SOIT: Spatial Ontologies Integration Tool and its Application to the Road Domain

Wassim Jaziri, Sana Châabane, Faïez Gargouri

Miracl Laboratory, ISIM Sfax,
Pôle Technologique, Route de Tunis Km 10,
B.P. 242, SFAX 3021, Tunisia
Emails: (jaziri.wassim, sana.chaabane, faiez.gargouri)@gmail.com

***Abstract*—Designed as semantic structures to support the sharing and reuse of geographic data, spatial ontologies have recently gained attention within the geo-information community. Geographic ontologies are designed to provide a common understanding of the structure of geographic models, and to support the development of geographic information systems that are conceptually complex. This paper proposes an approach for merging spatial ontologies based on three complementary modules: matching, mapping and merging. A Spatial Ontologies Integration Tool (SOIT) is also developed and applied to the road domain.**

Keywords-Spatial ontologies; SOIT; Integration tool; Geographic Information systems; Road domain.

I. INTRODUCTION

In the applications with spatial vocation (e.g., geographic domain), ontologies are an effective solution in particular to ensure interoperability and semantic cooperation between Geographic Information systems (GIS). Spatial ontologies offer a relevant solution for the sharing and the integration of geographic data.

The problem of heterogeneity of geographic ontologies is more complex than that of other domain ontologies [5]; because it is necessary to take into account the spatial and temporal aspects as well as rules governing the data evolution.

Spatial data interoperability allows simplifying and enhancing the sharing, reuse and integration of geographic data. However, semantic heterogeneity [8] is a major obstacle to the interoperability of geographic data [9]. Indeed, the implementation of a geographic ontology can manage and structure multiple data sets that can be grouped according to geographical criteria. Its objective is not to only describe the list of existing geographic objects (territory, boundary, road network, etc.) but to identify classes; to define the relationships may exist between them, and to describe the attributes in order to obtain the knowledge base.

In this paper, we aim to resolve the problem of heterogeneity of geographic ontologies by proposing a merging approach.

This paper is organized as follows: Section 2 presents an overview about tools and techniques for merging ontologies. The Third Section details the proposed approach for merging geographical ontologies. A spatial ontologies integration tool is described in Section 4. Section 5 presents the application of our approach and tool on the road domain. The conclusion and outlook of our work are listed in Section 6.

II. ONTOLOGY MERGING TOOLS AND TECHNIQUES

Several ontologies merging tools were developed in literature, such as [6]: Chimaera, FCA-Merge, PROMPT, OntoMorph and ONIONS.

- Chimaera: It is an interactive ontologies merging tool that allows the diagnosis, the test and the edition of the merging result [4]. It helps user to find the best term by proposing a list of the used terms while helping to resolve the terminological difficulties. This tool, based on the Ontolingua ontology editor, offers a support for the merging process to enable the collection of ontologies expressed in different formalisms. It makes the translation at the language level and uses heuristics to find the parts of the ontology to be reorganized.

- PROMPT: Based on a semi-automatic merging approach, it allows making certain tasks automatically and helps the user along the merging process [5]. PROMPT determines possible filminesses in the state of the ontology resulting from user's actions and suggest solutions for them.

- OntoMorph is based on a merging approach which is similar to the two previous tools [2]. An expert uses an initial list of correspondences between concepts of the source ontologies: the user defines a set of operators that are applied to ontologies for resolving inconsistencies.

- FCA-Merge: It uses a formal, bottom-up method of ontology merging based on the extraction of concepts from textual documents [7]. It applies natural language processing and generates a "concept trellis" from “FCACore” algorithm. This trellis is transformed subsequently into domain ontology by an expert of the domain.

- ONIONS (ONtological Integration Of Naive Sources) is a method designed for the conceptual analysis and ontological integration of terminologies [3]. This method

 ISBN: 978-1-61208-104-5

consists of two steps: (1) A reengineering step which consists in the extraction, formatting, analysis and formalization of data; (2) A merging step which allows the merging of ontologies using an algorithm based on algebra.

The developed tools do not consider the spatial aspect of objects describing the geographical domain. The spatial dimension as well at the intrinsic level of the concepts at the level of the spatial and semantic relationships were lacking to these tools. This limits their applicability in geographical ontologies.

III. SPATIAL ONTOLOGIES MERGING APPROACH

We developed an approach for geographical ontologies merging, based on two criteria:

- *The identity search*: the search for relationships of spatial identity and total identity between concepts of the initial ontologies.

Definition 1. Two objects are spatially identical if they are located in the same place but having a different characteristic such as the instance name or the acquisition date.

Definition 2. Two objects are totally identical if they are spatially and semantically identical. By semantic identity, we mean that both objects have the same name and the same properties.

This criterion allows obtaining the skeleton of the ontology result of merging process. We thus join the not identical individuals of the candidate ontologies to serve as entries to the second step of the merging process.

- *The search for enrichment relationships*. Enrichment relationships have two types: the semantic relationships such as *equivalence* and *part-of* and the spatial relationships such as *adjacency, intersection, joint, junction* etc.

The proposed approach is based on three main modules: (1) matching module, (2) mapping module and (3) merging module (Figure 1). The first module consists in determining the matching process between candidate ontologies. The output of this phase is a list of matching functions. The second phase allows finding correspondences between concepts of candidate ontologies. The result of this phase is two lists: a list of matches between candidate concepts and a second list of concepts without correspondences. The third phase is merging which is based on merging rules. It produces as result a comprehensive ontology spatially and semantically richer than the candidate ontologies.

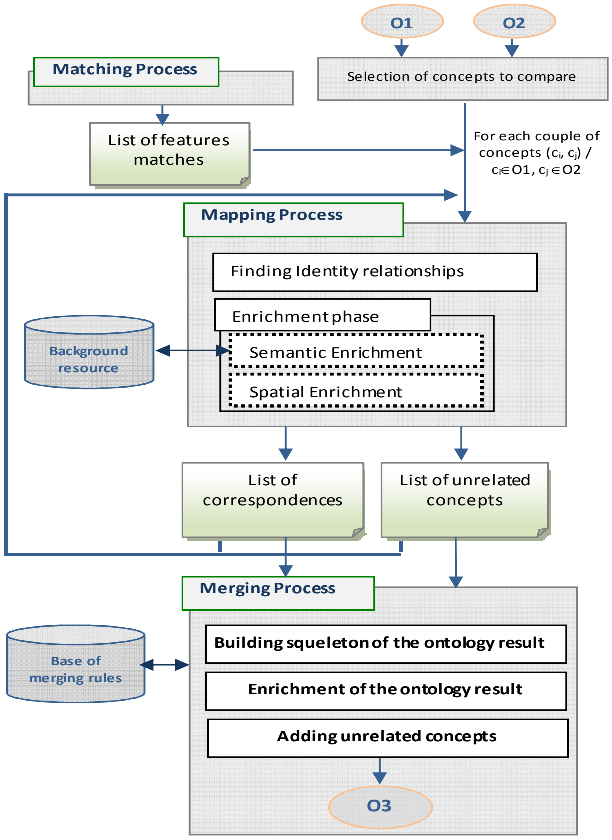


Figure 1. The proposed approach for merging spatial ontologies.

3.1 The matching phase

A matching process defines a set of functions which specifies correspondences between terms of ontologies. This phase gives as a result a list of features matches. There are two types of relationships considered in our approach:

- Connecting relationships: they are the relating points between two ontologies. We distinguish spatial identity relationships and semantic identity relationships.

- Enriching relationships: we distinguish semantic relationships and topological relationships (intersection, union, etc.).

We use two types of matching: spatial and semantic matching. In the semantic matching, we define two functions: the first one defines semantic identity relationship (Idsem) and the second function defines semantic enrichment relationships between candidate concepts. In the spatial matching, we define two functions: the spatial identity relationship (Idspa) and the spatial enrichment relationships.

The semantic Identity: Idsem means that two concepts have the same name and same properties. We use the syntactic technique to derive such relationship. We use the edit distance of Levenshtein to calculate the similarity between concepts names and their properties. This measure of edit distance ed represents the minimum number of insertions, deletions or substitutions necessary to transform one string x into another y. Similarity s(x, y) normalized to [0,1] is defined as follows:

s=1-ed/max(| x |,| y |).

We consider that two concepts C1 and C2 admit a semantic identity relationship Idsem(C1, C2) if and only if:

s(C1.name,C2.name) = 1

and for every attribute atti of C1, there exists an attribute attj of C2 where s(C1.atti, C2.attj)=1, and vice versa, for every attribute of C2, there is an attribute of C1 where s(C2.attj,C1.atti) = 0.

The spatial identity relationship Idspa relates only to geographical concepts. A spatial object is described according to its graphical form: GF (point, line or polygon), semantics data (eg name, nature, appearance, various characteristics) and localization data (position on the surface). The search for Idspa relationship between concepts of candidate ontologies consists in comparing localization characteristics of concepts.

We formally define the relationships considered by our approach. We present the following formal definitions defined.

Definition 1: Two objects are spatially identical if they are located in the same place but having a different characteristic such as the instance name or the acquisition date.

Definition 2: Two objects are totally identical if they are spatially and semantically identical.

Let us consider the concepts C1(X1, Y1) and C2(X2, Y2) with X1, Y1 and X2, Y2 are coordinates of C1 and C2 respectively and C1.GF=point and C2.GF=point. We consider Idspa(C1, C2) if and only if Euclidean distance dE(C1,C2) =0. The function of spatial identity relationship is defined as follows:

$$\mathbf{IDspa:=}\left\{\begin{matrix}\mathbf{C1 \in O1, C2 \in O2 \wedge C1.FG = point \wedge} \\ \mathbf{C2.FG = point \wedge dE = 0}\end{matrix}\right\}. \quad (1)$$

$$\mathbf{dE = \sqrt{(x1 - x2)^2 + (y1 - y2)^2}}. \quad (2)$$

3.2 The mapping phase

The input of this phase is two concepts from both candidate ontologies. The mapping process is iterative and consists of two steps. The first step is to investigate Identity relationships and the second step is to investigate Enrichment relationships between candidate concepts. The search for enrichment relationships is performed on non-identical concepts selected at the previous phase. The compared concepts and their correspondences are stored in a base of matches. Concepts which the mapping algorithm found no connections between them, i.e. the concepts that do not verify any type of relationship between them (called unrelated concepts) are stored in a base of unrelated concepts.

We have defined rules to optimize the number of comparisons of concepts in order to avoid a randomly process. For example, to research identity relationship, we rely on the type of the graphic form of the concept to make comparisons.

3.3 The merging phase

The merging phase consists in building the ontology result. The input of this phase is composed of bases of correspondences and unrelated concepts. The aim of this module is to apply the correspondence links stored at the correspondences bases (semantic and geographical) in accordance with merging techniques. The merging process creates a new geographical ontology from two candidate ontologies connected by identity concepts which are used as connected points between the two ontologies and enriched by the semantic and geographical relationships.

Rules for merging candidate concepts are defined. These rules are of two types: rules for semantic relationships and rules for spatial relationships. The merging rules are applied to concepts accepting connections between them. Unrelated concepts are transmitted in the ontology result without any treatment.

IV. SOIT: SPATIAL ONTOLOGIES INTEGRATION TOOL

We have developed the SOIT tool (Spatial Ontologies Integration Tool) (Figure 2) based on Java language and the integrated development environment (IDE) NetBeans. The tool is designed for automatically merge two spatial ontologies. SOIT takes as input two spatial ontologies written in OWL and produces as result an ontology spatially and semantically richer. In addition, SOIT allows other functions: it can perform two types of matching candidate ontologies and see the result of the matching process. It can also generate the graph of an ontology written in OWL and view or print one ontology in the form of a text or a graph. We model a use case diagram of UML language representing the various functionalities of the SOIT tool (Figure 2).

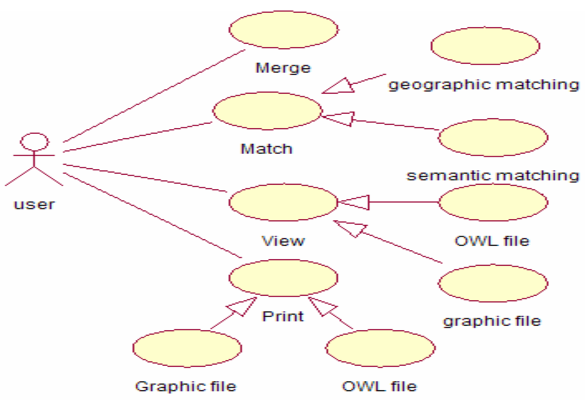


Figure 2. The Use Case diagram of SOIT.

The host interface of SOIT includes a menu bar contains five menus: “File”, “View”, “Match”, “Merge” and “Help”.

The matching process starts with the introduction of two candidate ontologies (figure 3).

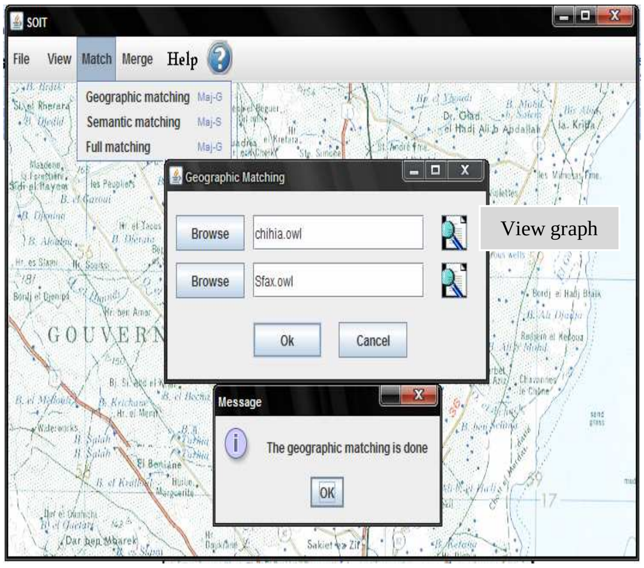
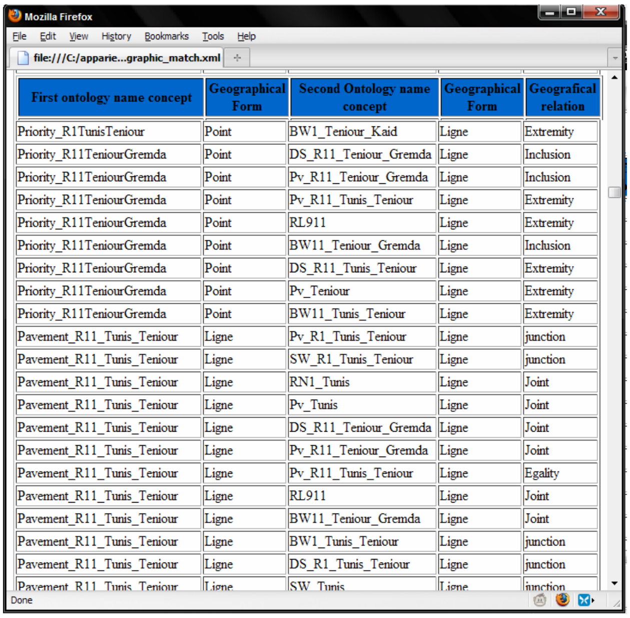


Figure 3. The functionalities of the "Match" menu.

The graph and the OWL file of a candidate ontology can be viewed through the button "view graph" (figures 4).

Figure 4. The OWL file of the candidate ontology.

After running the matching process, the system displays the list of matches found. This functionality is performed using XSLT style sheets (figure 5). For example, we have identified a relationship of type *Extremity* between individuals: *Priority_R1_Tunis_Teniour* and *BW1_Teniour_Kaied*.



First ontology name concept	Geographical Form	Second Ontology name concept	Geographical Form	Geografical relation
Priority_R1TunisTeniour	Point	BW1_Teniour_Kaid	Ligne	Extremity
Priority_R11TeniourGremda	Point	DS_R11_Teniour_Gremda	Ligne	Inclusion
Priority_R11TeniourGremda	Point	Pv_R11_Teniour_Gremda	Ligne	Inclusion
Priority_R11TeniourGremda	Point	Pv_R11_Tunis_Teniour	Ligne	Extremity
Priority_R11TeniourGremda	Point	RL911	Ligne	Extremity
Priority_R11TeniourGremda	Point	BW11_Teniour_Gremda	Ligne	Inclusion
Priority_R11TeniourGremda	Point	DS_R11_Tunis_Teniour	Ligne	Extremity
Priority_R11TeniourGremda	Point	Pv_Teniour	Ligne	Extremity
Priority_R11TeniourGremda	Point	BW11_Tunis_Teniour	Ligne	Extremity
Pavement_R11_Tunis_Teniour	Ligne	Pv_R1_Tunis_Teniour	Ligne	junction
Pavement_R11_Tunis_Teniour	Ligne	SW_R1_Tunis_Teniour	Ligne	junction
Pavement_R11_Tunis_Teniour	Ligne	RN1_Tunis	Ligne	Joint
Pavement_R11_Tunis_Teniour	Ligne	Pv_Tunis	Ligne	Joint
Pavement_R11_Tunis_Teniour	Ligne	DS_R11_Teniour_Gremda	Ligne	Joint
Pavement_R11_Tunis_Teniour	Ligne	Pv_R11_Teniour_Gremda	Ligne	Joint
Pavement_R11_Tunis_Teniour	Ligne	Pv_R11_Tunis_Teniour	Ligne	Egality
Pavement_R11_Tunis_Teniour	Ligne	RL911	Ligne	Joint
Pavement_R11_Tunis_Teniour	Ligne	BW11_Teniour_Gremda	Ligne	Joint
Pavement_R11_Tunis_Teniour	Ligne	BW1_Tunis_Teniour	Ligne	junction
Pavement_R11_Tunis_Teniour	Ligne	DS_R1_Tunis_Teniour	Ligne	junction
Pavement_R11_Tunis_Teniour	Ligne	SW_Tunis	Ligne	junction

Figure 5. Geographic matching result.

For merging two ontologies, the user has to introduce two geographic ontologies instances of the same model, by clicking on the button “Browse”. The following window displays (Figure 6).

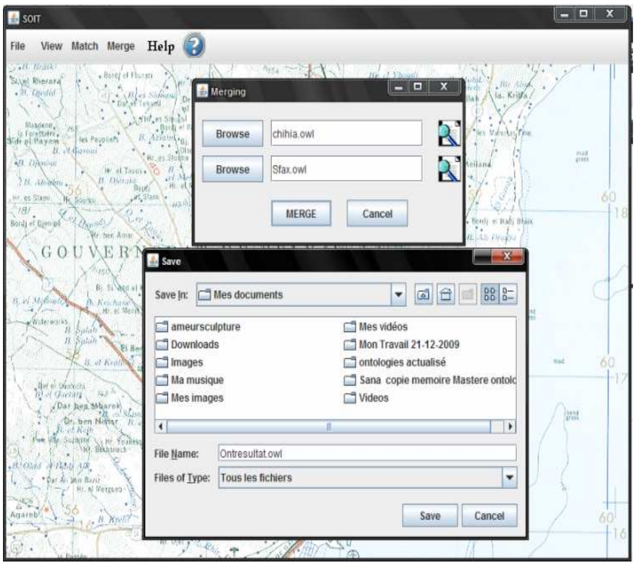


Figure 6. Selection of a candidate ontology.

Finally, by clicking on the button "Merge", the user can visualize the concepts and the individuals of the ontology result (figure 7).

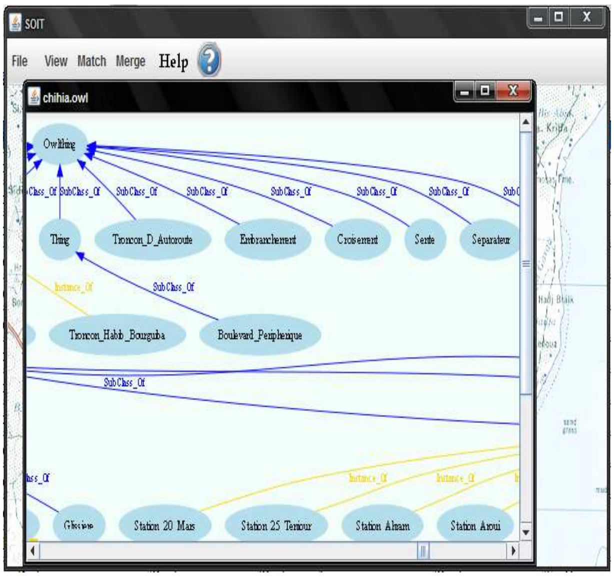


Figure 7. Graph of the ontology result.

V. APPLICATION TO THE ROAD DOMAIN

The application domain of the developed spatial ontologies integration tool is the road domain. We developed two spatial ontologies related to the city of Sfax (Tunisia), called respectively *ontoRoadChihia.owl* and *ontoRoadSfax.owl* instances of the OntoRoad ontology [1] which is developed to model the road domain concepts .

The studied corpus is composed of topographic maps. The instantiation of the *OntoRoad* ontology is made by geographical zone. Both candidate ontologies subject of experiment cover different geographical zones from the city of *Sfax* (Tunisia). We extract all the objects of the considered zone and we attribute them to their corresponding classes. For example, the object *Hedi_Chaker* is a Street; the street *Ibn_kholdoun* is one-way.

The following extract presents the modelling of the individual “*RL911*” of the concept “*Local_Road*” in the ontology *ontoRoadChihia.owl* (Table I).

TABLE I. EXTRACT OF THE ONTOLOGY ONTOROADCHIHIA.OWL.

<Local_Road rdf:about="**#RL911**">
<Position_Route rdf:datatype="&xsd;string">**0.0**</Position_Route>
<Debut_De_Section_Voie rdf:datatype="&xsd;string"
>**Carrefour_G_3Chemins**</Debut_De_Section_Voie>
<Fin_De_Section_Voie
rdf:datatype="&xsd;string">**Km12**</Fin_De_Section_Voie>
<Forme_geometrique
rdf:datatype="&xsd;string">**Ligne**</Forme_geometrique>
<Nom_Route
rdf:datatype="&xsd;string">**Route_Teniour**</Nom_Route>
<Rencontre_Voie_Voie rdf:resource="**#Av_7Novembre**"/>
<A_Droite_De rdf:resource="**#Av_7Novembre**"/>
<Rencontre_Voie_Voie rdf:resource="**#Av_Afrique**"/>
<A_Droite_De rdf:resource="**#Av_Afrique**"/>
<A_Droite_De rdf:resource="**#Av_Teboulbi**"/>
<A_Droite_De rdf:resource="**#BW11_Teniour_Gremda**"/>
<Rencontre_Voie_Voie
rdf:resource="**#BW11_Teniour_Gremda**"/>
<A_Gauche_De rdf:resource="**#BW11_Tunis_Teniour**"/>
<Rencontre_Voie_Voie rdf:resource="**#BW11_Tunis_Teniour**"/>
<Rencontre_Voie_Voie rdf:resource="**#BW1_Teniour_Kaid**"/>
<A_Droite_De rdf:resource="**#BW1_Teniour_Kaid**"/>
<Rencontre_Voie_Voie rdf:resource="**#BW1_Tunis_Teniour**"/>
<A_Gauche_De rdf:resource="**#BW1_Tunis_Teniour**"/>
<A_Gauche_De rdf:resource="**#RL921**"/>
<Rencontre_Voie_Voie rdf:resource="**#RL921**"/>
<Adjacence_Route_Trottoir rdf:resource="**#SW_Teniour**"/>
<Rencontre_Voie_Voie rdf:resource="**#S_Khaledwalid**"/>
<A_Droite_De rdf:resource="**#S_Khaledwalid**"/>
<Rencontre_Voie_Voie rdf:resource="**#S_Tina**"/>
<A_Droite_De rdf:resource="**#S_Tina**"/>
</Local_Road>

The following extract presents the modelling of the same individual in the *ontoRoadSfax.owl* (Table II).

TABLE II. EXTRACT OF THE ONTOLOGY ONTOROADSFAX.OWL.

<Local_Road rdf:about="**#RL911**">
<Position_Route rdf:datatype="&xsd;string">**0.0**</Position_Route>
<Debut_De_Section_Voie rdf:datatype="&xsd;string"
>**Carrefour_G_3Chemins**</Debut_De_Section_Voie>
<Fin_De_Section_Voie
rdf:datatype="&xsd;string">**Km12**</Fin_De_Section_Voie>
<Forme_geometrique
rdf:datatype="&xsd;string">**Ligne**</Forme_geometrique>
<Nom_Route
rdf:datatype="&xsd;string">**Route_Teniour**</Nom_Route>
<Rencontre_Voie_Voie rdf:resource="**#Av_5Aout**"/>
<Rencontre_Voie_Voie rdf:resource="**#Av_Majida_Boulila**"/>
<A_Droite_De rdf:resource="**#BW11_Teniour_Gremda**"/>
<Rencontre_Voie_Voie
rdf:resource="**#BW11_Teniour_Gremda**"/>
<A_Gauche_De rdf:resource="**#BW11_Tunis_Teniour**"/>
<Rencontre_Voie_Voie rdf:resource="**#BW11_Tunis_Teniour**"/>
<Rencontre_Voie_Voie rdf:resource="**#BW1_Teniour_Kaid**"/>
<A_Droite_De rdf:resource="**#BW1_Teniour_Kaid**"/>
<Rencontre_Voie_Voie rdf:resource="**#BW1_Tunis_Teniour**"/>
<A_Gauche_De rdf:resource="**#BW1_Tunis_Teniour**"/>
<Connexion_Extremite-Noeud rdf:resource="**#GCR_3Chemins**"/>
<A_Droite_De rdf:resource="**#RL_Kaid**"/>
<A_Gauche_De rdf:resource="**#RN1_Tunis**"/>
<Adjacence_Route_Trottoir rdf:resource="**#SW_Teniour**"/>
</Local_Road>

The following extract presents the result of merging of these two ontologies (Table III).

TABLE III. EXTRACT OF THE ONTOLOGY RESULT.

```
<Local_Road rdf:about="#RL911">
	<Position_Route rdf:datatype="&xsd;string">0.0</Position_Route>
		<Debut_De_Section_Voie rdf:datatype="&xsd;string"
	  >Carrefour_G_3Chemins</Debut_De_Section_Voie>
	<Fin_De_Section_Voie
	 rdf:datatype="&xsd;string">Km12</Fin_De_Section_Voie>
	<Forme_geometrique
rdf:datatype="&xsd;string">Ligne</Forme_geometrique>
	<Nom_Route
rdf:datatype="&xsd;string">Route_Teniour</Nom_Route>
	<Rencontre_Voie_Voie rdf:resource="#Av_5Aout"/>
	    <Rencontre_Voie_Voie rdf:resource="#Av_7Novembre"/>
	<A_Droite_De rdf:resource="#Av_7Novembre"/>
	<Rencontre_Voie_Voie rdf:resource="#Av_Afrique"/>
	<A_Droite_De rdf:resource="#Av_Afrique"/>
	  <Rencontre_Voie_Voie rdf:resource="#Av_Majida_Boulila"/>
	<A_Droite_De rdf:resource="#Av_Teboulbi"/>
	    <A_Droite_De rdf:resource="#BW11_Teniour_Gremda"/>
	<Rencontre_Voie_Voie
rdf:resource="#BW11_Teniour_Gremda"/>
	    <A_Gauche_De rdf:resource="#BW11_Tunis_Teniour"/>
	<Rencontre_Voie_Voie rdf:resource="#BW11_Tunis_Teniour"/>
	  <Rencontre_Voie_Voie rdf:resource="#BW1_Teniour_Kaid"/>
	      <A_Droite_De rdf:resource="#BW1_Teniour_Kaid"/>
	<Rencontre_Voie_Voie rdf:resource="#BW1_Tunis_Teniour"/>
	     <A_Gauche_De rdf:resource="#BW1_Tunis_Teniour"/>
	<Connexion_Extremite-Noeud rdf:resource="#GCR_3Chemins"/>
	<A_Gauche_De rdf:resource="#RL921"/>
	<Rencontre_Voie_Voie rdf:resource="#RL921"/>
	<A_Droite_De rdf:resource="#RL_Kaid"/>
	<A_Gauche_De rdf:resource="#RN1_Tunis"/>
	    <Adjacence_Route_Trottoir rdf:resource="#SW_Teniour"/>
	     <Rencontre_Voie_Voie rdf:resource="#S_Khaledwalid"/>
	<A_Droite_De rdf:resource="#S_Khaledwalid"/>
	<Rencontre_Voie_Voie rdf:resource="#S_Tina"/>
	<A_Droite_De rdf:resource="#S_Tina"/>
</Local_Road>
```

VI. CONCLUSION AND FUTURE WORK

The need to combine ontologies developed in an independent way and containing heterogeneity, raised problems from the point of view of the ontological language, the conceptualization and the specification. The heterogeneity between the knowledge expressed within each of the ontologies treating the same domain must be resolved. Several solutions to produce much more successful ontologies were proposed and varied techniques were developed for the adaptation, the merging and the integration. The integration is the construction of a new ontology reusing the other available ontologies which will be a part of the new ontology. The logical integration of two ontologies supplies to the user a vision unified by various sources.

In this paper, we have presented an approach for merging geographic ontologies. This approach consists of three processes: (1) the matching process, (2) the mapping process and (3) the merging process. We also developed SOIT: a tool for spatial ontologies integration. The application of this tool has been made on the road domain. Our ongoing work are to evaluate “SOIT” by comparing the result produced by this tool with the one developed by an expert in the field. In future work, we aim at extending this tool with functionalities for query ontological data bases.

REFERENCES

[1] Chaabane, S. and Jaziri, W., OntoRoad : Spatial ontology for the road domain, Journées Francophones sur les Ontologies (JFO’2008), pages 135-144, December1-2, 2008, Lyon, France.

[2] Chalupsky, H., OntoMorph: A translation System For Symbolic Kcowledge, In proceedings of the 7th international conference on principles of knowledge representation and reasoning (KR’2000), pages 471-482, Colorado, USA, 2000.

[3] Gangemi, A., Pisanelli, D.M. and Steve, G., Ontology Integration: Experience with medical terminologies, In N. Guarino (ed.), Formal Ontology in Information Systems, pages 163-178. IOS Press, 1998.

[4] McGuinness, D.L., Fikes, R., Rice, J. and Wilder, S., The Chimaera ontology Environment. In proceeding of the seventeenth national conference on artificial intelligence (AAAI 2000), 2000.

[5] Noy, N.F. and Musen, M.A., PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment. In proceeding of the 17th national conference on artificial intelligence (AAAI 2000), 2000.

[6] Poulain, T. and Cullot, N., Ontology Mappings for Information Systems Cooperation. Deuxièmes Journées Francophone sur les Ontologies (JFO2008), ACM Edition, pages 47-52, 2008.

[7] Stumme, G. and Maedche, A., FCA-merge: Bottom-Up Merging of Ontologies, Proceedings of the International Conference on Artificial Intelligence (IJCAI’01), pages 225-230, USA, 2001.

[8] Villie M. and Félix S., Ontologies: solving semantic heterogeneity in a federated spatial database system. Proceeding of 5th International Conference on Enterprise Information System, pages 347-352, Angers, France, April 2003.

[9] Sheth A.P. and Larson J.A.., Federated database systems for managing distributed heterogeneous and autonomous databases. ACM Computing Surveys, 1990.