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]	81525	78.299830
1	[Git, Commit]	10604	10.184500
2	[File, Git]	5495	5.277615
3	[Git, Tag]	1313	1.261057
4	[Author, Git, Person]	1219	1.170776
5	[Json, Key]	668	0.641574
6	[Json, Value, Scalar]	603	0.579145
7	[Committer, Git, Person]	371	0.356323
8	[NPM, Dependency]	338	0.324629
9	[Type, TS, Primitive]	291	0.279488
10	[Type, TS, Declared]	276	0.265081
11	[TS, ExternalDeclaration]	214	0.205534
12	[Type, TS, Literal]	136	0.130620
13	[Json, Value, Object]	133	0.127738
14	[Type, TS, Union]	119	0.114292
15	[Type, TS, ObjectMember]	101	0.097004
16	[NPM, Script]	91	0.087400
17	[TS, Property]	65	0.062429
18	[TS, Function]	47	0.045141
19	[Type, TS, FunctionParameter]	40	0.038418
20	[Type, Object, TS]	39	0.037457
21	[Git, Branch]	35	0.033615
22	[File, Directory]	34	0.032655
23	[Type, TS, Function]	34	0.032655
24	[TS, Parameter]	33	0.031695
25	[Package, File, Json, NPM]	29	0.027853
26	[TS, Variable]	24	0.023051
27	[TS, ExternalModule]	23	0.022090
28	[Value, TS, Literal]	20	0.019209
29	[jQAssistant, Rule, Concept]	19	0.018248

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']
- ['Git', 'Commit']
- ['File', 'Git']
- others
- ['Git', 'Tag']
- ['Author', 'Git', 'Person']
- ['Json', 'Key']
- ['Json', 'Value', 'Scalar']

78.30% (81525)
0.58% (603)
0.64% (668)
1.17% (1219)
1.26% (1313)
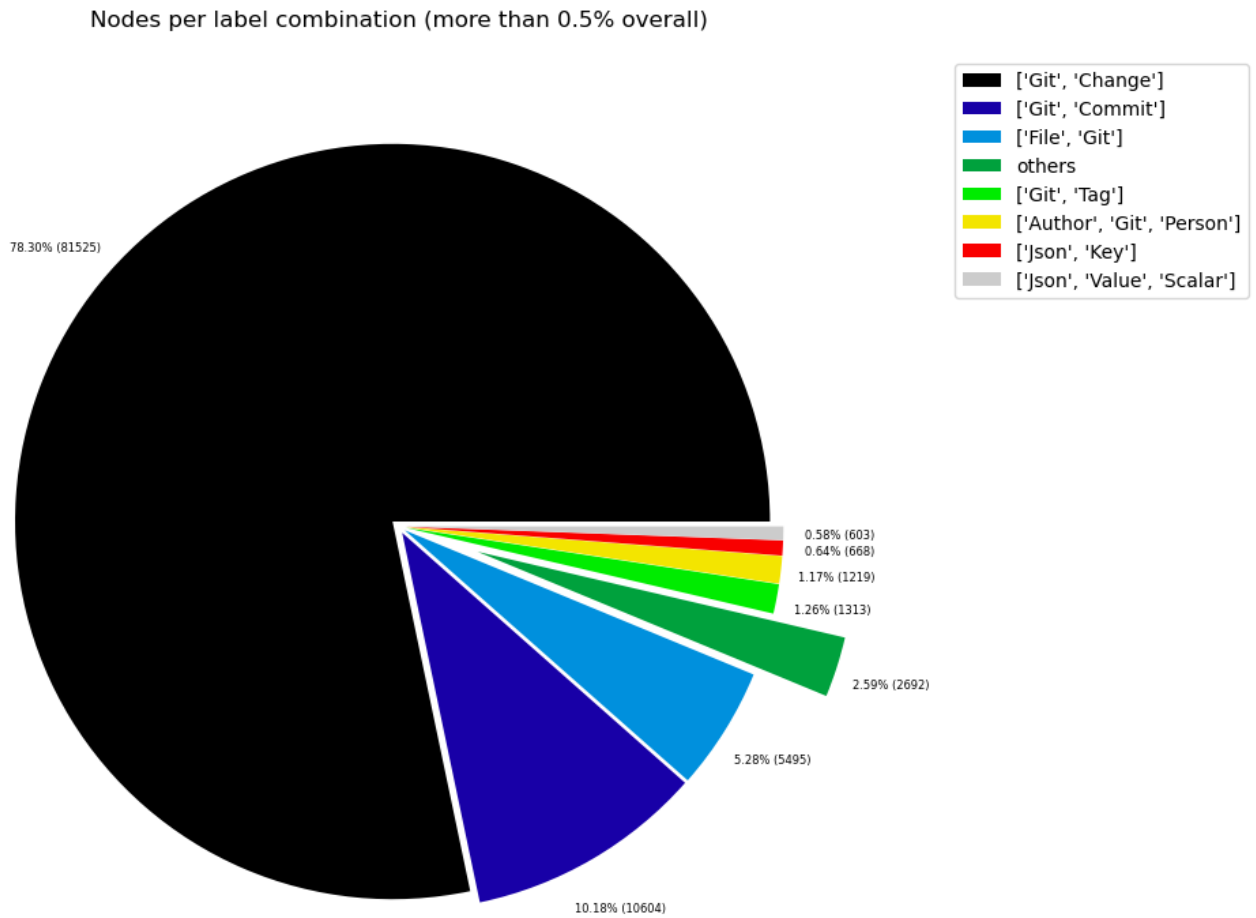2.59% (2692)
5.28% (5495)
10.18% (10604)

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	0.000960
1	[File, TS, Scan]	1	0.000960
2	[TS, Method]	1	0.000960
3	[Repository, File, Git]	1	0.000960
4	[TS, Constructor]	1	0.000960
5	[Value, TS, ObjectMember]	1	0.000960
6	[TS, Class]	1	0.000960
7	[TS, Enum]	2	0.001921
8	[Value, Object, TS]	3	0.002881
9	[Type, TS, Tuple]	3	0.002881
10	[Value, TS, Function]	4	0.003842
11	[TS, TypeParameter]	4	0.003842
12	[Value, TS, Complex]	5	0.004802
13	[NPM, Engine]	6	0.005763
14	[Project, TS]	6	0.005763
15	[File, Local]	6	0.005763
16	[Value, TS, Call]	6	0.005763
17	[Value, TS, Member]	6	0.005763
18	[File, TS, Local, Module]	6	0.005763
19	[Type, TS, TypeParameterReference]	6	0.005763
20	[TS, EnumMember]	8	0.007684
21	[Type, TS, NotIdentified]	11	0.010565
22	[Json, Value, Array]	12	0.011525
23	[Value, TS, Declared]	13	0.012486
24	[TS, TypeAlias]	16	0.015367
25	[File, Directory, Local]	16	0.015367
26	[TS, Interface]	17	0.016327
27	[Type, TS, Intersection]	17	0.016327
28	[jQAssistant, Rule, Concept]	19	0.018248
29	[Value, TS, Literal]	20	0.019209

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:
- ['Committer', 'Git', 'Person']
- ['NPM', 'Dependency']
- ['Type', 'TS', 'Primitive']
- ['Type', 'TS', 'Declared']
- ['TS', 'ExternalDeclaration']
- ['Type', 'TS', 'Literal']
- ['Json', 'Value', 'Object']
- ['Type', 'TS', 'Union']
- ['Type', 'TS', 'ObjectMember']
- ['NPM', 'Script']
- others
- ['TS', 'Property']
- ['TS', 'Function']
- ['Type', 'TS', 'FunctionParameter']
- ['Type', 'Object', 'TS']
- ['Git', 'Branch']
- ['Type', 'TS', 'Function']
- ['File', 'Directory']
- ['TS', 'Parameter']
- ['Package', 'File', 'Json', 'NPM']
- ['TS', 'Variable']
- ['TS', 'ExternalModule']
- ['Value', 'TS', 'Literal']
- ['jQAssistant', 'Rule', 'Concept']
- ['Type', 'TS', 'Intersection']
- ['TS', 'Interface']
- ['File', 'Directory', 'Local']
- ['TS', 'TypeAlias']
- ['Value', 'TS', 'Declared']
- ['Json', 'Value', 'Array']
- ['Type', 'TS', 'NotIdentified']

Data labels on chart:
13.78% (371), 12.56% (338), 10.81% (291), 10.25% (276), 7.95% (214), 5.05% (136), 4.94% (133), 4.42% (119), 3.75% (101), 3.38% (91), 2.90% (78), 2.41% (65), 1.75% (47), 1.49% (40), 1.45% (39), 1.30% (35), 1.26% (34), 1.26% (34), 1.23% (33), 1.08% (29), 0.89% (24), 0.85% (23), 0.74% (20), 0.71% (19), 0.63% (17), 0.63% (17), 0.59% (16), 0.59% (16), 0.48% (13), 0.45% (12), 0.41% (11)
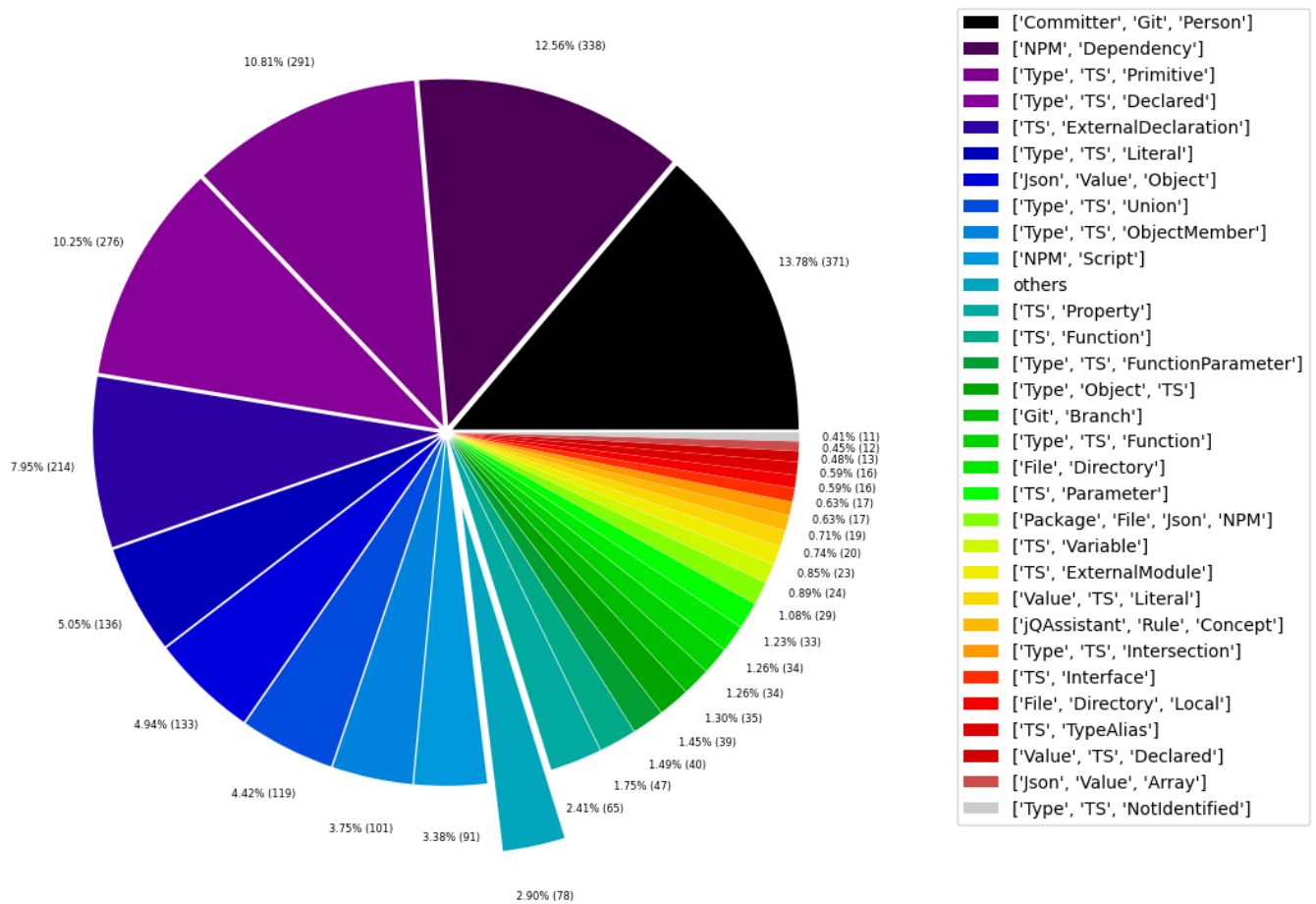
Table 1c - Highest node count by single label

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	100563	96.584677
1	Change	81525	78.299830
2	Commit	10604	10.184500
3	File	5588	5.366936
4	TS	1600	1.536703
5	Person	1590	1.527099
6	Json	1445	1.387835
7	Tag	1313	1.261057
8	Author	1219	1.170776
9	Type	1073	1.030552
10	Value	806	0.774114
11	Key	668	0.641574
12	Scalar	603	0.579145
13	NPM	464	0.445644
14	Committer	371	0.356323
15	Dependency	338	0.324629
16	Primitive	291	0.279488
17	Declared	289	0.277567
18	ExternalDeclaration	214	0.205534
19	Object	175	0.168077
20	Literal	156	0.149829
21	Union	119	0.114292
22	ObjectMember	102	0.097965
23	Script	91	0.087400
24	Function	85	0.081637
25	Property	65	0.062429
26	Directory	50	0.048022
27	FunctionParameter	40	0.038418
28	Branch	35	0.033615
29	Parameter	33	0.031695
30	Package	29	0.027853
31	Local	28	0.026892
32	Variable	24	0.023051
33	ExternalModule	23	0.022090
34	jQAssistant	20	0.019209
35	Concept	19	0.018248
36	Rule	19	0.018248
37	Interface	17	0.016327
38	Intersection	17	0.016327
39	TypeAlias	16	0.015367

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
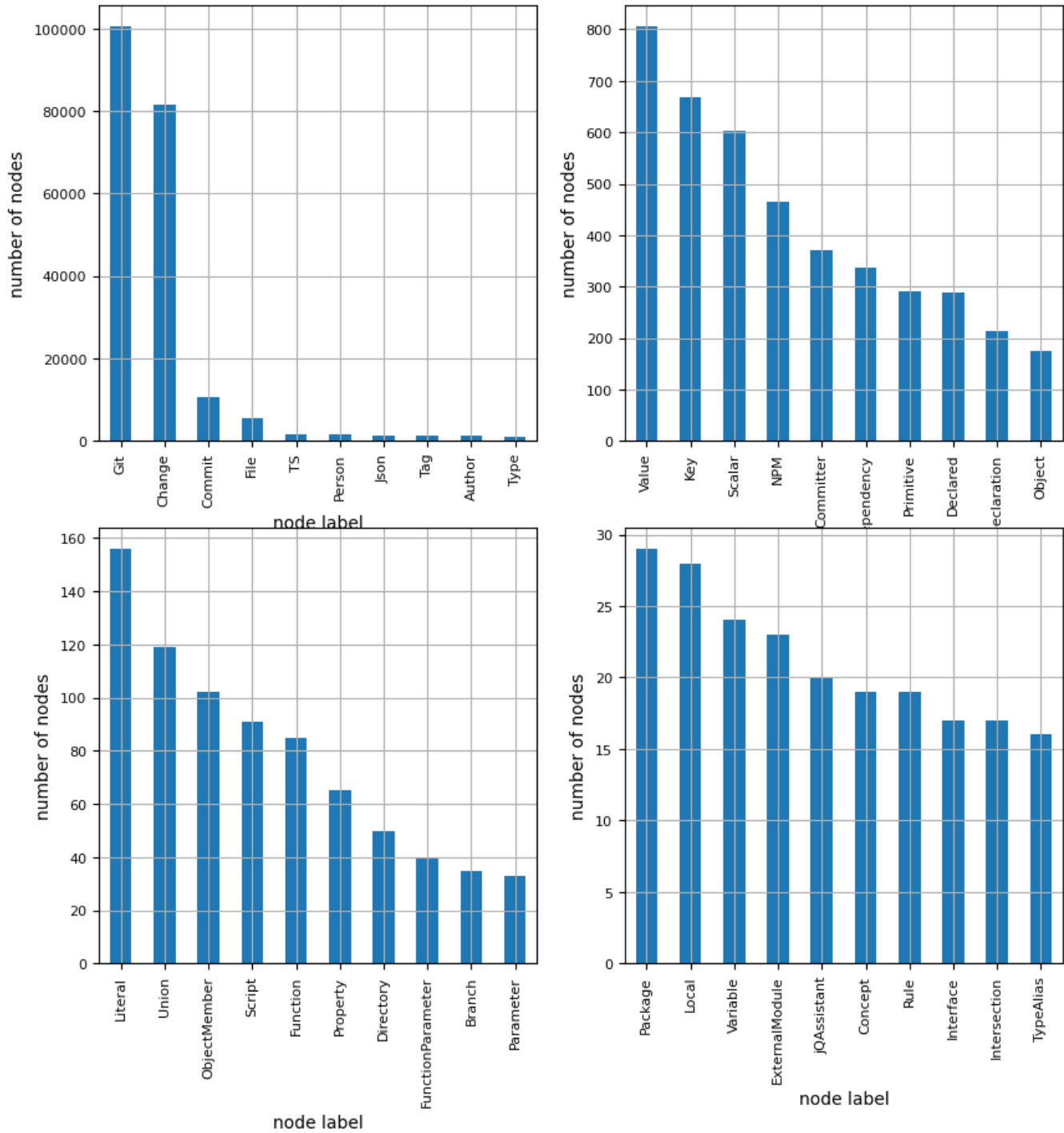
Relationship Types

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: 311021
```

	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	CONTAINS_CHANGE	81525	26.212056
1	MODIFIES	81525	26.212056
2	UPDATES	53337	17.149003
3	COMMITTED	21208	6.818832
4	CREATES	19662	6.321760
5	DELETES	11723	3.769199
6	HAS_PARENT	11650	3.745728
7	HAS_COMMIT	10604	3.409416
8	HAS_FILE	5495	1.766762
9	RENAMES	3197	1.027905
10	HAS_NEW_NAME	1726	0.554946
11	HAS_TAG	1313	0.422158
12	ON_COMMIT	1313	0.422158
13	HAS_AUTHOR	1219	0.391935
14	DEPENDS_ON	961	0.308982
15	HAS_KEY	668	0.214776
16	HAS_VALUE	668	0.214776
17	CONTAINS	594	0.190984
18	HAS_COMMITTER	371	0.119285
19	OF_TYPE	337	0.108353
20	EXPORTS	283	0.090991
21	REFERENCES	197	0.063340
22	DECLARES	186	0.059803
23	DECLARES_DEV_DEPENDENCY	169	0.054337
24	DECLARES_DEPENDENCY	161	0.051765
25	HAS_MEMBER	102	0.032795
26	HAS_TYPE_ARGUMENT	94	0.030223
27	DECLARES_SCRIPT	91	0.029258
28	RETURNS	82	0.026365
29	HAS_PARAMETER	73	0.023471

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.
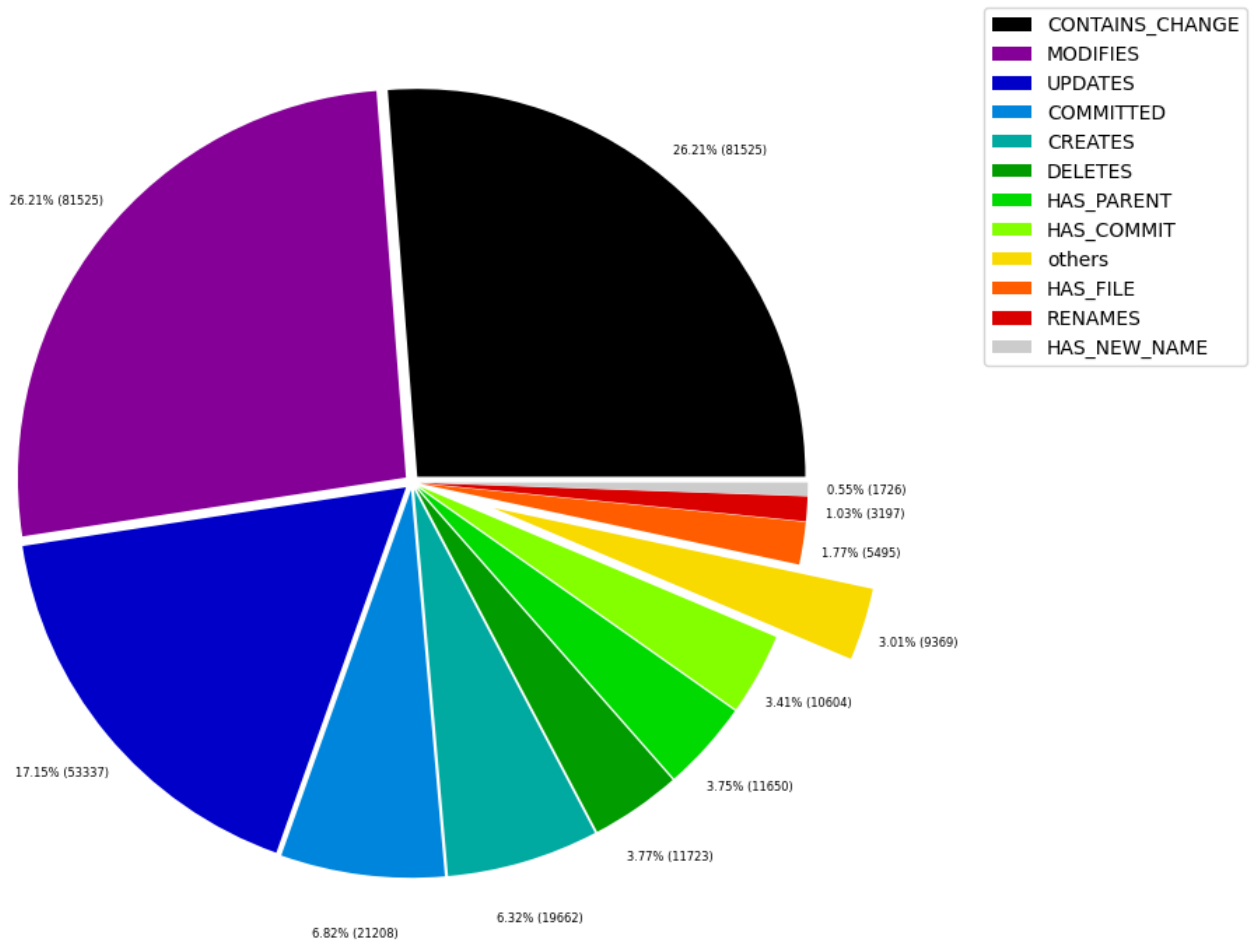
```
<Figure size 640x480 with 0 Axes>
```

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.
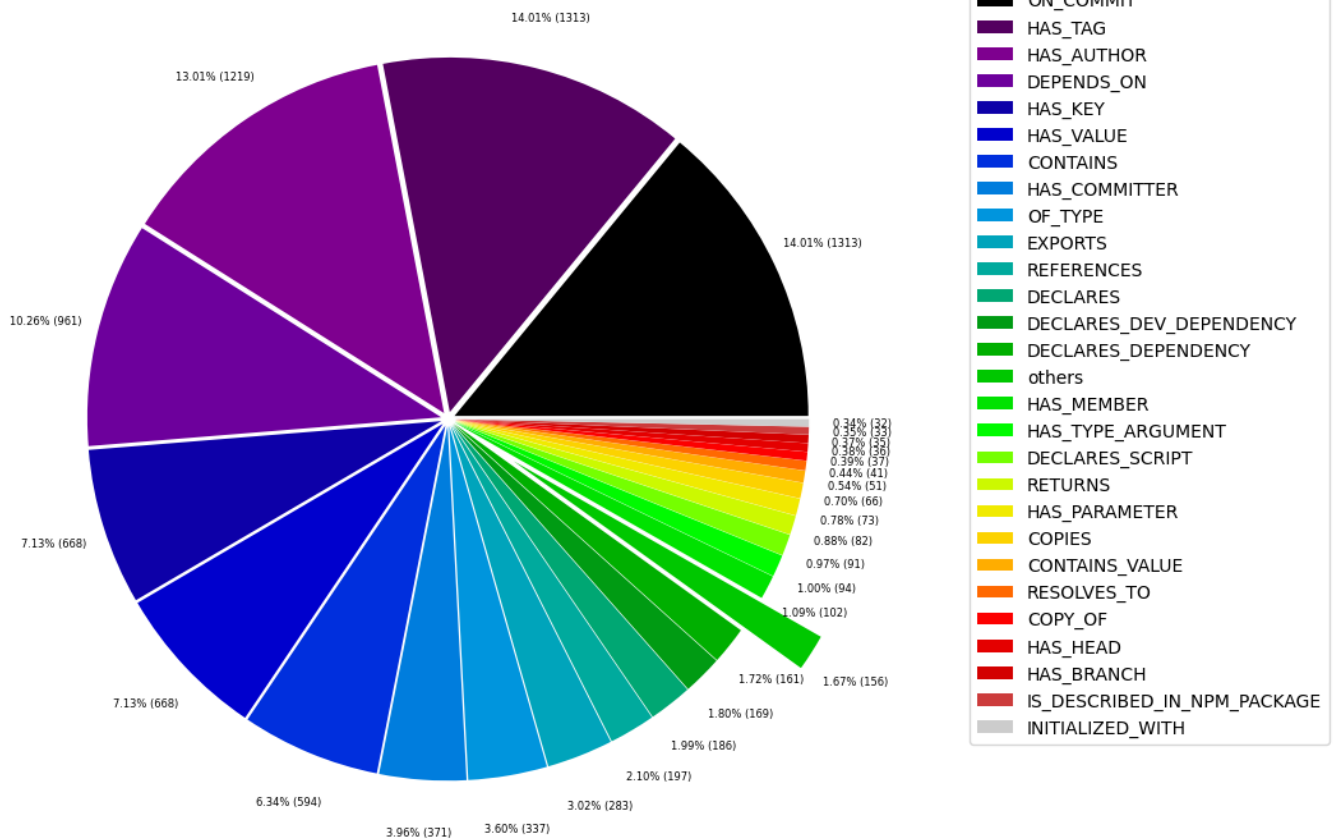
	relationshipType	nodesWithThatRelationshipType	nodesWithThatRelationshipTypePercent
0	PROVIDED_BY_NPM_DEPENDENCY	1	0.000322
1	IS_IMPLEMENTED_IN	2	0.000643
2	CONSTRAINED_BY	4	0.001286
3	REFERENCED_PROJECTS	5	0.001608
4	CONTAINS_PROJECT	6	0.001929
5	DECLARES_ENGINE	6	0.001929
6	EXTENDS	6	0.001929
7	HAS_ARGUMENT	6	0.001929
8	CALLS	6	0.001929
9	HAS_NPM_PACKAGE	6	0.001929
10	HAS_ROOT	6	0.001929
11	MEMBER	6	0.001929
12	PARENT	6	0.001929
13	HAS_CONFIG	6	0.001929
14	SIMILAR	6	0.001929
15	DECLARES_PEER_DEPENDENCY	8	0.002572
16	INCLUDES_CONCEPT	19	0.006109
17	USES	23	0.007395
18	REQUIRES_CONCEPT	28	0.009003
19	INITIALIZED_WITH	32	0.010289
20	IS_DESCRIBED_IN_NPM_PACKAGE	33	0.010610
21	HAS_BRANCH	35	0.011253
22	HAS_HEAD	36	0.011575
23	COPY_OF	37	0.011896
24	RESOLVES_TO	41	0.013182
25	CONTAINS_VALUE	51	0.016398
26	COPIES	66	0.021220
27	HAS_PARAMETER	73	0.023471
28	RETURNS	82	0.026365
29	DECLARES_SCRIPT	91	0.029258

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	numberOfNodes
0	[Git, Commit]	CONTAINS_CHANGE	[Git, Change]	81525	10604	
1	[Git, Change]	MODIFIES	[File, Git]	81525	81525	
2	[Git, Change]	UPDATES	[File, Git]	53337	81525	
3	[Git, Change]	CREATES	[File, Git]	19662	81525	
4	[Git, Change]	DELETES	[File, Git]	11723	81525	
5	[Git, Commit]	HAS_PARENT	[Git, Commit]	11650	10604	
6	[Repository, File, Git]	HAS_COMMIT	[Git, Commit]	10604	1	
7	[Author, Git, Person]	COMMITTED	[Git, Commit]	10604	1219	
8	[Committer, Git, Person]	COMMITTED	[Git, Commit]	10604	371	
9	[Repository, File, Git]	HAS_FILE	[File, Git]	5495	1	
10	[Git, Change]	RENAMES	[File, Git]	3197	81525	
11	[File, Git]	HAS_NEW_NAME	[File, Git]	1726	5495	
12	[Repository, File, Git]	HAS_TAG	[Git, Tag]	1313	1	
13	[Git, Tag]	ON_COMMIT	[Git, Commit]	1313	1313	
14	[Repository, File, Git]	HAS_AUTHOR	[Author, Git, Person]	1219	1	
15	[Json, Value, Object]	HAS_KEY	[Json, Key]	668	133	
16	[Json, Key]	HAS_VALUE	[Json, Value, Scalar]	552	668	
17	[Repository, File, Git]	HAS_COMMITTER	[Committer, Git, Person]	371	1	
18	[TS, Function]	DEPENDS_ON	[TS, ExternalDeclaration]	289	47	
19	[File, TS, Local, Module, Mark4ModuleWeaklyCon...]	DEPENDS_ON	[TS, ExternalDeclaration]	232	4	
20	[TS, ExternalModule]	EXPORTS	[TS, ExternalDeclaration]	214	23	
21	[Package, File, Json, NPM]	DECLARES_DEV_DEPENDENCY	[NPM, Dependency]	169	29	
22	[Package, File, Json, NPM]	DECLARES_DEPENDENCY	[NPM, Dependency]	161	29	
23	[Type, TS, Union]	CONTAINS	[Type, TS, Primitive]	147	119	
24	[Type, TS, Declared]	REFERENCES	[TS, ExternalDeclaration]	142	276	
25	[TS, Function]	DEPENDS_ON	[TS, ExternalModule]	132	47	
26	[Type, TS, Union]	CONTAINS	[Type, TS, Literal]	119	119	
27	[Json, Key]	HAS_VALUE	[Json, Value, Object]	104	668	
28	[Type, Object, TS]	HAS_MEMBER	[Type, TS, ObjectMember]	101	39	
29	[Package, File, Json, NPM]	DECLARES_SCRIPT	[NPM, Script]	91	29	

Graph Density

```
total_number_of_nodes (vertices): 104119
total_number_of_relationships (edges): 311021
-> total directed graph density: 2.869022194355019e-05
-> total directed graph density in percent: 0.0028690221943550188
```