Description of title sentence attributes used in classification

James M. Eales, George Demetriou, Robert Stevens

1 Attribute details

Table 1 lists some of the attributes calculated for each article title, the remaining attributes are described in section 1.1.

1.1 Additional attributes

1.1.1 Typed dependencies

We also include a set of values which provide the proportion of dependency links between phrases in the title for each of the standard Stanford parser dependency types (see http://nlp.stanford.edu/software/ dependencies_manual.pdf). The names of these attributes follow the pattern 'prop' short dependency name (first letter capitalised) 'Dep'. Examples are 'propNnDep', 'propDobjDep' and 'propPartmodDep'.

1.1.2 Part-of-speech tag n-grams (POS-grams)

The sequence of part-of-speech tags, from each title, are first shortened to a base form (e.g. NNP becomes NN and VBZ becomes VB) and then split into bigrams and trigrams and these are screened for the presence of 717 unique (shortened) bi- and trigrams that occur commonly in titles. Therefore the values for the POS-gram attributes are numerical integers that show the number of times a given POS-gram appears in the title.

good TitleOrNot Finominal (Class) ("good", "bad") porTagOfficit Token Nominal Part-of-speech tag for first token porTagOfficit Token Nominal Part-of-speech tag for first token porTagOfficit Token Nominal Does the title contain a colon containsColon Binominal Does the title contain a colon containsQuestionMark Binominal Does the title contain a newformark containsQuestionMark Binominal Does the title contain a newformark containsUpten Binominal Does the title contain a popten lengthInCharacters Numerical Number of tokens (similar to number of words) in the title numberOfSentences Numerical Number of stotences in title, determined by OpenNLP sentences piltter numberOfInetences Numerical Number of unique proteins found by whatizit pipeline "whatizitStwissport" numberOfInetences Numerical Number of unique giventities numberOfInetences Numerical Number of unique seases found by whatizit pipeline "whatizitStwissport" numberOfInetences Numerical Number of unique seases found by unimat cell types Mol Statizet biasead* numberOfCellTypeE	Attribute name	Attribute type	Description
posTagOfFirstToken Nominal Part-of-speech tag for first token posTagOfFicenonToken Nominal Part-of-speech tag for scond token posTagOfFindTloken Nominal Part-of-speech tag for scond token containsVerb Binominal Does the title contain a verb containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Binominal Does the title contain a typhen lengthInCharacters Numerical Number of tokens (similar to number of words) in the title determined by OpenNLP scatteres pitter numberOfSentences Numerical Number of settence splitter numberOfProteinEntities Numerical Number of unique proteins found by four different named entity recognition tools numberOfDiseaseEntities Numerical Number of unique small molecules found by whatizit pipeline 'whatizitDisease' found by unimar cell types Web Service proportionOfNours Numerical Proportion of title tokens that are nouns (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of title tokens that are hedge words (e.g., 'may', 'potentiat', 'could') proportionOfAdjectiveChunks Numerical Proportion of adjective pirases in title (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective pirases in title (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective pirases in title (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective pirases in title (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective pirases in title (determined by OpenNLP POS tagger) proportionO	goodTitleOrNot	Binominal (Class)	('good', 'bad')
posTagOffScondToken Nominal Part-of-speech tag for third token containsColon Binominal Does the title contain a verb containsColon Binominal Does the title contain a verb containsQuestionMark Binominal Does the title contain a verb containsQuestionMark Binominal Does the title contain a newsion mark containsQuestionMark Binominal Does the title contain a newsion mark containsQuestionMark Binominal Does the title contain a newsion mark lengthInCharacters Numerical Number of tokens (similar to number of words) in the title lengthInCharacters Numerical Number of tokens (similar to number of unique spatians found by whatizit pipeline "whatizitStwissport" numberOfSmallMoleculeEntities Numerical Number of unique small molecules found by whatizit pipeline "whatizitDisease" numberOfEllTypeEntities Numerical Number of unique exit are verbs (determined by OpenNLP POS tagger) proportionOfNours Numerical Proportion of title tokens that are nouns (determined by OpenNLP POS tagger) proportionOfAdjectives Numerical Proportion of title tokens that are nouns (determined by OpenNLP POS tagger) proportionOfAdjectiv	posTagOfFirstToken	Nominal	Part-of-speech tag for first token
posTagOfThirdToken Nominal Part-of-speech tag for third token containsColon Binominal Does the title contain a veb containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Number of tokens (similar to number of words) in the title lengthInCharacters Numerical Number of suences splitter numberOfRoneese Numerical Total number of anared entities found by four different named entity recognition tools numberOfSnallMoleculeEntities Numerical Number of unique gentence splitter numberOfDiseaseEntities Numerical Number of unique seall molecules found by whatizit pipeline "whatiztDisease" proportionOfNouns Numerical Number of title tokens that are ours (determined by OpenNLP Postager) proportionOfAdjectives Numerical Proportion of title tokens that are hedge words (e.g. 'may', 'potential', 'could') proportionOfAdjectiveChunks Numerical Proportion of	posTagOfSecondToken	Nominal	Part-of-speech tag for second token
containsVerb Binominal Does the title contain a verb containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Binominal Does the title contain a question mark containsItyphen Binominal Does the title contain a question mark lengthInTokens Numerical Number of text characters in the title numberOfSentences Numerical Number of sentences in title, determined by OpenNLP sentence splitter numberOfNamedEatities Numerical Number of unique proteins found by whatizit pipcline "whatizitSyspord" numberOfSmallMoleculeEntities Numerical Number of unique small molecules found by whatizit pipeline "whatizitChebiDict" numberOfCellTypeEntities Numerical Number of unique grotes found by whatizit pipcline "whatizitDisease" proportionOfNouns Numerical Number of unique types found by uninan cell types Web Service proportionOfAlgectives Numerical Proportion of title tokens that are nouns (de termined by OpenNLP POS tagger) proportionOfAdjectives Numerical Proportion of title tokens that are hedge words (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of tithe tokens in tithe (deter- mined by OpenNLP POS tagger) <	posTagOfThirdToken	Nominal	Part-of-speech tag for third token
containsQuestionMark Binominal Does the title contain a question mark containsQuestionMark Binominal Does the title contain a question mark containsItyphen Binominal Does the title contain a question mark lengthInTokens Number of tokens (similar to number of words) in the title lengthInCharacters Numerical Number of text characters in the title numberOISentences Numerical Number of sentences in title, determined by OpenNLP sentence splitter numberOIFrotcinEntities Numerical Number of unique proteins found by four different named entity recognition tools numberOISmallMoleculeEntities Numerical Number of unique seases found by whatizit pipeline 'whatizitChebiDict' numberOfCellTypeEntities Numerical Proportion of title tokens that are nouns (determined by OpenNLP POS tagger) proportionOfMons Numerical Proportion of title tokens that are adjectives (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective phrases in title (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective phrases in title (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proporti	containsVerb	Binominal	Does the title contain a verb
containsQuestionMark Binominal Does the title contain a question mark containsHyphen Binominal Does the title contain a hyphen lengthInCharacters Numerical Number of tokens (similar to number of words) in the title numberOfSentences Numerical Number of sentences in title, determined by OpenNLP sentence splitter numberOfProteinEntities Numerical Number of mance entities found by whatizit pipeline 'whatizitChebiDict' numberOfDiseaseEntities Numerical Number of unique proteins found by whatizit pipeline 'whatizitChebiDict' numberOfDiseaseEntities Numerical Number of unique conteins found by uninan cell types found by uninan cell types found by uninan cell types Web Service proportionOfNouns Numerical Number of unique cell types found by uninan cell types Web Service proportionOfAdjectives Numerical Proportion of title tokens that are nouns (determined by OpenNLP POS tagger) proportionOfAdjectives Numerical Proportion of adjectic phrases in title (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjectic phrases in title (determined by OpenNLP POS tagger) proportionOfAdverbChunks Numerical Proportion of adjectic phrases in title (determined by OpenNLP chunker)	containsColon	Binominal	Does the title contain a colon
containsHyphen Binominal Does the title contain a hyphen lengthInTokens Numerical Number of tokens (similar to number of words) in the title numberOfSentences Numerical Number of sentences in title, determined by OpenNLP sentence splitter numberOfSentences Numerical Number of unique proteins found by four different named entity recognition tools numberOfProteinEntities Numerical Number of unique genesses found by whatizit pipeline 'whatizitChebiDict' numberOfDiseaseEntities Numerical Number of unique desses found by wintzit pipeline 'whatizitChebiDict' numberOfEllTypeEntities Numerical Number of unique desses found by wintzit pipeline 'whatizitChebiDict' numberOfEllTypeEntities Numerical Proportion of title tokens that are nouns (determined by OpenNLP POS tagger) proportionOfIverbs Numerical Proportion of title tokens that are adjectives (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of title tokens that are adjectives (determined by OpenNLP chunker) proportionOfAdjectiveChunks Numerical Proportion of adjective phrases in title (determined by OpenNLP chunker) proportionOfAdjectiveChunks Numerical Proportion of adjective phrases in title (determined by OpenNLP chun	containsQuestionMark	Binominal	Does the title contain a question mark
lengthInTokens Numerical Number of tokens (similar to number of words) in the title lengthInCharacters Numerical Number of text characters in the title numberOfSentences Numerical Number of sentences in title, determined by OpenNLP sentence splitter numberOfProteinEntities Numerical Total number of nanced entity recognition tools numberOfProteinEntities Numerical Number of unique proteins found by whatizit pipeline 'whatizitSwissprot' numberOfSmallMoleculeEntities Numerical Number of unique sensal molecules found by whatizit pipeline 'whatizitChebiDict' numberOfDiseaseEntities Numerical Number of unique sensal molecules found by whatizit pipeline 'whatizitSwissprot' numberOfEcllTypeEntities Numerical Number of unique coll spasses of und by whatizit pipeline 'whatizitSwissprot' proportionOfNouns Numerical Number of unique coll spasses of und by uniman cell types Web Service proportionOfAdjectives Numerical Proportion of title tokens that are nouns (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of ditle tokens that are hedge words (e.g. 'may', 'potential', 'could') proportionOfAdjectiveChunks Numerical Proportion of num phrases in title (determinined by OpenNLP chun	containsHyphen	Binominal	Does the title contain a hyphen
Image: constraint of the second sec	lengthInTokens	Numerical	Number of tokens (similar to number of
lengthfnCharacters Numerical Number of text characters in the title numberOfSentences Numerical Number of sentences in title, determined by OpenNLP sentence splitter numberOfNamedEntities Numerical Total number of named entities found by four different named entities found by whatizit pipeline 'whatizitSwissprot' numberOfProteinEntities Numerical Number of unique proteins found by whatizit pipeline 'whatizitSwissprot' numberOfSmallMoleculeEntities Numerical Number of unique sensalt molecules found by whatizit pipeline 'whatizitChebiDict' numberOfDiseaseEntities Numerical Number of unique diseases found by whatizit pipeline 'whatizitDisease' numberOfOfNouns Numerical Number of unique cell types found by uniman cell types Web Service proportionOfVerbs Numerical Proportion of title tokens that are enouns (de- termined by OpenNLP POS tagger) proportionOfAdjectiveS Numerical Proportion of title tokens that are explex (de- termined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective phrases in title (deter- mined by OpenNLP chunker) proportionOfAdverbChunks Numerical Proportion of nor phrases in title (deter- mined by OpenNLP chunker) proportionOfPrepositionChunks Numerical <td< td=""><td></td><td></td><td>words) in the title</td></td<>			words) in the title
numberOfSentences Numerical Number of sentences in title, determined by OpenNLP sentence splitter numberOfNamedEntities Numerical Total number of named entities found by four different named entity recognition tools numberOfProteinEntities Numerical Number of unique proteins found by whatizit pipeline 'whatizitSwissprot' numberOfSmallMoleculeEntities Numerical Number of unique small molecules found by whatizit pipeline 'whatizitOhebiDict' numberOfCellTypeEntities Numerical Number of unique cilcases found by uninan cell types Web Service proportionOfNouns Numerical Proportion of title tokens that are nouns (de termined by OpenNLP POS tagger) proportionOfVerbs Numerical Proportion of title tokens that are verbs (de termined by OpenNLP POS tagger) proportionOfAdjectives Numerical Proportion of title tokens that are dectives (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective plrases in title (deter- mined by OpenNLP chunker) proportionOfAdjectiveChunks Numerical Proportion of adjective plrases in title (deter- mined by OpenNLP chunker) proportionOfPrepositionChunks Numerical Proportion of numerical in the (deter- mined by OpenNLP chunker) proportionOfPrepositionChunks Numeric	lengthInCharacters	Numerical	Number of text characters in the title
numberOfNamedEntitiesNumericalOpenNLP sentence splitternumberOfProteinEntitiesNumericalTotal number of named entity recognition toolsnumberOfProteinEntitiesNumericalNumber of unique proteins found by whatizit pipeline 'whatizitSwissprot'numberOfDiscaseEntitiesNumericalNumber of unique small molecules found by whatizit pipeline 'whatizitChebiDict'numberOfDiscaseEntitiesNumericalNumber of unique discases found by whatizit pipeline 'whatizitChebiDict'numberOfCellTypeEntitiesNumericalNumber of unique cell types found by uniman cell types Web ServiceproportionOfNounsNumericalProportion of title tokens that are nouns (de- termined by OpenNLP POS tagger)proportionOfVerbsNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfAdjectiveChunksNumericalProportion of title tokens that are hedge words (e.g. 'may', 'potential', 'could')proportionOfAdverbChunksNumericalProportion of adverb phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of rule tokens (stemmed) found in the or- dered list of title tokens (stemmed lound in the or- dered list of title tokens (stemmed lound in	numberOfSentences	Numerical	Number of sentences in title, determined by
numberOfNamedEntities Numerical Total number of named entities found by four different named entity recognition tools numberOfProteinEntities Numerical Number of unique proteins found by whatizit pipeline 'whatizitChebiDict' numberOfSmallMoleculeEntities Numerical Number of unique diseases found by whatizit pipeline 'whatizitChebiDict' numberOfCllTypeEntities Numerical Number of unique cll types found by whatizit pipeline 'whatizitDisease' numberOfCllTypeEntities Numerical Number of title tokens that are nouns (determined by OpenNLP POS tagger) proportionOfNouns Numerical Proportion of title tokens that are verbs (determined by OpenNLP POS tagger) proportionOfAdjectives Numerical Proportion of title tokens that are verbs (determined by OpenNLP POS tagger) proportionOfIIedges Numerical Proportion of adjective phrases in title (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective phrases in title (determined by OpenNLP chunker) proportionOfAdverbChunks Numerical Proportion of adjective phrases in title (determined by OpenNLP chunker) proportionOfAverbChunks Numerical Proportion of proportion of phrases in title (determined by OpenNLP chunker) proportionOfVerbChunks Numerical </td <td></td> <td></td> <td>OpenNLP sentence splitter</td>			OpenNLP sentence splitter
numberOfProteinEntitiesNumericalNumber of unique proteins found by whatizit pipeline 'whatizitSwissprot'numberOfSmallMoleculeEntitiesNumber and the statist pipeline 'whatizitChebiDict'numberOfDiseaseEntitiesNumericalNumber of unique small molecules found by whatizitThebiDict'numberOfCellTypeEntitiesNumericalNumber of unique cell types found by uniman cell types Web ServiceproportionOfNounsNumericalProportion of title tokens that are nouns (de- termined by OpenNLP POS tagger)proportionOfVerbsNumericalProportion of title tokens that are verbs (de- termined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are dijectives (determined by OpenNLP POS tagger)proportionOfAdjectiveChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfAdjectiveChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion o	numberOfNamedEntities	Numerical	Total number of named entities found by four
numberOfProteinEntities Numerical Number of unique proteins found by whatizit pipeline 'whatizitSwissprot' numberOfSmallMoleculeEntities Numerical Number of unique small molecules found by whatizit pipeline 'whatizitChebiDict' numberOfDiseaseEntities Numerical Number of unique diseases found by whatizit pipeline 'whatizitDisease' numberOfCellTypeEntities Numerical Number of unique cell types found by uniman cell types Web Service proportionOfNouns Numerical Proportion of title tokens that are nouns (de- termined by OpenNLP POS tagger) proportionOfVerbs Numerical Proportion of title tokens that are verbs (de- termined by OpenNLP POS tagger) proportionOfHedges Numerical Proportion of title tokens that are adjectives (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective phrases in title (deter- mined by OpenNLP Pounker) proportionOfAdverbChunks Numerical Proportion of adjective phrases in title (deter- mined by OpenNLP chunker) proportionOfFvepositionChunks Numerical Proportion of preposition phrases in title (deter- mined by OpenNLP chunker) proportionOfVerbChunks Numerical Proportion of preposition phrases in title (deter- mined by OpenNLP chunker) proportionOfVerbChunks Nume			different named entity recognition tools
pipelinewhatizitSvissprot'numberOfSmallMoleculeEntitiesNumericalNumber of unique small molecules found by whatizit pipeline 'whatizitChebiDict'numberOfDiseaseEntitiesNumericalNumber of unique cell seases found by whatizit pipeline 'whatizitDisease'numberOfCellTypeEntitiesNumericalNumber of unique cell seases found by uniman cell types Web ServiceproportionOfNounsNumericalProportion of title tokens that are nouns (de termined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are verbs (de termined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are verbs (de termined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are hedge words (e.g. 'may', 'potential', 'could')proportionOfAdjectiveChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProport	numberOfProteinEntities	Numerical	Number of unique proteins found by whatizit
numberOISmallMoleculeEntities Numerical Number of unique small molecules found by whatizit pipeline 'whatizitChebiDict' numberOfDiseaseEntities Numerical Number of unique diseases found by whatizit pipeline 'whatizitDisease' numberOfCellTypeEntities Numerical Number of unique cell types found by uniman cell types Web Service proportionOfNouns Numerical Proportion of title tokens that are nouns (de- termined by OpenNLP POS tagger) proportionOfAdjectives Numerical Proportion of title tokens that are verbs (de- termined by OpenNLP POS tagger) proportionOfAdjectives Numerical Proportion of title tokens that are hedge words (determined by OpenNLP POS tagger) proportionOfAdjectiveChunks Numerical Proportion of adjective phrases in title (deter- mined by OpenNLP chunker) proportionOfAdverbChunks Numerical Proportion of adverb phrases in title (deter- mined by OpenNLP chunker) proportionOfNounChunks Numerical Proportion of preposition phrases in title (deter- mined by OpenNLP chunker) proportionOfVerbChunks Numerical Proportion of verb phrases in title (deter- mined by OpenNLP chunker) proportionOfVerbChunks Numerical Proportion of verb phrases in title (deter- mined by OpenNLP chunker) proportionOfVerbChunks Numerical			pipeline 'whatizitSwissprot'
numberOfDiseaseEntitiesNumericalNumericalNumber of unique diseases found by whatizit pipeline 'whatizitDisease'numberOfCellTypeEntitiesNumericalNumber of unique cell types found by uniman cell types Web ServiceproportionOfNounsNumericalProportion of title tokens that are nouns (de- termined by OpenNLP POS tagger)proportionOfVerbsNumericalProportion of title tokens that are verbs (de- termined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfHedgesNumericalProportion of title tokens that are hedge words (e.g. 'may', 'potential', 'could')proportionOfAdjectiveChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of nadverb phrases in title (deter- mined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of nor phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of or phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of or phrases in title (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe first verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in th	numberOfSmallMoleculeEntities	Numerical	Number of unique small molecules found by
numberOfDiseaseEntitiesNumericalNumericalNumber of unque diseases found by whatizit pipeline 'whatizitDisease'numberOfCellTypeEntitiesNumber of unque cell types found by uniman cell types Web ServiceproportionOfNounsNumericalProportion of title tokens that are nouns (de- termined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfAdjectiveChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe structure (stilks between phrases) in the title, value derived using Stan- for			whatizit pipeline 'whatizitChebiDict'
numberOfCellTypeEntitiesNumericalNumericalNumber of unique cell types found by uniman cell types Web ServiceproportionOfNounsNumericalProportion of tille tokens that are nouns (de- termined by OpenNLP POS tagger)proportionOfVerbsNumericalProportion of tille tokens that are verbs (de- termined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of tille tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfAdjectiveChunksNumericalProportion of tille tokens that are hedge words (e.g. 'may', 'potential', 'could')proportionOfAdverbChunksNumericalProportion of adjective phrases in tille (deter- mined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of adverb phrases in tille (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in tille (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in tille (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in tille (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in tille (deter- mined by OpenNLP chunker)verbTwoNominalThe first verb (stemmed) found in the ordered list of tille tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of tille tokens (stemmed) using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumeric <t< td=""><td>numberOfDiseaseEntities</td><td>Numerical</td><td>Number of unique diseases found by whatizit</td></t<>	numberOfDiseaseEntities	Numerical	Number of unique diseases found by whatizit
number of Unique cell types found by unmanical cell types Web ServiceNumanical cell types Web Serviceproportion OfNounsNumericalProportion of title tokens that are nouns (de- termined by OpenNLP POS tagger)proportion OfVerbsNumericalProportion of title tokens that are verbs (de- termined by OpenNLP POS tagger)proportion OfAdjectivesNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportion OfHedgesNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportion OfAdjectiveChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumeric<		NT • 1	pipeline 'whatizitDisease'
proportionOfNounsNumericalProportion of title tokens that are nouns (de- termined by OpenNLP POS tagger)proportionOfVerbsNumericalProportion of title tokens that are verbs (de- termined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfAdjectiveChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of adverb phrases in title (deter- mined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumericNumericNumericNumericNumericNumericNumericNumeric	numberOfCellTypeEntities	Numerical	Number of unique cell types found by uniman
proportion of INOLINS Numerical Proportion of tube tokens that are nouns (determined by OpenNLP POS tagger) proportion Of Verbs Numerical Proportion of title tokens that are verbs (determined by OpenNLP POS tagger) proportion OfAdjectives Numerical Proportion of title tokens that are adjectives (determined by OpenNLP POS tagger) proportionOfAdjectives Numerical Proportion of title tokens that are hedge words (e.g. 'may', 'potential', 'could') proportionOfAdjectiveChunks Numerical Proportion of adjective plarases in title (determined by OpenNLP chunker) proportionOfAdverbChunks Numerical Proportion of adjective plarases in title (determined by OpenNLP chunker) proportionOfAdverbChunks Numerical Proportion of noun plarases in title (determined by OpenNLP chunker) proportionOfPrepositionChunks Numerical Proportion of adjective plarases in title (determined by OpenNLP chunker) proportionOfVerbChunks Numerical Proportion of preposition phrases in title (determined by OpenNLP chunker) verbOne Nominal The first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser) verbTwo Nominal The third verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)			cell types web Service
proportionOfVerbsNumericalProportion of title tokens that are verbs (de- termined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfHedgesNumericalProportion of title tokens that are hedge words (e.g. 'may', 'potential', 'could')proportionOfAdjectiveChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of num phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbTwoNominalThe third verb (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumericNumericNumericNumeric of typed dependencies (links between phrases) in the title, value derived using Stan- ford ParserproportionofTypedDependenciesNumericNumericNumber of typed dependenciesNumericNumericproper threeNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in	proportionOfNouns	Numerical	Proportion of title tokens that are nouns (de-
proportionOfVerbsNumericalProportion of the tokens that are verbs (determined by OpenNLP POS tagger)proportionOfAdjectivesNumericalProportion of title tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfHedgesNumericalProportion of title tokens that are hedge words (e.g. 'may', 'potential', 'could')proportionOfAdjectiveChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of adverb phrases in title (deter- mined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	nnon ontion Of Vonba	Numerical	Dependence of title telepis that are worked (de
proportionOfAdjectivesNumericalProportion of tile tokens that are adjectives (determined by OpenNLP POS tagger)proportionOfHedgesNumericalProportion of tile tokens that are hedge words (e.g. 'may', 'potential', 'could')proportionOfAdjectiveChunksNumericalProportion of adjective phrases in tile (determined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of adjective phrases in tile (determined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of noun phrases in tile (determined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of noun phrases in tile (determined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in tile (determined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in tile (determined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of tile tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe third verb (stemmed) found in the or- dered list of tile tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in the tile, calculated by Stanford parser	proportionOfverbs	Numericai	termined by OpenNLP POS tagger)
proportionOfHedgesNumericalProportion of title tokens that are hedge words (determined by OpenNLP POS tagger)proportionOfAdjectiveChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of adjective phrases in title (deter- mined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of noun phrases in title (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of preposition phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumericNumericaNumericNumericephraseStructureMaxDepthNumericNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	proportion Of A directives	Numerical	Proportion of title tokens that are adjectives
proportionOfHedgesNumericalProportion of tile tokens that are hedge words (e.g. 'may', 'potential', 'could')proportionOfAdjectiveChunksNumericalProportion of adjective phrases in tile (deter- mined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of adverb phrases in tile (deter- mined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of noun phrases in tile (deter- mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of preposition phrases in tile (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in tile (deter- mined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in tile (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of tile tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of tile tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericNumericMaximal depth of phrase structure in the tile, calculated by Stanford parser	proportionOnAujectives	Numericai	(determined by OpenNLP POS tagger)
proportion of Adjective Chunks Numerical Proportion of adjective phrases in title (determined by OpenNLP chunker) proportion Of Adverb Chunks Numerical Proportion of adverb phrases in title (determined by OpenNLP chunker) proportion Of NounChunks Numerical Proportion of noun phrases in title (determined by OpenNLP chunker) proportion OfPreposition Chunks Numerical Proportion of noun phrases in title (determined by OpenNLP chunker) proportion OfPreposition Chunks Numerical Proportion of preposition phrases in title (determined by OpenNLP chunker) proportion OfVerbChunks Numerical Proportion of verb phrases in title (determined by OpenNLP chunker) verbOne Nominal The first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser) verbTwo Nominal The third verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser) verbThree Nominal The third verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser) numberOfTypedDependencies Numeric Number of typed dependencies (links between phrases) in the title, value derived using Stanford Parser phraseStructureMaxDepth Numeric Number of typed dependencies (links between phrases) in the title, value derived using Stanford parser <td>proportionOfHedges</td> <td>Numerical</td> <td>Proportion of title tokens that are hedge words</td>	proportionOfHedges	Numerical	Proportion of title tokens that are hedge words
proportionOfAdjectiveChunksNumericalProportion of adjective phrases in title (determined by OpenNLP chunker)proportionOfAdverbChunksNumericalProportion of adverb phrases in title (determined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of noun phrases in title (determined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of preposition phrases in title (determined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (determined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (determined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	proportionomotificages	1 (difforficult	(e.g. 'may', 'potential', 'could')
Image: Propertion of Adverb ChunksNumericalProportion of adverb phrases in title (determined by OpenNLP chunker)proportion OfNounChunksNumericalProportion of noun phrases in title (determined by OpenNLP chunker)proportion OfPrepositionChunksNumericalProportion of preposition phrases in title (determined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of preposition phrases in title (determined by OpenNLP chunker)verbOneNominalProportion of verb phrases in title (determined by OpenNLP chunker)verbTwoNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	proportionOfAdjectiveChunks	Numerical	Proportion of adjective phrases in title (deter-
proportionOfAdverbChunksNumericalProportion of adverb phrases in title (determined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of noun phrases in title (determined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of preposition phrases in title (determined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (determined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		mined by OpenNLP chunker)
mined by OpenNLP chunker)proportionOfNounChunksNumericalProportion of noun phrases in title (determined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of preposition phrases in title (determined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (determined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	proportionOfAdverbChunks	Numerical	Proportion of adverb phrases in title (deter-
proportionOfNounChunksNumericalProportion of noun phrases in title (determined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of preposition phrases in title (determined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (determined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser			mined by OpenNLP chunker)
mined by OpenNLP chunker)proportionOfPrepositionChunksNumericalProportion of preposition phrases in title (de- termined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	proportionOfNounChunks	Numerical	Proportion of noun phrases in title (deter-
proportionOfPrepositionChunksNumericalProportion of preposition phrases in title (de- termined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumber of typed dependencies (links between phrases) in the title, value derived using Stan- ford parserphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser			mined by OpenNLP chunker)
mathematicaltermined by OpenNLP chunker)proportionOfVerbChunksNumericalProportion of verb phrases in title (determined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford Parser morphological analyser)numberOfTypedDependenciesNumericNumber of typed dependencies (links between phrases) in the title, value derived using Stanford parserphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	proportionOfPrepositionChunks	Numerical	Proportion of preposition phrases in title (de-
proportionOfVerbChunksNumericalProportion of verb phrases in title (deter- mined by OpenNLP chunker)verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser			termined by OpenNLP chunker)
werbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumber of typed dependencies (links between phrases) in the title, value derived using Stan- ford parserphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	proportionOfVerbChunks	Numerical	Proportion of verb phrases in title (deter-
verbOneNominalThe first verb (stemmed) found in the ordered list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumber of typed dependencies (links between phrases) in the title, value derived using Stan- ford parserphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser			mined by OpenNLP chunker)
list of title tokens (stemmed using Stanford parser morphological analyser)verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumericphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	verbOne	Nominal	The first verb (stemmed) found in the ordered
verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumber of typed dependencies (links between phrases) in the title, value derived using Stan- ford parserphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser			list of title tokens (stemmed using Stanford
verbTwoNominalThe second verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumber of typed dependencies (links between phrases) in the title, value derived using Stan- ford parserphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser			parser morphological analyser)
dered hst of title tokens (stemmed using Stan- ford Parser morphological analyser)verbThreeNominalThe third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser)numberOfTypedDependenciesNumericNumber of typed dependencies (links between phrases) in the title, value derived using Stan- ford parserphraseStructureMaxDepthNumericMaximal depth of phrase structure in the title, calculated by Stanford parser	verbTwo	Nominal	The second verb (stemmed) found in the or-
verbThree Nominal The third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser) numberOfTypedDependencies Numeric Number of typed dependencies (links between phrases) in the title, value derived using Stan- ford parser phraseStructureMaxDepth Numeric Maximal depth of phrase structure in the title, calculated by Stanford parser			dered list of title tokens (stemmed using Stan-
verbinge Nominal The third verb (stemmed) found in the or- dered list of title tokens (stemmed using Stan- ford Parser morphological analyser) numberOfTypedDependencies Numeric Number of typed dependencies (links between phrases) in the title, value derived using Stan- ford parser phraseStructureMaxDepth Numeric Maximal depth of phrase structure in the title, calculated by Stanford parser		NT1	The third event (to 1) for the th
numberOfTypedDependencies Numeric Number of typed dependencies (links between phrases) in the title, value derived using Stanford parser phraseStructureMaxDepth Numeric Maximal depth of phrase structure in the title, calculated by Stanford parser	verbinree	Nominal	I ne third verb (stemmed) found in the or-
numberOfTypedDependencies Numeric Number of typed dependencies (links between phrases) in the title, value derived using Stanford parser phraseStructureMaxDepth Numeric Maximal depth of phrase structure in the title, calculated by Stanford parser			ford Parson morphological analyzer)
number of typed dependencies Numeric phraseStructureMaxDepth Numeric Numeric Maximal depth of phrase structure in the title, calculated by Stanford parser	numberOfTupedDependencies	Numoria	Number of typed dependencies (links between
phrases) in the title, value derived using staff- ford parser phraseStructureMaxDepth Numeric Maximal depth of phrase structure in the title, calculated by Stanford parser	number of rypedDependencies		phrases) in the title, value derived using Stan
phraseStructureMaxDepth Numeric Maximal depth of phrase structure in the title, calculated by Stanford parser			ford parser
calculated by Stanford parser	phraseStructureMaxDepth	Numeric	Maximal depth of phrase structure in the title
	F		calculated by Stanford parser

Table 1: Table of article title attributes and their description