The official link of the article titled “A survey of automatic term extraction for Brazilian Portuguese” is:

http://www.journal-bcs.com/content/20/1/12

Abstract

**Background**
Term extraction is highly relevant as it is the basis for several tasks, such as the building of dictionaries, taxonomies, and ontologies, as well as the translation and organization of text data.

**Methods and Results**
In this paper, we present a survey of the state of the art in automatic term extraction (ATE) for the Brazilian Portuguese language. In this sense, the main contributions and projects related to such task have been classified according to the knowledge they use: statistical, linguistic, and hybrid (statistical and linguistic). We also present a study/review of the corpora used in the term extraction in Brazilian Portuguese, as well as a geographic mapping of Brazil regarding such contributions, projects, and corpora, considering their origins.

**Conclusions**
In spite of the importance of the ATE, there are still several gaps to be filled, for instance, the lack of consensus regarding the formal definition of meaning of ‘term’. Such gaps are larger for the Brazilian Portuguese when compared to other languages, such as English, Spanish, and French. Examples of gaps for Brazilian Portuguese include the lack of a baseline ATE system, as well as the use of more sophisticated linguistic information, such as the WordNet and Wikipedia knowledge bases. Nevertheless, there is an increase in the number of contributions related to ATE and an interesting tendency to use contrasting corpora and domain stoplists, even though most contributions only use frequency, noun phrases, and morphosyntactic patterns.

Keywords:
Automatic term extraction; Statistical; Linguist; Hybrid knowledge