

What useR! deals with : a text mining over the ages

Bertrand Vautier^{1*}

1. Student in applied statistics, agrocampus-ouest

* Contact author: bertrand.vautier@free.fr

Keywords: text mining, useR, abstracts, lexical field

Since 2004 useR! Conferences gather contributors, users, programmers and everyone in touch with the R project. Now is the time to sum up the subjects presented at useR!

The aim of this work is to give an overview of the subjects presented at useR! To show the evolution of attendance and of interests over years at useR! This work has been done using the abstract books of 2004, 2006 and 2008 and the power of text mining.

Text mining applied to these abstracts showed interesting results : in addition to the expected growth of attendance, the list of participants nationality got bigger over years and, in the group of the most represented countries, each country was quite specialized on a specific subject. This method allowed to draw a list of lexical fields, which were the main subjects developed at useR!, and it showed a particular result : some subjects were not only separated from the mass of words but also opposed to specific fields.

In this work, text mining showed a great aptitude to summarize more than 400 texts dealing with various subjects. It revealed itself as a powerful analysis tool whose scope goes beyond texts, it analyses also the evolution and policy of R-project and highlights the work of thousands of people.

References

useR!2004 Abstract book (2004).

<http://www.r-project.org/conferences/useR-2004/abstracts/Abstracts.pdf>

useR!2006 Abstract book (2006).

<http://www.r-project.org/useR-2006/Abstracts/Abstracts.pdf>

useR!2008 Abstract book (2008).

http://www.statistik.uni-dortmund.de/useR-2008/abstracts/_Abstracts.pdf