

Index

A

abbreviation identification 360
 ABNER (A Biomedical Named Entity Recognizer) 315
 ABNER's graphical user interface 316
 ABNER's two built-in tagging models 317
 ABNER distribution 317
 acronyms, integration of 25
 administrator physician 364
 agriculture NEs (ANEs) 331
 ambiguity 142
 ambiguous terms 151
 American National Library of Medicine (NLM) 222
 anatomy 349
 ANN 260
 annotation for knowledge extraction 333
 annotation process 333
 annotation schema 333
 annotation schema for causality extraction 338
 anonymisation 366
 ANR VODEL project 78
 application programming interface (API) 317
 artificial neural networks (ANNs) 345
 asthma 197
 attributes 2
 auditor physician 365
 automatic indexing 11

B

base classifiers, adaptive selection of 50
 Bigrams 209
 bilingual sentence alignment 81
 BioCaster 327
 BioLiterature 303

biological databases, automatic maintenance of 321
 biomedical applications 142
 biomedical applications, word sense disambiguation in 142
 biomedical databases 39
 biomedical domain 106
 biomedical homograph 63
 biomedical information extraction 315
 biomedical information extraction, software tool for 315
 biomedical language processing 224
 biomedical language processing (BLP) 222
 biomedical literature 163
 biomedical literature, sequence variants of genes 289
 biomedical natural language processing 12
 biomedical ontologies 124
 biomedical ontologies, lexical enrichment of 124
 biomedical ontologies with WordNet 126
 biomedical terminological resources 58
 biomedical terminologies 60
 biomedical terms, translation of 106
 biomedical text mining, classification in 149
 biomedicine 1
 biomedicine, text mining for 1
 biomedicine text processing, problems in 227
 BioThesaurus 60
 bottom-up population 67
 British National Corpus (BNC) 128

C

candidate sets 203
 case recommendations, definition 280

causality tags 333
 causation, definition 280
 CCHMC 44
 CCHMC collection 38, 48
 cellular components 304
 CE text, outcome detection in 211
 CE Text, polarity classification in 211
 change phrases 209
 characterisation algorithm 232
 ChEBI 61
 chemical terminologies 60
 CISMeF 81
 classification-based WSD 147
 classifiers 281
 clinical case reporting 274
 clinical case studies 274
 routine patient case reports 274
 clinical case studies 273
 Clinical Data Analytics Language (CliniDAL)
 372
 clinical evidence (CE) 191
 clinical literature, analyzing the text 190
 clinical notes, analysing 357
 clinical outcome detection 208
 clinical outcomes, polarity of 207
 clinical QA 194
 CLIR experiment 118
 clustering-based WSD 147
 co-occurrence 345
 coherence check 230
 collaborative PMM process 330
 collective intelligence analysis 325
 common hypernyms (HYPE) 89
 common hyponyms (HYPO) 89
 common synonyms (SYN) 89
 common synonyms and hyponym/hypernym
 inclusion (SHH) 89
 compound segmentation 22
 concept recognition 361
 conceptual roles 194
 conditional random fields (CRFs) 315
 confidential information 274
 context features 209
 contextual problem 186
 contextual tagging 177
 contextual tagging, functional 182

contextual tagging, morpho-syntactic 182
 cores, importance of 197
 corpus, descriptive statistics of 280
 corpus of narrative notes 366
 corpus selection and analysis 350
 corresponding finite-state machine 319
 CorTag 177, 180
 CorTag, overview of 180
 CorTag's primary feature 179
 CRF graphs 319

D

data collections 44
 data mining (DM) 3
 data retrieval system 358
 data set 201
 definition representation, levels of 88
 derivational patterns 20
 diagnosis, definition 279
 diagnostic procedures 279
 diagnostic test, definition 279
 dictionaries 84
 dictionary-based methods 144
 disambiguation 134
 disambiguation algorithm 134
 disambiguation algorithm, refined 135
 discourse markers 233
 discourse processing 222
 disease warning 330
 document categorization 320
 document classification 222
 domain-independent polysemy 62
 domain-specific semantic categories 273
 DrugBank 60

E

elliptic coordinations 23
 entities 304
 EntrezGene 61
 events 304
 evidence-based medicine (EBM) 191
 evidence tagging 175
 evolutionary algorithm 78
 evolutionary algorithms (EAs) 91
 evolutionary alignment model 97
 exceptionality, definition 280

Index

exclusion rules, definition 280
external lexical resources 21
extra-clausal processing 169
extraction set 214

F

FACTA 6
factoid question answering (FQA) 192
facts 2
fact types 2
feature set 281
feature set, impact of 156
feature sets 202
FMA 136
Foundational Model of Anatomy (FMA) 128
frame constructor 358
frames 194
frame type 193
fusion 170

G

GATE tool 180
Gene Ontology (GO) 60, 64, 75
gene ontology (GO) 41, 60, 305
generalized non-linear relevance feedback 262
genes 304
genetic variation 289
genre specific semantics 279
genuine polysemy 62
GO (gene ontology) 302
GO, subgraph of 306
GOA (GO annotation) 302
GOAnnotator 306
GPSDB 60
graphical user interface (GUI) 317
growth 197

H

head-modifier normalization 18
hidden Markov models (HMMs) 318, 345
higher-level information extraction 320
human anatomy ontology, lexical enrichment of 128
human diseases 325
human expert information extraction 331

hypernym 132
hyponym 132

I

identical molecules 344
IE template filling 3
implicit semantic relations 63
indexing 72
inflection normalization 18
information (medical content) retrieval systems 255
information extraction 163, 222
information extraction (IE) 2, 195, 316, 320
information extraction of protein phosphorylation 163
information overload 1
information retrieval 58, 222
information retrieval (IR) 2, 190, 320
information retrieval, efficient 78
information retrieval systems 79
information systems 125
interactive framework 257
IR systems 58

J

Java API 315

K

k-Nearest Neighbour classifier 90
kernel formalism 93
k Nearest Neighbour algorithm 78
kNN-based alignment model 95
know-how extraction 337
know-why extraction 334
knowledge-based WSD 146
knowledge discovery 1
knowledge discovery, semantic metadata for 3
knowledge extraction experiment 340
knowledge integration 328
knowledge integration, medical resource to 42

L

labeled input sequence 319
language engineering 325
language pairs 116

- LDI 85
- lemmatisation 86
- lexica, assembling 367
- lexical analyzer 358
- lexical form 358
- lexical granularity 11
- lexical knowledge 137
- lexical resources, compilation of 59
- lexical variations 21
- lexical verification 360, 366
- lexicon refinement and enrichment 65
- life science data 11
- life science data resources 59
- life sciences, word ambiguity in 145
- life science texts, WSD in 148
- LingInfo, ontologies with 137
- linguistic description 87
- linguistic metadata 66
- linguistic processing 83, 86
- linguistic research 224
- literature-based discovery 222
- local PMM process 329
- M**
- machine-learning techniques 90
- machine learning approach 142
- machine learning systems 283
- manual annotation 273
- manual annotations 281
- maximum entropy (MaxEnt) classifier 273
- maximum entropy modeling (MaxEnt) 281
- medical brand names 21
- medical content retrieval systems, relevance feedback measures 256
- medical content visualization 30
- medical documents, treemap visualization of 31
- medical information retrieval systems 255
- medical information retrieval systems, performance of 259
- medical ontologies 38, 276
- medical ontology 278
- Medical Subject Headings (MeSH) 60
- medical terminologies, automatic alignment of 78
- medical test, definition 279
- MEDIE 4
- MEDIE, user interface 5
- MedLEE 362
- MEDLEX corpus 17
- MEDLINE 4, 163, 274, 290, 303
- Medline, outcome detection in 212
- Medline, polarity classification in 212
- Medline Plus 60
- MEMODATA 81
- meronym 132
- MeSH 14, 17, 84
- MeSH, descriptors in 16
- MeSH, English 16
- MeSH, Swedish 16
- MeSH, updating medical brand names 21
- MeSH concepts 14
- MeSH paths of Agnosia 15
- MeSH spelling error 21
- MeSH terms, structure of 14
- Message Understanding Conference (MUC) 331
- MetaMap 362
- MetaMap modules 47
- MLTC 51
- molecular biology ontology, lexical enrichment of 131
- molecular biology research projects 301
- morphological descriptors 236
- morphological descriptors, adequacy of 236
- morphological descriptors, distribution of 237
- morphological descriptors, reliability of 236
- morphological normalization 19
- multi-label text categorization 47
- multi-label text categorization (MLTC) system 38
- multi-label text categorization system 38
- multi-scale text processing 228
- multi-word expressions (MWE) 360
- myocardial infarction 194
- N**
- naïve Bayes classifiers (NB) 345
- named entities 320
- Named entity recognition (NER) 316
- named entity recognition (NER) 290, 315
- name entity recognition 331

name mapping 175
narrative notes 357
narrative patient records 274
National Centre for Text Mining 7
natural-language processing 346
Natural Language Processing (NLP) 256, 275
natural language processing (NLP)
 39, 224, 225, 316
natural language processing (NLP) in biomedicine 316
natural language processing (NLP) techniques 224
natural language processing systems 347
NCBI taxonomy 60
negation 362
NE recognition experiment 340
NER system 293
neural network (NN) architecture 255
neural network approach 255
NLPBA shared task 316
NLP in the hospital ward 371
NLP services for clinical work 363
NLP tasks 357
non-contextual problem 186
non-factoid QA (NFQA) 192
non-linear relevance feedback 255

O

OBO foundry 350
OHSUMED 45
OHSUMED collection 48
online information 190
online Rlims-P tool, development of 173
ontologies, building from texts 187
ontology alignment 82
ontology class 133
organism 152
organism function 210
orthologous genes 63
OSIRIS 289
OSIRIS approach 293
OSIRIS approach, perspectives 297
OSIRISv1.2, application of 297
OSIRISv1.2, evaluation of 296
OSIRISv1.2 system workflow 295

P

paragraph tagging 232
parameter setting 94
parser 358
parthenogenesis 349
pathologic function 210
patient case profiles 274
 case recommendations 280
 causation 280
 diagnosis 279
 diagnostic test 279
 exceptionality 280
 exclusion 280
 medical test 279
 patient demographics 280
 patient health profile 280
 referral 279
 sign 279
 symptom 279
 treatment 279
patient case profiles, extracting 273
patient demographics, definition 280
patient health profile, definition 280
pattern matching 167
patterns 167
performance bottleneck 215
performance measures 94
PGN 62, 63
phosphorylation annotation ranking 174
phrasal level 360
phrase processing 360
physiology 349
plagiarism detection 222
PMID mapping 175
PMM analyzer with language engineering 330
PMM map 326
pPMM map generation scenario 327
PMM process architecture 331
point-of-care clinician 364
polarity classification 208
polarity classification, performance bottleneck
 in 213
polysemous entries 61
polysemy, reducing 65
polysemy indicators, set of 65
POSEDIT 180
PoS tagging 198

precision of alignments 94
 preprocessing 198, 281
 problem-solving method extraction 334
 problems-solving map extraction 325
 problem solving maps 325
 procedural tags 333
 processing, stages of 366
 ProMED-PLUS 327
 propagation rules 230
 proposed alignment models 94
 proposed architectural approach 259
 protein 152
 protein entity mapping 175
 protein interactions 304
 protein localisation 304
 protein phosphorylation 163
 protein posttranslational modification (PTM)
 163
 proteins 304
 PROTO-AGENT 194
 PROTO-PATIENT 194
 PubChem 61

Q

QA system, components of 196
 query drift 70
 query expansion (QE) 68
 query expansion, document-analysis for 68
 query expansion, using terminological resources 69
 query refinement 70
 query reformulation 67
 query reformulation, terminological resources
 for 67
 question-answering (QA) systems 191
 question answering (QA) 190
 question answering, clinical literature 190

R

recall of alignments 94
 recursive similarity measure adaptation 264
 reduced gradient method 266
 redundancy 61
 referral, definition 279
 relationships 304

relevance feedback 255, 256, 257, 258, 259,
 260, 261, 262, 263, 266, 268, 269,
 270, 271
 relevance feedback algorithms 258
 relevant information feedback 256
 researcher physician 364
 retrieval performance, objective evaluation
 criteria for 268
 rice and human diseases 325
 rice related information 329
 RLIMS-P 163
 RLIMS-P extraction system 165
 RLIMS-P system 165
 RLIMS-P text-mining 174
 rules, lattice of 111
 rules, rewriting 106, 109

S

search engines 2
 section tagging 232
 SEEK-CLIPS 347
 Seekbio 344
 SEEKbio, location hierarchies 348
 SEEKbio graphs editor 345
 SEEKbio tool 345
 segmentation 86
 semantic class analysis 194
 semantic classes, cores of 196
 semantic classification 166
 semantic description 87
 semantic indexing 71
 semantic indexing, information extraction for
 71
 semantic interpreter 358
 semantic lexicon 124
 semantic relation mining 28
 semantic roles 194
 semantic tag set, justification of the need 277
 semantic types 210
 semi-supervised model 204
 sense disambiguation 126
 sense disambiguation algorithm 127
 sentence classification system 273
 sentence classification task 280
 sequence-labeling task 318
 sequence variants of genes 289

Index

sequence variants of genes, identification of 289
shallow parsing 165
sign, definition 279
similarity measure, minimal modification of 265
similarity measures 94, 260
similarity measure type estimation 263
similarity measure type modeling 262
Single Nucleotide Polymorphisms (SNPs) 290
site rule (SR) 169
slots 2
SNAP/SPAN BFO ontology 348
SNAP entities 350
SNOMED-CT (Systematized Nomenclature of Medicine - Clinical Terms) 60
SNPs 289
sophisticated IR systems 320
sorting 186
SPAN entities 350
sparse training data, enrichment of 153
spatial relations, retrieval of 344
Specialist lexicon 61
speech tagging 360
steroids 197
stop-word filter 86
string matching 24
stroke 197
structural analyzer 358
substance ontologies 135
support vector machine (SVM) 281
support vector machine (SVM) classifier 273
support vector machine algorithm 78
support vector machines (SVMs) 92, 205, 345
SVM, application of 154
SVM-based alignment model 99
SVM-S7 model 101
SVM algorithm 92
SVM classifier 78
SVMs 204
SVMs, basic principles of 92
Swedish MeSH® 11
Sydney hospital, intensive health services 357
symptom, definition 279
synsets 152
syntactic labelling 86

Systematized Nomenclature of Medicine (SNOMED CT) 276
system biology 344

T

tagging 177
template 2
terminologies, thesauri, ontologies and classifications (TTOCs) 363
terminology integration 79
text, recall and abstraction 64
text categorization (TC) 39
text categorization, evaluating 49
text descriptors 233
text linguistics 224
text mining 1, 6, 224, 304
text mining, biomedicine 1
text mining, discourse processing for 222
text mining system 290
text parsing 228
text processing 22
text representation 230, 232
textual data mining experiments 237
textual knowledge 1
thrombolysis 194
training data, enrichment of 155
translation experiments 113
translation research, analysing clinical notes 357
translation technique 109
treatment, definition 279
treatment frame 194
treatment procedures 279
tree kernel 282
treemap visualization 31

U

UMLS 42
UMLS (Unified Medical Language System) 38
UMLS, expanding terms with 46
UMLS elements 42
UMLS knowledge sources 43
uncurated protein annotations 301
uncurated protein annotations, verification of 301

Unified Medical Language System (UMLS)
60, 210, 275

unigrams 208

UNITEX system 180

V

valuable knowledge 325

vector representation 261, 270

vertical polysemy 63

VIDAL 84

VIDAL dictionary 81

W

Ward Rounds Information System (WRIS) 371

Wikipedia 85, 133

Wikipedia translations, biomedical ontologies
with 133

WordNet 40, 124

WordNet, biomedical ontologies with 126

WordNet, human anatomy ontology with 128

WordNet, molecular biology ontology with
131

word sense disambiguation 142

word sense disambiguation (WSD) 142, 362

WSD-problem 144

WSD in life science texts 148

WSD problem, approaches to solving 146

WSD problem, classification-based approach
149