

User profile Ontology for the Personalization approach

Youssef EL ALLIOUI
USMBA University
B.P 1796 – Fes
Fez - Morocco

Omar EL BEQQALI
USMBA University
B.P 1796 – Fes
Fez - Morocco

ABSTRACT

The personalization is to facilitate the expression of the user's need and enable him to obtain relevant information in these information system accesses. The relevance of information is defined by a set of customizable preferences specific to each user or user community. Data describing users are often grouped as profiles. The content of a user's profile varies according to approaches and applications.

Existing approaches solves partially problems related to personalization, but it lacks a model giving a complete overview of all aspects of taking into account the preferences of users. In this paper we propose a generic model of profile that includes all aspects of personalization. The proposed model will be the basis of building an ontology called O'Profil able to store all this information, personalize the content and to instantiate the user profile.

Keywords

Personalization, Ontology, User profile Context, Preference, Profile model

1. INTRODUCTION

This paper rests on a central problematic: the personalization of information access. Indeed, face to considerable growth of data, the heterogeneity of the roles and needs and to the rapid development of mobile systems, it becomes important to propose a personalized system able to provide user with relevant and adapted information to his need and use daily. This system must take into account the different characteristics of the user and all contextual situations that influence his behavior during his interaction with information system.

The system of personalized access to information on which we work is based on the principle of profile management. A generic model of profile is proposed in order to run all facets of personalization. This model is able to collect user's information, his context of use and represents any need for personalization once modeled.

The generic model of profile according to which the personalization system is articulated based mainly on three elements: the profile, the context and user's preferences. The profiles are containers of knowledge about the user. The context defines a set of parameters that characterize the environment of the system. The User's preferences represent the expectations of the user.

On the other hand, ontology is a good candidate for representing knowledge about users, to have a shared understanding between people or software agents of the sense of terms and their relations, a controlled vocabulary and common and so on, which implies a formal definition of

concepts. In addition to their contribution in terms of reusability, of modularity and knowledge sharing, ontology is used to define a precise vocabulary, which is the basis of communication between different users. So we will use them for the representation of user characteristics and their activities, modeling of neurological domain and the management of personalization.

This work is part of TENEMO¹ project. It is a multidisciplinary project that aims to design and develop a solution of expertise and decision support for diagnosis in the field of neurology and its emergencies.

Our contributions to this project focus on the mechanisms and strategies for personalized access to heterogeneous and distributed neurological data in a mobile environment. It is also three main contributions: 1) modeling knowledge about user, 2) modeling Neurological knowledge [28] and 3) Propose a strategy of personalization and adaptation of neurological responses on the basis of user profile. In this paper we will settle for the first contribution.

This paper is organized as follows: Section (2) summarizes the outline of the art state of personalization, the management of user profiles, the works oriented to design a new generation of personalization systems based on context and the use of ontology to define and build profiles. In section (3) we present the models of profile, context and preferences that we propose and relationships between them. In section (4) we describe the two main operators of profile management; the instantiation and the matching. Section (5) is a comparison between a generic model of UML and ontology. In Section (6), we will show the process of building the O'Profil, while following the method of Noy.

2. RELATED WORK

The current information systems provide access to large number of heterogeneous and distributed sources. As the sources and the volume of data are increasing, the user is faced with information overload in which it is difficult to distinguish relevant information from secondary information. In addition, the evaluation of a query is getting usually regardless of the context and / or needs of user who issued it. The same query, submitted by two different users, produces the same results even if these users have different expectations. For example, for the query "Apple", some users are interested in finding the results dealing with computers branded "Apple", while others are interested in finding the results dealing with the hidden meaning of the fruit "Apple". Studies in [1] show that the fault of such systems lies partly in the fact that they consider the information needs of the user is

¹<http://homepages.laas.fr/khalil/wiki/index.php?n=PROJETS.TENEMO>

completely represented by his query and do not take into account the user in the chain of access to information.

To address the problems of information overload and to be able to discriminate users based on their specific needs, some systems propose personalization techniques based on user profile. The notion of profile models interests centers of user and his main preferences in terms of filtering and quality of information, in terms of access to systems and delivery of results, in terms of the spatio-temporal context in which it is located and the device it uses. All this knowledge is defined in totality or in part directly by the user or derived by algorithms learning from past actions of the user.

The classification, organization and structuring of profile data is a key element of personalization. Various studies have addressed this aspect without the cover as a whole. For example, P3P [2] standard for secure the profiles, allows to define classes that distinguish between demographic attributes, the professional attributes and the attributes of behavior. In [3], the authors propose a profile model for users of a digital library, consisting of five categories of information: personal data, collected data, delivery data, behavior data and the safety data. These attempts of structure are laudable but insufficient to cover the personalization domain. Moreover, they simply categorize profile information, but are extensible with difficulty.

The personalization of information is a major challenge for the computer industry [4]. Whether in the context of enterprise information systems, electronic commerce, access to knowledge or even leisure, the relevance of the provided information, its intelligibility and its adaptation to the uses and preferences of customers are key factors for success or rejection of these systems. Personalization of information was specifically addressed in three technological domains: the Human-Computer Interaction (HCI) [5], the Information Retrieval (IR) [6], and Databases (DB) [7] [8].

Whatever the domain of technology, personalization of information can be exploited according to two modes of management: by recommendation or by interrogation. Recommender systems [9] exploit user profiles or user communities for disseminate offers targeted on the interests centers and preferences of the latter. This procedure is also called "push mode". User feedback is very important to refine their profiles and increase the efficiency of the system. Personalization by interrogation [7] is to adapt the evaluation of a query in relation to the characteristics and preferences of the user who sent it. In this context, the system reacts to a specific request of the user in enriching his request so as to make it more accurate [7], in choosing data sources based on the quality requirements of the user [10] or personalizing the display of results [11]. This procedure is also called "pull mode". The work described in this paper fits into this context.

The implementation of personalization systems is getting mainly in two phases:

- 1) Representation of the user profile: There are several techniques of representation of profile and of these interests' centers. Indeed, they can be represented as vectors of weighted terms [12], or semantically weighted according to the concepts of a general ontology [13][14], or according to matrices of concepts [15].
- 2) Integration of user profile in the process of personalizing the data: the profile must be exploit in one of the main phases of the evaluation of the

request: reformulation [12], calculating the score of relevance of information [16] or presentation of research results [14][17].

On the other hand, several ontologies of specific domains have been designed with the object is to establish a conceptual research for finding ways of simplifying navigation through the semantic categories of the hierarchy used. This requires the use of a knowledge representation language [18] which specifies the content of web pages according to taxonomy of concepts. We give among the search engines that exploit ontologies, "Google" and "Yahoo". All of these techniques already mentioned are not dedicated to recognize users while remaining independent of the context of the user.

For this, the works are oriented to design of a new generation of personalization systems based on context, aimed at delivering information relevant and appropriate to the context of the user who issued the request. In [19], a contextual personalization system is defined as follows: "Combine search technologies and knowledge about the query and user context into a single framework in order to provide the most appropriate answer for a user's information needs". The works of [20] placed the notion of context and situation, without distinction, where the context describes the intentions of the user on the one hand, and research environment on the other. There are many definitions of context discussed in the literature that differ primarily by elements of the context. A multi-dimensional definition of context [21] adds to the situation notion the characteristics related of a part in the temporal aspect of information needs and the type of research asked the other. Although the authors do not converge to the same definition of context, however there are common dimensions descriptive such as cognitive environment, the need for information, etc. Into a contextual personalization system, stress is laid on using a user's model previously constructed called "profile" [22]. The first systems are designed based on collaborative filtering. These systems such as Grouplens [23] exploit the collaborative profile linked to a group of users sharing common interests and persistent and returns the user to the information meeting the criteria of the profile of the group to which it belongs. On the other, personal agents of data personalization are then developed as the system Letizia [24] which is a personal assistant to using the web, able to propose the information without explicit request by the user. Other systems [13] [15] explore different techniques for learning of user profile that is subsequently used in one of the phase of the personalization process.

We focused in this paper to present an extension of approaches to implicit construction of user profile previously developed in the literature with the object of build a profile. Our approach is based on the using of ontology. This profile will be exploited in a personalization process of neurological responses.

3. MULTIDIMENSIONAL MODEL FOR REPRESENTATION OF PROFILE, CONTEXT AND PREFERENCES

To solve the problem of ambiguity between the concepts; profile, context and preference appeared in some personalization systems, we make a clear distinction between them. User profile is a container of information describing the user. The context represents the data describing the environment of interaction between a user and a system. Preference is an expression that prioritizes the importance of information in a profile or context.

In our approach we will study the concepts of profile, context and preferences independently before integrating them through the semantic relations that clearly show their combined use. Indeed, we will propose the models of profile, context and preferences (Fig 1) which are both hierarchical, generics and evolutionary. The term "generic" does not mean an exhaustive set of attributes, but rather a list of high-level concepts which can be specialized, refined and instantiated independently of the system wherein they can be integrated. To achieve this goal, we draw on works already done in this context (see the related work above) and the proposed models should be:

- Open: Easily extensible to other kinds of knowledge and other types of preferences.
- Flexible: Able to acquire the main categories of knowledge used in the current personalization systems.
- Multi-faceted: to be analyzed from different angles (dimensions, attributes, complex and / or simple, etc.).
- Evolutionary: Can be evolving over time, i.e. they should allow any changes or updates as soon as necessary.
- Independent of any data management system and of any technology,
- Specialization, generalization and instantiation of these models should be easy.

In summary, the main objective of the model we propose is to clarify the concepts; profile, context and preferences through a classification of knowledge that constitute them, and establish a framework to define and manipulate personalized access systems to masses of data. This framework takes the form of a generic model instantiable and adaptable to each type of application. In the following sections, we describe the three concepts of this user model, namely profile, context and preferences.

3.1 Modeling profile

Fig 1 shows a representation of the model profile by objects. This representation can be structured in the form of hierarchy of classes which are:

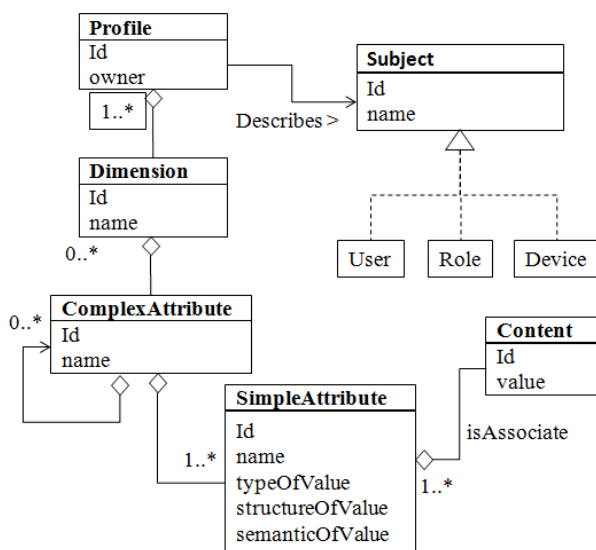


Fig. 1: Profile model

Subject is the object that is described by a profile. This object can be the user, device, role, or others. The objects for which the profiles are defined can be organized into a set of classes (User, Device, etc.). So the class "Subject" will be the superclass of these classes.

Profile describes the criteria that characterize a given subject as the user. It consists of several dimensions.

Dimension describes information characterizing a profile; this information can be grouped and structured in the form of a set of open dimensions. Each dimension consists of a set of attributes. These attributes can be simple or complex.

Complex Attribute: Information that is grouped and structured at dimensions will be in turn organized into complex attributes. For example, the dimension "Personal Data" may be characterized by the complex attribute "Address". A complex attribute is a set of simple attributes that are semantically related (e.g. address consists of street number, street name, postal code etc.).

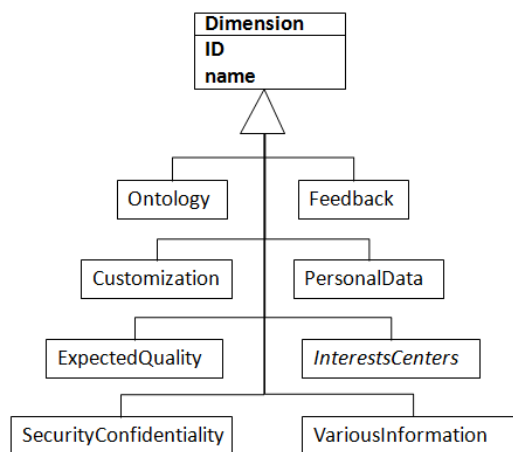


Fig. 2: Set of dimensions characterizing user profile

Simple Attribute: This class represents an elementary attribute or non-decomposable which a content is assigned (name, size, color, etc.). Each simple attribute is characterized by a name, a type of values, a structure of values (single value, set of values, interval, vector of weighted terms, etc.) and a semantic of value. The semantic of value can define the concept represented by the attribute when its meaning is related to the context of use. Each simple attribute can be associated with one or more content (s).

Content: describes the possible values that are associated with each simple attribute.

Pursuer the classification of Amato in [3], we propose a set of dimensions that can accommodate the different information characterizing a profile. We need to distinguish between the following dimensions (Fig 2):

3.1.1 Personal data

Personal data (Fig 3) will have a dual focus, firstly the management of user identification (name, surname, etc.). And secondly, they allow categorizing the user based on characteristics such as the authentication attributes (login, password, IP address, etc.) demographic factors (age, gender, first language, place of birth, etc.), the personal and/or professional user contacts and other information such as blood group, bank account number, etc.

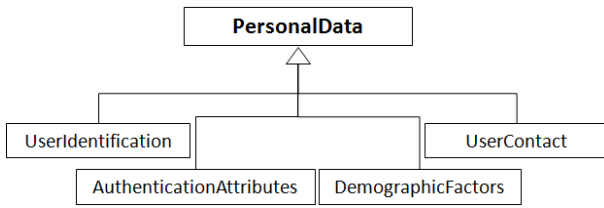


Fig. 3: Dimension of personal data

3.1.2 Interest center

The interest center is the main dimension of the user profile. This dimension includes all attributes that concern the content objects (target information). It can define the expertise domain, the perimeter exploration and the qualification level of the user in a particular domain that the content he is interested. The interest center can be represented in different ways (Fig 4): A set of keywords, keywords structured as ontology, a set of predicates or as a list of examples of objects seen by the user with actions that he made it. And we can use multiple representations in the same profile, e.g. ontology that complete a set of predicates.

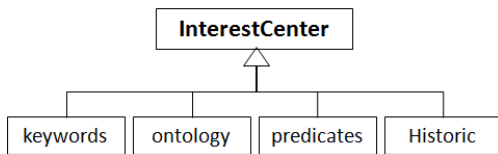


Fig. 4: Dimension of interest center

3.1.3 The domain ontology

The domain ontology completes the definition of interest center with additional specification of semantics of certain terms or certain operators related to the field of user activity (Fig 5).

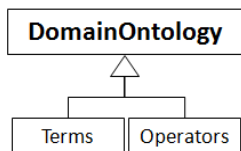


Fig. 5: Dimension of domain ontology

3.1.4 Expected quality of results delivered

Quality is one of key factor of personalization; it can express extrinsic preferences as information origin, accuracy, freshness, its period of validity, the time required to produce it or the credibility of its source. The attributes of this dimension express the expected quality by the user. The expected quality can be refined into several sub-dimensions (Fig 6) that characterize the data (e.g. freshness and accuracy of data), data sources (e.g. level of confidence in the source, reliability, etc.) and the production process (e.g. to measure process performance, their ability to produce all relevant results, etc.).

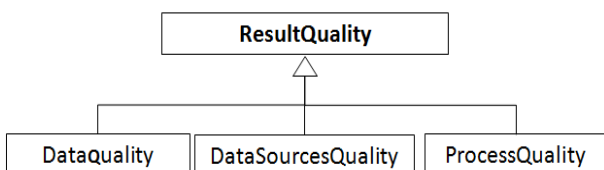


Fig. 6: Dimension of quality

3.1.5 Customization

It concerns matters associated with the presentation of results depending on the platform, the nature and volume of data delivered, the user's visual preferences. It also concerns the execution of the request, the moment of execution (push mode or pull mode), how notify the results (delayed, immediate, etc.) and the amount of results you wish to receive (Top k, first calculated, etc.).

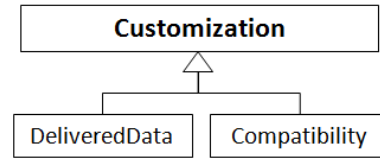


Fig. 7: Dimension of customization

It is worth noting that the delivery data depends on the context of interaction and the results delivered are not presented in the same way on a PDA, a laptop or a mobile phone.

3.1.6 Security and confidentiality

It may relate to data that are queried or modified, the calculated information, the user's requests or the other dimensions of the profile. We distinguish three types of security (Fig 8):

- *Access rights security*: it is through the management of access permissions to profile information.
- *Data security*: it concerns data (content) delivered to the user. It is expressed by indicating the level of security results (on a predefined scale) or through expressions in an existing formalism.
- *Process security*: it expresses the will of the user to hide sensitive treatment that it performs.

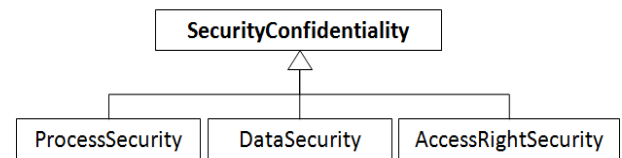


Fig. 8: Dimension of security and confidentiality

3.1.7 Feedback

This dimension is an alternative collection of information on user preferences to build a profile in a dynamic way. At first, the user provides a set of keywords describing his preferences to allow the system to build an initial profile. When a new data arrive, the system uses the user profile to select data that could be interesting for him. With the feedback, In addition to relevant data, the user also indicates those that are not. This information is being provided explicitly or implicitly by user.

3.1.8 Various information

As its name indicates, this characteristic includes the different attributes that cannot be included in any of the dimensions that we cited, for example, the bandwidth assigned to profile manager or disconnection of profile for some queries.

In this section, we have put forward a multidimensional model of the profile. A profile contains characteristics that describe a user, an information system or a source. These characteristics may depend on the context in which they are exploited. The following section presents the context model.

3.2 Modeling context

Drawing on the different works of context modeling, we propose a context model (Fig 9) represented by two main dimensions spatiotemporal dimension (location and time) and device (software and / or hardware).

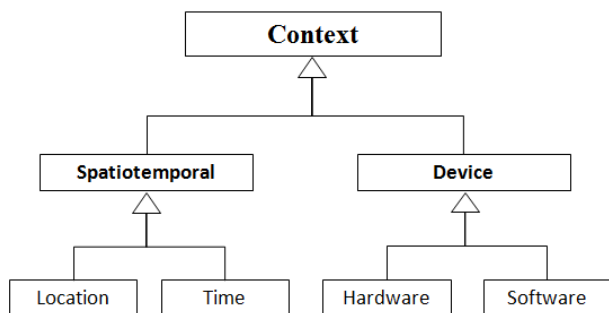


Fig. 9: Context Model

The dimension "Time" allows locating the user in the time and thus indicates the moment of the interaction between the user and the system.

The dimension "Location" means the place where the user is located. It may contain the precise coordinates of the user or the wording of location in which he is located (work, home, etc.).

The dimension "Device" allows us to know the characteristics of the equipment used, and consider in order making the results exploitable on the platform of the user. The device can be represented by characteristics of the hardware (screen size, memory capacity, etc.) and/or software (operating system, office software, etc.).

The choice of these two dimensions (spatiotemporal and device) is justified by the expected goal of the personalization of information. So these two dimensions must be able to meet the needs and preferences of the user in all its contextual situations.

3.3 Modeling Preferences

One of personalization definitions is a set of individual preferences represented by pairs (attribute, value), an organization of criteria or semantic rules specific to each user or user community. Preference is an expression for prioritize the importance of information about profile and/or context. The fact that user preferences are part of their profile and even the description of the user and their preferences can be changing depending on the context in which it evolves gives peculiarity to preference model to be instantiable independently of profile model and / or context model. The preference model is then a range of wishes or favors that can be used to prioritize the importance of knowledge defined in the profiles and contexts.

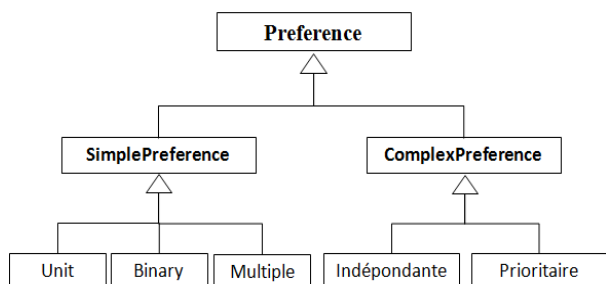


Fig. 10: Preferences Model

In the preference model that Kostadinov proposed [25], we identify a typology of two preferences: simple and complex preferences (Fig 10):

Simple preference is defined by a single concept characterizing one or more objects. Depending on the number of objects characterized, simple preferences can be **unit** (they apply to a single object), **binary** (comparison or favoritism between two elements) or **multiple** (they apply to a set of objects).

Complex preference is a combined expression based of two or more intermediate preferences (simple preferences and / or composed preference). This combination can be done **independently** or **priority**. In the first case, intermediate preferences have the same degree of importance. but in the second case, the order in which intermediate preferences are applied is important.

To better illustrate this categorization of preferences, we suggest the following list of examples (Fig 11):

Simple preference	Complex preference
Unit	Independently
Film.kind = 'action' The user is only interested to action films	((action>>drama) ⊗ (English>>French)) The user is interested at action movies, and English movies with the same degree of importance.
Binary	Priority
action >> drama The user prefers to see action movies more then the drama movies	((action>>drama) & (English>>French)) The user likes to watch action movies and if they are in English will be better for him
Multiple	
LOWEST (price) The user searches the lowest price	

Fig. 11: Typology of preferences

3.4 Relationship between Profile, context and preferences

In the previous sections, we studied the three concepts (profile, context and preferences) separately in order to better study them. While the description of the user, the representation of context and the expressions of preferences is done only by identifying the relationships between them. Indeed, a complete representation of the user and context does not just contain elements of description, but also preferences. In addition, the definition of a user profile may depend on the context in which it operates. Fig 12 shows the main links between the profile, context and preferences.

A user profile consists of a set of elements profile and preferences. These elements can be dimensions, complex attributes, simple attributes, etc. Preferences can be applied to all of these elements, provided that the combined elements must be the same type (e.g. two values of the same attribute).

As is the case with the profile, the context consists of a set of elements of context and preferences. And thus, preferences can be applied to all of these elements to adapt the results or responses issued depending on the context of the user.

In addition to being defined on elements of the profile and context, preferences can be defined on other preferences to form more complex expressions. This justifies the reflexive association on preferences.

The contents of the profile may depend on the context. This dependence is expressed by the association "Contextualized According To". And by inheritance, an element of the profile or preference can be contextualized.

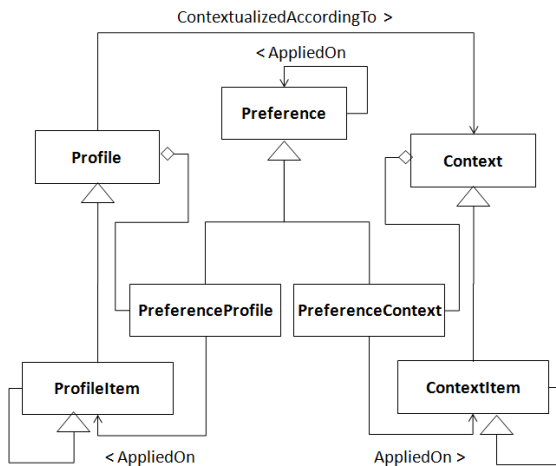


Fig. 12: Relationship between Profile, Context and Preference

4. FROM ONTOLOGY TO GENERIC MODEL

Ontology, such as generic models of UML, they model the universe of discourse by means of hierarchical classes by subsumption and associated to properties. However, ontology present five differences compared with the generic models of UML. These differences represent the basis of the contribution that ontology may provide in the problematic of modeling:

Semantics: Ontology originates from the world of artificial intelligence to accurately capture knowledge. But the generic models of UML originate from the world of software engineering to simplify the description of the software. Indeed, ontologies are used in an Open World Assumption (OWA), which enable discovery of new knowledge by the automated reasoning. In contrast, UML models are used in a Closed World Assumption (CWA) in which information defined prescriptively for construction.

Goal: In ontology, the emphasis is generally on concepts and relations of a domain, independently of each specific application and any system in which it is likely to be used. In contrast, a generic model of UML emphasized the operational side of a particular computer system in order to deal with a given application need.

Atomicity: Unlike a generic model of UML, where each concept only makes sense in the context of the model in which it is defined, in ontology each concept is individually identified and is a basic unit of knowledge. We can therefore extract from ontology only the concepts (classes and properties) relevant to our target application. We can also respecting their semantics (e.g. subsumption), organize them fairly different from their organization in ontology, the reference to the ontological concept to define precisely the meaning of the referencing entity.

Consensuality: In ontology of domain, concepts are represented as consensus for a community. The use of ontology in a design phase (by developers) or operational phase (by user), allows a natural and ergonomic access to

domain information when designers and users within the community covered by the ontology. Similarly, the semantic integration of all systems based on the same ontology can be made easily if the references to the ontology are explained.

Canonicity: Unlike generic models that use a language (UML) Semi-formal semantics (metamodel) non-redundant to describe domain information, ontologies use language (OWL) formal semantics (description logic) which allowed to define, in more primitive concepts, defined concepts which therefore provide alternative access to the same information, which allows to extend the usability access to information.

The next step is to transform the classes and attributes of the profile model, context model and preferences model to concepts and slots of an ontology that we called O'Profil.

5. BUILDING O'PROFIL ONTOLOGY

Ontology is good candidate for representing knowledge about users. In addition to their contribution in terms of reusability, of modularity and knowledge sharing, ontologies are used to define a precise vocabulary, which is the basis of communication between different users. We can, therefore, use them for the representation of user characteristics, their activities, management of personalization and scientific and technical documents, representation of their content and their manipulation, etc. Representation of the contents can manipulate it to make the search by specialization, by similarity, by analogy, etc.

Using these tools of information processing allows effective sharing of knowledge in the organization (enterprise, research team, etc.).

In this section, on the basis of generic models that we proposed in section 3, we will show the process of building O'Profil, while following the Noy's method.

5.1 Driven Ontology design - Noy's approach

There are many methods for ontology engineering. However, there is no agreement on the principles that should guide the ontological modeling. It appears that these approaches are heterogeneous and none is yet ready. In our case, we opted for Noy's method proposed by Stanford University because it comprises steps clear, simple and easy to understand. Add to this the fact that the tool with which we will build the ontology in this case PROTÉGÉ is developed by the same university.

The Noy's method process can be summarized in the following steps:

- 1) Determine the domain and scope of the ontology
- 2) Consider reusing existing ontologies
- 3) Enumerate important terms in the ontology
- 4) Define the classes and the class hierarchy
- 5) Define the properties of classes - slots
- 6) Define the facets of the slots
- 7) Create instances

We note that we can't separate these steps because it is a nonlinear process, several back and forth were made during the development and incremental refinement of the ontology to takes into account the errors detected in each step. We will detail these steps later.

5.2 Implementation of O'Profil

Most ontology was created using the OWL (Web Ontology Language). OWL is designed for use by applications that need to process the content of information instead of just presenting information to humans. OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF, and RDF Schema (RDF-S) by providing additional vocabulary along with a formal semantics. OWL has three increasingly-expressive sublanguages: OWL Lite, OWL DL, and OWL Full.s

Proposals for the design of ontologies on the theoretical and practical view can be found in [26]. Ontology O'Profil was designed following the best practices of the Semantic Web [27] and implemented by using OWL-DL. Briefly, we follow the Noy's methodology described in [26]. The ontology design process can be summarized in the following steps:

Step 1. Determine the domain and scope of the ontology :

This step identifies the goal of ontology, its domain, its users and its scope. The result of this step can be summarized in the specifications points (with natural language) as follows:

The domain of ontology is going to cover the modeling of the user profile (users and technical resources).

The purpose of the using our ontology is to describe the neurological service users (Profile, context and preferences) to capitalized them in the personalization process of returned neurological responses .

The ontology must respond to user requests, taking into account the semantics of those requests.

The ontology will be used by the system of decision support itself.

The ontology is maintained by experts in the field.

Step 2. Consider reusing existing ontologies :

Our research of existing ontologies in the field of user profile management did not lead to satisfactory results. So we will build our ontology fully by conducting interviews with experts in the field and with the help of our internet research, especially the articles and works with a research subject is modeling of user profiles.

Step 3. Enumerate important terms in the ontology :

Our study on the field of user profile management allowed us to identify a list of important terms. We can't list all of them, we just a few examples: Profile, dimensions, attributes, interest center, History, personal data, Quality, Delivered Data, Security, etc.

Step 4. Define the classes and the class hierarchy :

Classes provide an abstraction mechanism for grouping resources with similar characteristics. Like RDF classes, every OWL class (owl:class) is associated with a set of individuals called the extension of class. Fig 8 shows a summary of the model of O'Profil (classes and class hierarchy).

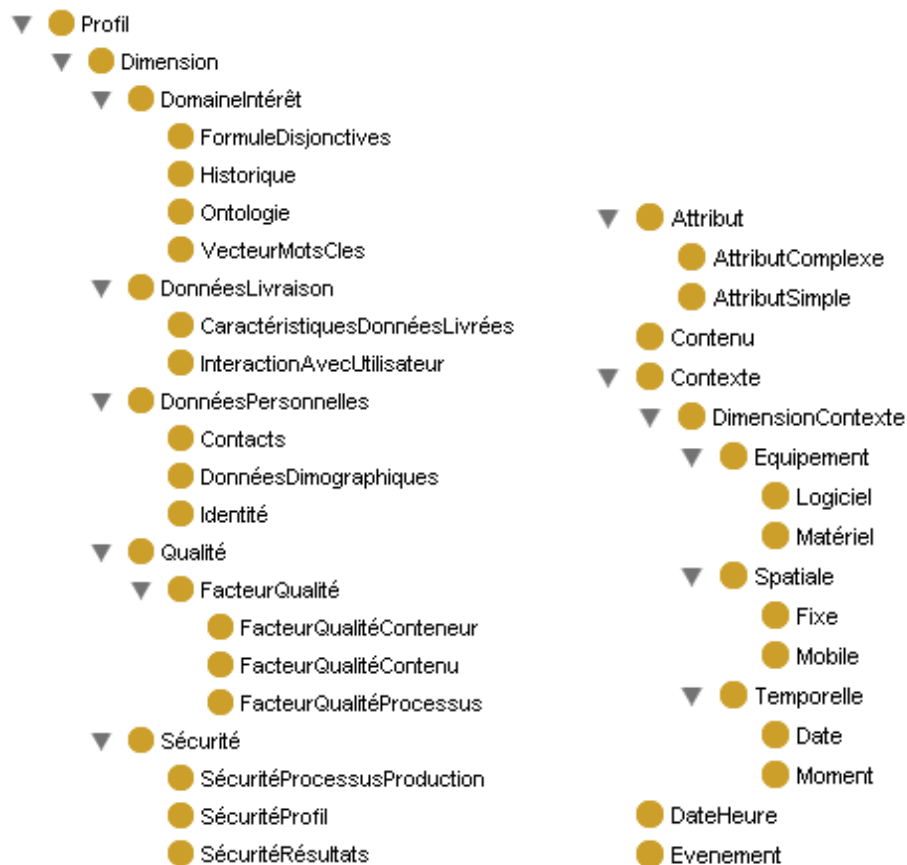


Fig. 13: Hierarchy of O'Profil concepts by PROTÉGÉ

Step 5. Define the properties of classes – slots :

A property can be considered as an attribute or as a binary relation between two classes. OWL distinguishes two kinds of properties:

- owl:DatatypeProperty: The properties of data type (class properties) have a range of data value, and binds individuals to data values. For example: property "name" for class "Identity Dimension".

```
<owl:onProperty>
  <owl:DatatypeProperty rdf:about="#nom"/>
</owl:onProperty>
```

- owl:ObjectProperty: Object properties (relationship) have a range of values of class individuals. For example the property "Contextualized according to" applies to an object of class called "Profile element" and takes its values in the class called "Context Element".

```
<owl:ObjectProperty rdf:ID="Défini_sur">
  <rdfs:domain
rdf:resource="#ElementProfil"/>
  <rdfs:range
rdf:resource="#ElémentContexte"/>
</owl:ObjectProperty>
```

Step 6. Create instances :

The instances are all objects of classes. These objects are very numerous; Fig 9 shows the creation of the instance "Identité_1" of Miss Salma Bellakhdar².

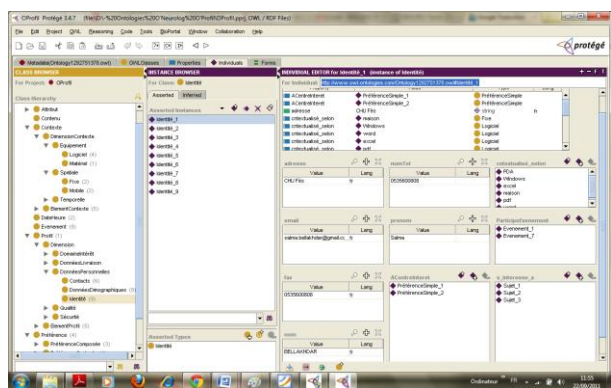


Fig. 14: Creating instance by PROTÉGÉ

The code generated by PROTÉGÉ is:

```
...
<Identité rdf:ID="Identité_1">
  <ACentreInteret
rdf:resource="#Pr#233;f#233;renceSimple_1"/>
  <ACentreInteret
rdf:resource="#Pr#233;f#233;renceSimple_2"/>
  <adresse xml:lang="fr">CHU
F#232;s</adresse>
  <contextualisé_selon rdf:resource="#excel"/>
  <contextualisé_selon
rdf:resource="#maison"/>
  <contextualisé_selon rdf:resource="#PDA"/>
  <contextualisé_selon rdf:resource="#pdf"/>
  <contextualisé_selon
rdf:resource="#Windows"/>
```

² Neurologist - neurological department of teaching hospital Hassan II - Fez, Morocco

```
<contextualisé_selon rdf:resource="#word"/>
<email xml:lang="fr"
>salma.bellakhdar@gmail.com</email>
  <fax xml:lang="fr">0535600808</fax>
  <nom xml:lang="fr">BELLAKHDAR</nom>
  <numTel
rdf:datatype="xsd:string">0535600808</numTel>
  <ParticipeEvenement
rdf:resource="#Evenement_1"/>
  <ParticipeEvenement
rdf:resource="#Evenement_7"/>
  <prenom
rdf:datatype="xsd:string">Salma</prenom>
  <s_interesse_a rdf:resource="#Sujet_1"/>
  <s_interesse_a rdf:resource="#Sujet_2"/>
  <s_interesse_a rdf:resource="#Sujet_3"/>
</Identité>
...
```

5.3 O'Profil validation

To test the O'Profil, we use the Pellet reasoner available directly from PROTÉGÉ. Pellet is a complete and capable OWL-DL reasoner with acceptable to very good performance, extensive middleware, and a number of unique features [29]. Pellet is written in Java and is open source under a very liberal license. It is used in a number of projects, from pure research to industrial settings.

Pellet is the first sound and complete OWL-DL reasoner with extensive support for reasoning with individuals (including nominal support and conjunctive query), user-defined datatypes, and debugging support for ontologies [29]. It implements several extensions to OWL DL including combination formalism for OWL-DL ontologies, a non-monotonic operator, and preliminary support for OWL/Rule hybrid reasoning. It has proven to be a reliable tool for working with OWL-DL ontologies and experimenting with OWL extensions.

In this section we describe three tests provided by Pellet: consistency checking, classification test and queries test.

5.3.1 Consistency test

Consistency checking provided by Pellet is made based on the class description, which ensures that ontology does not contain any contradictory facts, i.e. check for each class, there must be at least one individual member of this class. A class is considered inconsistent if it cannot have any instance. The result of this test, as shown in Fig 14 indicates that all classes are consistent.

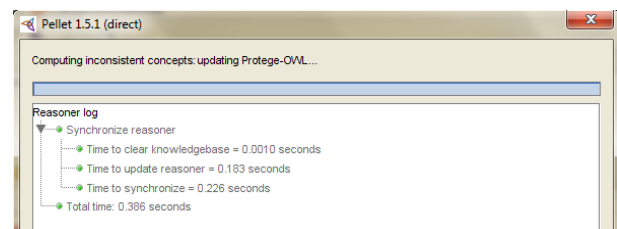


Fig. 15: Consistency test results of O'Profil ontology

5.3.2 Classification test

The classification test can check whether a class is a subclass of another class or not. It computes the subclass relations between every named class to create the complete class hierarchy. The class hierarchy can be used to answer queries such as getting all or only the direct subclasses of a class.

When this test is invoked, the consistency test is first performed for all classes of the ontology, because inconsistent classes cannot be classified correctly. Once the classification test is performed on the class hierarchy containing the logical expressions, it is possible for the classifier to infer a new hierarchy "inferred ontology class hierarchy" which is a hierarchy where classes are classified according to the relationship superclass / subclasses.

In this case the classification test shows that no suggestion has been produced by the reasoner "Pellet" and that "Asserted hierarchy" and "Inferred hierarchy" are identical, indicating the validity of classification of our ontology.

5.3.3 Test queries

PROTÉGÉ (specially the queries Tab) allows querying your project and locating all instances that match the criteria you specify. You can create a simple query, or combine multiple

criteria to restrict or expand your results. Queries are not part of your knowledge base, but are a way to identify the instances in your project, based on class and slot properties. Creating queries in PROTÉGÉ is guided by a graphical user interface (Fig 15). We give some examples of query that we have created:

- What are the users interested in subject called "S"?
- What are the interest's centers of a user "X"?
- Find documents written by user "X"
- What are the users who participate in the event "X"?

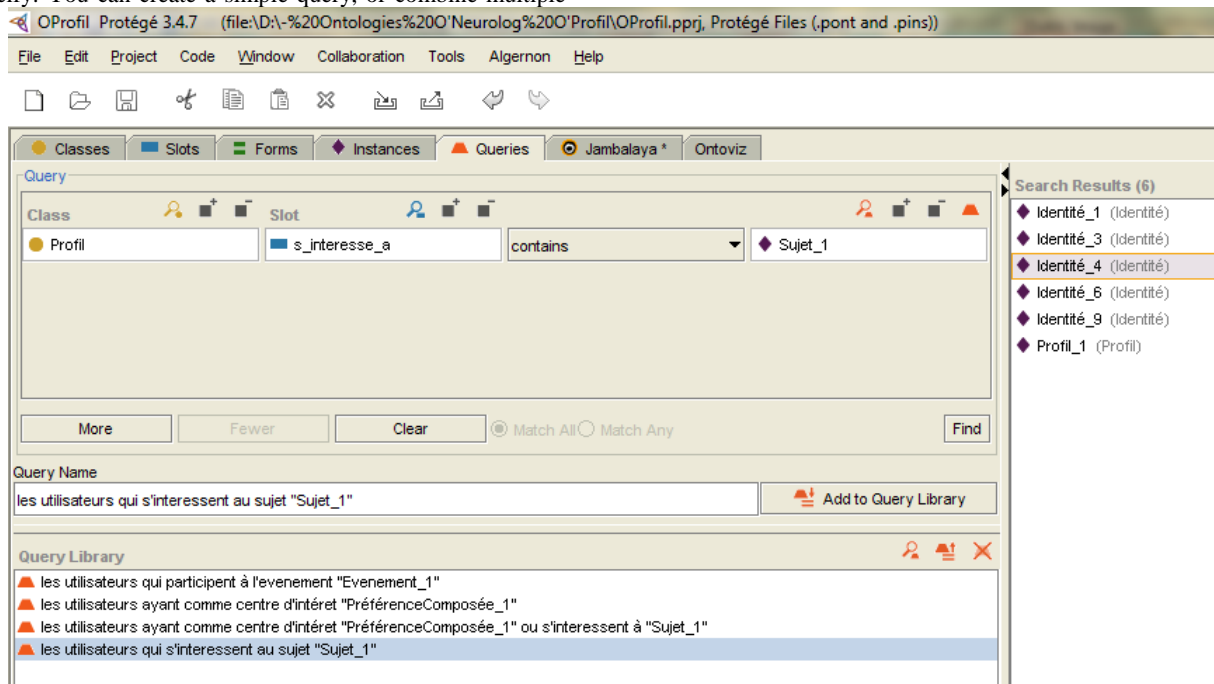


Fig. 16: Queries Tab screenshot

5.4 Discussion

We have created ontology for representing, under one entity, all characteristics of a user. There is a need to filtering and selecting relevant information for user, but also to personalize his system use. The major advantage of O'Profil ontology is its coherence and its consistency shown by tests on it. Thus, it is ready for future use in any personalization system that is based on the user profile.

6. CONCLUSION

The development of a generic profile model is the first step towards the construction of personalization systems native able to search and retrieve quality information according to user preferences. In this paper, we have defined a generic profile model that is used to structure all necessary information for the description of user preferences. A profile includes eight dimensions comprising all necessary parameters for personalization. We have shown what information types are contained in these dimensions and how the model can be instantiated to define user-specific profiles.

On the other hand, our approach should not stop at the generic profile model. Indeed, we defined the conceptual structure that we have chosen for representation and manipulation of this model. We have clarified the organization of model knowledge and show that ontologies meet our expectations. We therefore addressed more formally the expected features of our approach based on ontologies. This is the building of the User ontology called O'Profil. The purpose of this ontology is to allow a user community to share knowledge of their specific domain and capture the characteristics of users in language that allows concepts and their relations to be interpreted by computers.

In this paper we have set a terminology, are the fundamental concepts that will provide a basis for the construction of users ontology, O'Profil. While this ontology can be represented, as a single entity, all characteristics of user, we believe it is necessary for filtering and selection of information relevant to user, and also to personalize its use of system.

7. REFERENCES

- [1] Budzik J. Hammond K. "Users interactions with everyday applications as context for just-in-time

- information access". Proceedings of the 5th international conference on intelligent user interfaces, p. 41-51, 2000.
- [2] Cranor L., Dobbs B., Egelman S., Hogben G., Humphrey J., Langheinrich M., Marchiori M., Presler-Marshall M., Reagle J., Schunter M., The Platform for Privacy Preferences 1.1 (P3P1.1) Specification. W3C Working Draft 1 July (2005), <http://www.w3.org/TR/2005/WD-P3P11-20050701/>
- [3] Amato G., Staraccia U., "User profile modelling and applications to digital libraries", Proceedings of the 3rd European Conference on Research and advanced technology for digital libraries, p. 184-187, 1999.
- [4] Bouzeghoub M., Kostadinov D., L'art et définition d'un modèle flexible de profils, Actes de la 2ème conférence en personnalisation de l'information : aperçu de l'état de recherche d'informations et applications CORIA'2005, Grenoble, France, 2005
- [5] Shearin S., Lieberman H., Intelligent Profiling by Example, In Proceedings of the 2001 International Conference on Intelligent User Interfaces, Santa Fe, USA, p. 145-151, 2001.
- [6] Zemirli N., Tamine-Lechani L., Boughanem M., Présentation et évaluation d'un modèle d'accès personnalisé à l'information basé sur les diagrammes d'influence, Dans les Actes du XXVème congrès INFORSID, Perros-Guirec, France, p. 89-104, 2007.
- [7] Koutrika G., Ioannidis Y. E., Personalized Queries under a Generalized Preference Model, In Proceedings of the 21st International Conference on Data Engineering (ICDE 2005), Tokyo, Japan, p. 841-852, 2005.
- [8] Rocacher D., Liétard L., Préférences et quantités dans le cadre de l'interrogation flexible: sur la prise en compte d'expressions quantifiées, Dans les actes des 22e Journées Bases de Données Avancées (BDA), Lille, France, 2006.
- [9] Nguyen A.-T., Denos N., Berrut C., Exploitation des données « disponibles à froid » pour améliorer le démarrage à froid dans les systèmes de filtrage d'information, Dans les Actes du XXIVème Congrès INFORSID, Hammamet, Tunisie, p. 81-95, 2006.
- [10] Naumann, F., Freytag J.C., Spiliopoulou M., Quality Driven Source Selection Using Data Envelope Analysis, In Proceedings of the MIT Conference on Information Quality (IQ'98), Cambridge, USA, p. 137-152, 1998.
- [11] Pletschner A., Gauch S., Ontology Based Personalized Search, In Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'99), Chicago, USA, p. 391-398, 1999.
- [12] Sieg A., Mobasher B., Lytinen S., Burke R., « Using Concept Hierarchies to Enhance User Queries in Web-based Information Retrieval », Artificial Intelligence and Applications(AIA), 2004.
- [13] Sieg A., Mobasher B., Burke R., Prabu G., Lytinen S., « Representing user information context with ontologies », uahci05, 2005.
- [14] Challam et al., 2007: Challam V., Gauch S., Chandramouli A., « Contextual Search Using Ontology-Based User Profiles », Proceedings of RIAO 2007, Pittsburgh USA, 30 may - 1 june, 2007.
- [15] Liu F., Yu C., Meng W., « Personalized Web Search For Improving Retrieval Effectiveness », IEEE Transactions on Knowledge and Data Engineering, vol. 16, n° 1, p. 28-40, 2004.
- [16] Tamine L., Boughanem M., Zemirli W., « Exploiting Multi-Evidence from Multiple User's Interests to Personalizing Information Retrieval », IEEE International Conference on Digital Information Management(ICDIM 2007), 2007a.
- [17] Ma Z., Pant G., Sheng, « Interest-based personalized search », ACM Transactions on Information Systems, 2007.
- [18] Lassila O., Swick R., « Resource Description Framework (RDF) Model and Syntax Specification », August, 1998.
- [19] Allan et al., 2002: Allan J., al., « Challenges in information retrieval and language modelling », Workshop held at the center for intelligent information retrieval, Septembre, 2002.
- [20] Saracevic T., « The stratified model of information retrieval interaction : extension and applications », Proceedings of the 60th annual meeting of the American Society for Information Science, Medford, NJ, p. 313-327, 1997.
- [21] Fuhr N., « information retrieval : introduction and survey, post-graduate course on information retrieval, university of Duisburg-Essen, Germany », 2000.
- [22] Tamine L., Calabretto S., Recherche d'information contextuelle et Web, Ouvrage intitulé recherche d'information sur le web, édition hermes, à paraître, IRIT, France, 2007b.
- [23] Miller B., Konstan J., Matz D., Herlocker J., Gordan L., Riedl A., « GroupLens : applying collaborative filtering Usenet news, Communications of ACM », March, 1997.
- [24] Lieberman H., « Autonomous interface agents », ACM Conference on Human-Computer Interface, p. 67-74, March, 1997.
- [25] Dimitre Kostadinov, Thèse : Personnalisation de l'information : une approche de gestion de profils et de reformulation de requêtes. Version 1 - 22 Sep 2008, UNIVERSITE DE VERSAILLES SAINT-QUENTIN-EN-YVELINES
- [26] Noy NF, McGuinness DL. Ontology development 101: a guide to creating your first ontology; 2001. Report No.: Stanford Knowledge Systems Laboratory Technical Report KSL-01-05, Stanford Medical Informatics Technical Report SMI-2001-0880.
- [27] Smith B, Ceusters W, Klagges B, Kohler J, Kumar A, Lomax J, et al. Relations in biomedical ontologies. Genome Biol 2005;6(5):R46.
- [28] El Alloui Y., El Beqqali O. : O'Neurolog - Building an Ontology for Neurology in Mobile Environment. IJCSNS - International Journal of Computer Science and Network Security, 2010. Vol. 10 No. 9 pp. 188-197.
- [29] E. Sirin, B. Parsia, B. Grau, A. Kalyanpur, Y. Katz. Pellet: A practical OWL-DL reasoner. Web Semantics: Science, Services and Agents on the World Wide Web In Software Engineering and the Semantic Web, Vol. 5, No. 2. (June 2007), pp. 51-53. doi:10.1016/j.websem.2007.03.004 Key: citeulike:2615678.