

LinkMapia: Uma Abordagem para Converter Dados Geográficos Livrementemente Anotados em Dados Ligados

Juarez A. P. Sacenti¹, Renato Fileto¹

¹Dep. de Informática e Estatística – Universidade Federal do Santa Catarina (UFSC)
Caixa Postal 476 – 88040-900 – Florianópolis – SC – Brazil

juarezsacenti@inf.ufsc.br, r.fileto@ufsc.br

Abstract. *Collecting geographic data by using satellites, global positioning system (GPS) and even crowdsourcing with free annotations often gathers data with little semantic, with ambiguities and little interoperability. This paper proposes a process to convert textual annotated geographic data to a linked geographic data collection. This proposal is based in a process that filters and cleans data, and then applies several techniques to align this data with existing linked data collection. Experiments applying the proposed process to Wikimapia and Linked Geo Data data generate linked data that support several useful SPARQL queries.*

Resumo. *A coleta de dados geográficos, utilizando tecnologias de sensoriamento (e.g., GPS, câmeras) e mesmo sistemas de anotação colaborativa, frequentemente geram dados com semântica insuficiente, com ambiguidades e pouca interoperabilidade. Este trabalho propõe um processo para converter dados geográficos anotados textualmente em dados ligados. A proposta baseia-se em um processo que filtra e limpa dados, e então aplica diversas técnicas para alinhar estes dados a coleções existentes de dados ligados. A aplicação do processo proposto aos dados do Wikimapia e Linked Geo Data em experimentos gerou dados ligados que suportam diversas consultas SPARQL úteis.*

1. Introdução

Dados associados a características espaciais, que referenciam uma localização na superfície da Terra, são chamados de dados geográficos [Casanova et al. 2005]. O espaço, ocupado por elementos (e. g., construções, rios e estradas) e fenômenos (e. g., massas de ar, dissiminação de doenças e temperatura), é representado computacionalmente por formas vetoriais (e. g., pontos, retas, polígonos) ou matriciais (e. g., *raster*). Estas representações são os atributos espaciais dos dados geográficos.

Porém, dados geográficos possuem informações alfanuméricas, tão importantes quanto atributos espaciais: os atributos descritivos [Rigaux et al. 2000]. Descrever dados geográficos depende da interpretação humana do elemento ou fenômeno representado. A coleta de dados geográficos utilizando tecnologias de sensoriamento (e.g., GPS, câmeras) é insuficiente para levantar estas informações.

Sistemas de coleta colaborativa de dados geográficos (e. g., Wikimapia¹, *Open Street Map*² - OSM) são providos de recursos para permitir principalmente a coleta de

¹Disponível em: <http://wikimapia.org/about>. Acesso em mar. 2014.

²Disponível em: <http://www.openstreetmap.org/about>. Acesso em mar. 2014.

atributos descritivos, geralmente de forma voluntária e via *web*. Entretanto, dados anotados desta forma, *i. e.*, *volunteer geographic information* (VGI) [Goodchild 2007], geralmente não apresentam semântica suficientemente formal para ser utilizada computacionalmente e suficientemente detalhada para suprir as necessidades das aplicações. As anotações livres, *i. e.*, anotações que não apresentam nenhum tipo de estruturação explícita de seu conteúdo (*e. g.*, vocabulário comum, glossário), herdam os problemas da manipulação de textos livres: sinônimos, ambiguidades e erros ortográficos.

Nestes sistemas sociais de anotação (*Social Tagging Systems* - STS), algumas anotações livres são utilizadas para posteriormente recuperar os objetos anotados. A estrutura conceitual destas anotações, emergente de STSs, é chamada de *folksonomia* [García-castro and García 2011].

A adoção e manutenção de convenções de anotação, *e. g.*, enciclopédias (*thesaurus*), glossários, dicionários geográficos (*gazetters*), em sistemas com grande número de voluntários são custosas. As soluções da *web* semântica, como as anotações semânticas [Berners-Lee et al. 2001], ontologias [Guarino 1998] e dados ligados (*linked data*) [Heath and Bizer 2011], podem contribuir para resolver os problemas de ambiguidade e interoperabilidade das anotações colaborativas.

Para obter dados ligados de um sistema colaborativo, é necessário converter a *folksonomia* em ontologia. Este trabalho propõe um processo para conversão de dados geográficos textualmente anotados em dados ligados, e aplica-o a dados do projeto Wikimapia, um sistema colaborativo ainda não conectado a *web* de dados.

O restante deste artigo é disposto em diferentes seções. A seção 2 descreve o processo proposto. A seção 3 apresenta a implementação de cada passo deste processo. A seção 4 vislumbra a implicação do enriquecimento semântico na recuperação dos dados geográficos. A seção 5 relata alguns trabalhos correlacionados, que discorrem sobre VGI, dados geográficos ligados e alinhamento ontológico. Finalmente, a seção 6 apresenta conclusão e trabalhos futuros, acrescido de agradecimentos.

2. Proposta

Este trabalho define um processo de conversão de dados geográficos livremente anotados em dados ligados, composto por cinco etapas ilustradas na figura 1. Este processo não apenas triplifica os dados geográficos como também cria uma ontologia a partir dos termos usados na categorização destes dados, correlacionando suas categorias com entidades de ontologias como Linked Geo Data³ (LGD) e GeoNames⁴.

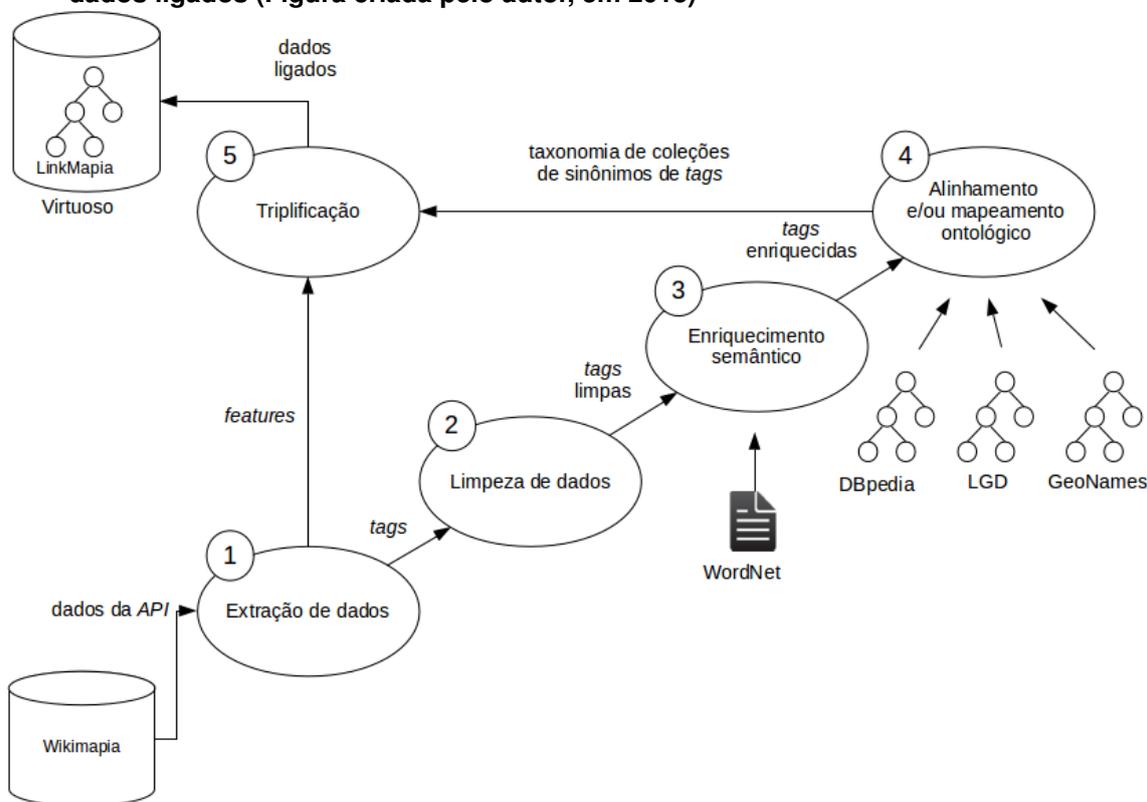
A primeira etapa deste processo realiza a extração (ou coleta) de objetos geográficos (*features*) de sistemas de anotação livre de dados geográficos, como os sistemas colaborativos OSM e Wikimapia. As anotações destes sistemas são utilizadas para categorizar e recuperar objetos geográficos. A extração de categorias é necessária para elaborar uma ontologia dos objetos extraídos.

Na segunda etapa, os nomes de categorias são tratados lexicamente. Nomes compostos são separados em termos (*tokenização*). A remoção de *stopwords* exclui termos

³Disponível em: <http://linkedgeo.org/About>. Acesso em nov. 2013.

⁴Disponível em: <http://www.geonames.org/>. Acesso em nov. 2013.

Figure 1. Processo proposto - conversão de dados geográficos anotados em dados ligados (Figura criada pelo autor, em 2013)



semanticamente irrelevantes (*e.g.*, artigos, preposições). O passo de *stemming* reduz lexicamente termos restantes a radicais. O tratamento léxico é necessário para obter melhores resultados em comparações sintáticas.

Na terceira etapa, os termos são comparados sintaticamente a termos de dicionários léxicos, como o WordNet⁵. O enriquecimento semântico explicita o conceito da categoria, antes implícito nos diferentes significados das palavras anotadas. A categoria passa a ser representada por lista de conceitos (conjunto de sinônimos). Por exemplo, o termo *service* é expandido em seu conjunto de sinônimos:

[*service, work, assist, assistance, help, activity, care, maintenance, upkeep*]

A quarta etapa, o alinhamento ou mapeamento ontológico, relaciona as categorias com classes de uma dada ontologia geográfica (*e.g.*, GeoNames, LGD). As categorias (lista de conceitos) são comparadas com os rótulos (*label*) de classes da ontologia. Esta comparação é inspirada do processo de alinhamento ontológico LOM [Li 2004].

LOM é uma técnica de *matching* de ontologias que considera apenas a representação léxica (rótulos) de classes, desconsiderando a estrutura relacional da ontologia. Deste modo, é possível adaptar LOM para o alinhamento de termos da *folksonomia* para conceitos de ontologias. Este processo identifica conceitos equivalentes, que facilitam a conversão da *folksonomia* em uma nova ontologia e integra-a a ontologias existentes.

⁵<http://wordnet.princeton.edu/>. Acesso em nov. 2013.

LOM compara os conjuntos dos rótulos dos conceitos das duas ontologias que se pretende alinhar. Tendo em mãos os dois conjuntos, são aplicados quatro processos de pareamento de conceitos: pareamento de termo completo, pareamento de termos (*tokens*) que constituem os termos completos, pareamento de conjunto de sinônimos correspondentes aos termos completos e pareamento de tipos (não explorado nesse trabalho).

Figure 2. Fluxo de pareamento entre *folksonomia* e ontologia (Figura criada pelos autores, em 2013)

LOM Adaptado - Etapas de comparação

1. Matching de termo completo (tags).
BANK = bank
2. Matching das palavras (tokens) que compõem os termos.
BANKS -> 'banks' x 'bank' <- bank = 0 %
BANKS -> 'banks' x 'saving' + 'bank' <- saving bank = 0 %
3. Matching das stems (radicais dos tokens) que compõem os termos.
BANKS -> 'bank' x 'bank' <- bank = 100 %
BANKS -> 'bank' x 'save' + 'bank' <- saving bank = 50 %
4. Matching dos conjuntos de sinônimos das palavras.
STREAM BANK -> 'stream' + 'bank' 'bank' <- bank
synset:stream + synset:bank x synset: bank = 50 %

Nesta proposta, o pareamento de categoria e conceitos é realizado por diferentes níveis: termo completo, *tokens*, *stemming tokens* e conjunto de sinônimos (figura 2). O pareamento de termo completo envolve a comparação do texto sem tratamento. Em caso de igualdade o alinhamento é positivo e os termos envolvidos removidos das listas de candidatos a alinhamento.

Os rótulos de cada classe devem seguir os passos da etapa de limpeza para realizar os pareamentos seguintes. O pareamento de *stemming tokens* é adicionado a adaptação do LOM em virtude de radicais possuírem maior grau de generalização, facilitando a comparação léxica. Ambas as etapas utilizam uma métrica para avaliar a similaridade entre categorias da *folksonomia* (A) e conceitos de ontologia (B), tal qual:

$$sim(A, B) = \frac{quantidade_termos_iguais}{tamanho_maior_conjunto(A,B)}$$

A comparação léxica pode utilizar métricas de similaridade léxica (*e.g.*, *soft TFIDF* [Cohen et al. 2003]). A categoria mais similar (categoria alinhada) a uma classe é associada a esta classe por meio da URI da classe na ontologia. As categorias menos similares àquela classe, que respeitem um limite inferior de similaridade, são relacionadas à classe alinhada como conceitos próximos.

Na quinta e última etapa, triplificação, tanto os objetos geográficos quanto as categorias são convertidas para o formato *Resource Description Framework*⁶ (RDF). Primeiro, as categorias são convertidas em classes gerando uma ontologia rasa (*i.e.*, uma árvore de altitude igual a 1, cuja raiz é a classe *Place*). Segundo, classes de categorias

⁶<http://www.w3.org/RDF/>. Acesso em mar. 2014.

alinhadas a classes externas são ligadas a estas classes. Terceiro, conceitos próximos encontrados pelo alinhamento são ligados as classes. Finalmente, os objetos geográficos são convertidos em RDF e associados aos conceitos (pelas URIs) da ontologia formada.

Os dados podem então ser publicados em *end-points* SPARQL⁷ (e.g., Strabon [Kyzirakos et al. 2012], Parliament⁸ e Virtuoso⁹), possibilitando consultas integradas na coleção de dados ligados gerada e nas ontologias externas.

3. Implementação

Nesta seção é ilustrado o desenvolvimento do protótipo do processo proposto para a conversão de dados geográficos livremente anotados em dados ligados. O sistema de coleta colaborativa de dados geográficos deste caso de estudo é o Wikimapia, um projeto cujo objetivo é criar e manter um mapa atualizado, completo e multilíngue de todo o mundo, com potencial para enriquecer o conteúdo da *web* de dados geográficos (LGD, GeoNames, GeoLinkedData¹⁰, dentre outros). Os objetos geográficos deste caso de estudo são restritos aos limites da cidade de Milão, Itália, para análises espaço-temporais de trabalhos futuros. Contudo, o protótipo permite a conversão de dados de qualquer localidade.

Extração de dados. A etapa de extração de dados pode ser dividida na extração dos objetos geográficos e das categorias utilizadas na classificação destes objetos. Este conjunto de categorias forma a *folksonomia* do Wikimapia.

Objetos Geográficos - O projeto Wikimapia disponibiliza seus dados a partir de sua *Application Programming Interface* (API)¹¹. Para realizar a extração, é necessário informar o menor retângulo que contém a cidade de Milão (*minimal bounding rectangle*) é definido pelos pontos (45.388039, 9.043907) e (45.536266, 9.278963), de latitude e longitude respectiva, segundo o polígono dos limites políticos da cidade na base de dados *Database of Global Administrative Areas*¹² (GADM).

A extração de objetos geográficos (*features*) utiliza as funções *box* e *object* da API da Wikimapia. A função *box* resultou em uma lista 1276 ids de objetos, separados por páginas, sendo a variável *count* o limite de resultados por página, em formato XML¹³ ou JSON¹⁴. A função *object* obteve a descrição detalhada de cada objeto, dado seu *id*. Os dados extraídos apresentaram 243 categorias e 1513 relações de anotação. O Sistema de Gerenciamento de Banco de Dados (SGBD) utilizado para armazenar dados extraídos foi o PostgreSQL¹⁵ 9.2.4 com a extensão geográfica PostGIS¹⁶ 2.0.3-1.

Folksonomia do Wikimapia - A função *Category.GetAll* retorna todas as categorias do Wikimapia, compondo a *folksonomia* do sistema colaborativo. A extração obteve 8615 categorias: *ids*, nomes e frequência de uso (tabela 1).

⁷<http://www.w3.org/TR/rdf-sparql-query/>. Acesso em mar. 2014.

⁸Disponível em: <<http://http://parliament.semwebcentral.org/>>. Acesso em nov. 2013.

⁹Disponível em: <<http://http://virtuoso.openlinksw.com/>>. Acesso em nov. 2013.

¹⁰Disponível em: <http://geo.linkeddata.es/web/guest/home>. Acesso em nov. 2013.

¹¹Disponível em: <http://wikimapia.org/api>. Acesso em mar. 2014.

¹²Disponível em: <http://www.gadm.org/>. Acesso em mar. 2014.

¹³Disponível em: <http://www.w3.org/XML>. Acesso em mar. 2014.

¹⁴Disponível em: <http://www.json.org/xml.html>. Acesso em mar. 2014.

¹⁵Disponível em: <http://www.postgresql.org/about/>. Acesso em mar. 2014.

¹⁶Disponível em: <http://postgis.net/>. Acesso em mar. 2014.

tados nas comparações entre termos da ontologia e da *folksonomia*, os *tokens* resultantes do processo de limpeza são submetidos como parâmetros de consulta à base do WordNet e recuperados os conjuntos de sinônimos de cada termo. A comunicação com a API do WordNet é realizada pelo conversor.

Alinhamento e/ou Mapeamento Ontológico. O alinhamento das categorias do Wikimapia com conceitos de ontologias (*e.g.*, DBpedia, GeoNames e LGD) apresentada neste trabalho é uma adaptação das idéias apresentadas no artigo sobre LOM [Li 2004]. Foram selecionadas 101 categorias sob o critério de maior número de objetos anotados. Os conceitos de ontologias selecionados foram 143 conceitos do GeoNames²⁰, já mapeados para conceitos do DBpedia e LGD. A tabela 2 mostra o número de alinhamentos obtidos nos diferentes níveis de pareamento.

RESULTADOS

NÚMERO DE CATEGORIAS UTILIZADAS DO WIKIMAPIA	101
NÚMERO DE CONCEITOS DE ONTOLOGIAS EXTERNAS	143
NÚMERO DE ALINHAMENTOS POR TEXTO ORIGINAL	15
NÚMERO DE ALINHAMENTO POR TOKENS	7
NÚMERO DE ALINHAMENTO POR STEMMING TOKENS	2
NÚMERO DE ALINHAMENTO POR SYMSET	0
TOTAL DE ALINHAMENTOS	24

Table 2. Tabela de resultados do alinhamento (tabela criada pelo autor, em 2013)

A similaridade entre rótulos é um número real entre 0 e 1. As relações de equivalência e proximidade de conceitos são obtidas classificando os resultados, baseando-se no valor de similaridade dos rótulos (*e.g.*, resultados com valor de similaridade dentro do intervalo (1; 0,9) são considerados idênticos, no intervalo [0,9; 0,5) são considerados similares, e resultados menores que 0,5 não são considerados similares).

Triplificação. O processo de transformação dos dados do Wikimapia em uma coleção RDF, chamada LinkMapia, iniciou-se com a conversão de sua *folksonomia* em uma ontologia rasa, *i.e.* uma árvore de altura 1. Foi escolhida a classe *Place* como derivada direta da classe *Thing*. Todas as categorias da *folksonomia* foram inicialmente classificadas como subclasses de *Place*. O esquema *Web Ontology Language*²¹ (OWL) foi baseado em um protótipo gerado automaticamente pela ferramenta Protege²², versão 4.3.0.

A coleção de dados resultante ainda não está ligada a outras fontes. A ontologia precisa ser correlacionada com ontologias externas como LGD e GeoNames. O mapeamento da ontologia gerada para as ontologias externas é realizado convertendo as relações de equivalência de rótulos, obtidas na etapa de alinhamento e/ou mapeamento ontológico, em propriedades *rdfs:equivalentClass*, enquanto que relações de proximidade de conceito são convertidas em propriedades temporárias *nearConcept* para, posteriormente, facilitar a estruturação da ontologia pela comunidade do sistema colaborativo, como o caso da identificação da correlação entre *Hotel* e *Motel*, ilustrado na figura 4. A figura 5 ilustra o RDF de exemplo gerado pelo Protege, utilizado como esquema para a conversão.

²⁰Disponível em: http://www.geonames.org/ontology/mappings_v3.01.rdf. Acesso em nov. 2013.

²¹Disponível em: <http://www.w3.org/2001/sw/wiki/OWL>. Acesso em mar. 2014.

²²Disponível em: <http://protege.stanford.edu/>. Acesso em mar. 2014.

Figure 4. Triplificação de categoria (figura criada pelo autor, em 2013)

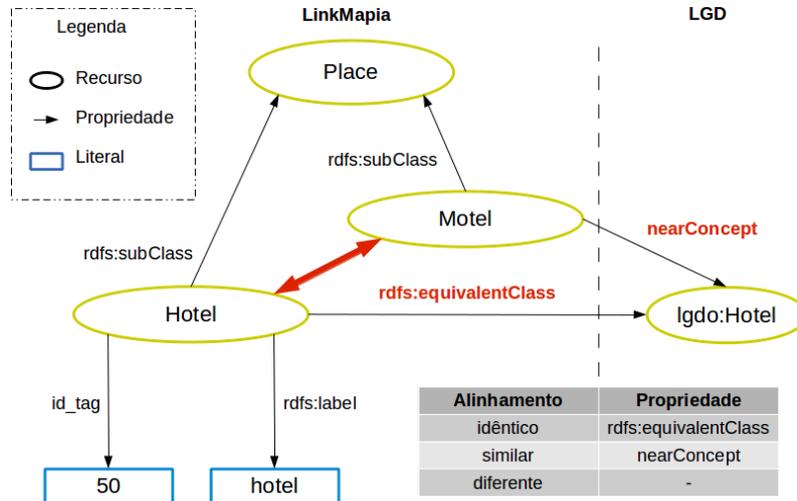
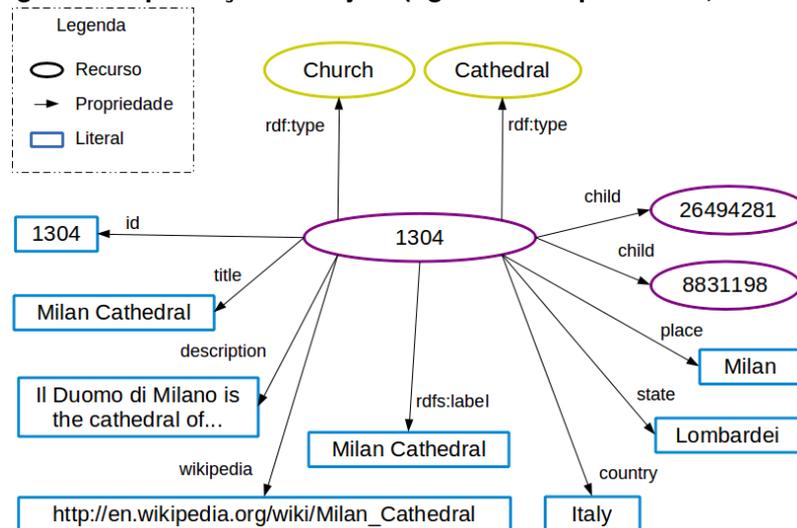


Figure 5. Triplificação de objeto (figura criada pelo autor, em 2013)



4. Resultados esperados

Com a nova coleção de dados LinkMapia, é possível responder perguntas como: **Quais são as opções de locais para dormir (como hotel, pensão, albergue) a até quinhentos metros do centro tecnológico da UFSC?** Dependendo de como a ontologia foi estruturada pela comunidade, é possível utilizar termos mais abrangentes (locais para dormir) para considerar na consulta uma coleção de conceitos (hotel, pensão, albergue), possibilitando diferentes granularidades para a categoria do objeto espacial.

O endpoint SPARQL permite consultas como: **Qual o restaurante conhecido como "Universitário" a até 200m do Centro de Eventos?** (figura 6). Esta consulta considera não apenas objetos geográficos do LGD classificados como *Igdo:Restaurant* como também da Wikimapia, anotados como *restaurant* e conceitos mais específicos (*seafood restaurant*, *drive-in restaurant*, dentre outros).

Figure 6. Consulta SPARQL - Qual o restaurante conhecido como "Universitário" a até 200m do Centro de Eventos?

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

SELECT DISTINCT ?local ?nomeLocal {
  ?local
    a lgdo:Restaurant;
    rdfs:label ?nomeLocal ;
    geom:geometry [ ogc:asWKT ?geo ] .
  FILTER( regex(?nomeLocal, "Universitário") &&
    bif:st_intersects ( ?geo,
      bif:st_point (-48.52015,-27.60225), 0.2) ).
}
```

5. Trabalhos Correlatos

Os principais trabalhos relacionados são o Projeto SEEK²³, o Linked Geo Data [Stadler et al. 2012], GeoLinkedData²⁴ (Espanha), o Linked Wikimapia²⁵, o *OurMap* [Gonzalez et al. 2013], o FolksOntology [Damme et al. 2007] e o LOM [Li 2004].

A iniciativa deste trabalho foi a obtenção de dados geográficos anotados para realizar análises semânticas de trajetórias de objetos móveis. Pretende-se com este trabalho apoiar trabalhos futuros, como o projeto *Semantic Enrichment of trajectory Knowledge discovery* (SEEK) e Baquara [Fileto et al. 2013].

O enriquecimento semântico de VGI já foi realizado em dados do projeto OSM [Stadler et al. 2012]. Outra coleção de dados geográficos ligados é o GeoLinkedData, uma iniciativa do *Ontology Engineering Group* (OEG) destinada ao enriquecimento da *Web* de dados com dados geoespaciais da Espanha. Há outro projeto de publicação de dados do Wikimapia, o Linked Wikimapia. Entretanto, são publicados dados ligados de objetos geográficos, sem ontologia para representar a *folksonomia*. Os dados são ligados com o DBpedia a partir da conversão dos *links* já existentes para páginas do Wikipedia.

Uma diferente abordagem é seguida no sistema de coleta colaborativa *OurMap* [Gonzalez et al. 2013], onde o sistema de coleta colaborativa utiliza anotações semânticas e o usuário tem a opção de gerenciar a ontologia.

O processo de derivação de uma ontologia a partir de uma *folksonomia* é discutido por [Damme et al. 2007]. Este trabalho inspirou o processo proposto. Outro trabalho que ajudou na definição e implementação do processo foi o de [Li 2004].

6. Conclusões e Trabalhos Futuros

Este trabalho apresenta uma proposta de conversão de dados geográficos coletados voluntariamente em dados ligados, descreve uma abordagem de implementação e explora os resultados esperados pela conversão. Trabalhos futuros incluem análises de outros sistemas

²³Disponível em: <http://www.seek-project.eu/>. Acesso em nov. 2013.

²⁴Disponível em: <http://geo.linkeddata.es/web/guest/home>. Acesso em nov. 2013.

²⁵Disponível em: <http://openeanwrap.appspot.com/>. Acesso em nov. 2013.

colaborativos, elaboração e implementação de novas propostas de processo e análises de novos casos de estudo para determinar a melhor maneira de converter *folksonomias* em ontologias. A elaboração do processo proposto, sua implementação e resultados preliminares demonstram que a abordagem proposta é promissora.

Agradecimentos. Este trabalho contou com o apoio do projeto European Union's IRSES-SEEK (concessão 295179), do CNPq (concessão 478634/2011-0), da CAPES, e da FEESC. Agradecimentos também a comunidade Wikimapia, e a Willian Ventura Koerich e André Salvaro Furtado pelo apoio técnico.

References

- Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The Semantic Web: Scientific American. *Scientific American*, 284(5).
- Casanova, M., Câmara, G., Davis, C., Vinhas, L., and de Queiroz, G. R., editors (2005). *Bancos de Dados Geográficos*. MundoGEO.
- Cohen, W. W., Ravikumar, P. D., and Fienberg, S. E. (2003). A comparison of string distance metrics for name-matching tasks. In *IWeb*, pages 73–78.
- Damme, C. V., Hepp, M., and Siorpaes, K. (2007). Folksonology: An integrated approach for turning folksonomies into ontologies. In *Proceedings of the ESWC Workshop Bridging the Gap between Semantic Web and Web 2.0*. Springer.
- Fileto, R., Krüger, M., Pelekis, N., Theodoridis, Y., and Renso, C. (2013). Baquara: A holistic ontological framework for movement analysis using linked data. In *ER*, volume 8217 of *Lecture Notes in Computer Science*, pages 342–355. Springer.
- García-castro, L. J. and García, E. (2011). Folksonomies behind the scenes.
- Gonzalez, A., Izidoro, D., Willrich, R., and Santos, C. (2013). Representação Aberta e Semântica de Anotações de Incidentes em Mapas Web. In *Simpósio Brasileiro de Sistemas Multimídia e Web*, pages 1–12.
- Goodchild, M. F. (2007). Citizens as voluntary sensors: spatial data infrastructure in the world of web 2.0. *International Journal of Spatial Data Infrastructures Research*, 2:24–32.
- Guarino, N. (1998). Formal ontology and information systems. pages 3–15. IOS Press.
- Heath, T. and Bizer, C. (2011). *Linked Data: Evolving the Web into a Global Data Space*. Morgan & Claypool, 1st edition.
- Kyzirakos, K., Karpathiotakis, M., and Koubarakis, M. (2012). Strabon: A semantic geospatial dbms. In *International Semantic Web Conference (1)*, volume 7649 of *Lecture Notes in Computer Science*, pages 295–311. Springer.
- Li, J. (2004). Lom: A lexicon-based ontology mapping tool. In *Proceedings of the Performance Metrics for Intelligent Systems (PerMIS)*, page 2004.
- Rigaux, P., Scholl, M., and Voisard, A. (2000). *Introduction to Spatial Databases: Applications to GIS*. Morgan Kaufmann.
- Stadler, C., Lehmann, J., Höffner, K., and Auer, S. (2012). Linkedgeodata: A core for a web of spatial open data. *Semantic Web Journal*, 3(4):333–354.