Git, Change]	CREATES	[File, Git]	61715	261797	
4	[Git, Change]	DELETES	[File, Git]	38029	261797	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	22255	13428	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	15590	12887	
7	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	12887	1	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	12887	252	
9	[Author, Git, Person]	COMMITTED	[Git, Commit]	12887	298	
10	[Repository, File, Git]	HAS_FILE	[File, Git]	10887	1	
11	[Git, Change]	RENAMES	[File, Git]	10562	261797	
12	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8495	13428	
13	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8420	13428	
14	[File, Git]	HAS_NEW_NAME	[File, Git]	6240	10887	
15	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6160	13296	
16	[Java, ByteCode, Bound]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3465	7235	
17	[Java, ByteCode, Bound, ParameterizedType]	OF_RAW_TYPE	[Type, File, Java, ByteCode, JavaType]	3110	7279	
18	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, Bound, WildcardType]	2947	7279	
19	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound, ParameterizedType]	2695	13296	
20	[Java, ByteCode, Member, Constructor, Method]	WRITES	[Java, ByteCode, Member, Field]	2519	2128	
21	[Java, ByteCode, Bound, ParameterizedType]	HAS_ACTUAL_TYPE_ARGUMENT	[Java, ByteCode, TypeVariable, Bound]	2474	7279	
22	[Java, Value, ByteCode, Annotation]	OF_TYPE	[Type, File, Java, ByteCode, ExternalType, Ext...	2445	2931	
23	[Java, ByteCode, Member, Constructor, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2195	2128	
24	[Xml, Element]	HAS_ELEMENT	[Xml, Element]	2144	2162	
25	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2113	13428	
26	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2107	13428	
27	[Java, ByteCode, Member, Constructor, Method]	HAS	[Java, ByteCode, Parameter]	2075	2128	
28	[Java, ByteCode, Parameter]	OF_GENERIC_TYPE	[Java, ByteCode, Bound]	2056	13296	
29	[Java, ByteCode, Member, Method, Lambda]	INVOKES	[Java, ByteCode, Member, Method]	1954	972	

Graph Density

```
total_number_of_nodes (vertices): 353311
total_number_of_relationships (edges): 1096433
-> total directed graph density: 8.783528393377125e-06
-> total directed graph density in percent: 0.0008783528393377125
```