

Graph Serialization Method

1. Variable: Tool name with full types including core concept, layer type, measurement level of input data and output data.

Tool Name with Full Types			
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution113	471
	expert2	solution42	127
	expert3	solution133	200
	expert4	solution137	110
	expert5	solution113	297
	expert6	solution154	150
	expert7	solution137	134
	expert8	solution114	331
	expert9	solution56	114
	expert10	solution89	640
	expert11	solution29	209
	expert12	solution137	106
	expert13	solution104	266
	expert14	solution112	98
Longest Distance	expert1	solution24	616
	expert2	solution89	328
	expert3	solution89	300
	expert4	solution89	372
	expert5	solution90	405
	expert6	solution89	299
	expert7	solution89	301
	expert8	solution90	471
	expert9	solution89	327
	expert10	solution24	818
	expert11	solution89	334
	expert12	solution89	376
	expert13	solution90	344
	expert14	solution89	270
Median Distance	expert1		557.0

	expert2	194.0
	expert3	226.0
	expert4	179.0
	expert5	359.0
	expert6	200.0
	expert7	162.0
	expert8	402.0
	expert9	166.0
	expert10	751.0
	expert11	249.0
	expert12	180.0
	expert13	302.0
	expert14	171.0
Mean Distance	expert1	557.07
	expert2	188.77
	expert3	229.37
	expert4	190.67
	expert5	353.65
	expert6	204.12
	expert7	173.64
	expert8	402.86
	expert9	176.31
	expert10	745.25
	expert11	249.77
	expert12	191.61
	expert13	300.08
	expert14	172.33

2. Variable: Tool name with core concept and layer type of input data and output data.

Tool Name with Coreconcept and Layer Type			
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution117	377
	expert2	solution42	85
	expert3	solution167	141

	expert4	solution24	85
	expert5	solution113	224
	expert6	solution148	103
	expert7	solution134	103
	expert8	solution89	257
	expert9	solution50	80
	expert10	solution113	533
	expert11	solution29	137
	expert12	solution24	84
	expert13	solution85	204
	expert14	solution106	48
Longest Distance	expert1	solution90	522
	expert2	solution89	234
	expert3	solution113	211
	expert4	solution114	279
	expert5	solution90	322
	expert6	solution113	200
	expert7	solution85	211
	expert8	solution90	395
	expert9	solution117	220
	expert10	solution118	672
	expert11	solution89	244
	expert12	solution114	281
	expert13	solution90	273
	expert14	solution31	192
Median Distance	expert1		460.0
	expert2		143.0
	expert3		170.0
	expert4		146.0
	expert5		267.0
	expert6		160.0
	expert7		129.0
	expert8		326.0
	expert9		126.0
	expert10		606.0
	expert11		195.0
	expert12		147.0

	expert13		227.0
	expert14		131.0
Mean Distance	expert1		462.34
	expert2		132.85
	expert3		170.84
	expert4		152.08
	expert5		265.33
	expert6		156.17
	expert7		132.23
	expert8		328.20
	expert9		128.94
	expert10		606.61
	expert11		193.39
	expert12		153.81
	expert13		225.33
	expert14		123.89

3. Variable: Tool name with coreconcept and measurement type of input data and output data.

Tool Name with Coreconcept and Measurement Type			
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution89	293
	expert2	solution42	68
	expert3	solution29	147
	expert4	solution137	65
	expert5	solution117	209
	expert6	solution29	124
	expert7	solution166	83
	expert8	solution29	210
	expert9	solution154	82
	expert10	solution113	351
	expert11	solution164	125
	expert12	solution137	63
	expert13	solution114	202
	expert14	solution112	66

Longest Distance	expert1	solution109	354
	expert2	solution89	220
	expert3	solution85	199
	expert4	solution89	248
	expert5	solution90	304
	expert6	solution89	201
	expert7	solution89	213
	expert8	solution90	282
	expert9	solution89	217
	expert10	solution39	441
	expert11	solution114	209
	expert12	solution89	250
	expert13	solution90	259
	expert14	solution89	187
Median Distance	expert1		320.0
	expert2		118.0
	expert3		167.0
	expert4		109.0
	expert5		273.0
	expert6		154.0
	expert7		107.0
	expert8		247.0
	expert9		107.0
	expert10		407.0
	expert11		152.0
	expert12		110.0
	expert13		227.0
	expert14		119.0
Mean Distance	expert1		320.91
	expert2		120.65
	expert3		167.51
	expert4		117.22
	expert5		268.12
	expert6		153.39
	expert7		116.38
	expert8		246.34
	expert9		115.71

	expert10	403.47
	expert11	152.07
	expert12	115.75
	expert13	228.39
	expert14	118.25

4. Variable: Tool name with layer and measurement type of input data and output data.

Tool Name with Layer and Measurement Types				
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance	
Shortest Distance	expert1	solution113	376	
	expert2	solution42	100	
	expert3	solution145	171	
	expert4	solution109	84	
	expert5	solution113	234	
	expert6	solution148	127	
	expert7	solution118	111	
	expert8	solution89	284	
	expert9	solution50	93	
	expert10	solution113	529	
	expert11	solution29	151	
	expert12	solution109	84	
	expert13	solution104	224	
	expert14	solution106	86	
Longest Distance	expert1	solution24	520	
	expert2	solution89	283	
	expert3	solution86	237	
	expert4	solution89	316	
	expert5	solution24	334	
	expert6	solution89	235	
	expert7	solution89	260	
	expert8	solution24	399	
	expert9	solution89	262	
	expert10	solution24	692	
	expert11	solution89	270	

	expert12	solution89	315
	expert13	solution109	268
	expert14	solution89	217
Median Distance	expert1		459.0
	expert2		155.0
	expert3		190.0
	expert4		153.0
	expert5		292.0
	expert6		177.0
	expert7		141.0
	expert8		335.0
	expert9		139.0
	expert10		626.0
	expert11		205.0
	expert12		154.0
	expert13		244.0
	expert14		146.0
Mean Distance	expert1		460.32
	expert2		155.39
	expert3		191.63
	expert4		158.78
	expert5		288.30
	expert6		177.09
	expert7		149.21
	expert8		340.05
	expert9		145.03
	expert10		625.93
	expert11		208.66
	expert12		160.39
	expert13		243.91
	expert14		145.22

5. Variable: Tool name with only coreconcept type of input data and output data.

Tool Name with Only Coreconcept Type

Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution89	196
	expert2	solution42	55
	expert3	solution168	89
	expert4	solution137	45
	expert5	solution113	130
	expert6	solution142	85
	expert7	solution166	48
	expert8	solution114	135
	expert9	solution167	54
	expert10	solution113	233
	expert11	solution31	86
	expert12	solution137	41
	expert13	solution114	128
	expert14	solution106	29
Longest Distance	expert1	solution90	246
	expert2	solution89	131
	expert3	solution90	126
	expert4	solution85	154
	expert5	solution90	217
	expert6	solution90	133
	expert7	solution114	122
	expert8	solution90	203
	expert9	solution113	127
	expert10	solution118	286
	expert11	solution114	130
	expert12	solution114	159
	expert13	solution90	182
	expert14	solution114	107
Median Distance	expert1		220.0
	expert2		76.0
	expert3		104.0
	expert4		63.0
	expert5		181.0
	expert6		106.0
	expert7		68.0

	expert8	167.0
	expert9	67.0
	expert10	266.0
	expert11	98.0
	expert12	66.0
	expert13	149.0
	expert14	76.0
Mean Distance	expert1	217.53
	expert2	78.43
	expert3	102.21
	expert4	72.26
	expert5	178.22
	expert6	105.60
	expert7	70.21
	expert8	166.45
	expert9	68.98
	expert10	261.66
	expert11	98.93
	expert12	73.78
	expert13	149.47
	expert14	71.56

6. Variable: Tool name with only layer type of input data and output data.

Tool Name with Only Layer Type			
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution117	303
	expert2	solution42	72
	expert3	solution166	110
	expert4	solution90	63
	expert5	solution113	180
	expert6	solution148	89
	expert7	solution166	79
	expert8	solution89	217
	expert9	solution50	59
	expert10	solution113	435

	expert11	solution29	102
	expert12	solution118	64
	expert13	solution162	151
	expert14	solution106	41
Longest Distance	expert1	solution90	420
	expert2	solution117	189
	expert3	solution113	179
	expert4	solution113	223
	expert5	solution90	239
	expert6	solution31	163
	expert7	solution113	174
	expert8	solution24	316
	expert9	solution117	182
	expert10	solution90	549
	expert11	solution90	195
	expert12	solution114	227
	expert13	solution90	194
	expert14	solution29	148
Median Distance	expert1		370.0
	expert2		120.0
	expert3		130.0
	expert4		122.0
	expert5		205.0
	expert6		137.0
	expert7		106.0
	expert8		264.0
	expert9		96.0
	expert10		496.0
	expert11		155.0
	expert12		125.0
	expert13		170.0
	expert14		108.0
Mean Distance	expert1		372.16
	expert2		110.76
	expert3		131.93
	expert4		123.0
	expert5		204.23

	expert6	134.90
	expert7	108.69
	expert8	265.54
	expert9	101.21
	expert10	493.86
	expert11	158.24
	expert12	124.91
	expert13	168.89
	expert14	101.02

7. Variable: Tool name with only measurement type of input data and output data.

Tool Name with Only Measurement Type			
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution31	209
	expert2	solution42	65
	expert3	solution127	114
	expert4	solution109	44
	expert5	solution117	152
	expert6	solution165	104
	expert7	solution166	69
	expert8	solution142	159
	expert9	solution50	65
	expert10	solution117	237
	expert11	solution162	100
	expert12	solution109	43
	expert13	solution15	152
	expert14	solution106	52
Longest Distance	expert1	solution24	251
	expert2	solution89	175
	expert3	solution113	152
	expert4	solution89	194
	expert5	solution53	229
	expert6	solution86	151
	expert7	solution89	175

	expert8	solution28	207
	expert9	solution89	167
	expert10	solution39	316
	expert11	solution89	166
	expert12	solution89	195
	expert13	solution39	188
	expert14	solution114	142
Median Distance	expert1		228.0
	expert2		99.0
	expert3		126.0
	expert4		78.0
	expert5		206.0
	expert6		128.0
	expert7		88.0
	expert8		183.0
	expert9		86.0
	expert10		285.0
	expert11		112.0
	expert12		78.0
	expert13		168.0
	expert14		91.0
Mean Distance	expert1		229.60
	expert2		102.03
	expert3		126.97
	expert4		86.70
	expert5		201.36
	expert6		128.36
	expert7		95.80
	expert8		182.95
	expert9		90.34
	expert10		284.60
	expert11		113.67
	expert12		86.53
	expert13		168.40
	expert14		90.24

8. Variable: Full types of input data and output data without tool name.

Full Types				
Measurements	Expert flow	Work- flow	Generated Workflow	Levenshtein Dis- tance
Shortest Distance	expert1		solution117	343
	expert2		solution28	56
	expert3		solution127	142
	expert4		solution109	89
	expert5		solution127	205
	expert6		solution154	69
	expert7		solution109	86
	expert8		solution114	214
	expert9		solution53	75
	expert10		solution89	503
	expert11		solution29	133
	expert12		solution109	87
	expert13		solution104	164
	expert14		solution110	78
Longest Distance	expert1		solution24	476
	expert2		solution89	242
	expert3		solution89	204
	expert4		solution89	273
	expert5		solution118	269
	expert6		solution89	209
	expert7		solution89	222
	expert8		solution24	351
	expert9		solution89	239
	expert10		solution24	665
	expert11		solution117	237
	expert12		solution89	270
	expert13		solution118	236
	expert14		solution89	204
Median Distance	expert1			424.0
	expert2			137.0
	expert3			166.0
	expert4			132.0
	expert5			231.0
	expert6			116.0

	expert7		110.0
	expert8		296.0
	expert9		110.0
	expert10		603.0
	expert11		184.0
	expert12		134.0
	expert13		199.0
	expert14		113.0
Mean Distance	expert1		419.91
	expert2		127.71
	expert3		164.52
	expert4		141.47
	expert5		231.53
	expert6		121.62
	expert7		120.71
	expert8		291.95
	expert9		120.42
	expert10		601.36
	expert11		183.36
	expert12		140.91
	expert13		197.90
	expert14		116.98

9. Variable: Coreconcept and layer types of input data and output data without tool name.

Coreconcept and Layer Types			
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution117	259
	expert2	solution50	42
	expert3	solution162	71
	expert4	solution109	60
	expert5	solution117	125
	expert6	solution148	33
	expert7	solution106	43
	expert8	solution117	149
	expert9	solution24	37

	expert10	solution31	397
	expert11	solution29	74
	expert12	solution109	60
	expert13	solution113	89
	expert14	solution106	26
Longest Distance	expert1	solution118	376
	expert2	solution117	154
	expert3	solution31	133
	expert4	solution114	177
	expert5	solution118	186
	expert6	solution114	114
	expert7	solution31	131
	expert8	solution118	266
	expert9	solution117	134
	expert10	solution118	516
	expert11	solution109	164
	expert12	solution114	177
	expert13	solution118	162
	expert14	solution31	118
Median Distance	expert1		327.0
	expert2		84.0
	expert3		99.0
	expert4		105.0
	expert5		144.0
	expert6		64.0
	expert7		72.0
	expert8		219.0
	expert9		77.0
	expert10		464.0
	expert11		128.0
	expert12		105.0
	expert13		122.0
	expert14		64.0
Mean Distance	expert1		328.73
	expert2		82.94
	expert3		95.64
	expert4		104.08

	expert5	144.42
	expert6	66.71
	expert7	73.40
	expert8	217.14
	expert9	75.06
	expert10	464.58
	expert11	127.99
	expert12	104.08
	expert13	119.20
	expert14	64.88

10. Variable: Coreconcept and measurement types of input data and output data without tool name.

Coreconcept and Measurement Types				
Measurements	Expert	Work-	Generated	Levenshtein Dis-
	flow	flow	Workflow	tance
Shortest Distance	expert1		solution89	160
	expert2		solution21	29
	expert3		solution29	83
	expert4		solution109	44
	expert5		solution117	96
	expert6		solution57	41
	expert7		solution109	38
	expert8		solution114	109
	expert9		solution29	26
	expert10		solution117	203
	expert11		solution168	57
	expert12		solution7	43
	expert13		solution89	98
	expert14		solution43	44
Longest Distance	expert1		solution118	209
	expert2		solution89	128
	expert3		solution116	121
	expert4		solution89	148
	expert5		solution39	172
	expert6		solution89	111

	expert7	solution89	133
	expert8	solution109	159
	expert9	solution89	129
	expert10	solution39	285
	expert11	solution89	118
	expert12	solution89	144
	expert13	solution115	152
	expert14	solution89	119
Median Distance	expert1		188.0
	expert2		55.0
	expert3		101.0
	expert4		64.0
	expert5		151.0
	expert6		61.0
	expert7		55.0
	expert8		139.0
	expert9		52.0
	expert10		256.0
	expert11		84.0
	expert12		61.0
	expert13		132.0
	expert14		61.0
Mean Distance	expert1		186.77
	expert2		61.02
	expert3		101.67
	expert4		70.98
	expert5		147.88
	expert6		65.52
	expert7		63.12
	expert8		135.40
	expert9		57.87
	expert10		252.00
	expert11		83.40
	expert12		69.19
	expert13		130.02
	expert14		66.01

11. Variable: Layer and measurement types of input data and output data without tool name.

Layer and Measurement Types			
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution89	245
	expert2	solution56	47
	expert3	solution145	91
	expert4	solution109	59
	expert5	solution117	134
	expert6	solution148	48
	expert7	solution109	55
	expert8	solution89	159
	expert9	solution39	53
	expert10	solution89	396
	expert11	solution29	78
	expert12	solution109	59
	expert13	solution113	121
	expert14	solution106	63
Longest Distance	expert1	solution24	379
	expert2	solution89	190
	expert3	solution31	158
	expert4	solution89	216
	expert5	solution24	196
	expert6	solution89	143
	expert7	solution89	174
	expert8	solution24	281
	expert9	solution89	175
	expert10	solution24	539
	expert11	solution109	178
	expert12	solution89	213
	expert13	solution53	165
	expert14	solution89	152
Median Distance	expert1		327.0
	expert2		99.0
	expert3		125.0
	expert4		105.0

	expert5		165.0
	expert6		87.0
	expert7		86.0
	expert8		226.0
	expert9		88.0
	expert10		484.0
	expert11		137.0
	expert12		105.0
	expert13		146.0
	expert14		85.0
Mean Distance	expert1		325.66
	expert2		97.07
	expert3		123.46
	expert4		111.60
	expert5		164.39
	expert6		87.30
	expert7		93.07
	expert8		228.63
	expert9		90.29
	expert10		483.36
	expert11		138.52
	expert12		110.56
	expert13		142.41
	expert14		88.33

12. Variable: Only coreconcept type of input data and output data without tool name.

Only Coreconcept Type			
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution89	245
	expert2	solution56	47
	expert3	solution145	91
	expert4	solution109	59
	expert5	solution117	134
	expert6	solution148	48

	expert7	solution109	55
	expert8	solution89	159
	expert9	solution39	53
	expert10	solution89	396
	expert11	solution29	78
	expert12	solution109	59
	expert13	solution113	121
	expert14	solution106	63
Longest Distance	expert1	solution24	379
	expert2	solution89	190
	expert3	solution31	158
	expert4	solution89	216
	expert5	solution24	196
	expert6	solution89	143
	expert7	solution89	174
	expert8	solution24	281
	expert9	solution89	175
	expert10	solution24	539
	expert11	solution109	178
	expert12	solution89	213
	expert13	solution53	165
	expert14	solution89	152
Median Distance	expert1		327.0
	expert2		99.0
	expert3		125.0
	expert4		105.0
	expert5		165.0
	expert6		87.0
	expert7		86.0
	expert8		226.0
	expert9		88.0
	expert10		484.0
	expert11		137.0
	expert12		105.0
	expert13		146.0
	expert14		85.0
Mean Distance	expert1		325.66

	expert2	97.07
	expert3	123.46
	expert4	111.60
	expert5	164.39
	expert6	87.30
	expert7	93.07
	expert8	228.63
	expert9	90.29
	expert10	483.36
	expert11	138.52
	expert12	110.56
	expert13	142.41
	expert14	88.33

13. Variable: Only layer type of input data and output data without tool name.

Only Layer Type			
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution114	188
	expert2	solution50	25
	expert3	solution162	42
	expert4	solution109	38
	expert5	solution114	62
	expert6	solution148	18
	expert7	solution109	24
	expert8	solution117	112
	expert9	solution24	22
	expert10	solution31	307
	expert11	solution29	35
	expert12	solution109	38
	expert13	solution162	51
	expert14	solution106	19
Longest Distance	expert1	solution24	277
	expert2	solution117	98
	expert3	solution31	96

	expert4	solution113	124
	expert5	solution118	104
	expert6	solution35	79
	expert7	solution35	91
	expert8	solution24	196
	expert9	solution117	89
	expert10	solution109	392
	expert11	solution109	129
	expert12	solution113	124
	expert13	solution35	91
	expert14	solution35	83
Median Distance	expert1		241.0
	expert2		52.0
	expert3		63.0
	expert4		74.0
	expert5		80.0
	expert6		43.0
	expert7		45.0
	expert8		158.0
	expert9		47.0
	expert10		354.0
	expert11		91.0
	expert12		74.0
	expert13		70.0
	expert14		45.0
Mean Distance	expert1		240.17
	expert2		52.59
	expert3		63.60
	expert4		76.23
	expert5		81.41
	expert6		44.22
	expert7		49.24
	expert8		157.48
	expert9		49.76
	expert10		353.06
	expert11		93.05
	expert12		76.23

	expert13	68.09
	expert14	46.21

14. Variable: Only measurement type of input data and output data without tool name.

Only Measurement Type			
Measurements	Expert Work-flow	Generated Workflow	Levenshtein Distance
Shortest Distance	expert1	solution141	80
	expert2	solution42	27
	expert3	solution120	44
	expert4	solution109	16
	expert5	solution117	49
	expert6	solution57	25
	expert7	solution109	21
	expert8	solution117	56
	expert9	solution151	13
	expert10	solution89	88
	expert11	solution168	29
	expert12	solution1	16
	expert13	solution127	53
	expert14	solution57	21
Longest Distance	expert1	solution24	111
	expert2	solution89	96
	expert3	solution39	77
	expert4	solution89	94
	expert5	solution39	104
	expert6	solution89	66
	expert7	solution89	93
	expert8	solution39	88
	expert9	solution89	83
	expert10	solution39	161
	expert11	solution89	69
	expert12	solution89	89
	expert13	solution39	90
	expert14	solution89	79

Median Distance	expert1		95.0
	expert2		40.0
	expert3		58.0
	expert4		35.0
	expert5		84.0
	expert6		40.0
	expert7		34.0
	expert8		75.0
	expert9		34.0
	expert10		138.0
	expert11		45.0
	expert12		33.0
	expert13		71.0
	expert14		36.0
Mean Distance	expert1		95.75
	expert2		44.18
	expert3		59.32
	expert4		39.22
	expert5		81.56
	expert6		39.60
	expert7		40.33
	expert8		73.71
	expert9		35.53
	expert10		134.85
	expert11		44.75
	expert12		37.11
	expert13		69.71
	expert14		37.74

Graph2vec Method

1. Variable: Tool name with full types including core concept, layer type, measurement level of input data and output data.

Tool Name with Full Types

Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution55	0.702201
	expert2	solution38	0.647654
	expert3	solution148	0.551233
	expert4	solution144	0.678818
	expert5	solution39	0.715467
	expert6	solution153	0.621692
	expert7	solution131	0.658949
	expert8	solution38	0.587662
	expert9	solution27	0.655151
	expert10	solution39	0.625597
	expert11	solution112	0.703269
	expert12	solution59	0.713255
	expert13	solution144	0.616552
	expert14	solution34	0.597709
Smallest Similarity	expert1	solution82	-0.224159
	expert2	solution72	-0.267458
	expert3	solution159	-0.168330
	expert4	solution72	-0.189649
	expert5	solution14	-0.189417
	expert6	solution159	-0.165104
	expert7	solution72	-0.224331
	expert8	solution8	-0.269929
	expert9	solution66	-0.239710
	expert10	solution159	-0.125021
	expert11	solution81	-0.144989
	expert12	solution155	-0.195514
	expert13	solution66	-0.201910
	expert14	solution66	-0.231455
Median Similarity	expert1		0.350206
	expert2		0.203074
	expert3		0.289530
	expert4		0.241284
	expert5		0.343645
	expert6		0.239933
	expert7		0.272546

	expert8		0.311812
	expert9		0.217373
	expert10		0.310664
	expert11		0.301810
	expert12		0.254272
	expert13		0.329221
	expert14		0.266847
Mean Similarity	expert1		0.288471
	expert2		0.219490
	expert3		0.255715
	expert4		0.224858
	expert5		0.301580
	expert6		0.218203
	expert7		0.248108
	expert8		0.263249
	expert9		0.218044
	expert10		0.283327
	expert11		0.278202
	expert12		0.258588
	expert13		0.280003
	expert14		0.237808

2. Variable: Tool name with coreconcept and layer type of input data and output data.

Tool Name with Coreconcept and Layer Type			
Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution142	0.630135
	expert2	solution52	0.729467
	expert3	solution89	0.787529
	expert4	solution137	0.597835
	expert5	solution122	0.819309
	expert6	solution145	0.673605
	expert7	solution7	0.714759
	expert8	solution142	0.919864

	expert9	solution50	0.876922
	expert10	solution52	0.697767
	expert11	solution52	0.727939
	expert12	solution15	0.530121
	expert13	solution7	0.617803
	expert14	solution106	0.709707
Smallest Similarity	expert1	solution116	-0.024103
	expert2	solution100	-0.065246
	expert3	solution59	-0.138250
	expert4	solution101	-0.150094
	expert5	solution57	-0.064527
	expert6	solution6	-0.147261
	expert7	solution17	-0.063534
	expert8	solution24	-0.144984
	expert9	solution166	-0.091799
	expert10	solution100	-0.009502
	expert11	solution100	-0.040455
	expert12	solution106	-0.034335
	expert13	solution28	0.007556
	expert14	solution168	-0.037891
Median Similarity	expert1		0.320492
	expert2		0.367857
	expert3		0.418267
	expert4		0.307804
	expert5		0.382470
	expert6		0.314752
	expert7		0.261454
	expert8		0.321319
	expert9		0.423424
	expert10		0.382828
	expert11		0.380461
	expert12		0.286261
	expert13		0.327302
	expert14		0.332486
Mean Similarity	expert1		0.310635
	expert2		0.368106
	expert3		0.376779

	expert4	0.281144
	expert5	0.405516
	expert6	0.293735
	expert7	0.267765
	expert8	0.319250
	expert9	0.410881
	expert10	0.384832
	expert11	0.381748
	expert12	0.272186
	expert13	0.316251
	expert14	0.344727

3. Variable: Tool name with coreconcept and measurement type of input data and output data.

Tool Name with Coreconcept and Measurement Type			
Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution54	0.718636
	expert2	solution35	0.826677
	expert3	solution139	0.666877
	expert4	solution26	0.647709
	expert5	solution63	0.637097
	expert6	solution122	0.594375
	expert7	solution26	0.746767
	expert8	solution137	0.686386
	expert9	solution140	0.562329
	expert10	solution63	0.705894
	expert11	solution63	0.792947
	expert12	solution63	0.776705
	expert13	solution53	0.580849
	expert14	solution53	0.587184
Smallest Similarity	expert1	solution101	0.068571
	expert2	solution92	0.075973
	expert3	solution4	0.139836
	expert4	solution147	-0.052925

	expert5	solution95	0.009377
	expert6	solution147	-0.014290
	expert7	solution142	-0.095514
	expert8	solution142	0.021251
	expert9	solution146	-0.070326
	expert10	solution142	0.014537
	expert11	solution109	0.061422
	expert12	solution92	-0.098568
	expert13	solution6	0.078354
	expert14	solution6	0.071164
Median Similarity	expert1		0.376071
	expert2		0.339421
	expert3		0.370946
	expert4		0.278906
	expert5		0.353257
	expert6		0.293029
	expert7		0.273242
	expert8		0.327721
	expert9		0.275062
	expert10		0.388823
	expert11		0.380683
	expert12		0.308180
	expert13		0.332068
	expert14		0.347206
Mean Similarity	expert1		0.360885
	expert2		0.376588
	expert3		0.376042
	expert4		0.282548
	expert5		0.333234
	expert6		0.309987
	expert7		0.294916
	expert8		0.351848
	expert9		0.280778
	expert10		0.374683
	expert11		0.372881
	expert12		0.299196
	expert13		0.342555

	expert14		0.353301
--	----------	--	----------

4. Variable: Tool name with layer and measurement type of input data and output data.

Tool Name with Layer and Measurement Type			
Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution5	0.633041
	expert2	solution24	0.724598
	expert3	solution63	0.574906
	expert4	solution154	0.629295
	expert5	solution63	0.648764
	expert6	solution154	0.729779
	expert7	solution27	0.525829
	expert8	solution152	0.654824
	expert9	solution154	0.675442
	expert10	solution154	0.797842
	expert11	solution154	0.756925
	expert12	solution154	0.723650
	expert13	solution1	0.644714
	expert14	solution24	0.705925
Smallest Similarity	expert1	solution21	0.039570
	expert2	solution46	-0.119925
	expert3	solution62	0.073993
	expert4	solution167	-0.118420
	expert5	solution120	-0.019253
	expert6	solution55	-0.145015
	expert7	solution90	0.051810
	expert8	solution167	-0.092715
	expert9	solution167	-0.126126
	expert10	solution46	0.019291
	expert11	solution46	-0.046015
	expert12	solution101	-0.147281
	expert13	solution120	0.035635
	expert14	solution113	0.084681

Median Similarity	expert1		0.391548
	expert2		0.295525
	expert3		0.339571
	expert4		0.223930
	expert5		0.348837
	expert6		0.277477
	expert7		0.325952
	expert8		0.325633
	expert9		0.293577
	expert10		0.346866
	expert11		0.290064
	expert12		0.183100
	expert13		0.385612
	expert14		0.347732
Mean Similarity	expert1		0.393571
	expert2		0.287489
	expert3		0.344284
	expert4		0.242243
	expert5		0.340306
	expert6		0.268428
	expert7		0.323550
	expert8		0.331241
	expert9		0.283905
	expert10		0.339083
	expert11		0.297527
	expert12		0.214824
	expert13		0.365631
	expert14		0.356986

5. Variable: Tool name with only coreconcept type of input data and output data.

Tool Name with Only Coreconcept Type			
Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution41	0.666878

	expert2	solution40	0.793503
	expert3	solution131	0.689610
	expert4	solution41	0.601446
	expert5	solution123	0.849760
	expert6	solution166	0.491881
	expert7	solution167	0.844320
	expert8	solution144	0.771579
	expert9	solution56	0.933171
	expert10	solution137	0.738544
	expert11	solution41	0.834198
	expert12	solution166	0.494536
	expert13	solution136	0.659132
	expert14	solution101	0.738668
Smallest Similarity	expert1	solution22	-0.099393
	expert2	solution149	-0.152390
	expert3	solution59	-0.197197
	expert4	solution28	-0.265970
	expert5	solution87	0.018190
	expert6	solution122	-0.043218
	expert7	solution115	0.012800
	expert8	solution149	-0.171258
	expert9	solution165	-0.002691
	expert10	solution101	-0.084791
	expert11	solution101	-0.144166
	expert12	solution28	-0.147701
	expert13	solution22	-0.053668
	expert14	solution62	-0.062013
Median Similarity	expert1		0.311606
	expert2		0.356590
	expert3		0.401834
	expert4		0.200611
	expert5		0.403699
	expert6		0.245310
	expert7		0.255702
	expert8		0.303005
	expert9		0.402431
	expert10		0.360150

	expert11		0.384383
	expert12		0.168129
	expert13		0.350045
	expert14		0.310746
Mean Similarity	expert1		0.294642
	expert2		0.343802
	expert3		0.354705
	expert4		0.187807
	expert5		0.416298
	expert6		0.249210
	expert7		0.294653
	expert8		0.294229
	expert9		0.398608
	expert10		0.371656
	expert11		0.369754
	expert12		0.174189
	expert13		0.330036
	expert14		0.351115

6. Variable: Tool name with only layer type of input data and output data.

Tool Name with Only Layer Type			
Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution22	0.626089
	expert2	solution142	0.547154
	expert3	solution103	0.717462
	expert4	solution6	0.586448
	expert5	solution122	0.820681
	expert6	solution53	0.568542
	expert7	solution60	0.494499
	expert8	solution142	0.903505
	expert9	solution53	0.904458
	expert10	solution135	0.669705
	expert11	solution51	0.646682
	expert12	solution22	0.655117

	expert13	solution138	0.647182
	expert14	solution112	0.900962
Smallest Similarity	expert1	solution101	-0.147645
	expert2	solution101	-0.055522
	expert3	solution143	-0.179905
	expert4	solution33	-0.150201
	expert5	solution167	0.036058
	expert6	solution122	-0.014391
	expert7	solution101	-0.009776
	expert8	solution101	-0.110602
	expert9	solution58	-0.027904
	expert10	solution28	0.104535
	expert11	solution28	0.023928
	expert12	solution101	-0.010945
	expert13	solution163	-0.132317
	expert14	solution28	-0.032681
Median Similarity	expert1		0.340753
	expert2		0.302225
	expert3		0.413709
	expert4		0.207231
	expert5		0.331604
	expert6		0.324417
	expert7		0.281993
	expert8		0.294773
	expert9		0.440358
	expert10		0.382565
	expert11		0.360290
	expert12		0.246091
	expert13		0.313070
	expert14		0.413133
Mean Similarity	expert1		0.340971
	expert2		0.295805
	expert3		0.368525
	expert4		0.225743
	expert5		0.363823
	expert6		0.319613
	expert7		0.283568

	expert8	0.292044
	expert9	0.425782
	expert10	0.385944
	expert11	0.344254
	expert12	0.267013
	expert13	0.317627
	expert14	0.439584

7. Variable: Tool name with only measurement type of input data and output data.

Tool Name with Only Measurement Type			
Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution62	0.722744
	expert2	solution27	0.668794
	expert3	solution139	0.643330
	expert4	solution154	0.603017
	expert5	solution138	0.623682
	expert6	solution27	0.577986
	expert7	solution58	0.556299
	expert8	solution152	0.708336
	expert9	solution26	0.547051
	expert10	solution154	0.826386
	expert11	solution154	0.807324
	expert12	solution154	0.651972
	expert13	solution154	0.720682
	expert14	solution28	0.692411
Smallest Similarity	expert1	solution120	0.105672
	expert2	solution106	0.017930
	expert3	solution3	0.177415
	expert4	solution107	0.104920
	expert5	solution166	0.194137
	expert6	solution120	0.027251
	expert7	solution3	0.023459
	expert8	solution92	0.089207

	expert9	solution108	0.061653
	expert10	solution167	-0.025253
	expert11	solution167	0.071835
	expert12	solution120	0.090386
	expert13	solution167	0.010703
	expert14	solution57	0.080625
Median Similarity	expert1		0.457332
	expert2		0.327260
	expert3		0.389485
	expert4		0.355430
	expert5		0.411255
	expert6		0.326331
	expert7		0.295167
	expert8		0.389645
	expert9		0.332461
	expert10		0.349977
	expert11		0.412990
	expert12		0.356794
	expert13		0.394811
	expert14		0.383141
Mean Similarity	expert1		0.447054
	expert2		0.326926
	expert3		0.396839
	expert4		0.346694
	expert5		0.410508
	expert6		0.320167
	expert7		0.293449
	expert8		0.380815
	expert9		0.314605
	expert10		0.340801
	expert11		0.407919
	expert12		0.351031
	expert13		0.386932
	expert14		0.382097

8. Variable: Full types of input data and output data without tool name.

Full Types				
Measurements	Expert flow	Work-Generated Workflow	Similarity Value	
Largest Similarity	expert1	solution20	0.685854	
	expert2	solution32	0.665157	
	expert3	solution61	0.682807	
	expert4	solution58	0.669155	
	expert5	solution154	0.660022	
	expert6	solution140	0.632223	
	expert7	solution116	0.546165	
	expert8	solution20	0.691765	
	expert9	solution16	0.621509	
	expert10	solution122	0.621068	
	expert11	solution40	0.598849	
	expert12	solution20	0.644071	
	expert13	solution99	0.650329	
	expert14	solution20	0.643061	
Smallest Similarity	expert1	solution123	0.040798	
	expert2	solution129	0.045122	
	expert3	solution3	0.052986	
	expert4	solution107	-0.003844	
	expert5	solution3	0.071726	
	expert6	solution3	-0.025481	
	expert7	solution166	-0.021766	
	expert8	solution101	0.104213	
	expert9	solution38	0.029930	
	expert10	solution166	0.030496	
	expert11	solution129	0.107629	
	expert12	solution107	0.084202	
	expert13	solution60	0.037157	
	expert14	solution37	-0.048117	
Median Similarity	expert1		0.381479	
	expert2		0.357076	
	expert3		0.407769	
	expert4		0.305559	
	expert5		0.348456	
	expert6		0.268243	

	expert7		0.321449
	expert8		0.351065
	expert9		0.264713
	expert10		0.380928
	expert11		0.384845
	expert12		0.297575
	expert13		0.363433
	expert14		0.334929
Mean Similarity	expert1		0.370699
	expert2		0.374180
	expert3		0.384429
	expert4		0.322088
	expert5		0.350452
	expert6		0.271946
	expert7		0.305382
	expert8		0.356783
	expert9		0.283791
	expert10		0.368537
	expert11		0.378560
	expert12		0.315458
	expert13		0.356937
	expert14		0.325797

9. Variable: Coreconcept and layer types of input data and output data without tool name.

Coreconcept and Layer Types			
Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution148	0.699657
	expert2	solution49	0.713416
	expert3	solution86	0.770164
	expert4	solution38	0.733081
	expert5	solution151	0.870862
	expert6	solution151	0.924490
	expert7	solution35	0.750983

	expert8	solution142	0.891940
	expert9	solution26	0.927882
	expert10	solution40	0.594660
	expert11	solution49	0.699934
	expert12	solution63	0.605190
	expert13	solution148	0.855109
	expert14	solution97	0.750705
Smallest Similarity	expert1	solution100	-0.125994
	expert2	solution141	-0.218258
	expert3	solution60	-0.370817
	expert4	solution145	-0.094327
	expert5	solution16	-0.006367
	expert6	solution62	0.001863
	expert7	solution148	-0.178299
	expert8	solution5	-0.127045
	expert9	solution4	-0.153623
	expert10	solution141	-0.042805
	expert11	solution141	-0.103044
	expert12	solution141	-0.115209
	expert13	solution60	-0.160674
	expert14	solution61	-0.158069
Median Similarity	expert1		0.248544
	expert2		0.341720
	expert3		0.236582
	expert4		0.436434
	expert5		0.491614
	expert6		0.434610
	expert7		0.344048
	expert8		0.266000
	expert9		0.404252
	expert10		0.320086
	expert11		0.387795
	expert12		0.267187
	expert13		0.388714
	expert14		0.369952
Mean Similarity	expert1		0.259948
	expert2		0.334282

	expert3	0.221468
	expert4	0.396234
	expert5	0.476927
	expert6	0.450582
	expert7	0.331996
	expert8	0.255314
	expert9	0.429844
	expert10	0.323791
	expert11	0.381200
	expert12	0.270616
	expert13	0.348145
	expert14	0.343410

10. Variable: Coreconcept and measurement types of input data and output data without tool name.

Coreconcept and Measurement Types			
Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution140	0.703248
	expert2	solution34	0.732865
	expert3	solution139	0.644339
	expert4	solution47	0.678722
	expert5	solution88	0.557166
	expert6	solution58	0.601908
	expert7	solution140	0.572893
	expert8	solution140	0.667768
	expert9	solution139	0.644444
	expert10	solution134	0.595683
	expert11	solution140	0.616336
	expert12	solution55	0.792956
	expert13	solution137	0.649990
	expert14	solution139	0.632397
Smallest Similarity	expert1	solution106	-0.011858
	expert2	solution93	-0.034642
	expert3	solution106	0.105551

	expert4	solution90	-0.134225
	expert5	solution165	0.050883
	expert6	solution133	0.099795
	expert7	solution165	0.132124
	expert8	solution106	0.015675
	expert9	solution133	0.056549
	expert10	solution167	0.062901
	expert11	solution166	0.021102
	expert12	solution90	-0.101664
	expert13	solution51	0.044125
	expert14	solution106	0.010608
Median Similarity	expert1		0.327962
	expert2		0.307489
	expert3		0.351332
	expert4		0.323888
	expert5		0.319687
	expert6		0.379292
	expert7		0.369486
	expert8		0.293826
	expert9		0.370667
	expert10		0.379870
	expert11		0.324308
	expert12		0.388037
	expert13		0.325553
	expert14		0.302611
Mean Similarity	expert1		0.335862
	expert2		0.349602
	expert3		0.343438
	expert4		0.324449
	expert5		0.318395
	expert6		0.373420
	expert7		0.364763
	expert8		0.299485
	expert9		0.343782
	expert10		0.366672
	expert11		0.326764
	expert12		0.388093

	expert13	0.328619
	expert14	0.311572

11. Variable: Layer and measurement types of input data and output data without tool name.

Layer and Measurement Types			
Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution154	0.945767
	expert2	solution40	0.629040
	expert3	solution27	0.736896
	expert4	solution154	0.662358
	expert5	solution27	0.618076
	expert6	solution58	0.572643
	expert7	solution1	0.708289
	expert8	solution154	0.830874
	expert9	solution154	0.536125
	expert10	solution54	0.683852
	expert11	solution54	0.644630
	expert12	solution136	0.716111
	expert13	solution24	0.591592
	expert14	solution54	0.581672
Smallest Similarity	expert1	solution106	0.042087
	expert2	solution113	0.018473
	expert3	solution126	-0.082970
	expert4	solution124	-0.065891
	expert5	solution166	0.049946
	expert6	solution166	0.042678
	expert7	solution90	0.000555
	expert8	solution120	0.039893
	expert9	solution101	0.073189
	expert10	solution109	0.125643
	expert11	solution113	0.103878
	expert12	solution141	-0.121105
	expert13	solution126	0.066336

	expert14	solution113	0.021606
Median Similarity	expert1		0.348557
	expert2		0.362934
	expert3		0.336744
	expert4		0.225096
	expert5		0.322521
	expert6		0.307487
	expert7		0.305464
	expert8		0.388146
	expert9		0.315939
	expert10		0.395825
	expert11		0.388919
	expert12		0.296247
	expert13		0.338377
	expert14		0.328764
Mean Similarity	expert1		0.346139
	expert2		0.341674
	expert3		0.355929
	expert4		0.234763
	expert5		0.333888
	expert6		0.308785
	expert7		0.321798
	expert8		0.374826
	expert9		0.305425
	expert10		0.389963
	expert11		0.378982
	expert12		0.310178
	expert13		0.345383
	expert14		0.322943

12. Variable: Only coreconcept type of input data and output data without tool name.

Only Coreconcept Type				
Measurements	Expert	Work-	Generated	Similarity Value
	flow		Workflow	

Largest Similarity	expert1	solution58	0.610505
	expert2	solution57	0.655166
	expert3	solution100	0.733549
	expert4	solution5	0.554003
	expert5	solution151	0.821256
	expert6	solution141	0.841156
	expert7	solution151	0.800473
	expert8	solution152	0.781265
	expert9	solution151	0.811028
	expert10	solution115	0.887082
	expert11	solution115	0.862003
	expert12	solution5	0.567887
	expert13	solution2	0.707015
	expert14	solution100	0.825319
Smallest Similarity	expert1	solution146	-0.385532
	expert2	solution100	-0.527276
	expert3	solution59	-0.418211
	expert4	solution100	-0.263712
	expert5	solution87	-0.232639
	expert6	solution89	-0.072768
	expert7	solution100	-0.074046
	expert8	solution3	-0.265700
	expert9	solution87	-0.092167
	expert10	solution100	-0.154280
	expert11	solution100	-0.099495
	expert12	solution40	-0.318376
	expert13	solution148	-0.114321
	expert14	solution115	-0.430438
Median Similarity	expert1		0.204645
	expert2		-0.019159
	expert3		0.418569
	expert4		0.047064
	expert5		0.363366
	expert6		0.412669
	expert7		0.329850
	expert8		0.304066
	expert9		0.445015

	expert10		0.289268
	expert11		0.422921
	expert12		0.015897
	expert13		0.323952
	expert14		0.457315
Mean Similarity	expert1		0.176384
	expert2		0.019276
	expert3		0.326827
	expert4		0.097225
	expert5		0.340816
	expert6		0.399828
	expert7		0.314526
	expert8		0.282697
	expert9		0.429184
	expert10		0.294740
	expert11		0.406060
	expert12		0.044471
	expert13		0.310676
	expert14		0.370031

13. Variable: Only layer type of input data and output data without tool name.

Only Layer Type				
Measurements	Expert flow	Work- flow	Generated Workflow	Similarity Value
Largest Similarity	expert1		solution145	0.742726
	expert2		solution44	0.715800
	expert3		solution89	0.757187
	expert4		solution16	0.664472
	expert5		solution168	0.834000
	expert6		solution151	0.914276
	expert7		solution145	0.605257
	expert8		solution145	0.814816
	expert9		solution27	0.664751
	expert10		solution153	0.934472

	expert11	solution153	0.817007
	expert12	solution95	0.539640
	expert13	solution139	0.678048
	expert14	solution120	0.624414
Smallest Similarity	expert1	solution101	-0.078892
	expert2	solution117	-0.159497
	expert3	solution43	-0.136186
	expert4	solution31	-0.189306
	expert5	solution59	-0.085852
	expert6	solution59	-0.106000
	expert7	solution101	-0.213645
	expert8	solution101	-0.241230
	expert9	solution141	-0.002780
	expert10	solution59	0.047935
	expert11	solution141	-0.091529
	expert12	solution31	-0.194355
	expert13	solution87	-0.201612
	expert14	solution141	-0.159142
Median Similarity	expert1		0.272498
	expert2		0.439269
	expert3		0.238634
	expert4		0.375316
	expert5		0.435982
	expert6		0.484113
	expert7		0.319574
	expert8		0.217872
	expert9		0.374260
	expert10		0.487257
	expert11		0.474991
	expert12		0.224278
	expert13		0.408656
	expert14		0.355051
Mean Similarity	expert1		0.273866
	expert2		0.394518
	expert3		0.240428
	expert4		0.327806
	expert5		0.409176

	expert6	0.453519
	expert7	0.293173
	expert8	0.218351
	expert9	0.363054
	expert10	0.478980
	expert11	0.446074
	expert12	0.216534
	expert13	0.373110
	expert14	0.330221

14. Variable: Only measurement type of input data and output data without tool name.

Only Measurement Type			
Measurements	Expert Work-flow	Generated Workflow	Similarity Value
Largest Similarity	expert1	solution154	0.874659
	expert2	solution154	0.833838
	expert3	solution61	0.709234
	expert4	solution139	0.676246
	expert5	solution22	0.660192
	expert6	solution22	0.714969
	expert7	solution63	0.720205
	expert8	solution154	0.726319
	expert9	solution61	0.777027
	expert10	solution139	0.677122
	expert11	solution154	0.625293
	expert12	solution139	0.666416
	expert13	solution139	0.598307
	expert14	solution22	0.660428
Smallest Similarity	expert1	solution27	-0.049111
	expert2	solution27	-0.024789
	expert3	solution106	0.049943
	expert4	solution7	0.093353
	expert5	solution4	0.067599
	expert6	solution4	0.128641

	expert7	solution95	0.010980
	expert8	solution106	-0.056003
	expert9	solution122	-0.031667
	expert10	solution93	0.080061
	expert11	solution93	0.032419
	expert12	solution96	-0.008460
	expert13	solution121	0.024136
	expert14	solution121	0.038007
Median Similarity	expert1		0.389185
	expert2		0.368097
	expert3		0.398989
	expert4		0.329378
	expert5		0.361438
	expert6		0.431351
	expert7		0.367077
	expert8		0.364714
	expert9		0.378630
	expert10		0.353568
	expert11		0.408238
	expert12		0.362075
	expert13		0.315985
	expert14		0.372805
Mean Similarity	expert1		0.384570
	expert2		0.362699
	expert3		0.385926
	expert4		0.337316
	expert5		0.361915
	expert6		0.427275
	expert7		0.373165
	expert8		0.353291
	expert9		0.363012
	expert10		0.362700
	expert11		0.395071
	expert12		0.337199
	expert13		0.315018
	expert14		0.354551

Graph2vec Method- Average Data

Table 1: Variable: Tool name with full types including core concept, layer type, measurement level of input data and output data

Tool Name with Full Types	
Measurements	Similarity Value
Average Largest Similarity	0.65
Average Smallest Similarity	-0.20
Average Median Similarity	0.28
Average Mean Similarity	0.256

Table 2: Variable: Tool name with coreconcept and layer type of input data and output data

Tool Name with Coreconcept and Layer Type	
Measurements	Similarity Value
Average Largest Similarity	0.72
Average Smallest Similarity	-0.07
Average Median Similarity	0.34
Average Mean Similarity	0.34

Table 3: Variable: Tool name with coreconcept and measurement type of input data and output data

Tool Name with Coreconcept and Measurement Type	
Measurements	Similarity Value
Average Largest Similarity	0.68

Average Smallest Similarity	0.01
Average Median Similarity	0.33
Average Mean Similarity	0.34

Table 4: Variable: Tool name with layer and measurement type of input data and output data

Tool Name with Layer and Measurement Type	
Measurements	Similarity Value
Average Largest Similarity	0.67
Average Smallest Similarity	-0.04
Average Median Similarity	0.31
Average Mean Similarity	0.31

Table 5: Variable: Tool name with only coreconcept type of input data and out- put data

Tool Name with Only Coreconcept Type	
Measurements	Similarity Value
Average Largest Similarity	0.72
Average Smallest Similarity	-0.10
Average Median Similarity	0.32
Average Mean Similarity	0.32

Table 6: Variable: Tool name with only layer type of input data and output data

Tool Name with Only Layer Type	
Measurements	Similarity Value
Average Largest Similarity	0.69
Average Smallest Similarity	-0.05
Average Median Similarity	0.33
Average Mean Similarity	0.33

Table 7: Variable: Tool name with only measurement type of input data and output data

Tool Name with Only Measurement Type	
Measurements	Similarity Value
Average Largest Similarity	0.67
Average Smallest Similarity	0.07
Average Median Similarity	0.37
Average Mean Similarity	0.36

Table 8: Variable: Full types of input data and output data without tool name

Full Types	
Measurements	Similarity Value
Average Largest Similarity	0.64
Average Smallest Similarity	0.04
Average Median Similarity	0.34

Average Similarity	Mean	0.34
-----------------------	------	------

Table 9: Variable: Coreconcept and layer types of input data and output data without tool name

Coreconcept and Layer Types		
Measurements		Similarity Value
Average Similarity	Largest	0.77
Average Similarity	Smallest	-0.13
Average Similarity	Median	0.35
Average Similarity	Mean	0.34

Table 10: Variable: Coreconcept and measurement types of input data and out- put data without tool name

Coreconcept and Measurement Types		
Measurements		Similarity Value
Average Similarity	Largest	0.65
Average Similarity	Smallest	0.02
Average Similarity	Median	0.34
Average Similarity	Mean	0.34

Table 11: Variable: Layer and measurement types of input data and output data without tool name

Layer and Measurement Types		
Measurements		Similarity Value

Average Similarity	Largest	0.68
Average Similarity	Smallest	0.02
Average Similarity	Median	0.33
Average Similarity	Mean	0.33

Table 12: Variable: Only coreconcept type of input data and output data without tool name

Only Coreconcept Type		
Measurements		Similarity Value
Average Similarity	Largest	0.75
Average Similarity	Smallest	-0.25
Average Similarity	Median	0.29
Average Similarity	Mean	0.27

Table 13: Variable: Only layer type of input data and output data without tool name

Only Layer Type		
Measurements		Similarity Value
Average Similarity	Largest	0.74
Average Similarity	Smallest	-0.13
Average Similarity	Median	0.36
Average Similarity	Mean	0.34

Table 14: Variable: Only measurement type of input data and output data with- out tool name

Only Measurement Type		
Measurements		Similarity Value
Average	Largest Similarity	0.71
Average	Smallest Similarity	0.03
Average	Median Similarity	0.37
Average	Mean Similarity	0.37