# GUMO – the General User Model Ontology

Dominik Heckmann, Tim Schwartz, Boris Brandherm, Michael Schmitz,
Margeritta von Wilamowitz-Moellendorff

Saarland University, Saarbrücken, Germany
{dominik,schwartz,brandherm,schmitz}@cs.uni.sb.de

**Abstract.** We introduce the general user model ontology GUMO for the uniform interpretation of distributed user models in intelligent semantic web enriched environments. We discuss design decisions, show the relation to the user model markup language USERML and present the integration of ubiquitous applications with the `u2m.org` user model service.

**Keywords:** user model ontology, semantic web, ubiquitous user model service, intelligent environments, user model markup language

## 1 Motivation and Introduction

A commonly accepted top level ontology for user models could be of great importance for the user modeling research community. This ontology should be represented in a modern semantic web language like OWL and thus via internet be available for all user-adaptive systems at the same time. The major advantage would be the simplification for exchanging user model data between different user-adaptive systems. The current problem of syntactical and structural differences between existing user modeling systems could be overcome with a commonly accepted ontology, specialized for user modeling tasks. We are suggesting a user model ontology rather than a user model*ing* ontology, which would additionally include the inference techniques or knowledge about the research area in general. We are collecting the user's dimensions that are modeled within user-adaptive systems like the *user's heart beat*, the *user's age*, the *user's current position*, the *user's birthplace* or the *user's ability to swim*. Furthermore, the modeling of the user's interests and preferences like *reading poems*, *playing adventure games* or *drinking certain French Bordeaux wines* is analyzed.

### 1.1 Choosing OWL as Ontology Language for GUMO

Ontologies provide a shared and common understanding of a domain that can be communicated between people and heterogeneous and widely spread application systems, as pointed out in [3]. Since ontologies have been developed and investigated in artificial intelligence to facilitate knowledge sharing and reuse, they should form the central point of interest for the task of exchanging user models. XML is designed to serve for weakly structured data as an interchange format. The user model markup language USERML is defined as an XML application, see [4]. However, XML is purely syntactic and structural in nature. The RDF standard has been proposed as a data model for representing meta data by [8]. Nonetheless, the web ontology language OWL has more facilities for expressing semantics, [10], and it has a greater machine interpretability than XML and RDF. It adds more vocabulary for describing properties and classes. OWL can be used to explicitly represent the meaning of terms in vocabularies and the relationships between those terms. OWL is a revision of the DAML+OIL web ontology language in which we presented the first user model ontology[1]. To summarize, OWL is our choice for the representation of user model terms and their interrelationships.

### 1.2 GUMO is Influenced by UserML, SUMO and UbisWorld

The main conceptual idea in USERML's approach of SITUATIONALSTATEMENTS, see [5], is the division of user model dimensions into the three parts: `auxiliary`, `predicate` and `range` as shown below.

$$\texttt{subject}\left\{\textit{UserModelDimension}\right\}\texttt{object}$$
$$\Downarrow$$
$$\texttt{subject}\left\{\texttt{auxiliary},\ \texttt{predicate},\ \texttt{range}\right\}\texttt{object}$$

If one wants to say *something about the user's interest in football*, one could divide this into the `auxiliary`=*hasInterest*, the `predicate`=*football* and the `range`=*low-medium-high*. If one wants to express something like *knowledge about symphonies*, one could divide this into the `auxiliary`=*hasKnowledge*, the `predicate`=*symphonies* and the `range`=*poor-average-good-excellent*. GUMO is designed according to this USERML approach. Approximately 1000 groups of `auxiliaries`, `predicates` and `ranges` have so far been identified and inserted into the ontology[2]. However, it turned out that actually everything can be a `predicate` for the `auxiliary` *hasInterest* or *hasKnowledge*, what leads to a problem if one does not work modularized. The suggested solution is to identify basic user model dimensions on the one hand while leaving the more general world knowledge open for already existing other ontologies on the other hand. Candidates are the general suggested upper merged ontology SUMO, see [9] and the UBISWORLD ontology[3] to model intelligent environments. This insight leads to a modular approach which forms a key feature of GUMO. Nevertheless, since no top level user model ontology has been proposed so far, it is done so in this paper. Which groups of user dimensions can be identified? In [6] and [7] rough classifications for such categories can be found.

## 2 Defining GUMO Auxiliaries and Predicates

Identified user model `auxiliaries` are *hasKnowledge*, *hasInterest*, *hasBelieve*, *hasPlan*, *hasProperty*, *hasGoal*, *hasPlan*, *hasRegularity* and *hasLocation*. This listing is not intended to be complete, but it is a start with which a lot of user facts can be realized. We restrict ourself in this paper to present user model `predicates` that fit to the `auxiliary`: *hasProperty*, the so called *BasicUserDimensions*.

[1] First user model ontology in DAML: `http://www.daml.org/ontologies/444`
[2] GUMO homepage: `http://www.gumo.org`
[3] UbisWorld homepage: `http://www.ubisworld.org`

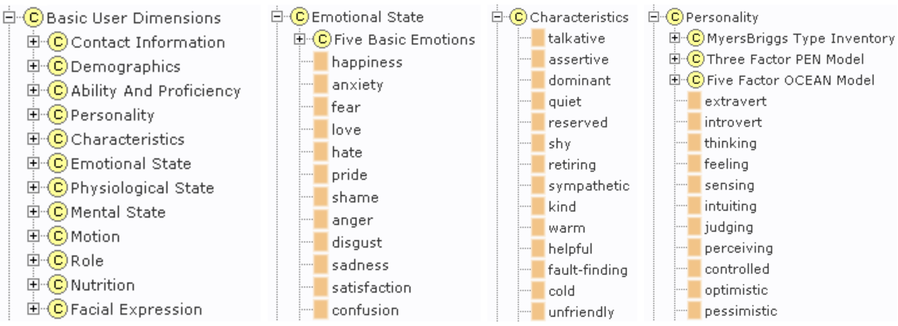


**Fig. 1.** Some *BasicUserDimensions*: Emotional States, Characteristics and Personality. The complete ontology can be inspected with a foldable tree browser at `www.gumo.org`

The following listing presents the concept *PhysiologicalState* defined as `owl:Class`. It is defined as a subclass of *BasicUserDimensions*. A class defines a group of individuals that belong together because they share some properties. Classes can be organized in a specialization hierarchy using `rdfs:subClassOf`.

```
<owl:Class rdf:ID="PhysiologicalState.700016">
  <rdfs:label> Physiological State </rdfs:label>
  <rdfs:subClassOf rdf:resource="#BasicUserDimensions.700002" />
  <gumo:identifier> 700016 </gumo:identifier>
  <gumo:lexicon>state of body or bodily functions</gumo:lexicon>
  <gumo:privacy> high.640033 </gumo:privacy>
  <gumo:website rdf:resource="&GUMO;concept=700016" />
 </owl:Class>
```

Every concept has a unique `rdf:ID`, that can be resolved into a complete URI. Since the handling of these URIs could become very unhandy, a short identification number was introduced, the so called `gumo:identifier`. The identification number has the advantage of freeing the textual part in the `rdf:ID` from the need of being semantically unique. Apart from solving the problem of conceptual ambiguity, this number facilitates the work within relational databases, which is important for the implementation. The lexical entry `gumo:lexicon` is defined as *the state of the body or bodily functions*, while it could also be realized through a link to an external lexicon like WORDNET. The attribute `gumo:privacy` defines the default privacy status for this class of user dimensions. It can be overridden in the concrete SITUATIONALSTATEMENT. The attribute `gumo:website` points towards a web site, that has its purpose in presenting this ontology concept, to a human reader. The abbreviation `&GUMO;` is a shortcut for the complete URL to the GUMO ontology in the semantic web. The next listing defines the dimension *Happiness* as an `rdf:Description`. The attribute `gumo:expiry` provides a default value for the average expiry which carries the qualitative time span of how long the statement is expected to be valid. In most cases when user model dimensions are measured, one has a rough idea about the expected expiry. For instance, emotional states hold normally no longer than 15 minutes, however personality traits won't change within months. Since this qualitative time span is dependent from every user model dimension, it should be defined within GUMO.

```
<rdf:Description rdf:ID="Happiness.800616">
  <rdfs:label> Happiness </rdfs:label>
  <rdf:type rdf:resource="#EmotionalState.700014" />
  <rdf:type rdf:resource="#FiveBasicEmotions.700015" />
  <gumo:expiry> 15 minutes </gumo:expiry>
  <gumo:image rdf:resource="http://u2m.org/img/happiness.gif" />
</rdf:Description>
```

Another important point that is shown here is the ability of multiple-inheritance in OWL. In detail, *happiness* is defined as `rdf:type` of the class *EmotionalState* and *FiveBasicEmotions*. Thus OWL allows to construct complex, graph-like hierarchies of user model concepts, which is especially important for ontology integration. Some examples of rough expiry-classifications are:

– physiologicalState.heartbeat - can change within seconds
– mentalState.timePressure - can change within minutes
– characteristics.inventive - can change within months
– personality.introvert - can change within years
– demographics.birthplace - can't normally change at all

The idea behind `gumo:expiry` is that if no new actual value is available on the user model server after a while, one can still work with old values, probably combined with reduced confidence values. The presented new GUMO vocabulary for the user model ontology language consists of `gumo:identifier`, `gumo:expiry`, `gumo:image`, `gumo:privacy`, `gumo:website`, `gumo:image` and `gumo:lexicon`. To support the distributed construction and refinement of GUMO, we developed a specialized online editor to introduce new concepts, to add their definitions and to transform the information automatically into the required semantic web language.

## 3 UserModelService and Ubiquitous Applications Using GUMO

A user model *service* manages information about users and contributes additional benefit compared to a user model *server*. The `u2m.org` user model service is an application-independent server with a distributed approach for accessing and storing user information, the possibility to exchange and understand data between different applications, as well as adding privacy and transparency to the statements about the user. The key feature is that the semantics for all concepts is mapped to the GUMO ontology. Applications can retrieve or add information to the server by simple HTTP requests, alternatively, by an USERML web service. A basic request looks like:

```
http://www.u2m.org/UbisWorld/UserModelService.php?
subject=Peter&auxiliary=hasProperty&predicate=Happiness
```

The ALARMMANAGER, see [1], is a notification service for instrumented environments that adapts the presentation of announcements to the user's state of arousal and the user's location. Both are retrieved from the GUMO enabled `u2m.org` user model service. The location is derived from a POSITIONINGSERVICE application, see [2]. This service runs on the user's PDA and uses infrared beacons and active RFID tags that are installed in the environment to estimate the location of the user which is then send via WiFi to the user model service.

## 4 Summary

We presented the general user model ontology GUMO, discussed why we have used the ontology language OWL to define it and showed by integrating ubiquitous user-adaptive applications that the interaction of the ontology with the exchange language `UserML` and the `u2m.org` user model service is promising.


### Acknowledgements

This research is being supported by the German Ministry of Education and Research (BMB+F) under grant 524-40001-01 IW C03, the project SPECTER; by the German Science Foundation (DFG) in its Collaborative Research Center on Resource-Adaptive Cognitive Processes, SFB 378, Project EM 4, BAIR, and its Transfer Unit on Cognitive Technologies for Real-Life Applications, TFB 53, Project TB 2, RENA. Special thanks go to Vadim Chepegin and Lora Aroyo for fruitful discussions about GUMO.


## References


1. Brandherm, B., Schmitz, M.: Presentation of a modular framework for interpretation of sensor data with dynamic Bayesian networks on mobile devices. In: LWA 2004, Lernen Wissensentdeckung Adaptivität, Berlin, Germany (2004) 9–10
2. Brandherm, B., Schwartz, T.: Geo referenced dynamic bayesian networks for user positioning on mobile systems. In: Proceedings of the International Workshop on Location- and Context-Awareness (LoCA), Munich, Germany (2005)
3. Fensel, D.: Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce. Springer-Verlag Berlin Heidelberg (2001)
4. Heckmann, D., Krüger, A.: A user modeling markup language (UserML) for ubiquitous computing. Lecture Notes in Artificial Intelligence **2702** (2003) 393–397
5. Heckmann, D.: Introducing situational statements as an integrating data structure for user modeling, context-awareness and resource-adaptive computing. In: ABIS2003, Karlsruhe, Germany (2003) 283–286
6. Jameson, A.: Systems That Adapt to Their Users: An Integrative Perspective. Habil, Saarbrücken, Germany (2001)
7. Kobsa, A.: Generic user modeling systems. User Modelling and User-Adapted Interaction Journal **11** (2001) 49–63
8. Ora Lassila, R.R.S.: Resource Description Framework (RDF) Model and Syntax Specification. W3C. (1999) W3C recommendation.
9. Pease, A., Niles, I., Li, J.: The suggested upper merged ontology: A large ontology for the semanticweb and its applications. In: AAAI-2002Workshop on Ontologies and the Semantic Web. Working Notes (2002)
10. McGuinness, D.L., van Harmelen, F.: OWL web ontology language overview. http://www.w3.org/TR/owl-features/ .W3C Recommendation (2003)