Providing Web Credibility Assessment Support

by

sonal aggarwal, Herre Van Oostendorp, Y.Raghu Babu Reddy, Bipin Indurkhya

in

European Conference on Cognitive Ergonomics (ECCE)

Report No: IIIT/TR/2014/-1

Centre for Software Engineering Research Lab
International Institute of Information Technology
Hyderabad - 500 032, INDIA
September 2014

Providing Web Credibility Assessment Support

Sonal Aggarwal
IIIT, Hyderabad
500 032, AP, India
sonalaggarwal1@gmail.com

Herre Van Oostendorp
Utrecht University
Princetonplein 5,
3584CC Utrecht,
the Netherlands
h.vanoostendorp@uu.nl

Y. Raghu Reddy
IIIT, Hyderabad
500 032, AP, India
raghu.reddy@iiit.ac.in

Bipin Indurkhya
IIIT, Hyderabad
500 032, AP, India
bipin@iiit.ac.in

ABSTRACT
Presence of information from multiple sources on the internet requires evaluating the credibility of the information, before its utilization. Researchers have suggested that internet users experience difficulty in accessing necessary information and do not pay enough attention to its credibility. We present here the design and implementation of an automated Web Credibility Assessment Support Tool (WebCAST) that considers multiple factors (type of website, popularity, sentiment, date of last update, reputation and review based on users' ratings reflecting personal experience) for assessing the credibility of information and returns a summary indication of the credibility of a website. We use Potentially All Pairwise RanKings of all possible Alternatives (PAPRIKA) method of Multi-Criteria Decision Analysis (MCDA) to give weights to the scale values on each factor, representing the relative importance of the attributes. An empirical evaluation of the tool was conducted by computing the correlation between the tool-generated credibility scores and that of human judges. The correlation was found to be 0.89, thus verifying the validity of the tool. In the future the proposed tool can be made useful to students in their learning process of credibility assessment.

Author Keywords
Websites; Web Trust; Credibility; Automation; Support tool.

INTRODUCTION

Nowadays we rely on the internet for most of the information we need. However, when the information needed is available on multiple websites, how do we trust one website over the other? Corritore *et al.* [5] suggest that the trust for online information is influenced by three major factors: perception of credibility, ease of use, and risk that the user's vulnerabilities will be exploited (*i.e.* a user expects his or her interests to be respected by the website.) In this study, we focus on the credibility. Credibility becomes important when the information presented on the website is used by the end-user for further processing or decision-making. We define credibility as the level of confidence a user puts on the information available on a given website based on various objective and subjective factors.

Various factors can influence credibility assessment. In our previous work [1], we automated credibility assessment based on four factors: type of the website, date of the most recent update, result of sentiment analysis, and Google page rank. We also mentioned other influencing factors like information about the author, presence of advertisements, affiliations, and domain expert's view that need to be considered for credibility assessment. While evaluating a previous version of our tool, we found a low correlation (0.484) between the system-generated credibility scores and scores given by human judges. A plausible reason for this weak relation may have been due to the difference in opinion between human judges about the various influencing factors. This suggested that the perception of credibility by the end-user might vary based on the criteria they set. Another reason might be that the implementation of the tool was based on only four factors. To overcome this limitation, we developed an automated web credibility assessment support tool (WebCAST) that incorporates additional factors relevant for credibility assessment, i.e., reputation and review based on users' rating reflecting personal experience with the website. Thus, we use six criteria in all now instead of four in the previous work. We make use of the online available real-time databases that have information related to several factors that affect credibility. We hypothesize that this tool will provide more valid credibility scores as compared to our previous system, and will correlate positively and substantially when compared with human judges.

The rest of the paper is organized as follows. We first review recent research related to user's perception of credibility, factors that affect credibility, and tools for automation of credibility assessment. Then, we describe the methodology of designing WebCAST. Further we present

ECCE '14, September 01 - 03 2014, Vienna, Austria

ACM978-1-4503-2874-6/14/09…$15.00.
http://dx.doi.org/10.1145/2637248.2637260

an evaluation of WebCAST in terms of it's validity (that is, it's compatibility with human judgments). Finally, we provide some conclusions and suggestions for future research.

BACKGROUND AND RELATED WORK

User perspective for credibility assessment of websites

According to the 3S-model proposed by Lucassen *et al.* [15], there are three strategies for credibility evaluation of a website: i) considering semantic features like accuracy or neutrality (which depends on domain expertise), ii) surface features like the design of the website, website aesthetics, coloring of text and iii) past experience with the source of information, which refers to the details about the owner or the author of the website and their affiliation. They also suggest that semantic features like accuracy of information influence credibility only when the users have domain expertise pertinent to the information. Surface features like text color influence credibility when users have high information skills, i.e. these users have experience in processing and evaluating documents and websites but possibly in other domains.

Studies support the notion that the source of a website affects its credibility only for familiar users [14]. When the information comes from an unknown source, users rely mainly on references, images and textual features [13]. However, the results shown in their work was restricted to Wikipedia articles. In general this may not be valid for all websites. Based on the studies on users' perspective for credibility evaluation, Lucassen *et al.* suggest that surface features could be taken into account while developing credibility assessment support systems for novice (unfamiliar) users, but familiar users do not have any specific preference for semantic or surface features while assessing credibility [12]. In our proposed tool, we take into consideration mainly the semantic features like the neutrality of the website by doing a sentiment analysis or using the reputation and users' ratings reflecting personal experience with the website.

Impact of credibility assessment

Studies show that college students increasingly depend on internet for general-purpose information as well as research-related information, but do not explicitly evaluate the source of information [16]. Case studies with undergraduates on critical analysis of internet and scholarly sources suggest that they are unable to discriminate between credible and non-credible sources [2]. However those students who assess the credibility, and use source characteristics, achieve better comprehension of the content [25, 26]. Providing a novice user with credibility scores adds to the confidence of the user about the selected information [4]. Studies also show that as people gain experience, they are more likely to evaluate the credibility of the web sites [17]. Thus, as the user becomes more knowledgeable, their perception of credibility becomes better. The WebCAST tool proposed in this paper is designed to serve this purpose of providing users feedback on credibility of websites.

Criteria of credibility assessment

A thorough study of literature on credibility assessment led to the following criteria:

- *Type* of website as reflected by the top-level domain or domain suffix of the website address (URL) [8]. Usually a website which is meant for non-commercial purpose tends to be more credible. For example .com in the URL refers to business or commercial website while .gov refers to a government website.
- The last *updated date* of a website impacts the credibility of a website. A website which is more recently updated is likely to have more recent and credible information. This is useful when the information being hosted on the website is dynamic and changes with time [19].
- Whether a web-document is a *primary* or *secondary source*. If a website is the primary source for the information, it is likely to be more credible, while there are more possibilities of fact manipulation if it is a secondary source.
- Availability of *contact information* (address and/or email id) of the owner of the website. This information is important because it shows that the author/owner of the website is known and could be contacted for any clarifications related to its authenticity.
- Analyzing the *link integrity* of the website [10]. A website with a balance in number of internal and external links is more credible. If a website is referring more to the external links, it is basically redirecting the user to another domain whose credibility needs to be assessed again by the user. Also, a website should not have any broken links as it negatively affects its credibility.
- Analyzing the header/footer of the website for any *affiliation* (if available) [3]. Usually the header of a webpage contains the logo of the concerned organization. Also the footer gives information about its sponsors. This can be used to assess trust of the given website.
- *Completeness*, *accuracy and unambiguity* of the information make it more credible. This leaves little to no room for interpretation or assumptions about the given information.
- Authors' *expertise* on the provided content. Their *connection* to the *source* of publication; *connection* to the intended audience. Their *point of view*, objectivity and impartiality. Doing a "sentiment analysis" of the website content, it can be decided whether the author's opinion is biased (positive or negative) or unbiased (neutral). In addition, authors' *credentials* like institutional affiliation

(where they work), educational background, past writings, or experience.

- *Purpose* of the website. This shows the intent of the website owner for which the website has been created. It can be inferred to a certain extent by the type of website. For example, a commercial website is mainly meant for selling products while the purpose of an educational website is more to provide knowledge to its visitors.
- *Interactivity and usability* of the website. A user-friendly website makes it easy for the user to access it, thereby making it more preferable.
- *Website design* and *aesthetics* also add to the credibility of the website. Its structure in terms of graphics and text should be appropriate [10].
- *Quality* of information on the website: elementary, technical, or advanced. Users will find a website with good quality of information more credible.
- *Tone* of the website: ironic, humorous, exaggerated or overblown arguments need to be considered while interpreting the credibility of a website [24]. For example a website with overblown arguments with no supporting evidence is worthless to read.
- Determining whether the *same person* or *organization* is supplying the advertising as well as the informational content. If so, advertising is likely to bias informational content. Usually advertisements tend to divert the user's attention, so its presence or absence reduces/adds to the credibility.
- Determining any *software requirements* that may limit access to web information could be another factor.
- *Ranking* of the website by various search engines can also be a criterion of source credibility assessment. Google Page ranks and Alexa popularity ranks are few of such databases that could be utilized for this purpose. These rankings help to determine the popularity of a given website among its users, which in turn is dependent on its number of page-visits and unique visitors.
- *Domain experts' view* on the credibility of any web-document could be helpful to a novice user for assessing a website. Experts develop different strategies while seeking information from internet. These strategies can also be helpful to a novice user. An expert can assess the credibility of a web document on the basis of two dimensions of credibility: trustworthiness (well-intentioned, unbiased) and expertise (knowledgeable, competent) [9]. Ratings given by experts or previous users of the website, when given to a novice user, who is new to the website, can be helpful in her or his perception of credibility.

Automating the process of credibility assessment

A number of solutions have been proposed by various authors for automating credibility assessment [1, 2, 20, 23]. While some of them are generic in nature, others are specific to a set of websites. Credibility scores may be based on assessing the structure of a website's links, weighted approach to combining various features or machine learning techniques. Another automated approach for credibility assessment is based on finding the disputed information on the internet for given information [7]. Other major techniques for credibility assessment are using the social popularity of a web page based on Facebook [18], or using a decentralized approach as proposed by Thanasis *et al.* [20]. In [21], it is argued that augmenting the search results with few web-page features leads to more effective credibility assessment at the user end. We incorporate this in our approach by rearranging the Google and/or Wikipedia search results based on their assessed credibility scores. The website with the highest credibility score is presented on top (see Figure 2).

In our work, we give a relative credibility scoring, that is, we assign credibility scores to presented websites with respect to each other. So the highest-scored website is considered to be the most credible for a given search topic. We use a search topic as input (e.g. "balanced diet"), while the search results provide a list of website URLs. Applying the tool on these websites and ranking them in terms of credibility provides useful output in the form of a summary indication of the credibility of the websites evaluated. Alternatively the tool can be modified to assess the credibility, given a list of URLs instead.

For the evaluation of the tool, either we can compare it with some predefined credibility scores, or with human judges. Alexandra *et al.* used Microsoft dataset containing 1000 URLs along with the credibility scores for their tool evaluation [18]. To evaluate our tool, we use human judges and determine the correlation between the scores of the automated tool and those given by human judges.

DESIGN AND IMPLEMENTATION

Design

In previous section we indicated various factors that need to be considered for credibility assessment. An automated tool that considers all these factors would be ideal; however it is infeasible given the difficulty of quantitatively measuring all these factors. Hence, we have automated credibility scoring using a few major factors that can be assessed quantitatively as well as automatically. Moreover, to develop our tool, we used various Application Programming Interfaces (APIs) provided by the real-time databases available on internet, such as Alchemy API for sentiment analysis, Google API for search results, Web Of Trust (WOT) API for users rating and Alexa web request for getting popularity ranks. As these online APIs are frequently used by many different users and also

compatible with the research work related to credibility assessment, we consider using this information as a justifiable step in tool development, even though strict information about their reliability is lacking.

Figure 1 shows our process diagram for credibility assessment. The user gives the search topic and the search engine name as input. Alternatively, a list of URLs for the websites, the credibility of which has to be assessed, is given as input. Next, the computation of the following 6 factors is done in parallel: a) type of website, b) date of update, c) sentiment analysis, d) website popularity, e) reputation based on users' ratings and f) review based on users' ratings. We use multithreading to minimize the computation time. To obtain a credibility score based on these six factors, we perform Multi-Criteria Decision Analysis (MCDA) using the **P**otentially **A**ll **P**airwise **R**an**K**ings of all possible **A**lternatives (PAPRIKA) method [11], and determine weights of each attribute (category) on the six factors. The credibility score for each website is finally calculated using these weights.

Implementation

WebCAST has been implemented using Python due to its robustness, flexibility, and ease of use. Additionally, all the real-time datasets used for the credibility assessment tool provide access through API in Python. The step-wise procedure of implementation is described below.

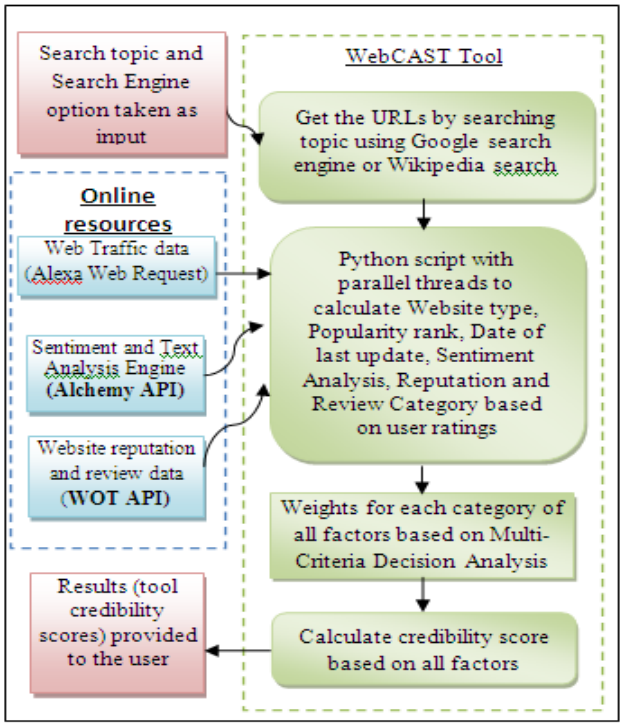

Figure 1. Process diagram

a) Finding the search results with date of last update

The input to WebCAST is a search topic of user's interest. The user can choose either Google as a search engine, or external links of Wikipedia, or both. Google search results are limited to 64 (as provided by the Google API), while Wikipedia search gives all available external links. We use Pattern software, a web-mining module for the Python programming language, for obtaining the search results programmatically [6]. The Python script function returns search URL along with the title, description, author name and date of last update of the website, when such information is available.

b) Computing the Website type

As mentioned previously, the domain suffix of the URL can be used to assess the type of the website. We compute the website type using a Python script function considering the categories as shown in Table 1.

Domain suffix	Website Type
*.gov	Government
*.edu, *.ac	Educational
*.org	Non-profit Organization
*.info	Informational
*net	Network
*.com	Commercial

Table 1. Website Types based on domain suffix

c) Computing the Popularity Rank

A popularity rank indicates how popular a website is. This ranking is most of the time based on web traffic data. To compute it we use Alexa.com, which is a leading provider for web metrics. The rank is calculated using a combination of average daily visitors on a given website and its page views in past three months. The site with highest combination of visitors and page-views is ranked first. The request to obtain the rank is made as shown below:
http://data.alexa.com/data?cli=10&dat=snbamz&url=[URL]
The response to the above request is XML-based, which is parsed to get the popularity rank. The above approach of obtaining the popularity rank is based on a real-time database. The rank is given to a website and not to a specific page of website. The rank obtained is a numeric value. We have identified five intervals for this ranking: rank 1 to 100, 101 to 1000, 1001 to 10.00, 10.001 to 50.000 and above 50.000.

d) Computing the Sentiment Analysis

Sentiment Analysis is a way to determine the bias in the author's opinion. If the sentiment is neutral, it implies that the author's opinion is unbiased. We use an online

sentiment and text-analysis engine: Alchemy API [22], which looks for words that carry a positive or negative connotation, and then figures out which person, place or thing is being referred to. It also takes into consideration negations (like 'not') and modifiers (like 'good'). The request to the Alchemy API is made as shown below, and the results are parsed to get positive, negative or neutral sentiments:
http://access.alchemyapi.com/calls/url/URLGetTextSentiment?apikey=[API_KEY]&sourceText=raw&url=[URL]

e) Computing Reputation and f) Review Category based on user ratings

Reputation indicates how much a website could be trusted based on the opinion of users. If some users have accessed a particular website previously, and have provided their views/ratings, these reputation ratings can be helpful to a user who is new to the website. Web Of Trust (www.mywot.com) is based on crowd sourcing approach that collects ratings and reviews from a global community of millions of users who rate and comment on websites based on their personal experiences. The reputation ratings are complemented by the review of users. Instead of a numerical rating users have here to review and assess the websites in terms of verbal categories as Negative, Questionable, Neutral or Positive. The request to the API is made as shown below and the result in JavaScript Object Notation (JSON) format is parsed to obtain the reputation (varying from excellent to poor) and review category:
http://api.mywot.com/0.4/public_link_json2?hosts=[URL]/&callback=process&key=[key]
This way we obtain the values for the two factors reputation and review. Reputation is a numeric value (maximum value 100) which is categorized into five intervals (0-19, 20-39, 40-59, 60-79, 80-100 corresponding to very poor, poor, unsatisfactory, good, excellent reputation respectively) and the review category could be negative, questionable, neutral or positive.

url	alexa_rank	reputatior	category	websitetype	sentime	update_(	credibility
http://en.wik	7	Excellent	Positive:Good	Organizational	positive		70
http://www.t	26151	Excellent	Positive:Good	Organizational	positive	4/27/2010	66
http://www.s	123	Excellent	Positive:Good	Networking	positive	8/5/2010	63
http://wikitra	2818	Excellent	Positive:Good	Organizational	positive		61
http://travel.'	4	Excellent	Positive:Good	Commercial	positive		57
http://www.t	28094	Excellent	Positive:Good	Commercial	positive	10/27/2010	54
http://www.t	32649	Good	Positive:Good	Organizational	positive		51
http://thingst	432	Excellent	Positive:Good	Commercial	positive		50
http://www.t	198	Excellent	Positive:Good	Commercial	positive		50
http://www.t	968	Excellent	Positive:Good	Commercial	positive		50
http://www.v	5639	Excellent	Positive:Good	Commercial	positive		48
http://www.r	38448	Excellent	Positive:Good	Commercial	positive		42
http://www.b	29270	Excellent	Positive:Good	Commercial	positive		42

Figure 2. Result database file created after computation through WebCAST

Finding weights for the attributes on each factor

Once the scores corresponding to the various attributes or categories on the factors are obtained independently (See Figure 2), we compute the final composite scores based on the weight for these attributes on each factor. We use 1000minds (www.1000minds.com) software, which implements the PAPRIKA method. This method is based on choice-based conjoint analysis technique [11]. Stimuli consisting of multiple combinations of the attributes of the independent variables (factors) are presented and participants are asked for their preference. In this method, a pair-wise ranking of all possible alternatives is done to give dominant attributes more priority. The software automatically ranks all pairs, also the pairs that are logically implied by transitivity and eliminates them from presentation. In order to achieve this, the system requires in this case answering 822 questions (combining two factors at a time) to rank one item of a given pair over the other out of which 78 questions were explicitly answered and remaining 744 questions were resolved automatically. These questions refer to all hypothetical possibilities taking two categories (of two factors) at a time. Thus by answering these questions, the system computes and identifies the dominating attributes and ranks them accordingly via a mathematical procedure. For example, one such question is: Which is more credible: "A less popular educational website or a very popular commercial website?" The answers could be either of them is more credible or both are equally credible. For this study, we answered all these questions ourselves. In subsequent research it is worthwhile to repeat this, and to perform the rating process by a team of experts. After answering the questions, the system generates based on mathematical methods ranking of the categories (attributes) of each factor in the form of percentages that are used to get the weights. The percentages along with the scales of each factor are given in Appendix 1. We calculated weights from these percentages, by assigning the weight 100 to the scale value with the highest percentage and then assigning proportional weights to the other scale values. For instance, PAPRIKA assigned to the attribute Governmental website a credibility percentage of 18.8%, which is the highest, so a weight of 100 on credibility whereas the attribute commercial website only got 0.7% and a weight of 3.571 on credibility.

Calculating the composite credibility score and storing in Database

Once we have all factors (and categories) computed, the composite credibility score for each website is calculated as below:

$$\textit{Credibility score} = (x_1 + x_2 + x_3 + \ldots x_n) / n$$

where x_n is the weight for the computed scale of the n^{th} factor and n is the total number of factors (six). Refer Appendix 1 for the respective weights of the attributes. The computed factors and the composite credibility scores are

stored in a Microsoft Access Database file. Figure 2 shows a snapshot of the database file. The composite credibility scores (ranging from 0-100) are shown in the right-most column of the database file. For example a website as Wikipedia.org is categorized as an organizational website, having no information on the last update, having a positive sentiment, popularity rank of 7, excellent reputation and positive review category will be scored based on the weights in Appendix 1 as below:
Credibility score = (78.57 + 0 + 41.89 + 100 + 100 + 100)/6 = 70.07

EMPIRICAL EVALUATION

We carried out an empirical evaluation of WebCAST by conducting an experiment with some websites and human judges (students). Our working hypothesis was that the tool-generated composite credibility scores will be compatible and correlate with the credibility scores of human judges.
The variables measured were: credibility and usability on a scale of 1 (low) to 10 (high); we included usability as a variable in order to make the study more generic, and not let the participants know the objective of the test. We generated composite credibility score results through WebCAST for the topic *Balanced Diet*. We selected ten websites from the output set where the credibility scores vary from 0 to 70. The participants were asked to inspect these websites and to assess their credibility and usability. A mixture of twenty five undergraduate and graduate students (19 males and 6 females, mean age = 23.64 years, sd = 4.59) from IIIT Hyderabad participated in the experiment. The participation was on a voluntary basis. The participants were given an online form, which asked them to visit each URL.

After collecting the data, we performed in view of our hypothesis a statistical analysis to compute the correlation between the tool-generated credibility scores and the average credibility scores given by the participants. As mentioned previously, the usability related ratings were used only for making the experiment generic (and not to focus their attention only to credibility), and we do not use those data for further analysis. The Pearson product-moment correlation coefficient of credibility scores with the tool-generated scores is 0.8988 ($p < .01$). The positive correlation indicates that the tool performance is compatible with the credibility judgments of users. In our previous work [1], we had found a low (and not significant) correlation (0.484, $p < 0.10$) between the human judges and the proposed tool. Now with the enhanced version of the tool (WebCAST) that takes into consideration additional factors, particularly reputation and review ratings based on users' personal experience, the correlation is positive and considerably higher.

CONCLUSIONS AND DISCUSSION

We presented here an automated website credibility assessment support tool (WebCAST). This tool utilizes the existing real-time databases like Alexa for popularity ranking, and Web of Trust for user ratings and review data. The tool provides a completely automated practical approach for credibility assessment. WebCAST is developed using open-source solutions and therefore can be made freely available to the users.

The correlation value between the tool-generated scores and human judges shows that the tool is able to generate credibility scores compatible with human judgments, supporting our hypothesis.

The main goal of the study was to provide a proof of concept that it is possible to design a tool which automatically computes and summarizes credibility of websites. Future research plans include modifying and refining the tool to automatically retrieve more credibility-related factors specifically when the website is Web 3.0 compliant. Features such as primary or secondary sources, availability of contact information, and author credentials (e.g. affiliation) are promising next candidates. We also plan to conduct more experiments by increasing the number of participants, and by using a broader set of websites. In the current experiment we used students as human judges due to their availability at our institute. Further experiments could be conducted by employing experts for credibility assessment and also taking into consideration the inter-individual variability factors. Finally, implementing machine-learning techniques to take user's feedback on the system will probably improve its performance.

An application area where WebCAST in the future can be used is for teaching or training purposes. As indicated in the introduction, students often experience problems with evaluating information and sites when searching for information on the web [9, 25, 26]. We can imagine that the WebCAST tool can be used as component in a course for e.g. high school/bachelor students on learning how to use information from the internet and to make informed decisions. Further, information on how to evaluate websites would be part of this course. The tool could be presented to the students at the beginning of their learning process, and summary information of the tool could be provided to them as a model. As the students progress in learning how to use internet information, the tool could be gradually withdrawn (fading out), as we might expect the students to learn about the credibility criteria, and apply them in a systematic way themselves. This study has demonstrated that at least it is possible to design a valid tool, further development is needed to design an interface that is appropriate to this context.

REFERENCES

1. Aggarwal, S. and Van Oostendorp, H. An attempt to automate the process of source evaluation. *ACEEE International Journal on Communication* 2, no. 2 (2011).

2. Amin, A., Zhang, J., Cramer, H., Hardman, L., and Evers, V. The effects of source credibility ratings in a cultural heritage information aggregator. In *Proc. 3rd workshop on Information credibility on the web*, ACM Press (2009), 35-42.
3. Beck, S. The Good, The Bad & The Ugly: or, Why It's a Good Idea to Evaluate Web Sources. (1997). http://lib.nmsu.edu/instruction/evalcrit.html.
4. Bråten, I., Strømsø, H. I., and Britt, M. A. Trust matters: Examining the role of source evaluation in students' construction of meaning within and across multiple texts. *Reading Research Quarterly* 44, no. 1 (2009), 6-28.
5. Corritore, C. L., Kracher, B., and Wiedenbeck, S. On-line trust: concepts, evolving themes, a model. *International Journal of Human-Computer Studies* 58, no. 6 (2003), 737-758.
6. De Smedt, T., and Daelemans, W. Pattern for Python. *Journal of Machine Learning Research*, 13: 2031–2035, (2012).
7. Ennals, R., Byler, D., Agosta, J. M., and Rosario, B. What is Disputed on the Web? In *Proc. workshop on Information credibility*, ACM Press (2010), 67-74.
8. Evaluating Internet Information http://www.usg.edu/galileo/skills/unit07/internet07_08.phtml
9. Fogg, B. J., and Tseng, H. The elements of computer credibility. In *Proc. SIGCHI conference on Human factors in computing systems*, ACM Press (1999), 80-87.
10. Grassian, E. Thinking Critically about World Wide Web Resources. UCLA Library (1995) http://www2.library.ucla.edu/libraries/college/11605_12337.cfm
11. Hansen, P., and Ombler, F. A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *Journal of Multi-Criteria Decision Analysis* 15.3-4 (2008), 87-107.
12. Lucassen, T., and Schraagen, J. M. The role of topic familiarity in online credibility evaluation support. In *Proc. Human Factors and Ergonomics Society Annual Meeting*, vol. 56, no. 1, SAGE Publications (2012), 1233-1237.
13. Lucassen, T., and Schraagen, J. M. Trust in wikipedia: how users trust information from an unknown source. In *Proc. 4th workshop on Information credibility*, ACM Press (2010) 19-26.
14. Lucassen, T., and Schraagen, J. M. The influence of source cues and topic familiarity on credibility evaluation. *Computers in Human Behavior* 29, no. 4 (2013), 1387-1392.
15. Lucassen, T., Muilwijk, R., Noordzij, M. L., and Schraagen, J. M. Topic familiarity and information skills in online credibility evaluation. *Journal of the American Society for Information Science and Technology* 64, no. 2 (2013), 254-264.
16. Metzger, M. J., Flanagin, A. J., and Zwarun, L. College student Web use, perceptions of information credibility, and verification behavior. *Computers & Education* 41, no. 3 (2003), 271-290.
17. Metzger, M.J. Understanding how Internet users make sense of credibility: A review of the state of our knowledge and recommendations for theory, policy, and practice. In *Symposium on Internet Credibility and the User*, 2005.
18. Olteanu, A., Peshterliev, S., Liu, X., and Aberer, K. Web credibility: Features exploration and credibility prediction. In *Advances in Information Retrieval*, Springer Berlin Heidelberg (2013), 557-568.
19. Pattanaphanchai, J., O'Hara, K., and Hall, W. Trustworthiness criteria for supporting users to assess the credibility of web information. In *Proc. 22nd international conference on World Wide Web companion*, (2013).
20. Ranvier, J. E. M., Olteanu, A., Aberer, K., and Papaioannou, T. G. A decentralized recommender system for effective web credibility assessment. In *Proc. International conference on Information and Knowledge Management*, ACM Press (2012), 704-713.
21. Schwarz, J., and Morris, M. Augmenting web pages and search results to support credibility assessment. In *Proc. SIGCHI Conference on Human Factors in Computing Systems*. ACM Press (2011).
22. Sentiment and Text Analysis Engine - AlchemyAPI. http://www.alchemyapi.com/api
23. Sondhi, P., Vydiswaran, V. V., and Zhai, C. Reliability prediction of webpages in the medical domain. In *Advances in Information Retrieval*, Springer Berlin Heidelberg (2012), 219-231.
24. UC Berkeley - Teaching Library Internet Workshops, Finding Information on the Internet: A Tutorial (2010) http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/Evaluate.html
25. Walraven, A., Brand-Gruwel, S., and Boshuizen, H. How students evaluate information and sources when searching the World Wide Web for information. *Computers & Education* 52, no. 1 (2009), 234-246.
26. Walraven, A., Brand-Gruwel, S., and Boshuizen, H. Fostering students' evaluation behaviour while searching the internet. *Instructional Science* 41 (2013), 125-146.

APPENDIX 1

Weights for all attributes or categories on each of the six factors used for Credibility assessment. Please note that the percentage values are shown only up-to two decimals.

Type of Website	Percentage	Weight
others	0%	0
com	0.7%	3.571
info, net	3.5%	18.57
edu, ac, org	14.8%	78.57
gov	18.8%	100
Date of update		
not available	0%	0
>5 year	0.1%	2.381
> 1 year and <5 year	3.8%	66.67
less than 1 year	5.6%	100
Sentiment		
negative	0%	0
positive	4.2%	41.89
neutral	9.9%	100

Alexa Popularity Rank		
50.001 and above	0%	0
10.001 -50.000	2.7%	12.42
1001-10.000	10.1%	46.58
101-1000	11.8%	54.66
1-100	21.6%	100
Reputation (based on user's rating)		
0-19 (very poor)	0%	0
20-39 (poor)	7.2%	24.66
40-59 (unsatisfactory)	21.5%	73.06
60-79 (good)	21.7%	73.97
80-100 (excellent)	29.4%	100
Review Category (based on user's rating)		
negative	0%	0
questionable	3.2%	22.02
neutral	10.7%	73.39
positive	14.6%	100