International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:11, No:5, 2017

Ontology for a Voice Transcription of OpenStreetMap Data: The Case of Space Apprehension by Visually Impaired Persons

Said Boularouk, Didier Josselin, Eitan Altman

***Abstract*—In this paper, we present a vocal ontology of OpenStreetMap data for the apprehension of space by visually impaired people. Indeed, the platform based on produsage gives a freedom to data producers to choose the descriptors of geocoded locations. Unfortunately, this freedom, called also *folksonomy* leads to complicate subsequent searches of data. We try to solve this issue in a simple but usable method to extract data from OSM databases in order to send them to visually impaired people using Text To Speech technology. We focus on how to help people suffering from visual disability to plan their itinerary, to comprehend a map by querying computer and getting information about surrounding environment in a mono-modal human-computer dialogue.

***Keywords*—Ontology, OpenStreetMap, visually impaired people, TTS, taxonomy.

I. Introduction

THE emergence of new participatory on-line activity to create and access information, to share knowledge and content production, strongly links information production and use. Crowdsourcing was defined by Enrique Estells-Arolas and Fernando Gonzlez Ladrn-de-Guevara from university of Valencia as follows:

*”Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task; of variable complexity and modularity, and; in which the crowd should participate, bringing their work, money, knowledge **[and/or]** experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcing will obtain and use to their advantage that which the user has brought to the venture, whose form will depend on the type of activity undertaken”*.

The OSM and Wikipedia platforms are the most known examples of participatory models of information production and usage. This new model allows to enrich existing content with the goal of future improvement. It is a kind of hybridization between production and use, *coined by Australian media scholar Axel Bruns [1] as ”produsage”* because producers are themselves users [2].

Produsage is based on collaboration and consensus. The rules are generally applied by the community rather than individuals. Producers are driven by constructive motivations (altruism, intellectual stimulation, personal expression, social recognition, etc.) or sometimes by harmful motivations (mischief, criminal intent, etc.) [3], [1]. This new concept of produsage, generally under a free licence of use and modification, finds a balance between license with copyright and those without. This model has advantages and disadvantages. For visual impaired application, OpenStreetMap may be a good opportunity since using and sharing data is free of charge and data are very rich. On another hand, logistic products based on new technologies for visually impaired are very expensive (guide-dogs, GPS or specific object detection tool, electronic cane). The use of OpenStreetMap as a free crowdsourcing may lead to less expensive connected objects since data are free. In addition, spatial data got from produsage are rich and varied since Producers describe their places of life with an accurate granularity. In this context we propose an ontology able to extract useful data and build an efficient vocal application able to help visually impaired for (mind) map representation.

Our goal is to build a taxonomy to exploit disambiguated geographic information by overcoming disadvantages of openStreetMap as a crowdsourcing platform. The solution we propose is a taxonomy to respond to a qualitative query and to rebuild spatial information in a simple and memorable way [1]. Those data can be collected by visually impaired persons from their desk or during a real move. Spatial data must be transmitted by speech synthesis in a hierarchical, comprehensive and effective way in a very short time [4].

II. Geographic Informations Produsage

The geographic information OSM is most of the time populated by voluntary data submitted by citizens (professionals, amateurs or interested experts), equipped with a computer and an internet connection. This free information handling *crowdsourcing* induces a large data traffic, which quickly enhances the OSM data. This information is in the frame of ODBL (Open Database License), promoting a free flow of data and allowing users to share, modify and freely use the geographical database, while maintaining the same freedom for all.

Said Boularouk is a Ph.D. candidate in the LIA, University of Avignon, France (e-mail: said.boularouk@univ-avignon.fr).
Didier Josselin is a research director in CNRS, UMR ESPACE 7300, University of Avignon, France (e-mail: didier.josselin@univ-avignon.fr).
Eitan Altman is a research director in the LIA-INRIA, Sophia Antipolis, France (e-mail: eitan.altman@sophia.inria.fr).

Fig. 1 Different descriptors identifying a bus stop

Gare
public_transport: stop_positin
wheelchair:limited
operator: sncf
railway: stop
train: yes
name= avignon centre

Restaurant
amenity: restaurant
old_name: "Les domaines"
cuisine: steak_house
name: "Restaurant La Boucherie Avignon Centre"
url: http://www.la-boucherie.fr/restaurants/avignon-centre/

Marché
amenity: marketplace
name: "Les Halles"

Distributeur
amenity: post_office
stamping_machine: yes
ref:FR:LaPoste: 15854A
change_machine: yes
copy_facility: yes
addr:postcode: 84000
wheelchair: no
operator: "La Poste"
phone: 3631
name: "Avignon Saint-Ruf"
atm: yes

Fast food
old_amenity: fast_food
shop: bakery
addr:housenumber: 25
opening_hours: "Mo-Sa 06:30-19:30"
addr:postcode: 84000
addr:street: "Rue de la Croix"
addr:city: Avignon
old_name: "Mezzo di pasta"
name: "Marie Blachère"

Arrêt de bus
highway: bus_stop
shelter:operator: JCDecaux
tactile_paving: no
shelter:colour: brown
waste_basket: yes
shelter:ref: 96
wheelchair: no
bench: yes
name: Saint-Lazare

Fig. 2 Taxonomy of places

The collection of geographic information is performed at a very fine granularity as contributors tend to describe their place of daily life (zone, district, place of professional or commercial activity). This is an advantage insofar since the information collected can be related to a particular type of population or application, which is the case for visually impaired people. At the same time, the representation that amateur produsers give about environment which they are familiar with, expresses more field experience than general cartographic expertise. However, verification and validation of the data produced is impossible due to a lack of means (time, human resources), looking the importance of the information generated [1].

Ultimately, a number of problems taint the quality of OpenStreetMap (OSM) data and therefore the use of OSM data such as spatial accuracy, semantic accuracy, arbitrary selection of fields to fill for the same types of objects, continuity and comprehensiveness of information, frequent lack of metadata. Whereas crowdsourcing is an opportunity, problems listed before represent a lock to the fundamental issue of how to help blind people in their mobility needs. On the one hand, we can advantageously benefit from collaboration to share free spatial information with visually impaired people. On the other hand, we find ourselves faced to a large geocoded data imbroglio from which it is necessary to extract the pertinent information. In addition, we may have difficulties with data format, leading to syntactic heterogeneity or semantic heterogeneity. That is why we propose a spatial ontology [5] for the voice transcription of OpenStreetMap data for visually impaired people.

Several experiments already exist around the ontology OSM data. One of them is Osmonto, it was developed at the University of Bremen and DFKI Bremen by Mihai Codescu Gregor Horsinka Oliver Kutz, Till Mossakowski and Rafaela Rau [6]. It is an ontology of tags from OpenStreetMap. Its purpose is to facilitate maintenance of all *tags*, that leads to enrich the tag semantics in relation with other ontologies. DO-ROAM is a web application designed to facilitate the spatial and temporal planning activities and trips. The tool can display the sites where certain activities are held at a given time. Data is retrieved from the OSM Tags reorganized into an ontology [7].

III. Ontological Exploration of Metadata

The basic components of OpenStreetMap database are:

- les *nodes* (defining points in space);
- les *ways* (defining linear features and boundaries of the area);
- les *relations* (which are sometimes used to explain how other elements interact).

These components are mainly related to citizen data by *tag* freely chosen from a list, to describe the meaning of the particular identified item. This consists of keywords in free text containing a *Key* and an attribute *value* whose number of characters does not exceed 255.

As said before, in this process, we can speak of *folksonomy* [1] in opposition to *taxonomy*: contributors are not constrained by a predefined terminology, but they can adopt the terms they wish to classify their information. This hardens the problem mentioned above of *tags* quality related to the lack of expressiveness of the metamodel. In fact, the produser has no constraint when describing the content with free text tags.

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:11, No:5, 2017

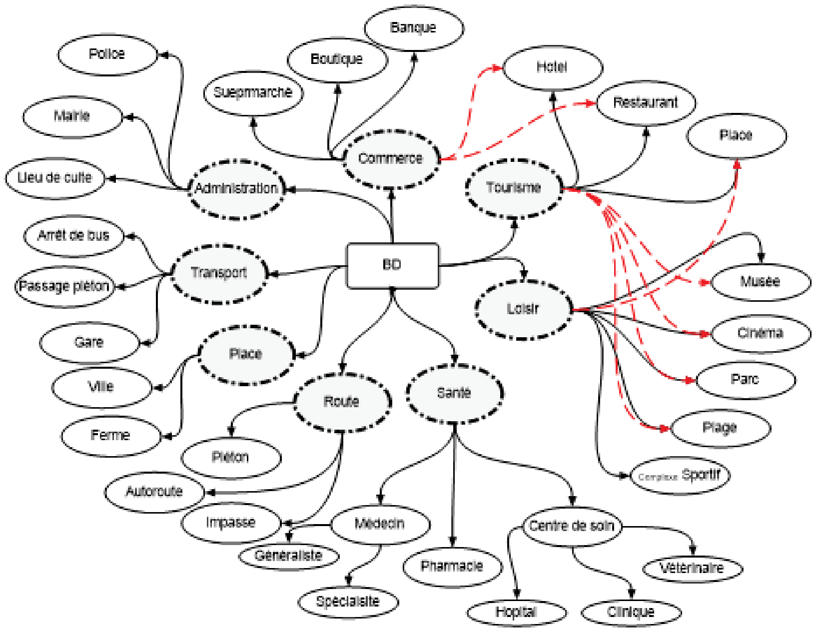

Fig. 3 Tree of places

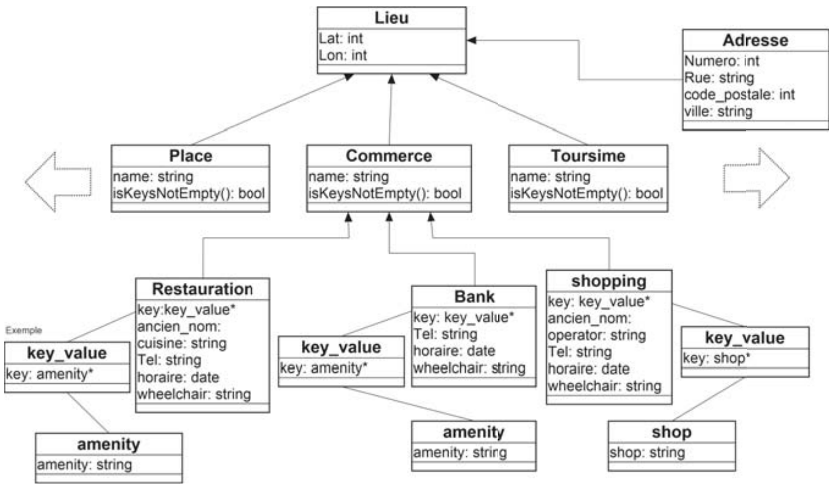

Fig. 4 Diagram UML

Paradoxically, due to the degree of freedom given to data production, users may experience difficulties when searching for information. This difficulty is further complicated in case of communicating geographical data using synthesis voice as proposed in this paper. In addition, the human machine dialogue itself still suffers from many technical difficulties.

We can gather problems resulting from OSM produsage within syntactic heterogeneity and semantic heterogeneity. To resolve syntactic heterogeneity, we parse data to a *Postgresql* database with a spatial extension. We focus on the table of nodes, since we try to send informations about places to blinds people. In this table, we have four columns*(id, latitude, longitude, tags)*. The number of stored lines depends on the size of geographical zones. We have more than 23 millions records for the French *PACA region*. Each column *Tags* represents a vector of tags data $\{K_1, V_1, K_2, V_2, K_3, V_3, ...K_n, V_n\}$. a Key has an odd position, a value has an even position. To solve semantic heterogeneity, we use a simple taxonomy according to the cognitive capacity of the different types and ages of users. However, the OSM linguistic ontology is an English words set. As we have to reduce the semantic distance between users non-English speakers and the Human-Machine interface, we need a multi-lingual translator API.

IV. Taxonomy and Classification of Places

On a map, a place is represented by GPS coordinates (latitude, longitude) and materialized by a point. Each place

Fig. 5 Some restaurants in Avignon

belongs to well-defined categories or classes (transport, trade, tourism, etc). For example, a bus stop can be identified in several ways: shelter arranged for handicapped person, a bench, a label and its time card. From a semantic point of view, a bus stop is a place characterized by the elements shown in the graph in Fig. 1 and defined by corresponding *Tags*. It is noted that the OSM extracted description could be more explicit in expressing the color of the cover, the type of equipment for handicapped person or advertising operator, etc.

Here is the result of an SQL query to find bus stop within 1,000 meters around the University of Avignon. It gives distances from each bus stop and details from tag vector. We put a | between each paired key-value.

1) $869.49m\ shelter : yes|\ highway : bus\_stop|\ bench : yes|\ name : Gerfauts|$

2) $161.67m\ shelter : operator : JCDecaux|\ tactile\_paving : no|\ shelter : colour : brown|\ waste\_basket : yes|\ shelter : ref : 96|\ wheelchair : no|\ highway : bus\_stop|\ bench : yes|\ name : Saint-Lazare|$

3) $553.31m\ shelter : yes|\ highway : bus\_stop|\ name : Carmes|$

The description of a place or an urban component depends on the good will of the contributor. The question is whether such an explicit information is really useful for visually impaired users or not. We need to build a schema able to make a suitable sorting of useful data. The situation is more complex for some places, for instance, if we search for a specialist physician or a restaurant with a special cuisine.

For optimizing the search of places which belong to different categories, we firstly build a taxonomy to organize places in a non-hierarchical tree, since a place could belong to multiple categories (e.g a Restaurant belongs to a Tourism Commerce). This approach allows to lighten the voice recognition system and to soften the redundancy of words recognized by the Speech To Text (STT) module.

To convey the information to blind users exactly as the producer describes his data requires a very precise and detailed architecture. Under conditions of current use, the application may generate confusion and panic in the blind movement because of the complexity and the slowness of the system responses. It is therefore more appropriate to describe the scene in a conventional way, not on the way that produsers fill this information. This brings us to build a rhetorical model of a man-machine dialogue system whose main purpose is a semantic understanding of the request and a rhetoric transmission (i.e. sending information in a constructive but non responsive way). For example, the conventional description of a store is: store name, address, marketed products, if it belongs to a franchise of a national or international company and finally its opening hours. A simple query to our system can provide this information easily and naturally, if it is available in the database. However, a more difficult question to solve is the following: what is merely responding to the blind ”this place is ...”, ”this place is located within such a distance”. Does it work by adding ”this place consists of, is made of and is produced by ...”. In this case, more descriptive subclasses places are required, but this may make vocal synthesis more complex, slower and subsequently decrease the tool usability.

Based on the taxonomy of Fig. 2 we thus propose a simple ontology of three intelligiblelevels (see Fig. 3) for visually impaired users, with possible redundancy of terms between taxonomy branches to adapt to the way producers fill their information into OSM. With this structure, the response of the system to users questions is fast and efficient. Additionally, the UML diagram in Fig. 4 describes the different objects classes associated to the most used OSM keys.

To illustrate the results of a query, a graphic example of the places found is provided in Fig. 5 about the restaurants in a certain direction and at a known distance. The list of restaurants is given to the blind people by a woman synthesis voice. It gives to the user the name of all selected restaurant, their direction and relative location. User can navigate in the places list by saying (*previous/following*)

V. Discussion

The produsage concept in OpenStreetMap may complicate data searching because of the *folksonomy* when introducing data. With Text To Speech ans Speech To Text technology, it may increase the query processing time and may make it more sensitive because of words redundancy, voice recognition and synthesis time delay. It is therefore essential to build an efficient method to extract data and to fix rules on how to build useful requests. The taxonomy presented here is simple, easy to memorize. It takes into account only the most visited locations during the daily life of each individual, if they beloni to one of domains described in Fig. 3. This looks to be a solution to circumvent the negative effects of free choice of tags during the process of data production. The proposed solution is thus suitable for this vocal HMI dedicated to blinds.

References

[1] Anthony Hombiat, Marlne Villanova-Oliver, and Jrme Gensel, "Un nouveau mta-modle pour rapprocher la folksonomie et lontologie dOSM," in *Colloque International de Gomatique et d'Analyse Spatiale (SAGEO 2015), Nov 2015, Hammamet, Tunisia. 2015*, 2015.

[2] Axel Bruns, "The Future Is User-Led: The Path towards Widespread Produsage," *Fibreculture Journal*, 2008. (Online). Available: http://eprints.qut.edu.au/12902/1/12902.pdf.

[3] David J. Coleman, Yola Georgiadou, and Jeff Labonte, "Volunteered Geographic Information: The Nature and Motivation of Produsers ," *International Journal of Spatial Data Infr astructures Research, Vol. 4*, 2009. (Online). Available: http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/viewFile/140/223.

[4] Said Boularouk, Didier Josselin, and Eitan Altman, "Interaction Homme Machine pour l'orientation spatiale des malvoyants (article soumis)." Nice, SAGEO'2016.

[5] C. V. (Ed), *Les ontologies spatiales*. Lavoisier, Paris, Revue Internationale de Gomatique, Volume 14 N 2/2004, 2004.

[6] Mihai Codescu, Gregor Horsinka, Oliver Kutz, Till Mossakowski, and Rafaela Rau, "OSMonto - An Ontology of OpenStreetMap Tags," 2011. (Online). Available: http://www.informatik.uni-bremen.de/ till/papers/osmonto.pdf.

[7] Mihai Codescu, Daniel Couto Vale, Oliver Kutz, and Till Mossakowski, "Ontology-based Route Planning for OpenStreetMap," 2012. (Online). Available: https://www.semanticscholar.org/paper/Ontology-based-Route-Planning-for-Openstreetmap-Codescu-Vale/33a9ea38469a222a621fa1ac7cade89f0d301a5a.