International Cartographic Association
Association Cartographique Internationale

From Conceptual Models to Ontologies: Achieving Knowledge Management in SDIs

Fabíola Andrade Souza [a,b,*], Leonardo Scharth Loureiro Silva [b,c,d], Adriana Alexandria Machado [b], Sabrina de Andrade Oliveira Santos [e], Érika do Carmo Cerqueira [e], Silvana Philippi Camboim [b]

[a] *Polytechnic School, Federal University of Bahia (UFBA), Salvador-BA, Brazil - fabiola.andrade@ufba.br*
[b] *Graduate Program in Geodetic Sciences (PPGCG), Federal University of Paraná (UFPR), Curitiba-PR, Brazil – adri.alexandria@gmail.com; silvanacamboim@ufpr.br*
[c] *Brazilian Institute of Geography and Statistics (IBGE), Rio de Janeiro-RJ, Brazil – leonardo.l.silva@ibge.gov.br*
[d] *Department of Geoenvironmental Analysis (GAG), Federal Fluminense University (UFF), Niterói-RJ, Brazil.*
[e] *Department of Geosciences, Federal University of Bahia (UFBA), Salvador-BA, Brazil - sabrina.andradeoliveira@gmail.com; erika.cerqueira@ufba.br*

* Corresponding author

Keywords: ontology, semantic web, geospatial data models.

Abstract:

Despite their role in consolidating legal and technological frameworks for data sharing, Spatial Data Infrastructures (SDIs) are still adapting to an era of management by intelligent agents and adequate circulation of spatial knowledge (Varanka, 2021). Such initiatives must evolve to achieve their role in the current global challenges scenario, and a vital aspect is knowledge management and semantic interoperability. While an essential step in creating SDIs has been the development of agreed conceptual models, as in the ET-EDGV, the Brazilian standard for topographic data, the current scenario highlights the need to harmonise such concepts between models. Incidentally, the semantic web has established itself as a technological paradigm to advance the compatibility of meanings between heterogeneous sources. In Big Data, such advance is essential because if the volume and velocity of data are large, there is minimal space for manual compatibilization. If the variety is significant, even more emphasis on harmonisation is imperative.

This work extends an existing semantic harmonisation between the ET-EDGV and the conceptual model of OpenStreetMap. Firstly, the choice of such models is due to the pressing need to use collaborative data to keep up to date with the fragile cartographic coverage in Brazil. Machado (2020) performed the first semantic alignment, and Silva (2022) extended the study by pointing out and detailing the 15 classes with the most significant potential for data usage and included QGIS scripts to import data from OSM in a database in the ET-EDGV model. Key issues for semantic interoperability include how conceptual models are represented and how they relate. Formally describing the semantics of geospatial terms and sharing their meanings is still challenging (Janowicz et al., 2013). Using ontologies provides tools to understand the geographic universe and facilitate information exchange and knowledge reuse. Thus, research in ontology compatibility is promising for interoperability. Geospatial data fusion and the development of context-embedded bases are also approaches that contribute significantly (Novack et al., 2019).

Therefore, to show the potential of such integration, we illustrate here the number of features that can update the official local government database in Salvador-Bahia. These initial tests indicated the need to improve the class correspondences, including at a geometry level. For example, substation, helipad and pitch in ET-EDGV are presented as polygons, while in the OSM, they are represented by points. So, there was a need for manual treatment, converting the polygons to their centroids before import. In addition, road data are separated into "trecho_arruamento" (internal urban road) and "trecho_rodoviario" (intercity roads) and were manually unified before importing (to associate as "highway"). The results in Table 1 show that, even so, the feature number increase is substantial, reinforcing the importance of this data-updating strategy based on diverse sources.

However, the key to this importation is in a long table of correspondences created based on models, legislations, and dictionaries. Moreover, this strategy demands non-automated manipulation and analysis for processing the data import scripts, which is inadequate for search by semantic web mechanisms. This work proposes a step forward by implementing such a knowledge base as an ontology. The goal is to extend the capacity for automating the construction of this knowledge base, taking it out of the relational sphere and bringing it into a semantic structure. This structure is a first step so that in future developments, it could be possible to develop automation mechanisms for the conceptual integration of different models. In order to achieve this goal, ontologies must be formalised through specific software and, whenever possible, aligned with other bases that present similar concepts, expanding the possibilities.

Class	Official Base existent features	OSM new features	Total
Substation	23	24	47
Park	311	291	602
Pitch	2522	212	2734
Helipad	9	19	28
Highway	26971	3002	29973

Table 1. Number of features in the official database and imported from OSM.

Improved correspondence between classes can occur using ontologies for representing the models. For example, we used the WebProtégé (open-source OWL editor) to represent the semantic alignment between ET-EDGV and OSM models through knowledge graphs for 5 of the 15 classes that were pointed out by Silva (2022): substation, park, pitch, helipad, and motorway (part of highway). As a result, the initial studies produced semantic alignment graphs (Figure 1). Classes and subclasses in the same model are related through the yellow arrow; semantic alignment between models is represented by the blue arrows "correspondsTo". However, the partial results indicate the need for evolution in representing other model elements in the ontology, such as geometries and spatial relations between objects.

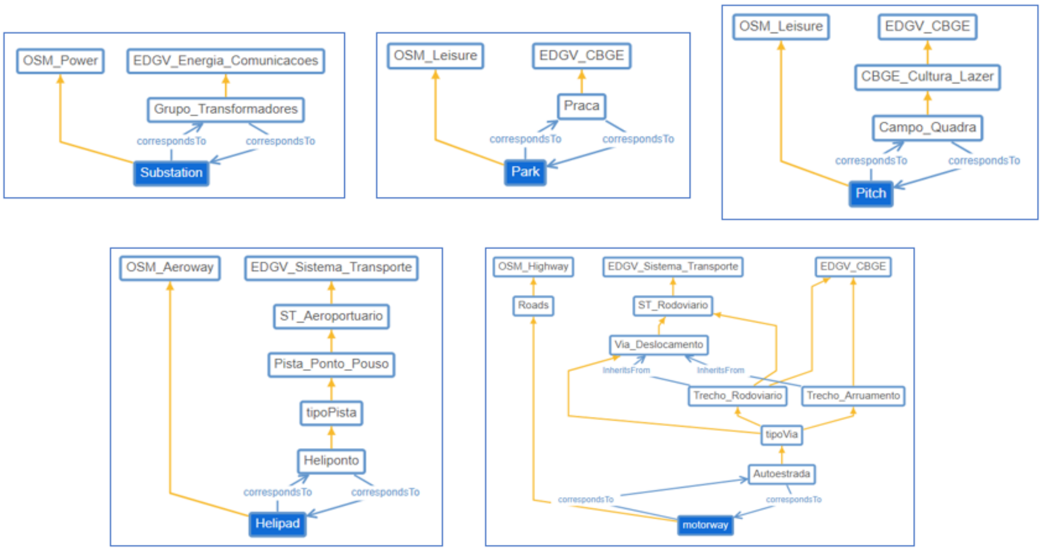

Figure 1. Graphs of the semantic alignment for substation, park, pitch, helipad and motorway (ET-EDGV classes in Portuguese).

Here we report the experience in development to implement semantic tools in a large and diverse country like Brazil. The evolution of this work intends to use the represented ontologies for automated integration between the models, seeking to overcome the problems of lack of data compatibility and to turn to the construction of decentralised and open solutions to exchange spatial knowledge.

References

Janowicz K.; Scheider S.; Adams B. A Geo-semantics Flyby. In: (Eds) Rudolph S.; Gottlob G.; Horrocks, I.; Van Harmelen; F. *Reasoning Web: Semantic Technologies for Intelligent Data Access.* Reasoning Web 2013. Lecture Notes in Computer Science, v. 8067. Springer, Berlin, Heidelberg, p. 230-250, 2013.

Machado, A. A. *Compatibilização Semântica entre o Modelo de Dados do Openstreetmap e a Especificação Técnica para Estruturação de Dados Geoespaciais Vetoriais (ET-EDGV)*. PhD Thesis in Geodetic Sciences. Universidade Federal do Paraná. Curitiba-PR-Brazil. 2020.

Novack, T.; Grinberger, A. Y.; Schultz, M.; Zipf, A.; Mooney, P. The Geographical and Cultural Aspects of Geo-Information: An Introduction. Proceedings of the *"Geographical and Cultural Aspects of Geo-Information: Issues and Solutions"*. AGILE 2019 Workshop, Limassol, Cyprus, June 17th, 2019.

Silva, L. S. L. *Integração de Dados Provenientes de Mapeamento Colaborativo na Cartografia de Referência do Brasil.* PhD Thesis in Geodetic Sciences. Universidade Federal do Paraná. Curitiba-PR-Brazil. 2022.

Varanka, D. E. (2021). *Data Properties. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2021 Edition)*, John P. Wilson (Ed.). DOI: 10.22224/gistbok/2021.1.15.

Abstracts of the International Cartographic Association, 6, 243, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-243-2023 |