Author's response to reviews

Title: CDAPubMed: a browser extension to retrieve EHR-based biomedical literature

Authors:

David Perez-Rey (dperezdelrey@fi.upm.es)
Ana Jimenez-Castellanos (ajimenez@infomed.dia.fi.upm.es)
Miguel Garcia-Remesal (mgremesal@fi.upm.es)
Jose Crespo (jcrespo@fi.upm.es)
Victor Maojo (vmaojo@fi.upm.es)

Version: 2 Date: 1 September 2011

Author's response to reviews: see over
Cover letter for second (revised) submission to BMC Medical Informatics & Decision Making

Title of the Manuscript: CDAPubmed: a browser extension to retrieve EHR-based biomedical literature

Authors: David Perez-Rey, Ana Jimenez-Castellanos, Miguel Garcia-Remesal, Jose Crespo and Victor Maojo

Madrid, Spain, 1st September, 2011

Dear Editor,

We are submitting a revised version of our manuscript entitled “CDAPubmed: a browser extension to retrieve EHR-based biomedical literature” to be considered for a “software” publication at BMC Medical Informatics & Decision Making.

We would like to thank the three referees for their valuable comments regarding our work. We have followed their recommendations and rewritten a significant part of the paper. We hope that our paper will be now suitable for publication in BMC Medical Informatics and Decision Making. Changes have been highlighted in red to facilitate their task.

The main concerns presented in the review were related to previous works in the area, the use of a relevance concept and the evaluation of CDAPubMed. As one of the reviewers correctly pointed out, we did not focus our paper on issues directly centered on research evaluation since we submitted our paper as a “Software” publication. In the new version of the manuscript, we clarified that the main contribution of our work is the new functionality that CDAPubMed provides to generate EHR-based literature queries. Thus, one of the results is the open source tool that we have made available (following the BMC Medical Informatics and Decision Making recommendations, at the moment of the publication, CDAPubMed will also be published on Sourceforge).

Anyway, as suggested by the reviewers, we have extended the results with a larger dataset of EHRs, and one of the co-authors (Prof. Maojo, who is also a medical doctor), has evaluated the keyword identification process. If further evaluation on the results are required, or reviewers suggest to change this work to a “research” type of publication, please let us know it. In this case, we can include other experts in the area for the kind of evaluation that should be necessary for such different type of publication.

Bellow, we present a point-by-point description of the changes made following the reviewers suggestions.

Major issues:

1. Previous work such as “InfoButtons” (and the MEDLINE button) has been discussed in the Background section compared to the semi-automatically generated queries of CDAPubMed. The corresponding references have been included.

2. We have extended the test set to the 17 English EHRs from the public HL7 dataset.
3. We have carried out an evaluation of the identification process —with human judgment. Results are presented in a new Table (1). Additional information regarding the experiment can be included as additional material.

4. HL7 CDA has been further commented in the Architecture and Technologies section.

5. New versions or other EHR formats have been discussed in the Discussion and Conclusions sections.

6. The description of Figure 2 has been extended with configuration examples. We understand that figures 3 and 4 (CDAPubMed screenshots) may also help to understand the process.

7. OpenNLP has been briefly discussed in the Keyword Identification section.

8. Regarding a possible weakness of retrieving zero results by adding too many keywords, it has been clarified, in the Functionality and Discussion section. Only keywords that retrieve at least one publication, when added to the query, can be selected —the number of potential publication is displayed as a superscript of the keyword.

9. The Keyword Identification process has been commented in the Discussion section regarding previous works on assigning MeSH terms to clinical text. These methods and MeSH major topics could be used in the future to provide a meaningful order to EHR keyword candidates. The corresponding references have been included.

10. The relevance metric has been removed.

Minor issues

1. The Background section of the abstract has been shortened

2. We have changed every “Pubmed” occurrence by “PubMed”

3. Most “—” (“dash” characters) have been deleted

4. PubMed-EX reference has been included in the discussion section

5. Other spelling and other errors has been solved (“such (as) HubMed”, “a(n) NLP”, “PubMed (MEDLINE)”, etc.)

6. Resolution of Figure 3 has been increased

7. CDAPubMed can be downloaded from the project home page http://porter.dia.fi.upm.es/cdapubmed, included in the Availability and Requirements section

Regarding the quality of writing, our English editor was on vacation on August, but she would review the paper again before its publication.

We are looking forward to hearing from you.

Sincerely yours,
David Perez-Rey, PhD
Associate Professor
http://norbert.dia.fi.upm.es
Departamento de Inteligencia Artificial
Facultad de Informática, Universidad Politécnica de Madrid
Campus de Montegancedo S/N
28660 Boadilla del Monte, Madrid, SPAIN
Phone: +34-91-336-74-45
Fax: +34-91-352-48-19
e-mail: dperezdelrey@fi.upm.es