How to Assess and Rank User-Generated Content on Web?

A Survey on Available Methodologies and Frameworks

Elaheh Momeni
University of Vienna
Vienna, Austria
elaheh.momeni-ortner@univie.ac.at

Claire Cardie
Cornell University
Ithaca, USA
cardie@cs.cornell.edu

Nicholas Diakopoulos
Northwestern University
Evanston, USA
nad@northwestern.edu

ABSTRACT

User-generated content (UGC) on the Web, especially on social media platforms, facilitates the association of additional information with digital resources and online social topics and it can provide valuable supplementary content. However, UGC varies in quality and, consequently, raises the challenge of how to maximize its utility for a variety of end-users, in particular in the age of misinformation. This study aims to provide researchers and Web data curators with answers to the following questions: (1) What are the existing approaches and methods for assessing and ranking UGC? (2) What features and metrics have been used successfully to assess and predict UGC value across a range of application domains? This survey is composed of a systematic review of approaches for assessing and ranking UGC: results obtained by identifying and comparing methodologies within the context of short text-based UGC on the Web. This survey categorizes existing assessment and ranking approaches into four framework types and discusses the main contributions and considerations of each type. Furthermore, it suggests a need for further experimentation and encourages the development of new approaches for the assessment and ranking.

CCS CONCEPTS

• **Information systems → Web searching and information discovery**;

KEYWORDS

User-Generated Content, Assessment and Ranking, Social Media, Crowd-Based, Adaptive and Interactive

ACM Reference Format:
Elaheh Momeni, Claire Cardie, and Nicholas Diakopoulos. 2018. How to Assess and Rank User-Generated Content on Web?: A Survey on Available Methodologies and Frameworks. In *WWW '18 Companion: The 2018 Web Conference Companion, April 23–27, 2018, Lyon, France.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3184558.3186239

1 INTRODUCTION

User-generated content (UGC) on the Web, and on social media platforms in particular, is a vital part of the online ecosystem. UGC is the foremost mechanism for participants to comment on, enhance, and augment online social topics and objects ranging from online videos, images, and audio fragments to more classic news articles. Perhaps not surprisingly, the growing popularity and availability of UGC on the Web has generated exciting new opportunities for actively using information technologies to understand the opinions of others as well as to benefit from the diversity of their knowledge. Moreover, UGC can be employed to aid and improve machine-based processes such as recommendation, retrieval, and search systems. However, managing, hosting, and making sense of this content can be costly, time consuming or even harmful in the age of digital misinformation. As a result, the owners of platforms that host UGC wish to sort and filter contributions according to their value – their truthfulness, credibility, helpfulness, diversity, and so forth – so as to create the best experience possible for viewers. In addition, as the volume of UGC increases, the ability to perform this assessment and ranking automatically becomes increasingly important. However, the task of assessing and ranking UGC is a relatively complex one. This is because (1) UGC, a relatively broad term, can encompass different application domains (e.g.,tags, product reviews, postings in the questions and answers (Q&A) platforms, and comments on digital resources), and each type of UGC has different characteristics; (2) the definition of value varies with regard to different characteristics of application domains and specific tasks in hand (e.g., extracting relevant posts—such as tweets—related to a specific news topic is an important value in microblogging platforms, whereas extracting truthful product reviews is a value in product reviews); and (3) a particular value can be assessed and maximized in different ways due to the different characteristics of UGC. For example, assessing the truthful of product reviews requires different features and methods compared to extracting truthful postings in microblogging platforms. Product reviews can be long, and authors can write false reviews on purpose to deceive the reader. Therefore, the features related to the text of a review are important features to assess the credibility of a review [25, 36]. Instead, postings in microblogging platforms are sometimes short, and features related to texts alone cannot help to assess the credibility of postings. Hence, features need to be included that relate to the activities and backgrounds of authors for a more accurate assessment [9].

This paper is an extended abstract of our published survey paper [30], which aims to explore and shed light on the methods and frameworks for assessment and ranking of different types of UGC by presenting a unifying scheme that includes the commonly used definitions of values and methods for maximizing the value in existing research. A systematic review of existing approaches and methodologies for assessing and ranking UGC are put forward to achieve this goal. The focus is, in particular, on the short, text-based UGC typically found on the Web. In general, the main methods

WWW '18 Companion, April 23–27, 2018, Lyon, France

ACM ISBN 978-1-4503-5640-4/18/04.
https://doi.org/10.1145/3184558.3186239

utilized for assessment and ranking decisions can be categorized in two groups:

Human-centered: This method enables a crowd of end-users, an end-user, or a platform designer to interact with system and specify default rankings or settings. Examples of systems, which use this method are: a system giving each end-user the possibility to assess the quality and vote on the content provided by other end-users or a system enabling each end-user to interact with the platform and rank content with regard to the value in the her mind and task at hand.

Machine-centered: This method utilizes computational methods, in particular machine-learning methods (supervised, semi-supervised, or unsupervised), to develop a ranking and assessment function that learns from the assessment and ranking behavior of the three above mentioned entities, a crowd of humans (external or internal crowd, such as training a classification function, which learns from helpfulness votes of the involved community of end-users), a particular end-user (preferences, background, or online social interactions, which is also called personalization); and the designer (such as providing balanced views of UGC around a political issue on online news platforms).

With regard to these high-level observations, in this study we categorize available frameworks related to assessment and ranking of UGC in the following groups:

Community-based framework: Approaches that fall under this group use the human-centered or machine-centered methods to classify, cluster, and rank UGC based on the majority preferences (or an appropriate metric of preference) of the crowd of humans mainly with regard to a particular definition of value (e.g., truthfulness [25, 39], credibility [6, 8, 9], popularity [20, 21, 42, 46]) and for a particular domain of an application (e.g., forum posts, product reviews). Examples include distinguishing helpful versus non-helpful product reviews [16, 29, 43], classifying useful and non-useful comments on social media objects (e.g., YouTube videos, News articles) [31, 37, 40], or identifying credible postings in online forums.

End-user–based framework: Approaches that use this framework aim to accommodate individual differences in the assessment and ranking of UGC through human-centered or machine-centered methods, thus offering an individual user the opportunity to explore content, specify his or her own notion of value, or interact with the system to modify the display of rankings and assessments in accordance with preferences expressed, behaviors exhibited implicitly, and details explicitly indicated by individual users. Examples include generating a set of content rankers by clustering sub-communities of the user's contact (based on the common content produced) to help users find content more relevant to their interest on their feeds without using explicit user input [7].

Designer-based framework: Approaches that fall under this group mainly use machine-centered methods to encode the software designer's values in the ranking scheme. Examples include an approach that provides balanced political views around an issue [34, 35].

Hybrid framework: The three previous groups of approaches are not necessarily exclusive and often overlap each other. Therefore, there are bodies of assessment and ranking approaches that do not fall explicitly under any of the previous groups. Nevertheless, they take advantage of different categories and are hybrid approaches. Examples include an approach that learns from community behaviors to develop higher-level computational information cues that are useful and effective for adaptive and interactive systems for a single end-user [13] combination of community-based and end-user–based approaches.

2 MAIN OBSERVATIONS

In this section, we discuss the main observations, with a focus on the above mentioned four frameworks concerning three aspects: values, applied methods, and application domains.

Observations for Community-Based Framework We have observed that most of the available research works related to community-based ranking and assessing of UGC utilize machine-centered methods. Nevertheless, default methods utilized by many platforms are human centered. In general, these works focus on:

Motivation and incentivizing: It is important to consider that when human-centered methods are utilized for ranking and assessment systems, participation and contribution in the human-centered methods are basically voluntary, and accordingly, methods to incentivize contributors need to be developed to allocate rewards [17, 44]. Motivation and incentivizing loops can make designing effective human-centered methods challenging in practice.

Bias of judgments of a crowd of end-users: Examining machine-centered methods more closely reveals that for creating a ground truth, many machine-centered assessment approaches completely exclude end-user ratings due to potential biases and wrong effects. Three reasons articulated in the literature for excluding human ratings include the following: (1) Different biases of crowd-based approaches, such as "imbalance voting", "winner circle", and "early bird" [28, 33]. For example, the context or users' awareness of previous votes by a crowd of end-users on particular elements of content (e.g., a product review) needs to be taken into consideration in that it affects the quality of the new vote [12, 41]. Understanding the nature of biases in such human-produced training data is essential for characterizing how that bias may be propagated in a machine-centered assessment approach. (2) A lack of an explicit definition of value that may be requested by the crowd to assess some application domains. For example, many assessment approaches for classification of product reviews with regard to helpfulness as the value have used either a human-centered or a combination of human- and machine-centered approaches. This is because many product review platforms have explicitly defined and asked a crowd of end-users to assess the helpfulness of product reviews. However, most approaches related to assessment of credibility exclude judgments of a crowd of end-users because no platforms have asked them for credibility judgments. (3) Human judgments cannot be as precise as machine-centered judgments in the case of some application domains and values, such as rating the truthfulness of a product review [36].

Different methods for creating ground truths for machine-centered methods: Approaches that exclude judgments of crowds of end-users mainly utilize two methods to create a ground truth or training set: (1) using an external crowd (e.g., using crowd-sourcing platforms) that independently judges content with regard to a particular value [31] and (2) developing their own coding system for collecting independent judgments from a closed set of users. Both of these methods may of course introduce their own sets of biases in the training data [19].

The importance of different dimensions of quality (as a value) varies according to the application domain: With regard to different values that are expected to be maximized, many approaches appear to maximize quality in general, applying a human-centered method as a default ranking method. Nevertheless, with quality being a very general value, some approaches focus on more sophisticated definitions of value and take into consideration different dimensions of quality. In addition, what is defined as value varies with regard to different application domains and specific tasks at hand. Finally, it is observed that most of the recent available approaches focus on maximizing different dimensions of quality in particular truthfulness for microblogging platforms (e.g., Twitter) [10, 15, 39]. This is perhaps due to the very simple and structured characteristics of these platforms. Figure 1 shows focus of number of available approaches for different values for various application domains.

Different machine-centered methods are appropriate for different values and application domains: With regard to the application domain, a more detailed examination leads us to discover that many proposed machine-centered assessment approaches utilize supervised methods. However, when interconnectedness and interdependency between sets of entities in an application domain (e.g., interdependency between Questions, Answers, and Users in a Q&A domain) occur, assessment and ranking approaches mainly utilize semi-supervised learning methods such as co-training or mutually reinforcing approaches [24]. A set of approaches related to community-based frameworks that provide relevant content mainly utilize unsupervised machine learning methods [3, 4]. Nevertheless, as supervised and semi-supervised methods require adequate amounts of labeled data for an accurate training, the development of adaptive machine-centered methods that can be utilized in different application domains is challenging in practice. Therefore, finding a way to optimize the process of labeling and improve the accuracy of hard machine-centered judgments is essential.

Influential features vary for different values and application domains: It is observed that approaches employing machine-centered methods for different application domains use similar sets of content and context features related to three different entities of social media platforms. These three entities are Author, User-Generated Content, and Resource (the media object or topic on which authors generate content). The influence of the features changes with regard to the application domain and definition of the value to be maximized. We group influential features in nine different groups. Figure 2 provides a short overview of the examination of influential features for various values, demonstrated by available research results. A more detailed examination of features leads us to discover that many text- and semantic-based features are important for classifying and clustering UGC in all application domains. Next, similar to the assessment of quality, popularity requires many features to be used in its assessment, and some of the more important ones include authors' activities, background, networks and structures, and propagation and interactive features [29, 38, 46]. These features related to authors' activities and networks also play an important role when assessing credibility, because features simply related to texts cannot help to assess the credibility of postings. However, in the case of assessing spam and deceptive content, authors can write fake reviews written to appear true and deceive the reader [25, 36]. It is worth noting that time-based features play an important role in assessing helpfulness, usefulness, and popularity.

Observations for End-User Framework: A detailed exploration of available approaches for end-user assessment and ranking reveals that most of the available end-user assessment and ranking approaches focus on maximizing different values mainly for two application domains: postings in microblogging platforms and forums. For this framework, the main difference with regard to human- and machine-centered methods, is that human-center (interactive and adaptive) approaches in contrast to machine-centered (personalized) approaches, do not explicitly or implicitly use a user's previous common actions and content to assess and rank the content. However, they provide individual end-users with opportunities to interact with the system and explore the ranked content to find content to match their requirements. Interfaces, created by interactive and adaptive approaches, permit users to browse their feeds more efficiently by providing ready access to all content in a user's feed and enabling users to find content related to their own interests [14]. In particular, some interactive and adaptive ranking solutions have focused on topic based [1, 5, 32] browsing, which groups the comments into coherent topics, creating interfaces that allow users to browse their feeds more efficiently by exploring clusters of topics. Personalized approaches are based on an algorithm that learns from a particular end-user's preferences, background, or online social interactions and connections to personalize the ranking and assessment process [2, 7, 11, 45]. Nevertheless, the number of research results for development of advanced personalized approaches is very low. Furthermore, available approaches are mainly based on the concept of collaborative filtering, which isolates users in information bubbles. Therefore, their views on a particular issue will be influenced and, even worse, more distorted by these filters. This may be difficult to correct. Thus, developing systems and algorithms that provide users more diverse and balanced information may be an interesting challenge.

Observations for Designer-Based Framework: The decision of the platform's designer partially influences the definition of the value for every type of assessment and ranking framework (community based or end-user based). Designers choose definitions for values either because they can be understood and rated by a community or because they can be operationalized for machine-centered methods. Designers can also introduce other objectives into rankings, including desires to optimize more than one value simultaneously. For instance, a systems that encourage users towards more diverse exposure to content [26, 34, 35]. Although rankings seek to optimize a set of objects along a dimension of interest, diversity instead seeks to introduce another dimension or dimensions into that ranking so that a set of content achieves balance across the values of interest. Approaches in this category mainly utilize machine-centered methods and primarily focus on (1) development of alternative measures of diversity for UGC sets, (2) development of algorithms and systems for selecting content sets that jointly optimize diversity of presentation, or (3) development of alternative selection and presentation methods on users' desire to apply their exposure or an aggregation service to diverse opinions.

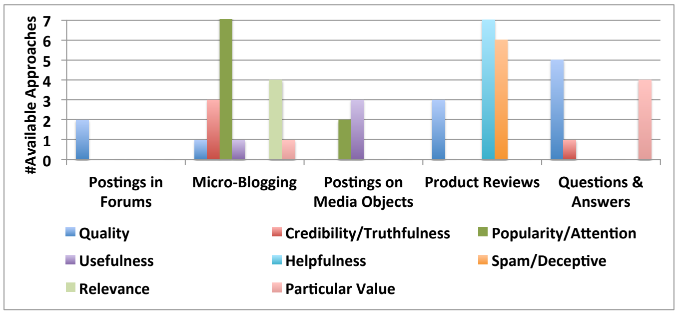

Figure 1: Values that are important and assessed by different application domains

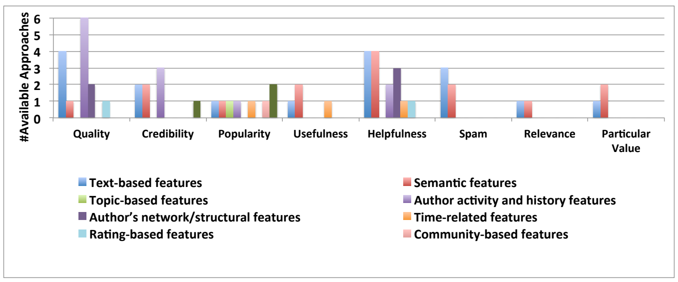

Figure 2: Influential features sets for assessment and ranking of different values

Observations for Hybrid Framework: The advantage of the hybrid framework is that it has different categories and combined approaches. Accordingly, it has high potential for developing more sophisticated and useful techniques. Nevertheless, it has received inadequate attention. Current approaches mainly focus on the following: (1) leveraging a community-based framework for an end-user framework, such as (a) a Web service that provides "neighborhood-specific" information based on Twitter posts for an individual end-user by utilizing activities of a crowd of end-users [13, 22], or (b) a platform that examines how journalists filter and assess the variety of trustworthy tweets found on Twitter and gives an individual end-user (a journalist) the chance to interact with the system and explore a number of computational information cues, trained using a crowd of humans [13], (2) leveraging the end-user framework for the designer-based framework, such as (a) approaches that leverage the patterns of assessment and rank settings by end-users to minimize the cost of changing settings for another end-user [23], (b) a framework that codes designer value by leveraging previous activities of the end-user and diversifies user comments on news articles [18].

3 CHALLENGES AND OPPORTUNITIES

Based on the aforementioned observations and analyses of results, we then list several open issues and limitations of the available approaches. Addressing these issues and limitations creates natural avenues and opportunities for future work:

Bridging the conceptual gap between human-centered and machine-centered approaches receives little attention, triggering many technical challenges. These include how to develop algorithms and methods for mitigating biases of the crowd (e.g., leveraging the blockchain technology) or how to take advantage of semi-supervised learning such as active learning for efficient integration of the crowd into machine-centered approaches.

Maximizing some values related to various dimensions of quality for some application domains receives less consideration. In other words, some dimensions of quality are analyzed only for limited application domain. For example, truthfulness are mainly discussed and analyzed in the domain of microblogging platforms [10, 15, 39] and partially in product reviews [25, 27]. Nevertheless, in the age of misinformation credibility and identifying truthfulness of content may be an important value for other application domains, such as commenting systems for news articles or answers in Q&A platforms. Therefore, it is important to find out to what extent the impact of the influential features for different dimensions of quality vary with regard to various application domains.

The development of methods to incentivize high-quality UGC has not been completed. This triggers challenges such as how advancement of computational methods (e.g.,game-theoretic foundations) can help incentivize high-quality UGC and advanced development of assessment and ranking approaches [44]. For example, for a more accurate incentivizing model, the temporal aspect of UGC may be taken into consideration (a sequential model may be better suited to many UGC environments) [17].

Approaches aiming to accommodate individual differences in the assessment and ranking of UGC and in general end-user-based frameworks have received inadequate attention. In other words, how can we help people make personal assessments of a particular value rather than rely on particular sources as authorities for ground truth? Most of the available approaches rely on particular sources of ground truth and do not enable users to make personal assessments of a particular value. For example, most of the work on identification of helpfulness of product reviews creates and develops prediction models based on a set of majority-agreement labeled reviews. However, helpfulness is a subjective concept that can vary for different individual users, and therefore it is important that systems help individuals make personal assessments of a particular value.

The additional presentation techniques receive scant attention. These include more sophisticated displays of challenging and diverse content for supporting balanced or diverse views surrounding an issue. In line with this challenge, there are bodies of works in the context of information retrieval that maximize diversity among the displayed items so that certain elements are not redundant. However, there is lack of attention given to such work for UGC.

Finally, combined and hybrid approaches deserve more attention. We believe that combined and hybrid approaches have significant potential for further development because they can benefit from the advantages of various frameworks discussed in this article to develop more sophisticated and advanced techniques for assessment and ranking of UGC, such as the development of systems that learn from crowd behaviors to personalize assessment and ranking of content or the development of personalized models for a smaller crowd (geographically or other demographically driven measures that produce individualized/adapted content for a crowd).

REFERENCES

[1] Fabian Abel, Ilknur Celik, Geert-Jan Houben, and Patrick Siehndel. 2011. Leveraging the Semantics of Tweets for Adaptive Faceted Search on Twitter. In *Proceedings of the 10th International Conference on The Semantic Web - Volume Part I (ISWC'11)*. Springer-Verlag, Berlin, Heidelberg, 1–17.

[2] Jisun An, Meeyoung Cha, Krishna Gummadi, Jon Crowcroft, and Daniele Quercia. 2012. Visualizing Media Bias through Twitter. In *Proc. of ICWSM*. AAAI.

[3] Hila Becker, Dan Iter, Mor Naaman, and Luis Gravano. 2012. Identifying content for planned events across social media sites. In *Proceedings of the fifth ACM international conference on Web search and data mining (WSDM '12)*. ACM, 10.

[4] Hila Becker, Mor Naaman, and Luis Gravano. 2011. Selecting Quality Twitter Content for Events.. In *ICWSM*. The AAAI Press.

[5] Michael S. Bernstein, Bongwon Suh, Lichan Hong, Jilin Chen, Sanjay Kairam, and Ed H. Chi. 2010. Eddi: interactive topic-based browsing of social status streams. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology (UIST '10)*. ACM, New York, NY, USA, 303–312.

[6] Jiang Bian, Yandong Liu, Ding Zhou, Eugene Agichtein, and Hongyuan Zha. 2009. Learning to recognize reliable users and content in social media with coupled mutual reinforcement. In *Proceedings of the 18th international conference on World wide web (WWW '09)*. ACM, New York, NY, USA, 51–60.

[7] Matthew Burgess, Alessandra Mazzia, Eytan Adar, and Michael Cafarella. 2013. Leveraging Noisy Lists for Social Feed Ranking. In *The 7th International AAAI Conference on Weblogs and Social Media (ICWSM2013)*. AAAI, Boston, USA.

[8] Kevin R. Canini, Bongwon Suh, and Peter L. Pirolli. 2011. Finding Credible Information Sources in Social Networks Based on Content and Social Structure. In *2011 IEEE Third Int'l Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third Int'l Conference on Social Computing*.

[9] Carlos Castillo, Marcelo Mendoza, and Barbara Poblete. 2011. Information credibility on twitter. In *the 20th international conference (WWW)*. http://dx.doi.org/10.1145/1963405.1963500

[10] Nikan Chavoshi, Hossein Hamooni, and Abdullah Mueen. 2017. Temporal Patterns in Bot Activities. In *Proceedings of the 26th International Conference on World Wide Web Companion (WWW '17 Companion)*.

[11] Jilin Chen, Rowan Nairn, and Ed Chi. 2011. Speak Little and Well: Recommending Conversations in Online Social Streams. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA.

[12] Cristian Danescu-Niculescu-Mizil, Gueorgi Kossinets, Jon Kleinberg, and Lillian Lee. 2009. How opinions are received by online communities: a case study on amazon.com helpfulness votes. In *Proceedings of the 18th international conference on World wide web (WWW '09)*. 10.

[13] Nicholas Diakopoulos, Munmun De Choudhury, and Mor Naaman. 2012. Finding and assessing social media information sources in the context of journalism. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems (CHI '12)*. ACM, 10. http://doi.acm.org/10.1145/2208276.2208409

[14] Siamak Faridani, Ephrat Bitton, Kimiko Ryokai, and Ken Goldberg. 2010. Opinion Space: A Scalable Tool for Browsing Online Comments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 1175–1184.

[15] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2016. The Rise of Social Bots. *Commun. ACM* 59, 7 (June 2016).

[16] Anindya Ghose and Panagiotis G. Ipeirotis. [n. d.]. Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics. *IEEE Transactions on Knowledge and Data Engineering* 23, 10 ([n. d.]).

[17] Arpita Ghosh and Patrick Hummel. 2012. Implementing optimal outcomes in social computing: a game-theoretic approach. In *Proceedings of the 21st international conference on World Wide Web (WWW '12)*. ACM, New York, NY, USA, 10.

[18] Giorgos Giannopoulos, Ingmar Weber, Alejandro Jaimes, and Timos Sellis. 2012. Diversifying User Comments on News Articles. 7651 (2012), 100–113.

[19] F. Maxwell Harper, Daniel Moy, and Joseph A. Konstan. 2009. Facts or friends?: distinguishing informational and conversational questions in social Q&A sites. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.

[20] Liangjie Hong, Ovidiu Dan, and Brian D. Davison. 2011. Predicting popular messages in Twitter. In *Proceedings of the 20th international conference companion on World wide web (WWW '11)*. ACM, New York, NY, USA, 57–58.

[21] Chiao-Fang Hsu, Elham Khabiri, and James Caverlee. [n. d.]. Ranking Comments on the Social Web. In *Proceedings of the 2009 International Conference on Computational Science and Engineering - Volume 04*. IEEE Computer Society, Washington, DC, USA.

[22] Yuheng Hu, Shelly Diane Farnham, and Andrés Monroy-Hernández. 2013. Whoo.ly: facilitating information seeking for hyperlocal communities using social media. In *CHI*.

[23] Travis Kriplean, Jonathan Morgan, Deen Freelon, Alan Borning, and Lance Bennett. 2012. Supporting Reflective Public Thought with Considerit. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 265–274.

[24] Baoli Li, Yandong Liu, and Eugene Agichtein. 2008. CoCQA: co-training over questions and answers with an application to predicting question subjectivity orientation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP '08)*. Association for Computational Linguistics, Stroudsburg, PA, USA.

[25] Jiwei Li, Myle Ott, Claire Cardie, and Eduard Hovy. [n. d.]. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 1566–1576.

[26] Q. Vera Liao and Wai-Tat Fu. 2013. Beyond the Filter Bubble: Interactive Effects of Perceived Threat and Topic Involvement on Selective Exposure to Information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 2359–2368.

[27] Ee-Peng Lim, Viet-An Nguyen, Nitin Jindal, Bing Liu, and Hady Wirawan Lauw. 2010. Detecting Product Review Spammers Using Rating Behaviors. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM '10)*. 939–948.

[28] Jingjing Liu, Yunbo Cao, Chin Y. Lin, Yalou Huang, and Ming Zhou. 2007. Low-Quality Product Review Detection in Opinion Summarization. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*.

[29] Yue Lu, Panayiotis Tsaparas, Alexandros Ntoulas, and Livia Polanyi. 2010. Exploiting social context for review quality prediction. In *Proceedings of the 19th international conference on World wide web (WWW '10)*. 10.

[30] Elaheh Momeni, Claire Cardie, and Nicholas Diakopoulos. [n. d.]. A Survey on Assessment and Ranking Methodologies for User-Generated Content on the Web. *ACM Comput. Surv.* ([n. d.]).

[31] Elaheh Momeni, Claire Cardie, and Myle Ott. 2013. Properties, Prediction, and Prevalence of Useful User-generated Comments for Descriptive Annotation of Social Media Objects. In *The 7th International AAAI Conference on Weblogs and Social Media (ICWSM2013)*. AAAI, Boston, USA.

[32] Elaheh Momeni, Reza Rawassizadeh, and Eytan Adar. 2017. Leveraging Semantic Facets for Adaptive Ranking of Social Comments. In *Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval (ICMR '17)*.

[33] Lev Muchnik, Sinan Aral, and Sean J. Taylor. 2013. Social Influence Bias: A Randomized Experiment. *Science* 341, 6146 (09 Aug. 2013), 647–651.

[34] Sean A. Munson, Stephanie Y. Lee, and Paul Resnick. 2013. Encouraging Reading of Diverse Political Viewpoints with a Browser Widget.. In *ICWSM*, Emre Kiciman, Nicole B. Ellison, Bernie Hogan, Paul Resnick, and Ian Soboroff (Eds.). The AAAI Press. http://dblp.uni-trier.de/db/conf/icwsm/icwsm2013.html#MunsonLR13

[35] Sean A. Munson and Paul Resnick. 2010. Presenting Diverse Political Opinions: How and How Much. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 1457–1466.

[36] Myle Ott, Claire Cardie, and Jeff Hancock. 2012. Estimating the prevalence of deception in online review communities. In *Proceedings of the 21st international conference on World Wide Web (WWW '12)*. ACM, New York, NY, USA, 201–210.

[37] Tim Paek, Michael Gamon, Scott Counts, David Maxwell Chickering, and Aman Dhesi. 2010. Predicting the Importance of Newsfeed Posts and Social Network Friends.. In *AAAI*. AAAI Press.

[38] Matthew Rowe, Sofia Angeletou, and Harith Alani. 2011. Predicting Discussions on the Social Semantic Web. In *Extended Semantic Web Conference*.

[39] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake News Detection on Social Media: A Data Mining Perspective. *SIGKDD Explor. Newsl.* 19, 1 (Sept. 2017), 22–36.

[40] Stefan Siersdorfer, Sergiu Chelaru, Wolfgang Nejdl, and Jose San Pedro. 2010. How useful are your comments?: analyzing and predicting youtube comments and comment ratings. In *Proceedings of the 19th international conference on World wide web (WWW '10)*. ACM, 10.

[41] Ruben Sipos, Arpita Ghosh, and Thorsten Joachims. 2014. Was This Review Helpful to You?: It Depends! Context and Voting Patterns in Online Content. In *Proceedings of the 23rd International Conference on World Wide Web (WWW '14)*.

[42] Gabor Szabo and Bernardo A. Huberman. 2010. Predicting the popularity of online content. *Commun. ACM* 53, 8 (Aug. 2010), 80–88.

[43] Oren Tsur and Ari Rappoport. 2009. RevRank: A Fully Unsupervised Algorithm for Selecting the Most Helpful Book Reviews. In *International Conference on Weblogs and Social Media (ICWSM09)*.

[44] Don Turnbull. 2007. Rating, Voting & Ranking: Designing for Collaboration & Consensus. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems (CHI EA '07)*. ACM, New York, NY, USA, 2705–2710.

[45] Ibrahim Uysal and W. Bruce Croft. 2011. User oriented tweet ranking: a filtering approach to microblogs.. In *CIKM*, Craig Macdonald, Iadh Ounis, and Ian Ruthven (Eds.). ACM, 2261–2264. http://dblp.uni-trier.de/db/conf/cikm/cikm2011.html#UysalC11

[46] Claudia Wagner, Matthew Rowe, Markus Strohmaier, and Harith Alani. 2012. What Catches Your Attention? An Empirical Study of Attention Patterns in Community Forums. In *ICWSM*.