

MOODLE TEXT-CLOUDS

Riccardo Mazza¹, Lorenzo Sommaruga¹, Andrea Baldassari¹, Luca Mazzola²

¹SUPSI – Scuola Universitaria Professionale della Svizzera italiana
Lugano – Switzerland

²eLab – eLearning Lab @ USI – Università della Svizzera italiana
Lugano – Switzerland

{riccardo.mazza, lorenzo.sommaruga, andrea.baldassari}@supsi.ch
luca.mazzola@usi.ch

PAPER

THEMATIC AREA: Technical aspects

Abstract

We have created a plugin for Moodle that generates a text cloud of the most popular keywords of the course. Text keywords are automatically extracted from course resources and correspond to the most frequent words that appear in resources. The size of keywords in the visual representation reflects the frequency of the keywords among the resources of the course. Keywords in visual representations are hyperlinks. When the user clicks on a keyword, a small pop-up shows a collection of resources that contain that keyword. The user can then select and immediately access the requested resource.

Keywords - Text-Cloud, classification of resources, navigation inside course materials, block

1 RATIONALE

According to a statistic conducted in the academic year 2008/09 at the University of Lugano and at the University of applied sciences of Southern Switzerland on 1743 Moodle courses, the average number of resources inserted in a course is about 20. We found also that 196 out of the 1743 existing Moodle courses provided on-line more than 50 resources, and the maximum number of resources found in a single course was 259. When the number of resources is high (above 20), it became difficult to the learner to find a particular resource that he/she might need for his/her study. The learner has to sequentially browse among all the available resources to find the right one. The worst (and not so uncommon) case is when the title of the resource is not meaningful (such as *slides lesson 1*, *slides lesson 2*, etc.), in this case the learner has to download and open each single file and read all the contents, in order to find the right resource. We want to improve the usability of Moodle, and allow learners to reach the desired resources in fewer clicks. We believe that by improving the usability of Moodle, we can stimulate and encourage its use in all universities (in Switzerland and abroad). A way to facilitate learners in finding a particular resource with a single click could be the inclusion on Moodle interface of a text cloud of the most important keywords.

2 THE TEXTCLOUD BLOCK

The purpose of this paper is to present a new block for Moodle that generates a text cloud with the most popular keywords of the course. Text keywords are automatically extracted from course resources (HTML and text pages, PDF files, WORD and PPT documents, ODF documents), and correspond to the most frequent words that appear in course resources. Keywords are represented in the typical visual representation of tag cloud arrangement with alphabetical sorting in a sequential line-by-line layout (see Figure 1). The size of keywords in the visual representation is proportional to the number of times that the keyword appears in the documents of the course. Keywords in visual representations are hyperlinks. When the user click on a keyword, a small pop-up shows a collection of resources that are

analysis computer course
data design
examples fig font
graphical **graphics**
http knowledge le maps
network number
organization process
profiles
representation space
tables techniques time
tio tree tuft variables
visual will

Fig. 1: Example of tag cloud

associated with a keyword. The user can then select and immediately access the requested resource.

This functionality improves the usability of Moodle, as it allows learners to reach the desired resources in fewer clicks: in this way, we can stimulate and encourage its usage. The software is released as Free software under the GPL software license [1].

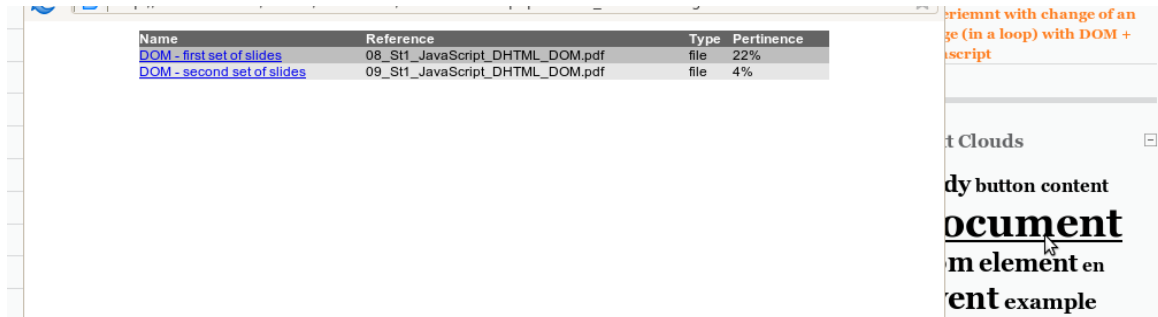


Fig. 2: User clicked on a tag: all the resources connected are presented (with the relative relevance)

2.1 TECHNICAL DESCRIPTION

The feature is implemented as additional add-in for the Moodle platform. The instructors can include this block in their courses, and the students will see the visualization that can be used for a direct

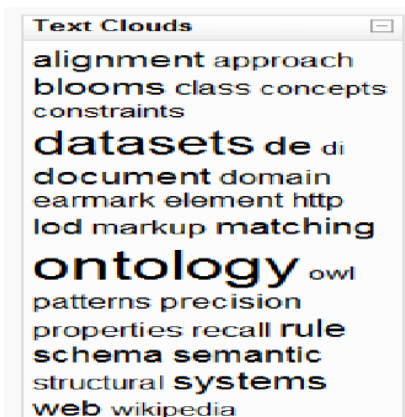


Fig. 3: Example of block

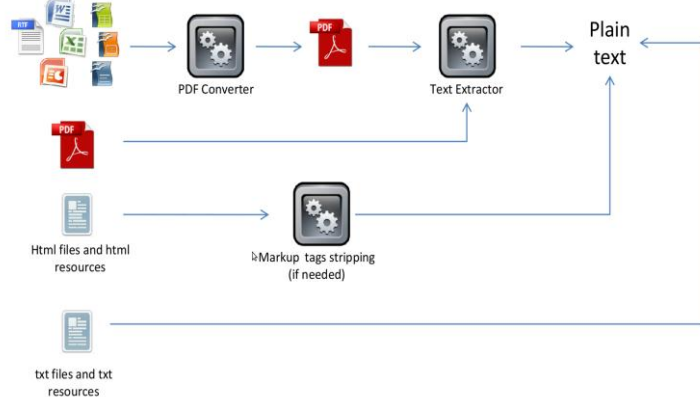


Fig. 4: The Process for creating source of TagCloud block

access to the resources that are associated with a keyword. It basically performs the following operations:

- 1)_ On each cron cycle, the block search for new resources included in the course. If new resources are found, it process the resource for indexing
- 2)_ The indexing process distinguishes between different resources types. MS Office and Open Document Format documents are converted into PDF format by the Unoconv application [2]. PDF files are converted into PlainText Files. HTML files are converted into Text files via a specific function that removes HTML tags and other formatting text.
- 3)_ Text of PlainText files is extracted, and processed for text analysis. This involves other further steps: a) Language detection; b) stop-words filtering; c) stemming algorithm application.
- 4)_ Last step consists on calculating frequencies and other statistics to determine the weight of every tag.

Moodle text cloud block is able to work with different languages. The current version comes with 11 analysis languages (danish, dutch, english, french, german, italian, norwegian, polish, portuguese, spanish and swedish), 4 of these (english, german, italian and spanish) contains also the stemming engine. New language analyser and stemming engine can be added very easily.

Riferimenti bibliografici

[1] <http://sourceforge.net/projects/moodletextcloud/>

[2] <http://dag.wieers.com/home-made/unoconv/>