

Credibility Analysis for Available Information Sources on the Web: A Review and a Contribution

Irvin Dongo

Univ. Bordeaux, ESTIA, Bidart, France
Electrical and Electronics Engineering Department
Universidad Católica San Pablo, Arequipa, Perú
irvin.dongo@ucsp.edu.pe

Yudith Cardinale

Dpto. de Computación y T.I
Universidad Simón Bolívar
Caracas, Venezuela
ycardinale@usb.ve

Ana Aguilera

Escuela de Ingeniería Informática
Facultad de Ingeniería
Universidad de Valparaíso, Chile
ana.aguilera@uv.cl

Abstract—In this paper, we present a study about the most popular information sources available on the Web (e.g., Google, Facebook, Twitter) and the available methods to verify their publications. We propose a generic credibility analysis model for social information sources, which is instantiated for *Twitter*. We show a proof of concepts through the development of *World White Web*, a Google Chrome extension application that implements the model to analyze *tweets* in real time, using *web scraping*. Our study demonstrates the feasibility and suitability of creating efficient methods to analyze the credibility of publication on social networks and points out the challenges and open problems that should be overcome in future solutions in this area.

Keywords—API, Google, Facebook, Twitter, Information Sources, Fake News, Credibility

I. INTRODUCTION

The Web is an open environment for publishing and exchanging data that is constantly developing. The information is generated through millions of Web pages, blogs created by users and organizations, as well as through social networks and initiatives such as *Open Data*. In this way, information is shared in free environments that can be used in various areas ranging from everyday life to the development of new technologies [1], [2]. However, the amount of information sources is huge and in many cases, they are not documented or validated, which makes it difficult to use and analyze.

All platforms on the Internet, which allow users to communicate, share, and generate information without formal references to sources, became popular in the early 1990s. Since then, the concept of *credibility* has become essential in various disciplines and from different perspectives, such as information engineering, business administration, journalism, information retrieval, human-computer interaction.

That is why the term *credibility* is confused with the term *trustness* or *truthfulness*. However, in the literature its differences have been established. While *credibility* is associated with the **level of belief that is perceived about (how credible it is)** a person, object, or process, *trustness* has to do with the **positive belief about the perceived confidence (reliability)** in a person, object, or process [3], [4].

Depending on the information source, different levels of credibility can be deduced, for example search engines such as Google and Bing, use positioning algorithms that show on the first pages the most relevant information (e.g., media, high-impact blogs) to a given query, while in social networks there

is no information classification. All these aspects are reflected in the perception of the credibility of the information sources in people. In this sense, some questions arise: *How much can we trust the information obtained from the Web? What is the credibility level perceived by users of these information sources? Can the credibility level of the information on the Web be calculated?* These and other concerns lead to the need of mechanisms to validate and classify such information.

In this context, this work has three main contributions: (i) we analyze the information sources with the greatest impact on the Web (e.g., *Indexers, Facebook, Twitter, Instagram, Web pages*), from a technical point of view (accessibility, format, tools offered for their manipulation, analysis and verification) and from the point of view of the credibility perception by users; (ii) we propose a credibility analysis model of publications on information sources, adaptable to various social networks; the credibility analysis is based on three measures: text credibility, user credibility (based on attributes such as creation date, verified account), and social credibility (based on attributes such as *followers* and *following*); and (iii) we show a proof of concepts through the development of *World White Web*, an application that implements the model, as an extension of Google Chrome; it performs the credibility analysis, in real time, of *tweets* using *web scraping*¹ since *Twitter* is the most used source of information in the literature.

This work demonstrates the feasibility of creating effective credibility analysis methods and exposes some challenges in this area, such as: (i) the variety of information sources, which makes it difficult to use general criteria of credibility; and (ii) the fact that information sources are limiting access to data, adopting forms of interaction such as *pay-to-use*.

II. INFORMATION SOURCES ON INTERNET

The information available on the Web is accessible by millions of people mainly through web browsers, such as Safari, Chrome, Edge. However, due to the enormous amount of information, users use search engines (e.g., Google, Bing) whose function is to index the information on the Web, acting as a filter and providing relevant information to a requested context. With the advance of the technology, the information

¹*Web scraping* is a technique for extracting information, focusing on the generation of structured data.

TABLE I
INFORMATION SOURCES AVAILABLE ON THE WEB

Source	Access Method	Level of Credibility (technical)	Access Limits	Frequency	Response Format	
Search Engine	Google	API	Low (Based on Ranking)	- 100 query/day (free) - 1000 queries (5 dollars) - 10000 query/day limited	-	JSON XML
	Bing	API	Low (Based on Ranking)	- 3000 query/month (free) - 1000 queries (7 dollars) and several options	3RPS 250RPS	JSON
Twitter	API	Zero	- 100 tweets per query - 500 tweets per query	30 RPM, 10 RPS (free) 60 RPM, 10 RPS (premium)	JSON	
Facebook	Graph API	Zero	-	-	-	
Instagram	Graph API	Zero	-	-	-	
Web Page	LOD	Download Link	High	Unlimited	Unlimited	RDF/XML
	General	Web Scraping	Zero	Unlimited	Unlimited	Text (Adapted functions)

shared has a more structured and detailed format, for example by using the Extensible Markup Language (XML), which allows an easy interpretation by computers of the represented data, adding basic information through descriptors (tags). XML is used to propose new formats in different contexts, such as sound synthesis, interactive videos [5]–[7].

Due to the enormous amount of information and data that can be extracted from the Web, other models of information representation have been proposed with the idea to be processed automatically. Resource Description Framework (RDF) is one of the most used models by the scientific community. From triplets (subject, predicate and object), real resources are described on the Web. The semantic is present in this model, allowing the inferring of more information [8]. Knowledge models as Web Ontology Language (OWL) allow the modeling of concepts and the relationship between them [9].

Some works are focused on the normalization of the information available on the Web through the use of order rules, reduction of redundant resources [10]–[12]. Other studies address the problem of similarity among resources for interoperability and the reuse of information such as in [13] [14], where XML documents are syntactically and semantically compared; in [15] [16] for information based on the RDF model and in [17] [18] for the ones based on ontologies. These models of information representation, are adopted not only by scientific communities, but by governmental organizations as well, that are part of open and linked data initiatives (*Linked Open Data - LOD*). Moreover, several works allow converting unstructured information into semantic models to increase productivity and reduce costs in services [19]. However, when sharing, the conversion to structured data and the similarity of the information do not ensure its veracity; therefore, its use for research is compromised by this vulnerability [20] [21]. There are works focused on measuring the level of credibility of the information applied on specific data sources, such as *Twitter* [3] [22]–[28], but nonexistent for others, as *Facebook*.

In this section, we first present a brief analysis of information sources on Internet, then we present a study of user perception of credibility on such information sources.

A. Analysis of Information sources on Internet

Table I shows the different information sources considered in this study. *Google* offers several possibilities to access the information, ranging from free-use but limited in handling, up to custom and manipulable searches and results. It provides a JSON API to perform queries. However, in order to manipulate

the information, customizable queries have a cost of 5\$ per one thousand requests. *Bing* allows queries limited to the number of transactions per second (*Requests per Second - RPS*), offering 3 RPS for free, but limited to 3000 per month. It offers several configurations, for example 30 RPS, 100 RPS, 150 RPS, ranging from 1\$ up to 7\$ per one thousand transactions. The query and response formats are JSON-like.

As a result of the latest events related to privacy, information sources as *Facebook* and *Twitter* have hardened their access policies to information, forcing users who wish to use such APIs to request explicit permissions. These requests for use of APIs are reviewed in detail by peers of the owners of the information sources, with a high rejection rate. However, *Twitter* is the most used information source for the analysis of social media information, such as sentiment analysis. It offers free access to public *tweets* maximum of 7 days and 30 days for categories *premium* and *enterprise* (*tweets* since 2006 for some configurations). From 100 to 500 *tweets* are obtained by each query, depending on the category level. Moreover, a maximum of 60 RPM (*Request per Minute*) and 10 RPS are allowed for the category *premium*. *Facebook* offers an API called *Graph API* that allows the creation of applications that request permissions from users (e.g., access to friends) to be able to access and analyze the information. *Facebook* does not specify the number of queries to the information. *Instagram*, being part of *Facebook*, has the same access policies.

The Web pages related to the initiatives of LOD (e.g., DBpedia), provide information in formats that allow the exchange and reuse of information on the Web such as RDF. The access is available through download links and its use is not restricted. In this context, *web scraping* is an alternative for data management, independent of the information source; however, the use of the values obtained from the Web pages may be protected by copyright or another type of restrictions. In addition, this technique must be adapted for each scenario, i.e., an implementation for *Facebook* is not compatible with one developed for *Twitter* due to different html structures.

At the level of technical credibility, which in this work is defined as the quality of the information provided by the sources of information (see third column in Table I), search engines such as *Google*, *Bing* use ranking based on search algorithms, showing relevant information towards a given search (e.g., news, high traffic pages). However, ranking techniques optimize Web pages to have a high value in the search engine ranking, reducing the quality of the information

shown and thus, needing a credibility analysis mechanism. For information sources such as *Twitter*, *Facebook*, and *Instagram*, the verification of the information is "zero", i.e., there is no a technique, ranking, etc., allowing us to infer a certain credibility level of information. In the case of *LOD*, these databases are created by government organizations, research teams, and other institutions that support the credibility of the information. In conclusion, the analysis of information from social media is essential to provide a level of credibility to users, where in recent years they have become the most used scenarios to share data but lack technical credibility criteria.

B. Human-Credibility Perception for Information Sources

Our initial hypothesis was that "the credibility perception of the information sources is directly related to the topic of information and to the resident country". In order to analyze our hypothesis, we conducted a survey in five different countries: France, Mexico, Peru, Venezuela, and Chile. Two questions were asked to measure the level of credibility in two different contexts: Q1: *What is your credibility confidence for topics related to political aspects (e.g., Donald Trump)?*; and Q2: *What is your credibility confidence for topics related to natural disasters (e.g., Amazon fire)?*. A total of 80 people per country participated in this activity with 88.30% of engineering and 52.10% of participants of Computer Science of the total².

Table II shows the results obtained from this survey. The information source with the highest credibility perception is *Google*, while the one with the lowest value is *Facebook*. It is clear that the latest news regarding privacy and the dissemination of false information against *Facebook* have directly impacted in the users. Moreover, we can observe that Q1, related to the political context has less credibility than Q2, which is focused on the context of natural disasters. In general, France has the lowest credibility level while Mexico the highest. A highlight is that Venezuela has a high credibility value except for *Facebook* where it only has 37.59% and 45.35% for Q1 and Q2, respectively. *Bing* is only 62.84% known by the users (see Table II column *Known*), being mostly known in Mexico as it is near to United States. On the other hand, it is unfortunately that the information source *LOD*, which has a high level of technical credibility, is only known by 66.85% of the users, which clearly evidences a lack of knowledge of the information source even more, being more than the half of users of Computer Science.

III. RELATED WORK

Twitter has been used in different studies to calculate the credibility of the shared information. In this section, we first reference studies that carry out an extensive review of works focused on credibility analysis in social networks, then we describe works specifically focused on credibility analysis on *Twitter* and highlight the differences with our proposal.

A. Surveys on credibility of information

In a recent study presented in [29], authors conducted an extensive review of papers that assess the credibility of the

information. There were reviewed 192 papers, from which around 92 papers dealt with social media analysis. From these 92 works, 90% use *Twitter* as the information source, 83% only perform an analysis on the text (10% perform analysis on text and images), 57% use the *Twitter* API, while the remaining 43% use commercial software to collect required *tweets* data, 77% of these works are focused on English language, 19% include Arabic, Japanese, and Chinese and only 4% consider other languages.

In another recent study presented in [30], a review of 211 works on false information in Web ecosystems is reported. This is achieved from a perspective of computational approaches focused to the problem of false information on the Web and its understanding. This work considers four aspects: (i) users' perception of false information; that is, how users perceive and interact with false information; the review includes works that use different methodologies including analysis of large databases from social networks, questionnaires/interviews, and crowdsourcing platforms; (ii) dynamic propagation of false information; this study intends to understand the dynamics of the flow through works that include data analysis techniques or mathematical and statistical approaches, until efforts that provide systems with visualization capabilities about this dynamic propagation; (iii) detection and contention of false information; Machine Learning techniques are used to detect false information and contention is fought with diffusion of correct information or information that refutes the false information; both cases require strong human-machine interaction; and (iv) false information in politics; after the US elections in 2016, the problem of spreading false information has aroused great interest in the scientific community, where political campaigns are responsible for the substantial dissemination of false political information and bots are frequently used. Authors emphasize that *Twitter* is one of the most used tools to disseminate false information and rumors. They also point out that it is possible to use automatic techniques for false information detection and credibility analysis.

These works demonstrate the trends in the use of information generated from social networks, in particular from *Twitter*, the growing interest of works focused on carrying out credibility analysis, and the challenges that have to be overcome in this area.

B. Credibility Analysis on Twitter

Concerning on how to do the credibility analysis, the existing works consider the extraction and analysis of different types of information to calculate different credibility measures. Thus, several terms of *credibility* have been proposed [24], [29], [31]. Inspired by these works, we propose the following classification of credibility terms in social networks:

Text Credibility (Post Credibility): measures the level of relevance and accuracy of the text, independent of the referenced topic [24] or with respect to a certain topic [29]. It is calculated through text analysis techniques.

User Credibility: calculates the user account credibility based on attributes that describe it. It can be calculated based on, for

²The survey is available at: <https://forms.gle/9yVGxT8tAuuwS5aAA>

TABLE II
HUMAN-CREDIBILITY PERCEPTION FOR SOME INFORMATION SOURCES

Country	Google			Bing			Twitter			Facebook		
	Known	Perception		Known	Perception		Known	Perception		Known	Perception	
		Q1	Q2		Q1	Q2		Q1	Q2		Q1	Q2
France	100%	55.63%	68.13%	56.25%	44.44%	47.00%	93.75%	48.67%	59.33%	97.77%	25.00%	35.00%
Mexico	97.78%	74.54%	76.67%	73.33%	53.33%	57.43%	91.11%	65.12%	69.55%	96.31%	51.70%	61.00%
Peru	100%	64.44%	71.55%	66.67%	46.67%	50.00%	88.99%	56.25%	61.00%	96.54%	42.22%	46.44%
Venezuela	98.90%	72.78%	80.90%	60.44%	50.72%	57.29%	97.80%	59.66%	68.09%	95.60%	37.59%	45.35%
Chile	100%	67.44%	72.79%	57.50%	43.04%	48.44%	86.25%	52.03%	60.72%	93.75%	42.27%	52.89%
Average	99.34%	66.97%	74.00%	62.84%	47.64%	52.03%	91.58%	56.35%	63.74%	95.99%	39.7%	48.14%

Country	Instagram			Any Website			LOD		
	Known	Perception		Known	Perception		Known	Perception	
		Q1	Q2		Q1	Q2		Q1	Q2
France	94.50%	32.33%	38.67%	93.75%	30.00%	42.63%	75.00%	50.00%	53.33%
Mexico	93.77%	52.00%	58.89%	95.56%	55.58%	62.44%	77.78%	56.86%	57.47%
Peru	96.54%	37.50%	39.33%	88.89%	50.13%	58.34%	63.54%	48.57%	51.24%
Venezuela	95.60%	44.25%	53.29%	90.11%	45.85%	52.98%	60.44%	52.18%	54.07%
Chile	93.75%	45.87%	53.42%	92.50%	45.68%	54.32%	57.50%	53.04%	55.00%
Average	94.83%	42.44%	48.72%	92.16%	45.45%	54.14%	66.85%	52.13%	54.22%

example, the account creation date, if the account is verified, user gender and age.

Social Credibility: calculates the credibility of a publication, related or not to a topic, based on the available metadata that describe the social impact of the user account and the *post* itself, with respect to other users. It is calculated based on data such as number of *followers*, number of *following*, *retweets*.

Topic-level Credibility: measures the level of acceptance of the topic or event referenced in the text. It consists of identifying if the text refers to a specific topic or not, usually through text and sentiment analysis techniques. Together, all these credibility measures attribute a global credibility level of a publication in a social information source. The following works are evaluated based on whether or not they consider these credibility levels.

Authors in [23] evaluate a *tweet* content through automatic text analysis using a Web miner and text sources in German (text credibility). In [22] the *Twitter* API is used to implement an algorithm that combines text search on a specific topic with social aspects of the account (*followers*), to measure the *tweet* credibility. In [3], in addition to considering the text content, characteristics associated with the user account are evaluated, such as its verified quality and its *Twittergrader.com*³ (text, user, and social credibility). In [24], authors mix characteristics associated with the message (text and sentiment analysis), the users (registration age), the social impact (number of *followers*, number of *following*, number of *tweets* the user has authored in the past, *retweets*), and a particular topic. In [25] two measures are proposed to calculate the topic-level credibility, one considers the *tweet* credibility based on the positive and negative opinions of the topic and the other one considers the author expertise. A combination of text content, account aspects, and social impact are applied on *TwitterBot* [26].

In [27], [28] the impact of the verification of an account is analyzed, concluding that it does not affect the user perception but should be taken into account in automatic credibility analysis systems (user credibility). In [32], [33], *TweetCred* and *CredFinder* are respectively described, two practical solutions

³*Twittergrader.com* measures the power, scope, and authority of a *Twitter* account based on: number of *followers*, *followers* impact, updates, credibility of news, *followers/following* relationship, and commitment.

proposed as Google Chrome extensions, which calculate the *tweet* credibility in real time, considering the content in the text, attributes of the *tweet* (publication time, source from which the *tweet* was posted), and social impact. Both use the *Twitter* API. Another practical solution for *Twitter*, which discredits rumors (rumor debunking) in real time is presented in [34]. This work takes into account attributes, extracted with the *Twitter* API associated with the user, the account, and the text. Additionally they use *web scraping* to search *tweets* similar to the analyzed rumor.

Table III summarizes the comparison of the reviewed works according to the four credibility levels and the analysis or implementation technique used (API, *web scraping*). The common characteristic of most works is that they do not consider the four levels of credibility. The only work that considers the four credibility levels is the one presented in [24], however, since it uses the *Twitter* API, which is now restricted, the work is no longer available. It should be noted that only 2.5% of the works cited in the study presented in [29] use mixed strategies.

TABLE III
RELATED WORK COMPARISON

Works	Credibility Measures				Extraction Techniques/Implementation
	Text	User	Social	Topic-level	
[23]	Yes	-	-	Yes	Machine Learning/Web Mining
[22]	-	-	<i>followers</i>	Yes	API
[3]	Yes	Account verification	<i>followers, following</i>	-	API
[24]	Yes	Age	<i>followers, following, retweets</i>	Yes	API, Machine Learning, Sentiment Analysis
[25]	-	User Expertise	-	Yes	Sentiment Analysis
[26]	Yes	Account verification	<i>followers, following, Location</i>	Yes	API, Text Analysis, Machine Learning
[27]	-	Account verification	-	-	Label Access
[28]	-	Account verification	-	-	Label Access
Real-Time Credibility Analysis					
[32]	Yes	Time of tweet Source	<i>followers, following, retweets</i>	-	API, Machine Learning, Text/Sentiment Analysis, Google Chrome extension
[33]	Yes	Time of tweet Source	<i>followers, following, retweets</i>	-	API, Text Analysis, Google Chrome extension
[34]	Yes	Account verification Source, Name	<i>retweets, Location</i>	-	API, SVM, Web scraping (to search similar tweets)
Ours	Yes	Account verification, Account creation (year)	<i>following, followers</i>	-	Web scraping, Google Chrome extension

From the works showed in the Table III, *TweetCred* [32],

CredFinder [33] and Liu et al. proposal [34] are similar to our proposal: they are solutions that analyze credibility in real time, consider the credibility of text, user, and social, independent of the topic, and are proposed as extensions of Google Chrome (except the one of Liu et al. [34], which is a stand alone application). However, the three solutions have lost their validity and are not available because they use the *Twitter* API that was open at that time, but from 2016 its use was restricted. In contrast, our proposal is based on *web scraping* to extract the necessary attributes to calculate the three credibility measures. The analysis of the text is done through filters that detect SPAM, bad words, misspelling. We do not consider yet including Machine Learning algorithms or Sentiment Analysis. However, such algorithms in the context of specific topics are under consideration in our ongoing research work. The proposed model is flexible enough to easily incorporate them, thereby completing all four aspects of credibility of a *tweet*.

IV. MODEL TO CALCULATE CREDIBILITY IN SOCIAL NETWORKS: OUR PROPOSAL

We propose a model to measure the credibility level of texts from sources of information in social networks. First, we present our formal and general model, then we describe how to apply it for the credibility analysis of *tweets*.

A. Model Formalization

According to the different credibility measures presented in Section III, our proposed model for calculating the credibility level of a *post* considers text credibility, user credibility, social credibility, regardless of the topic treated. First we establish, in Def. 1, the different elements that can be associated and extracted from a *post*, in any social network.

Definition 1: Post (p). A *post*, denoted as p , is a structure composed by:

- a text, denoted as $p.text$, that announces a fact, an event, a comment, etc.;
- an account to which it belongs, denoted as $p.user$ and consisting of a tuple with elements that describe the user (e.g., name, age, sex, account creation year);
- a set of metadata that describes its social impact, denoted as $p.social$; it considers social aspects related to the account ($p.social_{user}$) and related to the *post* ($p.social_{post}$).

The definition of the structure of a *post* adapts to the social network in question. For example, in the case of *Twitter* a *post* corresponds to a *tweet*, say t , with $t.social_{post} = \langle retweets, likes \rangle$. In addition, this definition can be extended to consider that the content of the *post* is not just text and may include multimedia data. These aspects are being considered in our future work. Def. 2 presents the formal definition of credibility level of a *post*.

Definition 2: Post Credibility Level ($PCred$). Given a *post* p , the *Post Credibility Level* is a function, denoted as $PCred(p)$, that returns a measure $\in [0, 100]$, of its level of credibility, defined as:

$$PCred(p) = weight_{text} \times TextCred(p.text) +$$

$$weight_{user} \times UserCred(p.user) +$$

$$weight_{social} \times SocialCred(p.social), \text{ where:}$$

- $weight_{text}$, $weight_{user}$, and $weight_{social}$ represent the weights that the user gives to text credibility, user credibility, and social credibility, respectively, such as: $weight_{text} + weight_{user} + weight_{social} = 1$;
- $TextCred(p.text)$, $UserCred(p.user)$, and $SocialCred(p.social)$ represent the credibility measure related to the text, the user, and the social impact of p , respectively.

The text credibility is based on measures that analyze the text, as presented in Def. 3.

Definition 3: Text Credibility ($TextCred$). Given a text of a *post*, $p.text$, the *Text Credibility* is a function, denoted as $TextCred(p.text)$, that returns a measure $\in [0, 100]$, derived from text analysis techniques.

Our model allows to incorporate any text analysis technique, such as Machine Learning, Sentiment Analysis, or Syntax Analysis. As stated by Def. 4, the user credibility is based on attributes that describe the user and his account.

Definition 4: User Credibility ($UserCred$). Given an account of a *post*, $p.user$, the *User Credibility* is a function, denoted as $UserCred(p.user)$, that returns a measure $\in [0, 100]$, derived from the attributes that describe the user account from which p was published.

The social credibility level is measured based on the available metadata, both related to the user account and the analyzed *post*, which describe the social impact or relate the user to the other users of the social network (see Def. 5).

Definition 5: Social Credibility ($SocialCred$). Given a set of social metadata of a *post*, $p.social$, the *Social Credibility* is a function, denoted as $SocialCred(p.social)$, that returns a measure $\in [0, 100]$, derived from the attributes that describe the social impact of the user account ($p.social_{user}$) and the *post* itself ($p.social_{post}$).

The attributes of interest to compute of the user credibility ($UserCred$) and the social credibility ($SocialCred$), depend on the analyzed social network, the technique used to extract the used attributes, and the interest of the developer of the credibility analysis system. Consequently, the respective functions that calculate these credibility measures ($UserCred(p.user)$ and $SocialCred(p.social)$) depend on the attributes considered.

For social information sources such as *Facebook* and *Twitter*, the numerical relation between data such as *followers* and *following* (related to the user account), and *likes* and *retweets* (related to the publication), certainly have an impact on the credibility level of a publication. In the next section, we show the functions defined for both cases, in this study.

B. Model Application on Twitter

Based on the credibility analysis model proposed in the previous section, we propose an application on *Twitter*. Thus, the definitions that conform the model are adapted to our particular case, as shown below.

Def. 6 adapts Def. 1 to describe the structure of a *tweet*, since the *posts* on *Twitter* are the so-called *tweets*.

Definition 6: Tweet (t). A *tweet*, denoted as t , is a structure composed by:

- a *text*, denoted as $t.text$, that announces a fact, an event, a comment, etc.;
- an *account* to which it belongs, denoted as $t.user$ and consisting of a tuple with elements that describe the user account, such as: $t.user = \langle user_id, year_of_creation, validated/non_validated \rangle$;
- a *set of metadata* that describes its social impact, denoted as $p.social$; it considers social aspects related to the account ($t.social_{user} = \langle followers, following \rangle$) and related to the post ($t.social_{tweet} = \langle retweets, likes \rangle$).

The *tweet* credibility level is calculated according to Def. 2. Each credibility measure is calculated according to the respective definitions and with specific attributes and functions.

1. Text Credibility: As stated by Def. 3, it is necessary to decide the text analysis technique to be used and the function that returns the credibility measure according to that analysis. In this study, it was decided to use syntactic analysis techniques through SPAM, bad words, and misspelling filters. **SPAM Filter:** SPAM messages are usually unwanted advertising messages, which usually use hyperbolic language, excess capitalization, and accentuation. If the analyzed text has these characteristics, it may not contain useful information, so its credibility level may decrease. Def. 7 introduces the SPAM filter function.

Definition 7: SPAM Filter ($isSPAM$). Given a text of a *tweet*, $t.text$, the SPAM filter is a function, denoted as $isSPAM(t.text)$, that returns a measure $\in [0, 100]$, indicating the probability of $t.text$ being a SPAM, defined as:

$$is_SPAM(t.text) = detect_spam(t.text)$$

where $detect_spam$ is an algorithm that analyzes $t.text$ and returns the probability of $t.text$ being a SPAM.

There are many free and open source algorithms to detect SPAM that can be used to implement this $isSPAM$ filter.

Bad Words Filter: It is used to detect those publications that have a high content of bad words, in which case their credibility should be reduced. Def. 8 presents how this filter is used to calculate the proportion of bad words in the text.

Definition 8: Bad Words Filter (bad_words). Given a text of a *tweet*, $t.text$, the bad words filter is a function, denoted as $bad_words(t.text)$, that returns a measure $\in [0, 100]$, indicating the proportion of bad words in $t.text$, defined as:

$$bad_words(t.text) = 100 - \left(\frac{extract_bw(t.text)}{extract_tw(t.text)} \right) \times 100$$

where $extract_bw$ and $extract_tw$ are algorithms that counts the number of bad words and total words in $t.text$, respectively.

There are also many free and open source algorithms in this domain, which can be used to implement this filter.

Misspelling Filter: It is used to detect syntax errors in writing text (Def. 9). Credibility will decrease as the number of misspellings found.

Definition 9: Misspelling Filter ($misspelling$). Given a text of a *tweet*, $t.text$, the misspelling filter is a function, denoted as $misspelling(t.text)$, that returns a measure $\in [0, 100]$, indicating the proportion of misspelled words in $t.text$, defined as:

$$misspelling(t.text) = 100 - \left(\frac{extract_mp}{extract_tw(t.text)} \right) \times 100$$

where $extract_mp$ and $extract_tw$ are algorithms that counts the number of misspelled words and total words in $t.text$, respectively.

As in the previous cases, there are many free and open source algorithms that analyze the spelling correctness of a text. Based on these three filters, the text credibility is defined as shown by Def. 10. User decides the importance of each filter.

Definition 10: Text Credibility ($TextCred$). Given a text of a *tweet*, $t.text$, the Text Credibility is a function, denoted as $TextCred(t.text)$, that returns a measure $\in [0, 100]$, defined as: $TextCred(t.text) = w_{SPAM} \times isSpam(t.text) + w_{BadWords} \times bad_words(t.text) + w_{MisspelledWords} \times misspelling(t.text)$

where w_{SPAM} , $w_{BadWords}$, and $w_{MisspelledWords}$ represent the weights that the user gives to each filter, respectively, such as $w_{SPAM} + w_{BadWords} + w_{MisspelledWords} = 1$.

2. User Credibility. To calculate this credibility measure on *Twitter*, we consider whether the account is verified or not (Def. 11) and the activity time of the account since its creation (Def. 12).

Definition 11: Verified Account Weight ($Verif_Weight$). Given an account of a *tweet*, $t.user$, the weight of a Verified Account is a function, denoted as $Verif_Weight(t.user)$, that returns a measure $\in \{0, 50\}$, defined as:

$$Verif_Weight(t.user) = \begin{cases} 50 & \text{if } t.user \text{ has been verified,} \\ 0 & \text{otherwise.} \end{cases}$$

Definition 12: Account Creation Weight ($Creation_Weight$). Given an account of a *tweet*, $t.user$, the weight of the Account Creation is a function, denoted as $Creation_Weight(t.user)$, that returns a measure $\in [0, 50]$, defined as: $Creation_Weight(t.user) = \frac{Account_Age(t.user)}{Max_Account_Age(t.user)} * 50$ where:

- $Account_Age(t.user) = Current_Year - Year_Joined(t.user)$
- $Max_Account_Age(T) = Current_Year - Twitter_Creation_Year$ where $Twitter_Creation_Year$ is the year in which *Twitter* was created (2006).

While the year account was joined ($Year_Joined$) is closer to the *Twitter* creation date (2006), the account credibility is greater. The maximum points obtained by this criterion is 50, since the other 50 is for the verified account weight ($Verif_Weight$). This attribute should have a high degree of importance as evidenced by recent work [27], [28]. With these two measures, we adapt Def. 4 as shown by Def. 13.

Definition 13: User Credibility ($UserCred$). Given an account of a *tweet*, $t.user$, the User Credibility is a function, denoted as $UserCred(t.user)$, that returns a measure $\in [0, 100]$, defined as: $UserCred(t.user) =$

TABLE IV
FOLLOWERS AND FOLLOWING FOR VERIFIED *Twitter* ACCOUNTS

Country	Name	Twitter ID	Followers	Following	Category
France	Élysée	@Elysee	2.3M	293	Nacional Institution
	Équipe de France	@equipefrance	4.5M	349	International Institution
	Louane	@louane	1M	1403	Artist
	Antoine Griezmann	@AntoGriezmann	6.5M	7	Soccer Player
Mexico	Fiscalía General de la República	@FGRMexico	706k	243	Nacional Institution
	National Soccer Selection	@miseleccionmx	6M	1467	International Institution
	Maná	@manaoficial	2.8M	69	Artist
	Chicharito Hernández	@CH14_	9.1M	486	Soccer Player
Peru	Presidencia Perú	@presidenciaperu	699.5K	323	Nacional Institution
	Federación Peruana de Fútbol	@TuFPF	1.1M	395	International Institution
	Gian Marco	@gianmarcomusica	3.7M	137	Artist
	Luis Advíncula	@luisadvincula17	271.2K	194	Soccer Player
Venezuela	Asamblea Nacional	@AsambleaVE	1.1M	419	Nacional Institution
	Miss Venezuela	@MissVzla	949.9K	268	International Institution
	Nacho	@nacholacriatura	2.5M	219	Artist
	Tomás Rincón	@TomasRincon5	701.7K	206	Soccer Player
Chile	Gobierno de Chile	@GobiernodeChile	1.2M	325	Nacional Institution
	Uber Chile	@Uber_Chile	51.6K	3704	International Institution
	Javiera Mena	@javieramena	171.4K	784	Artist
	Alexis Sánchez	@Alexis_Sanchez	2.9M	36	Soccer Player
Average	-	-	2.7M	566	-

$$Verif_Weight(t.user) + Creation_Weight(t.user)$$

3. Social Credibility. The behavior of accounts that are considered reliable and other untrustworthy was observed. Thanks to that, it was determined that untrusted accounts usually follow a greater number of users than the number of users that follow it. Similarly, trusted accounts have a large followers and follow a few people. It also was noted that the difference for these attributes in both types of accounts is extremely large. Table IV shows the *followers* and *following* of some verified accounts on *Twitter* from different countries and categories. They have an average of 2.7M of *followers*, while only to follow 566 accounts (*following*). Based on this observation, we consider that if the number of *followers* is around 2 million, an account can be considered highly reliable. In this work, we only consider the social impact of the account ($t.social_{user}$) based on the number of *followers* and *following*.

We calculate the influence of an account based on the number of *followers*, as shown in Def. 14.

Definition 14: Followers Impact (*FollowersImpact*). Given a set of social metadata of an account from a tweet t is published, $t.social_{user}$, the social *Followers Impact* is a function, denoted as $FollowersImpact(t.social_{user})$, that returns a measure $\in [0, 50]$, defined as:

$FollowersImpact(t.social_{user}) = \frac{t.social_{user}.followers}{MAX_followers} \times 50$ where $t.social_{user}.followers$ is the number of *followers* of $t.user$ and $MAX_followers$ is a user parameter.

The maximum value for $MAX_followers$ is a parameter provided by the system user. For example, according to our analysis shown in the Table IV, in this work we consider $MAX_followers = 2millones$. The influence of accounts that exceed the value of *followers* is calculated based on that maximum established value. We calculate the ratio between *followers* and *following*, as shown in Def. 15.

Definition 15: Followers-Following Proportion (*FFProportion*). Given a set of social metadata of an account from tweet t is published, $t.social_{user}$, the *Followers-Following Proportion* is a function, denoted as $FFProportion(t.social_{user})$, that returns a measure $\in [0, 50]$, defined as: $FFProportion(t.social_{user}) = \frac{t.social_{user}.followers}{t.social_{user}.followers + t.social_{user}.following} \times 50$ where $t.social_{user}.followers$ and $t.social_{user}.following$ are the number of *followers* and *following* of $t.user$, respectively.

Finally, social credibility is calculated based on the influence of the account (Def. 14) and the proportion of *followers* and *following* (Def. 15), each one with a maximum weight of 50. Def. 16 shows the calculation of social credibility.

Definition 16: Social Credibility (*SocialCred*). Given a set of social metadata of a tweet, $t.social$, the *Social Credibility* is a function, denoted as $SocialCred(t.social)$, that returns a measure $\in [0, 100]$, defined as: $SocialCred(t.social) = FollowersImpact(t.social_{user}) + FFProportion(t.social_{user})$

With the three credibility measures the credibility level of a tweet is calculated, as shown in Def. 17 adapted from Def. 2.

Definition 17: Tweet Credibility Level (*TCred*). Given a tweet, t , the *Tweet Credibility Level* is a function, denoted as $TCred(t)$, that returns a measure $\in [0, 100]$, of its level of credibility, defined as:

$$TCred(t) = weight_{text} \times TextCred(t.text) + weight_{user} \times UserCred(t.user) + weight_{social} \times SocialCred(t.social)$$

where:

- $weight_{text}$, $weight_{user}$, and $weight_{social}$ represent the weights that the user gives to text credibility, user credibility, and social credibility, respectively, such as: $weight_{text} + weight_{user} + weight_{social} = 1$;
- $TextCred(t.text)$, $UserCred(t.user)$, and $SocialCred(t.social)$ represent the credibility measure related to the text, the user, and the social impact of t , respectively.

The parameters considered in the filters and the weight that each filter represents in the final credibility calculation are values that users can configure in the system. In the next section we describe the implementation of our credibility model applied to *Twitter*, as an extension of Google Chrome.

V. CREDIBILITY ANALYSIS IN REAL-TIME: THE *World White Web* APPLICATION

This section presents the proof of concepts of the model proposed in the previous section, through the implementation of an application able of performing credibility analysis of *tweets* in real time. To develop this application, it was necessary to decide how to obtain the attributes to calculate the different credibility measures. From the analysis of the different information sources and the kind of available attributes, we first planned to use the *Twitter* APIs to extract such as attributes and calculate the user and social credibility measures. However, due to the recent privacy and security policies of *Facebook* and *Twitter*, it is currently not possible to easily access its API.

It is necessary for *Twitter* to grant permission as a developer. Even when we tried to obtain these permissions, our requests were repeatedly rejected when explaining our research project. Thus, due to these restrictions, we decided to use the technique of *web scraping* on the Web pages of the corresponding social network. Even when this technique is susceptible to updates made by social networks, it has advantages in terms of the possibility of real-time analysis without relying on proprietary licenses or access restrictions.

The credibility analysis application on *Twitter*, called *World White Web*, was implemented as an extension of Google Chrome. It includes a view of options, which allows to customize the criteria used when calculating credibility of the information. These criteria are related to: (i) the weights that the user wants to grant to each credibility measure (*TextCred*, *UserCred*, and *SocialCred*); (ii) the weights for the different filters used to calculate the credibility of the text (*isSpam*, *bad_words*, and *misspelling*); and (iii) the maximum number of *followers* (*MAX_Followers*) to calculate the social impact (*FollowersImpact*). For example, on *Twitter*, given the size restrictions of the *post*, users often use abbreviations and contractions, therefore the text credibility based on syntactic analysis should have a low weight or the misspelling filter could be overridden. The filters to calculate the text credibility were implemented as follows:

The **SPAM filter** uses a library called *Simple Spam Filter*, taken from a JavaScript development platform (NPM). The library offers the ability to process a text and obtain a *Boolean* value that describes whether it is SPAM or not. These values are translated in the filter as 0 or 50, if the *Boolean* is false or true, respectively (see Def. 7).

The **bad words filter** is another library published in NPM, called *Bad Words*, which censures the bad words achieved in the text. The filter uses this library to count the number of words that should be censored in the text and compares this number with the total number of words in it.

The **good spelling filter** works with the same calculation as the bad words filter. It gets the misspelled words and compares them to the total words present in the text. To determine which words are well written, we use another library published in NPM called *Nspell*.

The attributes to calculate the user and social credibility measures are extracted from the profile page of *Twitter*, through *web scraping*: amount of *followers*, amount of *following*, the joined year to *Twitter* of the account, and if it has the seal of *verified account*. The application offers two ways to analyze the information: (i) an interface to analyze any type of text, independent of the website where it is; and (ii) back-end that performs, in real time, the credibility analysis on *Twitter*. Each one is detailed below:

1. Text to Analyze: It allows to analyze a given particular text, inserted directly into the text box with the title *Text to Analyze*. The application returns a credibility value just below the title *Credibility* in the *pop-up* window. This analysis makes use of *isSPAM*, *bad_words*, and *misspelling* filters to process

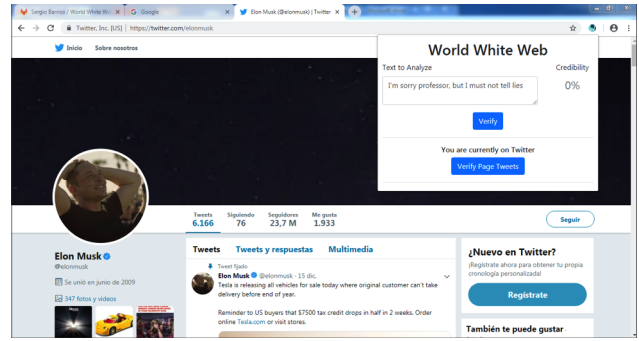


Fig. 1. Pop-up view of *World White Web* on *Twitter*

the text (*TextCred*), as long as it is not found on *Twitter*. If the browser is located on a *Twitter* profile, the information obtained from that page will also be used to calculate user and social credibility.

2. Verify Page Tweets: When browsing a user profile on *Twitter*, the *Verify Page Tweets* option appears in the *pop-up* window of the application (see Figure 1). When using this option, the *tweets* visible on the page are processed according to the credibility analysis model and the level of credibility value associated with each *tweet* is returned, inserting it as text in the body of each one. Figure 2 shows the credibility values obtained for the Elon Musk *Twitter* account. The color of this text is changed to reflect how close or far from 100% the value is: the closer to 100% it is, the greener the text, the closer to 0%, the redder the text is colored. The current version of *World White Web*, only supports the English language, due to the use of libraries that analyze the text. In our future work we plan to consider other languages, such as Spanish and French.

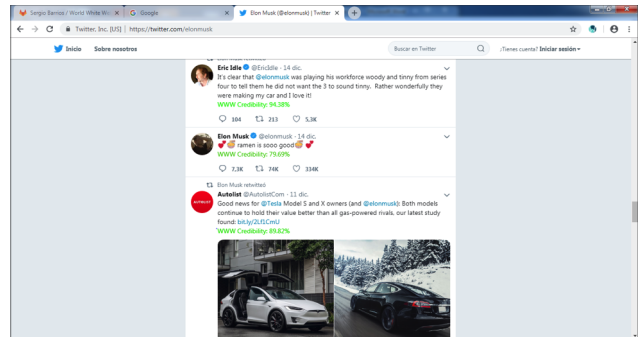


Fig. 2. Result of applying the option *Verify Page Tweets*

VI. GENERAL DISCUSSION

This work points out the feasibility and suitability of the implementation of a credibility analysis model for social networks, through the development experience of the application *World White Web*, as proof of concept of that model and applied to *Twitter*. This work, also highlights the most relevant aspects related to the information sources on the Web, the most used approaches to analyse their content and evaluate the publication credibility, as well as the challenges to be overcome in this matter.

A. The Credibility Analysis Model

In the proposed credibility analysis model, we consider several aspects, independently of the topic, that combine

structural, user, and social characteristics to measure the level of credibility of publications in Web information sources. It allows to set up different criteria to compute such as level of credibility, mainly by assigning weights for the different credibility aspects. These weights depend on the social information source and on the techniques used to perform the credibility analysis. For example, it is possible that the weight of the text credibility, *TextCred* (Def. 3) should be low for *Twitter*, since the length restrictions for the text forces users to use abbreviations and contractions; while for Facebook, this credibility aspect should have more impact in the global credibility level. Another aspect that can be extended in the model is regarding the analysis of images. The post structure (Def. 1) can be extended to consider not only text but multimedia content, and thus to consider image recognition techniques, text extraction from images, and their analysis with existing tools.

Furthermore, the model allows to integrate and define different text analysis techniques (e.g., Machine Learning, Sentiment Analysis) to refine such as credibility measure, as well as to consider different attributes and filters to calculate the user and social credibility measures, according to the social information source being analyzed. For example, for Facebook the attributes such as *likes* and number of comments should have an impact in the credibility measure; meanwhile for *Instagram*, attributes as *likes* and *comments* have more sense to be considered. All these possibilities, make the model flexible to be extended and adapted to any social network.

B. World White Web: Credibility Analysis on Twitter

Currently, *Twitter* is the social information source most popular and considered in recent studies in the credibility analysis area. There exist many works focused on analyzing *tweets* to verify their credibility, as well as in other domains, such as sentiment analysis and geolocation, thanks to the possibilities of extracting information from *Twitter*.

This generalized interest on *Twitter*, led us to extend the proposed credibility model to be applied on *Twitter* and to demonstrate the proof of concept with the development of the application *World White Web*. This application was developed as a Google Chrome extension, based on Node.js libraries and *web scraping* techniques to extract information and to interpret social attributes, in order to compute a representative measure of credibility in a *Twitter* account.

Recently, the owners of the main Web information sources (e.g., Google) and social networks (e.g., *Twitter*, *Facebook*) have restricted the information access. These restrictions are consequence of the recent privacy policies imposed by the European Union⁴ and due to several events that highlighted the lack of security and information control⁵. Although we agree with these privacy protection policies, we consider that it is important that these information owners offer options to allow studies and analysis to support research in this area.

In any case, APIs are restricted to be freely used, thus *web scraping* becomes a real possibility for such purposes.

World White Web instantiates the proposed model by using filters to analyze the text syntax (i.e., with *isSPAM*, *bad_words*, and *misspelling* filters), thus be able to obtain a measure of the text credibility, *TextCred*. It is possible to integrate more filters to increase the robustness of the credibility model. In particular, we consider that filters that work for other languages (e.g., Spanish, French) could be good complements. Another aspect in the text analysis, that is currently considered in recent works is related to the search of references in the *tweet*; i.e., determine if there are references to reliable information sources in the *tweet*. All these aspects, as well as aspects related to semantic analysis of the text (e.g., sentiment analysis, geolocation) to have measures of credibility related to a specific topic, can be easily integrated through the model. This is part of our future work.

Regarding user and social credibility, we agree with the state of the art, in the use of attributes such as account verification, *followers*, *following* (see Table III), which are easily extracted with *web scraping*. In our ongoing work, we are considering to incorporate other attributes describing the social impact of the *tweet*. Each *tweet* has information that can be analyzed independently of the user that posted it (*likes*, *retweets*, etc.).

The use of the *web scraping* technique and the HTML manipulation to implement *World White Web*, show the possibility to extend its functionality and usability, by considering the HTML structure for *Facebook*. The application will be able to automatically detect if the social network analyzed is *Facebook* or *Twitter* and act accordingly. *World White Web* is characterized by its simplicity, but it demonstrates the feasibility of implementing automatic approaches to analyze credibility in real time and with the available information.

C. Open Problems

According to the reviewed literature, besides the common used attributes (e.g., *followers*, *following*, *retweets*, *likes*) to compute the credibility level of publications in social networks, some studies have demonstrated the impact of considering if the account is verified [27], [28] and the location from which the post was done [26], [34]. In this work, we have only considered the first aspect (verified account); the second one (location) is being considered in our ongoing work.

An issue that has to be considered in any proposal of credibility analysis is the evaluation of the quality of their results. In this sense, the state of the art lacks of studies of the precision of the models due to the lack of datasets and benchmarks to perform an objective evaluation; hence, it is imperative the creation of testbeds and datasets.

Recent literature evidences that *Twitter*, as a source of information, has been widely studied and continues to be interesting for the scientific community, due to the enormous size of information available and its accessibility. However, there are many interesting challenges to overcome. There is no previous work that has rigorously assessed the credibility based on the user profile. Addressing the problem from the point of view of multiplatforms and multiformats is still on

⁴https://ec.europa.eu/info/privacy-policy_en

⁵<https://www.bbc.com/news/technology-46618582>

the early developing stage [29]. Authors in [30] state that there is not a decisive work that can establish a systematic credibility metric in social networks to detect and evaluate the erroneous and false information integrating reliability and reputation. There is neither a system or organization that analyzes, from a more systemic perspective, the messages shared in social networks, thus users believe, distribute, and accept the information as it is presented to them. In this sense, there are still many open problems that need to be addressed.

VII. CONCLUSIONS AND FUTURE WORKS

We analyze the available information sources on the Web, their access method, the frequency of access and their format. We also propose a model to calculate the credibility of *posts* in social networks, based on three credibility measures: text, account/user, and social impact. These measures are calculated based on filters that consider text analysis (SPAM, bad words, and good spelling) and account attributes (e.g., *creation date*, *followers*). A proof of concepts to show the feasibility of the model, called *World White Web*, was developed. This application uses the *web scraping* technique to extract the attributes used to compute the credibility measures. Even though the application approach was used only in the context of *Twitter*, it can be easily extended and adapted to other information sources.

Our future research is focused on improving our credibility model taking into account additional attributes. Also, we propose to create a database for *Twitter* in order to objectively evaluate our proposal against the state of the art.

REFERENCES

- [1] Y. Kammerer, E. Kalbfell, and P. Gerjets, "Is this information source commercially biased? how contradictions between web pages stimulate the consideration of source information," *Discourse Processes*, vol. 53, no. 5-6, pp. 430-456, 2016.
- [2] J. Slomian, O. Bruyère, J.-Y. Reginster, and P. Emonts, "The internet as a source of information used by women after childbirth to meet their need for information: A web-based survey," *Midwifery*, vol. 48, pp. 46-52, 2017.
- [3] H. S. Al-Khalifa and R. M. Al-Eidan, "An experimental system for measuring the credibility of news content in twitter," *Internat. Journal of Web Information Systems*, vol. 7, no. 2, pp. 130-151, 2011.
- [4] B. Hillgoss and S. Y. Rieh, "Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context," *Informat. Processing & Manag.*, vol. 44, no. 4, pp. 1467-1484, 2008.
- [5] K. Wouters, B. Preneel, A. I. González-Tablas, and A. Ribagorda, "Towards an xml format for time-stamps," in *Proc. of the ACM Workshop on XML Security*, 2002, pp. 61-70.
- [6] B. Torrente, S. J. Barro, A. Dapena, and C. J. Escudero, "Defining an xml format for sound synthesis," in *Proc. of the Internat. Conf. on Ubiquitous Information Management and Communication*, 2011, pp. 111:1-111:4.
- [7] B. Meixner and H. Kosch, "Interactive non-linear video: Definition and xml structure," in *Proc. of Symp. on Document Eng.*, 2012, pp. 49-58.
- [8] B. M. Graham Klyne, Jeremy J. Carroll, "RDF 1.1 Concepts and Abstract Syntax," <https://www.w3.org/TR/rdf11-concepts/>, 2014.
- [9] W3C-OWL-Working-Group, "OWL 2 Web Ontology Language - Document Overview (Second Edition)," <https://www.w3.org/TR/owl2-overview/>, 2012.
- [10] T. Pankowski and T. Pilka, "Dealing with redundancies and dependencies in normalization of xml data," in *Proc. of Internat. Multiconf. on Computer Science and Information Tech.*, Oct 2008, pp. 543-550.
- [11] M. Sridhar, M. R. S. Mitra, M. N. Santhi, M. V. P. Kumar, and P. R. Sudha, "Normalization in xml and data exchange," in *Proc. of Internat. Conf. on Distance Learning and Education*, 2010, pp. 232-236.
- [12] R. Ticona-Herrera, J. Tekli, R. Chbeir, S. Laborie, I. Dongo, and R. Guzman, "Toward rdf normalization," in *Proc. of Internat. Conf. on Conceptual Modeling*, Cham, 2015, pp. 261-275.
- [13] T. Kim, J. Lee, and J. Song, "Semantic structural similarity for clustering xml documents," in *Proc. of Internat. Conf. on Convergence and Hybrid Information Technology*, 2008, pp. 552-557.
- [14] R. Guzman, I. Dongo, and R. T. Herrera, "Structural and semantic similarity for xml comparison," in *Proc. of Internat. Conf. on Management of Emergent Digital EcoSystems*, 2013, pp. 177-181.
- [15] R. De Virgilio, A. Maccioni, and R. Torlone, "A similarity measure for approximate querying over rdf data," in *Proc. of the Joint EDBT/ICDT Workshops*, 2013, pp. 205-213.
- [16] P. Maillot and C. Bobed, "Measuring structural similarity between rdf graphs," in *Proc. of Symposium on Applied Comp.*, 2018, pp. 1960-1967.
- [17] K. N. Ismail and Z. A. Bakar, "Ontology structure comparison," in *Conf. on e-Learning, e-Management and e-Services*, Dec 2013, pp. 148-151.
- [18] Y. Ma, L. Liu, K. Lu, B. Jin, and X. Liu, "A graph derivation based approach for measuring and comparing structural semantics of ontologies," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 5, pp. 1039-1052, 2014.
- [19] K. Hassanzadeh, M. Reformat, W. Pedrycz, I. Jamal, and J. Berezowski, "T2r: System for converting textual documents into rdf triples," in *Proc. of Internat. Joint Conf. on Web Intelligence and Intelligent Agent Technologies*, vol. 3, 2013, pp. 221-228.
- [20] E. Go, K. H. You, E. Jung, and H. Shim, "Why do we use different types of websites and assign them different levels of credibility? structural relations among users' motives, types of websites, information credibility, and trust in the press," *Computers in Human Behavior*, vol. 54, pp. 231-239, 2016.
- [21] A. J. Flanagan, M. J. Metzger, R. Pure, A. Markov, and E. Hartsell, "Mitigating risk in ecommerce transactions: perceptions of information credibility and the role of user-generated ratings in product quality and purchase intention," *Electronic Commerce Research*, vol. 14, no. 1, pp. 1-23, 2014.
- [22] K. R. Canini, B. Suh, and P. Piroli, "Finding relevant sources in twitter based on content and social structure," in *NIPS Workshop*, 2010.
- [23] A. Juffinger, M. Granitzer, and E. Lex, "Blog credibility ranking by exploiting verified content," in *Proc. of Workshop on Information Credibility on the Web*, 2009, pp. 51-58.
- [24] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in *Proc. of Internat. Conf. on WWW*, 2011, pp. 675-684.
- [25] Y. Namihira, N. Segawa, Y. Ikegami, K. Kawai, T. Kawabe, and S. Tsuruta, "High precision credibility analysis of information on twitter," in *Proc. of Internat. Conf. on Signal-Image Technology & Internet-Based Systems*, 2013, pp. 909-915.
- [26] K. Lorek, J. Suehiro-Wiciński, M. Jankowski-Lorek, and A. Gupta, "Automated credibility assessment on twitter," *Computer Science*, vol. 16, no. 2, p. 157, 2015.
- [27] T. Vaidya, D. Votipka, M. L. Mazurek, and M. Sherr, "Does being verified make you more credible?: Account verification's effect on tweet credibility," in *Proc. of Conf. on Human Factors in Computing Systems*, 2019, pp. 525:1-525:13.
- [28] S. Edgerly and E. K. Vraga, "The blue check of credibility: Does account verification matter when evaluating news on twitter?" *Cyberpsychology, Behavior, and Social Networking*, vol. 22, no. 4, pp. 283-287, 2019.
- [29] M. Alrubaian, M. Al-Qurishi, A. Alamri, M. Al-Rakhami, M. M. Hassan, and G. Fortino, "Credibility in online social networks: A survey," *IEEE Access*, vol. 7, pp. 2828-2855, 2019.
- [30] S. Zannettou, M. Sirivianos, J. Blackburn, and N. Kourtellis, "The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans," *Journal of Data and Information Quality*, vol. 11, no. 3, pp. 10:1-10:37, 2019.
- [31] B. Kang, T. Höllerer, and J. O'Donovan, "Believe it or not? analyzing information credibility in microblogs," in *Proc. of Internat. Conf. on Advances in Social Networks Analysis and Mining*, 2015, pp. 611-616.
- [32] A. Gupta, P. Kumaraguru, C. Castillo, and P. Meier, "Tweetcred: Real-time credibility assessment of content on twitter," in *Proc. of Internat. Conf. on Social Informatics*, 2014, pp. 228-243.
- [33] M. AlRubaian, M. Al-Qurishi, M. Al-Rakhami, M. M. Hassan, and A. Alamri, "Credfinder: A real-time tweets credibility assessing system," in *Proc. of Internat. Conf. on Advances in Social Networks Analysis and Mining*, 2016, pp. 1406-1409.
- [34] X. Liu, A. Nourbakhsh, Q. Li, R. Fang, and S. Shah, "Real-time rumor debunking on twitter," in *Proc. of Internat. on Conf. on Information and Knowledge Management*, 2015, pp. 1867-1870.