

# Estudo de desenvolvimento de ontologia para instituição de ensino extendendo FOAF e RDFS

Felipe Hauschild Grings<sup>1</sup>

<sup>1</sup>Universidade do Vale do Rio do Sinos (UNISINOS)  
91.501-970 – São Leopoldo 93022-750 – RS – Brazil

**Abstract.** *This article develops a FOAF (Friend Of a Friend) and RDFS (Resource Description Framework Schema) ontology extension using a education institution as domain, where the goal is on the relationship between tasks, courses, people and associations. This paper follows the ontology development pattern described by Noy [8], Ontology Development Methodology 101. The paper presents a class and property diagram, as well as a graph diagram. At the end, the population of the database and an inference is performed. As result the ontology proposed on this paper shows as satisfactory for the proposals described.*

**Resumo.** *Este artigo realiza o desenvolvimento de uma ontologia para extensão de FOAF (Friend Of a Friend) e RDFS (Resource Description Framework Schema) utilizando o domínio de uma instituição de ensino, onde o foco é a relação entre tarefas, cursos, pessoas e organizações. O trabalho segue o padrão de desenvolvimento de ontologias descrito por Noy [8], Metodologia Ontology Development 101. O trabalho apresenta um diagrama de classes e propriedades, assim como o diagrama de instâncias. Ao final é realizado a população da base de dados e a inferência da mesma. A ontologia proposta neste trabalho apresentou-se satisfatória para as definições propostas.*

## 1. Introdução

A quantidade de dados científicos gerados diariamente está se tornando cada vez maior e a organização destas informações utilizando a correta estratégia pode garantir a sua integração progressiva com as informações já existentes, tornando-as disponíveis em formatos compreensíveis para os computadores e seres humanos. A integração e disponibilização desses dados faz com que resultados alcançados em Pequim, por exemplo, possam ser processados quase que instantaneamente em qualquer outro lugar do globo. Por essas razões, as informações científicas precisam ser armazenadas, padronizadas, processadas e disponibilizadas de uma forma que supere as idiosincrasias de seus criadores [1].

A padronização por ontologias e web semântica iniciaram com o intuito de resolução das divergências de estruturas para armazenamento de dados. Toda área de conhecimento cria ontologias para limitar a complexidade e organizar dados em informações e conhecimento. Novas ontologias estão surgindo e melhorando a resolução de problemas a partir do momento que os dados utilizam uma estrutura globalmente padronizada. Dentre as ontologias mais conhecidas na atualidade destacam-se a FOAF (Friend of a Friend) e RDFS (Resource Description Framework Schema), onde esse trabalho realiza a uma ontologia utilizando e integrando as duas citadas [1].

## 2. Conceitos Gerais

No contexto das ciências da computação e da informação, uma ontologia define um conjunto de primitivas representacionais com as quais pode-se modelar um domínio de conhecimento ou discurso. As primitivas representacionais são normalmente classes (ou conjuntos), atributos (ou propriedades) e relacionamentos (ou relações entre os membros da classe). As definições das primitivas representacionais incluem informações sobre seu significado e restrições em sua aplicação logicamente consistente. No contexto dos sistemas de banco de dados, a ontologia pode ser vista como um nível de abstração de modelos de dados, análogo aos modelos hierárquicos e relacionais, mas destinado a modelar o conhecimento sobre indivíduos, seus atributos e seus relacionamentos com outros indivíduos. Ontologias são normalmente especificadas em linguagens que permitem abstrações distantes de estruturas de dados e estratégias de implementação; na prática, as linguagens de ontologias estão mais próximas em poder expressivo da lógica de primeira ordem do que as linguagens usadas para modelar bancos de dados. Por esse motivo, as ontologias são consideradas no nível "semântico", enquanto os esquemas de banco de dados são modelos de dados no nível "lógico" ou "físico". Devido à sua independência de modelos de dados de nível inferior, as ontologias são usadas para integrar bancos de dados heterogêneos, permitindo a interoperabilidade entre sistemas distintos e especificando interfaces para serviços independentes baseados em conhecimento. Na pilha de tecnologia dos padrões da Web Semântica [1], as ontologias são chamadas como uma camada explícita. Agora existem linguagens padrão e uma variedade de ferramentas comerciais e de código aberto para criar e trabalhar com ontologias [2].

O conceito de linked data (dados ligados entre si) é um conjunto de práticas introduzidas por Tim Berners-Lee em suas notas sobre a arquitetura web "Linked Data" [3], com função de publicar e estruturar dados na Web. Essas práticas vêm sendo cada vez mais adotadas levando à criação do que conhecemos como web de dados [5]. No contexto de Web Semântica, a função não é somente disponibilizar os dados, mas também realizar de forma com que tanto as pessoas quanto as máquinas possam explorar a base de dados [5].

Um documento web é construído sobre um pequeno conjunto de padrões simples, utilizando URIs como mecanismo global e único de identificação, HTTP como mecanismo de acesso universal, e HTML como formato de conteúdo. A Web é construída sobre o princípio de manter hiperlinks entre documentos da Web. Baseando-se nisto, Berners-Lee criou suas notas, que ficaram conhecidas como "Os princípios da Linked Data".[5] Sendo elas:

- Use URIs para nomear as coisas
- Use URIs HTTP para que as pessoas possam procurar o desejado
- Quando alguém olha para um URI, forneça informações úteis, usando os padrões (RDF \*, SPARQL)
- Incluir links para outros URIs. Para que eles possam descobrir explorar mais as coisas

Diante dos padrões apresentados FOAF é dedica-se a vincular pessoas e informações. A ontologia integra três tipos de rede: redes sociais de colaboração humana, amizade e associação; redes representacionais que descrevem uma visão simplificada do universo dos desenhos animados em termos factuais e redes de informação que usam links baseados na Web para compartilhar descrições publicadas de forma independente

deste mundo interconectado. O FOAF não compete com sites de orientação social; em vez disso, fornece uma abordagem na qual diferentes sites podem contar diferentes partes de uma história maior, e pela qual os usuários podem manter algum controle sobre suas informações em um formato não proprietário.

No mesmo contexto também é apresentado RDF Schema que fornece um vocabulário de modelagem de dados para dados RDF. É complementado por vários documentos que descrevem os conceitos básicos e sintaxe abstrata de RDF, a semântica formal de RDF e várias sintaxes concretas para RDF, como Turtle, TriG, e JSON-LD. O RDF Primer fornece uma introdução informal e exemplos do uso dos conceitos especificados neste documento.

Com a apresentação dos conceitos utilizados para o desenvolvimento do trabalho, no próximo capítulo é realizada uma breve análise do estado da arte sobre ontologias e semânticas web.

### **3. Estado da Arte**

Dado o grande aumento no número de ontologias, que estão disponíveis online, seus desenvolvimento estão se tornando cada vez mais um processo centrado na reutilização. Em particular, o nível de reutilização pode variar significativamente, dependendo se diz respeito a outras ontologias, como DOLCE 1, SUMO e Kowien 2; módulos de ontologia; declarações de ontologia e padrões de projeto de ontologia; e recursos não ontológicos, como tesouros, léxicos, DBs, diagramas UML e esquemas de classificação. Assim, neste contexto, o desenvolvimento da ontologia pode ser caracterizado como o construção de uma rede de ontologias, onde os diferentes recursos podem estar gerenciados por pessoas diferentes, possivelmente em organizações diferentes. Dada esta nova visão da engenharia de ontologia por reutilização, torna-se importante fornecer um forte suporte metodológico para o desenvolvimento colaborativo de redes ontológicas [7].

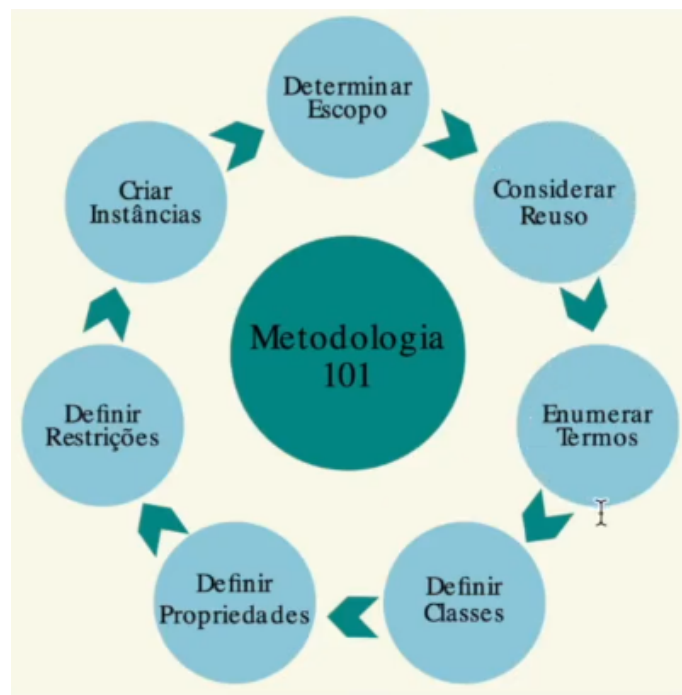
De acordo com Bursa [2], a agregação de ontologias já existentes é de extrema importância para o aprimoramento das ciências, principalmente em temas menos cotidianos das áreas de pesquisa. Bursa desenvolveu uma extensão onde são adaptadas as ontologias FOAF e BLOOD.HEALTH integrando com a busca por parentes e pessoas próximas. Já em Kassiri [4], foi desenvolvida a extensão de FOAF projetando uma unificação para análise de OSN (Online Social Network). Eles propuseram a integração das classes Pessoa para aprimoramento e detalhamento das informações relacionadas.

Com a análise e comparação de alguns dos métodos hoje estudados, apresenta-se uma melhor análise nos métodos menos probabilísticos.

#### **3.1. Tipo de aplicação desenvolvida**

Para realização do estudo foi desenvolvida uma ontologia para integração de FOAF e RDFS, utilizando o editor de ontologias de código aberto e framework para desenvolvimento de sistemas inteligentes, protege. A ontologia visa o detalhamento de um sistema educacional de ensino, no qual possa realizar a busca e o detalhamento das informações relacionadas as tarefas, pessoas e disponibilidades das universidades. O desenvolvimento foi seguido utilizando os princípios descrito por [5], conforme a Figura 1.

Para o desenvolvimento foi utilizado o seguinte contexto: "No domínio de uma instituição de ensino, existem cursos regulares e cursos em laboratório. Cursos regulares



**Figure 1. Ciclo de desenvolvimento de ontologias para Web Semantica**

são cursos teóricos e cursos de laboratório são cursos práticos. Matérias são parte dos cursos, os quais são organizados pelos professores. Entre os professores, existem os que trabalham um turno, chamados de Part Time e os que trabalham o dia completo Full Time. Os Part Time somente lecionam cursos de laboratório, enquanto os Full Time lecionam todos os cursos. A instituição de ensino é uma organização e o cargo representado no trabalho é de professor. Cada professor é uma pessoa onde deve ser representada como tal, com sua participação na instituição e suas tarefas a serem realizadas.” Onde o domínio é a Instituição de ensino e o escopo a recuperação da informação sobre a organização, os professores, cursos e matérias dos cursos. Com base nesta ontologia, visa-se responder as seguintes perguntas:

- Qual professor leciona qual curso?
- Qual professor organizou qual curso?
- Quais são as matérias de casa de um determinado curso?

Enumeração dos termos:

- Curso
- Curso Regular
- Curso Laboratório
- Matéria
- Professor
- Meio Turno
- Turno Completo
- ser/é
- é parte de
- é organizado
- leciona

- realiza

Com as seguintes prerrogativas:

- Existem cursos regulares
- Existem cursos em laboratório
- Lições de casa são parte dos cursos
- Entre os professores, existem os part time
- Entre os professores, existem os full time
- Os assistentes lecionam cursos de laboratório
- Os titulares lecionam todos os cursos

Definição das Classes:

- Curso
- Curso Regular
- Curso Laboratório
- Matéria
- Professor
- Part Time
- Full Time

As definições de propriedades e restrições são apresentadas nas Figuras 2 e 3.

Domain	Propriedade	Range
Todos conceitos	É	Todos Conceitos
Instituição de Ensino	É	Organização
Curso	É parte de	Instituição de Ensino
Matéria	É parte de	Curso
Full Time Professor	É um cargo de	Instituição de Ensino
Professor	Área de atuação de	Full Time Professor
Full Time Professor	Tem tarefas	Tarefas
Pessoa	É membro de	Organização
Pessoa	Realiza	Tarefas

**Figure 2. Definição de propriedades**

Curso Regular	É um	Curso
Curso Laboratório	É um	Curso
Full time	É um	Professor
Part time	É um	Professor
Organiza	É inverso	É organizado por

**Figure 3. Definição de restrições**

Após a realização das etapas descritas da Figura 1 foi desenvolvido a modelagem do diagrama para melhor visualização das informações, onde a Figura 4 apresenta em nível meta e a Figura 5 apresenta em nível de instâncias.

Após a configuração da ontologia proposta no protege e a população de um dataset para testes foi realizada a análise de resultados obtidos.

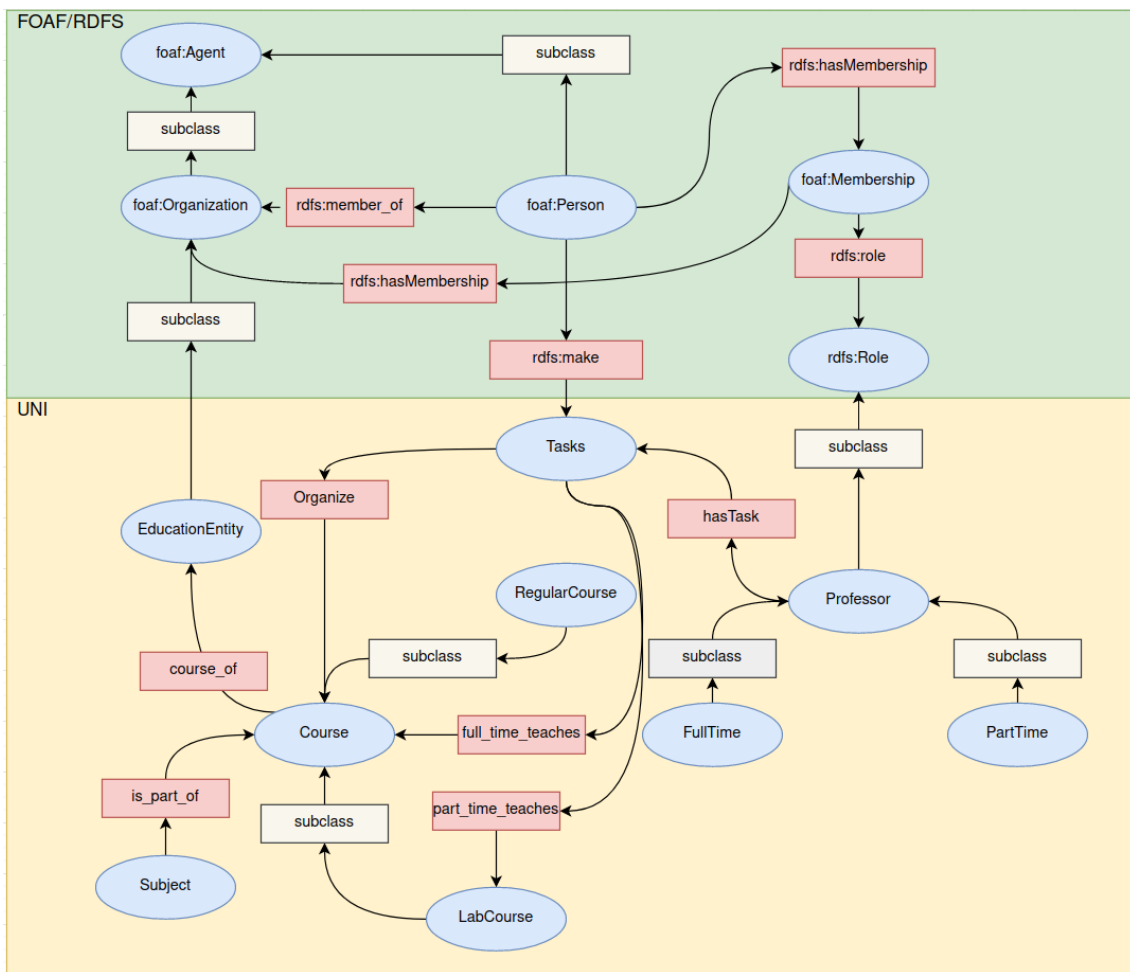


Figure 4. Diagrama em nível meta

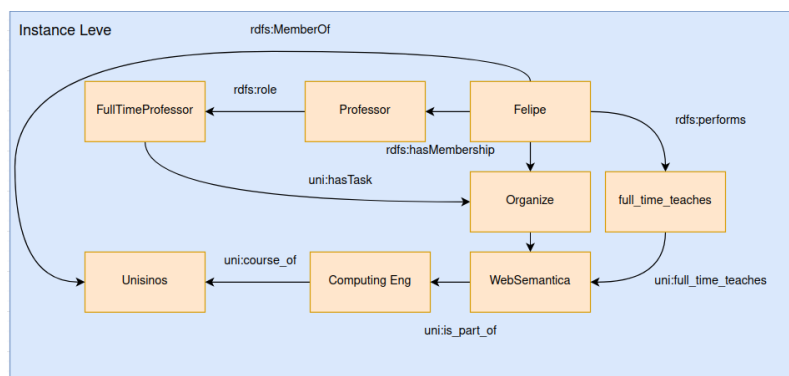
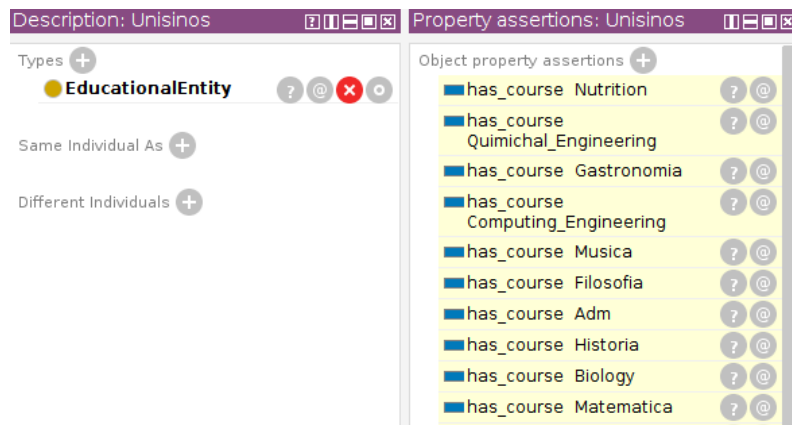


Figure 5. Diagrama em nível de instância

#### 4. Análise dos Resultados

Para análise de resultados foi populado o dataset utilizando as classes desenvolvidas e a importação de uma pequena base FOAF. A configuração das regras de inversão para as propriedades e equivalência para as classes apresentaram grande vantagem para o resultado disponibilizado pela inferência, apresentando um dataset mais completo e fiel as informações iniciais.



**Figure 6. Inferências realizadas**

Após a análise e a execução das inferências pode-se notar o correto funcionamento dos relacionamentos propostos e a alcance dos objetivos para este trabalho.

## 5. Conclusão

O desenvolvimento da ontologia apresentou-se válido para a proposta de estudo e integração de ontologias existentes. A complexidade de desenvolvimento e manipulação das informações foi um fator de grande relevância para o desenvolvimento do trabalho, pois o relacionamento entre entidades, classes, propriedades utiliza conceitos bem peculiares, dificultando a similaridade com bases de dados mais comuns nos campos da computação. Após os problemas resolvidos o trabalho foi considerado satisfatório.

## References

- Asmae El Kassiri, F.-Z. B. (2014). *A FOAF Ontology Extension to Meet Online Social Networks Presentation and Analysis*. Addison-Wesley, 15th edition.
- el al., B. (2003). Blood.health.foaf: Extending foaf with blood ontology.
- Robert Arp, B. S. and Spear, A. D. (2012). Building ontologies with basic formal ontology.
- Suárez-Figueroa, M. C. (2016). Ontology engineering in a networked world.
- [Robert Arp and Spear 2012], [Suárez-Figueroa 2016], [el al. 2003], [Asmae El Kassiri 2014].