

The MIT Press

Journals

[Sign In / Register](#)
[Books](#)
[Journals](#)
[Digital](#)
[Resources](#)
[About](#)
[Contact](#)


Home | Computational Linguistics | List Article navigation
of Issues | Volume 39 , No. 3 | Text
Representations for Patent Classification



Text Representations for Patent Classification

Quarterly (March,
June, September,
December)

160pp. per issue

6 3/4 x 10

Founded: 1974

2018 Impact

Factor: 1.319

2018 Google

Scholar h5-index:
32

ISSN: 0891-2017

E-ISSN: 1530-9312

Eva D'hondt, Suzan Verberne,
Cornelis Koster and Lou
Boves

Posted Online August 06, 2013

https://doi.org/10.1162/COLI_a_00149

© 2013 Association for Computational Linguistics

Computational Linguistics
Volume 39 | Issue 3 | September 2013
p.755-775

Journal Resources

Editorial Info

Abstracting and
Indexing

Release Schedule

Advertising Info

Author Resources

Submission

Guidelines

Publication

Agreement

Author Reprints


 **Download Options** >

Abstract **Full Text** **Authors**

With the increasing rate of patent application filings, automated patent classification is of rising economic importance. This article investigates how patent classification can be improved by using different representations of the patent documents. Using the Linguistic Classification System (LCS), we compare the impact of adding statistical phrases (in the form of bigrams) and linguistic phrases (in two different dependency formats) to the standard

Reader Resources

Rights and Permissions
Most Read
Most Cited

More About Computational Linguistics 

Metrics 



16 Total

citations

5 Recent

citations

8.36 Field Citation

Ratio

n/a Relative

Citation Ratio

Open Access 




Computational Linguistics
Computational Linguistics is Open Access. All content is freely available in electronic format (Full text HTML, PDF, and PDF Plus) to readers across the globe. All articles are


bag-of-words text representation on a subset of 532,264 English abstracts from the CLEF-IP 2010 corpus. In contrast to previous findings on classification with phrases in the Reuters-21578 data set, for patent classification the addition of phrases results in significant improvements over the unigram baseline. The best results were achieved by combining all four representations, and the second best by combining unigrams and lemmatized bigrams. This article includes extensive analyses of the class models (a.k.a. class profiles) created by the classifiers in the LCS framework, to examine which types of phrases are most informative for patent classification. It appears that bigrams contribute most to improvements in classification accuracy. Similar experiments were performed on subsets of French and German abstracts to investigate the generalizability of these findings.


Forthcoming

Most Read

[See More](#)

 **Lexicon-Based Methods for Sentiment Analysis** (13965 times)
Maite Taboada et al.
Computational Linguistics
Volume: 37, Issue: 2, pp. 267-307

 **Computational Linguistics and Deep Learning** (10500 times)
Christopher D. Manning
Computational Linguistics
Volume: 41, Issue: 4, pp. 701-707


 **Near-Synonymy and Lexical Choice** (3653 times)
Philip Edmonds et al.
Computational Linguistics
Volume: 28, Issue: 2, pp. 105-144


(Note that the Most Read numbers are based on the number of full text downloads over the last 12 months.)


Most Cited

[See More](#)

published under a [CC BY-NC-ND 4.0 license](#). For more information on allowed uses, please view the [CC license](#).
[Support OA at MITP](#)

 **Lexicon-Based Methods for Sentiment Analysis** (436 times)
 Maite Taboada et al.
 Computational Linguistics
 Volume: 37, Issue: 2, pp. 267-307

 **A Systematic Comparison of Various Statistical Alignment Models** (174 times)
 Franz Josef Och et al.
 Computational Linguistics
 Volume: 29, Issue: 1, pp. 19-51

 **Opinion Word Expansion and Target Extraction through Double Propagation** (147 times)
 Guang Qiu et al.
 Computational Linguistics
 Volume: 37, Issue: 1, pp. 9-27

(Note that the Most Cited numbers are based on Crossref's [Cited-by service](#) and reflect citation information for the past 24 months.)

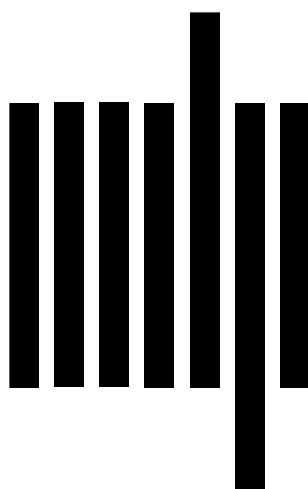
 **Download Options** >

Favorite  Sign up for Alerts 

Download Citation  RSS TOC 

RSS Citation  Submit your article

[Support OA at MITP](#) 



Journals

Terms & Conditions

Privacy Statement

Contact Us

Books

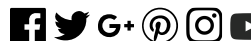
US

UK

Connect

One Rogers Street
 Cambridge MA
 02142-1209

Suite 2, 1 Duchess
 Street London,
 W1W 6AN, UK



© 2018 The MIT Press
 Technology Partner:
[Atypon Systems, Inc.](#)
[CrossRef Member](#)
[COUNTER Member](#)
 The MIT Press colophon is registered in the U.S. Patent and Trademark Office.
[Site Help](#)