

**Table S1. Biomedical information extraction-based features extracted by scispaCy.**

<b>Features</b>	<b>Models</b>
Entrez Gene and Protein Ontology	craft
Sequence Ontology	craft
NCBI Taxonomy	craft
Chemical Entities of Biological Interest	craft
Gene Ontology	craft
Cell Ontology	craft
DNA	jnlpba
Cell Type	jnlpba
Cell Line	jnlpba
RNA	jnlpba
Protein	jnlpba
Disease Names	bc5cdr
Chemicals Names	bc5cdr
Amino Acids	bionlp13cg
Anatomical Systems	bionlp13cg
Cancer	bionlp13cg
Cell	bionlp13cg
Cellular Components	bionlp13cg
Developing Anatomical Structure	bionlp13cg
Gene or Gene Product	bionlp13cg
Immaterial Anatomical Entity	bionlp13cg
Multi-tissue Structure	bionlp13cg
Organ	bionlp13cg
Organism	bionlp13cg
Organism Subdivision	bionlp13cg
Organism Substance	bionlp13cg
Pathological Formation	bionlp13cg
Simple Chemical	bionlp13cg

---

Tissue

bionlp13cg

---

**Table S2: Biomedical information extraction-based features extracted by MetaMap.**

<b>Features</b>	<b>Full Names</b>
<b>aapp</b>	Amino Acid, Peptide, or Protein
<b>acab</b>	Acquired Abnormality
<b>acty</b>	Activity
<b>aggp</b>	Age Group
<b>amas</b>	Amino Acid Sequence
<b>amph</b>	Amphibian
<b>anab</b>	Anatomical Abnormality
<b>anim</b>	Animal
<b>anst</b>	Anatomical Structure
<b>antb</b>	Antibiotic
<b>arch</b>	Archaeon
<b>bacs</b>	Biologically Active Substance
<b>bact</b>	Bacterium
<b>bdsu</b>	Body Substance
<b>bdsy</b>	Body System
<b>bhvr</b>	Behavior
<b>biof</b>	Biologic Function
<b>bird</b>	Bird
<b>blor</b>	Body Location or Region
<b>bmod</b>	Biomedical Occupation or Discipline
<b>bodm</b>	Biomedical or Dental Material
<b>bpoc</b>	Body Part, Organ, or Organ Component
<b>bsoj</b>	Body Space or Junction
<b>celc</b>	Cell Component
<b>celf</b>	Cell Function
<b>cell</b>	Cell

<b>cgab</b>	Congenital Abnormality
<b>chem</b>	Chemical
<b>chvf</b>	Chemical Viewed Functionally
<b>chvs</b>	Chemical Viewed Structurally
<b>clas</b>	Classification
<b>clna</b>	Clinical Attribute
<b>clnd</b>	Clinical Drug
<b>cnce</b>	Conceptual Entity
<b>comd</b>	Cell or Molecular Dysfunction
<b>crbs</b>	Carbohydrate Sequence
<b>diap</b>	Diagnostic Procedure
<b>dora</b>	Daily or Recreational Activity
<b>drdd</b>	Drug Delivery Device
<b>dsyn</b>	Disease or Syndrome
<b>edac</b>	Educational Activity
<b>eehu</b>	Environmental Effect of Humans
<b>elii</b>	Element, Ion, or Isotope
<b>emod</b>	Experimental Model of Disease
<b>emst</b>	Embryonic Structure
<b>enty</b>	Entity
<b>enzy</b>	Enzyme
<b>euka</b>	Eukaryote
<b>evnt</b>	Event
<b>famg</b>	Family Group
<b>ffas</b>	Fully Formed Anatomical Structure
<b>fish</b>	Fish
<b>fndg</b>	Finding
<b>fngs</b>	Fungus
<b>food</b>	Food

<b>ftcn</b>	Functional Concept
<b>genf</b>	Genetic Function
<b>geoa</b>	Geographic Area
<b>gngm</b>	Gene or Genome
<b>gora</b>	Governmental or Regulatory Activity
<b>grpa</b>	Group Attribute
<b>grup</b>	Group
<b>hcpp</b>	Human-caused Phenomenon or Process
<b>hcro</b>	Health Care Related Organization
<b>hlcs</b>	Health Care Activity
<b>hops</b>	Hazardous or Poisonous Substance
<b>horm</b>	Hormone
<b>humn</b>	Human
<b>idcn</b>	Idea or Concept
<b>imft</b>	Immunologic Factor
<b>inbe</b>	Individual Behavior
<b>inch</b>	Inorganic Chemical
<b>inpo</b>	Injury or Poisoning
<b>inpr</b>	Intellectual Product
<b>irda</b>	Indicator, Reagent, or Diagnostic Aid
<b>lang</b>	Language
<b>lbpr</b>	Laboratory Procedure
<b>lbtr</b>	Laboratory or Test Result
<b>mamm</b>	Mammal
<b>mbrt</b>	Molecular Biology Research Technique

<b>mcha</b>	Machine Activity
<b>medd</b>	Medical Device
<b>menp</b>	Mental Process
<b>mnob</b>	Manufactured Object
<b>mobd</b>	Mental or Behavioral Dysfunction
<b>moft</b>	Molecular Function
<b>mosq</b>	Molecular Sequence
<b>neop</b>	Neoplastic Process
<b>nnon</b>	Nucleic Acid, Nucleoside, or Nucleotide
<b>npop</b>	Natural Phenomenon or Process
<b>nusq</b>	Nucleotide Sequence
<b>ocac</b>	Occupational Activity
<b>ocdi</b>	Occupation or Discipline
<b>orch</b>	Organic Chemical
<b>orga</b>	Organism Attribute
<b>orgf</b>	Organism Function
<b>orgm</b>	Organism
<b>orgt</b>	Organization
<b>ortf</b>	Organ or Tissue Function
<b>patf</b>	Pathologic Function
<b>phob</b>	Physical Object
<b>phpr</b>	Phenomenon or Process
<b>phsf</b>	Physiologic Function
<b>phsu</b>	Pharmacologic Substance
<b>plnt</b>	Plant
<b>podg</b>	Patient or Disabled Group
<b>popg</b>	Population Group
<b>prog</b>	Professional or Occupational Group

<b>pros</b>	Professional Society
<b>qlco</b>	Qualitative Concept
<b>qnco</b>	Quantitative Concept
<b>rcpt</b>	Receptor
<b>rept</b>	Reptile
<b>resa</b>	Research Activity
<b>resd</b>	Research Device
<b>rnlw</b>	Regulation or Law
<b>sbst</b>	Substance
<b>shro</b>	Self-help or Relief Organization
<b>socb</b>	Social Behavior
<b>sosy</b>	Sign or Symptom
<b>spco</b>	Spatial Concept
<b>tisu</b>	Tissue
<b>tmco</b>	Temporal Concept
<b>topp</b>	Therapeutic or Preventive Procedure
<b>viru</b>	Virus
<b>vita</b>	Vitamin
<b>vtbt</b>	Vertebrate

**Table S3: Biomedical information extraction-based features extracted by spaCy customized model.**

<b>Features</b>
Pathogen
Medical Condition names
Medicine names

**Table S4: Named-entity recognition features.**

<b>Features</b>
Person
FAC
GPE
Product
Work of art
Language
Time
Money
Cardinal
NORP
Organization
Location
Event
Law
Date
Percent
Quantity
Ordinal

**Table S5: Linguistic features.**

<b>Features</b>
Num of stopwords
Num of @
Num of numeric values
Num of lowercase characters
Text language
Character count
Average word length
Num of ?
Num of /
Num of #
Num of uppercase characters
Num of all uppercase characters
Word count
Sentence count
Average sentence length

**Table S6: Sentiment Features.**

<b>Features</b>
Positive sentiment score
Negative sentiment score
Neutral sentiment score
Compound sentiment score
Text polarity