

Fake News and Aggregated Credibility: Conceptualizing a Co-Creative Medium for Evaluation of Sources Online

Montathar Faraon, Kristianstad University, Sweden

 <https://orcid.org/0000-0002-9740-2609>

Agnieszka Jaff, Kristianstad University, Sweden

Liegi Paschoalini Nepomuceno, Kristianstad University, Sweden

Victor Villavicencio, Art-O-Matic AB, Sweden

ABSTRACT

The accelerated spread of fake news via the internet and social media such as Facebook and Twitter have created a debate concerning the credibility of sources online. Assessing the credibility of these sources is generally a complex task and cannot solely rely on computer-based algorithms as evaluation still requires human intelligence. The research question guiding this article deals with the conceptualization of a theoretically anchored concept of a participatory and co-creative medium for evaluation of sources online. The concept-driven design research methodology was applied to address the research question, which consisted of seven activities that unify design and theory. The result of this article is a proposed concept that aims to support the assessment of the credibility of sources online using crowdsourcing as an approach for evaluation. The practical implications of the proposed concept could be to constrain the spread of fake news, strengthen online democratic discourse, and potentially improve the quality of online information.

KEYWORDS

Aggregated Credibility, Co-Creation, Community, Concept-Driven Design Research, Crowdsourcing, Fake News, Participation, Sources

1. INTRODUCTION

The assessment of the credibility of online sources (e.g., creators of and to content) is generally more complicated than in traditional media because “of the multiplicity of sources [such as contributors] embedded in the numerous layers of online dissemination” (Sundar, 2008, p. 74). Social media, such as Facebook and Twitter, have enabled information to flow faster than ever before, consequently increasing the speed in which false information can spread online (Allcott & Gentzkow, 2017). False information, often referred to as fake news, may relate to a vast array of phenomena, ranging from hoaxes to sensationalism, see Tandoc Jr, Wei Lim, and Ling (2018) for a review. In this article, we

DOI: 10.4018/IJACI.20201001.0a1

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

cautiously adopt the term ‘fake news’ as “the online publication of intentionally or knowingly false statements of fact” (Klein & Wueller, 2017, p. 6).

Fake news is regarded by the World Economic Forum (WEF) to be one of the biggest challenges to contemporary societies (Vicario et al., 2016). With the increasing spread of fake news, it has become necessary to identify ways to validate the information that users find online. Research has shown that social media shape our memory in a way where people tend to conform to a majority recollection, even if it proves to be wrong (Spinney, 2017). The challenge has become that people may not always be aware of what information should be regarded as credible or non-credible, especially with regards to deep fake video clips, i.e., making a person appear to say or do something they did not (Maras & Alexandrou, 2019). Information that is not credible may lead to inaccurate beliefs and result in misperception, which could undermine democratic decision-making processes (Allcott & Gentzkow, 2017; Hameleers & van der Meer, 2019).

Because it has become difficult for individuals to assess and evaluate information they encounter online (Allcott & Gentzkow, 2017; Metzger, Flanagin, Eyal, Lemus, & McCann, 2003; Robins & Holmes, 2008), a number of online services (e.g., Snopes, Hoax-Slayer, Politifact, Factcheck) have emerged in an effort to evaluate the credibility of online information through a process that involves fact-checking by an editorial team. However, research has shown that these services have a limited reach on consumers of non-credible information (Guess, Nyhan, & Reifler, 2018). As such, there is an increasing need to examine the credibility on a broader scale by using computer-assisted processing with human evaluation (Vosoughi, Roy, & Aral, 2018). The web serves as a prolific platform for collaboration and co-creation (Estellés-Arolas & González-Ladrón-de-Guevara, 2012), which has enormous potential to address the issue of online credibility assessment. By considering that a single task could be assigned to a large heterogeneous group, community-based crowdsourcing offers a broader way to evaluate the quality of content published online by exploiting collective competence and judgment (Hammon & Hippner, 2012).

While community-based approaches seem promising (Ishida & Kuraya, 2018), their primary focus has been on assessments of a particular type of content, for instance, images or online news articles. In contrast, the purpose of this article is to conceptualize a crowdsourcing medium, that is, a participatory and co-creative means for evaluating the credibility of *any* sources online. More specifically, the intended medium would focus on evaluating any source origin, whether it is a primary (e.g., original materials), secondary (e.g., reports of findings contained in primary sources), or tertiary source (e.g., the synthesizing of primary and secondary sources) that could be uploaded, found, or shared online (e.g., artifact, document, photo, video, audio).

Taking the aforementioned into consideration, the research question guiding this article is: *How can a participatory and co-creative medium be theoretically anchored and conceptualized to evaluate the credibility of sources online?* This question will be addressed by using the method of Stolterman and Wiberg (2010), namely, concept-driven design research, which will be further described in the methodology section. The remainder of this article proceeds as follows. In the next section, theoretical foundations related to fake news, crowdsourcing, and existing tools, as well as methods for assessing credibility, are described. After that, the methodology is described with regards to how it was applied for the proposed concept of a medium for evaluating the credibility of sources online. Following this, the results and analysis are presented, which includes the elaboration and the evaluation of the proposed concept. The final section offers a discussion of the proposed concept, potential implications, and suggestions for future work.

2. BACKGROUND

2.1. The Challenges of Fake News

Fake news has demonstrated an extensive influence on society, from affecting financial markets (Brigida & Pratt, 2017) and public opinion of climate change (van der Linden, Leiserowitz, Rosenthal,

& Maibach, 2017) to disrupting responses to terrorist attacks (Starbird, Maddock, Orand, Achterman, & Mason, 2014) and natural disasters (Gupta, Lamba, Kumaraguru, & Joshi, 2013; Mendoza, Poblete, & Castillo, 2010). The motivation behind fake news has been debated, and some have argued that monetary, social, and political benefits are among the driving forces (Zhang, Gupta, Kauten, Deokar, & Qin, 2019).

Social platforms, such as Facebook and Twitter, have been the primary venues for spreading fake news, which has posed a risk to political candidates in elections and business organizations in consumer markets (Gross, 2017). During the 2016 U.S. presidential election, fake news stories favoring either one of the presidential candidates were disseminated nearly 38 million times on Facebook (Allcott & Gentzkow, 2017). Despite Facebook's efforts to alter its algorithm to detect and constrain fake news, major fact-checking organizations have reported articles that have not been flagged by Facebook's system (Funke, 2018). It has been identified that visits to Facebook are more common than other platforms before visiting fake news articles, which suggests an influential role of the social network (Guess et al., 2018). When it comes to the social transmission of fake news, it has been contended that it was a relatively rare activity on Facebook during the 2016 U.S. presidential campaign (Guess, Nagler, & Tucker, 2019). At the same time, about 23% of social platform users have reported that they have shared articles, knowingly or not, that can be characterized as fake news (Barthel, Mitchell, & Holcomb, 2016).

In the case of Twitter, a large-scale study examined approximately 126,000 stories that were shared by about 3 million people over 4.5 million times between 2006 to 2017. Fake news reached 1000–100,000 people, while true ones rarely diffused to more than 1000 people. The researchers suggested that the novelty of fake news and the emotions they evoke (e.g., fear, disgust, and surprise) might explain the observed discrepancy (Vosoughi et al., 2018).

The advent of social platforms has created online environments for disintermediation, i.e., “cutting out the middleman,” such as reporters, editors, and media personalities (Eldridge II, García-Carretero, & Broersma, 2019; Fisher, Marshall, & McCallum, 2018). In the mentioned environments, anyone is given the possibility to write and publish anything, anytime, and anywhere. Consumers transitioned into producers of news—reliable, fake, or somewhere in between—that was shared through direct connections. This transition has ultimately created debates concerning the credibility of sources because it has increasingly become difficult for dopamine-driven smartphone individuals to verify facts before spreading them (Carr, 2011; Kim & Dennis, 2019).

Credibility, from Latin *credibilis*, means “worthy of being believed.” It may be defined as a perceived quality that consists of two key components, namely trustworthiness and expertise (Fogg & Tseng, 1999). While Fogg and Tseng's definition of credibility has been broadly cited, we agree with the criticism of Jessen and Jørgensen (2012) that “a great deal of information online is detached from these credentials and authority cues.” Specifically, it is difficult to find direct cues of expertise in user-generated platforms such as wikis, personal blogs, social media, and websites (Jessen & Jørgensen, 2012). At the same time, the weight of expertise as a guarantee of credibility is no more to be taken for granted. Hence, we instead use the definition proposed by the latter as aggregated trustworthiness, henceforth aggregated credibility, i.e., other people's collective judgment is perceived as a credibility metric.

By and large, credibility is related to the believability of “messages rather than speakers” (Metzger & Flanagin, 2013). In other words, it is defined by the evaluation of the source of information, the message itself, or a combination of the source and the message (Metzger & Flanagin, 2013). With regards to online sources, more specifically websites, it has been suggested that credibility is based on social factors such as friends' recommendations (Seckler, Opwis, & Tuch, 2015), users' positive interactions with a website (Fogg & Tseng, 1999), the expertise and credentials of a web site's author(s) and the accuracy, objectivity, and writing style of the information it contains (Jessen & Jørgensen, 2012), and finally, a web site's privacy and ability to secure valuable information (Metzger & Flanagin,

2013). Moreover, researchers have argued that the opinion of crowds could be a contributing factor that may increase or decrease the credibility of an online source (Giudice, 2010).

Previous research has suggested that credibility could be acknowledged in four ways (Fogg & Tseng, 1999), namely (1) when the referee of a website is someone trustworthy; (2) when a third party has authority in the field; (3) when the appearance of a website is appealing and; (4) when a user interacts with a website and the outcome is positive. However, Giudice (2010) argued that because of the rise of collaborative communities such as YouTube and Wikipedia, the evaluation of online sources does not solely rely on authorship or other methods for the judgment of the sources. Instead, it was suggested that a crowds' opinion is crucial for users' belief in the credibility of information on a website (Giudice, 2010). Similarly, Jessen and Jørgensen (2012) argued that perceived credibility is built based on three main factors, namely (1) social validation (e.g., made by others, such as comments, likes, ratings), (2) profile (i.e., having a known and verified identity plays an important role when assessing information) and (3) authority or trustee (i.e., a brand or authority on a matter).

Currently, there are no digital tools or online services that could help users evaluate the credibility of sources online by means of crowdsourcing. More specifically, it is meant that users' knowledge and expertise are taken into account while the medium is not under the control of a specific company or organization, that is, open source, but instead is maintained and controlled by users in a community that is based on the underpinnings of crowdsourcing.

2.2. Crowdsourcing for Evaluation of Sources Online

There are many definitions of 'crowdsourcing' since the practice has been adopted for different purposes (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). From an etymological point of view, it could be defined as follows: 'crowd' refers to individuals who anticipate an initiative, and 'sourcing' relates to different activities targeted at "finding, evaluating and engaging" (Estellés-Arolas & González-Ladrón-de-Guevara, 2012, p. 189). The phenomenon of engaging crowds is not novel, but with the inception of the Internet, modern-day, large-scale crowdsourcing was made possible. For this article, different definitions of crowdsourcing were elaborated into one to highlight essential elements that are of value to the research question. The elaborated definition of crowdsourcing is: *the deliberate mobilization of creative and contemporary ideas or stimuli throughout the world wide web* (Mazzola & Distefano, 2010) *with the intention to solve a problem or engage in tasks that require human intelligence* (Kazai, 2011), *in a process that is participative, where users are seeking the same outcome* (Estellés-Arolas & González-Ladrón-de-Guevara, 2012), *often in exchange for micro-payments, social recognition, or for entertainment purposes* (Kazai, 2011).

Crowdsourcing offers possibilities for commercial purposes. It could simultaneously be regarded as a powerful tool for government and non-profit sectors in problem-solving processes (e.g., Challenge, challenge.gov, Peer to patent, peertopatent.org) (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). Further, crowdsourcing addresses tasks that still need human intelligence and cannot be solved with full automation (Nassar & Karray, 2019). Examples of human intelligence tasks (HIT) include complex image categorization, survey completion (Brawley & Pury, 2016), and debunking misleading information (e.g., automatic systems fail in identifying whether a picture is fake or if it appears in a humorous context) (Boididou et al., 2018).

The web serves as a prolific arena for collaboration and co-creation (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). A task can be assigned to a heterogeneous group that, in turn, offers a large amount of knowledge and competence (Hammon & Hippner, 2012). A crowdsourcing process may be characterized by modularity and flexibility (Hammon & Hippner, 2012), which shortens the time needed to solve a particular task, and results in more innovative solutions, compared to traditional means (e.g., human resources available within a particular company). For example, the crowdsourcing innovation platform InnoCentive (innocentive.com) provides companies with crowdsourced solutions when their resources are insufficient.

For people to become a part of a crowdsourcing initiative, some incentives are often needed (Nassar & Karray, 2019). Studies that have attempted to define 'crowdsourcing' have often mentioned, in one form or another, the importance of compensation (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). While some people may translate it to *monetary profit*, others view it as *social recognition* and *entertainment*. Crowdsourcing gives opportunities to fulfill individual needs, such as developing creative skills, sharing knowledge, be a part of a community, enjoying being committed to solving a task, or having fun (Estellés-Arolas & González-Ladrón-de-Guevara, 2012). Many individuals feel prompted to contribute to something collective and of higher importance, which produces the sensation of belonging to a democratized web, because the outcome was developed and produced by users and for users (Brabham, 2013). For example, Stack Overflow (stackoverflow.com) is a popular question and answer service that is regarded as an invaluable source for users who share their knowledge and advance their skills in a range of topics, from math to programming (Kavaler & Filkov, 2018). Many users feel more prompted to engage when they receive a comment or other indicators that express the appreciation of their peers for their contribution (Brabham, 2013). Other examples include Amazon's Mechanical Turk (mturk.com) and Waze (waze.com). Amazon's Mechanical Turk is an online marketplace for finding human resources to solve human intelligence tasks (HIT) for compensation. Waze is a crowdsourcing traffic and navigation application designed for people all over the world to share real-time traffic and road information (Singh, Bansal, Sofat, & Aggarwal, 2017).

Several phenomena that originated from crowdsourcing can be identified, for instance, *crowd voting* (e.g., Reddit, Threadless), *crowdfunding* (e.g., Kickstarter, Indiegogo), *crowd recruiting* (e.g., Relode, Visage) and *crowd documentation* (e.g., Stack Overflow). Crowd voting is functionality that integrates users' viewpoints and needs into a project by allowing a crowd to evaluate proposed ideas (Hammon & Hippner, 2012). Crowdfunding aims to collect money to bring a specific project or product to a market. In contrast to traditional fundraising, crowdfunding is based on the idea of small amounts of money from a large group of people. Crowd recruiting focuses on recruiting freelancers from a crowd (Hammon & Hippner, 2012). Finally, crowd documentation is a collection of web resources curated by a large group, a crowd, which also contributes to the collection mentioned above. A crowd document is, therefore, different from other web resources, for instance, source code repositories, where curation is less dominant (Parnin, Treude, Grammel, & Storey, 2012).

Regarding a typical structure of any crowdsourcing initiative, Nassar and Karray (2019) suggest a division of the process into five modules, namely (1) designing incentives; (2) designing quality control methods for users, tasks, and collected data; (3) collecting data; (4) aggregating data; and (5) verifying received data (Nassar & Karray, 2019). The first module, *designing incentives*, indicates that incentives could be intrinsic (personal enthusiasm or altruism) or extrinsic (monetary reward) (Allahbakhsh et al., 2013). Examples of the former could be *entertainment* (e.g., games where users may earn points) and *social recognition* (e.g., attention from viewers on YouTube). For the latter, it could be *financial compensation* (e.g., money as a reward for solving tasks on Amazon's Mechanical Turk). The second module, *designing quality control methods for users, tasks, and collected data*, is necessary to verify aggregated data. Meek, Jackson, and Leibovici (2014) argued that scientists could discard crowd data because of quality issues; therefore, verification and validation of the collected data is of high importance. Examples of verification and validation include *quality controls of users* (accuracy and trustworthiness), *users' reputation* (community-based), and *experience* (task-dependent) (Nassar & Karray, 2019). Furthermore, the third module, collecting data, refers to finding and storing information to act on it in subsequent processes, such as aggregation and verification. The fourth module, *aggregating data*, indicates that the information collected from users should be aggregated in order to constitute meaningful outcomes (Nassar & Karray, 2019).

Researchers Hung, Tam, Tran, and Aberer (2013) provide two suggestions for techniques of aggregation, namely *non-iterative* and *iterative*. The former is based on heuristics to compute a single value of an object, while the latter is based on probability estimation to compute possible labels of

an object. The last module, *verifying received data*, implies different processes to assure the quality and reliability of data. Such processes may include *rewards and penalties* (e.g., monetary bonuses for quality annotators and penalties for false answers), *redundancy* (e.g., majority consensus), *manual verification* (e.g., in easy grading tasks where users can give a score), *automatic validation* (e.g., computational credibility check), and *comparison against authoritative data* (e.g., testing if workers are following a protocol by comparing their annotations to the gold standard) (Nassar & Karray, 2019). In addition, crowdsourcing solutions often adopt high-experienced *topical experts* to reduce costs and enhance the quality of the aggregated data (Nassar & Karray, 2019).

Despite many potential benefits of crowdsourcing initiatives, several risks can be identified. For instance, some of the risks with crowdsourcing include the danger of losing control that could manifest in boycott or obstruction within a crowd, vagueness of crowd structure, and troublesome communication with participants through feedback loops (Hammon & Hippner, 2012). Moreover, the majority of crowdsourcing platforms are based on the English language, which may create a barrier for users who experience language limitations. Also, crowdsourcing initiatives face a challenge for quality control, where aggregated data needs to undergo a verification process in order to establish the quality and reliability of the data (Nassar & Karray, 2019).

Different applications of crowdsourcing prove that the idea of involving crowds creates possibilities to achieve a variety of goals (e.g., *crowdfunding* to collect money; *crowd recruiting* to find a suitable workforce). A crowdsourcing approach that is based on the participative processes of online users could be exploited to improve the value of information shared online. In order to inspire and inform the design process for the proposed concept, existing approaches, and tools for assessing credibility online are presented and evaluated in the next section.

2.3. Approaches and Tools for Assessing Credibility Online

A growing number of approaches (e.g., checklist approach, fact-checking sites, collaborative filtering) and tools (e.g., Hypothes.is, Google PageRank, Alexa Rank Checker, MyWOT) are available when assessing credibility online. Traditional gatekeepers (e.g., reporters, editors, news organizations) are no longer sufficient in an Internet environment, which is characterized by a constant and uncontrolled flow of information (Shah, Ravana, Hamid, & Ismail, 2015; Westerwick, 2013). Existing techniques have attempted to address the previously mentioned issue, and those could broadly be classified into two fundamental approaches, namely (1) credibility evaluations by users and (2) credibility evaluations by computers (Choi & Stvilia, 2015; Shah et al., 2015).

One of the user evaluation approaches, referred to as the *checklist approach*, emerged as a web assessment method to propagate digital literacy (Breakstone, McGrew, Smith, Ortega, & Wineburg, 2018; Shah et al., 2015; Subramaniam et al., 2015). This approach relies on a list of guidelines that provide a course in web information evaluation to users online. However, while this method contributes to a more prudent usage of Internet resources, it is not applicable in many cases as checklist approaches are outdated (Breakstone et al., 2018), and excessively time-consuming (Shah et al., 2015; Subramaniam et al., 2015), with some achieving the mark of over 100 questions per web page visit (Shah et al., 2015).

Other information assessment approaches include fact-checking sites (e.g., Snopes, Hoax-Slayer, Politifact, Factcheck; see Duke Reporters' Lab, 2019 for an extensive list). These sites have risen as a new journalistic practice due to the backlash caused by the increase of the spread of misleading information online (Hameleers & van der Meer, 2019). Fact-checkers aim to improve political discourse and democratic accountability (Nyhan & Reifler, 2015) by guiding citizens to carefully examine the claims that appear in political news stories and highlight which story is to be trusted or not (Hameleers & van der Meer, 2019). Fact-checking could be regarded as a tool that strengthens democracy, and that emerges in those countries where democratic institutions are recognized as weak or threatened (Amazeen, 2020).

While fact-checkers contribute to a more informed society, researchers have identified several weaknesses with this journalistic practice. First, a majority of fact-checking sites focus on political issues, leaving a vast majority of other issues that appear in media uncovered (Lim, 2018), especially at the local level (Hassan, Arslan, Li, & Tremayne, 2017). Secondly, validation itself is intellectually demanding and laborious (Lim, 2018), which in turn makes the process costly and time-consuming (Hassan et al., 2017; Lim, 2018). Finally, citizens in many countries are still unaware of the existence of fact-checkers (Amazeen, 2020) because they use deprecated content management systems designed for traditional newspapers and blogs, which limits the spread of their content in various computational projects such as modern structured journalism (Hassan et al., 2017).

Another type of assessment approach, similar to fact-checkers, is based on peer-review collaboration and is called *collaborative filtering*, in which experts in a specific area evaluate websites or content (Shah et al., 2015). Such recommendations are considered to be highly trusted (Su & Khoshgoftaar, 2009). The main concern with this approach is that its effectiveness depends on the level of activity of the experts, that is, the credibility evaluation of content needs to be updated in a timely fashion in order to be successful (Shah et al., 2015). Considering the amount of new information that is continually being created and shared online, the task of continuous updates by experts may seem challenging. An example of collaborative filtering is the platform *Hypothes.is* that allows users to create web annotations on any web page or PDF file to facilitate discussions over a variety of topics (Perkel, 2015). With its functionalities of community-controlled annotation layers over the web, *Hypothes.is* creates an arena for fact-checking activities and civic engagement. The platform has proven successful by, for example, Climate Feedback, which is a group of high-profile climate scientists that use the platform to highlight inaccuracies and comment on articles that concern the topic of climate change.

Web annotations, as a collaborative activity, create a socio-technical environment for the exchange of thoughts, agreements, and disagreements, implementing the completion of everyday tasks associated with information literacy (Kalir & Dean, 2018). *Hypothes.is*'s interface refers to modern communication trends, which tend to be composed of overlapping *layers* of meaning, referents, and negotiations (Kalir & Dean, 2018). *Hypothes.is* is an example of an open-source community, which strengthens its transparency value, making it trustworthy due to the fact that in this kind of production, process and design are executed by individuals that cooperate in order to achieve common resources using their terms and needs, as a self-governing community (Brabham, 2013). The utilization of an *open source* approach grants actors and communities the freedom to customize applications in accordance with their needs and wishes (Faraon, 2018b; Faraon, Villavicencio, Ramberg, & Kaipainen, 2013). To improve its transparency, *Hypothes.is* takes advantage of a system that is called ORCID (Open Researcher and Contributor ID), which aggregates digital profiles of researchers that engage with the medium to verify users' identity (Perkel, 2015).

Although collaborative filtering platforms can be regarded as a promising assessment approach, they pose a threat of bias with their mechanism of engaging only specialists in a particular area. Inclusiveness of all the citizens should be acknowledged in order to avoid marginalization (Trechsel, 2007), and to ensure the transparency of the assessment process, which could be achieved by granting all users the ability to aggregate the credibility and validity of sources in the form of evaluations (Liaw, Zilnik, Baldwin, & Butler, 2013). Another evaluation method which deals with source credibility assessment and uses a mechanism of aggregating evaluations may be based on (a) rankings of websites that are computationally generated by search engines or (b) rankings aggregated by user ratings.

Google PageRank and Alexa Rank Checker are examples of search engines that generate rankings, which helps to determine the popularity of a website (Aggarwal, Van Oostendorp, Reddy, & Indurkha, 2014), and this, in turn, could aid users to estimate which websites could be trusted. Google PageRank works by calculating the number of pages that point to a specific website. Ishida and Kuraya (2018) indicated that Google has recently modified its ranking system to address the emergence of news that contains misleading information by identifying websites that aim to spread

such information. Westerwick (2013, p. 194) pointed out that “a high search-engine ranking increased sponsor credibility, and thus influenced information credibility indirectly.” Alexa Rank Checker is built on an algorithm that combines a number of average daily visitors on a given website and its page views in the past three months (Aggarwal et al., 2014). Those websites that have the highest combination of those two credentials are lowest in the overall rank. The main issue with automatic ranking systems is that popularity does not always imply importance in terms of high-quality content (Masterton & Olsson, 2018). Moreover, it is essential to point out that automatic ranking systems evaluate sources exclusively, which means that users, in order to check the veracity of a particular content online, still need to utilize other approaches that are available online.

On the other hand, web credibility evaluation systems (WCESs) take advantage of the wisdom of crowds and collect users’ ratings (Liu, Nielek, Adamska, Wierzbicki, & Aberer, 2015) instead of relying solely on calculations done computationally. Such systems, also called *social feedback systems*, employ the use of *collaborative editing tools* such as *comment boxes*, *rating tools*, *community-editable content*, and *collaborative linking*, which in turn allows users to approve information and share content (Shah et al., 2015). MyWOT (mywot.com) or WOT (Web of Trust) is an example of an existing WCES that assembles users’ ratings regarding web trustworthiness and child safety (Liu et al., 2015). Aggregated recommendations are visible next to search results in the form of traffic lights: green means that users have rated the site as reliable; yellow indicates caution while using a determined site, and red states a warning for possible threats. Additionally, by clicking the traffic lights, users can reach other users’ opinions or see more information about a site’s reputation. MyWOT takes advantage of *reputation systems* in which users are assigned scores when providing credibility ratings (Whiting et al., 2017), which is then translated into the weight of a user’s influence on the final credibility aggregation (Liu et al., 2015). While WCESs have demonstrated their usefulness, they face several challenges. An example connected with the passive character of rating collection is the fact that these are the systems that wait for users to submit their ratings (Liu et al., 2015). What this means is that most websites have no rating at all, and many who have, are based on a limited amount of ratings that are submitted by a small group of users. Moreover, there is a risk of misusing WCESs by malicious users who wish to submit fake ratings in order to destroy or build the reputation of a given website (Liu et al., 2015).

Most of the credibility evaluation methods or tools for the assessment of online content are text-based. The evaluation of the credibility of non-textual content proves to be even more challenging than that of textual content. In the digital era, misbehaving users manipulate both images and videos motivated by the desire to gain profit or to maximize or minimize the reputation of the media, which indirectly affects users’ level of trustworthiness (Rashed, Renzel, Klamma, & Jarke, 2014). It is typical that pictures are being manipulated visually or that they are published as the real ones but with false metadata. The most popular method to assess the validity of image sources is the so-called *reverse image search*. Engines that use reverse image search are based on advanced CBIR (content-based image retrieval) methods. To find potential matches in databases, the engines mentioned above create fingerprints of the searched image and exploit various machine learning algorithms (Mehrnezhad, Ghaemi Bafghi, Harati, & Toreini, 2017). The result of a search helps to estimate image credibility by analyzing when and where similar or alike images were published. The main drawback with the reverse image search method is accentuated if the image a system is trying to identify is extensively manipulated, which may culminate with the failure to find any matches. However, this problem may be solved with the involvement of human intelligence. Rashed et al. (2014) suggest a tool that combines automatic methods with community involvement to judge the credibility of multimedia. The authors underline that in their proposed tool, experts and media constitute a network of confidence relationships (Rashed et al., 2014). The proposed incentives rely on the idea of media authenticity rating mini-games that motivates users to participate in a fun way with a serious purpose.

A call for a practical credibility assessment of videos seems to be alarming in the era of an emerging threat called *deep fakes*. Deep fake videos are the outcome of artificial intelligence or machine-learning applications that operate by merging, combining, replacing, and superimposing images and video clips onto a new, fake video (Maras & Alexandrou, 2019). Maras and Alexandrou (2019) stressed that the most significant threat with manipulated videos is that they erode the trust in video evidence used in court. Moreover, there are other issues with deep fakes, such as the use of faces of public figures superimposed on the bodies of porn stars. Credibility assessment of video content can be achieved with the use of computational tools such as the InVID Verification Plugin (invid-project.eu). The tool is designed to support journalists in verifying video content on social networks, such as Facebook and YouTube. While it mostly builds on an analysis of metadata, it incorporates human input by letting users share their comments on the credibility of a particular.

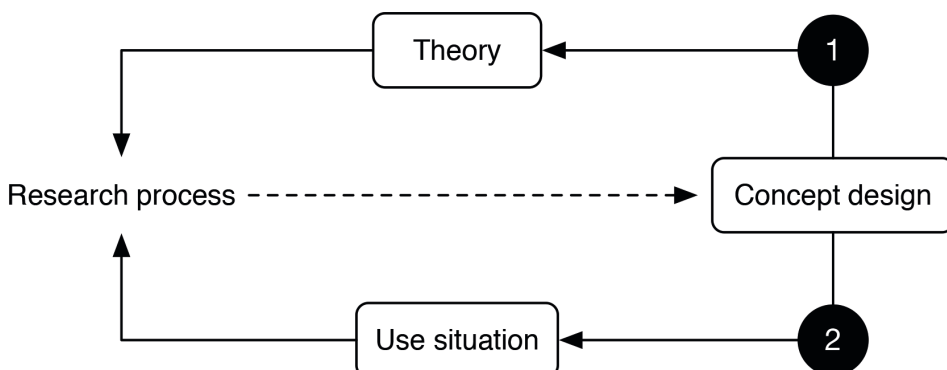
As discussed in this section, many online content assessment methods have proven to be useful, while at the same time, many have limitations. Scholars highlight the benefits of crowdsourcing methods and the need for computer-based assessment tools in combination with human intelligence. With the involvement of a vast number of users, many of the discussed issues, such as the time gap between when information is published and when professional editorial teams check it could thus vanish (Hassan et al., 2017). Based on previous considerations, the following section will describe the methodology of concept-driven design research that will be used to theoretically anchor and conceptualize a participatory and co-creative medium that aims to evaluate the credibility of sources online.

3. METHODOLOGICAL APPROACH

3.1. Concept-Driven Design Research

To focus on the creation of design concepts, this article has adopted the methodology of concept-driven design research proposed by Stolterman and Wiberg (2010). This method aims at “manifesting theoretical concepts in concrete designs” (Stolterman & Wiberg, 2010, p. 95). Furthermore, this method explicitly provides instruments for the design and creation of a concept and an artifact to manifest the desired theoretical ideas in their entirety (Stolterman & Wiberg, 2010). The method of concept-driven design research departs from theory/concept rather than empirical research. Furthermore, research makes the development of artifacts possible through immediate design. The advantage of this method is a design that is specific to an idea, a concept or a theory, rather than a specific problem, user or particular use context. The process of the concept-driven research approach is illustrated in Figure 1.

Figure 1. The process of the concept-driven design research approach in relation to theory and use situation, adopted from Stolterman and Wiberg (2010, p. 101)



Stolterman and Wiberg (2010) explicitly define the research approach as having a *conceptual/theoretical point of departure* rather than an empirical one (Arrow 1). Conceptual and theoretical explorations are aimed to produce a design concept that supports the *use situation* (Arrow 2). The end result—that is, the final design—is optimized in relation to a specific idea, concept, or theory rather than a specific problem, user, or a particular use context.

Additionally, the field of user interaction design and participatory design cannot be conceptualized and explained only through theoretical methods. Stolterman and Wiberg (2010, p. 99) argue, therefore, that “the theoretical advancement also has to be done through a more concrete and exploratory process, involving design and artifacts as significant elements.” The concept-driven design research method encompasses seven methodological activities: *concept generation*, *concept exploration*, *internal concept critique*, *design of artifacts*, *external design critique*, *concept revisited*, and *concept contextualization*. Each of these activities contributes to the process of concept generation and is further described in the following.

Concept generation consists in the discovery of new concepts using previous theoretical advancements in the field (Stolterman & Wiberg, 2010). The idea is to find the uniqueness of a concept using distinct methods such as metaphors or theories from other design fields, as well as combining new and old ideas into one. In other words, this stage supplies the researcher with crucial knowledge for the conceptualization of new ideas by using an array of methods that provide distinct inputs that may help generate possible concepts.

During *concept exploration*, different materials and content are used in order to generate new ideas. In this stage, the researcher creates prototypes, explores different materials, forms, and models et cetera in order to find new design spaces, and not merely refine existing ones. Mainly, this stage represents the possibility to “(...) find unseen parts of already known design spaces” (Stolterman & Wiberg, 2010, p. 110), which means that the researcher has the opportunity to explore and communicate with the design materials at hand toward finding possible concepts.

The third activity, *internal concept critique*, involves a process where a design and associated concepts are submitted to a theoretical foundation. To determine whether a concept is meaningful or not, three factors have to be taken into account: (a) the uniqueness of the chosen concepts; (b) the relation of the concepts to existing theory; and (c) how well these concepts can be expressed in a concrete design (Stolterman & Wiberg, 2010, p. 110). Strictly speaking, under this phase, the researcher must evaluate and determine whether or not a concept is purposeful enough to be made into a concrete design.

In the next activity, *design of artifacts*, a concrete artifact is produced that represents a design concept. Thus, this process constitutes a part of the design process and the theoretical advancements (Stolterman & Wiberg, 2010). The artifact must incorporate a design in its entirety, which means that under this phase, theory and craft—the making of the artifact—will merge to manifest a full concept.

Following the design of an artifact, *external design critique* is adopted to evaluate and validate an artifact in its entirety: idea, concept, and associated theoretical principles. In other words, it is under this activity that the whole concept is tested. Stolterman and Wiberg (2010) stated that testing means that the conceptual design is submitted to public evaluation and critiqued as a whole.

After this activity, reaching the *concept revisited* activity, the concept is revisited, revised, and the critique from the previous stage is used to guide additional designs. What this means is that it is difficult to determine in advance what reactions a concept will receive under evaluation (Stolterman & Wiberg, 2010). Therefore, it is essential to reexamine the critique gained to determine whether a concept has flaws and imperfections.

Finally, *concept contextualization* occurs after the revisited concept is defined, and the outcome is a prototype or an artifact (Stolterman & Wiberg, 2010). Fundamentally, in this activity, a new concept is juxtaposed with a currently updated body of concepts and theory in the field in order to be compared to similar concepts and theories and evaluate its contributions to previously realized work

(Stolterman & Wiberg, 2010). Namely, all knowledge gained under the whole process is gathered in order to contribute to further concepts.

The methodology of concept-driven design research has been adopted in several research studies (e.g., Eliasson, 2013; Faraon, 2018b; Johansson, Lassinantti, & Wiberg, 2015; Johansson & Wiberg, 2012; Nazzi, Bagalkot, Nagargoje, & Sokoler, 2012). For example, Eliasson (2013) applied the methodology in order to probe outdoor lessons on geometry and biology to evaluating proposed design tools. The outcome of this research was three design tools, namely design guidelines that guide the design of mobile technology fitted for outdoor lessons; a design model for evaluation and design of mobile technology for outdoor lessons; and design concepts that helps to reflect “on the placement of mobile technology in outdoor lessons” (Eliasson, 2013). Further, Johansson and Wiberg (2012) applied the method of concept-driven design to highlight a mobile IT concept that managed to be formulated, explored, and validated. Their research resulted in a concept referred to as CASAM, Context Awareness Supported Application Mobility, which could be used to support the work of home care service groups. Also, Johansson et al. (2015) utilized the methodology to develop a concept that could strengthen citizen involvement in the context of e-government. The result of this research was the creation of a widespread concept for the public sector service process that utilizes mobile e-services and open data as fundamental elements (Johansson et al., 2015).

In the following section, the seven previously mentioned activities associated with the concept-driven design research methodology are elaborated by way of how they have been applied in this article.

3.2. Application of Concept-Driven Design Research

This article utilizes the concept-driven design research approach in order to create a design concept that could evaluate the credibility of sources online. While there are many effective research methods in use that aid in the exploration of possibilities when designing digital artifacts, such as participatory design and contextual design, the concept-driven design research method was chosen because it is better suitable when the aim is to develop more *theoretical* and *conceptual* contributions (Stolterman & Wiberg, 2010). According to Stolterman and Wiberg (2010), the concept-driven design research approach encompasses seven activities that help the researcher find new ideas for new and never seen concepts. These activities are *concept generation*, *concept exploration*, *internal concept critique*, *design of artifacts*, *external design critique*, *concept revisited*, and *concept contextualization*.

The process of research using this method was conducted the following way: research started with the *concept generation stage*, which was a fundamental phase that helped build a theoretical foundation with the aid of research on three main areas: (1) credibility, (2) crowdsourcing and (3) existing tools and methods of credibility assessment online. Different perspectives on credibility with a focus on credibility *shaped in an online context* were explored. Crowdsourcing, with its many forms, was studied with emphasis on its possible advantages and disadvantages as well as on the design of a crowdsourcing process. Finally, the exploration of a multitude of approaches and tools for the evaluation of sources online was conducted. Under this stage, possible ideas for a concept started to be generated by examining theoretical concepts that could be adopted in a design concept.

During *concept exploration*, research was made using paper and pen, flowcharts, and mockups, in order to communicate ideas and explore new possibilities. For this article, this phase meant a deeper understanding of the problems with online credibility, identified in the background section, and how to assess them. Moreover, during this phase, the initial design of the proposed medium started to be idealized and formed. Different materials were applied using the knowledge acquired in the previous stage in order to externalize the first ideas for a concept.

In the *internal critique activity*, the developing concept was tested against existing approaches and tools to decide whether the proposed concept was unique or not. Strictly speaking, the emerging concept was tested with the help of theory to determine whether its foundations were well-grounded and how well they expressed earlier theoretical recommendations when designing a crowdsourced medium. This stage was also critical because it allowed for the verification of the concept’s concreteness. The

internal critique activity revealed that a few functions needed to be discarded as they did not appear to have the correct specifications from previous studies of a system that could be valuable for users. An example of a discarded function was the monetary compensation the proposed medium could offer its users. As previous research pointed out, assigning monetary compensations as a sort of incentive for people to partake in a specific activity can decrease the quality of the outcome (Nassar & Karray, 2019). Hence, many users could question the authenticity of the information evaluated using the proposed medium and jeopardize the quality of the concept.

Under the next activity, the *design of artifacts* activity, a storyboard was created, using an online tool named Plot, that allowed for the collaboration in the creation of storyboards online. The storyboard gathered all main functions and connected theory and design to manifest the concept. This phase was crucial because the proposed medium with its functions was synthesized for the first time and was prepared for the next activity where participants would finally evaluate it. The storyboard was chosen as a means of presentation of the concept because it contained pictures and informative text about the proposed medium. Participants were briefed about the issue of the assessment of the credibility of sources online and how the proposed medium could contribute to address this issue.

During the next activity, *external design critique*, the concept was exposed to participants to test and evaluate its underlying ideas as well as theoretical principles. All criticism gathered under this activity validates the theoretical and conceptual implications for the design (Stolterman & Wiberg, 2010). For this article, this phase was conducted using eight participants for one week. During this time, the participants, who consisted of five males and three females, with an age range of 23 to 55, were interviewed. In order to protect the participants' identities, the authors chose to use the labels P1-P8 ("P" for the participant) according to the order they were interviewed.

Furthermore, the authors have used convenience sampling; in other words, the choice of the participants who evaluated the concept was made from the population that was close to hand. The final selection resulted in a group of people that were involved with different kinds of occupations, ranging from students of design and behavioral studies and other occupations. The participants worked in pairs to address qualitatively self-generated questions, for instance, "*what is your general feeling about the medium,*" see the Appendix for a complete list of questions. Hence, participants were asked to discuss the mentioned questions openly in order to evaluate the proposed design concept, which was manifested via a short presentation containing visual and textual materials demonstrated by the authors. Four participants evaluated the concept physically in a booked classroom where they gathered to discuss the concept. The remaining four participants evaluated the concept using Skype as a means of communication. The choice of a qualitative questionnaire, as a means to support the evaluation of the emerging concept, created an opportunity for participants to discuss ideas and express their opinions more openly.

After the previously mentioned activities had been accomplished, the results of the discussions could then be utilized in the *concept revisited* activity. During this stage, the results of the study were analyzed and examined in order to detect possible flaws, and misunderstandings about the proposed design concept. All relevant feedback from the participants' evaluations were incorporated into the design concept.

Finally, one last activity remained: the *concept contextualization* activity. During this stage, all knowledge gained from theory, design, and evaluation with participants were synthesized in order to create a final design concept. Under this stage, focus laid on the uniqueness of the concept and how well it relates to previous research in the field.

3.3. Ethical Considerations

Ethical considerations, as recommended by the Swedish Research Council (2017), were followed in the research process that was conducted in this article. The ethical considerations dealt with the requirement of *information, consent, confidentiality, and usage*. Participants were informed about the purpose of this study, which was done at the beginning of each interview. Consent was acquired by

asking participants whether they agreed to participate in the interviews before starting. Furthermore, it was necessary to emphasize the confidentiality of the collected data. The only information about the participants that was gathered included their age, gender, and ethnic background, and this was done in order to ensure that the identities of the participants would not be revealed. Lastly, all information related to the interviews was promised to be used only in scientific research, which means that all information gathered during the study was utilized exclusively in this article.

4. RESULTS AND ANALYSIS

This section aims to describe and elaborate on the results of the concept-driven design research method that was adopted to conceptualize a crowdsourcing medium aimed to support the process of assessment of sources online. The concept, which will be elaborated in the next section, was evaluated by eight participants (P1-P8) as a part of the external design critique activity, which resulted in a revisited concept that is presented in the subsequent section.

4.1. Elaborating the Concept of Aggregated Credibility

Following the methodology of concept-driven design research (Stolterman & Wiberg, 2010), a concept of a participative and co-creative medium for aggregating evaluations of sources online was created. It is important to emphasize that the proposed medium has its foundations in the process of crowdsourcing. As previously mentioned, crowdsourcing is described as the deliberate mobilization of creative and contemporary ideas or stimuli throughout the world wide web (Mazzola & Distefano, 2010) to solve a problem or engage in tasks that require human intelligence (Kazai, 2011).

In line with Giudice's (2010) argumentation that the *crowds' opinion* is crucial for users' belief in the credibility of online information, and Jessen & Jørgensen's (2012) argument that perceived credibility is based on *social validation* (e.g., made by others, such as comments, likes, ratings), this article proposed a concept of a medium that aims to mobilize crowds with the *common purpose* to solve the issue of *assessment of the credibility of sources* online using participatory, co-creative and democratic processes. The credibility assessment process in the proposed medium could be described as following: (1) individuals discover a source, whether offline or online, and submit it for evaluation; (2) as more individuals become engaged, a process towards community building emerges with the aim to evaluate the submitted source (e.g., sharing own evaluations, taking part in other user's evaluation); (3) following iterative processes that include commenting and voting, consensus starts to take place (aggregating evaluations in order to estimate a majority solution); (4) the end result is the successful accomplishment of consensus, which could be described as the outcome of the evaluations collected from engaged individuals within the community.

The proposed medium aims to accentuate democratic values and is based on the notion that any individual who uses the Internet can participate by sharing *their evaluations* of a particular source/content, as well as take part in *other users' evaluations*. Further, to serve democratic functions, the proposed medium should adopt measures to avoid the marginalization of any groups of citizens; in other words, inclusiveness should be acknowledged (Trechsel, 2007). Additionally, the proposed medium may be used in both mobile devices and personal computers, which creates the possibility for all users to seek the evaluation of sources of all kinds, for instance, websites, blogs, pictures, and videos.

As for the first stage in any crowdsourcing initiative, *incentives* need to be designed (Nassar & Karray, 2019) in order to find ways to motivate the crowd to enroll in the initiative. By and large, the incentive for participating in the assessment of credibility of online content is the wish to contribute to 'something bigger', which increases the feeling of belonging to a 'democratized web', and because it is something that is made 'for us and by us' (Brabham, 2013, p. 91). Hence, the offered value for those who engage in the proposed medium would be *intrinsic* because users would receive *social recognition*, experience *personal enthusiasm*, and submit their evaluations for *altruistic* reasons (Nassar & Karray, 2019).

In order to evaluate sources online, the proposed medium would provide features that facilitate *knowledge sharing* and expertise amongst users. Such features are, for example, *crowd voting* and *collaborative editing tools*. *Crowd voting* is functionality that integrates users' viewpoints and perspectives into the project by allowing the crowd to evaluate ideas (Hammon & Hippner, 2012). In the case of the proposed medium, users must be allowed to express their opinions in order to evaluate a source or content. Crowd voting could be used as one of the main functionalities of the medium. Apart from the more extensive and qualitative ways of evaluating a source or content, such as commenting, linking to other relevant sources, and correcting, voting is a reasonable, time-effective, and democratic way of expressing opinions. The purpose of this feature is to create a filter that segregates online content into credible and non-credible to improve the quality of information that individuals discover over time. The voting system could be based on the idea of thumbs-up for the credible, and thumbs-down for the non-credible source/content (binarity), or more extended ranking systems (e.g., inspired by fact-checkers scales or expressing emotions as on Facebook). In order to explain one's choice about whether a determined source/content is credible or not, a user needs to fill in a short description of his/her argument before submitting the vote. Apart from expressing opinions about whether or not a source or a content is accurate, a third option, which could be labeled "needs evaluation," would provide the choice to mark a source or a content for other users to indicate that it needs/seeks to be evaluated, i.e., a request. Finally, the proposed medium could provide a feature to express approval or disapproval of a particular user in order to evaluate his/her online contributions. Such a feature would be visible to users in the form of a checkmark for the approved content, and an X for the disapproved content.

The proposed medium further aims to make *knowledge sharing* possible among users and should facilitate *collaboration* in the form of *collaborative editing tools* such as *comment boxes*, *community-editable content*, and *collaborative linking*, to allow users to approve information and share content (Shah et al., 2015). Strictly, the medium should contribute to collaboration among stakeholders, with users combining their resources to achieve common interests (Faraon, 2018a). For instance, with the comment function, users have the opportunity to present their arguments or to have a discussion with other users, and by linking, users can provide evidence using other sources as reference. Additionally, with the help of *web annotations*, as an example of community-editable function that "complements everyday activities associated with mediated information literacy" (Kalir & Dean, 2018, p. 18), users can point out possible inaccuracies in the content, or mark a part of the content that they wish to discuss or highlight. When a source needs correction, the medium should be able to allow users to propose these corrections directly with a single click. Such corrections should then be visible to all users without the dependence on the updates made by news agencies. Thus, the quality and validity of the source are enforced because they were independently generated from user to user (Liaw et al., 2013). These features contribute to a more nuanced expression of opinions and provide additional ways of sharing knowledge apart from simple "true or false" validation. Moreover, the application of the features mentioned above addresses one of the risks that many crowdsourcing initiatives face, namely troublesome communication with participants through feedback loops (Hammon & Hippner, 2012), which enables an immediate communication among users.

The proposed medium could adopt *simple visual feedback* to simplify the visual impact of aggregated credibility assessments. This could be based on colors or icons that would inform users of the current outcome of a voting process. For instance, in the medium MyWOT, aggregated recommendations are visible next to search results as traffic lights where green means that users have rated a website as reliable; yellow indicates caution while using a determined website, and red states a warning for possible threats. Additionally, a similar simple visual feedback could be extended to a more advanced visualization form that shows the details of the aggregated assessment, such as how many users voted, who voted, and what comments and corrections have been submitted.

Meek et al. (2014) argued that scientists could dismiss crowdsourced data because of related quality issues. Therefore, verification and validation of the collected data are of high importance.

In order to ensure that the proposed medium is reliable, it should contain a *reputation system* in which users are assigned scores while providing feedback to decide users' priority and the weight of the validity of their answers (Nassar & Karray, 2019). This reputation system could be used on the assessment of the reliability of individual users, having as a basis their past rating behavior, which is a *task-dependent* method of verifying if the collected data is to be trusted. Strictly speaking, the ratings of more reputable users impact more on the final credibility assessment (Liu et al., 2015).

Moreover, the reputation system would not only *aggregate* scores gained by an individual while evaluating sources, but also those gained in *peer assessment*, i.e., scores that other users addressed in an approval rating. In this way, quality control would be community-based at the same time (Nassar & Karray, 2019). Peer assessment should indulge two possibilities: to express (1) approval that would add points to a user's account (which would act as a reward) or (2) disapproval that would subtract points (a penalty). Additionally, *social recognition* works as a motivator for users to take part in a crowdsourcing initiative. Many users feel more prompted to engage when they receive a comment or other indicator that expresses the appreciation of their peers for their contribution (Brabham, 2013). A reputation system may potentially motivate users to engage in the proposed medium as they would continue to receive social recognition.

To strengthen the democratic value of the proposed medium, users could create a *ranking system* to suggest reputation levels. Additionally, being perceived as *an authority or trustee* in a community plays an essential role in how people assess the credibility of information (Jessen & Jørgensen, 2012). The rank system functions as signals for reputation or qualification on crowdsourcing platforms (Whiting et al., 2017), which may make it easier for users of a platform to judge each other's credibility.

In order to establish the quality and reliability of the aggregated data, the same needs to undergo a verification process (Nassar & Karray, 2019). One of the ways to improve the quality of the assessments is by the *discovery of topical experts*, in other words, searching for users with long experience in a particular topic who have both a high credibility score and are willing to act as moderators (Nassar & Karray, 2019). The proposed medium could discover patterns in users' activities by aggregating categories of validated sources (e.g., a user mostly comments on *#political #news* in *#USA*), together with users' validations to estimate their specialization area. That is referred to as *probability estimation*, which occurs when potential labels of an object are computed (Nassar & Karray, 2019). Users who get a tag of an expert in a specific topic have a higher impact on an aggregated evaluation within the proposed medium.

Further, a type of training should be available for all users involved in the proposed medium to improve *digital literacy* and to enhance the quality of individual credibility assessments. The training could have *the checklist approach as a basis*, which relies on a list of guidelines on how to evaluate particular online content. Users could test their knowledge, and, at the end of the training, collect badges, which later could be visible for other users in a specific user's profile.

Validation of users' profiles, not just in means of peer assessment, but to prevent the creation of multiple accounts is needed. Jessen (2012) points out that having a *known and verified identity* plays an important role when assessing information. Ways to validate the identities of users could include such methods as incorporated by *Hypothes.is*, which uses unique ORCID (Open Researcher and Contributor ID) digital profiles of researchers that engage in the medium (Perkel, 2015). Other ways to identify users are the involvement of a *public-key certificate* (e.g., e-signature, Bank-ID). That would limit the chances for a single user to create more than one account. Other ways of coping with the problem of account verification could be the use of an existing account on any social media platform. The chosen form of profile validation would be visible to all the members in order to communicate the value of a particular account better. Accounts verified with a public-key certificate would get a "verified" badge, which is the most trustworthy level of account, while those verified with, for instance, a Twitter account would only get a "Twitter verified" badge.

The proposed medium should be transparent by making it *open source*, which means that a community governs its activities and is not controlled by a government, institution, or organization,

which dictates its development. The utilization of an open source approach grants actors and communities the freedom to customize applications in accordance with their needs and wishes (Faraon, 2018b). In other words, as Brabham (2013) suggests, open-source productions promote cooperation amongst individuals with the purpose to produce common resources that are relevant for the same, using their terms and needs, as a self-governing community.

In the case of the proposed medium, it is crucial that it is created and controlled by members of a community to ensure that the intentions of the credibility assessments are not biased. *Transparency* could additionally be improved by allowing users to enhance the proposed medium, which could be made possible by adding a *suggestion box*. Moreover, a *conversation history* should be available to allow users to evaluate the whole process of credibility assessment, which means the possibility to see all the details of the impacts made by those users that were involved in an assessment.

During the external design critique activity, it was suggested the use of moderators who would aid in preventing different kinds of misbehavior. However, while the idea of moderation seems appealing, it may jeopardize the transparency of the system if, for instance, a moderator decides to block or punish a user without proper investigation of the situation at hand. Instead, we suggest the use of *block/unblock buttons* or even a “*mute*” button, leaving this kind of decision specifically at the hands of users.

The proposed medium should support *multilingualism* and *context-dependency*. Multilingualism is a necessary attribute for the reason that modern societies are linguistically diverse, which creates a need to support multilingual exchange among citizens (Faraon, 2018b). Multilingualism is also an important feature that creates inclusiveness and promotes consensus-seeking between users (Faraon, 2018b). Thus, the proposed medium should be able to provide the means for communication with the support of a translation mechanism, which could be adapted using third party services such as Google Translate, counting on the expertise of its users to improve the quality of the translation provided.

Context-dependency considers several aspects in the assessment process, namely *source* (e.g., individual, organization, company), *medium* (e.g., the Internet, mobile), and *content* (e.g., information). In the case of sources, the proposed medium provides ratings made by other users with all levels of knowledge and expertise. Trustworthiness and the expertness of the communicator are vital factors for the perceived credibility of a source, influencing other users' judgments on whether to accept or decline their messages (Choi & Stvilia, 2015). Moreover, the volume of ratings provided by other users is an important aspect that determines whether a source is credible or not (Metzger & Flanagin, 2013).

Further, concerning the aspect of the *medium*, it is not unusual that many users refer to “the Internet” or even “the computer” as a source of information (Choi & Stvilia, 2015). The Internet, as a medium, has gained remarkable trust since its inception. The credibility of the Internet as a source of information is because users have gained experience from using it regularly, which means that users have interacted with an array of online sources. Choi and Stvilia (2015, p. 2401) affirm that: “it is hardly possible to assign a value to the credibility of online information without having used the Web.” Hence, the proposed medium uses the Internet, not only as a source to find and seek information, but it allows other sources to be uploaded to it to improve the quality of the content that exists both offline and online.

Finally, concerning the aspect of content, the proposed medium offers the possibility to tailor its functionality to display only the content that is relevant for the users. Furthermore, the fact that the medium is open source, made by its users and for its users, makes it independent from any organizations or companies, which users perceive as unbiased, fair, and truthful (Choi & Stvilia, 2015). Besides, Fogg and Tseng (1999) point out that a site is perceived as credible when it contains links to external materials and sources. Therefore, the proposed medium should support the use of links that reinforce the statements contained in the users' reviews. Further, aggregated opinions seem to play an important role when judging the credibility of the sources. The higher the number of sites that contain the same piece of information, the higher is the number of users who find it credible (Choi & Stvilia, 2015).

With that in mind, the proposed medium should be designed to show the exact number of users who find a particular source credible or not credible.

4.2. Evaluating the Concept of Aggregated Credibility

Evaluation of the concept was conducted during the *external design critique*. The proposed concept of a medium for evaluation of sources online was presented to eight participants in order to evaluate and validate the artifact in its entirety: idea, concept, and associated theoretical principles. In the context of this article, the relevance of the collected feedback was determined by its capability to support the concept to the extent that it could be translated into a quality that improves the initial version of the proposed medium. This section focuses on the feedback gained through the concept-driven design research method, and its impact on the revisited design of the proposed medium for assessing the credibility of sources online. The obtained critique of the initial concept focused on issues such as validation of user account, anonymity, security, addressing eventual risks of users' misbehavior, and finally, on the possible improvements of the medium's functionalities.

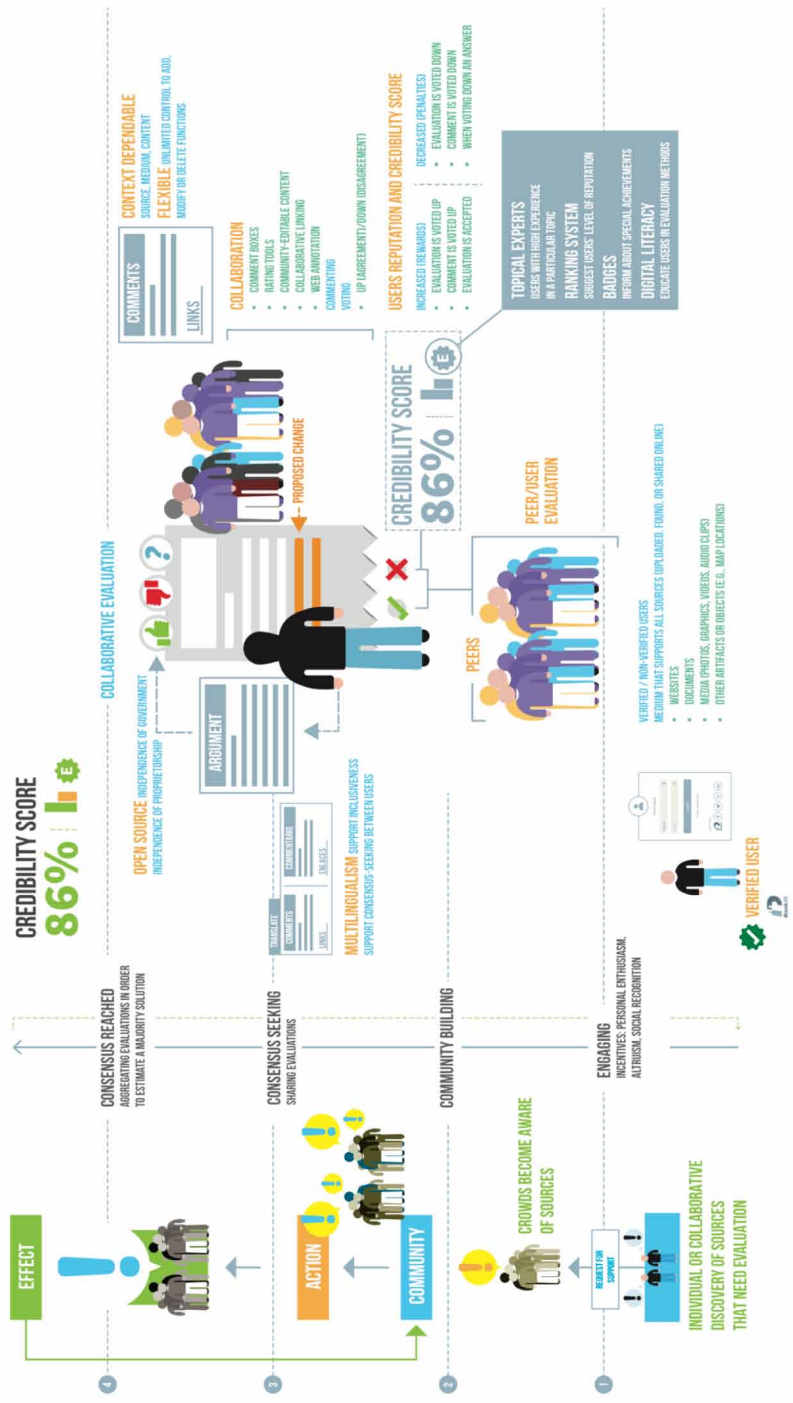
The validation of user accounts created a discussion that could be exemplified by the following quotation expressed by P1: "I would like to verify my account, but at the same time, I want to remain anonymous. I do not like the idea that other users can see my identity." P3 stated: "Good that the account is created with Bank-ID or other kinds of identification. But not so good if you create an account with social media login." At the same time, P2 proposed: "An ethical problem is how information from Bank-ID and login will be saved and handled." With all this said, the initial plan to include verification of a user account in the concept was partly reshaped to adjust to users' opinions. It was confirmed that other ways of validation of a user's profile are needed, such as validation with social media accounts. Moreover, the requirement to include the information on the kind of verification an account was submitted to was implemented in the revisited concept.

Concerns related to security was another topic that participants had the ambition to discuss. One particular commentary made by P3 was: "The medium can be unethical if it is used by pedophiles, criminals, etcetera." This statement signaled that the concept needed to be updated in order to allow users to express their disapproval in a situation of misbehavior. P5 suggested that a group of moderators should govern the system: "There could be moderators (as a kind of referee that tells people off and also educates people in good conduct) to keep order and a higher level of discussion." Thus, the initial concept has been updated with the notion of functionality that could answer the need for moderators.

Finally, some improvements in the functionalities of the proposed medium were suggested. Utilizing the idea of an open source feature that allows crowds to shape the proposed medium according to users' wishes and needs, it was proposed by P8: "A *suggestion box*: (like in companies), where people could contribute with ideas to make the tool better." Another participant, P5, suggested the use of *badges* as a way to ameliorate the medium: "You could also get badges. E.g., if you play a game or make a course you get badges that are added to your profile (bronze, silver, gold user instead of titles)", and P8 developed further the idea: "For the best contributions (in the suggestion box), the users could get points or badges." Another improvement, suggested by P2, included a *conversation history* in order to "check all corrections a post has been through."

Despite their concerns, all of the participants seemed to find the proposed medium an interesting option for checking the veracity of sources online. P6 expressed: "The functions of the proposed medium seem to complete each other and work harmoniously." P7 pointed out that: "It feels authentic. The rating system is nice. The proposed medium could be useful." In other words, the feedback gained in the *external design critique* confirmed that the proposed medium is well-grounded, but it signaled simultaneously that some changes needed to be done. Therefore, it was essential to reexamine the critique gained in order to determine whether a concept has flaws and imperfections. The suggestions mentioned above were taken into account in the revisited concept. In Figure 2, an abstract overview

Figure 2. On the right side, an abstract overview of the proposed medium in the form of a design artifact depicting the most crucial elements of the concept. On the left side, the underlying process of consensus-seeking in the credibility assessment process is depicted to highlight the concept's co-creative and participative values.



of the proposed medium is presented in the form of a design artifact depicting crucial elements in the revisited concept and the underlying process of consensus-seeking in the credibility assessment process.

5. CONCLUSION

Fake news is considered by the World Economic Forum (WEF) an alarming issue and to be one of the most significant challenges to society (Vicario et al., 2016). Research shows that we tend to easily conform to a majority recollection even if it proves to be wrong (Spinney, 2017). This, in turn, could lead to false beliefs and misperception, which may undermine democratic decision-making processes (Hameleers & van der Meer, 2019). Several methods and tools have emerged to address the issue of fake news, for example, fact-checkers in the form of editorial teams (e.g., Politifact), checklist approach, collaborative filtering (e.g., Hypothes.is), and web page ranking systems as Google Page Rank, Alexa Rank Checker, and MyWOT. However, the current body of methods and tools that aid users with the task of credibility assessment do not deal with the issue of evaluating any source or content that can be created or shared online.

The assessment of the credibility of sources online is a complex process because of the multitude of sources embedded in a large number of layers disseminated online (Sundar, 2008). Although previously mentioned methods and tools have improved political discourse and democratic accountability by guiding citizens to carefully examine the claims that appear in news stories and highlight which story is credible or not (Hameleers & van der Meer, 2019; Nyhan & Reifler, 2015), it is increasingly challenging for editorial teams to examine the majority of content that is continually being created, uploaded, and shared online.

The current work has showed how the concept-driven design research approach by Stolterman and Wiberg (2010) could be applied to anchor and underpin a design concept theoretically. The proposed concept of aggregated credibility in this article aims to support users through an online platform where individual assessments of all kinds of sources and content could be created and shared across the Internet in order to aggregate them in meaningful outputs. This may be achieved by facilitating possibilities for users to initiate and engage in massive collaboration to utilize the knowledge and competence of crowds to evaluate and improve the quality of content published online by exploiting collective judgment on matters of credibility.

The concept presented here is a good start in the right direction, judging from the reactions of the participants who evaluated the concept and expressed their support for the design of the proposed medium for online credibility assessment of sources. Future directions may expand on the proposed concept by practically advance it in the form of a prototype that should be evaluated with users. Such a prototype may consider incorporating existing models and frameworks related to trust management with the purpose to enhance reliability, privacy, and security (Mhetre, Deshpande, & Mahalle, 2016; Odella, 2016). The proposed concept may also be furthered by adopting an empirical perspective to collect data and use it to make theoretical contributions and concurrently optimize the concept. The challenges of fake news require a collective effort, which calls for different stakeholders to contribute to the practical development and evaluation of the concept.

REFERENCES

- Aggarwal, S., Van Oostendorp, H., Reddy, Y., & Indurkha, B. (2014). Providing web credibility assessment support. In *Proceedings of the 2014 European Conference on Cognitive Ergonomics* (pp. 1-29). New York, NY: ACM. <http://dx.doi.org/10.1145/2637248.2637260>
- Allahbakhsh, M., Benatallah, B., Ignjatović, A., Nezhad, H. R., Bertino, E., & Dustdar, S. (2013). Quality control in crowdsourcing systems: Issues and directions. *IEEE Internet Computing*, 17(2), 76–81. doi:10.1109/MIC.2013.20
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *The Journal of Economic Perspectives*, 31(2), 211–236. doi:10.1257/jep.31.2.211
- Amazeen, M. A. (2020). Journalistic interventions: The structural factors affecting the global emergence of fact-checking. *Journalism*, 21(1), 95–111. <https://doi.org/10.1177/1464884917730217>
- Barthel, M., Mitchell, A., & Holcomb, J. (2016). Many Americans believe fake news is sowing confusion. *Pew Research Center*. Retrieved from <https://www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/>
- Boididou, C., Papadopoulos, S., Zampoglou, M., Apostolidis, L., Papadopoulou, O., & Kompatsiaris, Y. (2018). Detection and visualization of misleading content on Twitter. *International Journal of Multimedia Information Retrieval*, 7(1), 71–86. doi:10.1007/s13735-017-0143-x
- Brabham, D. C. (2013). *Crowdsourcing*. MIT Press. doi:10.7551/mitpress/9693.001.0001
- Brawley, A. M., & Pury, C. L. S. (2016, January). Work experiences on MTurk: Job satisfaction, turnover, and information sharing. *Computers in Human Behavior*, 54, 531–546. doi:10.1016/j.chb.2015.08.031
- Breakstone, J., McGrew, S., Smith, M., Ortega, T., & Wineburg, S. (2018). Why we need a new approach to teaching digital literacy. *Phi Delta Kappan*, 99(6), 27–32. doi:10.1177/0031721718762419
- Brigida, M., & Pratt, W. R. (2017, November). Fake news. *The North American Journal of Economics and Finance*, 42, 564–573. doi:10.1016/j.najef.2017.08.012
- Carr, N. (2011). *The Shallows: What the Internet is doing to our brains*. W. W. Norton.
- Choi, W., & Stvilia, B. (2015). Web credibility assessment: Conceptualization, operationalization, variability, and models. *Journal of the Association for Information Science and Technology*, 66(12), 2399–2414. doi:10.1002/asi.23543
- Duke Reporters' Lab. (2019). *Fact-checking news*. Retrieved from <https://www.reporterslab.org/fact-checking/>
- Eldridge, S. A. II, García-Carretero, L., & Broersma, M. (2019). Disintermediation in social networks: Conceptualizing political actors' construction of publics on Twitter. *Media and Communication*, 7(1), 271–285. doi:10.17645/mac.v7i1.1825
- Eliasson, J. (2013). *Tools for designing mobile interaction with the physical environment in outdoor activities* (Doctoral Thesis). Stockholm University, Stockholm, Sweden.
- Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189–200. doi:10.1177/0165551512437638
- Faraon, M. (2018a). *Co-creating democracy: Conceptualizing co-creative media to facilitate democratic engagement in society* (Doctoral Thesis). Stockholm University, Stockholm, Sweden.
- Faraon, M. (2018b). Concept-driven design for democracy: Advancing co-creative media to support citizen participation and democratic engagement. *eJournal of eDemocracy and Open Government*, 10(1), 23–49. <https://doi.org/10.29379/jedem.v10i1.490>
- Faraon, M., Villavicencio, V., Ramberg, R., & Kaipainen, M. (2013). From mobilization to consensus: Innovating cross-media services to organize crowds into collaborative communities. In P. Parycek & N. Edelman (Eds.), *CeDEM13: Conference for E-Democracy and Open Government* (pp. 229–242). Edition Donau-Universität Krems.

- Fisher, C., Marshall, D., & McCallum, K. (2018). Bypassing the press gallery: From howard to hanson. *Media International Australia, Incorporating Culture & Policy*, 167(1), 57–70. doi:10.1177/1329878X18766077
- Fogg, B. J., & Tseng, H. (1999). The elements of computer credibility. In M. G. Williams, & M. W. Altom (Eds.), *CHI '99: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 80-87). New York: ACM Press. doi:10.1145/302979.303001
- Funke, D. (2018). Fact-checkers have debunked this fake news site 80 times. It's still publishing on Facebook. *Poynter.org*. Retrieved from <https://www.poynter.org/fact-checking/2018/fact-checkers-have-debunked-this-fake-news-site-80-times-its-still-publishing-on-facebook/>
- Giudice, K. D. (2010). Crowdsourcing credibility: The impact of audience feedback on web page credibility. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1–9. doi:10.1002/meet.14504701099
- Gross, M. (2017). The dangers of a post-truth world. *Current Biology*, 27(1), R1–R4. doi:10.1016/j.cub.2016.12.034
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5(1), eaau4586. doi:10.1126/sciadv.aau4586 PMID:30662946
- Guess, A., Nyhan, B., & Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. European Research Council, 9.
- Gupta, A., Lamba, H., Kumaraguru, P., & Joshi, A. (2013). Faking Sandy: Characterizing and identifying fake images on Twitter during hurricane Sandy. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 729-736). New York, NY: ACM. doi:10.1145/2487788.2488033
- Hameleers, M., & van der Meer, T. G. L. A. (2020). Misinformation and polarization in a high-choice media environment: How effective are political fact-checkers? *Communication Research*, 47(2), 227-250. <https://doi.org/10.1177/0093650218819671>
- Hammon, L., & Hippner, H. (2012). Crowdsourcing. *Wirtschaftsinformatik*, 54(3), 165–168. doi:10.1007/s11576-012-0321-7
- Hassan, N., Arslan, F., Li, C., & Tremayne, M. (2017). Toward automated fact-checking: Detecting check-worthy factual claims by ClaimBuster. *KDD: Proceedings / International Conference on Knowledge Discovery & Data Mining. International Conference on Knowledge Discovery & Data Mining*, 17, 1803–1812. <https://doi.org/10.1145/3097983.3098131>
- Hung, N. Q. V., Tam, N. T., Tran, L. N., & Aberer, K. (2013). An evaluation of aggregation techniques in crowdsourcing. In X. Lin, Y. Manolopoulos, D. Srivastava, & G. Huang (Eds.), *Lecture Notes in Computer Science: Vol. 8181. Web Information Systems Engineering - WISE 2013. WISE 2013*. Springer. doi:10.1007/978-3-642-41154-0_1
- Ishida, Y., & Kuraya, S. (2018). Fake news and its credibility evaluation by dynamic relational networks: A bottom up approach. *Procedia Computer Science*, 126, 2228–2237. doi:10.1016/j.procs.2018.07.226
- Jessen, J., & Jørgensen, A. H. (2012). Aggregated trustworthiness: Redefining online credibility through social validation. *First Monday*, 17(1-2). Advance online publication. doi:10.5210/fm.v17i1.3731
- Johansson, D., Lassnanti, J., & Wiberg, M. (2015). Mobile e-services and open data in e-government processes - concept and design. In M. Younas, I. Awan, & M. Mecella (Eds.), *Mobile Web and Intelligent Information Systems* (pp. 149–160). Springer. doi:10.1007/978-3-319-23144-0_14
- Johansson, D., & Wiberg, M. (2012). Conceptually advancing “application mobility” towards design: Applying a concept-driven approach to the design of mobile IT for home care service groups. *International Journal of Ambient Computing and Intelligence*, 4(3), 20–32. doi:10.4018/jaci.2012070102
- Kalir, J. H., & Dean, J. (2018). Web annotation as conversation and interruption. *Media Practice and Education*, 19(1), 18–29. doi:10.1080/14682753.2017.1362168
- Kavaler, D., & Filkov, V. (2018). Determinants of quality, latency, and amount of Stack Overflow answers about recent Android APIs. *PLoS One*, 13(3), e0194139. doi:10.1371/journal.pone.0194139 PMID:29547620

- Kazai, G. (2011). In search of quality in crowdsourcing for search engine evaluation. In P. Clough, C. Foley, C. Gurrin, G. J. F. Jones, W. Kraaij, H. Lee, & V. Mudoch (Eds.), *Lecture Notes in Computer Science: Vol. 6611. Advances in Information Retrieval. ECIR 2011* (pp. 165–176). Springer. doi:10.1007/978-3-642-20161-5_17
- Kim, A., & Dennis, A. R. (2019). Says who? The effects of presentation format and source rating on fake news in social media. *Management Information Systems Quarterly*, 43(3), 1025–1039. doi:10.25300/MISQ/2019/15188
- Klein, D. O., & Wueller, J. R. (2017). Fake news: A legal perspective. *Journal of Internet Law*, 20(10), 5–13.
- Liaw, R., Zilnik, A., Baldwin, M., & Butler, S. (2013). Maater: Crowdsourcing to improve online journalism. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)* (pp. 2549-2554). New York, NY: ACM. https://doi.org/10.1145/2468356.2468828
- Lim, C. (2018). Checking how fact-checkers check. *Research & Politics*, 5(3), 205316801878684. doi:10.1177/2053168018786848
- Liu, X., Nielek, R., Adamska, P., Wierzbicki, A., & Aberer, K. (2015, November). Towards a highly effective and robust web credibility evaluation system. *Decision Support Systems*, 79, 99–108. doi:10.1016/j.dss.2015.07.010
- Maras, M. H., & Alexandrou, A. (2019). Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos. *The International Journal of Evidence & Proof*, 23(3), 255–262. doi:10.1177/1365712718807226
- Masterton, G., & Olsson, E. J. (2018). PageRank's ability to track webpage quality: Reconciling google's wisdom-of-crowds justification with the scale-free structure of the web. *Heliyon*, 4(11), e00978. doi:10.1016/j.heliyon.2018.e00978 PMID:30761370
- Mazzola, D., & Distefano, A. (2010). Crowdsourcing and the participation process for problem solving: The Case of BP. In *Proceedings of ItAIS 2010 VII Conference of the Italian Chapter of AIS* (pp. 42-49). Academic Press.
- Meek, S., Jackson, M. J., & Leibovici, D. G. (2014). A flexible framework for assessing the quality of crowdsourced data. In J. H. Guijarro, S. Schade, & C. G. Canut (Eds.), *Connecting a Digital Europe through Location and Place. Proceedings of the AGILE'2014 International Conference on Geographic Information Science*. Academic Press.
- Mehrnezhad, M., Ghaemi Bafghi, A., Harati, A., & Toreini, E. (2017). PiSHi: Click the images and I tell if you are a human. *International Journal of Information Security*, 16(2), 133–149. doi:10.1007/s10207-015-0311-z
- Mendoza, M., Poblete, B., & Castillo, C. (2010). Twitter under crisis: Can we trust what we RT? In *Proceedings of the 1st workshop on social media analytics (SOMA '10)* (pp. 71-79). New York, NY: ACM. doi:10.1145/1964858.1964869
- Metzger, M. J., & Flanagin, A. J. (2013, December). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, 59, 210–220. doi:10.1016/j.pragma.2013.07.012
- Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D., & McCann, R. (2003). Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. In P. Kalbfleisch (Ed.), *Communication Yearbook 27* (pp. 293–335). Academic Press.
- Mhetre, N. A., Deshpande, A. V., & Mahalle, P. N. (2016). Trust management model based on fuzzy approach for ubiquitous computing. *International Journal of Ambient Computing and Intelligence*, 7(2), 33–46. doi:10.4018/IJACI.2016070102
- Nassar, L., & Karray, F. (2019). Overview of the crowdsourcing process. *Knowledge and Information Systems*, 60(1), 1–24. doi:10.1007/s10115-018-1235-5
- Nazzi, E., Bagalkot, N. L., Nagargoje, A., & Sokoler, T. (2012). Concept-driven interaction design research in the domain of attractive aging: The example of Walky. In P. Rodgers (Ed.), *Articulating design thinking* (pp. 227-245). Faringdon: Libris.
- Nyhan, B., & Reifler, J. (2015). The effect of fact-checking on elites: A field experiment on U.S. state legislators. *American Journal of Political Science*, 59(3), 628–640. doi:10.1111/ajps.12162
- Odella, F. (2016). Technology studies and the sociological debate on monitoring of social interactions. *International Journal of Ambient Computing and Intelligence*, 7(1), 1–26. doi:10.4018/IJACI.2016010101

- Parnin, C., Treude, C., Grammel, L., & Storey, M. A. (2012). *Crowd documentation: Exploring the coverage and the dynamics of API discussions on Stack Overflow*. Georgia Institute of Technology: Tech. Rep 11.
- Perkel, J. M. (2015). Annotating the scholarly web. *Nature*, 528(7580), 153–154. doi:10.1038/528153a PMID:26632591
- Rashed, K. A. N., Renzel, D., Klamma, R., & Jarke, M. (2014). Community and trust-aware fake media detection. *Multimedia Tools and Applications*, 70(2), 1069–1098. doi:10.1007/s11042-012-1103-3
- Robins, D., & Holmes, J. (2008). Aesthetics and credibility in web site design. *Information Processing & Management*, 44(1), 386–399. doi:10.1016/j.ipm.2007.02.003
- Seckler, M., Opwis, K., & Tuch, A. N. (2015). Linking objective design factors with subjective aesthetics: An experimental study on how structure and color of websites affect the facets of users' visual aesthetic perception. *Computers in Human Behavior*, 49, 375–389. doi:10.1016/j.chb.2015.02.056
- Shah, A. A., Ravana, S. D., Hamid, S., & Ismail, M. A. (2015). Web credibility assessment: Affecting factors and assessment techniques. *Information Research*, 20(1), 365–391.
- Singh, G., Bansal, D., Sofat, S., & Aggarwal, N. (2017, September). Smart patrolling: An efficient road surface monitoring using smartphone sensors and crowdsourcing. *Pervasive and Mobile Computing*, 40, 71–88. doi:10.1016/j.pmcj.2017.06.002
- Spinney, L. (2017). How facebook, fake news and friends are warping your memory. *Nature*, 543(7644), 168–170. doi:10.1038/543168a PMID:28277522
- Starbird, K., Maddock, J., Orand, M., Achterman, P., & Mason, R. M. (2014). Rumors, false flags, and digital vigilantes: Misinformation on Twitter after the 2013 Boston Marathon bombing. In *Proceedings of the 2014 iConference* (pp. 654-662). Academic Press.
- Stolterman, E., & Wiberg, M. (2010). Concept-driven interaction design research. *Human-Computer Interaction*, 25(2), 95–118. doi:10.1080/07370020903586696
- Su, X., & Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, 2009, 1-19. https://doi.org/10.1155/2009/421425
- Subramaniam, M., Taylor, N. G., St. Jean, B., Follman, R., Kodama, C., & Casciotti, D. (2015). As simple as that?: Tween credibility assessment in a complex online world. *The Journal of Documentation*, 71(3), 550–571. doi:10.1108/JD-03-2014-0049
- Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 73–100). MIT Press.
- Swedish Research Council. (2017). *Good research practice*. Swedish Research Council.
- Tandoc Jr, E. C., Wei Lim, Z., & Ling, R. (2018). Defining “fake news”. *Digital Journalism*, 6(2), 137–153. https://doi.org/10.1080/21670811.2017.1360143
- Trechsel, A. H. (2007). Inclusiveness of old and new forms of citizens' electoral participation. *Representation (McDougall Trust)*, 43(2), 111–121. doi:10.1080/00344890701363326
- van der Linden, S., Leiserowitz, A., Rosenthal, S., & Maibach, E. (2017). Inoculating the public against misinformation about climate change. *Global Challenges*, 1(2), 1600008. doi:10.1002/gch2.201600008 PMID:31565263
- Vicario, M. D., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. et al. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences of the United States of America*, 113(3), 554–559. doi:10.1073/pnas.1517441113 PMID:26729863
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. doi:10.1126/science.aap9559 PMID:29590045
- Westerwick, A. (2013). Effects of sponsorship, web site design, and google ranking on the credibility of online information. *Journal of Computer-Mediated Communication*, 18(2), 80–97. doi:10.1111/jcc4.12006

Whiting, M. E., Gamage, D., Gaikwad, S. S., Gilbee, A., Goyal, S., Ballav, A., & Bernstein, M. S. et al. (2017). Crowd guilds: Worker-led reputation and feedback on crowdsourcing platforms. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 1902-1913). New York, NY: ACM. doi:10.1145/2998181.2998234

Zhang, C., Gupta, A., Kauten, C., Deokar, A. V., & Qin, X. (2019). Detecting fake news for reducing misinformation risks using analytics approaches. *European Journal of Operational Research*, 279(3), 1036–1052. doi:10.1016/j.ejor.2019.06.022

APPENDIX: INTERVIEW QUESTIONS

- What is your general feeling about the system?
- What do you think is good about it?
- What problems do you identify with the system?
- Which parts could be improved and how?
- What kind of functions would you add/delete?
- Do you see any ethical problems with the use of the system?
- How do you think it could change your life and other people's lives?
- How could politicians and media influence by such system, provided that it became extremely popular?

Montathar Faraon is an interaction design researcher with an interest in co-creation, participatory processes, ICTs for democracy, and concept-driven design. He is an Assistant Professor in the Department of Design of Kristianstad University, Sweden.

Agnieszka Jaff holds both a M.Sc. in Journalism and Communication from the Jagiellonian University in Krakow, Poland and a B.Sc. in Informatics from Kristianstad University, Sweden. Her research interest lies in fake news, assessment of credibility online, and design to support users in their everyday activities.

Liegi Paschoalini Nepomuceno holds a B.Sc. in Informatics from Kristianstad University, Sweden. Her interests are digital creation, graphic design, and typography.

Victor Villavicencio is a graduate of Media Technology at Södertörn University, Sweden. He was the Swedish work package leader in the INTERREG IVA Central Baltic project IleP (Immigrants Inclusion by eParticipation, 2009-2012). He is a researcher, an experienced graphic designer with a demonstrated history of working in the design industry, and has a keen insight and knowledge of social media and digital collaboration tools.