

# Modelagem do conhecimento usando ontologias para avaliar a sustentabilidade de projetos construtivos

Marcello Peixoto Bax, Cristiano Geraldo Teixeira Silva

ECI - PPGGOC - Universidade Federal de Minas Gerais (UFMG)

Av. Antônio Carlos, 6627 - Pampulha - 31270-901

Belo Horizonte - Minas Gerais - Brazil

bax.ufmg@gmail.com, cgts.br@gmail.com

**Abstract.** *Measuring the sustainability performance of built environments is important for the increasingly deteriorating environment. However, most assessment tasks are still performed manually, consuming time and increasing the possibility of errors. The use of semantic technologies is an alternative to integrate data and make inferences in the decision-making process. This work presents the elaboration of an ontology to integrate data and support the evaluation of a construction regarding the fulfillment of the LEED certification requirements (Leadership in Energy and Environmental Design). Data from a project developed on a BIM (Building Information Modeling) platform are integrated with other tabulated open data. These datasets are semantically annotated following the semantic data dictionary approach to RDF conversion. Once the datasets are integrated into a triplestore, queries are performed to explore the linking of this data with the generation of inferences. The results suggest that the application of these technologies promotes the semantic extension of construction elements in BIM, facilitating their integration with other knowledge bases and conceptually organizes the data to better retrieve information.*

**Resumo.** *Mensurar o desempenho em sustentabilidade de ambientes construídos é importante para o meio ambiente em crescente deterioração. No entanto, a maioria das tarefas de avaliação ainda é realizada manualmente, consumindo tempo e aumentando a possibilidade de erros. O emprego de tecnologias semânticas é uma alternativa para integrar dados e realizar inferências no processo de tomada de decisões. Este trabalho apresenta a elaboração de uma ontologia para integrar dados e apoiar a avaliação de uma construção quanto ao atendimento dos requisitos da certificação LEED (Leadership in Energy and Environmental Design). Dados de um projeto desenvolvido em uma plataforma BIM (Building Information Modeling) são integrados a outros dados abertos tabulados. Esses conjuntos de dados são anotados semanticamente seguindo a abordagem de dicionário semântico de dados para conversão em RDF. Uma vez integrados os conjuntos de dados em um triplestore, são realizadas consultas para explorar a vinculação destes dados com a geração de inferências. Os resultados sugerem que a aplicação destas tecnologias promove a extensão semântica dos elementos construtivos em BIM, facilitando a sua integração com outras bases de conhecimento e organiza conceitualmente os dados para melhor recuperar informações.*

## 1. Introdução

A classificação de ambientes construídos sustentáveis como economicamente viáveis tem recebido atenção e, para obter um nível de classificação, a avaliação do projeto é indispensável [Azhar et al. 2011]. Um sistema empregado em vários países é a certificação LEED (*Leadership in Energy and Environmental Design*), concebida e concedida pela USGBC (*United States Green Building Council*) que classifica as construções em quatro níveis (certificado, prata, ouro e platina) e é amplamente utilizada nas fases de projeto, construção, operação e manutenção de ambientes construídos [USGBC 2020]. Entretanto, como a sua metodologia de avaliação é manual, trabalhosa e propensa a erros, métodos mais inteligentes para certificar a sustentabilidade de construções são necessários [Jiang et al. 2018].

Ontologias representam um domínio de conhecimento, estabelecendo uma estrutura com regras e critérios que regulam a combinação entre conceitos e relações no domínio. O W3C LBD CG (*Linked Building Data Community Group*) propõe ontologias que descrevem produtos da construção civil como soluções para representar a semântica dos dados [Rasmussen et al. 2018]. Na construção civil tem-se a ontologia ifcOWL BIM (*Building Information Modeling*), usada para simular um empreendimento em seu ciclo de vida [Baracho et al. 2017; Sacks et al. 2018].

Este trabalho visa contribuir para investigar a adoção de tecnologias semânticas para proporcionar maior eficiência na avaliação de projetos de ambientes construídos. Para tanto, elabora uma ontologia para organizar os diversos dados aplicados na avaliação de uma construção com relação à sua certificação verde, LEED (*Leadership in Energy and Environmental Design*).

Na Seção 2 do artigo apresenta-se o desenvolvimento do experimento com o emprego de tecnologias semânticas, a Seção 3 apresenta a demonstração do protótipo e a Seção 4 sintetiza as conclusões.

## 2. Desenvolvimento do Experimento

O fluxo de trabalho para o desenvolvimento do protótipo<sup>1</sup> exige o conhecimento de tecnologias e, ainda, pode ser considerado complexo para o uso de um especialista em certificação LEED. A implementação busca validar o experimento como solução para o problema desta pesquisa, com os devidos cuidados para que possa ser aprimorada e expandida em outras soluções.

No início deste fluxo, um arquivo no formato IFC (*Industry Foundation Classes*) de um projeto BIM é convertido em uma ontologia simplificada para avaliação dos critérios LEED (Seção 2.1). O IFC é o formato de troca de dados de um modelo BIM entre *softwares* que consiste em fornecer definições gerais e amplas de elementos de um ambiente construído. Este formato utiliza de termos como tipos, entidades e propriedades, e regras que devem ser usadas para criar um esquema específico [Sacks et al. 2018].

Para integrar dados externos, dados abertos, em formato CSV (*Comma-separated values*), são obtidos, mapeados para conceitos das ontologias e convertidos em grafos RDF, utilizando um dicionário semântico de dados (Seção 2.2). Os conjuntos de dados são

---

<sup>1</sup> Disponível em: <https://github.com/cgtsbr/sebim>.

inseridos no *triplestore* Parliament para realização das consultas e inferências em SPARQL (Seção 2.3). Uma consulta final é executada para apresentação do resultado das inferências sobre as pontuações obtidas para a certificação.

Este processo pode ser generalizado nas seguintes etapas: desenvolvimento da ontologia, extração de informações, alinhamento de classes e conversão para RDF e, consultas e inferências de regras.

## 2.1 Ontologia Simplificada

Uma ontologia simplificada é criada, reutilizando as ontologias ifcOWL [Pauwels e Terkaj 2016], SimpleBIM [Pauwels and Roxin 2016], BOT [Rasmussen et al. 2017] e BIMSO [Niknam and Karshenas 2017], para organizar semanticamente os dados BIM. Os critérios LEED abrangendo o uso de materiais, utilizações dos espaços, dentre outros, são considerados usando dados de um projeto BIM. Tais critérios, no contexto da ontologia, embasam as inferências que avaliam o estado da edificação. A ontologia simplificada SEBIM (*Semantic BIM*) foi criada no Protégé 5.5 (Figura 1), com classes que representam os dados necessários para organizar as informações e analisar os critérios da certificação LEED.

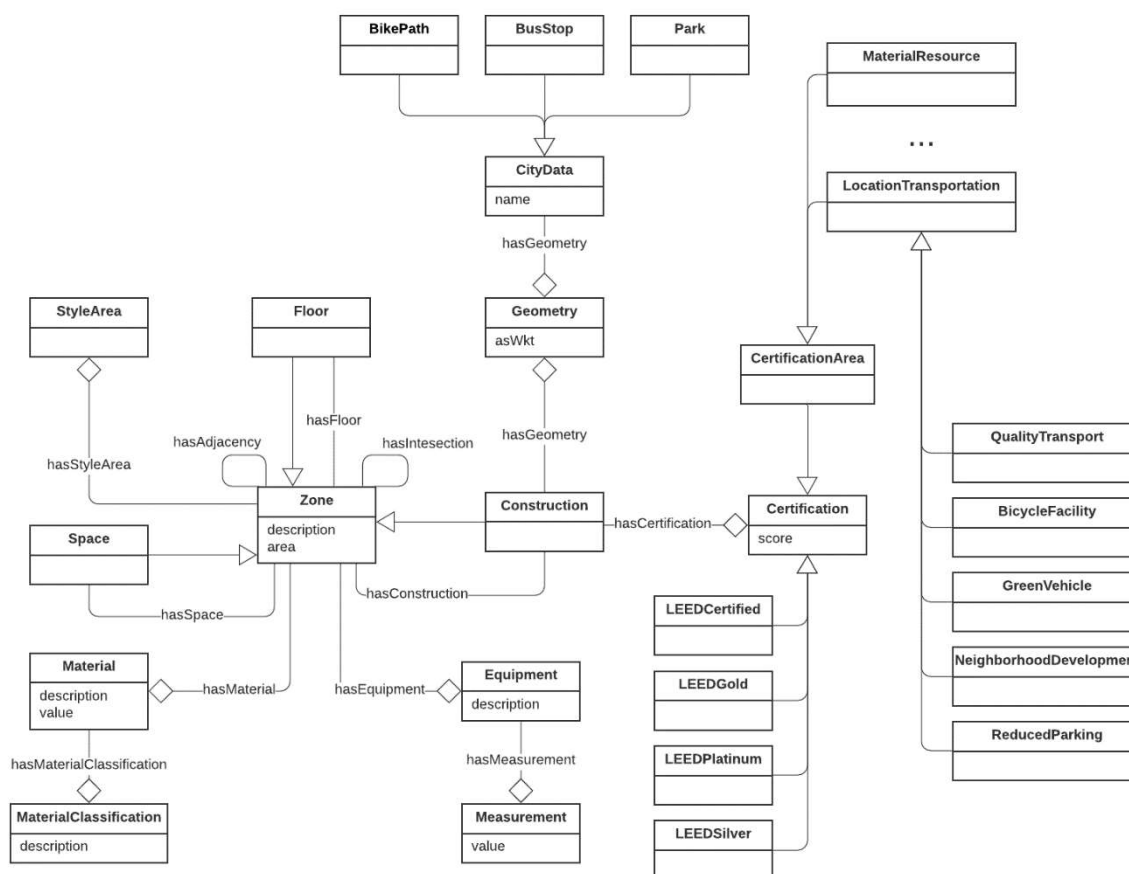


Figura 1. Ontologia simplificada

Os elementos do projeto de construção foram escolhidos para atender as análises dos critérios escolhidos para o experimento, mas, novas classes podem ser incluídas de acordo com outras análises, seguindo o mesmo método. As áreas (*Zone*) do ambiente construído são divididas nas subclasses *Construction*, *Space* e *Floor*. Estas zonas são relacionadas entre si através das propriedades de interseção (*hasIntersection*) e adjacência (*hasAdjacency*),

possibilitando a interconexão de zonas entre si. Por exemplo, instâncias de *Space* em um apartamento consistindo em zonas interconectadas: quartos, banheiro, sala de estar e cozinha. A classe *Zone* está relacionada à classe *StyleArea* e pode ter equipamentos instalados. A classe *StyleArea* é uma classificação importante para organizar e identificar áreas relevantes para avaliação ambiental, como áreas verdes, estacionamento de veículos verdes, entre outros.

Como muitas análises envolvem utilizar materiais e equipamentos, classes foram consideradas para abstrair esses dados, e o relacionamento de inclusão em uma Zona. A vinculação com equipamentos envolve a medição de dados e a localização interna de sensores, equipamentos de ar-condicionado, placas fotovoltaicas, entre outros. Já o relacionamento de zonas com materiais busca identificar o atendimento aos requisitos de sustentabilidade. Os materiais têm relação para a sua composição de matérias primas e para classificação, conforme o tipo de utilização. Tais relações e atributos correspondem às propriedades de dados e objetos.

Para aplicar os critérios da certificação, uma classe *Construction* se relaciona com a classe *Certification* (*hasCertification*). No manual LEED, os critérios são divididos em temas para avaliar a sustentabilidade. Assim, os critérios são subclasses especializadas conforme as classes de *LocationTransportation*, *WaterEfficiency*, *EnergyAtmosphere*, *ResourcesMaterials*, *RegionalPriorityCredit*, *IndoorEnvironmentalQuality* and *Innovation* classes. Between the *LocationTransportation* e *ResourcesMaterials*. Entre as classes de *LocationTransportation* e *MaterialResource* existem as demais classes, suprimidos no diagrama por limitação de espaço. As subclasses de *LocationTransportation*, bem como das outras subclasses, registram separadamente cada pontuação recebida para organizar as informações. Uma tripla de dados associada a uma subclasse será inserida ao conjunto de dados com a pontuação recebida. Destaca-se ainda que não foram realizados relacionamentos entre os critérios e as classes de componentes de uma construção, uma vez que as inferências realizam a recuperação destas informações.

Os atributos possuem valores numéricos (*Certification.score*, *Material.value*, *Zone.area* and *Measurement.value*), descrições alfanuméricas (*Zone.description*, *Material.description* and *Equipment.description*) e dados geoespaciais (*asWkt*). Sendo este último necessário para vincular os dados abertos sobre o entorno do ambiente construído.

## 2.2 Anotação de Dados Tabulares

Dados abertos sobre meios de transporte e parques, obtidos no site da prefeitura de Belo Horizonte, são utilizados para avaliação de critérios da classe *LocationTransportation*. Estes dados têm coordenadas geométricas de áreas e pontos da cidade, seguindo padronização da OGC (*Open Geospatial Consortium*), que define um vocabulário para representar dados geoespaciais. Esse formato de dados é necessário para realizar consultas, aplicando as funções GeoSPARQL, que retornam triplas conforme proximidades e relações entre coordenadas [Perry and Herring 2011].

Diante das opções para conversão de dados tabulares para RDF [Ding et al. 2011; Jeremy Tandy et al. 2015; Rashid et al. 2020], optou-se por Dicionários Semânticos de Dados, implementado na ferramenta *sdd2rdf*. Os objetos e seus atributos são representados e identificados por meio de ontologias relevantes que constituem essas informações de uma maneira formalmente precisa e legível por máquina [Rashid et al. 2020].



O grafo RDF gerado pela execução do *script* *ssd2rdf* contém a formalização dos dados e favorece a integração para consulta sobre locais da cidade por dados espaciais. Assim, dados sobre ciclovias, parques e pontos de ônibus são convertidos em grafos RDF.

### 2.3 Geração de Inferências

As inferências deste trabalho foram realizadas em SPARQL, aplicando um INSERT para inserir a tripla inferida no conjunto de dados, direcionando a pontuação à classe correspondente ao critério. A consulta (1) exibe a inferência realizada para avaliar e pontuar a instalação para bicicletas, calculada sobre a quantidade de unidades do prédio.

```
INSERT {?cr <http://www.semanticweb.org/SEBIM#score> "1"^^xsd:decimal}
WHERE {
  ?cr rdf:type sevim:BicycleFacilities.
  ?build rdf:type sevim:Material.
  ?build sevim:hasMaterialClassification sevim:RackBike.
  ?build sevim:value ?n.
  { SELECT (count(?s) as ?place) WHERE { ?s rdf:type sevim:Space.} }
}
GROUP BY ?cr ?build ?n ?place
HAVING (?place >= ?n)
```

(1)

O processo de inserir novas triplas via SPARQL automatiza a avaliação dos critérios da certificação LEED. As consultas foram agrupadas em dados internos ao projeto BIM, dados de simulações e dados externos ao BIM. A experimentação inicial baseou-se na criação de instâncias e atributos fictícios para validar as inferências dos critérios para avaliação de um projeto construtivo. Como a certificação LEED fornece critérios por tipologia da construção (novas construções, manutenções, hospitais e galpões), adotou-se a avaliação de uma nova construção. Diante da complexidade de informações necessárias para todos os critérios, selecionou-se aqueles que representassem um tipo de extração diferente. Assim, dos 57 critérios para uma nova construção, foram implementados 7 critérios, conforme explicado na próxima seção.

### 3. Demonstração do Protótipo

Para experimentar os métodos empregados no protótipo, foi escolhido um projeto acadêmico elaborado por alunos do curso de arquitetura da Universidade FUMEC. A modelagem foi criada no *software* Autodesk Revit<sup>2</sup> e contém edifícios envidraçados com 27 áreas (pavimentos). O Revit possibilita a modelagem tridimensional em BIM, oferece suporte a projetos, possibilita simulações de análises da construção e exporta dados para integração com outras ferramentas.

Os dados do projeto BIM (exportados do Revit no formato CSV) a ser avaliado como sustentável e os dados da prefeitura de Belo Horizonte (disponíveis no formato CSV) foram anotados semanticamente por meio do dicionário semântico e convertidos de forma integrada para o padrão RDF (com a ferramenta *sdd2rdf*) e, por fim, inseridos no *triplestore*. A execução das inferências via consultas SPARQL gerou novas triplas contendo a pontuação final para verificar se os critérios da certificação foram ou não atendidos. Os critérios analisados no projeto em questão são apresentados na Tabela 1. Pode-se concluir que apenas

<sup>2</sup> <https://www.autodesk.com.br/products/revit>

dois critérios não obtiveram nota máxima. Cabe ainda destacar que os critérios com nota máxima acima de 1, podem apresentar escalas de pontuação conforme o manual LEED.

Durante o levantamento bibliográfico deste trabalho, as dificuldades identificadas (Tabela 2) no processo tradicional de organização das informações para a certificação LEED são utilizadas para validar a pertinência da aplicação das tecnologias semânticas empregadas no protótipo proposto.

**Tabela 1. Pontuação dos critérios utilizados na experimentação**

<b>Categoria LEED</b>	<b>Critério</b>	<b>Nota Máxima</b>	<b>Pontuação</b>
Localização e Transporte	Acesso a transporte de qualidade	5	5
Localização e Transporte	Instalações para bicicletas	1	0
Localização e Transporte	Veículos verdes	1	0
Terrenos Sustentáveis	Espaço aberto	1	1
Energia e Atmosfera	Produção de energia renovável	3	0
Materiais e Recursos	Ingredientes do material	2	2
Qualidade do Ambiente	Iluminação interna	2	2

Fonte: Os autores.

**Tabela 2. Soluções para as dificuldades encontradas**

<b>Dificuldade / Desafio de pesquisa</b>	<b>Solução</b>
Extensão do formato IFC da tecnologia BIM	Conversão do IFC para a ontologia simplificada (SEBIM)
Integração com dados externos	Anotação semântica de dados tabulares
Organização das informações	Grafos de conhecimento
Análise dos dados (critérios LEED)	Inferências SPARQL
Recuperação de dados	Consultas SPARQL

Fonte: Os autores.

Na versão inicial do protótipo, buscou-se a abordagem baseada em ontologia monotônica, na qual a linguagem de regras SWRL foi usada para formalizar as regras para inferir as pontuações. Porém, esta abordagem apresentou algumas limitações, especialmente na modelagem de regras complexas e facilidade de manutenção. Assim, a implementação do modelo de raciocínio baseado em regras, utilizou consultas SPARQL, gerando novas triplas para inferências sobre a pontuação de cada critério.

Como o processo de avaliação abrange as fases de um projeto construtivo e exige revisar e atualizar o projeto para atender aos critérios, é importante considerar a organização histórica das informações. Assim, carregar uma nova versão de um projeto BIM permite comparar resultados e analisar a evolução no atendimento aos critérios para certificação.

O fluxo de trabalho apresentado permite a implementação de uma solução amigável para especialistas que desconhecem tecnologias semânticas. O experimento pode ser implementado em *frameworks* como *JenaSemanticWeb* e *VirtuosoOpenSource*, que suportam funções GeoSPARQL e recursos para as inferências.

#### 4. Conclusão

Este trabalho apresenta uma solução baseada em ontologias para integração de dados e inferências para automação da avaliação para certificação LEED. A tecnologia BIM, como referência atual na indústria da construção civil, é utilizada para criação de uma ontologia simplificada para integração com outros dados. Os dados de um projeto BIM e dados abertos da prefeitura de Belo Horizonte, em formato tabular, são anotados semanticamente com a utilização do dicionário semântico de dados. Uma vez os dados integrados em um *triplestore*,

consultas SPARQL forneceram inferências sobre o atendimento aos critérios para certificação LEED.

Novas experimentações proporcionarão o refinamento das consultas para a implementação de um sistema eficiente para automação de certificação LEED. Além disso, espera-se que aplicações similares possam favorecer à automação na avaliação de projetos construtivos como autorização de construção nas prefeituras, entre outras avaliações que exigem análises complexas.

## Referências

- Azhar, Salman, et al. “Automation in Construction Building information modeling for sustainable design and LEED ® rating analysis”. *Automation in Construction*, vol. 20, nº 2, Elsevier B.V., 2011, p. 217–24.
- Baracho, Renata M. A., et al. “Ontologia, Internet das Coisas e Modelagem da Informação da Construção (BIM): Estudo Exploratório e a Inter-relação entre as Tecnologias”. *CEUR Workshop Proceedings*, vol. 1908, 2017, p. 141–46.
- Ding, Li, et al. “TWC LOGD: A portal for linked open government data ecosystems”. *Journal of Web Semantics*, vol. 9, nº 3, Elsevier B.V., 2011, p. 325–33.
- Jeremy Tandy, et al. “Generating RDF from Tabular Data on the Web”. *W3C Recommendation*, nº December, 2015, p. 1–29, <https://www.w3.org/TR/csv2rdf/>.
- Jiang, Shaohua, et al. “Combining BIM and Ontology to Facilitate Intelligent Green Building Evaluation”. *Journal of Computing in Civil Engineering*, vol. 32, nº 5, 2018, p. 1–15.
- Niknam, Mehrdad, e Saeed Karshenas. “A shared ontology approach to semantic representation of BIM data”. *Automation in Construction*, vol. 80, Elsevier B.V., 2017, p. 22–36.
- Pauwels, Pieter, e Ana Roxin. “SimpleBIM: From full ifcOWL graphs to simplified building graphs”. *European Conference on Product & Process Modelling*, nº September, European Conference on Product & Process Modelling, 2016.
- Pauwels, Pieter, e Walter Terkaj. “EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology”. *Automation in Construction*, vol. 63, Elsevier B.V., 2016, p. 100–33.
- Perry, Matthew, e John Herring. “GeoSPARQL - A geographic query language for RDF data”. *Open Geospatial Consortium*, 2011, <http://www.opengeospatial.org/legal/>.
- Rashid, Sabbir M., et al. “The Semantic Data Dictionary – An Approach for Describing and Annotating Data”. *Data Intelligence*, 2020, p. 443–86.
- Rasmussen, Mads Holten, et al. “Web-based topology queries on a BIM model”. *5th LDAC workshop*, nº November, 2017.
- Sacks, Rafael, et al. *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers, and Contractors*. 3º ed, John Wiley & Sons, Inc., 2018.
- USGBC. “LEED V4.1: Building Design and Construction”. *US GREEN BUILDING COUNCIL*, nº January, 2020, p. 259.