| Paper | Year of publication | Dimension of Data Set (N of documents) | Pre-Processing | | | | | | | | NER Supervised Tools | | | | | | NER unsupervised tools | | | | | |
			Sentence splitting	Token	Lemma	Stemming	POS tagging	Chunk	Word frequency	Dependency parsing tree	SVM	CRF	NB	Dynamic Topic Modeling	MEM	Artificial Neural Networks	Clustering	LDA	PCA	Skip-Gram	CBOW	Experts
1) Extracting failure time data from industrial maintenance records using text mining	2016	11786	X	X					X		X		X									X
2) Developing a hybrid dictionary-based bio-entity recognition technique	2014	340		X		X	X															
3) Science foresight using life-cycle analysis, text mining and clustering A case study on natural ventilation	2017	92	X	X		X			X								X					X
4) Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis	2017	1560	X	X					X									X				X
5) A hybrid model for automatic identification of risk factors for heart disease	2015	1304	X	X		X	X		X			X	X	X								X
6) Big data driven co-occurring evidence discovery in chronic obstructive pulmonary disease patients	2017	64371 (summaries)	X	X			X		X													X
7) Revealing topics and their evolution in biomedical literature using Bio-DTM: a case study of ginseng	2017	5394	X	X		X			X					X								X
8) Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data	2016	151	X	X	X				X													
9) Evaluation of a rule-based method for epidemiological document classification towards the automation of systematic reviews	2017	51																				X
10) How to make the most of NE dictionaries in statistical NER	2008	2404 abstracts	X	x			X					X										
11) Effects of Semantic Features on Machine Learning-Based Drug Name Recognition Systems: Word Embeddings vs. Manually Constructed Dictionaries	2015	17,3 GB of texts	X	X			X	X				X					X			X		
12) Automatic information extraction from unstructured mammography reports using distributed semantics	2018	300300	X	X			X			X							X			X		X
13) Information extraction and knowledge graph construction from geoscience literature	2018	NA		X			X		X			X										
14) BCC-NER: bidirectional, contextual clues named entity tagger for gene/protein mention recognition	2017	20000 sentences	X	X	X		X	X	X			X							X			
15) Named Entity Mining from Click-Through Data Using Weakly Supervised Latent Dirichlet Allocation	2009	1000 queries	X															X				X
16) Named Entity Recognition for Automated Test Case Generation	2014	139	X				X								X							
17) Entity recognition in the biomedical domain using a hybrid approach	2017	97	X	X		X	X					X				X						
18) Recurrent neural networks with specialized word embeddings for health domain named-entity recognition	2017	426					X					X				X						
19) Latent semantics in Named Entity Recognition	2014	NA				X												X				
20) A novel approach for entity resolution in scientific documents using context graphs	2017	2000 abstracts	X	X					X		X											
21) Hybrid ontology-learning materials engineering system for pharmaceutical products: Multi-label entity recognition and concept detection	2017	1730 abstracts	X	X			X				X						X					X
22) PKDE4J: Entity and relation extraction for public knowledge discovery	2015	2384	X	X	X		X	X	X	X												X
23) Deriving technology intelligence from patents: Preposition-based semantic analysis	2018	2838	X				X		X	X							X					