Brazilian Cerrado: a case study of linked geographical and statistical data applied to Ecology

Adriano Souza[a,*], Oscar Corcho[b], Luis Vilches-Blázquez[b], Paulo Salles[a]
[a]*University of Brasilia, Institute of Biological Sciences, Graduate Program in Ecology, Campus Darcy Ribeiro, CEP: 70904-970, Brasilia, DF, Brazil.*
[b]*Ontology Engineering Group, Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, Spain.*

Abstract. This paper describes an ontology network that will serve as the basis for publishing and linking data about wood plant communities of the Brazilian Cerrado biome, obtained from scientific studies, meteorological and environmental data and geographical information (maps). Data Cube, Meteorological, GeoSPARQL and Time ontologies were used as infrastructure ontologies. In addition, two domain ontologies - Cerrado Concepts and Wood Plant Dynamics Ontology (Ccon) and Fire Ontology (Fire) - were developed to represent scientific knowledge about vegetation ecology focused on vegetation dynamics under different burning regimes. Datasets provided by Brazilian government agencies and those obtained from scientific literature, found in different formats, were transformed into RDF using Open Refine along with its RDF extension or, for shape files, using the geometry2RDF tool. A web-based application was deployed using Map4RDF, so as to provide a proper visualization of aggregated information, integrating map visualization using Google Maps API with ontology-based facet browsing. Ongoing work investigates possibilities of integrating these data with qualitative reasoning models. This research has potential to boost applications of linked geographical and statistical data technologies into ecological research and applications to conservation of biodiversity.

Keywords: Ontology network, Geolinked data, Ecology, Brazilian Cerrado

1. Introduction

Ecological research and its application on natural resources conservation and management require researchers to access data about distinct geographic regions, available on different sources and stored using different technologies, formats and models. It is also necessary to integrate and analyze the data; to produce predictive and explanatory models; and eventually use these results to develop public policy recommendations [12]. Moreover, the development of a robust ICT infrastructure, including standard vocabularies and ontologies, data access, share and reuse are considered important challenges faced by ecoinformatics researchers [24].

The system we are concerned about is the Cerrado, the second largest biome in Brazil, and one of the biggest savannas in the world, with about 2 million km^2 in the central area of the country. It is one of the 25 world hotspots of biodiversity [28], that is, a place with a high number of endemic species. Nevertheless, it is threatened by modern agricultural development and charcoal production for Brazil's steel industry [36], what is resulting in about 48.21% of the total area of Cerrado deforested until 2009 [25].

Considered a moist tropical savanna (annual rainfall $\geq$ 1500mm), the vegetation of the Cerrado biome shows a remarkable physiognomic variation [29], spanning from forest like (*Cerradão*) and savanna (*Cerrado sensu stricto* and *Campo cerrado*) to open areas (*Campo sujo* and *Campo limpo*). Several studies have been carried out aiming to understand the factors affecting the past and present geographical distribution of the physiognomies, their

* Corresponding author. E-mail: souzaajb@unb.br.

structure and composition [4,10,12-13,15,28,34,36]. It is essential to identify specific mechanisms that enable coexistence of trees and grasses, and determine the density of trees in the Cerrado in order to understand the dynamics of vegetation and to develop management strategies [14]. Four classes of hypotheses attempt to explain the coexistence of trees and grasses: (a) spatial separation niche; (b) phenological niche separation; (c) balanced competition; and (d) demographic bottlenecks. Due to the capacity for integrating disperse resources about ecological systems we believe that semantic technologies can successfully be applied to support ecologists in exploring these hypotheses about mechanisms of tree-grass coexistence in the Cerrado biome.

In this paper we describe an ontology network, following the principles of Open Linked Data and GeoLinked Data initiatives [40], and we present a use case interconnecting data about wood plant communities of Cerrado biome from scientific studies and linking them to other data sources such as meteorological, environmental and geospatial data.

The paper is organized as follows: initially we present an overview of our ontological model. Next, the infrastructure ontologies and domain ontologies are described. How the data was transformed is explained section 5, and in section 6 how they were linked and published. Related works are discussed in the following section and, finally, we present our conclusions.

2. Overview of our ontological model

There are important ongoing applications of semantic technology in biological sciences, such as The Open Biological and Biomedical Ontologies [1] (OBO Foundry), and the BioPortal [2] [2,45], which, currently, store about 368 ontologies.

In the ecology domain, we can highlight the Environmental Ontology (EnvO) [47] which provides a controlled, structured vocabulary designed to support the annotation of any organism or biological sample with environment descriptors. EnvO also contains terms for biomes, environmental features, and environmental material. Other ontologies include the Landscape Ontology [22]; Ecolingua [6]; the Ecological Concepts ontology [45], the Open Linked Amazon Vocabulary [18] and, finally, the Extensible Observation Ontology (OBOE) [23]. However, current available ontologies do not directly cover issues regarding ecological communities. The work presented here addresses specific concepts on the domain of ecological communities and processes that can lead to changes in their structure and composition over time, which are not well understood by ecologists.

Our Ontology Network was developed to represent scientific knowledge about vegetation ecology focusing on the dynamics of Brazilian Cerrado wood plants. This model includes ontologies from different domains, such as statistics, climate, geospace, time, plant communities and fire. With this network of ontologies it is possible to describe and connect data about the dynamics of wood plants and meteorological data, and data about fire events and maps, in different locations. It aims to store data and retrieve information about biodiversity, species composition and search for changes in Cerrado species populations over time. It can also be used to assess the influence of environmental variables such as precipitation and fire on community composition and dynamics.

The ontologies that constitute this model have been implemented in OWL [44] following good practices for publishing Linked Data [17]. Therefore, existing ontologies that match our ontology requirements were reused.

Figure 1 illustrates the two different layers the ontologies can be classified according to whether the ontology represents information required for the infrastructure level or domain specific information. These ontologies match the knowledge representation requirements acquired over the development of the model and the use case prototype, which are:

- To represent multi-dimensional statistical data. This requirement is covered by the W3C RDF Data Cube vocabulary [9].
- To represent meteorological measurements and their properties. The AEMET network of ontologies [33] covers this requirement.
- To represent geographical information and geospatial data. The GeoSPARQL ontology [30], which provides a vocabulary to express geospatial information covers this requirement.

  To represent time, instants and intervals. This requirement is covered by the W3C Time Ontology.

[1] http://www.bioontology.org
[2] http://bioportal.bioontology.org/

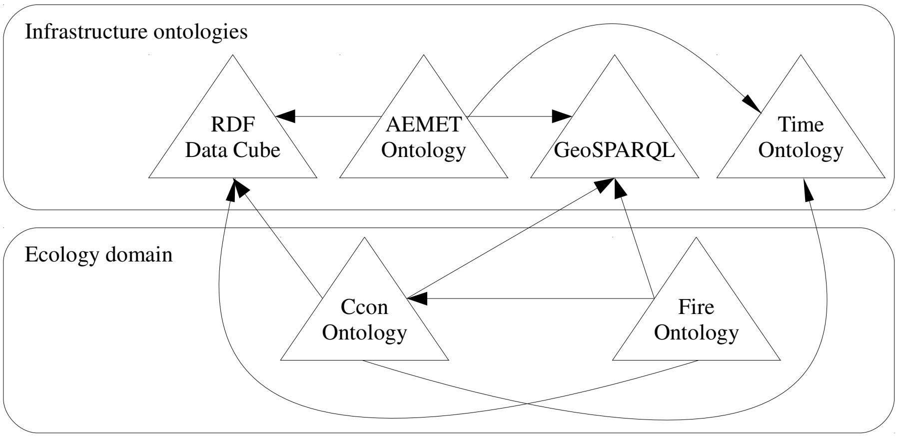

Fig. 1. Overview of the Brazilian Cerrado ontology network model.

- To represent ecological concepts such as dynamics of plant communities, and the different physiognomies of Cerrado biome. This requirement is covered by the Cerrado Concepts and Wood Plant Dynamics Ontology.
- To represent fire regimes and fire events. The Fire Ontology covers these requirements.

The following sections describe with more detail the Infrastructure and domain ontologies respectively.

3. Infrastructure ontologies

Infrastructure ontologies deal with the description of datasets, statistical, meteorological, geographical data and time information.

3.1. Data Cube Ontology

The RDF Data Cube vocabulary is focused on the publication of multi-dimensional or statistical data on the web [9]. The model underpinning the Data Cube vocabulary is compatible with the cube model that underlies SDMX (Statistical Data and Metadata eXchange), an ISO standard for exchanging and sharing data and metadata among organizations [9].

3.2. AEMET Ontology

The vocabulary used to describe meteorological variables is the AEMET Network of Ontologies [1]. It is a modular ontology, consisting of a central ontology that links a set of subdomain ontologies related to meteorological measurements modeling. Each measurement represents the atmospheric condition (humidity, temperature, precipitation etc.) at a place and time [1].

3.3. GeoSPARQL Ontology

GeoSPARQL is an OGC (Open Geospatial Consortium) standard that supports the representation and querying of geospatial data on the Semantic Web. GeoSPARQL defines a vocabulary for representing geospatial data in RDF, and it defines an extension to the SPARQL query language for processing geospatial data [30]. The GeoSPARQL is based on the OGC's Simple Features model, with some adaptations for RDF [3].

3.4. W3C Time Ontology

The W3C Time ontology provides a vocabulary for expressing facts about topological relations among instants and intervals, together with information about durations, and about datetime information [15].

4. Domain ontologies

This section describes the domain ontologies developed to represent the knowledge about Cerrado wood plant communities, their physiognomies and dynamics and also the influence of fire on these ecological systems. In order to accomplish this objective two OWL ontologies were developed.

4.1. Cerrado Concepts and Wood Plant Dynamics Ontology (Ccon)

Cerrado Concepts and Wood Plants Dynamics Ontology (Ccon) [3] represents scientific knowledge about vegetation ecology, focused on describing the dynamics of wood plants in the Cerrado. It aims to store data and collect information about biodiversity, composition and plant population dynamics, to manage data about cerrado plant community, and search for changes in Cerrado species populations over time.

The ontology describes how the Cerrado is structurally organized, in terms of biological organization levels. Many types of Cerrado physiognomies have been described and studied. However there is no consensus about the definitions and specific characteristics of each one. Thus, the concepts used in the ontology are the most accepted and used by researchers and research institutions and universities in Brazil, that is, the ones more frequently used in the scientific literature [6,9,12].

The ontology defines characteristics of the physiognomies, and also defines the concept of savanna, one of the major types of terrestrial ecosystems and the most representative ecosystem type in the Cerrado domain. Therefore, the ontology has distinct classes to represent each of the most representative physiognomies (Figure 2), and how the plant communities are organized to define them. In addition, the ontology defines classes to represent the biological diversity, reusing several terms from both Crop-Wild Relatives Ontology (CWR) [4] and from Environmental Ontology (EnvO) to describe the structure and dynamics of wood plant communities and the role of the driving forces acting to determine the structure and composition of these plant communities.

4.2. Fire Ontology (Fire)

The Fire [5] ontology was created in order to represent a set of concepts about fire occurring in natural vegetation, its characteristics, causes and effects, with a focus on the Cerrado vegetation domain. Fire plays a determinant role on the structure and composition of Cerrado physiognomies [7,25-26]. Events of fire in the vegetation are typically described in terms of burned area, duration, severity, frequency, along with other measurements when available [25,30]. Therefore, these variables where taken into account to build the ontology. In addition, relevant information includes causes and effects of fire events (Figure 3). For thousands of years, natural fires, during the wet season, and anthropogenic fires, during the dry season coexisted in the Cerrado region, suggesting that fire is one of the determinants of the Cerrado vegetation forms [26]. Currently, the main cause of fire in Cerrado is anthropogenic, that is, a human action that set fire in the vegetation [7,30-31].

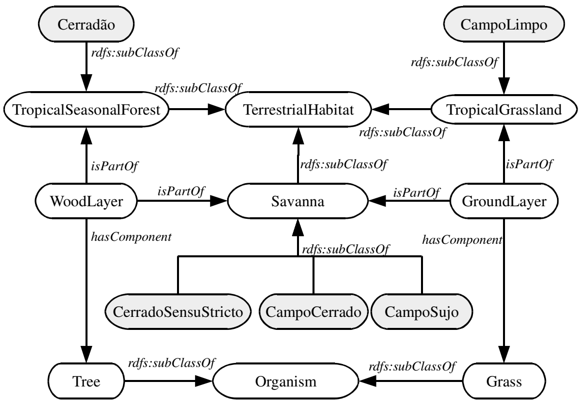

Fig. 2. Graphical representation of some classes and object properties of Ccon Ontology. The physiognomies are represented in grey.

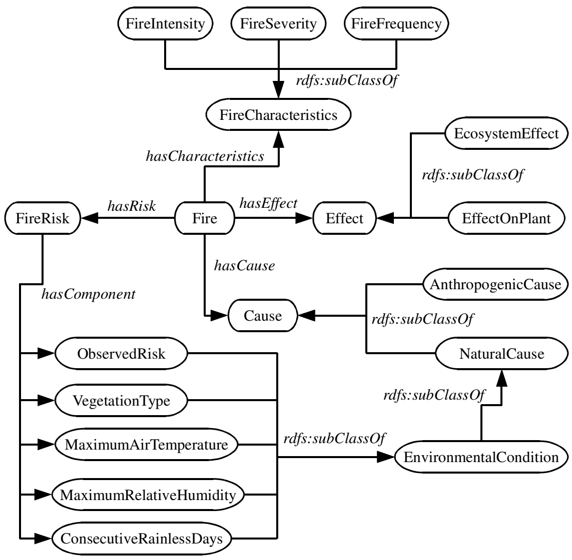

Fig. 3. Graphical representation of some classes and object properties of Fire Ontology.

[3] http://cerrado.linkeddata.es/ecology/ccon
[4] http://aims.fao.org/kos/crop-wild-relatives-ontology
[5] http://cerrado.linkeddata.es/ecology/fire

Some important effects of fire were considered, mainly those directly related to wood plant dynamics in Cerrado, which is the focus of the whole network ontology. Wildland fires are described as uncontrolled burn in opposition to controlled burn such as cool-season fires and high-intensity fires used for environmental management.

5. Ontology evaluation

The correct application of ontology development methodologies favors ontology quality. However its quality can be affected by difficulties involved in the modelling process which may cause the appearance of anomalies in ontologies [34].

Therefore, like any engineering artifact, an ontology needs to be thoroughly evaluated [43]. Moreover, it is necessary to identify which aspects of the ontologies are relevant to be evaluated, and select the most appropriate techniques to perform the evaluation. We used OOPS! (OntOlogy Pitfall Scanner)[6], an online service that supports automatic detection of pitfalls [34] which affect structure, functionality and usability of the ontologies.

5.1. OOPS! evaluation

OOPS! was used to keep quality control during the modeling process. That is, for the first ontology candidate version we performed the evaluation using OOPS! and all the pitfalls detected were counted. The results obtained were used as a reference to manually fix the errors using an ontology editor, and then, in an iterative process, new versions of the ontologies were generated. Pitfalls classified as critical and important were considered the most relevant to be fixed.

Figures 4 and 5 show the amount of pitfalls found in each version of the Ccon and Fire ontologies respectively. The amount of pitfalls decreases from the first to the last version. In the final version of both ontologies only one type of pitfall remains, which is missing comments on the properties.

5.2. Evaluation by domain experts

It is also important to evaluate the domain ontologies with regard to their conceptual correctness

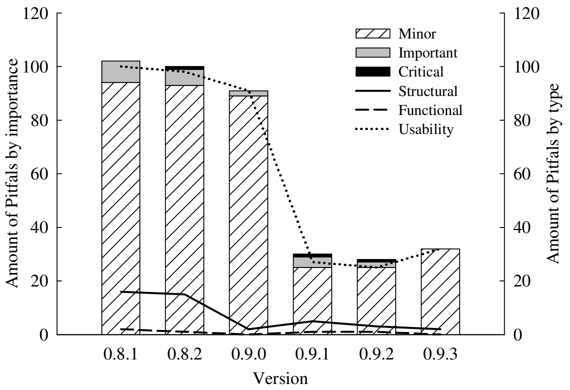

Fig. 4. Amount of Pitfalls by importance and by type, found in each version of Ccon Ontology, using OOPS!.

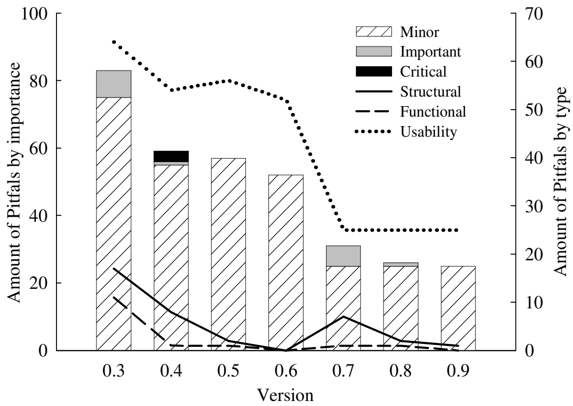

Fig. 5. Amount of Pitfalls by importance and by type, found in each version of Fire Ontology, using OOPS!.

and completeness. To perform such evaluation online questionnaires were developed, one for each domain ontology.

Likert scale and open ended questions types were included in the questionnaires. The questionnaires were sent by e-mail to experts in ecology and ontologies. Along with the questionnaires we provided the full documentation of the ontologies. These documents were relevant, since the experts could use them to inspect with details the structure and the concepts used in the ontologies.

Six and five researchers answered the questionnaires about Ccon ontology and Fire ontology, respectively, from which three are experts in ecology and three are experts in computer science.

In general, the experts are in accordance with most of the concepts used and they also suggested small changes, most of which were related to add or change terms. To test the changes, it would be required to

[6] http://oeg-lia3.dia.fi.upm.es/oops/index.jsp

reorganize and redo part of the ontologies, what we plan to perform in a future work.

5.2.1. Ccon responses

When the experts were asked how appropriate are the terms used in the ontology, most of them answered *highly appropriate* (33%), *appropriate* (50%) and *neutral* (17%). Asked about how appropriate the object properties and relationships are, most of the experts answered *highly appropriate* (17%), *appropriate* (50%) and *inappropriate* (33%). With regard to whether the concepts used are sufficient to describe the dynamics of plant communities in the Cerrado vegetation, 33% *fully agree*, 33% *agree* and 33% *disagree*.

Most comments pointed out to few missing concepts, such as: soil characteristics, soil depth, and soil moisture, to include more physiognomy types, such as moist fields, gallery forests, deciduous and semi-deciduous forests. It was suggested to divide the Diversity class into the following subtypes: structural, functional and compositional diversity, and few changes in how the classes and subclasses are organized.

5.2.2. Fire responses

When asked whether the concepts presented describe correctly the characteristics of a fire event, 25% *fully agree*, 50% *agree,* and 25% were neutral. When asked how appropriate the concepts used are, 50% of the experts answered they are *highly appropriate* and 50% answered *appropriate*. Asked about the adequacy of the relationships 25% answered the relations are *highly appropriate* and 75% of the experts answered the relations are *appropriate*. With regard to whether the concepts used are sufficient to describe the fire occurrence over the Cerrado vegetation, 25% of the experts *fully agree*, 50% *agree* and 25% were *neutral*.

In general, these results pointed out that the concepts and properties used in both domain ontologies developed in this work were considered appropriate and sufficient, moreover further improvements can be implemented.

6. Linked data generation and publishing

6.1. Dataset identification and transformation

Open data (Table 1) was collected from Brazilian government agencies such as INMET[7] (Brazilian National Institute of Meteorology), INPE[8] (Brazilian National Institute of Space Research) and IBGE[9] (Brazilian Institute of Geography and Statistics). Data about wood plant species occurrence and abundance were collected from different studies published in scientific journals, and MSc and PhD dissertations.

The datasets were found in different formats, such as spreadsheets, text files and shape files. In order to perform transformations of the datasets into RDF we used the Open Refine[10] with its RDF extension[11]. We also used geometry2RDF to transform shape files into RDF.

Table 1. The datasets used, their original format and the tools used to transform them into RDF format.

Dataset	Provenance	File format
Vegetation Dynamics	Scientific studies	Spreadsheet
Meteorological data	INMET	Text file
Fire Occurrence	INPE	Text file
Brazilian Soils Map	IBGE	Shape file
Brazilian Vegetation Map	IBGE	Shape file
Brazilian Biomes Map	IBGE	Shape file
Remaining areas of Cerrado Biome Map	IBGE	Shape file

6.2. URI Design

The base URI is *http://cerrado.linkeddata.es/*, this segment is common to all elements in the knowledge base. TBox, that is, classes and properties, and ABox, the instances, components were separated into two URI schemes, *http://cerrado.linkeddata.es/ecology/* and *http://cerrado.linkeddata.es/resource/* respectively.

The URI scheme of TBox components were separated into *http://cerrado.linkeddata.es/ecology/ccon/* and *http://cerrado.linkeddata.es/ecology/fire/* in order to specify elements that belong to Ccon and Fire ontology, respectively.

[7] http://www.inmet.gov.br/
[8] http://www.inpe.br/
[9] http://www.ibge.gov.br/
[10] http://openrefine.org/
[11] http://refine.deri.ie/

6.3. Linking datasets

To link the datasets we used an alignment approach based on the correct identification of *owl:sameAs* relations between entities. This procedure was performed using the Silk link discovery framework[12]. After that, the *sameAs* Link Validator tool[13] was used in order to validate the relationships discovered. Using this method, we enriched reference information (geometry) with data.

6.4. Publication and exploitation

In order to publish the RDF data on the web we used Virtuoso[14] and Pubby[15].

We deployed a web based application to provide a proper visualization experience of the aggregated information. Its interface integrates map visualization using Google Maps API with facet browsing and uses the Map4RDF environment. The application renders and displays on the map the distinct geometrical shapes of the features published as RDF. Moreover statistical data can be displayed on the map so that the user can observe and compare.

7. Related work

LinkedGeoData [20] was one of the first initiatives to enrich the Web of Data with spatial information. This work was responsible for the transformation and publication of the OpenStreetMap data according to the Linked Data principles, creating links with DBpedia and Geonames, and enabling new applications such as geo-data syndication or semantic-spatial searches [39].

The present study is based on the work performed by GeoLinkedData.es, an open initiative whose aim is to enrich the Web of Data with Spanish geospatial data [40], publishing diverse information sources belonging to the National Geographic Institute of Spain. The primary focus of GeoLinkedData.es was on hydrography features, but different statistical variables such as unemployment, population, dwelling, industry, and building trade were added and linked later on. Such data is provided by the National Statistics Institute in Spain [21]. The GeoLinkedData.es initiative also developed the geometry2RDF, a tool to transform shape files into RDF and Map4RDF a faceted browser to visualize and explore RDF geospatial data sources [21].

However those initiatives did not include information about biodiversity datasets. Linking datasets of occurrence and abundance of species based on scientific records was not on the scope of those woks. The present work may represent a starting point for not only linking species lists, but also associating them to geospatial information and other sources of environmental data.

The Linked Brazilian Amazon Rainforest Data initiative aimed to describe remote sensing observation data in the ecology domain about the Brazilian Amazon Rainforest, published as Linked Spatiotemporal Data [19]. Moreover, they show how these data can be further accessed and analyzed using the R statistical computing environment providing a tutorial to explore the data from R and plot it in maps. It uses semantic techniques to publish and explore data about deforestation in a biodiverse threatened biome and is a valuable application. Nevertheless, it does not provide an explicit visualization of the geospatial information.

LEAPS - Linked Entities for Algal Plant Sites [38] is a suite of linked datasets that collectively enable the evaluation of the potential of algal biomass production sites in North West Europe. The algal biomass production is a critical ecological issue in many regions and their work enables the screening of data and provides base data for more detailed planning purposes. The main limitation of this work is the low number of outgoing links it provides into other datasets [38].

8. Conclusion

This pilot study explores Brazilian maps and geographical information, linked with data about meteorological variables, wood plant species occurrence and dynamics, and fire occurrence, data provided by Brazilian governmental agencies and scientific studies.

A methodology based on a linked geographical and statistical data approach, successfully applied by [20,39–41] was adopted to create a case study aiming investigate the application of linked data principles to ecology.

A relevant question to be investigated in this pilot study is how to integrate geographical information, environmental and meteorological variables with

---

[12] http://www4.wiwiss.fu-berlin.de/bizer/silk/
[13] http://oeg-dev.dia.fi.upm.es:8080/sameAs/
[14] http://virtuoso.openlinksw.com/
[15] http://www4.wiwiss.fu-berlin.de/pubby

species occurrence and dynamics studies, and then be able to reason with the data and make inferences, and finally to show the results.

The present work contributes to: (1) Formalize into ontologies knowledge about Cerrado physiognomies, wood plant species dynamics and fire regimes. (2) Create links between data of different studies from distinct locations adding geospatial information to wood plant species occurrence and dynamics. (3) Create links between species occurrence of these studies and environmental variables, such as soil types, temperature, precipitation, fire events and deforestation. (4) Enable an alternative method to search for and detect biodiversity patterns. (5) Enable publishing and linking biodiversity data not only in terms of species occurrence, but also in terms of number of individuals over different periods of time. (6) Enable to approach questions with regard tree cover in Cerrado biome using a visual tool including geospatial dimension.

The topics addressed in the research described here have the potential to boost both applications of geolinked data technologies to new areas, and to open new perspectives for research involving ecological data management, integration and use. Ongoing work aims to add and integrate Qualitative Reasoning ecological models [4] to be linked to these datasets.

Acknowledgements

We would like to acknowledge CAPES-PDSE (*Programa de Doutorado Sanduíche no Exterior*) for the doctoral fellowship granted to the first author. AS and PS are grateful for the stimulating and supportive environment provided by the Ontology Engineering Group (OEG) at UPM.

References

[1] G. Atemezing, O. Corcho, D. Garijo, J. Mora, M. Poveda-villalón, P. Rozas, D. Vila-Suero, and B. Villazón-Terrazas, "Transforming meteorological data into Linked Data," *Semant. Web*, vol. 4, no. 3, pp. 285–290, 2013.

[2] J. B. L. Bard and S. Y. Rhee, "Ontologies in biology: design, applications and future challenges.," *Nat. Rev. Genet.*, vol. 5, no. 3, pp. 213–22, Mar. 2004.

[3] R. Battle and D. Kolas, "Enabling the geospatial semantic web with Parliament and GeoSPARQL," *Semant. Web*, vol. 3, no. 4, pp. 355–370, 2012.

[4] B. Bredeweg and P. Salles, "Qualitative models of ecological systems," *Ecol. Inform.*, vol. 4, no. 5–6, pp. 261–262, Nov. 2009.

[5] S. Bridgewater, A. Ibanez, J. A. Ratter, and P. Furley, "Vegetation classification and floristics of the savannas and associated wetlands of the Rio Bravo Conservation and Management Area, Belize," *Edinburgh J. Bot.*, vol. 59, no. 3, pp. 421–442, 2002.

[6] V. Brilhante, "Ecolingua: a formal ontology for data in ecology," *J. Brazilian Comput. Soc.*, vol. 11, no. 2, pp. 61–78, 2005.

[7] L. Coutinho, "O conceito de cerrado," *Rev. Bras. Botânica*, pp. 17–23, 1978.

[8] L. Coutinho, "Ecological effects of fire in Brazilian cerrado," *Ecol. Trop. savannas*, 1982.

[9] R. Cyganiak and D. Reynolds, "The RDF Data Cube Vocabulary," 2013. [Online]. Available: http://www.w3.org/TR/vocab-data-cube/.

[10] J. M. Felfili, M. C. da Silva Júnior, A. C. Sevilha, C. W. Fagg, B. M. Teles Walter, P. E. Nogueira, and A. V. Rezende, "Diversity, floristic and structural patterns of cerrado vegetation in Central Brazil," *Plant Ecol. Former. `Vegetatio'*, vol. 175, no. 1, pp. 37–46, Nov. 2004.

[11] R. Goodland, "A Physiognomic Analysis of theCerrado'Vegetation of Central Brasil," *J. Ecol.*, vol. 59, no. 2, pp. 411–419, 1971.

[12] J. L. Green, A. Hastings, P. Arzberger, F. J. Ayala, K. L. Cottingham, K. I. M. Cuddington, F. Davis, J. a. Dunne, M. Fortin, L. Gerber, and M. Neubert, "Complexity in Ecology and Conservation: Mathematical, Statistical, and Computational Challenges," *Bioscience*, vol. 55, no. 6, pp. 501–510, 2005.

[13] R. Henriques and J. Hay, "Patterns and dynamics of plant populations," in *The Cerrados of Brazil: ecology and natural history of a …*, no. chapter 5, P. S. Oliveira and R. J. Marquis, Eds. Columbia University Press, 2002, pp. 140–158.

[14] R. P. B. Henriques, "Influência da história, solo e fogo na distribuição e dinâmica das fitofisionomias no bioma do Cerrado," in *Cerrado: Ecologia, Biodiversidade e Conservaçao*, A. Scariot, J. C. Sousa-Silva, and J. M. Felfili, Eds. Brasília, DF: Ministério do Meio Ambiente, 2005, pp. 73–92.

[15] J. R. Hobbs and F. Pan, "Time Ontology in OWL," 2006. [Online]. Available: http://www.w3.org/TR/owl-time/.

[16] W. A. Hoffmann, R. Adasme, M. Haridasan, M. T. de Carvalho, E. L. Geiger, M. A. B. Pereira, S. G. Gotsch, and A. C. Franco, "Tree topkill, not mortality, governs the dynamics of savanna-forest boundaries under frequent fire in central Brazil.," *Ecology*, vol. 90, no. 5, pp. 1326–37, May 2009.

[17] B. Hyland, G. Atemezing, and B. Villazón-Terrazas, "Best Practices for Publishing Linked Data," 2014. [Online]. Available: http://www.w3.org/TR/ld-bp/.

[18] J. Jones and A. Sanchez, "Open Linked Amazon Vocabulary Specification," 2012. [Online]. Available: http://observedchange.com/amazon/ns/.

[19] T. Kauppinen, G. M. de Espindola, J. Jones, A. Sánchez, B. Gräler, and T. Bartoschek, "Linked brazilian amazon rainforest data," *Semant. Web J.*, vol. 5, no. 2, p. (in press), 2014.

[20] J. Lehmann, S. Hellmann, and S. Auer, "LinkedGeoData: Adding a spatial dimension to the Web of Data," *Semant. Web-ISWC 2009*, 2009.

[21] A. De León, V. Saquicela, L. M. Vilches-Blázquez, B. Villazón-terrazas, F. Priyatna, A. De León, and O. Corcho,

"Geographical linked data: a Spanish use case," in *Proceedings of the In I-SEMANTICS '10 6th International Conference on Semantic Systems*, 2010, pp. 1–3.

[22] C. A. Lepczyk, C. J. Lortie, and L. J. Anderson, "An ontology for landscapes," *Ecol. Complex.*, vol. 5, no. 3, pp. 272–279, Sep. 2008.

[23] J. S. Madin, S. Bowers, M. Schildhauer, S. Krivov, D. Pennington, and F. Villa, "An ontology for describing and synthesizing ecological observation data," *Ecol. Inform.*, vol. 2, no. 3, pp. 279–296, Oct. 2007.

[24] K. J. Metzger, R. Klaper, and M. a. Thomas, "Implications of informatics approaches in ecological research," *Ecol. Inform.*, vol. 6, no. 1, pp. 4–12, Jan. 2011.

[25] Ministério do Meio Ambiente - MMA and Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis - IBAMA, "Monitoramento do desmatamento nos biomas brasileiros por satélite: Monitoramento do bioma Cerrado 2008-2009," Brasília, DF, 2011.

[26] H. S. Miranda, M. N. Sato, W. Nascimento Neto, and F. S. Aires, "Fires in the cerrado, the Brazilian savanna," in *Tropical Fire Ecology: Climate Change, Land Use, and Ecosystem Dynamics*, vol. 2, Springer Berlin Heidelberg, 2009, pp. 427–450.

[27] A. G. Moreira, "Fire Protection and Vegetation Dynamics in the Brazilian Cerrado," Harvard University, 1992.

[28] N. Myers, R. Mittermeier, C. G. Mittermeier, G. A. B. Fonseca, and J. Kent, "Biodiversity hotspots for conservation priorities," *Nature*, vol. 403, no. 6772, pp. 853–8, Feb. 2000.

[29] A. Oliveira-Filho and J. Ratter, "Vegetation physiognomies and woody flora of the cerrado biome," in *The Cerrados of Brazil. Ecology and natural history of a neotropical savanna*, P. S. Oliveira and R. J. Marquis, Eds. New York: Columbia University Press, 2002, pp. 91–120.

[30] M. Perry and J. Herring, "OGC GeoSPARQL-A geographic query language for RDF data," 2012.

[31] V. R. Pivello, "Fire management for biological conservation in the Brazilian Cerrado," in *Savannas and dry forests: linking people with nature*, J. Mistry and A. Berardi, Eds. 2006.

[32] V. R. Pivello, "The Use of Fire in the Cerrado and Amazonian Rainforests of Brazil: Past and Present," *Fire Ecol.*, vol. 7, no. 1, pp. 24–39, Apr. 2011.

[33] M. Poveda, "The Aemet Network of Ontologies," 2011. [Online]. Available: http://aemet.linkeddata.es/ontology/.

[34] M. Poveda-Villalón, M. Suárez-Figueroa, and A. Gómez-Pérez, "Validating Ontologies with OOPS!," in *Knowledge Engineering and Knowledge Management SE - 24*, vol. 7603, A. Teije, J. Völker, S. Handschuh, H. Stuckenschmidt, M. d'Acquin, A. Nikolov, N. Aussenac-Gilles, and N. Hernandez, Eds. Springer Berlin Heidelberg, 2012, pp. 267–281.

[35] J. A. Ratter, S. Bridgewater, and J. F. Ribeiro, "Analysis of the Floristic Composition of the Brazilian Cerrado Vegetation III: Comparison of the Woody Vegetation of 376 Areas," *Edinburgh J. Bot.*, vol. 60, no. 01, pp. 57–109, Mar. 2003.

[36] J. Ratter, J. Ribeiro, and S. Bridgewater, "The Brazilian cerrado vegetation and threats to its biodiversity," *Ann. Bot.*, vol. 80, no. 3, pp. 223–230, 1997.

[37] M. F. Simon, R. Grether, L. P. de Queiroz, C. Skema, R. T. Pennington, C. E. Hughes, and L. P. De Queiroz, "Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 48, pp. 20359–64, Dec. 2009.

[38] M. Solanki, J. Skarka, and C. Chapman, "Linked data for potential algal biomass production," *Semant. Web*, vol. 4, no. 3, pp. 331–340, 2013.

[39] C. Stadler, J. Lehmann, K. Höffner, and S. Auer, "Linkedgeodata: A core for a web of spatial open data," *Semant. Web J.*, vol. 3, no. 4, pp. 333–354, 2012.

[40] L. M. Vilches-Blázquez, B. Villazón-Terrazas, O. Corcho, and A. Gómez-Pérez, "Integrating geographical information in the Linked Digital Earth," *Int. J. Digit. Earth*, no. April 2013, pp. 1–22, Apr. 2013.

[41] L. M. Vilches-blázquez, B. Villazón-terrazas, A. De Leon, F. Priyatna, and O. Corcho, "An Approach to Publish Spatial Data on the Web : The GeoLinked Data Case."

[42] L. M. Vilches-Blázquez, B. Villazón-Terrazas, V. Saquicela, A. de León, O. Corcho, and A. Gómez-Pérez, "GeoLinked data and INSPIRE through an application case," *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geogr. Inf. Syst. - GIS '10*, p. 446, 2010.

[43] D. Vrandečić, "Ontology evaluation," *Handb. Ontol.*, pp. 293–313, 2009.

[44] W3C OWL Working Group, "OWL 2 Web Ontology Language Document Overview (Second Edition)," 2012. [Online]. Available: http://www.w3.org/TR/owl2-overview/.

[45] R. J. Williams, N. D. N. D. Martinez, and J. Golbeck, "Ontologies for ecoinformatics," *J. Web Semant.*, vol. 4, no. 4, pp. 237–242, Dec. 2006.

[46] "The National Center for Biomedical Ontology." [Online]. Available: http://www.bioontology.org/. [Accessed: 30-Aug-2013].

[47] "Environment Ontology," *EnvO*. [Online]. Available: http://environmentontology.org/. [Accessed: 30-Sep-2013].