Exclude and Restrict to UMLS Semantic Types

Exclude Semantic Types

If certain UMLS Semantic Types (SemTypes) are not useful for your research or produce undesirable results, MetaMap will ignore those SemTypes if called with the --exclude_sts or -k command-line option. E.g., to exclude Drug Delivery Devices, Medical Devices, and Research Devices, call MetaMap with

-k drdd, medd, resd

Restrict to Semantic Types

Conversely, if your research is focused exclusively on certain SemTypes, MetaMap will restrict its results to those SemTypes if called with the --restrict_to_sts or -J command-line option. E.g., to restrict results to concepts in the Disease or Syndrome and Injury or Poisoning SemTypes, call MetaMap with

-J dsyn,inpo

The names of the SemTypes called with -k and -J are all lowercase.

More information about UMLS Semantic Types is available here.

Here are two examples of using --exclude_sts to reduce spurious ambiguity and/or limit false positives:

1. Ambiguous strings of SemType Gene or Genome (gngm): Many strings of the gngm SemType such as those below are ambiguous because they also denote UMLS concepts that are not primarily genetic. Excluding the gngm SemType may be useful to some researchers.

MIA	ARM	ART	BAR	CAMEL	CAR	CELL	CHAD	CLAMP	CLOCK	COIL
CRAMP	DREAM	DUET	FAN	FAT	FIAT	FIST	FLAP	FLOWER	GAS	GILT
GRASP	GRID	GRIT	HANG	HASH	HIP	HOP	HUB	HUGO	IMPACT	ITCH
JAM	JET	KID	KIR	LAMA	LAP	LARD	LIP	LOTUS	MAIL	MASK
MELT	MICA	MINERVA	MINT	MIST	MOLT	NAIL	NET	NODAL	NUMB	PAD
PALM	PARIS	PHI	PINCH	PION	PIP	PIT	POLE	PRUNE	PRY	RAGE
RAMP	SHARP	SHOT	SLUG	SOUL	SPA	SPICE	SPOT	STING	TACTILE	TAUT
TESMIN	TIP	TOPORS	TRAM	TRAP	TRUSS	TUB	TUNA	VERGE	WAVE	iCE

2. Strings of SemType Finding (fndg): We have found the fndg SemType to be especially noisy because many strings of that SemType, in particular numerical strings such as those below, are of dubious usefulness. Excluding the fndg SemType may be useful to some researchers.

0	0%	0.15	0.3	0.45	0.6	0.75	0.9	1
1 - 3	1+	1-10	1-2	1-24%	1.05	1.2	1.35	1.5
1.65	1.8	1.95	1/24	1/36	1/60	10%	10-14	11-12
11-16	13-14	15	15-19	15-21	16	17-24	19-26	2
2+	2.25	2.5	20%	20-24	20-25	29+	3	3 4
3 6	3+	3-4	3.5	3/12	3/18	3/24	3/36	3/4
3/6	3/60	3/9	30%	4	4+	4.5	400	4000
425	450	5	5 6	5-6	5-9	5.5	50%	500
550	575	6 9	6+11	6.5	6/12	6/18	6/24	6/3
6/30	6/36	6/4	6/45	6/5	6/6	6/60	6/9	60%
600	6000	65+	650	7-8	70%	75%)	80%	800
8000	850	9-10	9-18	90%	<50	+	++	+++

Suppressing mostly numerical concepts such as those shown above can also be accomplished using this option of MetaMap2016V2.