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]	235897	72.781434
1	[Java, ByteCode, Member, Method]	13171	4.063656
2	[Java, ByteCode, Parameter]	13000	4.010897
3	[Git, Commit]	11787	3.636650
4	[File, Git]	10598	3.269807
5	[Java, ByteCode, Bound, ParameterizedType]	7166	2.210930
6	[Java, ByteCode, Bound]	7112	2.194269
7	[Java, ByteCode, Member, Field]	3490	1.076772
8	[Java, ByteCode, Bound, WildcardType]	2906	0.896590
9	[Java, Value, ByteCode, Annotation]	2821	0.870365
10	[Xml, Element]	2144	0.661490
11	[Java, ByteCode, Member, Constructor, Method]	2040	0.629402
12	[Xml, Text]	1436	0.443050
13	[Java, ByteCode, Bound, TypeVariable]	1096	0.338149
14	[Java, ByteCode, Member, Method, Lambda]	936	0.288785
15	[Type, File, Java, ByteCode, ResolvedDuplicate...	900	0.277678
16	[Type, File, Java, ByteCode, Class]	861	0.265645
17	[Type, File, Java, ByteCode, JavaType]	746	0.230164
18	[Java, Value, ByteCode, Primitive]	650	0.200545
19	[Java, ByteCode, GenericDeclaration, Member, M...	569	0.175554
20	[Json, Key]	555	0.171234
21	[Value, Json, Scalar]	539	0.166298
22	[Type, File, Java, ByteCode, ExternalType]	396	0.122178
23	[Author, Git, Person]	289	0.089165
24	[Value, Array]	267	0.082378
25	[Committer, Git, Person]	243	0.074973
26	[Type, File, Java, ByteCode, Class, GenericDec...	236	0.072813
27	[Value, Property]	202	0.062323
28	[Type, File, Java, ByteCode, Interface]	187	0.057695
29	[Java, Value, ByteCode, Class]	172	0.053067

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']

72.78% (235897)
0.63% (2040)
0.66% (2144)
0.87% (2821)
0.90% (2906)
1.08% (3490)
2.19% (7112)
2.21% (7166)
3.27% (10598)
3.64% (11787)
3.70% (11985)
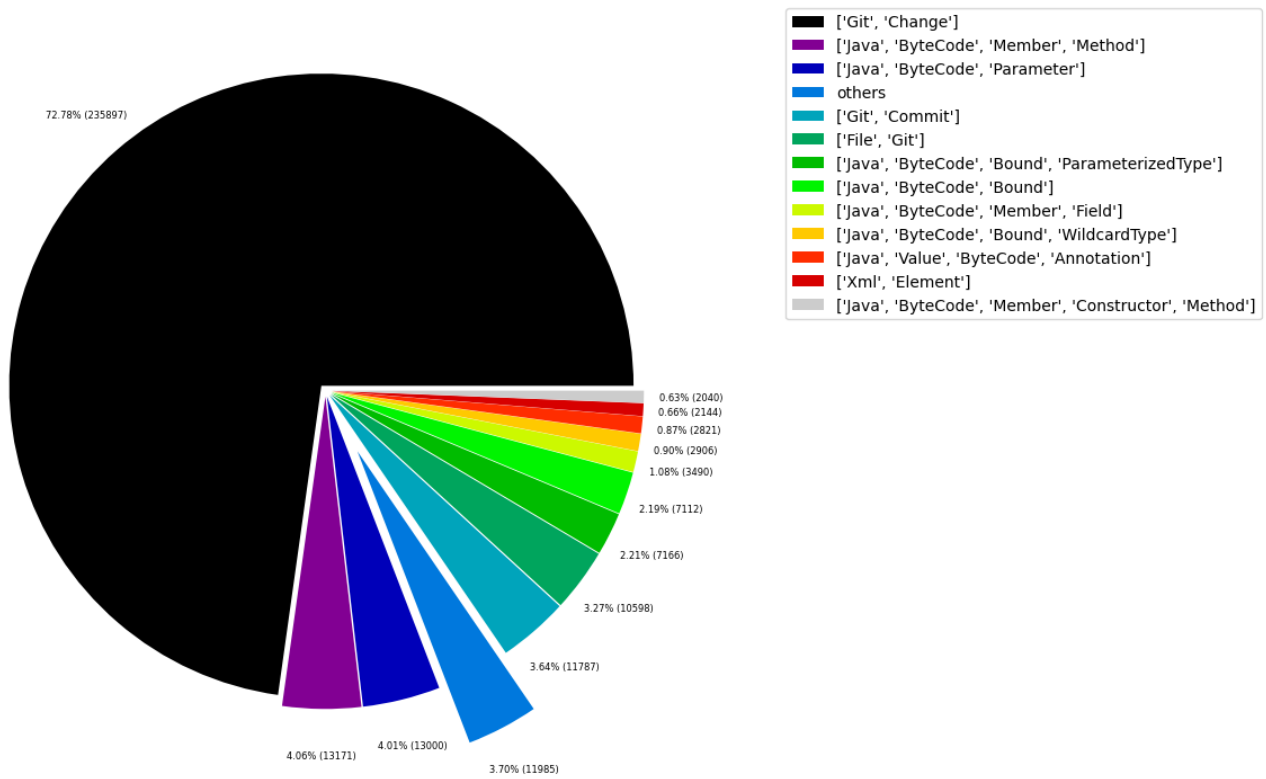4.01% (13000)
4.06% (13171)

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	0.000309
1	[File, TS, Scan]	1	0.000309
2	[File, Json]	2	0.000617
3	[File]	3	0.000926
4	[Java, ByteCode, GenericDeclaration, Member, C...	4	0.001234
5	[Maven, Exclusion]	5	0.001543
6	[Value, Array, Json]	6	0.001851
7	[Type, File, Java, ByteCode, Void]	9	0.002777
8	[File, Maven, Xml, Pom, Document]	9	0.002777
9	[Java, ManifestSection]	9	0.002777
10	[File, Java, Manifest]	9	0.002777
11	[Artifact, File, Jar, Archive, Zip, Java]	9	0.002777
12	[File, Java, ServiceLoader]	10	0.003085
13	[File, Java, Properties]	12	0.003702
14	[Maven, ExecutionGoal]	16	0.004936
15	[Maven, PluginExecution]	16	0.004936
16	[Xml, Attribute]	18	0.005554
17	[jQAssistant, Rule, Concept]	19	0.005862
18	[Maven, Plugin]	21	0.006479
19	[Maven, Configuration]	21	0.006479
20	[Git, Branch]	26	0.008022
21	[Type, File, Java, ByteCode, Enum]	28	0.008639
22	[Type, File, Java, ByteCode, PrimitiveType]	29	0.008947
23	[Xml, Namespace]	36	0.011107
24	[Type, File, Java, ByteCode, Annotation]	43	0.013267
25	[File, Directory]	44	0.013575
26	[Type, File, Java, ByteCode, GenericDeclaratio...	85	0.026225
27	[Artifact, Maven]	92	0.028385
28	[Type, File, Java, ByteCode, ExternalAnnotatio...	93	0.028693
29	[Value, ManifestEntry]	120	0.037024

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)

11.98% (1436)
9.14% (1096)
7.81% (936)
7.51% (900)
7.18% (861)
6.22% (746)
5.42% (650)
4.75% (569)
4.63% (555)
4.50% (539)
3.30% (396)
2.41% (289)
2.36% (283)
2.23% (267)
2.03% (243)
1.97% (236)
1.69% (202)
1.56% (187)
1.44% (172)
1.39% (167)
1.36% (163)
1.32% (158)
1.25% (150)
1.16% (139)
1.10% (132)
1.00% (120)
0.78% (93)
0.77% (92)
0.71% (85)
0.37% (44)
0.33% (39)
0.30% (36)

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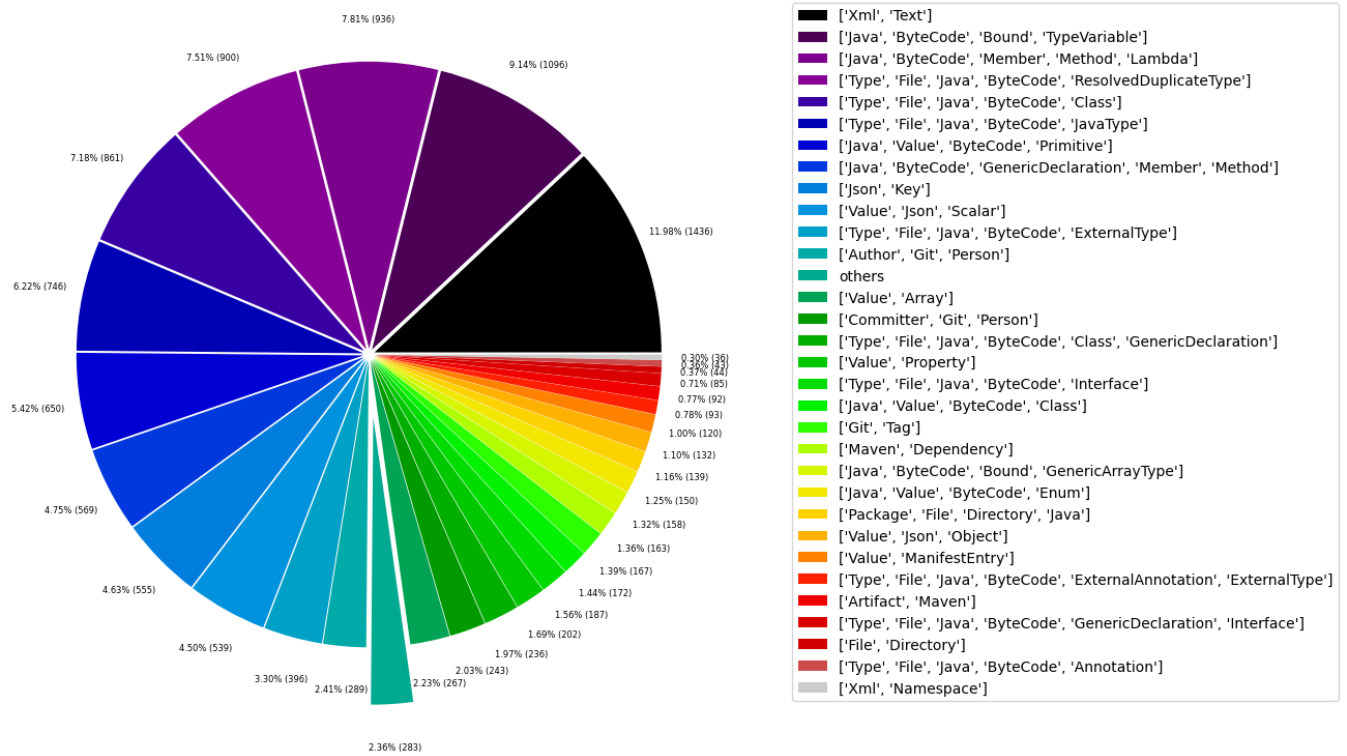- ['Xml', 'Namespace']

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	259007	79.911575
1	Change	235897	72.781434
2	Java	59242	18.277968
3	ByteCode	59054	18.219964
4	Member	20210	6.235403
5	Bound	18438	5.688686
6	Method	16720	5.158631
7	File	14449	4.457958
8	Parameter	13000	4.010897
9	Commit	11787	3.636650
10	ParameterizedType	7166	2.210930
11	Value	5059	1.560856
12	Xml	3643	1.123977
13	Type	3613	1.114721
14	Field	3490	1.076772
15	WildcardType	2906	0.896590
16	Annotation	2864	0.883632
17	Element	2144	0.661490
18	Constructor	2044	0.630636
19	Text	1436	0.443050
20	Class	1269	0.391525
21	Json	1234	0.380727
22	TypeVariable	1096	0.338149
23	Lambda	936	0.288785
24	ResolvedDuplicateType	900	0.277678
25	GenericDeclaration	894	0.275826
26	JavaType	746	0.230164
27	Primitive	650	0.200545
28	Key	555	0.171234
29	Scalar	539	0.166298
30	Person	532	0.164138
31	ExternalType	489	0.150871
32	Maven	343	0.105826
33	Author	289	0.089165
34	Array	273	0.084229
35	Interface	272	0.083920
36	Committer	243	0.074973
37	Property	202	0.062323
38	Directory	183	0.056461
39	Enum	178	0.054918

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
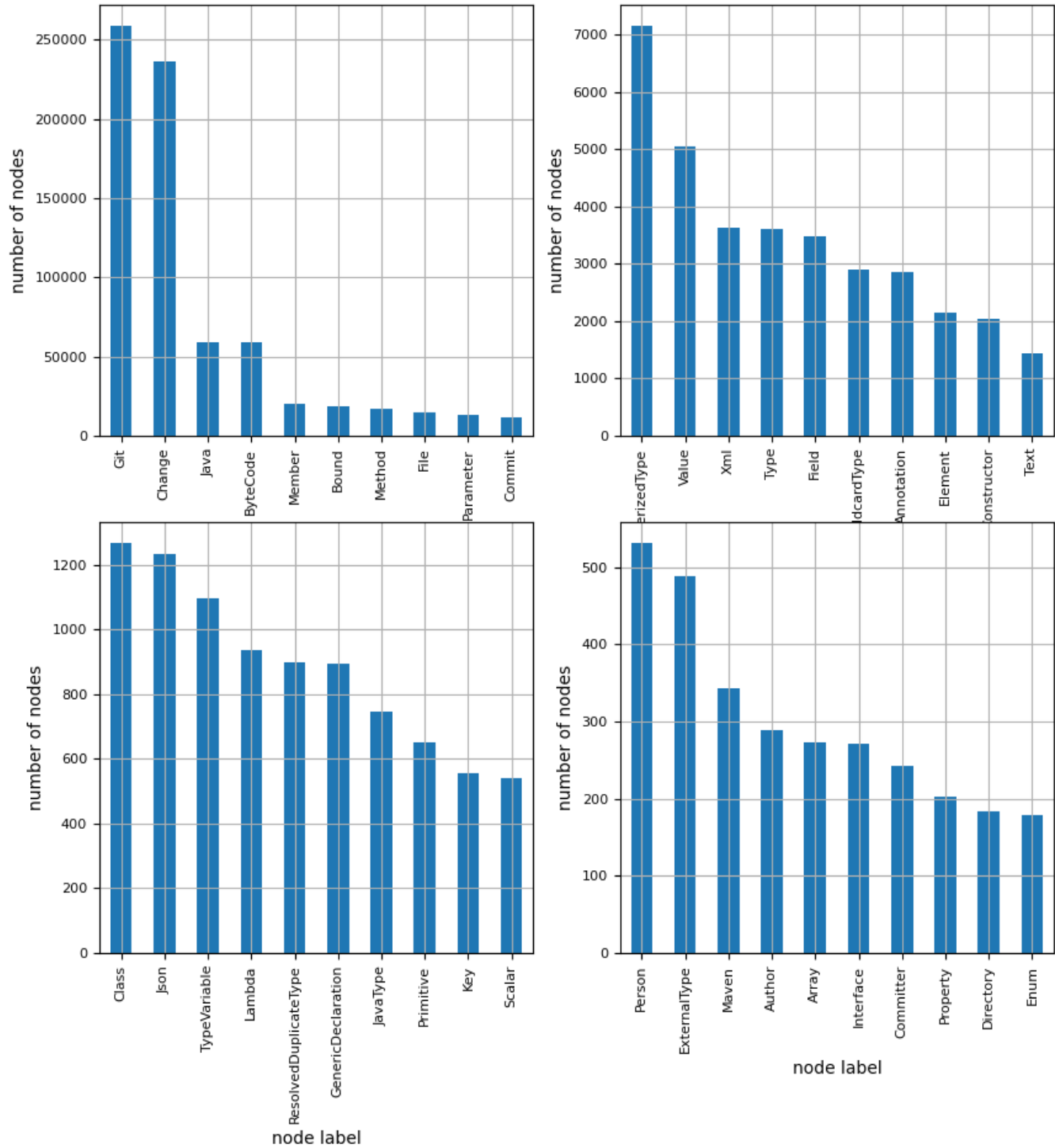
Relationship Types

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: 982675
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	235897	24.005597
1	MODIFIES	235897	24.005597
2	UPDATES	151660	15.433383
3	CREATES	58219	5.924543
4	DELETES	35985	3.661943
5	INVOKES	35594	3.622154
6	COMMITTED	23574	2.398962
7	DEPENDS_ON	21873	2.225863
8	OF_TYPE	21347	2.172336
9	DECLARES	20651	2.101509
10	OF_RAW_TYPE	17077	1.737808
11	HAS_PARENT	14327	1.457959
12	HAS	14094	1.434248
13	RETURNS	12547	1.276821
14	RENAMES	9967	1.014272
15	READS	9078	0.923805
16	HAS_ACTUAL_TYPE_ARGUMENT	8276	0.842191
17	HAS_NEW_NAME	6165	0.627369
18	OF_GENERIC_TYPE	5902	0.600605
19	RESOLVES_TO	5231	0.532322
20	SIMILAR	3968	0.403796
21	CONTAINS	3805	0.387208
22	WRITES	3803	0.387005
23	RETURNS_GENERIC	3543	0.360546
24	ANNOTATED_BY	2809	0.285852
25	REQUIRES	2173	0.221131
26	HAS_FIRST_CHILD	2144	0.218180
27	HAS_LAST_CHILD	2144	0.218180
28	HAS_ELEMENT	2126	0.216348
29	HAS_UPPER_BOUND	1521	0.154782

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
- others
- DELETES
- INVOKES
- COMMITTED
- DEPENDS_ON
- OF_TYPE
- DECLARES
- OF_RAW_TYPE
- HAS_PARENT
- HAS
- RETURNS
- RENAMES
- READS
- HAS_ACTUAL_TYPE_ARGUMENT
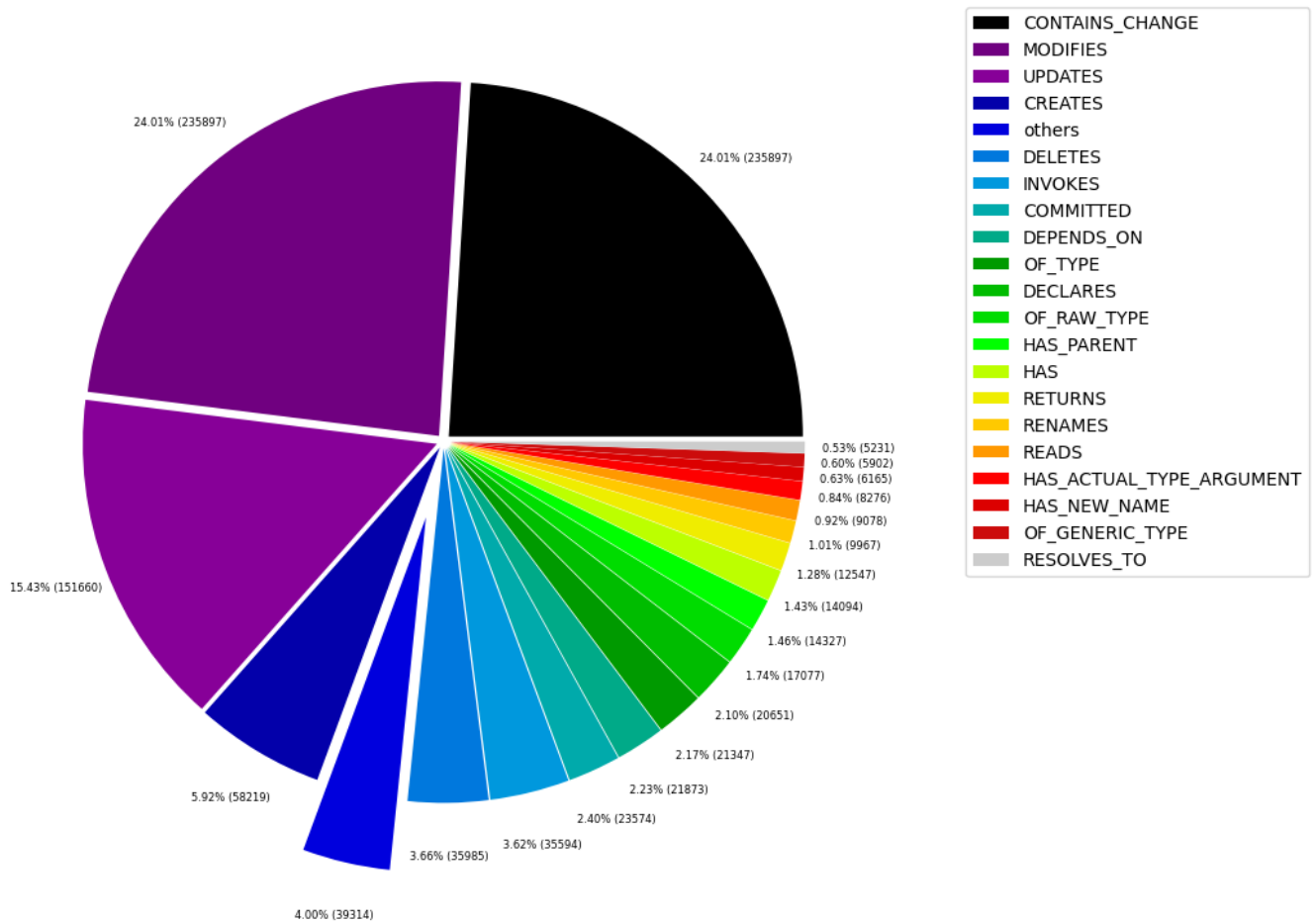- HAS_NEW_NAME
- OF_GENERIC_TYPE
- RESOLVES_TO

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	0.000102
1	THROWS_GENERIC	5	0.000509
2	EXCLUDES	5	0.000509
3	DESCRIBES	9	0.000916
4	HAS_ROOT_ELEMENT	10	0.001018
5	HAS_GOAL	16	0.001628
6	HAS_EXECUTION	16	0.001628
7	OF_NAMESPACE	18	0.001832
8	HAS_ATTRIBUTE	18	0.001832
9	INCLUDES_CONCEPT	19	0.001933
10	USES_PLUGIN	21	0.002137
11	REQUIRES_TYPE_PARAMETER	21	0.002137
12	IS_ARTIFACT	21	0.002137
13	HAS_CONFIGURATION	21	0.002137
14	HAS_HEAD	26	0.002646
15	REQUIRES_CONCEPT	28	0.002849
16	DECLARES_NAMESPACE	36	0.003663
17	HAS_DEFAULT	36	0.003663
18	CONTAINS_VALUE	120	0.012212
19	COPY_OF	125	0.012720
20	HAS_COMPONENT_TYPE	158	0.016079
21	TO_ARTIFACT	163	0.016587
22	DECLARES_DEPENDENCY	163	0.016587
23	ON_COMMIT	167	0.016994
24	COPIES	324	0.032971
25	IMPLEMENTS_GENERIC	373	0.037958
26	IS	377	0.038365
27	HAS_LOWER_BOUND	435	0.044267
28	EXTENDS_GENERIC	508	0.051696
29	HAS_VALUE	555	0.056478

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)
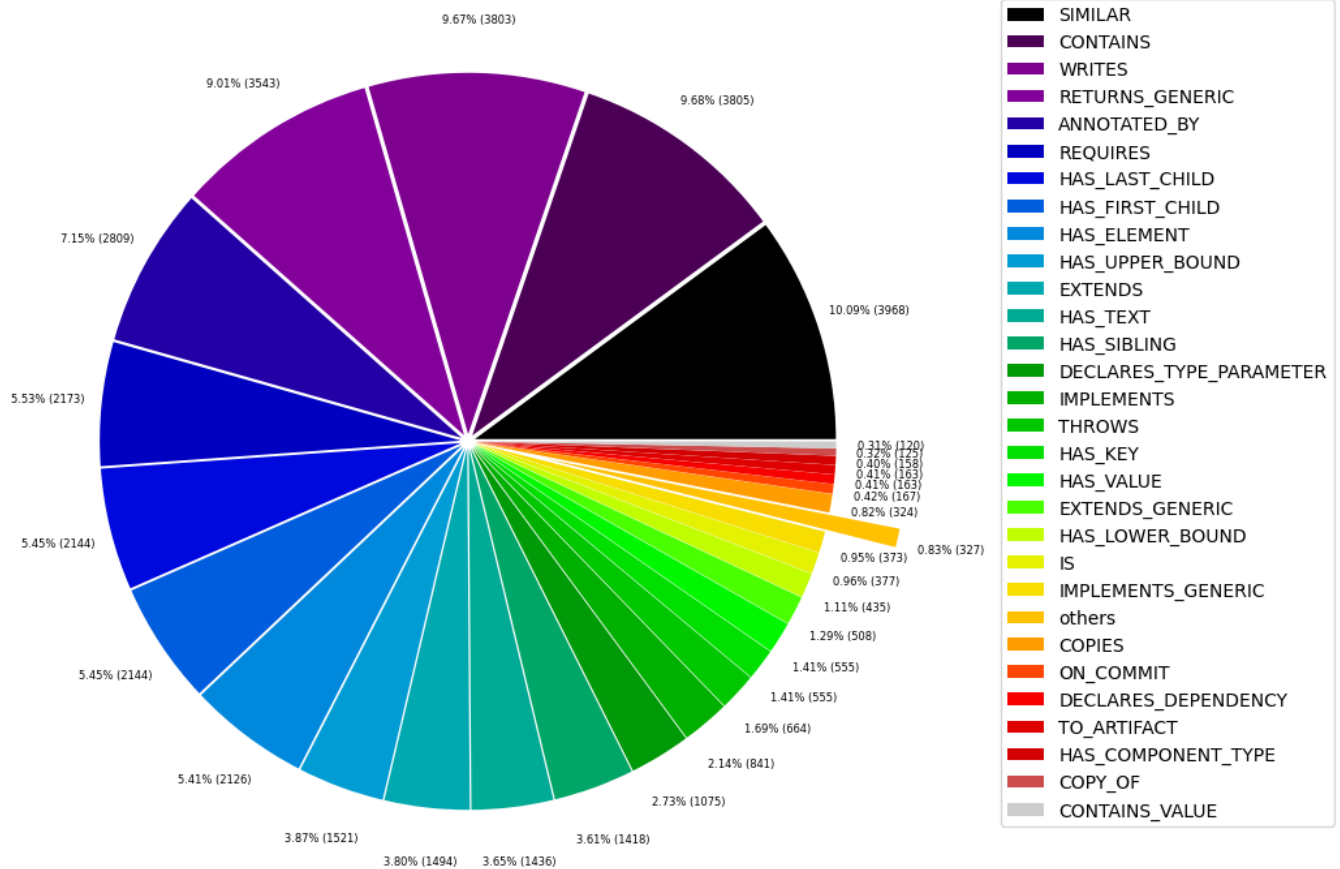
Node labels with their relationships

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]	235897	11787	
1	[Git, Change]	MODIFIES	[File, Git]	235897	235897	
2	[Git, Change]	UPDATES	[File, Git]	151660	235897	
3	[Git, Change]	CREATES	[File, Git]	58219	235897	
4	[Git, Change]	DELETES	[File, Git]	35985	235897	
5	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Method]	21567	13056	
6	[Git, Commit]	HAS_PARENT	[Git, Commit]	14318	11787	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	11787	289	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	11787	243	
9	[Git, Change]	RENAMES	[File, Git]	9967	235897	
10	[Java, ByteCode, Member, Method]	HAS	[Java, ByteCode, Parameter]	8316	13056	
11	[Java, ByteCode, Member, Method]	READS	[Java, ByteCode, Member, Field]	8129	13056	
12	[Java, ByteCode, Parameter]	OF_TYPE	[Type, File, Java, ByteCode, JavaType]	6360	13000	
13	[File, Git]	HAS_NEW_NAME	[File, Git]	6165	10598	
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]	2107	2040	
23	[Java, ByteCode, Member, Method]	RETURNS	[Type, File, Java, ByteCode, JavaType]	2086	13056	
24	[Java, ByteCode, Member, Method]	INVOKES	[Java, ByteCode, Member, Constructor, Method]	2030	13056	
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]	1885	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]	1796	2040	

Graph Density

```
total_number_of_nodes (vertices): 324117
total_number_of_relationships (edges): 982675
-> total directed graph density: 9.354221004728924e-06
-> total directed graph density in percent: 0.0009354221004728924
```