

Rough Analysis of v2_marked_and_linked_train.json

General Information

- Num. Entries: 63
- Tokenizer used: dmis-lab/biobert-v1.1

Definitions:

- Words represent the number of entities that have been considered distinct during the process of assigning tags to part of the text, it is essentially the text split on spaces
- Tokens are the actual pieces of data going into the model, punctuation is split from words and some large words are split down into smaller parts
- Total tags is the total number of tags across all words (If two or more words are part of the same tag this counts as a tag for each word)
- Unique tags is the total number of tags that are unique (not part of same tag)
- Max tags is the most tags a single word has
- Tagged words is the number of words that have at least one tag
- Tagged words % is the percentage of words that are tagged, it's given in brackets for individual entries.
- Avg tags is the average number of tags per word that is tagged
- MC words is the number of words with tags across multiple categories
- Total links is the total number of links across all tags (If two or more tags are linked this counts as a link for each one)
- Unique links is the total number of links that are unique (not part of the same link)
- Max links is the most amount of links a single tag or word has associated with it
- Linked tags, words is the number of tags or words that have links
- Linked % tags, words is the percentage of tags or words with links
- Avg links per tag, word is the average number of links per tag or word
- The schema used is BIES as it is all encompassing and models can convert to a lower resolution if wanted (Beginning, Inner, End, Singleton)

Sizes

- Average words: 213.1
- Maximum words: 290
- Minimum words: 109
- Average tokens: 406.6
- Maximum tokens: 607
- Minimum tokens: 229
- Entries with over 512 tokens: 12/63, 19.05%

Tags

Maximums

- Total tags: 61
- Unique tags: 35

- Max tags: 2
- Tagged words: 61
- Tagged words %: 26.32%
- Avg tags: 1.04
- MC words: 1

Averages

- Total tags: 24.78
- Unique tags: 12.48
- Max tags: 0.95
- Tagged words: 24.73
- Tagged words %: 11.44%
- Avg tags: 0.92
- MC words: 0.02

Links

Maximums

- Total links: 69
- Unique links: 7
- Max links per tag, word: 6, 6
- Linked tags, words: 33, 33
- Linked % tags, words: 100.0%, 14.68%
- Avg links per tag, word: 2.88, 2.88

Averages

- Total links: 16.6
- Unique links: 2.0
- Max links per tag, word: 1.67, 1.67
- Linked tags, words: 11.89, 11.87
- Linked % tags, words: 49.19%, 5.61%
- Avg links per tag, word: 1.21, 1.21

Schema

- Maximums (B, I, E, S): 16, 28, 16, 21
- Averages (B, I, E, S): 5.81, 6.49, 5.81, 6.67

Labels

- Label: Total, Maximum per entry, Average per entry

Categories

- Perturbing_Action: 696, 48, 11.05
- Context: 461, 28, 7.32
- Effect: 173, 11, 2.75
- Phenotype: 231, 12, 3.67

Perturbing_Action

- Gene Loss-Of-Function: 203, 15, 3.22
- Gene Gain-Of-Function: 111, 24, 1.76
- Rnai/Knockdown: 133, 20, 2.11
- Pharmacological Inhibition: 118, 20, 1.87
- Pharmacological Augmentation: 58, 15, 0.92

- Other: 73, 14, 1.16

Context

- Patient: 1, 1, 0.02
- Organism: 67, 9, 1.06
- Tissue/Organ: 23, 4, 0.37
- Neoplasm: 34, 4, 0.54
- Graft: 0, 0, 0.0
- Xenograft: 15, 6, 0.24
- Cells: 153, 15, 2.43
- Transformed Cells: 97, 14, 1.54
- Organoid: 3, 2, 0.05
- In Vitro: 28, 2, 0.44
- In Vivo: 40, 4, 0.63

Effect

- Positive: 63, 5, 1.0
- Negative: 81, 6, 1.29
- Regulates: 0, 0, 0.0
- Rescues: 11, 4, 0.17
- No Effect: 18, 5, 0.29

Phenotype

- Adhesion: 0, 0, 0.0
- Apoptosis: 8, 2, 0.13
- Anoikis: 1, 1, 0.02
- Autophagy: 54, 5, 0.86
- Cell Cycle Arrest: 6, 3, 0.1
- Cell Death: 6, 2, 0.1
- Cell Growth: 2, 2, 0.03
- Cell Survival: 0, 0, 0.0
- Colony Formation: 2, 2, 0.03
- Differentiation: 10, 2, 0.16
- Entosis: 0, 0, 0.0
- Epithelial-Mesenchymal Transition: 4, 4, 0.06
- Ferroptosis: 9, 8, 0.14
- Invasion: 1, 1, 0.02
- Metastasis: 30, 8, 0.48
- Migration: 4, 1, 0.06
- Mitophagy: 4, 2, 0.06
- Necroptosis: 0, 0, 0.0
- Necrosis: 2, 1, 0.03
- Oncosis: 0, 0, 0.0
- Proliferation: 17, 2, 0.27
- Pyroptosis: 0, 0, 0.0
- Quiescence: 2, 2, 0.03
- Self-Renewal: 1, 1, 0.02
- Senescence: 12, 4, 0.19
- Transformation: 4, 2, 0.06
- Tumour Growth: 24, 5, 0.38
- Tumourigenesis: 22, 6, 0.35
- Tumour Initiation: 0, 0, 0.0
- Tumour Progression: 2, 2, 0.03
- Tumour Regression: 4, 4, 0.06

Information on each entry

Title: 15_PMDID32249716.txt_CC2.xml, Words: 240, Tokens: 429

Tag data: Total: 26, Unique: 17, Max tags: 1, Tagged words: 26 (11%), Avg tags: 1.0, MC words: 0

Link data: Total: 20, Unique: 3, Max links (tag, word): (2, 2), Linked (tags, words): (18 (69.23%), 18 (7.5%))

Avg links (tag, word): (1.11, 1.11), Schema counts (B, I, E, S): (5, 4, 5, 12)

Title: 207_PMDID31488580.txt_CC2.xml, Words: 109, Tokens: 229

Tag data: Total: 21, Unique: 9, Max tags: 1, Tagged words: 21 (19%), Avg tags: 1.0, MC words: 0

Link data: Total: 16, Unique: 2, Max links (tag, word): (1, 1), Linked (tags, words): (16 (76.19%), 16 (14.68%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (6, 6, 6, 3)

Title: 51_PMDID31007149.txt_CC2.xml, Words: 267, Tokens: 568

Tag data: Total: 47, Unique: 18, Max tags: 1, Tagged words: 47 (18%), Avg tags: 1.0, MC words: 0

Link data: Total: 69, Unique: 6, Max links (tag, word): (4, 4), Linked (tags, words): (24 (51.06%), 24 (8.99%))

Avg links (tag, word): (2.88, 2.88), Schema counts (B, I, E, S): (6, 23, 6, 12)

Title: 185_PMDID32680921.txt_CC2.xml, Words: 160, Tokens: 317

Tag data: Total: 16, Unique: 7, Max tags: 1, Tagged words: 16 (10%), Avg tags: 1.0, MC words: 0

Link data: Total: 12, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (12 (75.0%), 12 (7.5%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (3, 6, 3, 4)

Title: 165_PMDID32303578.txt_CC2.xml, Words: 269, Tokens: 484

Tag data: Total: 27, Unique: 15, Max tags: 1, Tagged words: 27 (10%), Avg tags: 1.0, MC words: 0

Link data: Total: 10, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (10 (37.04%), 10 (3.72%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (8, 4, 8, 7)

Title: 120_PMDID32122909.txt_CC2.xml, Words: 288, Tokens: 599

Tag data: Total: 17, Unique: 9, Max tags: 1, Tagged words: 17 (6%), Avg tags: 1.0, MC words: 0

Link data: Total: 9, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (9 (52.94%), 9 (3.12%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (4, 4, 4, 5)

Title: 7_PMDID31679460.txt_CC2.xml, Words: 255, Tokens: 598

Tag data: Total: 53, Unique: 35, Max tags: 2, Tagged words: 52 (20%), Avg tags: 1.02, MC words: 1

Link data: Total: 17, Unique: 3, Max links (tag, word): (2, 2), Linked (tags, words): (15 (28.3%), 14 (5.49%))

Avg links (tag, word): (1.13, 1.21), Schema counts (B, I, E, S): (16, 2, 16, 19)

Title: 124_PMDID32665355.txt_CC2.xml, Words: 235, Tokens: 437

Tag data: Total: 36, Unique: 17, Max tags: 1, Tagged words: 36 (15%), Avg tags: 1.0, MC words: 0

Link data: Total: 30, Unique: 4, Max links (tag, word): (2, 2), Linked (tags, words): (25 (69.44%), 25 (10.64%))

Avg links (tag, word): (1.2, 1.2), Schema counts (B, I, E, S): (9, 10, 9, 8)

Title: 143_PMDID330223944.txt_CC2.xml, Words: 183, Tokens: 300

Tag data: Total: 8, Unique: 4, Max tags: 1, Tagged words: 8 (4%), Avg tags: 1.0, MC words: 0

Link data: Total: 8, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (8 (100.0%), 8 (4.37%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (3, 1, 3, 1)

Title: 142_PMDID33277366.txt_CC2.xml, Words: 185, Tokens: 326

Tag data: Total: 9, Unique: 5, Max tags: 1, Tagged words: 9 (5%), Avg tags: 1.0, MC words: 0

Link data: Total: 8, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (8 (88.89%), 8 (4.32%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (3, 1, 3, 2)

Title: 96_PMDID30669930.txt_CC2.xml, Words: 183, Tokens: 354

Tag data: Total: 12, Unique: 6, Max tags: 1, Tagged words: 12 (7%), Avg tags: 1.0, MC words: 0

Link data: Total: 7, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (7 (58.33%), 7 (3.83%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (4, 2, 4, 2)

Title: 155_PMDID33372040.txt_CC2.xml, Words: 222, Tokens: 376

Tag data: Total: 41, Unique: 18, Max tags: 1, Tagged words: 41 (18%), Avg tags: 1.0, MC words: 0

Link data: Total: 19, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (11 (26.83%), 11 (4.95%))

Avg links (tag, word): (1.73, 1.73), Schema counts (B, I, E, S): (13, 10, 13, 5)

Title: 42_PMID31276435.txt_CC2.xml, Words: 184, Tokens: 368
Tag data: Total: 12, Unique: 7, Max tags: 1, Tagged words: 12 (7%), Avg tags: 1.0, MC words: 0
Link data: Total: 5, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (5 (41.67%), 5 (2.72%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (4, 1, 4, 3)

Title: 43_PMID31204559.txt_CC2.xml, Words: 186, Tokens: 320
Tag data: Total: 29, Unique: 19, Max tags: 1, Tagged words: 29 (16%), Avg tags: 1.0, MC words: 0
Link data: Total: 14, Unique: 2, Max links (tag, word): (1, 1), Linked (tags, words): (14 (48.28%), 14 (7.53%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (7, 3, 7, 12)

Title: 16_PMID31232177.txt_CC2.xml, Words: 210, Tokens: 420
Tag data: Total: 32, Unique: 16, Max tags: 1, Tagged words: 32 (15%), Avg tags: 1.0, MC words: 0
Link data: Total: 35, Unique: 5, Max links (tag, word): (4, 4), Linked (tags, words): (21 (65.62%), 21 (10.0%))
Avg links (tag, word): (1.67, 1.67), Schema counts (B, I, E, S): (6, 10, 6, 10)

Title: 276_PMID31879538.txt_CC2.xml, Words: 182, Tokens: 322
Tag data: Total: 23, Unique: 12, Max tags: 1, Tagged words: 23 (13%), Avg tags: 1.0, MC words: 0
Link data: Total: 21, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (15 (65.22%), 15 (8.24%))
Avg links (tag, word): (1.4, 1.4), Schema counts (B, I, E, S): (6, 5, 6, 6)

Title: 248_PMID32217667.txt_CC2.xml, Words: 199, Tokens: 324
Tag data: Total: 0, Unique: 0, Max tags: 0, Tagged words: 0 (0%), Avg tags: 0, MC words: 0
Link data: Total: 0, Unique: 0, Max links (tag, word): (0, 0), Linked (tags, words): (0 (0%), 0 (0%))
Avg links (tag, word): (0, 0), Schema counts (B, I, E, S): (0, 0, 0, 0)

Title: 199_PMID32265225.txt_CC2.xml, Words: 290, Tokens: 548
Tag data: Total: 28, Unique: 11, Max tags: 1, Tagged words: 28 (10%), Avg tags: 1.0, MC words: 0
Link data: Total: 15, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (15 (53.57%), 15 (5.17%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (8, 9, 8, 3)

Title: 257_PMID28465358.txt_CC2.xml, Words: 182, Tokens: 332
Tag data: Total: 13, Unique: 9, Max tags: 1, Tagged words: 13 (7%), Avg tags: 1.0, MC words: 0
Link data: Total: 5, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (5 (38.46%), 5 (2.75%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (3, 1, 3, 6)

Title: 73_PMID30444165.txt_CC2.xml, Words: 260, Tokens: 518
Tag data: Total: 43, Unique: 24, Max tags: 1, Tagged words: 43 (17%), Avg tags: 1.0, MC words: 0
Link data: Total: 39, Unique: 5, Max links (tag, word): (2, 2), Linked (tags, words): (33 (76.74%), 33 (12.69%))
Avg links (tag, word): (1.18, 1.18), Schema counts (B, I, E, S): (8, 11, 8, 16)

Title: 166_PMID32060146.txt_CC2.xml, Words: 269, Tokens: 431
Tag data: Total: 24, Unique: 10, Max tags: 1, Tagged words: 24 (9%), Avg tags: 1.0, MC words: 0
Link data: Total: 29, Unique: 3, Max links (tag, word): (2, 2), Linked (tags, words): (22 (91.67%), 22 (8.18%))
Avg links (tag, word): (1.32, 1.32), Schema counts (B, I, E, S): (7, 7, 7, 3)

Title: 251_PMID30842218.txt_CC2.xml, Words: 160, Tokens: 262
Tag data: Total: 0, Unique: 0, Max tags: 0, Tagged words: 0 (0%), Avg tags: 0, MC words: 0
Link data: Total: 0, Unique: 0, Max links (tag, word): (0, 0), Linked (tags, words): (0 (0%), 0 (0%))
Avg links (tag, word): (0, 0), Schema counts (B, I, E, S): (0, 0, 0, 0)

Title: 228_PMID28698299.txt_CC2.xml, Words: 229, Tokens: 400
Tag data: Total: 21, Unique: 16, Max tags: 1, Tagged words: 21 (9%), Avg tags: 1.0, MC words: 0
Link data: Total: 6, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (6 (28.57%), 6 (2.62%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (4, 1, 4, 12)

Title: 115_PMID32661137.txt_CC2.xml, Words: 222, Tokens: 375
Tag data: Total: 49, Unique: 15, Max tags: 1, Tagged words: 49 (22%), Avg tags: 1.0, MC words: 0
Link data: Total: 24, Unique: 2, Max links (tag, word): (1, 1), Linked (tags, words): (24 (48.98%), 24 (10.81%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (14, 20, 14, 1)

Title: 125_PMIDS32213542.txt_CC2.xml, Words: 217, Tokens: 397
Tag data: Total: 18, Unique: 10, Max tags: 1, Tagged words: 18 (8%), Avg tags: 1.0, MC words: 0

Link data: Total: 21, Unique: 3, Max links (tag, word): (2, 2), Linked (tags, words): (16 (88.89%), 16 (7.37%))
Avg links (tag, word): (1.31, 1.31), Schema counts (B, I, E, S): (6, 2, 6, 4)

Title: 38_PMID31679456.txt_CC2.xml, Words: 256, Tokens: 486

Tag data: Total: 55, Unique: 27, Max tags: 2, Tagged words: 53 (21%), Avg tags: 1.04, MC words: 0

Link data: Total: 33, Unique: 3, Max links (tag, word): (2, 2), Linked (tags, words): (21 (38.18%), 21 (8.2%))

Avg links (tag, word): (1.57, 1.57), Schema counts (B, I, E, S): (7, 21, 7, 20)

Title: 247_PMID30366904.txt_CC2.xml, Words: 114, Tokens: 237

Tag data: Total: 0, Unique: 0, Max tags: 0, Tagged words: 0 (0%), Avg tags: 0, MC words: 0

Link data: Total: 0, Unique: 0, Max links (tag, word): (0, 0), Linked (tags, words): (0 (0%), 0 (0%))

Avg links (tag, word): (0, 0), Schema counts (B, I, E, S): (0, 0, 0, 0)

Title: 77_PMID30806153.txt_CC2.xml, Words: 190, Tokens: 422

Tag data: Total: 15, Unique: 9, Max tags: 1, Tagged words: 15 (8%), Avg tags: 1.0, MC words: 0

Link data: Total: 8, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (8 (53.33%), 8 (4.21%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (2, 4, 2, 7)

Title: 114_PMID32312833.txt_CC2.xml, Words: 176, Tokens: 367

Tag data: Total: 12, Unique: 8, Max tags: 1, Tagged words: 12 (7%), Avg tags: 1.0, MC words: 0

Link data: Total: 14, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (8 (66.67%), 8 (4.55%))

Avg links (tag, word): (1.75, 1.75), Schema counts (B, I, E, S): (3, 1, 3, 5)

Title: 44_PMID31818185.txt_CC2.xml, Words: 249, Tokens: 455

Tag data: Total: 31, Unique: 17, Max tags: 1, Tagged words: 31 (12%), Avg tags: 1.0, MC words: 0

Link data: Total: 6, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (6 (19.35%), 6 (2.41%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (8, 6, 8, 9)

Title: 32_PMID30290714.txt_CC2.xml, Words: 255, Tokens: 462

Tag data: Total: 26, Unique: 17, Max tags: 1, Tagged words: 26 (10%), Avg tags: 1.0, MC words: 0

Link data: Total: 13, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (9 (34.62%), 9 (3.53%))

Avg links (tag, word): (1.44, 1.44), Schema counts (B, I, E, S): (8, 1, 8, 9)

Title: 87_PMID32160082.txt_CC2.xml, Words: 258, Tokens: 538

Tag data: Total: 61, Unique: 23, Max tags: 1, Tagged words: 61 (24%), Avg tags: 1.0, MC words: 0

Link data: Total: 8, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (8 (13.11%), 8 (3.1%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (10, 28, 10, 13)

Title: 254_PMID32217666.txt_CC2.xml, Words: 204, Tokens: 302

Tag data: Total: 0, Unique: 0, Max tags: 0, Tagged words: 0 (0%), Avg tags: 0, MC words: 0

Link data: Total: 0, Unique: 0, Max links (tag, word): (0, 0), Linked (tags, words): (0 (0%), 0 (0%))

Avg links (tag, word): (0, 0), Schema counts (B, I, E, S): (0, 0, 0, 0)

Title: 235_PMID32217665.txt_CC2.xml, Words: 166, Tokens: 301

Tag data: Total: 16, Unique: 8, Max tags: 1, Tagged words: 16 (10%), Avg tags: 1.0, MC words: 0

Link data: Total: 11, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (11 (68.75%), 11 (6.63%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (4, 4, 4, 4)

Title: 5_PMID30870073.txt_CC2.xml, Words: 242, Tokens: 491

Tag data: Total: 14, Unique: 6, Max tags: 1, Tagged words: 14 (6%), Avg tags: 1.0, MC words: 0

Link data: Total: 5, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (5 (35.71%), 5 (2.07%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (3, 5, 3, 3)

Title: 35_PMID30894058.txt_CC2.xml, Words: 253, Tokens: 543

Tag data: Total: 37, Unique: 17, Max tags: 1, Tagged words: 37 (15%), Avg tags: 1.0, MC words: 0

Link data: Total: 6, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (6 (16.22%), 6 (2.37%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (9, 11, 9, 8)

Title: 168_PMID32900774.txt_CC2.xml, Words: 218, Tokens: 354

Tag data: Total: 22, Unique: 12, Max tags: 1, Tagged words: 22 (10%), Avg tags: 1.0, MC words: 0

Link data: Total: 15, Unique: 2, Max links (tag, word): (1, 1), Linked (tags, words): (15 (68.18%), 15 (6.88%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (6, 4, 6, 6)

Title: 45_PMID30686098.txt_CC2.xml, Words: 238, Tokens: 534
Tag data: Total: 11, Unique: 6, Max tags: 1, Tagged words: 11 (5%), Avg tags: 1.0, MC words: 0
Link data: Total: 7, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (7 (63.64%), 7 (2.94%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (4, 1, 4, 2)

Title: 109_PMID32217697.txt_CC2.xml, Words: 213, Tokens: 429
Tag data: Total: 16, Unique: 9, Max tags: 1, Tagged words: 16 (8%), Avg tags: 1.0, MC words: 0
Link data: Total: 21, Unique: 3, Max links (tag, word): (3, 3), Linked (tags, words): (11 (68.75%), 11 (5.16%))
Avg links (tag, word): (1.91, 1.91), Schema counts (B, I, E, S): (6, 1, 6, 3)

Title: 72_PMID31238788.txt_CC2.xml, Words: 178, Tokens: 368
Tag data: Total: 28, Unique: 17, Max tags: 1, Tagged words: 28 (16%), Avg tags: 1.0, MC words: 0
Link data: Total: 15, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (14 (50.0%), 14 (7.87%))
Avg links (tag, word): (1.07, 1.07), Schema counts (B, I, E, S): (8, 3, 8, 9)

Title: 24_PMID30859901.txt_CC2.xml, Words: 260, Tokens: 605
Tag data: Total: 36, Unique: 21, Max tags: 1, Tagged words: 36 (14%), Avg tags: 1.0, MC words: 0
Link data: Total: 22, Unique: 3, Max links (tag, word): (2, 2), Linked (tags, words): (19 (52.78%), 19 (7.31%))
Avg links (tag, word): (1.16, 1.16), Schema counts (B, I, E, S): (5, 10, 5, 16)

Title: 224_PMID30692208.txt_CC2.xml, Words: 206, Tokens: 380
Tag data: Total: 17, Unique: 5, Max tags: 1, Tagged words: 17 (8%), Avg tags: 1.0, MC words: 0
Link data: Total: 26, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (17 (100.0%), 17 (8.25%))
Avg links (tag, word): (1.53, 1.53), Schema counts (B, I, E, S): (5, 7, 5, 0)

Title: 100_PMID31462126.txt_CC2.xml, Words: 141, Tokens: 277
Tag data: Total: 35, Unique: 20, Max tags: 1, Tagged words: 35 (25%), Avg tags: 1.0, MC words: 0
Link data: Total: 14, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (8 (22.86%), 8 (5.67%))
Avg links (tag, word): (1.75, 1.75), Schema counts (B, I, E, S): (7, 8, 7, 13)

Title: 147_PMID31941699.txt_CC2.xml, Words: 180, Tokens: 295
Tag data: Total: 32, Unique: 12, Max tags: 1, Tagged words: 32 (18%), Avg tags: 1.0, MC words: 0
Link data: Total: 7, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (7 (21.88%), 7 (3.89%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (5, 15, 5, 7)

Title: 18_PMID31931659.txt_CC2.xml, Words: 247, Tokens: 534
Tag data: Total: 40, Unique: 24, Max tags: 1, Tagged words: 40 (16%), Avg tags: 1.0, MC words: 0
Link data: Total: 26, Unique: 3, Max links (tag, word): (2, 2), Linked (tags, words): (19 (47.5%), 19 (7.69%))
Avg links (tag, word): (1.37, 1.37), Schema counts (B, I, E, S): (6, 10, 6, 18)

Title: 226_PMID33303641.txt_CC2.xml, Words: 152, Tokens: 276
Tag data: Total: 40, Unique: 22, Max tags: 1, Tagged words: 40 (26%), Avg tags: 1.0, MC words: 0
Link data: Total: 39, Unique: 6, Max links (tag, word): (5, 5), Linked (tags, words): (22 (55.0%), 22 (14.47%))
Avg links (tag, word): (1.77, 1.77), Schema counts (B, I, E, S): (7, 11, 7, 15)

Title: 34_PMID31286822.txt_CC2.xml, Words: 264, Tokens: 535
Tag data: Total: 14, Unique: 8, Max tags: 1, Tagged words: 14 (5%), Avg tags: 1.0, MC words: 0
Link data: Total: 7, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (7 (50.0%), 7 (2.65%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (3, 3, 3, 5)

Title: 70_PMID31448672.txt_CC2.xml, Words: 179, Tokens: 365
Tag data: Total: 22, Unique: 10, Max tags: 1, Tagged words: 22 (12%), Avg tags: 1.0, MC words: 0
Link data: Total: 7, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (7 (31.82%), 7 (3.91%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (6, 6, 6, 4)

Title: 101_PMID32788173.txt_CC2.xml, Words: 177, Tokens: 346
Tag data: Total: 17, Unique: 11, Max tags: 1, Tagged words: 17 (10%), Avg tags: 1.0, MC words: 0
Link data: Total: 47, Unique: 7, Max links (tag, word): (6, 6), Linked (tags, words): (17 (100.0%), 17 (9.6%))
Avg links (tag, word): (2.76, 2.76), Schema counts (B, I, E, S): (5, 1, 5, 6)

Title: 161_PMID32816909.txt_CC2.xml, Words: 253, Tokens: 488
Tag data: Total: 39, Unique: 14, Max tags: 1, Tagged words: 39 (15%), Avg tags: 1.0, MC words: 0

Link data: Total: 14, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (14 (35.9%), 14 (5.53%))
Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (10, 15, 10, 4)

Title: 250_PMD30842218.txt_CC2.xml, Words: 160, Tokens: 262

Tag data: Total: 0, Unique: 0, Max tags: 0, Tagged words: 0 (0%), Avg tags: 0, MC words: 0

Link data: Total: 0, Unique: 0, Max links (tag, word): (0, 0), Linked (tags, words): (0 (0%), 0 (0%))

Avg links (tag, word): (0, 0), Schema counts (B, I, E, S): (0, 0, 0, 0)

Title: 149_PMD33093168.txt_CC2.xml, Words: 226, Tokens: 399

Tag data: Total: 20, Unique: 10, Max tags: 1, Tagged words: 20 (9%), Avg tags: 1.0, MC words: 0

Link data: Total: 7, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (7 (35.0%), 7 (3.1%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (7, 3, 7, 3)

Title: 84_PMD32075509.txt_CC2.xml, Words: 265, Tokens: 463

Tag data: Total: 42, Unique: 17, Max tags: 1, Tagged words: 42 (16%), Avg tags: 1.0, MC words: 0

Link data: Total: 20, Unique: 3, Max links (tag, word): (3, 3), Linked (tags, words): (12 (28.57%), 12 (4.53%))

Avg links (tag, word): (1.67, 1.67), Schema counts (B, I, E, S): (8, 17, 8, 9)

Title: 10_PMD30898012.txt_CC2.xml, Words: 239, Tokens: 436

Tag data: Total: 41, Unique: 13, Max tags: 1, Tagged words: 41 (17%), Avg tags: 1.0, MC words: 0

Link data: Total: 34, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (21 (51.22%), 21 (8.79%))

Avg links (tag, word): (1.62, 1.62), Schema counts (B, I, E, S): (6, 22, 6, 7)

Title: 119_PMD33023950.txt_CC2.xml, Words: 231, Tokens: 417

Tag data: Total: 28, Unique: 14, Max tags: 1, Tagged words: 28 (12%), Avg tags: 1.0, MC words: 0

Link data: Total: 41, Unique: 4, Max links (tag, word): (4, 4), Linked (tags, words): (17 (60.71%), 17 (7.36%))

Avg links (tag, word): (2.41, 2.41), Schema counts (B, I, E, S): (7, 7, 7, 7)

Title: 3_PMD31451060.txt_CC2.xml, Words: 210, Tokens: 398

Tag data: Total: 19, Unique: 12, Max tags: 1, Tagged words: 19 (9%), Avg tags: 1.0, MC words: 0

Link data: Total: 12, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (10 (52.63%), 10 (4.76%))

Avg links (tag, word): (1.2, 1.2), Schema counts (B, I, E, S): (5, 2, 5, 7)

Title: 86_PMD31865844.txt_CC2.xml, Words: 182, Tokens: 290

Tag data: Total: 21, Unique: 7, Max tags: 1, Tagged words: 21 (12%), Avg tags: 1.0, MC words: 0

Link data: Total: 41, Unique: 3, Max links (tag, word): (3, 3), Linked (tags, words): (17 (80.95%), 17 (9.34%))

Avg links (tag, word): (2.41, 2.41), Schema counts (B, I, E, S): (5, 9, 5, 2)

Title: 289_PMD30692206.txt_CC2.xml, Words: 174, Tokens: 353

Tag data: Total: 21, Unique: 9, Max tags: 1, Tagged words: 21 (12%), Avg tags: 1.0, MC words: 0

Link data: Total: 11, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (11 (52.38%), 11 (6.32%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (4, 8, 4, 5)

Title: 285_PMD29891559.txt_CC2.xml, Words: 188, Tokens: 345

Tag data: Total: 35, Unique: 15, Max tags: 1, Tagged words: 35 (19%), Avg tags: 1.0, MC words: 0

Link data: Total: 8, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (8 (22.86%), 8 (4.26%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (13, 7, 13, 2)

Title: 9_PMD31441382.txt_CC2.xml, Words: 261, Tokens: 607

Tag data: Total: 38, Unique: 28, Max tags: 1, Tagged words: 38 (15%), Avg tags: 1.0, MC words: 0

Link data: Total: 24, Unique: 4, Max links (tag, word): (4, 4), Linked (tags, words): (10 (26.32%), 10 (3.83%))

Avg links (tag, word): (2.4, 2.4), Schema counts (B, I, E, S): (7, 3, 7, 21)

Title: 53_PMD30774023.txt_CC2.xml, Words: 207, Tokens: 400

Tag data: Total: 25, Unique: 13, Max tags: 1, Tagged words: 25 (12%), Avg tags: 1.0, MC words: 0

Link data: Total: 12, Unique: 1, Max links (tag, word): (1, 1), Linked (tags, words): (12 (48.0%), 12 (5.8%))

Avg links (tag, word): (1.0, 1.0), Schema counts (B, I, E, S): (3, 9, 3, 10)

Title: 140_PMD33239431.txt_CC2.xml, Words: 229, Tokens: 431

Tag data: Total: 20, Unique: 11, Max tags: 1, Tagged words: 20 (9%), Avg tags: 1.0, MC words: 0

Link data: Total: 12, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (9 (45.0%), 9 (3.93%))

Avg links (tag, word): (1.33, 1.33), Schema counts (B, I, E, S): (8, 1, 8, 3)

Title: 128_PMID32193288.txt_CC2.xml, Words: 199, Tokens: 391

Tag data: Total: 10, Unique: 5, Max tags: 1, Tagged words: 10 (5%), Avg tags: 1.0, MC words: 0

Link data: Total: 14, Unique: 2, Max links (tag, word): (2, 2), Linked (tags, words): (10 (100.0%), 10 (5.03%))

Avg links (tag, word): (1.4, 1.4), Schema counts (B, I, E, S): (3, 2, 3, 2)