

RESEARCH ARTICLE

A Risk-Based Approach for Enhancing the Fitness of Use of VGI

TAREK SBOUI¹ AND SAIDA AISSI²¹GREEN-TEAM Laboratory, University of Tunis Al-Manar, Tunis 1068, Tunisia²SMART Laboratory, Higher Institute of Management, Tunis 2000, Tunisia

Corresponding author: Saida Aissi (saida.aissi@gmail.com)

ABSTRACT Volunteered Geographic Information (VGI) phenomena offer an alternative or supplement to the authoritative mechanism of geospatial data acquisition. It allows people without professional geospatial skills or knowledge to participate in the geospatial data collection. VGI has been boosted by recent advances in geospatial technology and applications. VGI applications have shown great potential in various areas such as disaster management and public health. However, VGI suffers from a lack of quality assurance, because VGI contributors may lack knowledge of the geospatial domain and credibility. Moreover, VGI data may have different levels of detail and precision, and may have been collected for different purposes. Appropriate VGI data for a specific application may be less appropriate for another application. End-users may use VGI data without being aware of its appropriateness to their requirements. This may cause a risk that arises when end-users inappropriately use, or be uncertain about using VGI data in their applications. This risk may undermine the VGI project in which end-users are involved. This paper proposes an approach that aims to enhance VGI quality assurance by measuring the spatio-semantic similarity between user requirements and provided VGI, and to evaluate the fitness-for-use of VGI. The proposed approach is based on an algorithm to help VGI end-users to deal with the risks related to the quality of VGI data. The approach helps end-user make appropriate decisions about VGI data (e.g., considering or not considering VGI data and being careful when using VGI data).

INDEX TERMS Fitness-for-use, risk management, volunteered geographic information.

I. INTRODUCTION

In the last decade, with the rapid development of geographic information systems (GIS) and remote sensing (RS) and the use of the Internet for GIS applications, there has been an important increase in the availability and openness of geospatial data. The availability of geospatial data has expanded with the emergence of volunteer geographic information (VGI). VGI refers to the use of the web to access, assemble, and disseminate geospatial data provided voluntarily by individuals [7], [15], [20]. VGI aims to allow people without professional geospatial skills or knowledge to participate in collecting geospatial data (e.g., OpenStreetMap). VGI has been widely adopted in many application areas such as disaster management, reporting the spread of avian influenza, and traffic management [2].

The associate editor coordinating the review of this manuscript and approving it for publication was Amjad Mehmood¹.

However, the quality of VGI data may be insufficient for a particular use. In fact, VGI contributors may provide geospatial data without necessarily having knowledge of the geospatial domain [4]. In addition, contributors may lack credibility (e.g., involved in vandalism or sending false information). Moreover, VGI data may be produced by many contributors who have used different technologies and tools. As such, VGI data may have different levels of detail and precision, and may have been collected for different purposes [43].

Consequently, VGI data may be characterized by uncertainty and insufficient appropriateness for a particular use [27], [39], [42]. Such uncertainty may cause a risk of misusing data and may undermine the VGI project in which end-users are involved [2].

To effectively respond to the uncertainty related to geospatial data in the context of VGI, end-users should be aware of the VGI data quality with regard to their application

(i.e., intended use). Thus, evaluating and enhancing the fitness-for-use of VGI has become increasingly important.

The aim of this paper is to provide a step forward in making users aware of how appropriate VGI data are for their intended use, and to enhance the fitness-for-use of VGI. We propose an approach to help end-users to deal with the risk of misusing VGI data. To this end, we first propose a spatio-semantic similarity measure between the user requirements and the provided VGI. Then, we evaluate a set of indicators to assess the fitness-for-use of VGI data. This evaluation will help end-users make appropriate decisions about geospatial data in the context of VGI. For example, if the quality of VGI data is very low, it may not be ideal to use it.

In the next section, we review some works that proposed approaches and algorithms to evaluate and enhance the quality of VGI data. In Section III, we discuss the risk of misuse of geospatial data in the context of VGI, and we present our proposed approach to enhance the fitness-for-use of VGI data. Our approach consists of a spatio-semantic similarity measure between user requirements and the provided VGI, an assessment of a set of indicators for VGI quality, and an algorithm that aims to help end-users make appropriate decisions with regard to VGI quality. In Section IV, we propose a prototype, we developed, to implement our approach. In section V, we present some experiments and results. Section VI presents a comparison with related state-of-the-art methods. Finally, we conclude the study and present further works.

II. RELATED WORKS

The quality of VGI data can be described by guidelines and indicators [5], [43]. Quality guidelines have been identified mainly by standardization organizations (e.g., the International Organization for Standardization (ISO)). Some researchers have based their work on such guidelines to reach a certain level of compatibility between VGI data and reality. This is done mainly by comparing the VGI data to authoritative data as a reference. Authoritative data normally follow standardization guidelines, and may be available from commercial or governmental sources [11], [16], [24], [29].

Regarding quality indicators, suggested methods aim to describe aspects of VGI data quality to enhance its reliability. The main standards organizations (e.g., ISO 19157, ICA, FGDC, and CEN) have defined a set of quality indicators. The commonalities among them are accuracy (thematic, positional, and temporal), consistency, and completeness [6], [22]. Completeness describes the extent to which elements of the reference dataset are present in geospatial dataset. This can be attributed to a lack of data. Consistency is the degree of adherence to the rules of data structure, attribution and relationships [28]. Accuracy is the closeness of the collected data to the truth [31].

Some researchers have based their work on the above mentioned indicators to evaluate the quality of data produced from the VGI process [17], [43]. To evaluate

completeness, researchers compared VGI with authoritative data. These studies are based on comparing the length in a grid map [18], [24], [46], comparing the total area [19], [30], or comparing VGI data with the ground truth [33].

Consistency is measured by comparing different dataset objects with other objects of the same theme (intra-theme consistency) or objects of other themes (inter-theme consistency) [5], [19]. Ali and Schmid [4] focused on detecting inconsistencies in VGI related to incorrect data classifications. Other studies have focused on assessing topological relationships in VGI, either based on mathematical techniques to determine topological consistency [12], or based on the similarity between spatial objects in multi-representation [26].

With regard to accuracy, some studies used trained volunteers to evaluate and enhance the closeness of the VGI data to the ground truth [36]. Other studies have assessed the accuracy of VGI with regards to authoritative data. Al-Bakri and Fairbairn [3] assessed spatial accuracy by comparing VGI and authoritative data, and found that the errors were quite high in the context of VGI. The authors attributed these errors to the fact that data contributors usually use low-precision instruments such as personal GPS devices to collect data.

In addition to the most common indicators (i.e., accuracy, consistency, and completeness), there are more abstract indicators to assess the quality of VGI, such as trustworthiness, credibility, and vagueness [43].

Trustworthiness can be expressed as the time at which the contributor has been registered with the VGI project [44]. With regard to VGI credibility, some studies have proposed crowd-intelligence-based approaches to evaluate and correct errors. Other works have proposed the intervention of gatekeepers (i.e., the social approach) [21], [25]. Regarding vagueness evaluation in the context of VGI, De Longueville *et al.* [13] proposed an approach that consists of automatically capturing the scale at which VGI data is produced. VGI produced at lower scales is classified as vague.

Previous works have mainly compared VGI data with authoritative data as a reference. However, authoritative data has limitations including the availability and accuracy. In fact, authoritative data may be unavailable or missing, and hence it is not possible to assess the quality of VGI. Accessing authoritative data, if available, is often restricted by constraints or it costs relatively much [5], [27].

In addition, while the aforementioned studies assessed some indicators to evaluate and enhance VGI quality, they did not focus on the context of a particular intended use. In fact, VGI data originally collected for a given purpose could be used for another purpose. Although it could be considered appropriate for the original purpose, it may be considered less appropriate for another purpose.

Also, existing studies have proposed solutions in a non-systematic manner. That is, they have not been based on predefined or ordered phases. As a result, end-users have to put extensive time and effort into identifying the risks related to quality of VGI data. Even with such intensive efforts, they

may fail to identify the severity of the risks of misusing VGI data. Consequently, existing approaches for evaluating VGI data quality remain vulnerable to the risk of data misuse.

Accordingly, and inspired by the methods of risk management in the field of project management [37], we defined an approach to deal with the risks of data misuse in VGI. The approach is based on evaluating a set of indicators of VGI quality and a risk management method that aims to facilitate decision-making regarding the risks of VGI misuse.

Our proposed approach to risk management consists of evaluating the risks related to VGI quality, and reacting to the risks by making appropriate decisions related to the use of VGI data.

III. ENHANCING THE FINTESS OF USE OF VGI

A. RISK RELATED TO VGI QUALITY

End-users typically do not have a proper idea of VGI data quality. Consequently, they may make incorrect assumptions regarding the use of VGI data. Such assumptions have the potential to expose end-users to the risk of data misuse. This risk is characterized by the probability that end-users inappropriately use or be uncertain about using VGI data in their applications (or integrating VGI data with their existing data), and by the damage that can be caused by such inappropriate use or uncertainty.

We illustrate the risk of data misuse with the following example: a VGI contributor provides georeferenced image taken by mobile phone (see Figure 1). This contributor sent the image with no detailed information about the geospatial features shown in the image (i.e., lack of metadata). Without more complete, clear, or detailed indications, the quality of VGI may be insufficient and lead users to false assumptions, such as what black polygon within a road refers to; a pothole or bulge on the road? Such uncertainty may cause a risk of inappropriate data usage.

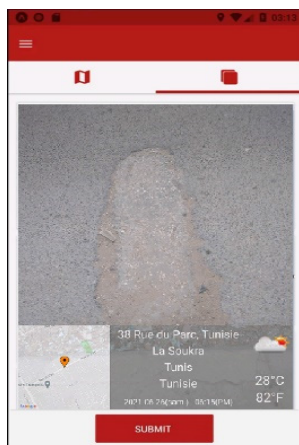


FIGURE 1. Uncertainty about geospatial features shown in the image.

Our method for identifying and responding to the risks of VGI misuse is based on evaluating the fitness-for-use, that is,

quality information about a dataset's suitability for a particular application or conformance to a set of requirements.

While a good fitness-for-use of VGI indicates that it is less likely to have a risk of data misuse; insufficient fitness-for-use indicates a higher risk. Consequently, evaluating the fitness-for-use of VGI allows to identify and evaluate of the risk of data misuse.

Moreover, evaluating the fitness-for-use of VGI facilitates the response to the risks of data misuse. Indeed, based on the fitness-for-use evaluation, end-users can be advised:

- Not to use a dataset that has low fitness-for-use.
- To use a dataset which has a good fitness-for-use.
- Carefully consider a dataset that has satisfactory fitness-for-use.

In our approach, we propose a set of indicators to evaluate the fitness-for-use of the VGI data with regard to user requirements. Therefore, we first propose a spatio-semantic similarity measure between user requirements and the provided VGI data.

B. SIMILARITY MEASURE BETWEEN USER REQUIREMENTS AND PROVIDED VGI

The idea is to identify the similarity between the provided VGI information and the explicit user requirements. The development of a similarity measure is then a fundamental step in VGI quality assessment. Two aspects of similarity were considered: (1) semantic similarity and (2) spatial similarity. In the following sections, we detail and propose a semantic similarity measure as well as a spatial similarity measure between user requirements and provided VGI description. The spatial similarity measure integrates a topological distance, a metric distance, and an orientation distance between user requirements and provided VGI description.

1) SEMANTIC SIMILARITY MEASURE BETWEEN USER REQUIREMENTS AND PROVIDED VGI

The proposal of the semantic similarity measure between user requirements and the provided VGI is detailed in the next section. We first present a formalization of some basic concepts.

a: USER REQUIREMENTS REFERENCES

We suppose that the user presents a query describing the VGI. Let Q : the set of the data elements presenting the user requirements regarding VGI.

We define the set of references corresponding to user requirements as $Q_R = \{R_Q, R_M \dots R_q^i\}$, where

$\{R_Q, R_M \dots R_q^i\}$ is the set of keywords (concepts) introduced in the user requirements; $1 \leq i \leq n$; i and n are positive integers.

b: PROVIDED VGI REFERENCES

We assume that the user can present a description of the VGI provided. This description is decomposed into a set of keywords describing the provided VGI.

Let M : the set of Provided VGI description.

We define the set of references corresponding to the provided VGI description as follows:

$$M_R = \{R_m^1, R_m^2 \dots R_m^j\} \text{ where:}$$

$\{R_m^1, R_m^2 \dots R_m^j\}$ is the set of keywords (concepts) introduced in the user requirements.

c: SEMANTIC SIMILARITY MEASURE BETWEEN USER REQUIREMENTS AND PROVIDED VGI

The similarity measures proposed in literature are based on the knowledge representation model presented in ontologies and semantic networks [41]. The concepts in our proposal are represented by references. To compute the semantic distance between references and concepts, we refer to an edge-counting method based on the Rada distance [40] using an application ontology representing the different concepts of the provided VGI. The Rada distance computes the minimum number of edges that separate concepts in the ontology. We opted for the Rada distance for semantic similarity evaluation because it is simple, accurate, and efficient [40], [41].

Let $A = (S_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ denote the matrix of the semantic distances between the references R_q^i ($1 \leq i \leq n$) of the user requirement and the references R_m^j ($1 \leq j \leq n$) of the VGI description; $1 \leq i \leq n$ and $1 \leq j \leq n$;

S_{ij} : Distance between the reference R_q^i of the user requirements (Q) and the reference R_m^j of the provided VGI description (M) using the Rada distance basing on the knowledge representation model offered by the application ontology.

The semantic distance between the user requirements (Q) and the VGI description (M), denoted as $D_{sem}(Q, M)$, is obtained as follows:

$$D_{sem}(Q, M) = \sum_{i=1}^n \sum_{j=1}^n S_{ij}; \quad 1 \leq i, j \leq n \quad (1)$$

The semantic similarity measure is derived from the semantic distance as follows:

$$Sim_{sem}(Q, M) = 1/(1 + D_{sem}(Q, M)). \quad (2)$$

2) SPATIAL SIMILARITY MEASURE BETWEEN USER REQUIREMENTS AND PROVIDED VGI

In the literature, spatial similarity assessment is mainly based on the topology, direction, and distance between spatial elements [8], [32]. Thus, to compute the similarity between user requirements and the provided VGI, we define a topological distance, metric distance, and directional distance between VGI user requirements and the provided VGI description.

In the next section, we propose and formalize three different distances between user requirements and provided VGI: an orientation distance, a topological distance, and a metric distance, as well as a spatial distance that integrates the three proposed distances.

a: ORIENTATION DISTANCE BETWEEN USER REQUIREMENTS AND PROVIDED VGI

To access the orientation distance between user requirements and the provided VGI, we propose computing the distance between the orientations of spatial objects presented in each of them. Therefore, to compute the orientation distance between spatial objects in user requirements and provided VGI, we use the graph of spatial directions and the costs of transformation defined in the TDD model [1], [32]. According to this model, nine types of orientations are defined (north, northwest, west, southwest, south, southeast, east, northeast, and equality), and the cost of moving from one direction into a close direction is equal to 2. Figure 2 shows the conceptual neighborhood and cost of transformation from one direction to another, as defined in the TDD model.

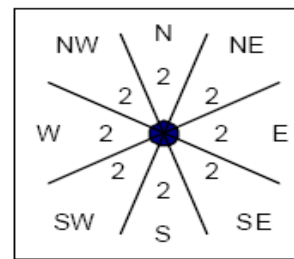


FIGURE 2. Graph of spatial directions and transformation costs [32].

Definition:

Let Q : the set of the data elements presenting user requirement's regarding VGI.

Let M : the set of Provided VGI description

Given user requirements (Q) and provided VGI description (M),

$OB_Q = (ob_q^1, ob_q^2, \dots, ob_q^i)$ where $ob_q^1, ob_q^2, \dots, ob_q^i$ are the spatial objects introduced in user requirements regarding VGI.

$OB_M = (ob_M^1, ob_M^2, \dots, ob_M^j)$ where $ob_M^1, ob_M^2, \dots, ob_M^j$: the spatial objects added in the Provided VGI description.

Let $B = B_{ij}$: Denote the matrix used to measure the orientation distance between the spatial objects in user requirements and the spatial objects in the provided VGI, $1 \leq i \leq n$, $1 \leq j \leq m$, and b_{ij} : the direction distance between object i in Q and object j in M ;

$a_{ij} = MCT(ob_q^i, ob_M^j)$: represents the minimum cost of transformation of the orientation of object i in Q to the orientation of object j in M using a conceptual neighborhood network.

The distance in terms of the direction between user requirements (Q) and the provided VGI (M), denoted as $D_{dir}(Q, M)$, is computed as follows:

$$D_{dir}(Q, M) = \sum_{i=1}^n \sum_{j=1}^m b_{ij}; \quad 1 \leq i \leq n \text{ and } 1 \leq j \leq m \quad (3)$$

b: TOPOLOGICAL DISTANCE BETWEEN USER REQUIREMENTS AND PROVIDED VGI

Computing the topological distance between user requirements and provided VGI description includes evaluating the topological distance between pairs of spatial objects in each of them.

To evaluate the topological distance between spatial objects, we proposed in [1] to use the conceptual neighborhood proposed by Li and Fonseca [32] in the TDD model for spatial similarity assessment. The graph is shown in Figure 3.

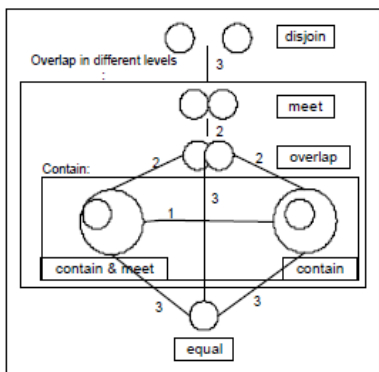


FIGURE 3. Conceptual neighborhood network and topological relationships [32].

In the TDD model, the authors propose to decompose the conceptual neighborhood graph into three groups of topological relationships. The distance between two arcs in the same group (intra-group) is equal to 2. However, the distance between two arcs belonging to two different groups is equal to 3. Except for the distance between the two nodes (meet and overlap), where the transformation cost is equal to one. Figure 3 presents the conceptual neighborhood network and the cost of transformation of topological relationships, as presented in the TDD model.

Definition: Given user requirements (Q) and provided VGI description (M),

$OB_q = (ob_q^1, ob_q^2, \dots, ob_q^i)$ where $ob_q^1, ob_q^2, \dots, ob_q^i$: the spatial objects adduced in user requirements.

$OB_M = (ob_M^1, ob_M^2, \dots, ob_M^j)$ where $ob_M^1, ob_M^2, \dots, ob_M^j$: the spatial objects adduced in the provided VGI description.

Let $T = t_{ij}$; denote the matrix used to measure the topological distance between spatial objects of user requirements (Q) and spatial objects provided VGI description (M), $1 \leq i \leq n, 1 \leq j \leq m$, and t_{ij} is the topological distance between the object i in Q and the object j in M using a conceptual neighborhood graph.

The topological distance between Q and M, denoted by $D_{topo}(Q, M)$, is computed as follows:

$$D_{topo}(Q, M) = \sum_{i=1}^n \sum_{j=1}^m t_{ij}; \quad 1 \leq i \leq n \text{ and } 1 \leq j \leq m \tag{4}$$

c: METRIC DISTANCE BETWEEN USER REQUIREMENTS AND PROVIDED VGI

To compute the metric distance between user requirements and VGI description, we propose to compute the thematic distance between spatial objects presented in each of them.

For this objective, we propose to use the traditional model [1], [32] composed of four possible situations for distances (equal, near, medium, and far). In this model, the cost of the transition from one situation to another is equal to one. Figure 4 shows various possible situations and the cost of transition from one situation to another.

Definition: Given user requirements (Q) and provided VGI description (M),

$OB_q = (ob_q^1, ob_q^2, \dots, ob_q^i)$ where $ob_q^1, ob_q^2, \dots, ob_q^i$: the spatial objects adduced in user requirement's regarding VGI.

$OB_M = (ob_M^1, ob_M^2, \dots, ob_M^j)$ where $ob_M^1, ob_M^2, \dots, ob_M^j$: the spatial objects adduced in the Provided VGI description.

Let $M = mt_{ij}$, denote the matrix used to measure the metric distance between spatial objects of user requirements (Q) and spatial objects provided in the VGI description (M), $1 \leq i \leq n, 1 \leq j \leq m$, and mt_{ij} : the metric distance between object i in Q and object j in M using a metric distance model.

The metric distance between Q and M, denoted by $D_{met}(Q, M)$, is computed as follows:

$$D_{met}(Q, M) = \sum_{i=1}^n \sum_{j=1}^m mt_{ij}; \quad 1 \leq i \leq n \text{ and } 1 \leq j \leq m \tag{5}$$

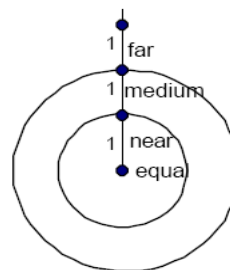


FIGURE 4. The metric distance network [32].

d: SPATIAL DISTANCE BETWEEN USER REQUIREMENTS AND PROVIDED VGI

The spatial distance between spatial objects integrates topological, orientation, and metric distances [1], [8], [32]. Thus, the spatial distance between the user requirements and the provided VGI description is computed as follows:

Definition: Given user requirements (Q) and the provided VGI description (M), the spatial distance between Q and M is computed as follows:

$$D_{spatial}(Q, M) = D_{topo}(Q, M) + D_{dir}(Q, M) + D_{met}(Q, M) \tag{6}$$

where:

- $D_{topo}(Q, M)$: the topological distance between Q and M ,
- $D_{dir}(Q, M)$: the orientation distance between Q and M .
- $D_{met}(Q, M)$: the metric distance between Q and M .

e: SPATIO-SEMANTIC SIMILARITY MEASURE BETWEEN USER REQUIREMENTS AND PROVIDED VGI DESCRIPTION

The principal objective of section B is to propose a similarity measure between user requirements and provided VGI data in order to evaluate the fitness-for-use of VGI. The similarity measure that we propose integrates semantic similarity as well as spatial similarity between user requirements and provided VGI description. Spatial similarity considers topological similarity, spatial similarity, and orientation similarity between spatial objects in user requirements and provided VGI.

The spatio-semantic distance between user requirements and provided VGI data reflects the degree of spatial similarity and semantic relatedness between them [1]. Therefore, the spatio-semantic distance between user requirements and VGI data is derived from the spatial and semantic distance as follows:

Definition: Given user requirements (Q) and a VGI description (M),

$D_{Sem}(Q, m)$ and $D_{Spatial}(Q, M)$ denote respectively the semantic and spatial distances between Q and M . The spatio-semantic distance and similarity between Q and M , denoted $D_{SPS}(Q, M)$ and $SM_{SPS}(Q, M)$, respectively, are computed as follows:

$$D_{SPS}(Q, M) = (D_{Sem}(Q, M) + D_{Spatial}(Q, M))/2 \quad (7)$$

$$SM_{SPS}(Q, M) = 1/(1 + D_{SPS}(Q, M)) \quad (8)$$

C. INDICATORS FOR VGI QUALITY ASSESSMENT

We propose a restricted set of indicators and quantitative approach for evaluating the fitness-for-use of VGI data. These indicators are VGI completeness and credibility.

It should be noted that these indicators do not aim to be complete or precise but rather indicate whether potential risks may occur. A small number of quality indicators provide synthetic key information about data fitness-for-use [14]. They allow end-users to have a global idea, and then they may or may not dig into details [14]. Typical decision-making processes use a small number of indicators [14].

Completeness is a widely used data quality indicator [29]. Measuring the completeness of VGI data is key for VGI to fit a particular usage. Credibility is an important issue in VGI as contributors, whether willfully or not, may provide false or misleading data. This may undermine the project in which VGI data is used and may cause a further lack of trust from end-users about the quality and usability of VGI [35].

For each indicator, the quality is assessed within the interval $[0, 1]$. A value of 1 indicates perfect fitness-for-use, whereas a value of 0 indicates very poor fitness-for-use.

1) COMPLETENESS OF ACCURACY

This indicator shows the number of VGI data elements with regard to the required elements. An element, whether data or metadata, may be a feature, attribute, or relationship with different elements. We represent a data element in the form {element nature, element type, element value}, where element nature indicates whether the element is related to data or metadata, and element type indicates whether the data is a feature, an attribute, or a relationship. On the other hand, the element value is the representation of the element. For example, the value of the feature *pothole* is a collection of pixels within the image. An example about a metadata element related to a data element *pothole* is the device used to take a picture of the pothole (e.g., a mobile phone).

We recognize thematic, spatial, and temporal completeness. The evaluation of completeness Cd is calculated as follows:

$$Cd = \begin{cases} 1; & \text{if } Nme \geq NmrandNse \geq NsrandNte \geq Ntr \\ (wm * Nme/Nmr + ws * Nse/Nsr + wt * Nte/Ntr)/3; & \text{Otherwise} \end{cases}$$

where Nme (Nse and Nte) is the number of thematic (spatial and temporal) data elements provided by a contributor. Nmr (Nsr , Ntr) is the number of thematic (spatial and temporal) data elements required for a particular use. Wm (ws , wt) is a predefined weight for thematic (spatial and temporal) data elements. This weight indicates the importance of each type in a particular use. If the number of available data elements is equal to or greater than the required number of elements, the VGI data is complete, and its value is set to 1. Otherwise, the ratio of existing elements to the required elements shows the degree of VGI completeness.

The number of elements (Nmr , Nsr , and Ntr) as well as the weights (wm , ws , and wt) can be predefined by VGI process analysts with end-users. In the example presented in Section II (Figure 1), the emergency application's users need, besides the data elements available (geospatial features), another element specifying the type of the feature shown in the image (e.g., is it a bulge? Is it a poorly repaired hole on the road?). Thus, if we suppose that the weight of spatial completeness is 1, then $C = 3/4$.

2) VGI CREDIBILITY

This indicator describes the degree of faith that we have regarding the data provided with regard to a particular usage. As such, an increase in the amount of negative feedback of the VGI data (negative report, or modification) lessens its credibility. Therefore, the greater is the number of reports and modifications, the lower is the VGI credibility. We evaluate VGI credibility using the following function:

$$Td = 1 - \left(\frac{N_{report} + N_{modify}}{Nf} \right)$$

where N_{report} (N_{modify}) is the number of reports (modifications) related to the VGI data, and Nf is the total feedback related to the same VGI data.

The overall quality Q_{vgi} is calculated as follows:

$$Q = \frac{\sum Wi * Qi}{Nq - (\sum(1 - Wiz))}$$

$Wiz = 1$ if $Wi < 0$, $Wiz = 0$ otherwise

where Qi is the i th quality indicator of the VGI data. where Wi is the i th predefined weight of the i th indicator. where Nq is the total number of indicators. Wiz is a binary number equal to zero if weight Wi is null. In this case we subtract this indicator from the total number of VGI indicators.

In the next section, we propose an algorithm to help users deal with the risk of misusing VGI in a systematic manner.

D. ALGORITHM TO MANAGE THE RISK RELATED TO VGI QUALITY

This section proposes an algorithm to help end-users make appropriate decisions about the risk of data usability in the context of VGI. The algorithm is based on the well-established possibilities of responding to risks (i.e., reducing, absorbing, or transferring risks).

Figure 5 illustrates the proposed algorithm. Based on the previously identified indicators of the fitness-for-use of VGI data, we start by identifying and assessing the risks of data misuse (e.g., incomplete VGI fitness-for-use or lack of credibility). Based on risk assessment, end-users will be advised to respond to the risks of data misuse in many ways:

- Suspending the VGI process if it presents a high risk of harmful consequences. In this case, end-users are invited not to use the VGI data.
- Carefully using VGI data if it has satisfactory fitness-for-use.
- Encouraging the use of VGI, if VGI data has good fitness-for-use.

The proposed indicators play a key role in the proposed algorithm. They allow to identify the risks of data misinterpretation and to draw conclusions about them. Specifically, these indicators have three principal aims:

First, to help end-users identify risks related to VGI data quality. In fact, in the proposed algorithm, while a good quality of VGI data indicates that it is less likely to have a risk of data usability, poor quality indicates a higher risk.

Second, to help end-users make appropriate decisions (e.g., to consider or not VGI data).

To help end-users intuitively understand the degree of risk related to VGI data, we provide a set of warnings based on standard danger symbols proposed by ANSI [47]. These symbols are enriched by well-known mapping markers to assist end-users of VGI data. The symbols show the degree of risk of VGI data misuse, and thereby stimulate appropriate responses to such a risk. For example, if the warning is 'Danger', it would be better to suspend the VGI data use process, which may lead to considerable risk. Table 1 shows an example of how warnings can be predefined with regard to the quantitative values of the fitness-for-use of VGI data.

In the following section, we present a prototype called *FitVGI*, which we developed to instantiate our approach.

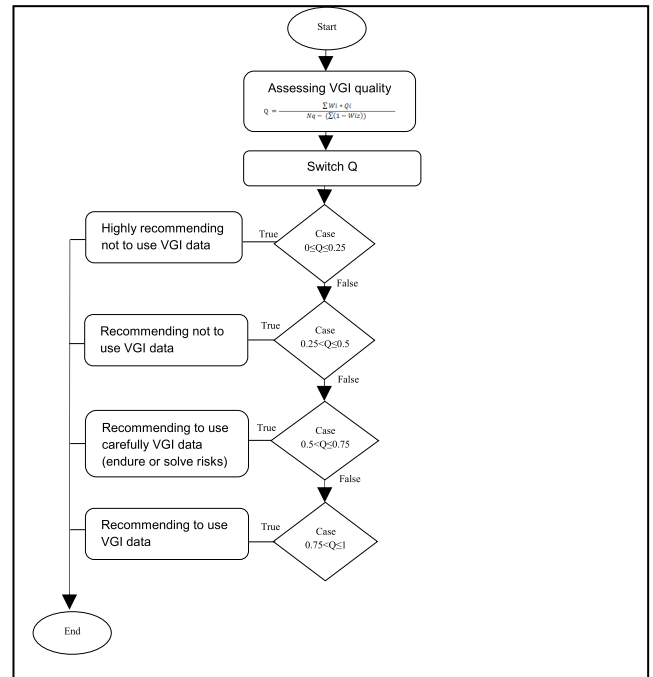


FIGURE 5. Algorithm flowchart for responding to risks of misusing VGI data.

TABLE 1. A definition of the symbols according to the VGI quality value.

Quality value	Representation of quality indicators	Significance of the symbol
$0 \leq Q \leq 0.25$		There is a high risk if end-users decide to use VGI data.
$0.25 < Q \leq 0.5$		End-users are warned about the existence of potential risks.
$0.5 < Q \leq 0.75$		End-users are invited to pay attention if they decide to use VGI data.
$0.75 < Q \leq 1$		Information will be shown to end-users.

Based on this prototype, we continue with the example introduced in section III. A. (see Figure 1) to show how the proposed approach can be used to address the risks related to VGI data usage. We then implement our *FitVGI* prototype in a case study of a road maintenance project.

IV. THE *FitVGI* PROTOTYPE

To implement our approach for managing the risk related to the VGI data quality, we developed a prototype called *FitVGI*. The *FitVGI* prototype was implemented in Python environment. We used some packages such as GeoPandas, Plotly, NumPy, SciPy, and Scikit-learn to develop our prototype.

The prototype implements three main functionalities: data collection from contributors, risk-based evaluation of VGI fitness-for-use, and visualization of fitness-for-use indicators.

To better illustrate our approach, we continue with the example of application presented in Figure 1. A Contributor takes a georeferenced image by a mobile phone and sends it to the *FitVGI* reception interface. Through this interface, the contributor can add some textual metadata (e.g., adding title, uploading a text file), as shown in Figure 6.

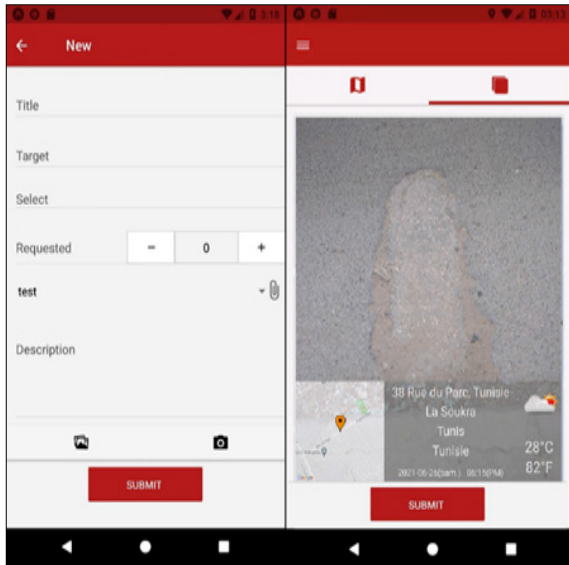


FIGURE 6. The prototype reception interface.

The VGI data has two features: a road and pothole (or a bulge on the road). In addition to these two elements, end-users need to know if there are buildings close to the pothole. Consequently, $Cd = 2/3$.

In addition, this data has been sent recently, and two other contributors have proposed some modifications (adding more information about this data, i.e., metadata). In addition, this data was sent three times through the prototype. Thus, $Td = (1 - (2/3)) = 1/3$.

Consequently, if we assume that the weight of each indicator is 1, the overall quality $Q = ((2/3) + (1/3)) / 2 = 0.5$.

Based on Table 1, the orange symbol is shown to end-users who should pay attention if they decide to use VGI data (see Figure 7).

V. EXPERIMENTS AND RESULTS

In order to show the utility of our approach, we implemented the *FitVGI* prototype in a case study of a road maintenance project in *Tunis* city (Tunisia). This ongoing project, managed by the Urban Planning Agency, aims at keeping the road structure as normal as possible and practicable. The project uses VGI data to locate and manage potholes and other road obstacles.

A. METHODOLOGY FOR APPROACH EVALUATION

We used a human evaluation technique to evaluate our approach. We believe that human experts in road maintenance are likely to be in a good position to evaluate the outcome

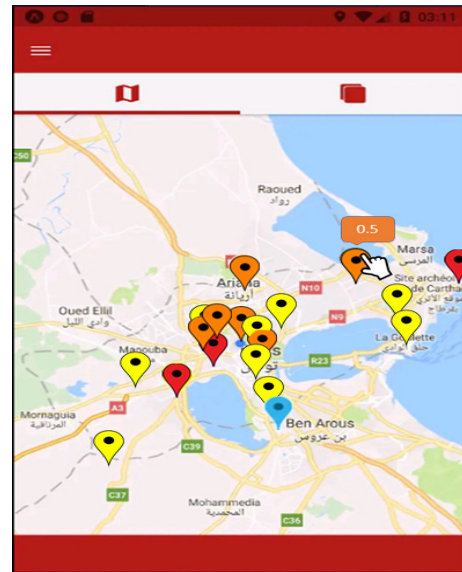


FIGURE 7. Visualization of quality indicators.

of our approach. Road maintenance experts are qualified to indicate if the assessment of VGI data quality provided by our systems is really useful to give good recommendation to end-users according to their needs and preferences.

Our evaluation technique consists of computing the correlation between (1) the quality assessment provided by our *FitVGI* prototype and (2) the quality assessment provided by road maintenance experts. To calculate this correlation, we adopted the widely-used Spearman correlation method [45]. This method is used to measure the strength and direction of a relationship between two ranked variables using a monotonic function. Many works have used the Spearman's rank correlation coefficient to explore and evaluate quality indicators [10], [23].

B. DATA PROCESSING

We used a set of georeferenced images taken by mobile phones, and received by the *FitVGI* prototype. The images, provided by VGI contributors, present potholes, bad roads and other road obstacles located in different areas of *Tunis*. Figure 8 shows some of the received images.

Using our *FitVGI* prototype, we evaluate the quality of each of those images. The selection of the images to be used in the evaluation process is made based on the methodology of Miller and Charles [34] in building the sample in the correlation evaluation process. In fact, the evaluation process is made with reference to 30 images: 10 images that have a high level of quality ($0.75 < Q \leq 1$), 10 images with a medium level of quality ($0.5 < Q \leq 0.75$) and 10 images with a lower level of quality ($Q < 0.5$); quality is assessed by our *FitVGI* prototype.

C. EVALUATION PROCESS AND RESULT ANALYSIS

We asked 15 road maintenance experts to visit the road where the VGI images were taken, and then to assign to each image

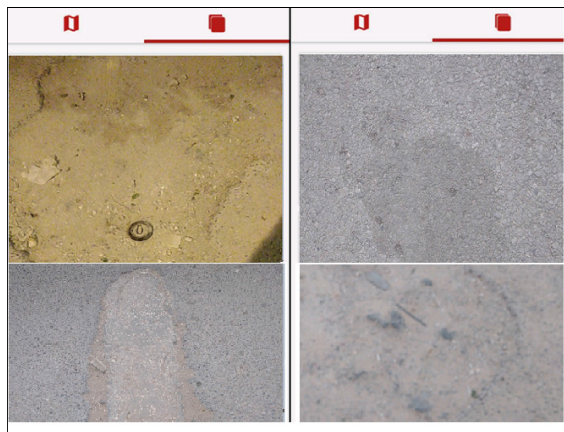


FIGURE 8. Georeferenced images taken by mobile phones.

a quality value that reflects its appropriateness to the requirements of the road maintenance project (i.e., how much the image is appropriate to be used for road maintenance project). The value of the quality indicator assigned to each image should be between 0 and 1.

After collecting quality values, we calculate Spearman’s coefficient that computes the correlation between the quality values of the *FitVGI* prototype and those provided by experts. We obtained 0.773 as a value of the Spearman’s correlation coefficient (see Figure 9).

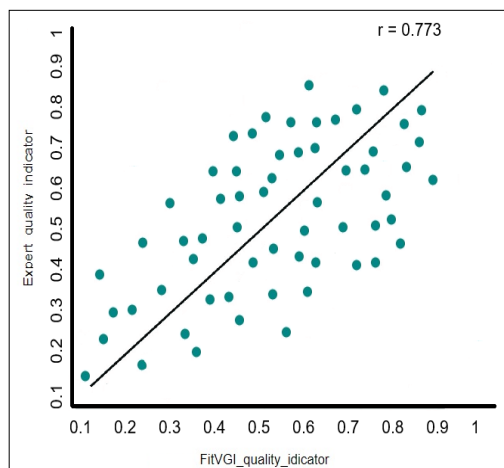


FIGURE 9. Correlation between quality assessment of *FitVGI* and quality assessment provided by road maintenance experts.

A value of the correlation coefficient equal to 0.773 presents a positive correlation between quality assessment provided by our *FitVGI* prototype and quality assessment provided by road maintenance experts. The result proves the ability of our approach to make useful quality evaluation of VGI data. In addition, and based on this evaluation, our approach proposes useful recommendations about using -or not using- VGI data.

VI. COMPARISON WITH RELATED STATE-OF-THE-ART METHODS

Previous works that proposed approaches and algorithms to enhance the quality of VGI may be categorized into three main groups: works proposing quality guidelines, works evaluating quality of VGI based on reference data, and works assessing VGI quality indicators.

The first group of works propose guidelines to enhance VGI quality (e.g., [28]). However, guidelines are often general and do not necessarily target a specific usage of data. On the other hand, our approach assesses the quality of VGI data in a particular usage (i.e., fitness-for-use).

For the second group, being based on reference data (basically authoritative data) may lead to limitations including the availability and the quality of reference data, and even if it is available, accessing reference data is often restricted by constraints or it may be expensive. In addition, these approaches do not take into account the intended use of VGI data. That is, even if VGI data is reflective to the reference data, it may be of poor quality in a particular usage. On the contrary, our approach does not refer to any authoritative data. It refers to the particular requirements of each end-user.

The third group includes algorithms that evaluate VGI quality indicators (e.g., [4], [9], [12], [27], [38], and [44]). These works focused on the internal quality of VGI data, and did not focus on the context of a particular usage. This may cause a risk of inappropriate usage of VGI data. In fact, VGI data may be appropriate for a specific usage, and less appropriate for another usage.

On the other hand, our approach takes into account the intended usage of VGI data. It evaluates the quality of VGI data for each particular usage.

Table 2 summarizes the comparison of our work with three groups of state-of-the-art works.

TABLE 2. A comparison of our approach with state-of-the-art works.

Related works	Main idea	Comparison with our approach
Works proposing quality guidelines (e.g., [28])	Proposing guidelines to reach a certain level of compatibility between VGI data and reality.	Guidelines are often general and do not necessarily target a specific use. Our approach targets a specific usage of VGI.
Works evaluating quality of VGI based on reference data (e.g., [30] and [19])	Comparing geospatial characteristics of VGI data with a reference data. VGI quality is evaluated to its closeness to the reference data.	It may not be possible to access the reference data. Our approach does not refer to any authoritative data.
Works assessing VGI quality indicators (e.g., [38] [44],[19] [27], [4] and [12]).	Assessing the quality of VGI data based on a set of indicators.	These works did not focus on the context of a particular use of VGI data. In our approach, we take into account a particular usage of VGI data.

In addition, based on risk management, our approach provides a recommendation for each usage of VGI data. As such, end-users can be advised either to use VGI data which has a good quality, not to use VGI data that has poor quality, or to carefully use a VGI data with not good enough quality.

VII. CONCLUSION

Uncertainty related to VGI data has the potential to expose end-users to the risk misusing such data. This risk may be due to the fact that VGI data may be of poor quality with respect to end-user requirements (i.e., fitness-for-use). In this work, we proposed a method to evaluate the fitness-for-use of VGI data based on a set of indicators and a spatio-semantic similarity measure between user requirements and the provided VGI. We also proposed a risk-based algorithm that consists of a systematic process to make end-users aware of potential risks, and help them make appropriate decisions about using or not using VGI data.

We developed a prototype called *FitVGI* to implement our proposed approach to enhance the fitness-for-use of VGI data. The prototype presents the quality indicators to the end-users in an intuitive way to help them make appropriate decisions about the risks related to VGI data.

Based on VGI data sent by contributors and a human evaluation approach, we demonstrated how our proposed algorithm, as well as the evaluation of the proposed fitness-for-use indicators, can help end-users to make appropriate decisions about using or not using VGI data, or to be careful when using it. We equate our approach implementation to a recommendation system about the usefulness of the provided VGI data with regards to particular requirements.

We should remind that the defined indicators do not aim at being complete, or completely eliminating the risk related to VGI data, but rather making users aware of such a risk.

Further work is required to enhance the approach and the *FitVGI* prototype by allowing interaction not only with end-users, but also with VGI contributors. Therefore, contributors may rectify or withdraw data if it does not fit a particular usage.

REFERENCES

- [1] S. Aissi, M. S. Gouider, T. Sboui, and L. B. Said, "A spatial data warehouse recommendation approach: Conceptual framework and experimental evaluation," *Hum.-Centric Comput. Inf. Sci.*, vol. 5, no. 1, pp. 1–18, Dec. 2015.
- [2] S. Aissi and T. Sboui, "Towards evaluating geospatial metadata quality in the context of VGI," *Proc. Comput. Sci.*, vol. 109, pp. 686–691, Jan. 2017.
- [3] M. Al-Bakri and D. Fairbairn, "Assessing the accuracy of 'crowdsourced' data and its integration with official spatial data sets," in *Proc. 9th Int. Symp. Spatial Accuracy Assessment Natural Resour. Environ. Sci.*, Leicester, U.K., Jul. 2010, pp. 317–320.
- [4] A. L. Ali and F. Schmid, "Data quality assurance for volunteered geographic information," in *Geographic Information Science*, M. Duckham, E. Pebesma, K. Stewart, and A. U. Frank, Eds. Berlin, Germany: Springer, 2014, pp. 41–126.
- [5] V. Antoniou and A. Skopeliti, "Measures and indicators of VGI quality: An overview," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. II-3/W5, pp. 345–351, 2015, doi: 10.5194/isprsannals-II-3-W5-345-2015.
- [6] J. Arsanjani, P. Mooney, A. Zipf, and A. Schauss, "Quality assessment of the contributed land use information from OpenStreetMap versus authoritative datasets," in *OpenStreetMap in GIScience (Lecture Notes in Geoinformation and Cartography)*. Cham, Switzerland: Springer, 2015, pp. 37–58.
- [7] G. Brown, M. Kelly, and D. Whittall, "Which 'public'? Sampling effects in public participation GIS (PPGIS) and volunteered geographic information (VGI) systems for public lands management," *J. Environ. Planning Manag.*, vol. 57, pp. 190–214, Feb. 2013.
- [8] T. Bruns and M. Egenhofer, "Similarity of spatial scenes," in *Proc. 7th Int. Symp. Spatial Data Handling*, Delft, The Netherlands, 1996, pp. 31–42.
- [9] A. Chehreghan and R. A. Abbaspour, "An evaluation of data completeness of VGI through geometric similarity assessment," *Int. J. Image Data Fusion*, vol. 9, no. 4, pp. 319–337, Oct. 2018.
- [10] H. Chen, D. Hailey, N. Wang, and P. Yu, "A review of data quality assessment methods for public health information systems," *Int. J. Environ. Res. Public Health*, vol. 11, no. 5, pp. 5170–5207, 2014.
- [11] D. Coleman, "Potential contributions and challenges of VGI for conventional topographic base-mapping programs," in *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*. Dordrecht, The Netherlands: Springer, 2013, pp. 245–264.
- [12] P. Corcoran, P. Mooney, and A. Winstanley, "Topological consistent generalization of OpenStreetMap," in *Proc. GIS Res. U.K. 18th Annu. Conf. (GISRUK)*, London, U.K., Apr. 2010, pp. 14–16.
- [13] B. De Longueville, N. Ostländer, and C. Keskitalo, "Addressing vagueness in volunteered geographic information (VGI)—A case study," *Int. J. Spatial Data Infrastruct. Res.*, vol. 5, 2010.
- [14] R. Devillers, Y. Bédard, R. Jeansoulin, and B. Moulin, "Towards spatial data quality information analysis tools for experts assessing the fitness-for-use of spatial data," *Int. J. Geograph. Inf. Sci.*, vol. 21, no. 3, pp. 261–282, 2007.
- [15] S. Elwood, M. F. Goodchild, and D. Z. Sui, "Researching volunteered geographic information: Spatial data, geographic research, and new social practice," *Ann. Assoc. Amer. Geographers*, vol. 102, no. 3, pp. 571–590, May 2012.
- [16] S. Elwood, M. F. Goodchild, and D. Sui, "Prospects for VGI research and the emerging fourth paradigm," in *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*. Berlin, Germany: Springer, 2013, pp. 361–375.
- [17] C. C. Fonte, V. Antoniou, L. Bastin, J. Estima, J. J. Arsanjani, J. C. L. Bayas, L. M. See, and R. Vateva, "Assessing VGI data quality," in *Mapping and the Citizen Sensor*, G. M. Foody, L. See, S. Fritz, C. C. Fonte, P. Mooney, A.-M. Olteanu-Raimond, and V. Antoniou, Eds. London, U.K.: Ubiquity Press, 2017, pp. 137–163.
- [18] M. Forghani and M. Delavar, "A quality study of the OpenStreetMap dataset for Tehran," *ISPRS Int. J. Geo-Inf.*, vol. 3, no. 2, pp. 750–763, May 2014.
- [19] J.-F. Girres and G. Touya, "Quality assessment of the French OpenStreetMap dataset," *Trans. GIS*, vol. 14, no. 4, pp. 435–459, Aug. 2010.
- [20] M. F. Goodchild, "Citizens as sensors: The world of volunteered geography," *GeoJournal*, vol. 69, no. 4, pp. 211–221, Aug. 2007.
- [21] M. F. Goodchild and L. Li, "Assuring the quality of volunteered geographic information," *Spatial Statist.*, vol. 1, pp. 110–120, May 2012.
- [22] S. C. Guptill and J. L. Morrison, *Elements of Spatial Data Quality*. Oxford, U.K.: Elsevier, 1995, pp. 31–58.
- [23] J. Hagenauer and M. Helbich, "Mining urban land-use patterns from volunteered geographic information by means of genetic algorithms and artificial neural networks," *Int. J. Geograph. Inf. Sci.*, vol. 26, no. 6, pp. 963–982, 2012.
- [24] M. Haklay, "Citizen science and volunteered geographic information: Overview and typology of participation," in *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, D. Z. Sui, S. Elwood, and M. F. Goodchild, Eds. Berlin, Germany: Springer, 2013, pp. 105–122.
- [25] G. B. Hall, R. Chipeniuk, R. D. Feick, M. G. Leahy, and V. Deparday, "Community-based production of geographic information using open source software and web 2.0," *Int. J. Geograph. Inf. Sci.*, vol. 24, no. 5, pp. 761–781, 2010.
- [26] P. Hashemi and R. A. Abbaspour, "Assessment of logical consistency in OpenStreetMap based on the spatial similarity concept," in *OpenStreetMap in GIScience*, J. J. Arsanjani, A. Zipf, P. Mooney, and M. Helbich, Eds. Cham, Switzerland: Springer, 2015, pp. 19–36.

- [27] K.-C. Hung, M. Kalantari, and A. Rajabifard, "Methods for assessing the credibility of volunteered geographic information in flood response: A case study in Brisbane, Australia," *Appl. Geography*, vol. 68, pp. 37–47, Mar. 2016.
- [28] *Geographic Information-Data Quality*, International Organization of Standardization (ISO), Geneva, Switzerland, ISO Standard 19157:2013, 2013.
- [29] S. Jackson, W. Mullen, P. Agouris, A. Crooks, A. Croitoru, and A. Stefanidis, "Assessing completeness and spatial error of features in volunteered geographic information," *ISPRS Int. J. Geo-Inf.*, vol. 2, no. 2, pp. 507–530, Jun. 2013.
- [30] M. Kalantari and V. La, "Assessing OpenStreetMap as an open property map," in *OpenStreetMap in GIScience SE* (Lecture Notes in Geoinformation and Cartography). Basel, Switzerland: Springer, 2015, pp. 255–272.
- [31] K. Kemp, *Encyclopedia of Geographic Information Science*. Newbury Park, CA, USA: Sage, 2008.
- [32] B. Li and F. T. Fonseca, "TDD: A comprehensive model for qualitative spatial similarity assessment," *Spatial Cognition Comput.*, vol. 6, no. 1, pp. 31–62, 2006.
- [33] A. Mashhadi, G. Quattrone, and L. Capra, "The impact of society on volunteered geographical information: The case of OpenStreetMap," in *OpenStreetMap in GIScience: Experiences, Research, and Applications*. Cham, Switzerland: Springer, 2015, pp. 125–141.
- [34] G. A. Miller and W. G. Charles, "Contextual correlates of semantic similarity," *Lang. Cogn. Processes*, vol. 6, no. 1, pp. 1–28, 1991.
- [35] P. Mooney, A. M. Olteanu-Raimond, G. Touya, N. Juul, S. Alvanides, and N. Kerle, "Considerations of privacy, ethics and legal issues in volunteered geographic information," in *Mapping and the Citizen Sensor*. London, U.K.: Ubiquity Press, 2017, pp. 119–135.
- [36] R. B. Moreira, L. C. Degrossi, and J. P. De Albuquerque, "An experimental evaluation of a crowdsourcing-based approach for flood risk management," in *Proc. 12th Workshop Exp. Softw. Eng. (ESELAW)*, Lima, Peru, Apr. 2015, pp. 1–11.
- [37] K. Morgan, "Development of a preliminary framework for informing the risk analysis and risk management of nanoparticles," *Risk Anal.*, vol. 25, no. 6, pp. 1621–1635, Dec. 2005.
- [38] R. G. Nejad, R. A. Abbaspour, and A. Chehreghan, "Spatiotemporal VGI contributor reputation system based on implicit evaluation relations," *Geocarto Int.*, vol. 37, pp. 1–28, May 2022.
- [39] J. Posetti, "The Twitterisation of ABC's emergency and disaster communication," *Austral. J. Emergency Manag.*, vol. 27, no. 1, pp. 34–39, 2012.
- [40] R. Rada, H. Mili, E. Bicknell, and M. Blettner, "Development and application of a metric on semantic nets," *IEEE Trans. Syst., Man, Cybern.*, vol. 19, no. 1, pp. 17–30, Jan./Feb. 1989.
- [41] K. Rezgüi, H. Mhiri, and K. Ghédira, "Theoretical formulas of semantic measure: A survey," *J. Emerg. Technol. Web Intell.*, vol. 5, no. 4, pp. 333–342, Nov. 2013.
- [42] S. Schade, L. Díaz, F. Ostermann, L. Spinsanti, G. Luraschi, S. Cox, M. Nuñez, and B. De Longueville, "Citizen-based sensing of crisis events: Sensor web enablement for volunteered geographic information," *Appl. Geomatics*, vol. 5, no. 1, pp. 3–18, Mar. 2013.
- [43] H. Senaratne, A. Mobasheri, A. L. Ali, C. Capineri, and M. Haklay, "A review of volunteered geographic information quality assessment methods," *Int. J. Geograph. Inf. Sci.*, vol. 31, pp. 139–167, May 2017.
- [44] J. Severinsen, M. de Roiste, F. Reitsma, and E. Hartato, "VGTrust: Measuring trust for volunteered geographic information," *Int. J. Geograph. Inf. Sci.*, vol. 33, no. 8, pp. 1683–1701, Aug. 2019.
- [45] C. Spearman, "The proof and measurement of association between two things," *Amer. J. Psychol.*, vol. 15, no. 1, pp. 72–101, Jan. 1904.
- [46] D. Zielstra and A. Zipf, "A comparative study of proprietary geodata and volunteered geographic information for Germany," in *Proc. 13th AGILE Int. Conf. Geographic Inf. Sci.*, Guimarães, Portugal, May 2010, pp. 10–14.
- [47] ANSI, *American National Standard for Product Safety Signs and Labels: Z535.4*, National Electrical Manufacturers Association, New York, NY, USA, 1991.

TAREK SBOUI received the Ph.D. degree in geomatic sciences from Laval University, Quebec City, QC, Canada, in 2010.

He was a Scientific Researcher at the Department of Geomatic Sciences, Laval University, from July 2010 to February 2011. He was a Scientific Researcher at the National Center for Scientific Research (NCSR), Paris, from February 2011 to September 2013. He has been an Assistant Professor at the Faculty of Science, University of Tunis El Manar, since January 2014. He has been teaching geomatic sciences. He is currently a member of the Research Laboratory "Integrated Management of Natural Resources: Remote Sensing, Spatial Analysis and Modeling (GREEN-TEAM)" at the National Agronomic Institute of Tunisia (INAT). He is the author of 15 chapters and peer-reviewed papers in various conferences, international journals, and books. His research interests include interoperability between geographic information systems, business intelligence in the geospatial context, spatial cognition, and spatial data quality. He has received several awards and contracts with industrial and government research centers.

SAIDA AISSI received the Ph.D. degree in computer sciences from the Higher Institute of Management, Tunis, Tunisia, in 2016. She has been an Assistant in computer science, since October 2007. She currently teaches at Time University, Tunisia. She is also a member of the SMART Laboratory, Higher Institute of Management of Tunis. Her research works is in the domain of data warehouses, geographic information systems, VGI data, and also in the context of SOLAP systems. She is the author of various papers in international journals and conferences dealing with SOLAP, OLAP, GIS personalization, and recommendation systems. She is the author of several peer-reviewed papers in various international conferences and journals.

• • •