Response to Reviewers

"A Study of Web Page Understandability for Consumer Health Search"

Joao Palotti, Guido Zuccon, Allan Hanbury

Dear reviewers and editors,

Thank you for your comments on our manuscript. A number of suggestions were provided and have been addressed in the revised submission. General comments and summary of modifications are provided at the end of this response letter.
Our detailed response to each reviewers' comments is provided below in the following format:

Original review comments are included verbatim in normal typesetting.

The authors response is provided with this typesetting.

Reviewer G:

General comments
=============
This highly technical paper discusses an important issue.

We thank the reviewer for highlighting the importance of the problem this article is tackling.

Specific comments
=============
Major comments
----
1. This is a very long paper, some of the information in tables could be placed in appendices which may improve the readability

We have followed the suggestion of the reviewer and have moved some of the context/tables in an appendix at the end of the article. To simplify the article, we have also replaced Table 14, which reported the detailed evaluation values and statistical analysis for the experiments, with a series of plots (Figures 4, 5, 6, 7 and 8), which make trends easier to understand. We have also removed the details of the statistical analysis from the body of the article and reported this analysis in the appendix.

2. in the section 'Integrating Understandability into Retrieval' I'm not clear if you're using readability and understandability as interchangeable terms or if you have a difference between them?

The original manuscript clarified this in the opening paragraph of the Related Work section:

"Understandability refers to the ease of comprehension of the information presented to a user. Put in other words, health information is understandable "when consumers of diverse backgrounds and varying levels of health literacy can process and explain key messages" [22]. Often the terms understandability and readability are used interchangeably: we use readability to refer to formulae that estimate how easy is to understand a text, usually based on its words and sentences. We use understandability to refer to the broader concept of ease of understanding: this is affected by text readability (as increasing readability tends to improve understanding), but may also be influenced by how legible a text is and its layout, including e.g., the use of images to explain difficult concepts."

From the reviewer's comment, we assume that, given the length of the article, the reader may gloss over the Related Work section, or may not recall what mentioned there when reading the later sections. To aid the reader we have included a reminder into the first paragraph of the 'Integrating Understandability into Retrieval' section: "Remember that, as described in the 'Related Work' section, readability formulae are a specific approach to estimate understandability."

3. Shortening the technical details and increasing the practical application in the discussion would help the non-expert reader to understand your key points more easily.

We decided to maintain the levels of technical details present in the original manuscript, because these allow for the reproduction of our study.

4. I didn't understand principal finding no 3. 'Learning to rank methods can be specifically trained to promote more understandable search results, while still providing an effective trade-off with topical relevance.' Learning by who – machine learning? Trained by who?

`Learning to rank` (in short: LTR or LETOR) is the expression used in Information Retrieval literature to refer to a family of machine learning approaches for which a *ranking* model (rather than a classification model) is learnt from training data.
To clarify this concept to the non-expert reader, we added the following sentence to the methods section:

"Learning to rank refers to a family of machine learning methods where ranking models are learnt from training data (and associated features)."

Minor comments

To follow on from this comment, we have added the following references:

[2] Fergus T, Spada M. Cyberchondria: Examining relations with problematic Internet use and metacognitive beliefs. Clinical psychology & psychotherapy; 2017; 24(6): 1322-1330. doi: 10.1002/cpp.2102.

[3] Diviani N, Putte B, Giani S, Weert J. Low Health Literacy and Evaluation of Online Health Information: A Systematic Review of the Literature. J Med Internet Res; 2015; 17(5): e112. doi: 10.2196/jmir.4018.

6. Figure 2 is very difficult to read, would reversing the axes help?

Following the suggestion by the reviewer, we reversed the axis, but the resulting figure was too wide to fit into a page. Therefore, we split Figure 2 from the original manuscript into the new Figures 2 (CLEF 2015) and 3 (CLEF 2016). We believe that the new version of the figures make results more readable.

7. Table 14 in the review copy was impossible to read

As already mentioned in answer to comment #1, we replaced Table 14 by a set of figures (Figures 4, 5, 6, 7 and 8). From these figures, the reader can more easily grasp the results and the trends we comment upon in the article. We have moved the detailed numerical results and the statistical analysis in an appendix.

Reviewer H:

General comments
=============
This paper....

Specific comments
=============
Major comments

None

Minor comments

ABSTRACT

1. Methods Line 2: …..on Web pages…..

The reviewer is referring to this sentence:

"In this article, we investigated methods to estimate Web page understandability"

We believe the current wording is correct, i.e. the intended meaning is that the methods estimate the understandability of a Web page.

2. Conclusion Line 1: …specialized….

We followed the British spelling of this word (i.e. specialised) in place of the suggested American spelling.

..

INTRODUCTION

3. Authors should consider to use 'relevance' in place of 'aboutness'.

The only mention of `aboutness' in the submitted manuscript was in the following sentence:

"Commonly, signals about the topicality or aboutness of a piece of information with respect to a query are used to estimate relevance"

We believe the use of aboutness in this context is appropriate: here we refer to the concept of the `information being about the query' (i.e. the information being on the topic of the query), while we recognise relevance being a broader and more complex than aboutness/topicality. This treatment of the concepts of aboutness, topicality and relevance follows the established literature in Information Seeking and Retrieval.

4. The subheading 'Related Work' should be expunged as the whole content under it forms part of the Introduction.

We thank the reviewer for this suggestion; however we prefer keeping the Introduction section more streamlined, as is in the original version of the manuscript, and rather introduced the content in section Related Work as indeed relevant, related work.

METHOD
5. Consider to review the word 'unaccessed'.

The reviewer is referring to the word `unassessed', rather than `unaccessed' (which is not mentioned in the manuscript.
Unassessed and unjudged are used interchangeably in the Information Retrieval literature to refer to documents for which an assessment, or judgement (typically of relevance) has not been collected or provided.

TABLES
6. Tables 13 and 14 are rather too tiny to read. As such, authors should consider making it clearer.

As per answer to reviewer's G comments, we replaced Table 14 with figures that makes the main results easier to read and follow.

CONFLICT OF INTEREST
7. Authors should make a statement on conflict of interest. If none, they should declare there is none.

We thank the reviewer for this suggestion; a declaration of no conflict of interest was added to the paper.

NOTE ALSO
8.   'analyzed' is in more common use than 'analysed'.
9.   So also, 'summarized' in place of 'summarised'.
10.   So also, 'specialized' in place of 'specialised'.
11.   So also, 'minimized' in place of 'minimised'.

We followed the British spelling of these words (e.g., 'analysed') in place of the American spelling (e.g., 'analyzed').

General Comments

In summary, the revisions broadly focused on improving the readability of the paper by: moving the statistical analysis from the main article to an appendix, moving part of the detailed empirical

evaluation results from a table in the main article to figures (and moving the detailed analysis to an appendix), and restating our interpretation of understandability and readability.

We thank the reviewers for their comments that have strongly contributed to an improved manuscript and look forward to your response on the revised submission.

Sincerely,

Joao Palotti, Guido Zuccon, Allan Hanbury