SIWAM: Using Social Data to Semantically Assess the Difficulties in Mountain Activities

Javier Rincón Borobia, Carlos Bobed, Angel Luis Garrido and Eduardo Mena
IIS Deparment, University of Zaragoza, Zaragoza, Spain

Keywords: Semantic Web, Information Extraction, Ontologies, Social Network.

Abstract: In the last few years, the amount of people moving to the mountains to do several activities such as hiking, climbing or mountaineering, is steadily increasing. Not surprisingly, this has come along with a raise in the amount of accidents, which are mainly due to the inexperience of the people, and the lack of information and proper planning. Although one could expect to find appropriate updated information about this issue on the Internet, most of the information related to mountain activities is stored in personal blogs, or in Web sites that are not exploiting the possibilities that the Semantic Web and the Social Web offer regarding content generation and information processing.

In this paper, we present SIWAM, a semantic framework oriented to share and evaluate the difficulties of mountain activities. It provides a thematic social network front-end to enable users to share their descriptions about their own experiences. Using text mining techniques on these descriptions, it extracts relevant facts about these experiences, which are used to evaluate the difficulty of the particular activity. The evaluation is done according to a well-established standard for evaluating the difficulty of mountain activities (MIDE), which is modeled in the system using ontologies.

1 INTRODUCTION

Mountain activities comprise a set of sports that are amongst the most practiced in the world. The amount of people practising them is increasing year by year all around the world (Global Industry Analyst, Inc., 2012; Jenkins, 2013). This steady increment of practitioners along with the fact that mountain activities are considered extreme sports make the security be an important and recurring matter of study. There are research areas (specially medical ones) (Chamarro and Fernández-Castro, 2009), and organizations, such as the International Mountaineering and Climbing Federation, which are specially concerned about mountain accidents (UIAA Mountaineering Commission, 2004). However, in spite of all the efforts, the amount of accidents is, not surprisingly, increasing as more and more people move to the mountains. From the yearly reports of the rescue groups in Spain (Ministerio del Interior, 2012), we can point out that the main problems regarding accidents are the inexperience of the people, and the lack of information and proper planning.

While mountain guide books can be a source of information about the tracks and the environment where the activity is held, it seems pretty safe to assume that, as the main cause of accidents is lack of information, people are not using them to plan their activity (when even they plan it). Instead, in these times, people go to the Internet to look for information as fast as possible, regardless the possible security implications. In the case of mountain activities, this information is mainly available as descriptions of the different experiences in text format (e.g., blog entries).

However, trusting these on-line descriptions might be a double-edged sword: On the one hand, they might be really useful as they hold information about the activity; but, on the other hand, they are describing unique experiences whose setup and conditions might not be applicable to other situations. A climbing route might have many variations, shortcuts, forks, etc., within the same path. In a mountain environment, a slight detour might turn an easy track into a challenging one. Besides, the weather conditions also affect strongly to the mismatch between the descriptions and the track difficulties. As an example, in high mountain, the orientation is mainly guided by milestones that may be covered by snow during winter season.

Moreover, the level of expertise of the person describing the activity might lead to dangerous biases,

Rincón Borobia J., Bobed C., Garrido A. and Mena E..
SIWAM: Using Social Data to Semantically Assess the Difficulties in Mountain Activities.
DOI: 10.5220/0004812600410048
In *Proceedings of the 10th International Conference on Web Information Systems and Technologies* (WEBIST-2014), pages 41-48
ISBN: 978-989-758-024-6

as, for an experienced mountaineer, several actions might be considered too easy to be worthy of being described, although they may suppose a danger for an amateur practitioner. Therefore, there is a need for methods to support the correct evaluation of the different activities taking into account the expertise and physical conditions of the practitioner.

Currently, in the mountain Web sites, the main method to introduce information about an activity is form-oriented, with several predefined fields frequently including a textual one, where users can provide a more detailed description of their experiences. In fact, the information that can make a difference may be hidden in these textual descriptions (e.g., "the track is quite easy, but there is a fork that might lead to a …"). An improvement could be to include more detailed forms; however, the amount of details to be included might be overwhelming for the users, resulting in a decrease of user collaboration.

In this paper, we present SIWAM[1], a system that provides a semantic framework to improve users' information in the domain of mountain activities with the main goal of improving their security. It consists of three main modules: 1) a social Web front-end (RSAM), where users can share their experiences and provide information about their profiles; 2) a text-extraction module (MECMIDE) that is in charge of detecting and extracting the relevant facts from the activity descriptions provided by the users; and, finally, 3) an activity semantic evaluator that, given the relevant facts, classifies the different activities according to a well-defined mountain activities standard (MIDE), which is modeled in the system by an ontology network. The complete set of individual evaluations can be taken into account to provide a more accurate evaluation of the different activities depending on different aspects such as the weather conditions, the season, the user's profile, and so on.

The rest of the paper is as follows. In Section 2, we present the architecture of our system, detailing the main modules. In Section 3, we introduce the system's ontology and the MIDE standard (Roche, 2002), and how we have modeled the latter to be exploited by SIWAM. Section 4 and Section 5 detail the information extraction module (MECMIDE) and the evaluation module (VALMIDE), respectively. A complete example of how SIWAM works is presented in Section 6. We discuss some related work in Section 7. Finally, the conclusions and future work are drawn in Section 8.

[1]SIWAM stands for *Sistema de Información Web para Actividades de Montaña* (in Spanish, which means Mountain Activities Web Information System).

2 ARCHITECTURE OF THE SYSTEM

In this section, we present a general overview of SIWAM, describing its aim and general architecture. SIWAM is conceived to be a social site where users of all expertise share their experiences performing different mountain activities. Using text-mining and Semantic Web techniques, SIWAM is capable of processing this information to assess the difficulty of the different tracks and activities according to a well-established mountain activities standard. To do so, SIWAM exploits the knowledge stored in the System Ontology (see Figure 1). In this way, SIWAM offers much more precise and updated information about the different activities, having as a final objective to improve the mountaineers security (no matter their experience levels).

As it can be seen in Figure 1, SIWAM consists of three main modules:

- RSAM (Social Network for Mountain Activities:) This is the Web front-end of SIWAM, and supports the common features of a social site (e.g., user profiles, instant messaging, groups, contacts, etc.). Its functionality is specifically extended with features oriented to the management and sharing of information about mountain activities (e.g., including new ones, adding descriptions and experiences, sharing maps, GPS routes, etc.). SIWAM processes these shared descriptions and experiences in background to obtain the information about the safeness of a particular activity.
- MECMIDE (Concept Extraction Module:) It is the module in charge of processing each of the descriptions and extracting the relevant facts out of them. To do so, it uses Freeling (Carreras et al., 2004) to analyze the texts, looking for relevant patterns which might contain information. These patterns are mapped to different concepts and properties of the System Ontology, so MECMIDE outputs a set of axioms in the form of RDF[2] triples which are correctly aligned with this ontology.
- VALMIDE (Evaluator Module:) This module takes as input the facts that MECMIDE has extracted from the text, and, with the help of a Description Logics reasoner (Baader et al., 2003) (DL reasoner from now on), evaluates the difficulty of the activity according to the System Ontology. This ontology models the MIDE (Roche, 2002) standard to evaluate the difficulty of mountain activities regarding several criteria. The

[2]http://www.w3.org/TR/rdf-primer/

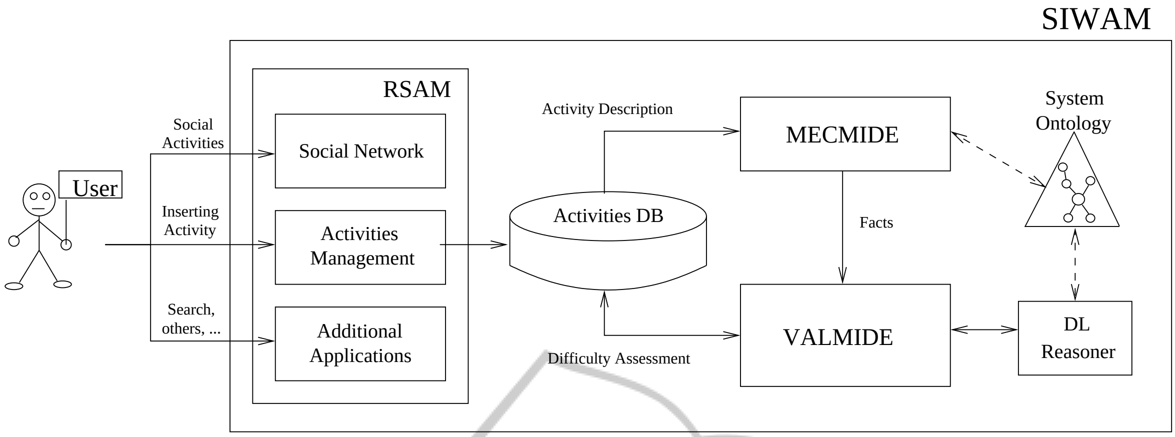

Figure 1: Architecture of SIWAM.

global evaluation of a particular activity is obtained combining the single evaluations of all the people that have done it. Considering all the descriptions makes the evaluation more robust against omitted information.

In the following, we will focus on the semantic part of SIWAM. First, we present how we have modeled the MIDE evaluation standard in the System Ontology. Then, we overview how MECMIDE extracts the information from the texts using the vocabulary defined in the used ontology. Finally, we explain how this knowledge is used by VALMIDE to assess the difficulty of a particular activity.

3 SYSTEM ONTOLOGY

The System Ontology stores all the information needed to perform the evaluation of the different activities. Following the directives given in NeON Methodology (Suárez-Figueroa, 2012), we have built an ontology network to make it possible to follow a modular development. In Figure 2, we can see the inner structure of this ontology:

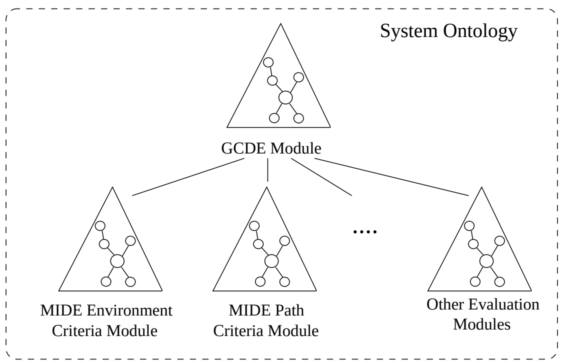

Figure 2: Inner structure of the ontology used by SIWAM.

- The Global Criteria Excursions Definition (GCDE) ontology module integrates the different modules that model different evaluation criteria, and stores information about how they must be used to evaluate an activity. In particular, for each of the activity types that SIWAM handles, GCDE stores the evaluation modules that are applicable to it and the methods to be used to do so. This enables the system to react automatically to the addition of new evaluation modules and methods.

  Currently, it only contains the description of one method of evaluation, which is the one applied to the MIDE evaluations (we will see it later in Section 5); however, it is interesting to specify its role in the system as it provides our system with a flexible method to add/remove evaluation modules.

- For each evaluation criteria to be used, there is an ontology module that models it in the system. In particular, we have selected the MIDE criteria for SIWAM, as it is a well-established standard for mountain activities evaluation.

The motivation for using ontologies to model the different evaluation criteria is two-fold: On the one hand, so far, there has not been not any initiative to model the domain of mountain activities and provide a shared vocabulary; on the other hand, using ontologies provides us with a logical framework to perform the activity evaluation in a more tractable and manageable way than using pure rule systems, as we will see in Section 5. Moreover, the use of an ontological model as a base for the text information extraction has been successfully applied in many works such as (Garrido et al., 2012; Vogrincic and Bosnic, 2011; Garrido et al., 2013; Kara et al., 2012).

Thus, in the rest of the section, we focus on how we have modeled the evaluation criteria knowledge in this ontology. To do so, we firstly overview the MIDE standard, and then we detail how we have captured its information in the appropriate ontology module.

3.1 MIDE Evaluation Standard

The MIDE standard (Roche, 2002) is focused on the evaluation of mountain tracks. To do so, MIDE proposes to tag a particular excursion (*experiences* using the SIWAM terminology) with two different kinds of information: 1) information about the track (name, type, accumulated height difference, etc.) and the conditions of the actual activity performing (hour, season, weather conditions, etc.); and 2) evaluation information, which provides difficulty values ranged in 1 – 5 for different aspects of the track, namely *Environment*, *Path*, *Journey*, and *Physical Effort*.

Most of the first kind of information can be captured using simple forms (e.g., times, height differences, distances, etc.). For the second type of information, MIDE provides guidelines to evaluate each of those values. While the Physical Effort can be obtained using different formulae, the rest of values (Environment, Path and Journey) are assessed using a set of criteria that might or might not be present in the track. Then, depending on the number of criteria the activity fulfils, it is assigned a difficulty value from 1 to 5 for that particular aspect.

Capturing these criteria using formularies is more complicated. They would require complex forms, which would be long and time-consuming to fill, a task that not all the users are prone to do. Fortunately, this is the kind of information that is usually comprised in the textual descriptions of the user's activities. Thus, the first step to perform an automatic evaluation of the track is to model these criteria to be used by SIWAM. In the following subsection, we present our approach to do this task.

3.2 Modeling MIDE with Ontologies

The development of the ontology modules that model the MIDE evaluation criteria was carried out in two well differentiated stages:

1. Modeling the mountain domain: The objective of this first stage was to obtain and organize the elements (concepts and properties) within the domain of the mountain activities. We did not aim at modeling the whole domain at once, but incrementally, taking each of the MIDE evaluation criteria as the competence questions (Grüninger and Fox, 1994) for each iteration. This way, we could assure that we had modeled all the elements needed to model the MIDE knowledge.

2. Modeling the criteria: In the case of MIDE, the evaluation is performed by checking whether a particular activity fulfils or not an specific (and complex) condition. From an ontological point of view, this is equivalent to say that the activity belongs to a particular type of activities defined by this condition. Thus, we used *defined concepts* to model them. In Description Logics (Baader et al., 2003), the underlying formalism of OWL[3], a defined concept provides a complete definition of its members, that is, it establishes *necessary and sufficient* conditions for an instance to belong to it. This kind of definitions enables DL reasoners to classify the instances according to them. In the following section, we will see how VALMIDE module takes advantage of this issue.

Moreover, we added also the information about how many criteria the activities must fulfil to be assigned each of the 1 – 5 values. We modeled this information also as definitions, but they do not affect to the reasoning process. They are just consulted by VALMIDE to obtain the mapping between the number of criteria and the final value.

We now present two different examples to illustrate how these definitions comprise the knowledge about the criteria. We focus on the subdomain of the Environment module, this is, criteria about the harshness of the environment:

- Criteria "crossing a place farther than 1 hour (walking time) from a inhabited place" is modeled as a new concept whose definition is

  ```
  (actionPerformed some CrossRemoteArea3H)
  ```

  along with the following definitions

  ```
  CrossRemoteArea3H equivalentTo
    (Cross
      and (actionPerformedIn some RemoteArea3H))

  RemoteArea3H equivalentTo
   (SingularElement
    and (distantFrom some InhabitedPlace)
    and (isWalkingDistance some integer[>=3]))
  ```

- Criteria "high probability of temperatures under 0°C" and "high probability of temperatures under 10°C" are modeled respectively as

  ```
  (hasMinTemperature some integer[<0])
  ```

  and

  ```
  (hasMinTemperature some integer[<-10])
  ```

[3] http://www.w3.org/TR/owl-primer/

These definitions enable VALMIDE to classify the texts with the help of a DL reasoner. In the following section, we present the MECMIDE module, that enables SIWAM to obtain the facts to be classified out from the text.

4 MECMIDE MODULE

The MECMIDE module is in charge of extracting the relevant information out from the descriptions provided by the users. It works as follows: For each fact or concept in the System Ontology that we want to be able to detect, MECMIDE has a list of search patterns that are to be looked up in the texts. For instance, the concept "Use hands on step" has a list of patterns such as "utilize hand", "use hand", "require hand use", etc. That list contains the lemmatized form of each word used. The patterns do not include articles, conjunctions, prepositions, or other words that lack of intrinsic semantic value.

To enrich the pattern-search analysis, MECMIDE includes two additional advanced semantic features:

- The processing window: MECMIDE searches each of the words composing the patterns in an area defined according to a given number of words. This feature enables MECMIDE to detect a pattern in the text regardless the word order. This is useful because sentences can hold different structures with the same words[4].
- The use of synonyms: Instead of looking just for the exact words of each pattern, MECMIDE considers also their synonyms. For example the word "way" has the same meaning as "traverse", "track", "trail", "path", "route", or "itinerary". The use of synonyms is solved using a lexical database, like WordNet (Miller, 1995), which allows MECMIDE to transform the words that integrate each pattern in *synsets*, i.e., the canonical form of its meaning. Then, the search is carried out using these synset, which broadens the vocabulary coverage of the patterns.

In many cases, it is not sufficient that a single pattern is recognized to deduce that the text is related to a particular concept. So, on top of these text patterns, MECMIDE uses a set of rules to decide if we can actually associate a text with a fact or concept in the VALMIDE concept (e.g., the need of a minimum number of patterns to deduce whether a particular description is related to an ontology concept). Thus, the process followed by MECMIDE to process each text is composed by three steps:

1. Lemmatization of the texts: Freeling (Carreras et al., 2004) is used as a tagger and lemmatizer, to filter stop words and to obtain the lemma of each word of the text, respectively.
2. Looking for patterns: To do this, the words that form the patterns are converted into *synsets* to improve the quality of searches. For example, "dangerous way" is equivalent to "slippery path". In our current implementation, we have chosen EuroWordNet (Vossen, 1998) as our lexical database because our prototype is in Spanish.
3. Applying the rules: The set of rules is evaluated to see whether the text must be linked to a concept, i.e., whether a fact can be derived from the text.

The result of this analysis is a set of concepts and facts that are associated to the text. In the following section, we present how this set of facts is used by VALMIDE to evaluate the activity according to the knowledge stored in the System Ontology.

5 VALMIDE MODULE

As we have seen in Section 3, the ontology used by SIWAM is an ontology network composed by two levels: One with information about the evaluation methods used for each evaluation criteria (GCDE Module), and the other with the actual knowledge needed to perform the actual evaluation (criteria modules). VALMIDE uses this information to perform the evaluation of each single activity. This process is composed by the following steps:

1. Deciding the evaluation modules: When it receives the description text along with the extracted facts, VALMIDE consults the GCDE module to see which evaluation modules are applicable to the type of activity that is being described.
2. Obtaining the evaluation methods: GCDE contains information about the specific method to be used for each evaluation module. VALMIDE consults it to know how to handle an specific module. This way, we can attach evaluation modules and methods in a flexible and decoupled manner.
3. Evaluating the activity description: For each module, VALMIDE asserts the facts in a copy of the evaluation module and classifies it. The classification and inferring capabilities of DL reasoners enable VALMIDE to make it explicit knowledge that otherwise would remain implicit. This information is used to calculate the actual evaluation.

[4]This structure richness is very typical when processing Spanish texts.

4. Combining evaluations: Depending on the evaluation module, VALMIDE can consider whether to just evaluate the activity as a single one, or to combine previous evaluations to achieve an agreed evaluation. The information of the method to use is stored in the GCDE module.

Regarding the MIDE Modules, we have developed an evaluation method based on the activity description classifications. As we have explained in Section 3, each MIDE criteria is modeled as a defined concept. The MIDE evaluation method obtains the agreed number of criteria fulfilled using the following formulae:

$$MIDE\ value(A_{instance}) = \sum_{i=1}^{|Crit|} fulfil(crit_i)$$

with

$$fulfil(crit_i) = \begin{cases} 1 & if \frac{\sum_{k=1}^{|Desc|} crit_i(desc_k)}{|Desc|} > CT \\ 0 & otherwise \end{cases}$$

where $A_{instance}$ is the activity we are evaluating, *Crit* is the set of criteria to be evaluated (the set of concept definitions), *Desc* is the set of individual descriptions for $A_{instance}$, CT is a confidence threshold in 0..1, and $crit_i(desc_k)$ is a function that evaluates whether a description is an instance of a particular criteria (this function is evaluated with the help of the DL reasoner, and returns 1 if the belonging relationship is entailed).

The above method counts how many of the criteria the activity fulfils by asking the DL reasoner whether the activity description belongs to each of the criteria definition concepts. Then, the MIDE information for each particular activity (the evaluation of each of its descriptions) is combined to establish whether a particular criterion is met. This is done by calculating whether it is present in the descriptions in a percentage above a particular confidence threshold. Finally, SIWAM translates the agreed number of criteria into the MIDE final value (ranged in 1 – 5 values) by consulting the information also stored in the module.

In the following section, we present two excerpts of an actual description to illustrate each of the steps that our system takes to process the information from text to the actual evaluation.

6 COMPLETE EXAMPLE

Although our system is in its early stages of implementation, we have already carried out some concept proofs that support the approach. In particular, we have used descriptions that correspond to real experiences of four different mountain ascensions located in the Pyrenees, a range of mountains that forms a natural border between France and Spain. The original texts are in Spanish, but we have translated the interesting excerpts used in this section to illustrate how SIWAM works.

In the ascension to the Petit Vignemale (3032 m), one of the mountaineers wrote:

> "... It took us a little more than 3 hours to reach the Refuge of Oulettes de Gaube, which was closed (we already knew it). This refuge is located at a height of 2151 m., in an isolated high mountain place, ...
>
> ... The temperature at such height is -18°C, with a 20 km/h wind, ..."

From these excerpts (*ex1*), MECMIDE extracts the following facts:

- From the subject of the description (is given by RSAM):

```
ex1 isA Hike
```

- From the location excerpt:

```
refugeOulettesGaube isA InhabitedPlace

zone2 isA SingularElement
zone2 distantFrom refugeOulettesGaube
zone2 isWalkingTime 3

crossZone2 isA Cross
crossZone2 actionPerformedIn zone2

ex1 actionPerformed crossZone2
```

  where *zone2* is the area which the mountaineers were traversing, and *crossZone2* is the action of traversing it.

- From the temperature excerpt:

```
ex1 hasMinTemperature -18
```

The instances created in the extraction are related to the activity instance, as MECMIDE assumes that a text contains the description of a single activity. Thus, VALMIDE asserts these axioms in the Environment module, and, with the help of a DL reasoner, infers the following (recall the definitions for the criteria presented in Section 3):

- From the location axioms:
  1. zone2 is a RemoteZone3H as it fulfils that is a SingularElement and is distant an InhabitedPlace (refugeOulettesGaube), and is at a walking distance of more than 3 hours.
  2. crossZone2 is a CrossRemoteZone3H as it is a Cross action performed in a RemoteZone3H.

3. ex1 is an instance of the concept definition associated to the MIDE criterion of crossing a remote area at 3 hours (walking distance) as it has an action that is a CrossRemoteZone3H.

- From the temperature axioms:
  1. ex1 is an instance of the concept definition associated to the criterion about temperatures below -10°C.
  2. ex1 is also an instance of the concept definition associated to the criterion about temperatures below 0°C.

VALMIDE thus calculates the final evaluation counting the criteria that ex1 is instance of[5]. Note how ex1 is added two points in the MIDE criteria due to its temperature as it fulfils two different criteria (below 0°C and below 10°C). This is coherent with the way the MIDE standard evaluates the activity.

7 RELATED WORK

The spread and presence of mountain activities thematic sites on the Web have been quite low compared to other fields. To the best of our knowledge, there is no such an approach as ours in any of the current Web sites about mountain activities.

Regarding the System Ontology, we have not found any other ontology modeling the domain of mountain activities. The closest works are related to the tourism domain (Fodor and Werthner, 2005; Prantner et al., 2007; Barta et al., 2009; Mouhim et al., 2011), although they have different aims as ours as they are mainly oriented to model the tourism domain within the context of the e-commerce. Anyway, we have considered all of them to capture the vocabulary of our ontology.

The most related works to SIWAM can be found in the Information Extraction (IE) field. In particular, according to (Wimalasuriya and D., 2010), SIWAM can be classified as an Ontology-Based Information Extraction (OBIE) system, which is extended with the reasoning capabilities of VALMIDE module. In this context, there are several approaches oriented to automatic content annotation (Cimiano et al., 2004; Buitelaar et al., 2008). PANKOW (Cimiano et al., 2004) processes Web pages looking for instances of a given ontology, thus automatically annotating the Web page with metadata about its content. SOBA (Buitelaar et al., 2008) is oriented to obtain structured information out from semi-structured resources (populate a knowledge base). However, these systems aim only at annotation and fact extraction, while SIWAM uses this extracted information to perform the evaluation of the different activities exploiting the model in the ontology and the DL reasoner classifying capabilities.

Without leaving the Information Extraction field, it is worthy mentioning several approaches (Wu et al., 2008; Cimiano and Völker, 2005) that have as its main goal to construct the ontology that is behind the processed information. For example, Kylin (Wu et al., 2008) uses extraction techniques against Wikipedia's articles to obtain a structured schema out from them. It uses also WordNet along with *machine learning* techniques to obtain the final ontology. However, constructing the ontology is not the objective of SIWAM. SIWAM exploits the domain model to detect important facts in the extraction stage, and then, the model is used to evaluate different aspects of each of the input activities (via their descriptions).

[5]In this example, we are considering just one description so the formulae in the previous section is reduced just to count the criteria concepts which *ex1* belongs to.

8 CONCLUSIONS AND FUTURE WORK

In this paper, we have presented SIWAM, a complete framework to share information about mountain activities. Apart from its social network features, SIWAM extracts and infers new information from the descriptions provided by the users to help assessing the difficulties of the different activities. This is done to improve the information for the practitioners in order to reduce the risks in the mountains. Moreover, to the best of our knowledge, we have developed the first ontology aimed at modeling mountain activities[6], and at modeling a standard evaluation method such as MIDE. Our system has the following features:

- It uses the descriptions provided by the users to extract relevant facts aligned to an ontology. This is a source of information which was almost unexploited in the mountain domain.
- It uses the extracted information to evaluate the agreed difficulty of each activity with the help of a DL reasoner. To do so, it exploits the inferring capabilities of the reasoner, along with the evaluation criteria definitions.
- It is completely ontology guided: the adoption of a modular evaluation method makes it possible to extend SIWAM with different evaluation standards in a flexible and efficient way.

[6]We cannot make it available by the time we are writing the paper due to several project restrictions.

Currently, the prototype is under development, although the preliminary results are very promising. As future work, apart from testing the user behaviour, we want to extend the modeled activities and include further evaluation methods taken from the field of recommender systems.

ACKNOWLEDGEMENTS

This research work has been supported by the CICYT project TIN2010-21387-C02-02 and DGA-FSE.

REFERENCES

Baader, F., Calvanese, D., McGuinness, D., Nardi, D., and Pastel-Scheneider, P. (2003). *The Description Logic Handbook. Theory, Implementation and Applications*. Cambridge University Press.

Barta, R., Feilmayr, C., Pröll, B., Grün, C., and Werthner, H. (2009). Covering the semantic space of tourism: An approach based on modularized ontologies. In *Proc. of the 1st Workshop on Context, Information and Ontologies (CIAO'09), Heraklion (Greece)*, pages 1–8. ACM.

Buitelaar, P., Cimiano, P., Frank, A., Hartung, M., and Racioppa, S. (2008). Ontology-based information extraction and integration from heterogeneous data sources. *International Journal of Human-Computer Studies*, 66(11):759–788.

Carreras, X., Chao, I., Padró, L., and Padró, M. (2004). FreeLing: An open-source suite of language analyzers. In *Proc. of the 4th Intl. Conf. on Language Resources and Evaluation (LREC'04)*, pages 239–242. European Language Resources Association.

Chamarro, A. and Fernández-Castro, J. (2009). The perception of causes of accidents in mountain sports: A study based on the experiences of victims. *Accident Analysis & Prevention*, 41(1):197–201.

Cimiano, P., Handschuh, S., and Staab, S. (2004). Towards the self-annotating web. In *Proc. of the 13th Intl. Conf. on World Wide Web (WWW'04), New York (NY, USA)*, pages 462–471. ACM.

Cimiano, P. and Völker, J. (2005). Text2Onto: A framework for ontology learning and data-driven change discovery. In *Proc. of the 10th Intl. Conf. on Natural Language Processing and Information Systems (NLDB'05), Alicante (Spain)*, pages 227–238. Springer Verlag.

Fodor, O. and Werthner, H. (2005). Harmonise: A step toward an interoperable e-tourism marketplace. *International Journal of Electronic Commerce*, 9(2):11–39.

Garrido, A. L., Buey, M. G., Ilarri, S., and Mena, E. (2013). GEO-NASS: A semantic tagging experience from geographical data on the media. In *Proc. of the 17th East-European Conf. on Advances in Databases and Information Systems (ADBIS'13), Genoa (Italy)*, pages 56–69. Springer Verlag.

Garrido, A. L., Gómez, O., Ilarri, S., and Mena, E. (2012). An experience developing a semantic annotation system in a media group. In *Proc. of the 17th Intl. Conf. on Natural Language Processing to Information Systems (NLDB'12), Groningen (The Netherlands)*, pages 333–338. Springer Verlag.

Global Industry Analyst, Inc. (2012). Extreme sports: A global industry outlook.

Grüninger, M. and Fox, M. S. (1994). The role of competency questions in enterprise engineering. In *Proc. of the IFIP WG5.7 Workshop on Benchmarking - Theory and Practice, Tronheim (Norway)*, pages 22–31. Springer.

Jenkins, M. (2013). Maxed out on Everest. *National Geographic*, 32(6):94–113.

Kara, S., Alan, Ö., Sabuncu, O., Akpınar, S., Cicekli, N. K., and Alpaslan, F. N. (2012). An ontology-based retrieval system using semantic indexing. *Information Systems*, 37(4):294–305.

Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, 38(11):39–41.

Ministerio del Interior (2012). Anuario estadístico del Ministerio del Interior. http://www.interior.gob.es/file/62/62261/62261.pdf, accessed September 13, 2013.

Mouhim, S., Aoufi, A. E., Cherkaoui, C. E., Douzi, H., and Mammas, D. (2011). A knowledge management approach based on ontologies: The case of tourism. *International Journal of Computer Science & Emerging Technologies*, 2(6):362–369.

Prantner, K., Ding, Y., Luger, M., Yan, Z., and Herzog, C. (2007). Tourism ontology and semantic management system: State-of-the-arts analysis. In *Proc. of IADIS Intl. Conf. WWW/Internet 2007, Vila Real (Portugal)*, pages 111–115. IADIS Press.

Roche, A. P. (2002). *MIDE: Método de Información de Excursiones*. Federación Aragonesa de Montañismo. http://www.montanasegura.com/MIDE/manualMIDE .pdf, accessed September 13, 2013.

Suárez-Figueroa, M. C. (2012). *NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse*. IOS Press.

UIAA Mountaineering Commission (2004). *To Bolt or not to Be*.

Vogrincic, S. and Bosnic, Z. (2011). Ontology-based multi-label classification of economic articles. *Computer Science and Information Systems*, 8(1):101–119.

Vossen, P. (1998). *EuroWordNet: a multilingual database with lexical semantic networks*. Kluwer Academic Boston.

Wimalasuriya, D. C. and D., D. (2010). Ontology-based information extraction: An introduction and a survey of current approaches. *Journal of Information Science*, 36(3):306–323.

Wu, F., Hoffmann, R., and Weld, D. S. (2008). Information extraction from Wikipedia: Moving down the long tail. In *Proc. of the 14th ACM Intl. Conf. on Knowledge Discovery and Data Mining (SIGKDD'08), Las Vegas (NV, USA)*, pages 731–739. ACM.