Multidisciplinary Research on Geographical Information in Europe and Beyond
Proceedings of the AGILE'2012 International Conference on Geographic Information Science, Avignon, April, 24-27, 2012
ISBN: 978-90-816960-0-5
Editors: Jérôme Gensel, Didier Josselin and Danny Vandenbroucke

Representing and Inferring Events from Deforestation Observations

Liqun Wu
Institute for Geoinformatics, University of Muenster
Weselerstr.253,48151
Muenster, Germany
tracia.wuliqun@uni-muenster.de

Abstract

Various techniques and approaches are used to monitor the deforestation in the Brazilian Amazon Rainforest. The result is a number of different datasets and data streams containing the results of the variety of observations. However, the heterogeneity in these observation datasets and the lack of linkage between them makes it difficult to have an idea about the actual events and change in the area. The Linked Data approach provides the possibility to connect different pieces of data. In this paper we describe an ontological structure with explicit and formal descriptions about the concepts in the tropic forest change monitoring domain, with a goal to support linking of spatial, temporal, thematic data of this domain. The idea is that such links would make it easier to query monitoring data, and extract interesting events from the data. We describe a tropic forest change monitoring to represent the forest change observations and the related domain knowledge, and also to represent tropic forest change events and involved objects. We will report how this ontology can be used to infer implicit events from the deforestation data. We finalize the paper by discussing the open issues about the representation of observations and events.
Keywords: Ontology, Linked Data, Amazon Rainforest, Forest Change Observation, Deforestation Event

1 Introduction

The large-scale deforestation emerging in the Brazilian Amazon Rainforest gained a lot of attention in recent years. It is because the loss of rainforest reduces earth's biodiversity and has a negative impact on climate change [1-4]. To monitor the forest change in this area, different techniques and approaches are used for collecting and processing the related data. However, the event representations remain implicit in the resulting datasets. Due to the heterogeneity among such datasets and ambiguous meanings of the data items, a lot of manual work is required if one wants to retrieve or figure out what is going on or happened at certain time in that area.

The Linked Data approach [5] provides the technologies to connect these data. Based on the Resource Description Framework (RDF) model in the form of <subject, predicate, object> and HTTP URIs for naming the resources, a machine can retrieve more information by following the links between two data items [6]. However, without explicating the underlying semantics such Linked Data alone is difficult to be used through machine reasoning, and would require precise patterns of the RDF graph for query [7].

In this paper we envision that the explicit and formal descriptions about the domain concepts' definition and characteristics, as well as the clear structure and relationships between them, would support to generate the links between the spatial, temporal, thematic data of the Brazilian Amazon Rainforest and to extract automatically the high-level descriptions about forest change. For this, we developed an ontology of the tropic forest monitoring. The idea is that different pieces of data can be annotated using the concepts of this ontology. By applying queries and reasoning rules, the implicit forest change events can be extracted from the data.

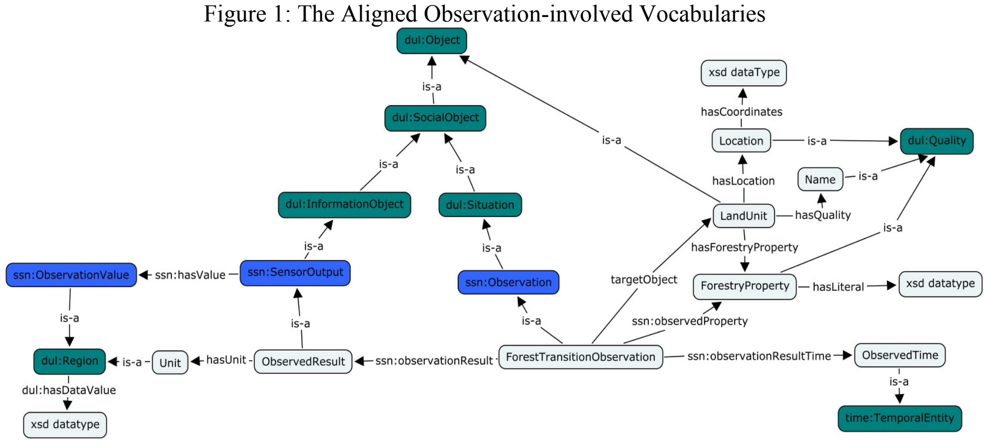
Upper-level ontologies offer the definition of the general occurrence cross the domain, serving as a semantic base for lower-lever ontologies[8]. The constructed ontology is aligned to the lightweight upper level ontology DOLCE+DnS Ultralite (DUL). The DUL simplifies and extends the DOLCE Lite-Plus Ontology and the Constructive Descriptions and Situation (DnS) Ontology with more intuitive concept names and simpler relations[9].It distinguishes the event from the object as the approach mentioned in [10,11,12]: Events are the occurrences happening over time, while object are entities exists over space. This corresponds to the idea of DOLCE to distinguish the enduring and perduring entities. The Quality concept is the perceivable and measurable entities and always presents with other entities. The value of a Quality describes its position within a quality space. This conceptual space falls into the Abstract class. [13] The DnS pattern makes a separation between the state of affairs and their interpretation which allows contextualized view on events [14].

Part of the Semantic Sensor Network Ontology (SSN) is also used for the alignment. The SSN ontology is developed for modelling sensor devices, systems and processes [15]. It provides a structure specially for representing the observation in general level. The SSN Ontology itself is aligned to the DUL ontology. Thus aligning the constructed ontology to the SSN will not introduce conflicts when using the DUL for alignment.

The constructed ontology borrows the OWL-Time Ontology [16] vocabulary for representing the temporal aspect. This ontology provides the vocabularies for describing temporal instants and intervals, as well as the topological relations among them, the duration expressions and the datetime information.

The structure of the paper is as follows. Section 2 introduces the tropic forest monitoring ontology built in this research. In Section 3 we illustrate how to derive event by using this ontology. The open issues are discussed in Section 4. We complete the paper with concluding remarks in Section 5.

Figure 1: The Aligned Observation-involved Vocabularies



2 Tropic Forest Monitoring Ontology

This section introduces the constructed tropic forest monitoring ontology. It has two parts: the ontological structure to represent the monitoring observations, and the domain knowledge vocabularies.

2.1 Ontology Construction Methodology

The ontology construction modifies the process of the METHONTOLOGY approach [17] to adapt to the light-weighted and application oriented ontology building with the following procedure:

1. Determine the scope: specify the intended use and the covered aspects of the ontology. This ontology aims to provide a structure with the vocabularies, to describe the observable events about the tropic forest change from the Brazilian Amazon monitoring data. It focuses on the physical phenomena about the increase or decrease of the forest on the earth surface.
2. Collect the terms: collect terms and their definitions which satisfy the defined scope. It includes the vocabularies of forest change events e.g. deforestation, and the involved objects and their characteristics, e.g. the geographic objects.
3. Formalize the inner structure: conceptualize the hierarchy and relations among the terms. This step results in a preliminary ontological structure covering the intended scope with only the collected terms from last step.
4. Align the terms to external vocabularies: the vocabularies determined in the previous steps were aligned to the DOLCE+DnS Ultralite ontology and SSN ontology according to the preliminary ontological structure. Some terms were replaced with the vocabularies in the aligned ontologies.
5. Implementation: the ontology is then encoded in the Web Ontology Language [18].

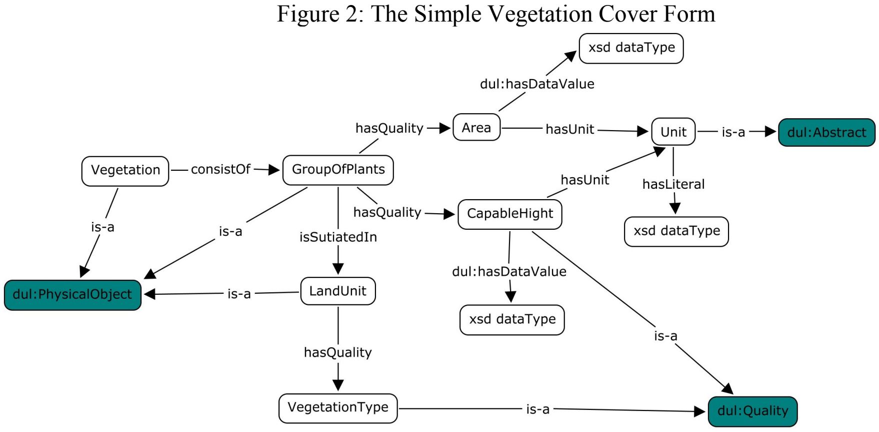
2.2 Forest Transition Observation

The concepts in the first part of the ontology are aligned to the ontologies introduced in Section 1 as in Figure 1.The prefixes used are: DUL as “dul”; OWL-time as “time”, SSN as “ssn”. Vocabularies without prefix come from the constructed ontology. The propertied are also aligned, e.g. the *targetObject* is a sub property of the *ssn:featureOfInterest*.

The *ForestTransitionObservation* is introduced as the central concept to connect different information. The observation structure is the essential base of the extraction of events. The description below is given from spatial, temporal, thematic aspects.

Spatial: It is needed to distinguish the location of the observed content from the one where the observation is taken. This ontology focuses on the observed content. The observation’s direct target here is the perceivable change of the plants. Since those plants are immovable objects situated on the earth surface, the collection of them can be grouped by geographic partition. The strategy is to relate the observation to the geographic objects which carry the spatial properties, treating them as the observed targets instead of the complex collection of the plants. A concept *LandUnit* is introduced.

Figure 2: The Simple Vegetation Cover Form



Multidisciplinary Research on Geographical Information in Europe and Beyond
Proceedings of the AGILE'2012 International Conference on Geographic Information Science, Avignon, April, 24-27, 2012
ISBN: 978-90-816960-0-5
Editors: Jérôme Gensel, Didier Josselin and Danny Vandenbroucke

The observation is related to it through the property *targetObject*. The *LandUnit* is defined as any distinguishable terrestrial part of the earth surface to satisfy the datasets under multiple scales. A LandUnit instance could exist in the dataset both as a point and with spatial extension. As long as it is geographically distinguishable, a LandUnit is allowed to be created by any kind of dividing standard. An instance of this class is determined purely by its geographic location. This ontology considers the location as a static characteristic of the *LandUnit*, which means the instance remains the same only when its location does not change. The *GeoLocation* of the *LandUnit* is treated as an entity. For avoiding too complex pattern, the value of the *GeoLocation* (described via coordinates) is represented using datatype instead of another entity.

Temporal: The constructed ontology only includes the time when the initial content is observed is. The temporal characteristics of the observations are also treated as entities. Strictly speaking, each observation is completely acquired at a certain time point when the content completes its present in reality. But a data item is also possible to reflect the situation over a period. The TemporalEntity concept from the OWL-time ontology is used as the range of the object property *observedResultTime*. It supports both time instant and interval.

Thematic content: This ontology relates an observation to the *ForestryProperty* (what it observed) and the *ObservationResult* (how the situation of that property is). The *ForestryProperty* is the quality of the *LandUnit*. The *ObservationResult* has *ObservedValue* and *Unit*. The *Unit* does not exist depending on the measurement. Different measurements may have the same unit. The *ObservedValue* and *Unit* are treated similarly as the other concept with quantitative or qualitative values: the vocabulary is included as a class with some data type property, mapped to some literal value.

2.3 Domain Knowledge Representation

The domain knowledge part provides the vocabularies and the hierarchy of the physical events which often appear in the Amazon rainforest, and are directly indicated by, or lead to the amount change of the plants. These vocabularies are aligned to DUL Event class. They are then being used to annotate the extracted events together with other part of the ontology. The temporal properties of these events are introduced based on the observation. The object property *isObservableAt* is used, when the observed duration is available. And the object properties *isObservableFrom*, *isObservableTo* are used, when the earliest and latest observed time is available.

This part also includes a simplified form to represent the situation of the vegetation covering a geographic object as shown in Figure 2. A *GroupOfPlants* can be situated in a *LandUnit*. Such a collection is internally heterogeneous, with different types of plants existing within it. But depending on the dominant plant type, the whole is recognized as certain kind of *Vegetation*. The object property *consistOf* is used to relate the *GroupOfPlants* to the *Vegetation*, when this *GroupOfPlants* instance is all the plants which constitute the *Vegetation*. The *VegetationType* is a quality of the *LandUnit*, whose value is determined by the Vegetation on this *LandUnit*. This *GroupOfPlants* is defined broadly. It could be under any level or standard of separation and the plants do not necessarily be geographically continuous. An instance of this class can have parts, which are other instances of this class. The property *isBiggestAmountIn* is introduced to indicate the total amount of the certain type of plants, which is part of a *GroupOfPlants* instance. This structure and vocabularies are the supplement of representation to describe the event-related objects.

3 Inferring an Events

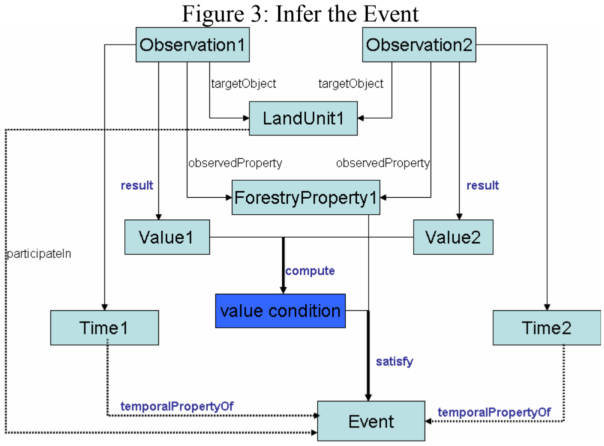
By associating the data with the constructed ontology, extraction of the implied event can be supported.

3.1 The Assumption

The constructed ontology can be used to infer the event instances which meet some requirements to represent events in [19]: distinguish the type of the event, associate it with the place, and express the temporal characteristics. As the focus is on the events denoted by the change of the plants, the assumption is that those events can be extracted via the comparison between two situations, which are recorded in two observations. How to infer an event from two observations follows the pattern in the Figure 3. The words along the lines are the relations. Among these, the blues ones are not the ontology vocabulary but for simplified description. The bolder lines refer to the inferring process. The dotted lines denote the inferred relations to the extracted event.

When two *ForestTransitionObservation* instances observe the same kind of *ForestryProperty* of the same *LandUnit* instance, the observed values (the value of the *ObservedResult* in the ontology) are computed according to the rules. Then the computed value condition, together with the observed *ForestryProperty* is checked. When they satisfy the condition of a certain type of event, an event instance of this type is generated. The target *LandUnit* is assigned as the participant of this event, and the time when these two observations were taken are related to its temporal characteristics.

Figure 3: Infer the Event



Multidisciplinary Research on Geographical Information in Europe and Beyond
Proceedings of the AGILE'2012 International Conference on Geographic Information Science, Avignon, April, 24-27, 2012
ISBN: 978-90-816960-0-5
Editors: Jérôme Gensel, Didier Josselin and Danny Vandenbroucke

3.2 A Pilot Study

This approach was applied in a pilot study to extract the deforestation event. The test dataset is the RDF triples generated from vector data [20]. It records the deforested area percentage from different years of 8250 cells created based on the pixels of the deforestation monitoring image, covering the whole Brazilian Amazon, each with the size of 25 km * 25 km. We annotated these RDF triples with the observation-related vocabularies in the constructed ontology.

The tests were formalized according to the assumption described above, and were run over the annotated observations. For example, 817 Deforestation instances in year 2008 are discovered and simply constructed as blank nodes for illustration using the SPARQL query [21] as (1). Their information matches the result of traditional query in the vector data with manual checking.

```
CONSTRUCT
 {
_:d             rdf:type                  tfc:Deforestation;
                dul:participant           ?cell;
                tfc:isObservbleAt         tfc:year2008.
 }
WHERE
 {
?observation1  ssn:observedProperty      ?p;
               ssn:observedResultTime    ?tfc:year2008;
               ssn:observationResult     ?result1;
               tfc:targetObject          ?cell.
?result1       ssn:hasValue              ?value1.
?value1        dul:hasDataValue          ?datavalue1.
?observation2  ssn:observedProperty      ?p;
               ssn:observedResultTime    ?tfc:year2007;
               ssn:observationResult     ?result12;
               tfc:targetObject          ?cell.
?result2       ssn:hasValue              ?value2.
?value2        dul:hasDataValue          ?datavalue2.
?p             rdf:type                  DeforestedPortion.
FILTER (?datavalue1-?datavalue2 > 0) }
```

(1)

Such work could support the historical deforestation event discovery by machine. In future work more rules need to be tested on the data for discovering different events with more complicated conditions, as well as for the optimization of the event instances such as merging event instance of same type and associating multiple participants in one event. After that, the next step would be testing this approach to the coming data stream to detect newly happened deforestation event, which can benefit the fast reaction to these events.

4 Discussion

The open issues regarding to the modelling of the ontology in this paper are discussed in this section.

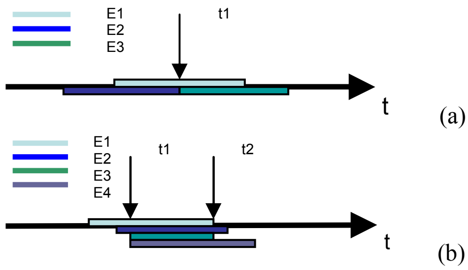
4.1 Modelling the Time of Events

When following the assumption in section 3.1 to infer event from the observations, how to assign suitable temporal representation to those events according to their different natures, is still need to be modeled. Some of the different temporal natures of the events are discussed as follows.

Through monitoring people may be interested in being notified the on-going situation ("what was/is/will be happening?"). This implies the events which at least fulfill the whole time extension being considered. Other situation is to tell the occurrence as a distinguishable entity ("what happened/happens/will happen?") that implies the events which at least not extend over the considered time. This difference results in the different temporal natures of the events. The following discussion uses present tense for simplification reason.

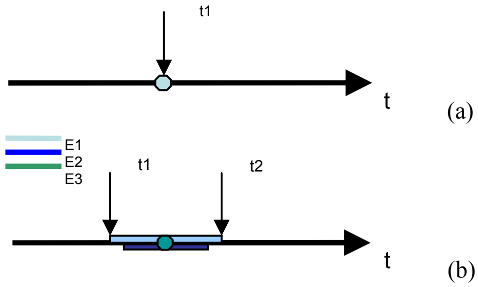
In the macroscopic world at low speed, considering time as an one-directional, homogeneous line is sufficient for representation. Under this condition, when describing on-going situation, two possibilities can be seen as in Figure 4. The E1~E4 show the temporal extensions of different events on the timeline. In Figure 4(a), people look at a certain time point t1 to see the on-going situation. The situation may probably pass through this point as E1. Or it may last till this point as E2, so at this point the condition of E2 still satisfies but at not any later point. The t1 can also be the earliest point as E3. In Figure 4(b), people look at a certain time period as the interval [t1, t2] shows. The on-going situation should satisfy the whole interval, which could be as one of the E1~E4 illustration.

Figure 4: The On-going Situation



Similarly, telling occurrence as a distinguishable entity divides into two situations, showed in Figure 5. Viewing from a certain time point t1 as Figures 5(a), if an event can be introduced as a distinguishable entity at this moment, it cannot last more than this instant point. In Figures 5(b), when the events do not cover the timeline over [t1, t2] interval, it can be viewed as whole entity within [t1, t2]. E1 has the time extension which exactly cover whole [t1, t2] is possible. Events as E2 also suit this situation, but have various forms which can "start", "during", or "finishes" [22] this time interval. An instantaneous event as E3 also can be the answer of "what happens" at [t1, t2] since it is fully within the time span.

Figure 5: Occurrence as a Distinguishable Entity



Also, the temporal resolution influences the event's behavior in the data. As Figure 3-b shows, when [t1~t2] (e.g. "one day") is the least time separation between the data records, the durations of E1, E2, E3 are undistinguishable. It turns that the "one day" does not really behave as an interval but reduces to an undividable zero dimension point on the timeline representation. Due to the various temporal resolutions, the same time length can be seen both as a point or interval on the timeline. However, it is not found that there is ontological support to represent the same temporal length under different temporal resolutions. The modeling of such temporal representation should gain further attention.

4.2 Conceptualization of *ForestyProperty*

Currently, the *ForestryProperty* is treated as the quality of the *LandUnit*. An instance of this class coexists with an instance of the *LandUnit* class, and can be the observed property of the observation. The specific observed aspects in the datasets are treated as sub classes of this class. When comparing different observation values, the machine checks the observed property instances to make the correct pair for comparison.

However, the conceptualization is arguable. The observations taken by certain kind of sensors observe the same aspect of some objects of the same type, for example, the temperature. The use of such an aspect is more similarly to an universal property whose behavior is homogenous to all involved objects. Then each observation associated with the same observed property instance may have the possibility to be paired. It needs further discussion.

5 Conclusions

This paper applied the ontological approach to make useful connections between the Amazon deforestation data. It associates the data with conceptual level descriptions. This approach enables the extraction of the implied forest change-related events, by matching the computed condition from observations with the condition which satisfies the events. It could support the historical event discovery, as well as the detection of newly happened event. Thus the domain experts' work such as analysis of the deforestation tendency and the deforestation notification can be enhanced by this approach. The following research should focus on to correctly relate the objects to the events according to their different temporal natural, and test this approach by integrating it into the applications.

6 Acknowledgement

This work has been funded through by the European research project ENVISION (FP7 249170).

References

[1] Jean-Paul Lanly. Deforestation and Forest Degradation Factors. In *Proceedings B of the XII World Forestry Congress*, Québec *City*, Canada, 2009.

[2] William F. Laurance, A. K. M. Albernaz and Carlos Da Costa. Is deforestation accelerating in the Brazilian Amazon? *Environmental Conservation*, 28(04): 305-311, 2001.

[3] S. F. D. Ferraz, et al.. Landscape dynamics of Amazonian deforestation between 1984 and 2002 in centra Rondônia, Brazil: assessment and future scenarios. *Forest Ecology and Management*, 204(1): 69-85,2005.

[4] Paulo Moutinho and Stephan Schwartzman. Tropical Deforestation and Climate Change. *Environmental Defense*, Washington, D.C, USA, 2005.

[5] Christian Bizer, Tom Heath and Tim Berners-Lee. Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*, 5(3): 1-22, 2009.

[6] Christian Bizer, Richard Cyganiak and Tom Heath. How to Publish Linked Data on the Web. [Online] Available http://www4.wiwiss.fu-berlin.de/bizer/pub /LinkedDataTutorial/, August 2011.

[7] Prateek Jain, et al.. Linked Data is Merely More Data. In *AAAI Spring Symposium "Linked Data Meets Artificial Intelligence"*, Menlo Park, USA, 2010.

[8] Viviana Mascardi, Valentina Cordì and Paolo Rosso. A comparison of Upper Ontologies. *Technical Report DISI-TR-06-21*, Genova, Italy, 2006.

[9] DOLCE+DnS Ultralite Wiki. Ontology: DOLCE+DnS Ultralite. [Online] Available http://ontologydesignpatterns.org/wiki/Ontology:DOLCE+Dn S_Ultralite, August 2011.

[10] Roberto Casati and Achille Varzi. Event. *The Stanford Encyclopedia of Philosophy*, *Spring 2010 Edition*, [Online] Available http://plato.stanford.edu/archives/spr2010/entries/ events/, August, 2011.

[11] David Hugh Mellor. On Things and Causes in Spacetime. British Journal for the Philosophy of Science, 31(3): 282-288, 1980.

[12] Anthony Quinton. Object and Events. *Mind*, 88: 197-214, 1979.

[13] Claudio Masolo, et al.. WonderWeb EU Project Deliverable 18: Ontology Library. 2003. [Online] Available

http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf, August 2011.

[14] Aldo Gangemi and Peter Mika. Understanding the Semantic Web through Descriptions and Situations. In *Proceedings of CoopIS/DOA/ODBASE*, Catania, Italy, 2003.

[15] Semantic Sensor Network Incubator Group. The Semantic Sensor Network Ontology. [Online] Available http://www.w3.org/2005/Incubator/ssn/wiki/Report_Work_on_the_SSN_ontology, August 2011.

[16] Jerry R. Hobbs and Feng Pan. Time Ontology in OWL. *W3C Working Draft*, 2006. [Online] Available http://www.w3.org/TR/owl-time/, August 2011.

[17] Asunción Gómez-Pérez, Mariano Fernández-Lopez and Oscar Corchoand. *Ontological Engineering*. Springer, London, 2004.

[18] W3C OWL Working Group. OWL 2 Web Ontology Language Document Overview. *W3C Recommendation*, 2009. [Online] Available http://www.w3.org/TR/owl2-overview/, August 2011.

[19] Ryan Shaw and Ray R Larson. Event Representation in Temporal and Geographic Context. In *Proceedings of the 12th European conference on Research and Advanced Technology for Digital Libraries*, Aarhus, Denmark, 2008.

[19] Ryan Shaw and Ray R Larson. Event Representation in Temporal and Geographic Context. In *Proceedings of the 12th European conference on Research and Advanced Technology for Digital Libraries*, Aarhus, Denmark, 2008.

[20] Tomi Kauppinen and Giovana Mira de Espindola. Linked Brazilian Amazon Rainforest Data. [Online] Available http://linkedscience.org/data/linked-brazilian-amazon-rainforest/, 2011.

[22] James F. Allen. Maintaining Knowledge About Temporal Intervals. *Communications of the ACM*, 26(11): 823-843, 1983.

Multidisciplinary Research on Geographical Information in Europe and Beyond
Proceedings of the AGILE'2012 International Conference on Geographic Information Science, Avignon, April, 24-27, 2012
ISBN: 978-90-816960-0-5
Editors: Jérôme Gensel, Didier Josselin and Danny Vandenbroucke