A Conversationalist Approach to Information Quality in Information Interaction and Retrieval

Workshop contribution for IQIIR 2022

Frans van der Sluis
Department of Communication
University of Copenhagen
Copenhagen, Denmark
frans@hum.ku.dk

ABSTRACT

Rather than using (proxies of) end user or expert judgment to decide on the ranking of information, this paper asks whether conversations about information quality might offer a feasible and valuable addition for ranking information. We introduce a theoretical framework for information quality, outlining how information interaction should be perceived as a conversation and quality be evaluated as a conversational contribution. Next, an overview is given of different systems of social alignment and their value for assessing quality and ranking information. We propose that a collaborative approach to quality assessment is preferable and raise key questions about the feasibility and value of such an approach for ranking information. We conclude that information quality is an inherently interactive concept, which involves an interaction between users of different backgrounds and in different situations as well as of quality signals on users' search behavior and experience.

1 INTRODUCTION

The Internet, and in particular Web 2.0, was initially thought to democratize knowledge. Online information bypasses traditional gatekeepers to knowledge, which allows more people to raise their voices and gain an opportunity to be heard [30]. Numerous social justice campaigns on social media illustrate that this has indeed, to some extent, been the case (e.g., #MeToo, #ferguson). At the same time, however, the responsibility for assessing the quality of information shifted increasingly to end users. Societal and academic discussions have pointed at a perceived lack of quality in both social media and search (e.g., the Google-Holocaust case [15]) as well as users' inability to recognise it [22]. It proves challenging for an individual user to assess the quality of information: Users

typically lack the ability or motivation to properly assess the quality of each document they encounter [22, 26, 29].

Search and filtering systems rely on (proxies of) both end user and expert judgment in their ranking of information. They include subjective quality signals derived from end users, including implicit feedback from clicks, likes, flags, and shares as well as explicit feedback from liking, flagging, and sharing. They furthermore include expert-style assessments of quality and/or relevance in the optimization of ranking algorithms. For example, Google employs expert readings of documents based on an objective checklist of qualities [14, sections 1-11] to optimize their ranking algorithm. However, experts tend to disagree [23, 50], lack domain expertise [17], and promote authoritative viewpoints [27] whereas an over-reliance on user feedback can promote certain information qualities over others, such as prioritizing popularity over veracity [e.g., popularity bias; 32]. This prioritization of authority and popularity by and large sustains existing authoritative frameworks, gatekeepers, and economic arrangements [30]. Those voices that were successful before are so again through contemporary ranking mechanics.

The two sources of quality signals have subjectivist and positivist tendencies. Positivism assumes the existence of an objective and singular reality that is independent of researchers. Here, expert assessment is generally needed to determine document quality and/or relevance, which is seen as objective and neutral [18]. Individual subjectivism assumes that there are as many realities as there are individuals and denies the existence of a shared truth. Here, relevance and/or quality is an individual, subjective experience that needs end-user assessment. Neither perspective assumes the existence of subjective, multiple realities that are nevertheless to some extent shared. Such interpretivist assumptions rather require a conversationalist approach to quality [25, 26], where a ranking is based on a collaborative assessment of quality that compares conflicting realities. It is clear that the subjectivist and objectivist approaches to ranking have limitations - where system developers currently need to strike a balance between existing authoritative frameworks and the risk for a "*cascade of misinformation*" [30, p. 20]. However, whether an alternative, conversationalist approach is feasible or valuable for ranking information is (yet) to be explored.

A precise implementation of a conversationalist approach to information quality for information interaction and retrieval is far from specified or specifiable. Rather, with this

preliminary paper we outline definitions and questions for discussion. In particular, we will address the following topics:

Quality and qualities What do we understand by information quality and how does it compare to the concept of relevance?

Ranking mechanisms What are current methods that include some form of conversations and agreement between users in ranking information?

Conducive conversations What are conversations about qualities and can they be afforded for ranking?

2 QUALITY AND QUALITIES

Information quality has been approached from a mostly pragmatic perspective. Over time dozens of checklists have been proposed, focusing on aspects of quality such as [14, 28]:

i) Information criteria such as accuracy, currency, usefulness, and importance;

ii) Source criteria such as trustworthiness, credibility, reliability, authoritativeness, and;

iii) Technical criteria such as findability, accessibility, and speed.

In addition, over 452 of empirical studies investigated users' perception of information quality online [13]. Most of the models and lists overlap. For example, criteria for information quality can similarly be formative factors for credibility [e.g., MAIN model 3, 44] as well as function as relevance criteria [38]. Even though these pragmatic approaches lead to workable lists of qualities, they consider quality to be an intrinsic property of the information. In doing so, these studies typically don't take essence into defining quality, rather noting that quality is an "*elusive concept*" [11].

Alternative conceptualizations of quality originate from the study of written communication. Central to these theories is that they consider information (pieces) as conversational acts. This means that any piece of information cannot be seen separately from its interlocutors, being both the authors and the searchers that consume the information as well as the context of the act. Seen as conversational acts, the quality of information is not simply regarded as a property of the information, such as correctness, but must be understood in concert with the situation of the searcher. What constitutes a good conversation can subsequently be considered from particularly two frameworks: speech act theory and the cooperative principle.

Habermas' speech act theory postulates that, in doing a speech act, a speaker raises three mutually irreducible types of validity claims: a claim to truth, to normative rightness, and to sincerity [12]. A claim to truth can be considered in line with a scientific perspective on quality, where the legitimacy of validity claims stems from an inter-subjective agreement on the justification procedure (e.g., methodology). A claim to normative rightness cannot be right or wrong, but rather appropriate or not. It is considered as an application of social-cultural norms to a speech situation and evaluates whether that what is said is context-appropriate. A claim to sincerity finally infers that a speech act is genuine: That it is a sincere expression of the speaker's mental state. Each of these claims can be assessed and challenged through *discourse*. Truth claims are justified through factual evidence and theoretical discourse about states of the world. Normative rightness claims are justified in practical discourse about the validity of the (moral/social) norms in question. Even sincerity claims are justifiable, although not by providing reasons but by one's subsequent actions in a aesthetic-existential discourse.

Grice's cooperative principle states that speakers (authors) must try to contribute meaningful, productive utterances to further the conversation. The cooperative principle postulates four maxims that a good conversation needs to uphold: *i)* Quantity, offering neither too little nor too much information; *ii)* Quality, offering a claim to truth for which the sender has adequate evidence; *iii)* Relation, being of relevance to the receiver; *iv)* Manner, speaking briefly and orderly whilst avoiding ambiguity and obscurity. These maxims, however, need to be considered within the messy nature of language and communication. They can never be uphold fully, but speakers are nonetheless asked to be as precise as possible. In comparison to Habermas, these maxims focus more on the relation between between sender and receiver, author and searcher. They are more pragmatic in nature and depend primarily on the searcher's situation. For example, information does not necessarily need to be up to date with the latest scientific developments in favor of its appropriateness for a (e.g., novice) target audience. And, what is good quality in one situation might not be so in another. The quality of information consequently depends on a complex relation between a sender's intentions, intertextual knowledge, socio-cultural facts, and the reader's activities and interests [26].

This definition of quality aligns it closely with the concept of relevance. Relevance is an inherently pragmatic concept that, similar to the presented conceptualizations, considers the relation between information and its consumer. Both in conceptualization and in practice, the concept of relevance includes a notion of quality. In contemporary conceptual models, relevance is considered equivalent to usefulness with respect to a task [36]. In practice, Web search typically prioritizes the precision of the documents retrieved where the goal is to retrieve documents good enough to fulfill a user's search query. In both cases, relevance implicitly subsumes a notion of quality: presumably, higher quality documents will also be more useful. Nonetheless, we identify two key differences between the two concepts. Firstly is that relevance, at least from a linguistic perspective, concerns the pragmatic dimension of information and less so its semantic or syntactic dimensions. This means that, for example, information can be relevant without being true [51]. Secondly we assert that relevance is more concerned with the potentiality of information, whilst quality is more concerned with its value conditional upon a conversational situation. A document that is of relevant *might* be of value, whereas a high-quality document *ought* to be of value.

From these conceptualizations of quality it follows that quality is a complex notion that is to be determined publicly and socially in shared forums [26]. At the same time, we

observe that the academic discussion is focused increasingly on users' (inaccurate) perception of credibility rather than on the study and development of such forums. Credibility denotes the factors that make a recipient find a message (or its source) credible [3, 44]. We trust documents that were produced in what we believe to be an appropriate way: Created by a trusted institution, written by people with credentials, or that went through a process that we deem reliable [30]. Social markers such as names and credentials lend a source credibility, but also reinforce the existing division of epistemic labor. They do not substantially allow for new voices beyond traditional arrangements that lend these markers. On the other hand, quality, and in particular the assessment thereof, can alternatively lend a message credibility. Given trust in the assessment procedure, an individual can rely on the outcomes of the quality assessment to gauge for its credibility .

3   RANKING MECHANISMS

The preceding puts conversations central to quality. Quality is defined in terms of the quality of a (written) conversation between author and searcher [26]. It furthermore puts forward the need for conversations to assess the complex relations between interlocutors, intertextual knowledge, and socio-cultural norms [26] as well as to challenge implicit claims to quality [12]. Given this arguable essentiality of conversations to quality and the assessment thereof, we explore whether some form of discourse can be afforded for ranking. Even though not many ranking systems currently incorporate discursive elements, already a fair degree of communication and even cooperation between users is common.

A useful framework for thinking about cooperation and collaboration is offered by Shirky [39, ch. 2]. This framework differentiates between four trins of shared social activity online [25]:

Sharing Sharing is the most accessible, least demanding form of group activity. It operates mostly in a take-it-or-leave-it fashion, which allows maximum individual freedom and minimal complications from group alignment. This is the most common mechanism online. Information systems typically aggregate individual feedback, such as 'likes' or tags. These aggregates foster a certain form of awareness of what others have liked or tagged, but not an agreement.

Cooperation Cooperation is a more demanding form of group activity as it involves some synchronization between people in a group, which in turn creates a form a group identity. Its most straightforward instance is a conversation. Conversations enforce an alignment, but not necessarily an agreement, as each individual still engages to advance their own personal goals.

Collaboration Collaboration is characterized by a shared goal. This goal can compete with individual goals whilst it enforces some degree of collective decision making. Wikipedia is a prototypical example of collaborative work which assures the co-creation of a single page per subject. Collaborative work can be more valuable, but takes more energy than aggregation as most decisions need to be negotiated.

Collective action Collective action is the "*hardest kind of group effort, as it requires a group of people to commit themselves to undertaking a particular effort together, and to do so in a way that makes the decision of the group binding on the individual members*" [39, p. 50]. This type of organization is atypical online, but rather found in for example unions where the members abide to the collective agreements.

Sharing actions are commonly used for ranking information. Democratic indexing and social tagging systems (folksonomies) are prototypical examples of this application. In it, users assign labels (tags) to documents in the system, where a document representation is formed by the aggregated labels [33]. Liking and voting mechanisms similarly operate at the individual sharing level. Likes and votes are commonly used to change the ranking of information, for example of posts on social media or of answers on question-answer systems [41]. These mechanisms allow users to exert some degree of (indirect) control over a ranking whilst these aggregates foster a certain form of communication between users by making users aware of what others have tagged or liked. They do not, however, support a conversation or agreement over how resources should be indexed and ranked, nor do they show users which considerations went into a particular ranking of information [10]. For example, folksonomies are characterized by a proliferation of tags whilst an individual's reasoning underlying a tagging or voting action remains opaque to others [33].

Cooperative systems are similarly common on the Internet. Reviews, comments, and forums all allow users to engage in a (form of) conversation with each other. The resulting user-generated content can subsequently be used to augment a document representation for indexing. Some examples exists where reviews are integrated in a search index and benefit subsequent performance of book searches [20] or where anchor text or link context are used to represent documents [7, Section 4.5]). In addition, user-generated content frequently describes some judgment of quality. This is intended for the case of reviews, but can also be observed on forums and in comments. Users actively discuss the quality of information on public fora [37] and talk pages (Wikipedia). Nevertheless, forum and comment threads do not promote a collaborative grounding between its participants. Initial findings even indicate that online discussions about information quality can be characterized as disputational talk [37] rather than leading to a consensus.

Collaborative systems are commonly seen for co-creating knowledge and software as well as for knowledge organization. Prototypical examples are:

Wikipedia , where editors discuss and decide on the quality of a document in a shared goal of improving the document [42, 43].

Question-answer sites , where contributors have the shared goal of answering questions posed by visitors and selecting the best answer to a question [41].

Open source projects , where contributors co-create a software program on code sharing platforms like Github. Often, these projects are characterized by a central ownership of the project [8].

Web directories , where volunteers categorize Web sources according in a large directory.

In all these systems, often detailed discussions occur about how to achieve the shared goal. Individual interests and perspectives differ and need to be aligned in order to proceed on the shared goal. These discussions are supported by various commenting tools and user roles, where some users are owners or can be elected as editors to support conducive discussions towards the shared goal. It remains hard, however, to appropriate these discursive elements for assessing and ranking information. Question-answer sites rank the best answer first on a per-question basis, but do not scale to the flexibility of ad-hoc queries. And, free text search and natural language indexing have prevailed as contemporary method to access online resources, replacing more effortful social systems of organization like Web directories.

4 TOWARDS CONDUCIVE CONVERSATIONS

The preceding exposition of different levels of collaboration indicated that user-generated contributions are mostly used for ranking at the sharing level with some further cases at the cooperation level. A step further on the ladder of social alignment would come if such a system would support "*back-and-forth talking and editing*" [39, p. 52]. The preceding furthermore showed the lack of success of cooperative and collaborative systems for indexing documents. Rather, ranking information is first of foremost based on natural language indexes coupled with tags or likes shared by users. These indexes are ideal for inferring relevance as they closely reflect the meaning of words whilst handling the messiness of language. Nonetheless, collaborative social systems seem particularly suitable to denote document quality complementary to its relevance. The conceptualizations of quality suggest that quality is to be determined publicly and socially in shared forums in order to capture and evaluate the complex web of interlocutors, intertextual knowledge, and socio-cultural norms. This raises the question as to whether a collaborative assessment of a document's quality and qualities is possible. We will explore this possibility through several (open) questions.

Can users collaboratively assess quality? Savolainen [37] showed that users already, without a shared goal, engage in quality assessments on public fora. Of the reviewed messages, 20.5% contained explicit judgments of information quality. These judgments were both positive and negative and covered

criteria such as usefulness, correctness, specificity, reputation, expertise, and honesty. Other studies looked into the agreeableness of credibility factors, noting that most factors receive a fair level of inter-rater agreement amongst users as well as amongst experts [3, 19]. These findings indicate that users can discuss and agree upon the assessment of qualities. Nonetheless, none of these studies made quality assessment a shared goal that enforces some form of alignment or agreement. It is likely that certain boundary conditions exist to the agreeableness of quality assessments. Quality is thought to be bounded by searchers' situations, expertise and beliefs as well as prevalent socio-cultural norms - which should surface in conversations but likely limits the degree of alignment that is feasible. Furthermore, the accuracy of user-based assessments is unknown. Kąkol et al. [19] indicates that lay users overall give highly positive ratings, whilst experts tend to be critical. In collaborative and dialectical situations, on the other hand, the necessary alignment might increase the accuracy of assessments.

Can assessments be afforded for ranking? The complexity of the notion of quality makes it hard to be afforded for ranking. In order to rank documents requires, eventually, to simplify a range of factors onto a single, sortable scale. A simple solution is to resort to an explicit feedback mechanism that coerces a complicated notion of quality into a single 'like' button or voting mechanism. The resulting feedback offsets text-based relevance estimates in deciding on a ranking. A method that remains closer to this complexity is to augment a document representation with the contents of assessments [20]. This method allows for searching not just in the contents of the document but also in the contents of the assessments, which can be particularly beneficial to delineate searchers' situations. A third and well-studied method is the use of tags for ranking [33]. When used as a free, unlimited vocabulary of tags, the proliferation and ambiguity of tags and lack of agreement amongst users about tag assignments is ideal for assigning meaning to documents. This ambiguity is, most likely, not suitable for denoting qualities. A good quality label is understandable and useable for retrieval purposes (end-user perspective), has epistemic or cognitive merit [34], and has a reasonable degree of agreeableness, all which favor some form of a controlled vocabulary of qualities. Even though these three methods illustrate the feasibility of incorporating aspects of quality assessments into rankings, their value in terms of ranking precision and user satisfaction as well as their scalability in terms of collection coverage and recall are unknown.

Can quality conversations engage users? Several studies show the dominance of ranking position over other cues in guiding search behavior. Users typically select search results from the top ranks, also after a results list was altered to rank lower quality results first (i.e., top-rank heuristic) [1, 31]. Users furthermore only allow for a limited influence of quality cues over ranking position on their behavior [16, 47]. Notwithstanding, an intermediate levels of uncertainty sparks curiosity. Feelings of uncertainty about, amongst others, the

completeness, coherence, or accuracy of one's knowledge [35] drive seeking behavior [5, 24] and cognitive engagement [46]. These findings suggest that changes in ranking position exert a strong influence on behavior, whilst uncertainty can be a strong motivator of user engagement. This indicates that quality signals, as part of a ranking mechanism or search result context [40], can positively influence users' engagement with and users' affective experience during information search [48, 49]. Of equal importance is whether users want to engage in quality discussions. There seems limited precedence to answer this question besides documented success stories of collective intelligence [45]. Cunningham's Law might offer some valuable insight though: "*The best way to get the right answer on the Internet is not to ask a question; it's to post the wrong answer.*"

5 DISCUSSION

The preceding sections introduced key conceptualizations of information quality, surveyed ranking mechanisms with some degree of social alignment, and raised questions on the feasibility and value of quality assessments for ranking information. These contributions offered a first approximation of what a a conversationalist approach to information quality in information interaction and retrieval would entail.

Even though this preliminary paper was initiated against the backdrop of information retrieval, the main questions raised are not about precision or recall. They are rather about users' ability to assess information quality collaboratively and (end) users' engagement with such assessments. This proposes information quality as an inherently interactive concept, as an interaction between users and with end users: Between users of different backgrounds and in different situations and with users on the cognitive-emotional effects of assessments on search behavior and experience.

Precision and recall will nonetheless be crucial for the likelihood of a conversationalist approach. An increased approximation of quality has the potential to improve precision. Coverage will be a limiting factor, however, both in research and in practice. Creating a test collection will demand a large amount of situated quality assessments, something only feasible with modern crowd sourcing [2, 9] solutions and properly set up tasks [6]. In practice, the limited coverage will mean such an approach will suffer from missing data, delayed uptake of new documents, and the cold start problem [21].

Whether collaborative assessments of quality could increase quality whilst simultaneously include new voices is, of course, currently too far-fetched to answer. Rather, market competition favors those solutions that offer speed and ease. Social systems, in particular at the level of collaboration, are slow and effortful. Practical feasibility is not the only possible impact, though. Empirical work on quality can redirect attention from a negative perspective of fake news, shallow novelty, and similar criticisms [4] towards a positive understanding of what information quality means. A systematic study of information quality can map boundary conditions to information quality and further our understanding of what

information quality means to different people in different situations. Eventually this improved understanding can inform the academic and societal discussion to consider in detail which qualities are and should be promoted when.

REFERENCES

[1] Eugene Agichtein, Eric Brill, and Susan Dumais. 2006. Improving web search ranking by incorporating user behavior information. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* (Seattle, Washington, USA). New York, NY, USA: ACM, 19–26.

[2] Omar Alonso. 2019. *The Practice of Crowdsourcing* (1 ed.). Synthesis Lectures on Information Concepts, Retrieval, and Services, Vol. 66. Morgan & Claypool Publishers. 151 pages.

[3] Alyssa Appelman and S. Shyam Sundar. 2016. Measuring Message Credibility. *Journalism & mass communication quarterly* 93, 1 (mar 2016), 59–79.

[4] D. Bawden and L. Robinson. 2008. The dark side of information: overload, anxiety and other paradoxes and pathologies. *Journal of Information Science* 35, 2 (21 nov 2008), 180–191.

[5] Daniel E. Berlyne. 1966. Curiosity and Exploration. *Science* 153, 3731 (1966), 25–33.

[6] Pia Borlund. 2016. A study of the use of simulated work task situations in interactive information retrieval evaluations. *Journal of Documentation* 72, 3 (9 may 2016), 394–413.

[7] Bruce Croft, Donald Metzler, and Trevor Strohman. 2009. *Search Engines: Information Retrieval in Practice* (1 ed.). Pearson, Boston. 552 pages.

[8] Paul B. de Laat. 2010. How can contributors to open-source communities be trusted? On the assumption, inference, and substitution of trust. *Ethics and information technology* 12, 4 (dec 2010), 327–341.

[9] Gianluca Demartini, Djellel Eddine Difallah, Ujwal Gadiraju, and Michele Catasta. 2017. An Introduction to Hybrid Human-Machine Information Systems. *Foundations and Trends®in Web Science* 7, 1 (2017), 1–87.

[10] Melanie Feinberg. 2006. An examination of authority in social classification systems. *Advances in Classification Research Online* 17, 1 (7 oct 2006).

[11] Noa Fink-Shamit and Judit Bar-Ilan. 2008. Information quality assessment on the Web - an expression of behaviour. *Information Research* 13, 4 (4 dec 2008).

[12] Barbara Fultner. 2014. Communicative action and formal pragmatics. In *Jürgen habermas: key concepts*, Barbara Fultner (Ed.). Acumen Publishing Limited, Durham, 54–73.

[13] Maryam Ghasemaghaei and Khaled Hassanein. 2016. A macro model of online information quality perceptions: A review and synthesis of the literature. *Computers in human behavior* 55 (feb 2016), 972–991.

[14] Google. 2020. Search Quality Rating Guidelines. Retrieved from https://static.googleusercontent.com/media/guidelines.raterhub.com/en//searchqualityevaluatorguidelines.pdf on Feb 11, 2021..

[15] The Guardian. 2017. Retrieved from https://www.theguardian.com/technology/2017/mar/15/google-quality-raters-flag-holocaust-denial-fake-news on Mar 4, 2021.

[16] Alexander Haas and Julian Unkel. 2017. Ranking versus reputation: perception and effects of search result credibility. *Behaviour & information technology* 36, 12 (2 dec 2017), 1285–1298.

[17] Birger Hjørland. 2005. Domain Analysis: A Socio-Cognitive Orientation for Information Science Research. *Bulletin of the American Society for Information Science and Technology* 30, 3 (31 jan 2005), 17–21.

[18] Birger Hjørland. 2009. The foundation of the concept of relevance. *Journal of the American Society for Information Science and Technology* 61, 2 (20 nov 2009), 217–237.

[19] Michał Kąkol, Michal Jankowski-Lorek, Katarzyna Abramczuk, Adam Wierzbicki, and Michele Catasta. 2013. On the subjectivity and bias of web content credibility evaluations. In *Proceedings of the 22nd International Conference on World Wide Web - WWW '13 Companion*. ACM Press, New York, New York, USA, 1131–1136.

[20] Marijn Koolen. 2014. *Advances in information retrieval*. Lecture notes in computer science, Vol. 8416. Springer International Publishing, Cham.

[21] Xuan Nhat Lam, Thuc Vu, Trong Duc Le, and Anh Duc Duong. 2008. Addressing cold-start problem in recommendation systems. In *Proceedings of the 2nd international conference on Ubiquitous information management and communication - ICUIMC '08*. ACM Press, New York, New York, USA, 208.

[22] Eun-Ju Lee and Soo Yun Shin. 2019. Mediated misinformation: questions answered, more questions to ask. *American Behavioral Scientist* 65, 2 (23 aug 2019), 259–276.

[23] Chloe Lim. 2018. Checking how fact-checkers check. *Research & Politics* 5, 3 (jul 2018), 205316801878684.

[24] Jordan Litman. 2005. Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition & Emotion* 19, 6 (2005), 793–814.

[25] Jens-Erik Mai. 2011. Folksonomies and the new order: authority in the digital disorder. *Knowledge Organization* 38, 2 (2011), 114–122.

[26] Jens-Erik Mai. 2013. The quality and qualities of information. *Journal of the American Society for Information Science and Technology* 64, 4 (apr 2013), 675–688.

[27] Robert Mejia, Kay Beckermann, and Curtis Sullivan. 2018. White lies: a racial history of the (post)truth. *Communication and Critical/Cultural Studies* 15, 2 (3 apr 2018), 109–126.

[28] Marc Meola. 2004. Chucking the Checklist: A Contextual Approach to Teaching Undergraduates Web-Site Evaluation. *portal: Libraries and the Academy* 4, 3 (2004), 331–344.

[29] Miriam J. Metzger, Andrew J. Flanagin, and Ryan B. Medders. 2010. Social and heuristic approaches to credibility evaluation online. *Journal of Communication* 60, 3 (19 aug 2010), 413–439.

[30] Nicola Mößner and Philip Kitcher. 2017. Knowledge, democracy, and the internet. *Minerva* 55, 1 (mar 2017), 1–24.

[31] Bing Pan, Helene Hembrooke, Thorsten Joachims, Lori Lorigo, Geri Gay, and Laura Granka. 2007. In google we trust: users'decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication* 12, 3 (apr 2007), 801–823.

[32] Bruno Pradel, Nicolas Usunier, and Patrick Gallinari. 2012. Ranking with Non-random Missing Ratings: Influence of Popularity and Positivity on Evaluation Metrics. In *Proceedings of the Sixth ACM Conference on Recommender Systems (RecSys '12)*. ACM, New York, NY, USA, 147–154. event-place: Dublin, Ireland.

[33] Pauline Rafferty. 2018. Tagging. *Knowledge Organization* 45, 6 (2018), 500–516.

[34] Nicholas Rescher. 2013. Chapter 1: epistemic merit. In *Epistemic Merit: And other Essays on Human Knowledge*. DE GRUYTER, 1–10.

[35] Derek D. Rucker, Zakary L. Tormala, Richard E. Petty, and Pablo Briñol. 2014. Consumer conviction and commitment: An appraisal-based framework for attitude certainty. *Journal of Consumer Psychology* 24, 1 (jan 2014), 119–136.

[36] Tefko Saracevic. 2007. Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: Nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology* 58, 13 (2007), 1915–1933.

[37] Reijo Savolainen. 2011. Judging the quality and credibility of information in Internet discussion forums. *Journal of the American Society for Information Science and Technology* 62, 7 (jul 2011), 1243–1256.

[38] Linda Schamber, Michael B. Eisenberg, and Michael S. Nilan. 1990. A re-examination of relevance: toward a dynamic, situational definition. *Information Processing & Management* 26, 6 (1990), 755 – 776.

[39] Clay Shirky. 2008. *Here Comes Everybody Power of Organizing Without Organizations (Hardcover, 2008)*. Penguin Press. 327 pages.

[40] Catherine L. Smith and Soo Young Rieh. 2019. Knowledge-Context in Search Systems: Toward Information-Literate Actions. In *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval - CHIIR '19*. ACM Press, New York, New York, USA, 55–62.

[41] Ivan Srba and Maria Bielikova. 2016. A comprehensive survey and classification of approaches for community question answering. *ACM Transactions on the Web* 10, 3 (16 aug 2016), 1–63.

[42] Besiki Stvilia, Michael B. Twidale, Les Gasser, and Linda C. Smith. 2005. *Information Quality in a Community-based Encyclopedia*. 101–113.

[43] Besiki Stvilia, Michael B. Twidale, Linda C. Smith, and Les Gasser. 2008. Information quality work organization in wikipedia. *Journal of the American Society for Information Science and*

Technology 59, 6 (2008), 983–1001.

[44] S. Shyam Sundar. 2008. The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility. *Digital Media, Youth, and Credibility* (2008), 73–100.

[45] Shweta Suran, Vishwajeet Pattanaik, and Dirk Draheim. 2020. Frameworks for collective intelligence. *Comput. Surveys* 53, 1 (29 may 2020), 1–36.

[46] Zakary L Tormala. 2016. The role of certainty (and uncertainty) in attitudes and persuasion. *Current opinion in psychology* 10 (aug 2016), 6–11.

[47] Julian Unkel and Alexander Haas. 2017. The effects of credibility cues on the selection of search engine results. *Journal of the Association for Information Science and Technology* 68, 8 (aug 2017), 1850–1862.

[48] F. Van der Sluis, E. L. Van den Broek, and E. M. A. G. Van Dijk. 2010. Information Retrieval eXperience (IRX): Towards a Human-Centered Personalized Model of Relevance. In *Proceedings of the Workshop on Web Information Retrieval Support Systems at the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, Vol. 3. 322–325.

[49] F. Van der Sluis, E. M. A. G. Van Dijk, and E. L. Van den Broek. 2010. Aiming for User Experience in Information Retrieval: Towards User-Centered Relevance (UCR). In *SIGIR 2010: ACM Proceedings of the 33rd International Conference on Research and Development in Information Retrieval* (Geneva, Switzerland), H-H. Chen, E. N. Efthimiadis, J. Savoy, F. Crestani, and S. Marchand-Maillet (Eds.). ACM, New York, USA, 924–924.

[50] Ellen Voorhees. 2002. The Philosophy of Information Retrieval Evaluation. In *Evaluation of Cross-Language Information Retrieval Systems*. Springer Berlin / Heidelberg, 143–170.

[51] T.D. Wilson, N.J. Ford, D. Ellis, A.E. Foster, and A. Spink. 2002. Information seeking and mediated searching: Part 2. Uncertainty and its correlates. *Journal of the American Society for Information Science and Technology* 53, 9 (2002), 704–715.