Virtual Activity Ontology Modeling: the Case of Sociocultural Knowledge Sharing

Kaladzavi Guidedi
University of Maroua
Maroua, Cameroon
kaladzavi@ymail.com

Papa Fary Diallo
University of Gaston Berger
Saint-Louis, Senegal
diallo.papa-fary@ugb.edu.sn

Kolyang
University of Maroua
Maroua, Cameroon
dtaiwe@gmail.com

Moussa LO
University of Gaston Berger
Saint-Louis, Senegal
moussa.lo@ugb.edu.sn

***Abstract*-In the light of the rhythm of the current cultural mixing, we believe that in the long term, culture of African people in particular may disappear. Some new computational techniques (semantic web technologies) are needed to manage the large repositories of sociocultural data and to discover useful patterns and knowledge from them. This paper presents a virtual activity ontology modeling approach, in the case of sociocultural knowledge sharing and co-construction named Ontoshare. Our modeling approach is based on Engeström's Human Activity Theory (HAT). With Ontoshare we designed how Internet users could build the content of a sociocultural Knowledge Management System (KMS); this vocabulary also organises data, facilitates information retrieval by introducing a semantic layer in social web platform, we plan to implement. The platform could be considered as a « collective memory » which will allow communities to share, co-construct and discover sociocultural knowledge in the Cameroonian context.**

Keywords-sociocultural knowledge; sharing; Human Activity Theory.

I. INTRODUCTION

Since 2005, UNESCO has projected that local knowledge will increasingly become the main point of social mutations; it leads economic, political, and sociocultural projects. Emerging societies must avoid being mere components of the "Global Information Society". Effective participation of African Countries in "Societies of knowledge" is required [1]. While, local knowledge is considered in the conception of different project development, B. Z. Deli deplores the fact that, in Africa, cultures are deteriorating and emptying of their meanings, their mellow content and values [2]. To promote indigenous knowledge, some media have been proposed: a permanent (re)education, radio broadcasting, and of course Internet. On the Web, any topics may be discussed. As a result, the Web constitutes the source of global information. However, when we consider how Internet works by focusing on data flow, social values such as equality, freedom, democracy which are supposed to characterise the "Global Information Society" appear to be an illusion. The unbalance between these social values is described by "*digital divide*" issue between Northern and Southern countries. The "*digital divide*" has many views. Out of these views, there are "*divide by access*", "*divide by use*", "*divide by decision*", and "*divide by content*". The "*divide by content*" is defined as the gap between Western culture over-represented and African culture under-represented. This gap is considerable [3]. To reduce this gap and promote the African culture, to refresh the memory of our citizens and show the transparent view of opportunities (unknown infrastructures, spaces, etc.) and challenges (investments unequally distributed, marginalised communities, etc.) through endogenous information (from involved actors), rather than external analysis, we propose, the implementation of ontologies-based web platform for sociocultural knowledge sharing and co-construction in the Cameroonian context.

By Web 2.0, Internet users are no longer just consumers but also authors of information. Semantic Web is a Web evolution, through ontologies; it improves data organisation, and information retrieval. Ontology is an explicit specification of a shared conceptualisation [4]. As such, ontologies are considered to be knowledge representation tools, transforming data into information and information into knowledge. They are recommended tools for knowledge reuse, organisation, interoperability, integration, and valorisation on the Web [5].

The objective of this paper is to present a virtual activity ontology modeling approach in the case of sociocultural knowledge sharing and co-construction named Ontoshare. Ontoshare modeling approach is based on Engeström's Human Activity Theory (HAT). HAT is a conceptual framework which, the foundational concept is "activity", understood as purposeful, transformative, and developing interaction between actors ("subjects") and the world ("objects") [6].

This paper is organised as follows: In Section 2, we present research work related to our domain. In Section 3, we shall explain our modeling approach, which is based on Engeström's Human Activity Theory and how we reused the related ontologies to solve interoperability issue. Moreover, we will describe in Section 4, how Ontoshare has been populated and improved through inference services provided by reasoning engine in Protégé. In the same section, we will

 ISBN: 978-1-61208-443-5

point out some SPARQL queries on a Knowledge Base related to the use cases to illustrate how the "collective memory" could be built. This will end with, a conclusion and future work.

II. RELATED WORK

To the best of our knowledge, only the Sociocultural Ontology [7][8] covers the sociocultural domain which has been developed in the Senegalese context. The modeling approach is based on Vygotsky framework [9]. This framework is considered as the first Generation of HAT. It is organised around the "*mediation*" concept and based on the idea that, human actions are mediated by cultural, symbolic or physical artifacts that enable man to act on his environment.

The main limitation of Vygotsky's framework is the fact that, it focuses solely on individual actions and not on individual actions within the community. The modeling approach proposed in [7][8] tried to solve this limitation by substituting *subject* by *community*. In this model, there are three main axes, which are:

1. *Community*: which is a group of people sharing a common interest in a sociocultural domain;
2. *Object*: considered as a locality or infrastructure where community evolves;
3. *Artifact*: it mediates the actions of community with object.

Translating *subject* by *community* has just solved the problem concerning community action but has failed to model the dynamics within the community. The authors of this model consider the community as an atomic entity. The modeled ontology in [7] hides some knowledge on the internal dynamics (collaborations, interactions, actors, roles, etc.) of the community and the contextual nature (regulations) while organising activities considering the fact that knowledge on internal dynamic within community will enrich our Knowledge Base and allow deep analysis of communities and activities.

Due to the limitations of sociocultural ontology [7], we proposed in our previous paper an improved sociocultural ontology named OntoSOC [10]. Its modeling approach considers the dynamics within communities and the contextual nature for organising activities. This vocabulary enabled us to semantically circumscribe the content of sociocultural knowledge.

However, on the Web, the flows have been reversed: Internet user is no longer passive (reader) but active (author). The transition from Web 1.0 to Web 2.0 done in 2004 was a decisive transition to social media paradigm. Social media includes all tools and applications that allow interaction between Internet users. Within that "*galaxy*" of social media, there are several "*planets*". Out of them, there are texts publishing tools (wikis, blogs, etc.), exchanging and sharing tools (YouTube for videos, Slideshare for sharing presentations, etc.), tools for discussion (Skype, Messenger, etc.), and networking tools (Twitter, Facebook, MySpace, etc.), etc., are in continuous supply.

Unfortunately, most of them are not sharing and discovering local knowledge oriented, except Wikipedia [11], the online encyclopedia which is inter-domain and inter-society. The second limitation and not of lesser importance, is the fact that they are not ontologies-based. In the case of Wikipedia, DBpedia [12] project has been launched to address this issue. DBpedia is an academic and community project for automatic data exploration from Wikipedia to propose a structured version in semantic web format (RDF) of data. Due to the previous limitations and the unknown of a sociocultural sharing ontology, designing ontology of how Internet users could interact to share knowledge is needed.

III. ONTOSHARE MODELING

In this section, we present, why and how we used Human Activity Theory to model Ontoshare.

A. Methodology

Practically, ontological engineering does not propose a standardised methodology for designing ontologies [13]. Our concern is to implement a "*collective memory*" system helping people to have a holistic view on local changes, while considering culture and historicity in our localities. Otherwise, we agree with P. Berger and T. Luckmann in [14] with the fact that reality is a social construction and that, the universe is evolving. These changes are driven by groups of individuals through their various activities. For analysing and understanding these changes and how they transform reality, some activity models have been proposed. There are three generations of this model: the first and the second are characterised by focusing on subject and its actions; the third generation is known as collective view framework making distinction between individual actions and collective activities. For these purposes, Engeström's model is more adapted and has been chosen as framework of our modeling process.

Engeström's model clearly points out distinction between individual action and collective activity. It is produced according to the historical and cultural view of activity. The model has six poles (*Subject, Object, Tools, Rules, Community and Division of Labour*) as shown in Figure 1.

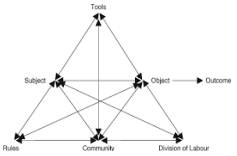

Figure 1. Engeström's Human Activity triangle

- *Subject*: represents the chosen individual to analyse;
- *Object*: environment transformation by activity (task to be performed, objective to be achieved);
- *Tools*: materials or symbolic tools that mediate activity ;
- *Comm*unity: set of individuals that share the same interests and thereby differ from other communities;
- *Division of Labour:* it considers, the horizontal distribution of actions among subjects, community members and the vertical hierarchy or responsibilities and status;
- *Rules*: they refer to implicit and explicit standards, conventions, habits, etc. that maintain and regulate actions and interactions within the community.

Many Human-Computer-Interaction (HCI) researchers were eager to move beyond the confines of traditional cognitive science to HAT direction [16]. Activity Theory is a powerful and clarifying descriptive tool rather than a strongly predictive theory. It incorporates strong notions of intentionality, history, mediation, collaboration and development in constructing conscientiousness. HCI research needs to understand and describe "*context*", "*situation*" and "*practice*". Considering these previous features of HCI, its objectives are related to HAT objectives. Although, the HAT is developed for offline communities, the concept of online community is still the set of people (Internet users) who share the same values and interests. Thus, to analyse how internet users will interact with computer to co-construct knowledge in our context, Human Activity Theory constitutes the powerful tool [17].

B. Concepts and Relationships identification process

In this section, we focus on how Engeström' HAT has been used to construct our ontology (Ontoshare). In other words, we want to show how we deduced concepts and relationships from HAT. To do that, due to "collaborative persona" approach, we simulated HAT in three use cases of online knowledge sharing. "*Persona method*" is a modeling strategy used by software architects. This idea was introduced by A. Cooper, software designer [18]. In Software Engineering, this approach is called "*Goal-directed design*". It represents patterns of users' behavior, goals and motivations, compiled in a fictional description of a single individual. It also contains made-up personal details, in order to make the person more "*tangible and alive*" for the development team. In our case, we use "*collaborative persona*" suitable for collaborative, participative and interactive context as community [19].

To simulate the online shared information, the following activities were chosen: a cultural event organised by *NakoSenda* community in Mokolo (locality), rural library building activity conducted by CDE-SAARE [20] in Kolara (locality) and a soccer tournament holidays organised by *Club 2-0-UMa.*

There are two types of Internet users: passive and active users. Active user is the one who edits (shares) information. For the purpose of this paper, active user is named contributor. Any contributor can be a member of a community which organises activity, or a witness who attends the activity or just anyone who has information (partial or complete) to share on an activity. For example, in the first case, a contributor named D, created in our platform a page named "*cultural event in Mokolo",* in which he edited information about the location (weather forecast, vegetation, etc.) of Mokolo. The second contributor named MP who is Nakosenda's member, shared information on the same page relative to *NakoSenda* (the founder, headquarter, regulations, type). The third contributor called M shared his impression on the event. All these contributors shared and co-constructed sociocultural knowledge on cultural event organised by *NakoSenda* in the respect of platform regulations. Thus, for each use case and for each contributor, we simulated separately, twelve triads within the overall triangle as shown in Figure 2.

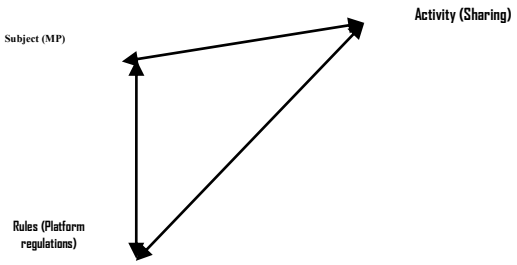

Figure 2. Subject-regulations-Activity triad

As seen in Figure 2, the above triad gave rise to the following triples:

- ***isRespectedBy (****contributor****, Paltform regulations****) ;***
- ***isParticipedby (****contributor****, sharing****) ;***
- ***isRegulatedBy (****sharing****, Paltform regulations****).***

Thirty triples in all were deduced from twelve triads. The connexity of some triads gave in some cases identical triples. For example: *Subject-Community-Division of labour* and *Community-Division of labour-Activi*ty producing identical triple, ***isCreatedBy*** (*Community*, *Division of labour*). After identifying and eliminating redundant triples, the following results were obtained:

- ***isUsedBy****(sociocultural Knowledge, Contributor);* ***isMemberOf*** *(Contibutor, community of Contributors);*
- ***isRegulatedBy (****Sharing****, Paltform regulations****);***
- ***isAllowedBy*** *(Contributions, Administrators);*
- ***isEditedBy*** *(Contibution,Contributor);*
- ***isConcernedBy*** *(Sharing, Contribution);*
- ***isMonitored*** *(Sharing, Administrators);*

 ISBN: 978-1-61208-443-5

- ***isRespectedBy*** *(Contributor, Paltform regulations);*
- ***isParticipedby*** *(Contributor, sharing).*

Subjects and objects of these triples model fundamental concepts and predicates representing relationships between them. Table 1 shows mapping carried out between poles of Engeström's model and upper-level concepts of our ontology.

TABLE I. MAPPING BETWEEN ENGESTRÖM'S MODEL POLES AND ONTOSHARE UPPER- LEVEL CONCEPTS

Engeström model poles	Ontoshare Upper-level concepts
Object	Sharing
Subject	Contributor
Rules	Platform regulations
Community	Community of Contributors
Division of Labour	Contributions
Tools	Sociocultural Knowledge

- ***Community of Contributors****:* set of contributors;
- ***Sharing:*** information co-construction activity ;
- ***Contributor:*** active Internet user;
- ***Platform Regulations****:* rules defined for guiding sharing activity and contributors;
- ***Contributions:*** shared information for describing any *Sociocultural Knowledge* on the platform. In others words, this concept models each piece of information edited (shared) by any contributor into the knowledge construction in platform;
- ***Sociocultural Knowledge:*** This concept models information about society and culture. According to [21], sociocultural knowledge concept, concerns all forms of human knowledge: objects that compound the real world, facts and events. This concept is complex. As a result, sociocultural ontology named OntoSOC proposed in our previous work [9] was reused. This vocabulary helped us to semantically circumscribe the content of sociocultural knowledge. To reuse it, *Sociocultural Knowledge* concept was used as a "bridge" concept.

C. Concepts and Relationships

Different concepts obtained in Table 1 represent fundamental classes of Ontoshare. There are seven upper-level concepts in all. Figure 3 illustrates these concepts and relationships between them. In fact, classes alone are not enough to define ontology; we need to define also relations between them and attributes to characterise classes. These two notions add semantics to ontologies. In our study, use cases were required in identifying the following relationships:

- ***isUsedBy****(Sociocultural Knowledge, Contributor) ;*
- ***MemberOf*** *(Contibutor, Community of Contributors) ;*
- ***isRegulatedBy*** *(Community of Contributors, Plateform-Regulations) ;*
- ***isAllowedBy*** *(Contribution, Administrators) ;*
- ***isEditedBy*** *(Contibution, Contributor) ;*
- ***isConcernedBy*** *(Sharing, Contribution) ;*
- ***isMonitored*** *(Sharing, Administrators).*

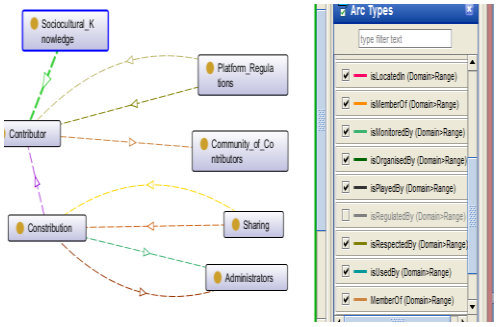

Figure 3. Ontoshare concepts and relationships

D. Hierarchy of classes

Out of seven concepts, three have variant depth, going from one to level six. There are many approaches to define hierarchy of classes: top-down, bottom-up and hybrid approach. The use of HAT not only enabled us to generate fundamental concepts but also marked the beginning of top-down method. Thereafter, to better define hierarchy, we intended to "*think up*" before making specifications. This is a top-down development process that begins with definition of the most general concepts in the domain and continues with sub-concepts specialisation. Figure 4 shows an overview of possible articulation between various generality levels of Ontoshare.

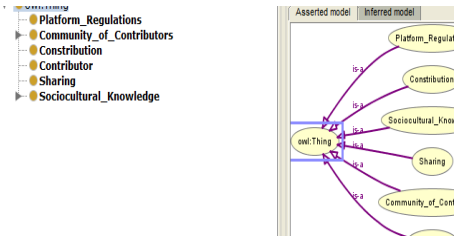

Figure 4. Ontoshare hierarchy extract

For example, *community* of *contributors* consisting of *Administrators*, *Community members* and *not-Community members* gives the overview of those who can share information on the platform. *Administrators* are particular members. In addition to their contributions, they have to work on monitoring edited contents by implementing and setting platform regulations. It clearly appears from this that, hierarchy of *sociocultural knowledge* provides information on different views (task, tools, community, actors, regulations, etc.) of data that each contributor can share through *sharing* concept.

 ISBN: 978-1-61208-443-5

E. *Ontoshare alignment*

For semantic web, alignment is a solution to the interoperability issue. It helps not to recreate those that exist but only to improve them. Ontoshare is an inter-domain vocabulary. It reuses some concepts of related ontologies (FOAF, Schema.org, DBpedia, and wai). Figure 5 presents alignment (manually) done between related vocabularies and some corresponding Ontoshare concepts.

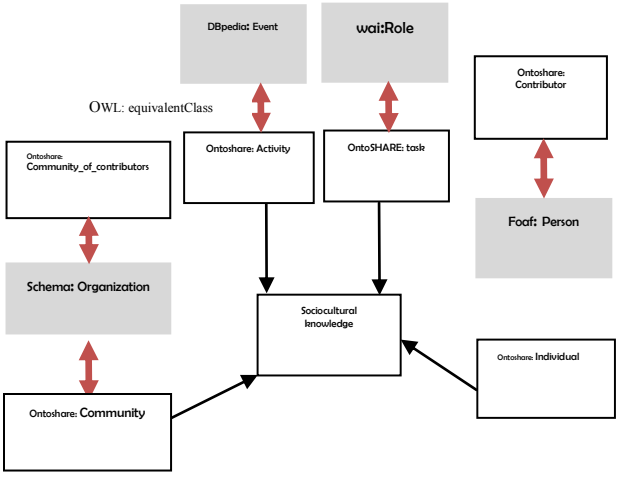

Figure 5. Ontoshare alignment

In recent years, the concept of Linked Open Data (LOD), and the so-called Web of Linked Data, has attracted tremendous attention from both the academic and real application world. The idea is, if we start to publish machine-readable data, such as RDF documents on the Web, and somehow make all these documents connected to each other, then we will be creating a Linked Data Web that can be processed by machines. Alignment technic participates to build the LOD project at two levels:

- Using *OWL:equivalentClass* property to connect Ontoshare concepts to standard vocabularies concepts and share semantics between these concepts. It is the case of *ontoshare:task* which is semantically equivalent to *wai:role and ontoshare:contributor* which is semantically equivalent to *foaf:person*;
- Using *is-a* property to add news concepts to LOD through standard vocabularies (FOAF, Schema.org, DBpedia, and wai). In our case, we added to LOD, a local concept as « *tontine* » which is the sub-concept of *ontoshare:community* class representing the local social network.

IV. ONTOSHARE POPULATING AND VALIDATION

For editing, we used "Protégé 5.0", an ontology development tool [22]. Ontoshare populating was done with data related to some activities (cultural events, sport events, and religious events). Three of them have been used previously for simulating HAT.

Protégé offers a number of reasoning engines and SPARQL endpoint in its standard distribution. A reasoner checks for consistency of description of class, subsumption between classes, taxonomy of class names (classification) and finds classes that match known instances. The performance of the proposed ontology has been evaluated at the following levels: classification, consistency checking using a reasoner, and competence question checking by SPARQL queries. According to classification checking, we tried to identify by classify function; if instances are automatically classified in a defined class. we did it for all concepts. For consistency checking, we aimed to verify, if there is any class which could never have an instance due to its definition. The competence question checking allowed by SPARQL endpoint enables to verify, if ontoshare can answer a competency question that guided its design. Thus, we focused on various activities organised in a specific locality to evaluate consistency of the Knowledge Base, and check for infinite query (eventually). The following query enabled to extract the relevant information on organised activities, resource used and tasks realised by subjects for any given community from our Knowledge Base.

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX OntoShare: <http://maroua-univ/ns/ OntoShare #>
SELECT  ?Communuties ?Activité ?task ?person ?tools
WHERE { ?task OntoShare:isUsedBy ?tools
OPTIONAL { ?Activité OntoShare:isRealizeBy ?task }
OPTIONAL { ?task OntoShare:isPlayedBy ?person }
OPTIONAL { ?task OntoShare:isCreatedBy ?Communuties }
} ORDER BY  ?Communuties
```

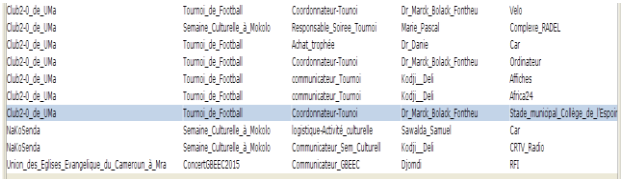

Figure 6. SPARQL query overview

The result of the SPARQL query is presented in Figure 6. It illustrates how the content of our KMS has been enriched by Internet users by editing some pieces of information related to sociocultural activities.

We would like to point out that, access features to ICT such as "*divide by access*", "*divide by use*", "*divide by decision*" in Cameroon must be improved, even if, sharing knowledge and discovering on that platform will not only be done by Cameroonians, but all Internet users worldwide. In fact, according to [23] "*digital divide*", statistics are very low. As a result, Cameroon would be the seventieth country over fifty four in Africa with about one million of Internet users. This number represents only 5% of its population. The same survey, recorded that the Internet penetration rate would be around 0.01 %.

 ISBN: 978-1-61208-443-5

V. CONCLUSION AND FUTURE WORK

In this paper, we presented Ontoshare modeling approach. Ontoshare is online activity ontology in the case of sociocultural knowledge sharing. To get there, we used Engeström's Human Activity Theory (HAT). To fix the interoperability problem, we have established mapping between Ontoshare concepts and those related. We populated our ontology with use cases and applied some SPARQL tests. Certainly, our use cases are far to be representative, but, their data helped us to eliminate or explain some inconsistencies and demonstrate how our "collective memory" could be co-constructed. It should be noted that redundant triples elimination process was done empirically. We have no guarantee of reaching minimal coverage. Nevertheless, reduction rate is considerable, about 60%. In addition, due to the fact that sociocultural knowledge is complex, we reused OntoSOC to circumscribe data type to be shared and co-constructed into our platform through sociocultural knowledge concept.

In perspective, we will focus on domain ontology and design the platform's architecture.

ACKNOWLEDGMENT

The authors would like to thank the University of Maroua which partially funded this work. Special thanks to LANI (Laboratoire d'Analyse Numérique et d'Informatique) of the University of Gaston Berger, Saint-Louis, in Senegal to our research position in the Laboratory.

REFERENCES

[1] UNESCO, "Toward the knowledge societies", world report of UNESCO, editions UNESCO, 2005.

[2] B. Z. Deli, "Western culture imperialism and the future of African culture: Challenges and Perspectives" GSSA, Maroua, Cameroon, 2008, pp15-20.

[3] http://www.enssib.fr/bibliotheque-numerique/notices/1948-de-la-fracture-numerique-a-la-fracture-cognitive-pour-une-nouvelle-approche-de-la-societe-de-l-information, [retrieved: 10, 2015].

[4] T. R. Gruber, "Toward Principles for the Design of Ontologies used for Knowledge Sharing", International Journal of Human Computer Studies, Special issue: the Role of Formal Ontology in the Information Technology, Vol. 43, No. 5, , 1995, pp.907-928.

[5] G. Kaladzavi, P. F. Diallo, Kolyang, and M. LO, "The role of ontologies in valuing African sociocultural knowledge on the Web ", 2nd Computer Science Reasearch Conference, Yaoundé, Cameroon, 2015, In press.

[6] https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/activity-theory, [retrieved: 10, 2015].

[7] P. F. Diallo, S. M. Ndiaye, and M. LO, "Study of sociocultural ontology", In the First International Conference on Social Economics and Informatics, Spain, October 2011, pp. 43-53.

[8] P. F. Diallo, O. Corby, M. LO, I. Mirbel, and S. M. Ndiaye, "Sociocultural Ontology: Upper-level and Domain Ontologies" in Actes JFO, Hammamet, Tunisia, november 2014, pp. 15-27.

[9] L. S. Vygotsky, "Mind in society: the development of higher psychological processes, "Cambridge: Harvard University Press.

[10] G. Kaladzavi, P. F. Diallo, Kolyang, and M. LO, "OntoSOC: Sociocultural knowledge ontology", International Journal of Web & Semantic Technology, vol 6, No 2, April 2015, DOI: 10.5121/ijwest.2015.6201.

[11] https://en.wikipedia.org/wiki/Main_Page, [retrieved: 09, 2015].

[12] http://mappings.dbpedia.org/server/ontology/classes/, [retrieved: 10, 2015].

[13] A. Gangemi, A. Prisco, M. T. Sagry, G. Steve, and D. Tiscornia,"Some Ontological Tools to Support Legal Regulatory Compliance, with a Case Study", In OTM Confederated International Workshops, 2003, pp. 1-2.

[14] http://mip-ms.cnam.fr/servlet/com.univ.collaboratif.utils.LectureFichiergw?ID_FICHIER=1295877017861, [retrieved: 09, 2015].

[15] L. S. Vygotsky, "Mind in society: the development of higher psychological processes." Cambridge: Harvard University Press.

[16] Y. Engeström, "Activity theory as a framework for analyzing and redesigning work", Ergonomics, 2000, pp. 960-974

[17] K. Kuutti, "Activity Theory as a potential framework for human-computer interaction research", Cambridge: MIT Press, 1995, pp. 17-44, [retrieved: 10, 2015].

[18] S. Blomkvis, "The User as a personality. Using Personas as a tool for design", Theoretical perspectives in Human-Computer Interaction, IPLab, KTH, September 3, 2002.

[19] A. Giboin, "From Individual to Collective Personas Modeling Realistic Groups and Com munities of Users (and not Only Realistic Individual Users) ", The Fourth International Conference on Advances in Computer-Human Interaction, 2011. pp. 32-35.

[20] J. P. Haton and M. C. Haton, "Knowledge Representation and Reasonings", in Artificial Intelligence, 15e ed. Paris, France: PUF, pp 39-67, 1989.

[21] cde-saare.de, [retrieved: 09, 2015].

[22] protege.stanford.edu/, [retrieved: 10, 2015].

[23] http://www.journaldunet.com/web-tech/chiffres-internet/cameroun/pays-cmr, [retrieved: 17, 2015].

 ISBN: 978-1-61208-443-5