## A Matter of Words - DTU Orbit (09/11/2017)

## A Matter of Words: NLP for Quality Evaluation of Wikipedia Medical Articles

Automatic quality evaluation of Web information is a task with many fields of applications and of great relevance, especially in critical domains, like the medical one. We move from the intuition that the quality of content of medical Web documents is affected by features related with the specific domain. First, the usage of a specific vocabulary (Domain Informativeness); then, the adoption of specific codes (like those used in the infoboxes of Wikipedia articles) and the type of document (e.g., historical and technical ones). In this paper, we propose to leverage specific domain features to improve the results of the evaluation of Wikipedia medical articles, relying on Natural Language Processing (NLP) and dictionaries-based techniques. The results of our experiments confirm that, by considering domain-oriented features, it is possible to improve existing solutions, mainly with those articles that other approaches have less correctly classified.

## **General information**

State: Published Organisations: Department of Applied Mathematics and Computer Science, Embedded Systems Engineering, IIT-CNR Authors: Cozza, V. (Ekstern), Petrocchi, M. (Ekstern), Spognardi, A. (Intern) Number of pages: 9 Pages: 448-456 Publication date: 2016

## Host publication information

Title of host publication: Web Engineering : 16th International Conference, ICWE 2016, Lugano, Switzerland, June 6-9, 2016. Proceedings Volume: 9671 Publisher: Springer ISBN (Print): 978-3-319-38790-1 ISBN (Electronic): 978-3-319-38791-8

Series: Lecture Notes in Computer Science ISSN: 0302-9743

Main Research Area: Technical/natural sciences

Conference: The 16th International Conference on Web Engineering (ICWE2016), USI Lugano, Switzerland, 06/06/2016 - 06/06/2016

Information Systems Applications (incl. Internet), Information Storage and Retrieval, Software Engineering, Computer Appl. in Administrative Data Processing, User Interfaces and Human Computer Interaction, Artificial Intelligence (incl. Robotics) DOIs:

10.1007/978-3-319-38791-8\_31 Source: FindIt Source-ID: 2306622913 Publication: Research - peer-review > Article in proceedings – Annual report year: 2016