

**UNIVERSIDADE FEDERAL DA GRANDE DOURADOS  
FACULDADE DE CIÊNCIAS EXATAS E TECNOLOGIA  
BACHARELADO EM SISTEMAS DE INFORMAÇÃO**

**ALEX YUDI OSHIRO  
SANDRO SEIYA OKAZACHI**

**PROTÓTIPO PARA RECOMENDAÇÃO DE ALIMENTOS  
A PARTIR DE SINTOMAS ASSOCIADOS A DOENÇAS NUTRICIONAIS**

**DOURADOS – MS**

**2016**

**ALEX YUDI OSHIRO**  
**SANDRO SEIYA OKAZACHI**

**PROTÓTIPO PARA RECOMENDAÇÃO DE ALIMENTOS**  
**A PARTIR DE SINTOMAS ASSOCIADOS A DOENÇAS NUTRICIONAIS**

Trabalho de Conclusão de Curso de graduação apresentado para obtenção do título de Bacharel em Sistemas de Informação pela Faculdade de Ciências Exatas e Tecnologia da Universidade Federal da Grande Dourados.  
Orientador: Prof.º Dr. Joinvile Batista Junior

**DOURADOS – MS**

**2016**

**ALEX YUDI OSHIRO**  
**SANDRO SEIYA OKAZACHI**

**PROTÓTIPO PARA RECOMENDAÇÃO DE ALIMENTOS**  
**A PARTIR DE SINTOMAS ASSOCIADOS A DOENÇAS NUTRICIONAIS**

Trabalho de Conclusão de Curso aprovado como requisito para obtenção do título de Bacharel em Sistemas de Informação na Universidade Federal da Grande Dourados, pela comissão formada por:

---

Orientador Prof. Dr. Joinvile Batista Junior  
FACET – UFGD

---

Prof. Me. Felipe José Carbone  
FACET – UFGD

---

Prof<sup>a</sup>. Me. Rosenilda Marques da Silva Felipe  
FACET – UFGD

Dourados, 24 de abril de 2016.

## RESUMO

Neste trabalho, é realizada a prototipação de um sistema para consultas de recomendação de alimentos, a partir de sintomas e fatores informados pelos pacientes. Fatores, como idade e ser praticante do tabagismo, são relevantes para a caracterização de doenças nutricionais, caracterizadas pela deficiência ou excesso de nutrientes na alimentação. Uma pequena ontologia é construída manualmente, a partir de uma fonte de informação majoritariamente não estruturada (sentenças em linguagem natural), para suportar a automação de consultas, cuja recomendação de alimentos é priorizada em função das informações prestadas pelo paciente.

**Palavras-Chave:** Processamento de Linguagem Natural, Ontologia.

## ABSTRACT

In this paper the prototyping of a query system for food recommendation is done, based in informed symptoms and factors by the patients. Factors as age and being a smoker are relevant for the characterization of nutritional diseases, characterized by the deficiency or excess of nutrients in the diet. A small ontology is manually constructed, based in a mostly unstructured information source (natural language sentences), in order to support the automation of queries, which prioritize a food recommendation accordingly to the information provided by the patient.

**Keywords:** Natural Language Processing, Ontology

## SUMÁRIO

<b>1. Introdução</b> .....	1
1.1. Histórico e Motivação .....	1
1.2. Oportunidades e Relevância .....	2
1.3. Objetivos do Trabalho .....	2
1.4. Metodologia Adotada .....	2
1.5. Conteúdo do Trabalho .....	3
<b>2. Fundamentação Teórica</b> .....	4
2.1. Metodologias para construção de ontologias .....	4
2.1.1. O que é ontologia.....	4
2.1.2. Como construir uma ontologia .....	4
2.1.3. Metodologias para construção de ontologias.....	6
2.1.3.1. Methontology .....	6
2.1.3.2. UPON (United Process for ONtologies) .....	7
2.1.3.3. ONIONS (ONthologic Integration of Naïve Source) .....	10
2.2. Aplicações de ontologias no domínio de nutrição.....	10
2.2.1. Projeto FOODS .....	11
2.2.2. Projeto de recomendação dietética para halterofilismo .....	12
<b>3. Desenvolvimento do Trabalho Proposto</b> .....	14
3.1. Seleção de subdomínio de interesse e fontes de informação.....	14
3.2. Extração de conceitos e relações relevantes a partir das fontes selecionadas .....	15
3.3. Construção manual de uma pequena ontologia para relacionar conceitos relevantes do subdomínio de interesse.....	17
3.4. Prototipação da automação de consultas a partir da ontologia construída .....	18
3.4.1. Ferramenta de população da base de dados .....	19

3.4.2. Funcionamento da ferramenta de população .....	21
3.4.3. Ferramenta para suporte de consultas.....	22
3.4.4. Consultas e resultados .....	23
<b>4. Considerações Finais .....</b>	<b>29</b>
4.1. Conclusões.....	29
4.2. Dificuldades Encontradas .....	29
4.3. Trabalhos Futuros.....	30
<b>REFERÊNCIAS .....</b>	<b>31</b>

## LISTA DE FIGURAS

Figura 1 .....	8
Figura 2 .....	9
Figura 3.....	12
Figura 4 .....	15
Figura 5 .....	19
Figura 6 .....	22
Figura 7 .....	24
Figura 8 .....	26

## **1. Introdução**

Neste trabalho é realizada a prototipação de consultas para recomendação de alimentos, a partir de sintomas e fatores informados pelos pacientes. Segundo Weininger(2007), fatores, como idade, sexo e ser praticante do tabagismo, são relevantes para a caracterização de doenças nutricionais, caracterizadas pela deficiência ou excesso de nutrientes na alimentação. A fonte primária das informações, utilizada majoritariamente neste trabalho, é não estruturada (textos em linguagem natural). Esta fonte é complementada por uma fonte secundária baseada em informação estruturada (tabelas) somente para relacionar nutrientes com alimentos, dado que a fonte primária provê pouca informação a respeito. São extraídas manualmente relações das sentenças de um texto sobre doenças nutricionais, para construir uma pequena ontologia para apoiar a realização automática de consultas. A ontologia é implementada a partir de uma base de dados e exhibe algumas características distintas das ontologias em geral. O usuário do protótipo seleciona um ou mais sintomas e, eventualmente uma doença, informa os fatores que fazem parte do questionário e recebe como resposta uma lista de alimentos recomendados. Na geração da resposta são priorizadas recomendações com um maior número de sintomas associados a uma provável doença nutricional, bem como aos fatores relevantes informados.

### **1.1. Histórico e Motivação**

De acordo com Jones, Bench-Capon e Visser (1998), a utilização de ontologias na área da computação, cresce em conjunto com a necessidade da criação de bases de conhecimento a respeito de um domínio. A construção de um modelo de domínio, representado por uma ontologia é a base para a realização de consultas a partir de informações extraídas de textos em linguagem natural. Com o grande número de informações disponíveis principalmente na Web, é imprescindível buscar uma maneira de interpretar, a nível semântico, essas informações escritas em linguagem natural, tornando-as interpretáveis por máquinas.

Utilizando-se de ontologias, diversos conceitos de aplicações se tornaram possíveis. Dentre eles, um sistema capaz de automatizar consultas a respeito de um domínio, baseando-se em informações fornecidas pela ontologia.

## 1.2. Oportunidades e Relevância

Textos que abordam a área nutricional, especificamente doenças nutricionais, apresentam, em abundância, informações relevantes e diversas relações entre os conceitos de seu domínio. Portanto, é possível a extração dessas relações afim da construção manual de uma ontologia. A partir da ontologia, é possível a criação de um sistema capaz de realizar consultas automatizadas, baseando-se em informações fornecidas pelo usuário. Dessa forma, informações como sintomas podem ser associadas a doenças nutricionais encontradas na ontologia, resultando em recomendações de alimentos.

## 1.3. Objetivos do Trabalho

O objetivo geral deste trabalho é o estudo sobre metodologias para construção de ontologias a fim da elaboração de um protótipo de um sistema que suporte consultas sobre doenças nutricionais.

Os objetivos específicos são:

1. Identificar heurísticas relevantes, nas metodologias estudadas, que possam ser utilizadas na construção de uma pequena ontologia, para suportar consultas.
2. Construir manualmente uma pequena ontologia a partir de informação majoritariamente não estruturada;
3. Prototipar a automação de consultas a partir da ontologia construída.

## 1.4. Metodologia Adotada

A metodologia adotada pode ser descrita em 5 etapas principais:

- **Primeira etapa:** Estudo de Metodologias para construção de ontologias extraídas a partir de artigos relacionados à área da computação. Assim como o estudo de aplicações no domínio de nutrição como base para correlacionar conceitos a partir de fontes textuais
- **Segunda etapa:** Seleção de subdomínio de interesse, de fonte primária não estruturada (texto em linguagem natural) e de fonte secundária estruturada (tabelas)
- **Terceira etapa:** Estudo, categorização e extração de conceitos e relações relevantes a partir das fontes selecionadas

- **Quarta etapa:** Construção manual de uma pequena ontologia para relacionar conceitos relevantes do subdomínio de interesse
- **Quinta etapa:** Prototipação da automação de consultas a partir da ontologia construída

### **1.5. Conteúdo do Trabalho**

No capítulo 2 é caracterizada a fundamentação teórica do trabalho.

No capítulo 3 são descritas cada uma das etapas realizadas durante o desenvolvimento do trabalho.

No capítulo 4 são comentados os resultados alcançados e propostas para trabalhos futuros.

## **2. Fundamentação Teórica**

Nesta seção são apresentados: o conceito de ontologia e metodologias para a sua construção.

### **2.1 Metodologias para a Construção de Ontologias**

Uma metodologia para construção de ontologias descreve o conjunto de heurísticas e técnicas abordadas durante a construção de uma ontologia. Foi realizado um levantamento de artigos relacionados à área de construção de ontologias, com o intuito de analisar e extrair informações que auxiliassem na construção de nossa ontologia.

#### **2.1.1. O que é ontologia**

De acordo com Corcho (2003), uma ontologia é uma descrição dos conceitos de um domínio em específico, contendo os seus atributos, restrições e relações entre eles. Além disso, uma ontologia deve ser interpretável por máquinas e suas informações devem refletir as noções aceitas por um grupo de pessoas, e não apenas um indivíduo.

Uschold e Grüninger (1996) descrevem a utilização de ontologias como uma ferramenta de comunicação, onde as ontologias podem ser usadas como uma ferramenta de transmissão de informações entre grupos de pessoas, uma vez que os conceitos abordados devem ser interpretados da mesma maneira, independente de fatores como língua ou classe social. Além disso, ontologias são capazes de estabelecer uma rede de relacionamentos entre seus conceitos, possibilitando a navegação e exploração do domínio de forma organizada, o que pode ser usado por outros softwares.

#### **2.1.2. Como construir uma ontologia**

Öhrngen e Sandkuhl (2005) descrevem a construção de ontologias como um processo que foi pesquisado ao decorrer dos anos. Dados os diversos domínios que podem ser abordados e inúmeras utilidades que a ontologia pode fornecer, é esperado que diferentes metodologias fossem desenvolvidas por pesquisadores de áreas de atuação distintas. Com o intuito de extrair informações de fontes estruturadas e não estruturadas, foi realizado um estudo a respeito das heurísticas e técnicas encontradas em metodologias para a construção de ontologias.

Inicialmente analisou-se várias metodologias em busca daquela que fosse mais apropriada para a nossa utilização. Para tanto, as metodologias foram analisadas sob a premissa da construção de uma ontologia, que iria se fundamentar na junção de informações extraídas de fontes estruturadas com não estruturadas. No entanto, notou-

se que as heurísticas e técnicas utilizadas pelas metodologias eram descritas, parcialmente, de forma superficial na literatura. Com isso, em busca de um processo de construção eficaz, adotou-se uma abordagem onde o enfoque dos estudos recaíram sobre as técnicas e heurísticas que se mostrassem relevantes ao contexto do trabalho a ser desenvolvido.

Abaixo, descreve-se um passo-a-passo simplificado da construção de uma ontologia segundo Noy e McGuinness (2001).

- 1. Determinar o escopo e domínio da ontologia:** especificação do domínio e finalidade da ontologia, simplificando e acelerando o desenvolvimento da mesma.
- 2. Considerar a reutilização de ontologias existentes:** verificar a existência de ontologias que abordem o mesmo domínio ou que possuam a mesma utilidade operacional da ontologia a ser construída. A reutilização é uma das técnicas que possuem maior potencial na redução de tempo para a construção da ontologia.
- 3. Enumerar os termos importantes da ontologia:** listar todos os conceitos relacionados ao domínio a ser abordado e seus atributos. Essa etapa visa adquirir o maior número de conceitos possíveis para elaboração de uma ontologia completa.
- 4. Definição de classes e hierarquias:** a organização de classes de forma hierárquica deve ser realizada observando a agregação de uma classe A à uma classe B. A hierarquia pode ser desenvolvida a partir dos termos mais abrangentes até os mais específicos (abordagem *top-down*), ou, analogamente agrupar os termos mais específicos inicialmente e posteriormente atribuí-los a um termo geral.
- 5. Definição de propriedade das classes:** definir atributos e restrições utilizados para descrever a classe. Esses atributos podem ser aqueles descritos no passo 3, e as restrições podem incluir restrições de cardinalidade, tipo de valor, entre outros.
- 6. Criação de instâncias:** a última etapa consiste da criação de uma classe ou mais classes dentro das hierarquias definidas na ontologia.

### 2.1.3. Metodologias para construção de ontologias

Durante o estudo das metodologias nos deparamos com heurísticas completamente diferentes e outras muito similares ao passo-a-passo descrito acima. Abaixo, são descritas as metodologias que melhor caracterizam os passos de seus processos.

#### 2.1.3.1. Methontology

Öhrgen e Sandkuhl (2005) ressaltam que a Methontology é considerada uma das metodologias mais maduras dentre as demais desenvolvidas, contudo, possui aspectos que podem ser aprimorados. É bastante utilizada na área de produtos químicos e medicina. Por ser a metodologia mais conhecida, percebe-se que outras metodologias (assim como o passo-a-passo descrito anteriormente) utilizam-na como base, possibilitando uma difusão e melhoria de heurísticas de acordo com o tipo de ontologia a ser construída. Fernández, Gómez-Perez e Juristo (1997), descrevem a Methontology como um processo que pode ser dividido em etapas específicas:

1. **Especificação:** etapa que consiste na definição do propósito, nível de formalidade, escopo e definição de questões de competência.
2. **Aquisição de conhecimento:** essa etapa visa abordar especialistas, livros, figuras, tabelas, manuais, ou outras ontologias como fontes de conhecimento que podem ser usadas em conjunto com técnicas como *brainstorming*, entrevistas, análises de textos formais ou informais para melhor aproveitamento.
3. **Conceitualização:** desenvolvimento do glossário de termos, elaborando e especificando todos os conceitos e possíveis relacionamentos existentes dentro do subdomínio.
4. **Integração:** com o intuito de acelerar o processo de construção da ontologia, avaliar a possibilidade de reutilização de definições encontradas em outras ontologias.
5. **Implementação:** implementação da ontologia com o auxílio do ferramental apropriado (específico para ontologias).

## 6. Avaliação:

- a. **Verificação** - é o processo técnico que garante a qualidade da ontologia como uma software.
- b. **Validação** - é o processo de análise quanto ao conteúdo da ontologia, com a utilização das questões de competência a ontologia tem seu desempenho avaliado.

Heurísticas relevantes incluem:

- Questões de competência: são questões elaboradas durante a definição do escopo. Essas questões devem ser elaboradas com o intuito de verificar a competência da ontologia para fornecer respostas de forma íntegra, ou seja, garantir que todas as informações que possam ser relevantes estejam presentes na ontologia.
- Avaliação: é o processo em que avalia-se se a ontologia fornece o que foi proposto pelas questões de competência.

### 2.1.3.2. UPON (United Process for ONtologies)

Segundo Nikola, Missikoff e Navigli (2009), UPON é uma metodologia incremental, baseando-se principalmente no Processo Unificado de Desenvolvimento de Softwares (UP) e a Linguagem de Modelo Unificada (UML). A metodologia é conhecida por ser utilizada em sistemas automatizados (serviços semânticos Web e agentes inteligentes).

Assim como na Engenharia de Software a UPON utiliza casos de uso e um processo iterativo cíclico incremental, possibilitando o detalhamento adicional da ontologia. Na UPON é possível observar a utilização de ciclos, fases, iterações e fluxos de trabalho (*workflow*) seguindo os conceitos do UP.

Cada ciclo é constituído por 4 fases:

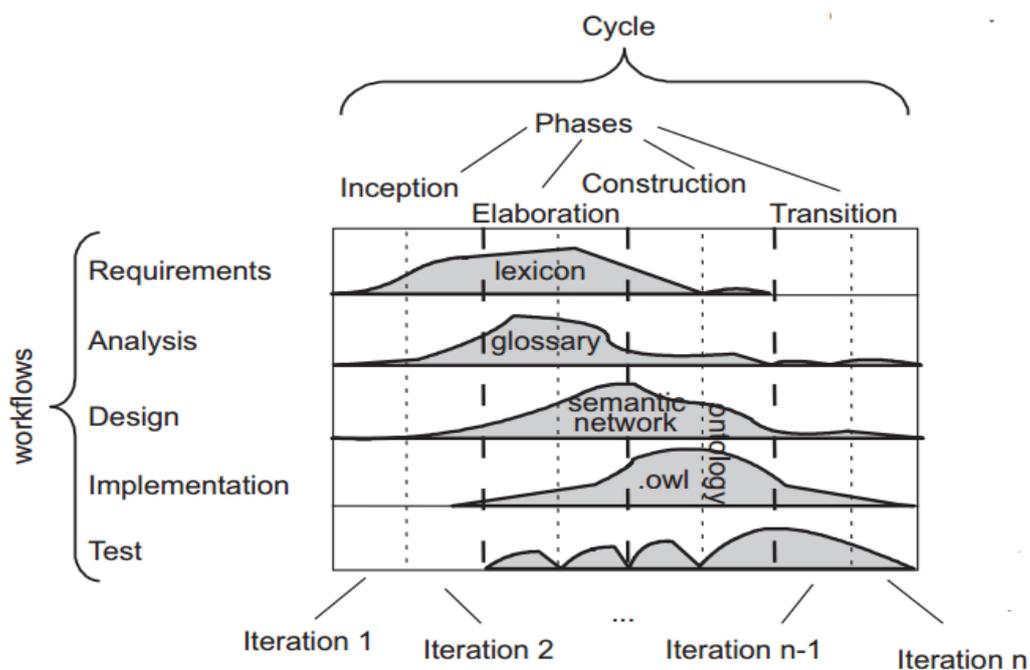
- Início (*Inception*);
- Elaboração (*Elaboration*);
- Construção (*Construction*);
- Transição (*Transition*).

Durante essas fases a ontologia é construída de forma incremental, inicialmente adicionando alguns conceitos e realizando testes de competência ao final de cada ciclo, o que levará a adição de novos conceitos de acordo com a necessidade da ontologia.

Essas fases estão presentes dentro dos 5 workflows:

- Workflow de requerimentos;
- Workflow de análise;
- Workflow de desenho;
- Workflow de implementação;
- Workflow de testes.

Figura 1: representação do ciclo de desenvolvimento da metodologia UPON



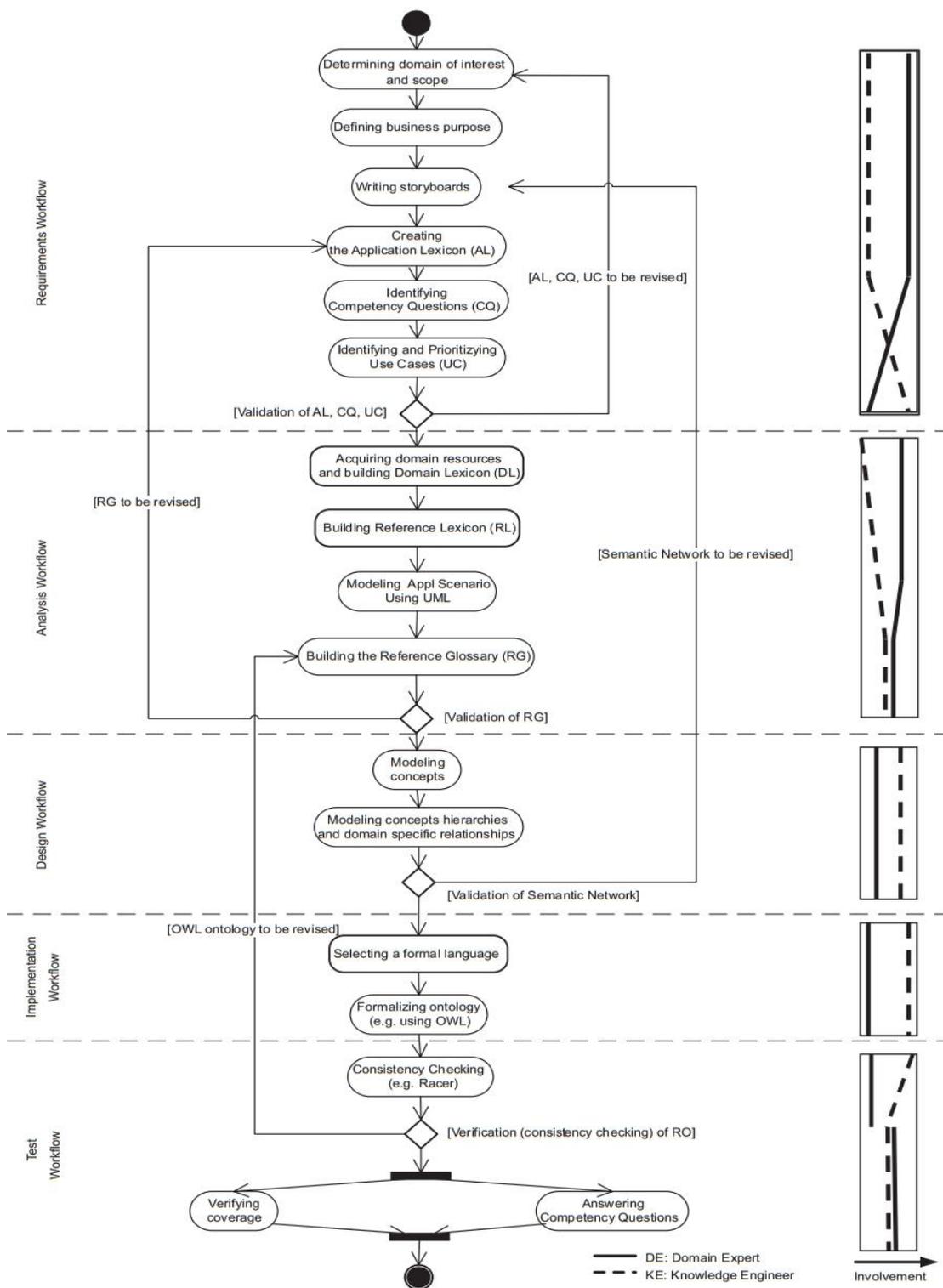
Fonte: NICOLA, MISSIKOFF, NAVIGLI, 2009.

A metodologia UPON é bastante pragmática, atribuindo diversas atividades relacionadas à construção da ontologia durante cada *workflow*. No entanto, como é possível observar na figura 1, assim como nas metodologias descritas na Engenharia de Software, os *workflows* podem ser utilizados em graus diferentes de acordo com a etapa do processo de desenvolvimento.

Na fase de Concepção (*Inception*), a ênfase é dada na definição do léxico para caracterizar os conceitos relevantes, com o início da construção de um glossário para caracterizar cada termo escolhido. Na fase de Elaboração (*Elaboration*) ocorrem a consolidação dos conceitos envolvidos (léxico) e do glossário desses conceitos, a construção de parte substancial da rede semântica envolvendo os conceitos e as relações relevantes e o início da representação da rede semântica na linguagem OWL; e se inicia o teste da implementação em OWL. Na fase de Construção (*Construction*) ocorre a consolidação da rede semântica e de sua representação na linguagem OWL, e continua o teste da implementação em OWL. Na fase de Transição (*Transition*) são feitas correções finais na implementação em OWL e completados os testes na implementação.

O *workflow* de Requisitos (*Requirements*), materializado na construção do léxico (conceitos relevantes), se inicia na fase de Concepção, se intensifica na fase de Elaboração e sofre alguma revisões na fase de Elaboração. O *workflow* de Análise, materializado na construção do glossário dos termos do léxico, ocorre em grande parte na fase de Elaboração, embora tenha se iniciado na fase de Concepção, com algumas atualizações nas fases posteriores. O *workflow* de Projeto, materializado na construção da rede semântica que representa a ontologia, ocorre basicamente nas fases de Elaboração e Construção. O *workflow* de Implementação se inicia na fase de Elaboração e finaliza na fase de Transição, mas ocorre majoritariamente na fase de Construção. Finalmente, o *workflow* de teste, ocorre em paralelo com o *workflow* de Implementação, com ênfase nas fases de Construção e Transição.

Figura 2: atividades presentes nos *workflows* da metodologia UPON.



Fonte: NICOLA, MISSIKOFF, NAVIGLI, 2009.

A figura 2 descreve todas as atividades presentes em cada *workflow*, seguindo a ordem do processo de construção da ontologia, assim como o envolvimento dos profissionais presentes durante o processo, que são o especialista do domínio estudado e o engenheiro de conhecimento.

Heurísticas relevantes:

- Processo incremental: utilização de processos incrementais para construção da ontologia. Bastante útil para quem irá construir sua primeira ontologia, uma vez que o *feedback* obtido no final de cada iteração permite que haja uma evolução por parte do construtor da ontologia.
- Modelagem da ontologia: o processo de elaboração em conjunto com o processo de desenho da ontologia é um conceito que pode ser utilizada para organizar e interpretar a ontologia com mais facilidade.

### 2.1.3.3. ONIONS (ONtologic Integration Of Naïve Source)

Segundo Steve, et al. (1997) essa metodologia foi criada com o intuito de analisar e integrar ontologias ou base de dados. É usada em áreas de subdomínios heterogêneos como a medicina.

É possível resumir as principais atividades da metodologia em alguns passos:

1. Criar um *corpus* de fontes de texto a respeito de um domínio. Esses textos devem possuir uma validação a respeito das noções do domínio a ser analisado.
2. Análise taxonômica: semelhante a definição de hierarquia nas metodologias vistas anteriormente.
3. Análise conceitual dos termos definindo atributos e restrições (baseando-se em informações fornecidas pelo corpus).
4. Integração com outras ontologias ou base de dados com o intuito de enriquecer ou corrigir as definições locais.
5. Implementação, classificação e validação do modelo construído.

Heurísticas relevantes:

- Integração entre ontologias: a possibilidade de integração completa entre ontologias é algo que vem sendo cada vez mais explorado, pois além de possibilitar a redução de trabalhos redundantes, evitando equívocos a respeito dos conceitos de nossas ontologias.

## 2.2. Aplicações de Ontologias no Domínio de Nutrição

Durante o desenvolvimento do trabalho, buscou-se explorar a utilização de ontologias no domínio de nutrição em outras aplicações, com o propósito de estudar e extrair práticas ou conceitos que se mostrassem relevantes ao nosso trabalho. O enfoque do estudo foi direcionado a duas aplicações:

### 2.2.1. Projeto FOODS (Food-Oriented Ontology-Driven System)

O projeto FOODS, proposto por Snae e Brückner (2008), tem como objetivo criar um sistema para auxiliar na recomendação de cardápios para consumidores e clientes de restaurantes. Para tanto, o usuário deve realizar o cadastro no sistema, informando as preferências de comida, ingredientes a serem evitados e doenças possuídas.

A aplicação utilizou como fonte de dados o banco de dados USDA (United States Department of Agriculture) National Nutrient Database for Standard e o vocabulário multilinguístico AGROVOC. Com a intenção de desenvolver a própria ontologia, utilizaram uma mescla das abordagens *bottom-up* e *top-down*. Além disso, a ferramenta Protégé, um editor e *framework* utilizado na construção de ontologias foi empregado em conjunto com a linguagem OWL (Web Ontology Language).

Durante a criação da ontologia, foi criado um questionário com questões de competência de dificuldade distintas. O objetivo desse questionário foi verificar a capacidade e complexidade das questões que o sistema deveria ser capaz de responder.

- Questão simples: *if users have entered a specific disease, such as diabetes, in the user interface, is it appropriate when they choose rice as the main staple?*  
Tradução para o português: Se usuários que entraram com uma doença específica, tal como diabetes, na interface do usuário, é apropriado quando escolhem arroz como prato principal?
- Questão complexa: *Which is the best choice to replace Brussels sprouts in the Chinese cuisine?*  
Tradução para o português: Qual é a melhor escolha para substituir couve de Bruxelas na culinária chinesa?

Com o propósito de aprofundar o entendimento sobre a utilização de ontologias, foram buscadas referências a respeito da ontologia usada e do sistema desenvolvido pelo projeto FOODS. No entanto, não foi possível encontrar muitas informações sobre suas práticas e conceitos, tanto no artigo quanto em outras fontes correlacionadas da literatura.

## 2.2.2. Projeto de recomendação dietética para halterofilismo

De acordo com Tumnark, et al. (2013), este projeto visava a criação de um sistema especialista capaz de realizar recomendações nutricionais personalizadas, possibilitando que os atletas alcançassem suas metas de treinamento. A ontologia utilizada na aplicação foi desenvolvida com a utilização da ferramenta Protégé em conjunto com a linguagem OWL e da SWRL (Semantic Web Rule Language). Modelando os conceitos relacionados aos atletas, alimentos, esportes praticados e nutrição.

Foi desenvolvida uma aplicação em Java para a inserção de dados e interação com a ontologia. A interface principal é capaz de receber dados relacionados ao usuário e seu objetivo de treinamento (sexo, peso, altura, massa muscular, gordura corporal, entre outros). Com essas informações, o sistema é capaz de realizar recomendações mais relevantes ao usuário, abrangendo áreas diferentes além da área nutricional.

Figura 3: exemplo de consulta realizado pelo sistema especialista.

Scenarios for energy needs	Queries
User A is male, prefers meats, has 63kg, 162cm, He is 25 years old, body fat is 13% muscle mass is 54.8kg, his objective is 62kg, the training phase is general preparation: hypotrophy and he wants main dish lunch menu.	If(gender = male), athletefav= meat and heighobj = (62-61) and age = 25 and bfat = 13 and mmass = 54.8 and train = HTrophy and menu = mdish and time = lunch)  Then recommendedmenu = (hasEnergy(600 <= energy <=660) and hasMeat)

Fonte: TURMNARK, et al., 2013.

A parte esquerda da tabela da figura 3 exibe um exemplo da entrada do usuário em forma não-estruturada, onde, através dessas informações o sistema especialista seria capaz de transformá-la em uma linguagem estruturada, o que é demonstrado na parte superior da direita da tabela. Logo abaixo, verifica-se a recomendação fornecida pelo sistema, onde se observa:

- Then recommended menu = (hasEnergy(600 <= energy <=660) and hasMeat). Traduzido para o português apresenta: então menu recomendado possui energia maior igual a 600 e menor igual a 660 e contém carne.

Novamente, com o intuito de detalhar o estudo a respeito da utilização de ontologias e suas práticas de criação, buscou-se, sem sucesso informações a respeito da ontologia e do sistema utilizado pelo projeto de recomendação dietética em outras fontes de informação.

### **3. Desenvolvimento do trabalho proposto**

A etapa de desenvolvimento foi dividida em duas partes:

- Construção manual de uma pequena ontologia para relacionar conceitos relevantes do subdomínio de interesse.
- Prototipação da automação de consultas a partir da ontologia construída.

A seguir são descritas as etapas e heurísticas aplicadas durante o desenvolvimento do trabalho.

#### **3.1. Seleção de subdomínio de interesse e de fontes de informação**

A seleção da fonte de informação é uma etapa fundamental durante a construção de uma ontologia. Para realização dessa etapa foi necessário, inicialmente, determinar o domínio e subdomínio que seriam abordados pela ontologia.

O domínio de interesse escolhido foi a nutrição, e o subdomínio a recomendação de alimento baseados em sintomas e fatores relacionados. Essa escolha foi realizada com a intenção de explorar a associação de textos em linguagem natural (fonte não estruturada) com tabelas (fonte estruturada), uma vez que textos desse subdomínio comumente possuem informações estruturadas e não estruturadas.

A fonte primordial deste trabalho é não estruturada (sentenças em linguagem natural): um texto da Enciclopédia Britânica abordando doenças nutricionais causadas por deficiência ou excesso de nutrientes.

A fonte estruturada é secundária e seu único objetivo é agregar informação qualitativa. A fonte selecionada foi o site “<http://www.healthaliciousness.com>”, que apresenta uma relação dos alimentos que possuem a maior quantidade do nutriente especificado. Essa informação é relevante, principalmente, na etapa final da consulta, onde a recomendação de alimento é realizada.

Figura 4: exemplo de relação de alimentos.



**#1: Sweet Potato (Cooked)**

Vitamin A in 100g	Per cup (200g)	Per medium potato (114g)
19218IU (384% DV)	38436IU (769% DV)	21909IU (438% DV)

**Other Types of Sweet Potato High in Vitamin A (%DV per cup):** Frozen Sweet Potato, cooked, cubed (578%), Canned Sweet Potato (444%), and Raw Sweet Potato, cubed (377%). [Click to see complete nutrition facts.](#)



**#2: Carrots (Cooked)**

Vitamin A in 100g	Per cup, sliced (156g)	Per carrot (46g)
17033IU (341% DV)	26572IU (532% DV)	7835IU (157% DV)

**Other Types of Carrot High in Vitamin A (%DV per cup):** Frozen Carrots, cooked, cubed (494%), and Raw Carrots, sliced (408%). [Click to see complete nutrition facts.](#)

Fonte: <http://www.healthaliciousness.com>, acesso em abril, 2016.

A figura 4 apresenta um exemplo da relação de alimentos disponíveis na fonte secundária. No exemplo o nutriente buscado foi a vitamina A, e como resultado foram retornados *Sweet potato* (batata doce) e *Carrots* (cenouras), assim como a quantidade de nutrientes dependendo do modo de preparo de cada alimento.

### 3.2. Extração de Conceitos e Relações Relevantes a partir das fontes selecionadas

Uma relação liga duas ou mais conceitos por meio de um relacionamento relevante ao domínio. Existem diferentes tipos de relações, que variam, principalmente, devido ao relacionamento entreos conceitos.

**Relação de associação:** é uma relação entre dois conceitos que não pertencem a uma hierarquia de especialização. (Conceito A -- relacionamento -- Conceito B).  
Exemplo: dietary fiber -- helps reduce -- blood cholesterol level

**Relação de especialização:** é uma relação entre um conceito genérico e um conceito específico. (Conceito genérico: Conceito específico).

Exemplo: protein : meat - poultry - seafood - beans - peas - eggs

**Relação de agregação:** é a relação entre um conceito total e seus conceitos parte. (Conceito total - Conceito parte)

Exemplo: apple: seed - skin - stem

As fontes de informações selecionadas possuem uma grande variedade de relacionamentos e conceitos. Portanto, depois de um estudo detalhado do texto, analisando os relacionamentos existentes entre seus conceitos, assim como, verificando possíveis novos encadeamentos, foi possível a categorização dos conceitos em *symptom* (sintoma), *disease* (doença), *cause* (causa), *nutrient* (nutriente) e *food* (alimento). Outro conceito considerado necessária foi o conceito *human factor* (fator humano), que apresenta fatores humanos que podem influenciar a caracterização da doença a partir de sintomas e, portanto, a recomendação de alimentos.

Comos conceitos categorizados, foi necessário também especificar os tipos de relação que seriam relevantes a ontologia. Utilizando das informações obtidas no texto dividiu-se as relações em dois grupos, com base na quantidade de conceitos envolvidos em uma relação: relações binárias e ternárias.

Relações binárias: são relações que descrevem um relacionamento entre dois conceitos.

Abaixo são citados cinco exemplos de **relações binárias**:

- symptom – disease  
→ symptom / weak bones - is caused by - disease / osteoporosis
- nutrient – food  
→ nutrient / vitamin D - can be found in - food / bread
- cause – disease  
→ cause / iodine deficiency - causes - disease / goitre
- human\_factor – disease  
→ human\_factor / diet without fruits and vegetables - increases risk of - disease  
/ scurvy
- symptom – cause  
→ symptom / night blindness - is caused by - cause / vitamin A deficiency

**Relações ternárias:** são relações que descrevem o relacionamentos entre três conceitos.

- cause - human\_factor - disease
  - ➔ cause / calcium deficiency - during - human\_factor / childhood - causes - disease / rickets
- cause - human\_factor - symptom
  - ➔ cause / severe iron deficiency - during - human\_factor / childhood - causes - symptom / growth impairment

### **3.3. Construção manual de uma pequena ontologia para relacionar conceitos relevantes do subdomínio de interesse**

O principal objetivo deste trabalho é gerar recomendações de alimentos a partir de sintomas e fatores humanos fornecidos pelo usuário. Para alcançar tal objetivo, durante a elaboração da ontologia, a heurística de elaboração de questões de competência foi substituída por uma metáfora de consulta a um nutricionista. Na consulta o paciente (usuário) fornece sintomas, doenças conhecidas e outros fatores, de acordo com uma lista fornecida, que podem influenciar no diagnóstico e, portanto, na recomendação de alimentos. Alguns exemplos de fatores incluem: praticante do tabagismo, gestante, pessoa idosa, pouca exposição ao sol, entre outros.

Encadeando conceitos, a partir de relações binárias, é possível recomendar alimentos a partir de um dado sintoma informado pelo usuário. Um resultado possível é:

- symptom : joint pain
- disease : scurvy
- cause : vitamin C deficiency
- nutrient : vitamin C
- food : orange

Fatores humanos podem se compor com as relações encontradas, reforçando indicação de uma dada doença ou causa. Por exemplo, a doença “scurvy” pode ser reforçada como alternativa de resposta à consulta, se o paciente informa o fator humano “diet with few fruits and vegetables”.

Durante as etapas iniciais do desenvolvimento do trabalho, foi utilizada uma padronização semelhante a linguagem natural.

Exemplos:

- vitamin D deficiency - causes - osteomalacia
- thiamin deficiency – is found in – people with diets of white rice and flour
- folic acid deficiency – may cause (in early pregnancy) – neural tube defects in the fetus

No entanto, o nome da associação, extraído originalmente do texto, pode ser substituído, com vantagem do ponto de vista de uma implementação genérica a partir de uma base de dados, pela composição do nome das conceitos envolvidas na relação. No exemplo “vitamin D deficiency – causes – osteomalacia” o nome da relação passa a ser: “cause-disease”.

Em geral ontologias são representadas a partir de relações binárias. No entanto, a partir da fonte utilizada foi possível identificar relações ternárias, com por exemplo:

- cause-symptom-human\_factor
  - ➔ cause / vitamin E deficiency – symptom / lungs oxidative damage – factor / smoker

Assim sendo, a ontologia foi implementada em uma base de dados com duas alterações fundamentais:

- O nome original das relações foi padronizado para indicar os conceitos envolvidos: symptom-disease, nutrient-food;
- Além de relações binárias, foram representadas relações ternárias.

### **3.4. Prototipação da Automação de Consultas a partir da Ontologia Construída**

Para a realização das consultas e do preenchimento da base de dados, foram criadas ferramentas com o uso da linguagem de programação Java, banco de dados MySQL e o ambiente de desenvolvimento integrado NetBeans.

A linguagem de programação Java foi selecionado para ser a linguagem utilizada nesse projeto devido, principalmente, a afinidade com a linguagem, e por ser uma linguagem onde é possível realizar o desenvolvimento tanto para a plataforma Unix quanto Windows, uma vez que o Java Virtual Machine que o próprio Java™ oferece, permite que a aplicação seja executada em ambos os ambientes.

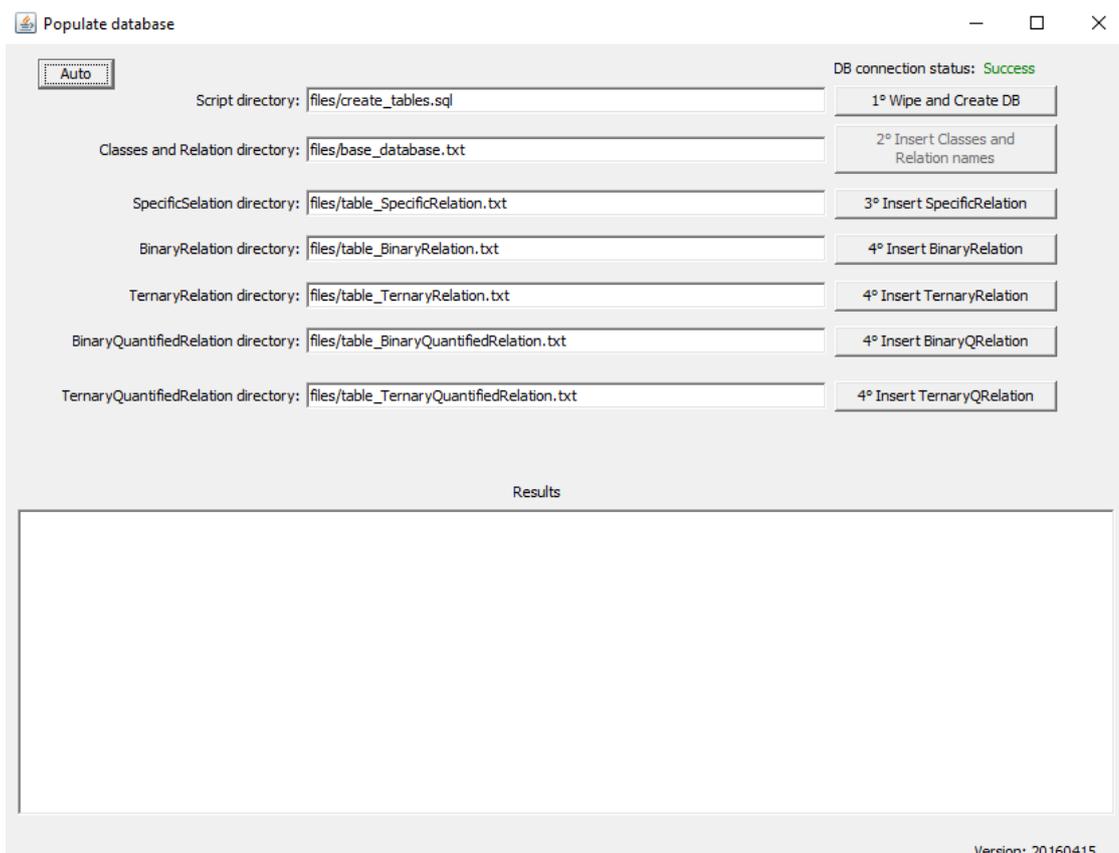
O sistema de gerenciamento de banco de dados MySQL foi selecionado, por ser gratuito sob a Licença Pública GNU e também pela familiaridade com a ferramenta.

Já o ambiente de desenvolvimento NetBeans foi selecionado para agilizar o desenvolvimento das ferramentas, tanto de população quanto de consultas. Os principais fatores que motivaram a escolha pelo NetBeans foram a sua acessibilidade, o suporte à linguagem Java e ao banco de dados MySQL e a oferta do GUI (Graphical User Interface) Builder, que facilita a criação da interface gráfica com usuário.

### 3.4.1. Ferramenta de população da base de dados

Após a extração e padronização das relações, o próximo passo é popular a base de dados. Foi criada, portanto, uma ferramenta com interface gráfica, demonstrada na figura 5, capaz de receber arquivos pré-formatados como entrada. Os arquivos apresentam tanto a configuração da base de dados quanto as informações extraídas das fontes de informações.

Figura 5: interface gráfica da ferramenta de população da base de dados



Fonte: elaborada pelo autor.

### **Exemplo de arquivo 1:**

Arquivo: create\_tables.sql

Possui linhas de comando utilizadas para a criação do banco de dados.

Exemplo de linha de comando:

```
CREATE TABLE ClassName(  
    id int NOT NULL PRIMARY KEY AUTO_INCREMENT,  
    class VARCHAR(255) NOT NULL  
);
```

Esse exemplo irá criar uma tabela no banco de dados que será utilizado para armazenar o nome das conceitos.

### **Exemplo de arquivo 2:**

Arquivo: base\_database.txt

Dentro do arquivo existem:

1. O nome dos tipos de relações -
  - a. RelationTypes : generalization
2. O nome das classes -
  - a. ClassName : vitamin C
3. O nome das relações -
  - a. RelationName : cause-disease

Esses exemplos apresentam as informações que atuam como suporte para a criação posterior de tuplas em outras tabelas, pois as tuplas descritas no exemplo de arquivo 2 possuem chaves estrangeiras que são fundamentais para o encadeamento de informações de outras tabelas.

### **Exemplo de arquivo 3:**

Arquivo: table\_BinaryRelation.txt

Dentro do arquivo existem:

1. As relações binárias a serem adicionadas no banco
  - a. nutrient-food#biotin#tomato

Onde: *nutrient-food* é o nome da relação.

*biotin* é o primeiro conceito da relação(nutrient).

*tomato* é o segundo conceito da relação(food).

### **3.4.2. Funcionamento da ferramenta de população**

Os passos a seguir devem ser seguidos para o funcionamento da ferramenta:

1. Criação da base de dados e tabelas
2. Realizar a inserção dos conceitos na base de dados.
3. Realizar a inserção da generalização dos conceitos adicionados.
4. Realizar a inserção de todos os tipos de relações anteriormente criadas.

Alternativamente o botão “Auto” foi adicionado, realizando todas as etapas listadas de forma automática.

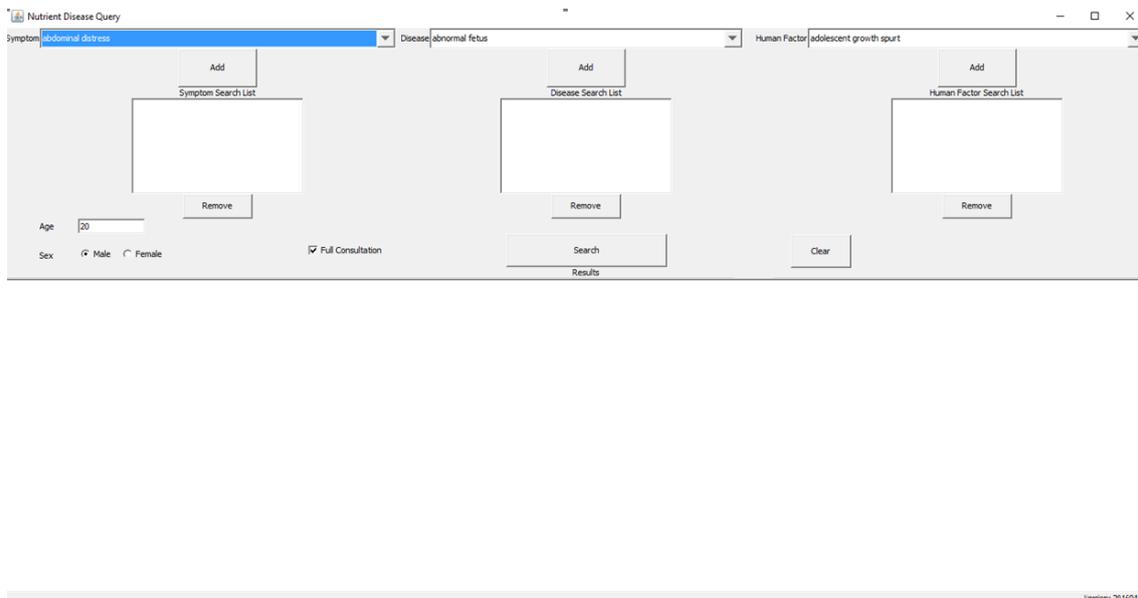
A base de dados representa a ontologia implementada, contendo diversos tipos de relações binárias e ternárias entre os conceitos, assim como uma hierarquia que é utilizada para auxiliar na consulta.

### **3.4.3. Ferramenta para suporte de consultas**

O desenvolvimento da ferramenta foi realizado com o intuito de simular uma consulta a um nutricionista, onde o usuário fornece informações para auxiliar o nutricionista a encontrar uma recomendação correspondente às informações fornecidas. A ferramenta dispõe, portanto, de uma interface gráfica, demonstrada na figura 6, com vários campos de entrada:

- Sintomas: uma lista dos possíveis sintomas armazenados na base de dados.
- Doenças: uma lista de doenças armazenados na base de dados.
- Fatores humanos: uma lista de fatores humanos armazenados na base de dados.
- Idade: idade do usuário
- Sexo: sexo do usuário.

Figura 6: interface gráfica da ferramenta para suporte de consultas



Fonte: elaborada pelo autor.

O usuário deve, então, selecionar itens através das listas disponibilizadas e adicioná-los a consulta; é necessário também, que a idade e o sexo sejam determinados antes da consulta. Durante o procedimento, o usuário pode remover conceitos indesejados individualmente ou coletivamente durante a consulta através dos botões “Remove” e “Clear” respectivamente, assim como alterar os dados idade e sexo.

Após a inserção dos dados, o usuário é capaz de realizar a consulta a partir do botão “Search”, que irá gerar os resultados em um campo de texto na parte inferior da interface.

#### 3.4.4. Consultas e resultados

Os resultados apresentados representam encadeamentos entre conceitos que estejam interligados por relações na ontologia. O ponto de partida é sempre o sintoma fornecido pelo usuário da consulta. O encadeamento mais longo existente é: symptom – disease – cause – nutrient – food. Uma resposta baseada neste encadeamento caracteriza que o sintoma que o usuário informou pode estar associado a uma doença, ocasionada pela deficiência de um nutriente, que pode ser encontrado em um alimento, rico neste nutriente.

Pelo fato das relações existentes na ontologia, serem provenientes do texto do corpus utilizado, nem sempre o encadeamento mais longo é retornado como resposta. Esta situação ocorre quando existem relações que interligam diretamente, por exemplo, o conceito sintoma a nutriente, ou conceito doença a alimento.

O usuário pode informar fatores humanos, que podem ser influenciar a composição dos encadeamentos. Um paciente com ingestão significativa de bebidas alcoólicas pode ter um propensão maior para doenças cardiovasculares. Uma mulher muçulmana, menos exposta a luz solar, pode apresentar maior propensão à deficiência de vitamina D. Alguns tipos de sintomas, doenças ou deficiências de nutrientes podem ser influenciados pela idade ou sexo do paciente. Nestes casos, a resposta leva em conta os fatores humanos informados, que tenham relações com os conceitos utilizados no encadeamento retornado. No entanto, a resposta é meramente qualitativa, pois a associação de probabilidades nas respostas fornecidas está fora do escopo deste trabalho, porque não está, em geral, disponível no corpus utilizado como fonte para a construção da ontologia.

Adicionalmente, uma das funcionalidades primordiais da ferramenta é a capacidade de priorizar recomendações. A priorização deve ocorrer em dois casos:

- Quando duas ou mais doenças são encontradas durante a consulta, aquela com o maior número de sintomas deve ser priorizada.
- Quando duas ou mais doenças são encontradas durante a consulta, aquela que possui um ou mais fatores humanos relevantes relacionados deve ser priorizada.

Figura 7: Exemplo de resultado da consulta com um único sintoma *Joint pain*.

```
symptom: joint pain
cause: fluoride excess
mineral: fluoride
food: carrot, spinach, asparagus, pickles, tomato, grape juice, apple juice, orange juice, grapefruit juice, beets

symptom: joint pain
cause: fluoride excess
disease: skeletal fluorosis
mineral: fluoride
food: carrot, spinach, asparagus, pickles, tomato, grape juice, apple juice, orange juice, grapefruit juice, beets

symptom: joint pain
disease: scurvy
vitamin: vitamin C
food: broccoli, kiwifruit, kale, green peas, yellow bell pepper, guavas, strawberry, tomato, papaya, orange

symptom: joint pain
disease: scurvy
cause: vitamin C deficiency
vitamin: vitamin C
food: broccoli, kiwifruit, kale, green peas, yellow bell pepper, guavas, strawberry, tomato, papaya, orange
```

---

Fonte: elaborada pelo autor

A figura 7 mostra um exemplo do resultado retornado através da consulta com o dado de entrada *Joint pain*.

São chamados de níveis a profundidade do encadeamento da pesquisa.

1. No primeiro nível da pesquisa por *Joint pain* no banco de dados a pesquisa encontrou os resultados *fluoride excess* e *scurvy*.
2. No segundo nível da pesquisa por *fluoride excess* são encontrados os resultados *fluoride* e *skeletal fluorosis*
3. No terceiro nível da pesquisa por *fluoride* os resultados encontrados foram os alimentos: *carrot, spinach, asparagus, pickles, tomato, grape juice, apple juice, orange juice, grapefruit juice* e *beets*. Como os resultados encontrados são alimentos esse encadamento acaba e retorna ao nível anterior.
4. No segundo nível da pesquisa por *skeletal fluorosis*, foi encontrado o resultado *fluoride*, e a partir das pesquisas realizadas com esse resultado, são encontrados os mesmos resultados do encadamento 3.
5. Após as pesquisas anteriores, a consulta retorna ao primeiro nível, porém, realizando então a pesquisa com o resultado *scurvy*.
6. No segundo nível da pesquisa por *scurvy* são encontrados os resultados *vitamin C* e *vitamin C deficiency*.
7. No terceiro nível da pesquisa por *vitamin C* os resultados encontrados são os alimentos: *broccoli, kiwifruit, kale, green peas, yellow bell pepper, guavas, strawberry, tomato, papaya* e *orange*. Como os resultados são alimentos esse

encadeamento acaba e retorna a algum o nível anterior que possua algum resultado a ser pesquisado.

8. No segundo nível da pesquisa por *vitamin C deficiency*, encontrou-se o resultado *vitamin C*, e a partir das pesquisas realizadas com esse resultado são encontrados os mesmos resultados do encadeamento 7.
9. Como não existem mais resultados para realizar encadeamentos, a consulta é encerrada e os resultados são exibidos ao usuário.

Figura 8: Exemplo de resultado da consulta por *bleeding gums, joint pain* e *weak bones* com a adição de fatores humanos *adult* e *diet with few fruits and vegetable*.

symptom: bleeding gums, joint pain  
disease: scurvy  
vitamin: vitamin C  
food: broccoli, kiwifruit, kale, green peas, yellow bell pepper, guavas, strawberry, tomato, papaya, orange  
Revelant: diet with few fruits and vegetables -> scurvy

symptom: bleeding gums, joint pain  
disease: scurvy  
cause: vitamin C deficiency  
vitamin: vitamin C  
food: broccoli, kiwifruit, kale, green peas, yellow bell pepper, guavas, strawberry, tomato, papaya, orange  
Revelant: diet with few fruits and vegetables -> scurvy

symptom: weak bones  
disease: osteomalacia  
cause: vitamin D deficiency  
food: bread, cereals, fortified milk, margarine  
Revelant: adult -> osteomalacia

symptom: joint pain  
cause: fluoride excess  
mineral: fluoride  
food: carrot, spinach, asparagus, pickles, tomato, grape juice, apple juice, orange juice, grapefruit juice, beets

symptom: joint pain  
cause: fluoride excess  
disease: skeletal fluorosis  
mineral: fluoride  
food: carrot, spinach, asparagus, pickles, tomato, grape juice, apple juice, orange juice, grapefruit juice, beets

---

Fonte: elaborada pelo autor

A figura 8 mostra um exemplo do resultado retornado através da consulta com os dados de entrada:

- *bleeding gums*
- *joint pain*
- *weak bones*
- *adult*
- *diet with few fruits and vegetables*

São chamados de níveis a profundidade do encadeamento da pesquisa e entre cada nível são realizadas comparações dos resultados com os fatores humanos inseridos pelo usuário para verificar sua possível relevância.

1. Primeiramente são realizadas pesquisas com os dados de entrada de sintomas, a fim de verificar quais possuem como resultado a mesma doença, e assim agrupá-

- las para mostrar ao usuário a relação das doenças encontradas com os sintomas inseridos.
2. No primeiro nível da pesquisa por *bleeding gums* no banco de dados a pesquisa encontrou o resultado *scurvy*.
  3. No segundo nível da pesquisa por *scurvy* são encontrados os resultados *vitamin C* e *vitamin C deficiency*, e também é encontrada a relevância com o fator humano, *diet with few fruits and vegetables*, inserido pelo usuário. Sendo assim todos os encadeamentos posteriores serão priorizados no durante a exibição dos resultados.
  4. No terceiro nível da pesquisa por *vitamin C* os resultados encontrados são os alimentos: *broccoli, kiwifruit, kale, green peas, yellow beel pepper, guavas, strawberry, tomato, papaya* e *orange*. Como os resultados são alimentos, esse encademento acaba e retorna a algum nível anterior que precise realizar uma pesquisa.
  5. No segundo nível da pesquisa por *vitamin C deficiency*, foi encontrado o resultado *vitamin C* e a partir das pesquisas realizadas com esse resultado, são encontrados os mesmos resultados do encadeamento 4.
  6. Após as pesquisas anteriores, a consulta retorna ao primeiro nível, porém, realizando então a pesquisa com a entrada *joint pain*, encontrando os resultados *scurvy* e *fluoride excess*.
  7. Contudo, o resultado *scurvy* encontrado por *joint pain* é o mesmo resultado encontrado pela pesquisa *bleeding gums*, sendo assim realiza-se o agrupamento do sintoma *joint pain* ao sintoma *bleeding gums*.
  8. No segundo nível da pesquisa por *fluoride excess*, são encontrandos os resultados *fluoride* e *skeletal fluorosis*
  9. No terceiro nível da pesquisa por *fluoride*, os resultados encontrados foram os alimentos: *carrot, spinach, asparagus, pickles, tomato, grape juice, apple juice, orange juice, grapefruit juice* e *beets*. Como os resultados encontrados são alimentos, esse encademento acaba e a busca retorna ao nível anterior.
  10. No segundo nível da pesquisa por *skeletal fluorosis*, foi encontrado o resultado *fluoride*. A partir das pesquisas realizadas com esse resultado são encontrados os mesmos resultados do encadeamento 9.
  11. Então, ainda no primeiro nível é realizada a pesquisa com a entrada *weak bones*.

12. No segundo nível da pesquisa por *weak bones* é encontrado o resultado *osteomalacia*, e também é encontrada a relevância com o fator humano, *adult*, inserido pelo usuário. Assim todos os encadeamentos posteriores serão priorizados durante a exibição dos resultados.
13. No terceiro nível da pesquisa por *osteomalacia* é encontrado o resultado *vitamin D deficiency*.
14. No quarto nível da pesquisa por *vitamin D deficiency*, os resultados encontrados são os alimentos: *bread*, *cereals*, *fortified milk* e *margarine*. Como os resultados são alimentos esse encadamento acaba e retorna a algum nível anterior que precise realizar uma pesquisa.
15. Como não existem mais resultados para realizar encadeamentos, a consulta é encerrada e os resultados são exibidos ao usuário.

#### **4. Considerações Finais**

Nesta seção são apresentadas: as conclusões, as dificuldades encontradas e sugestões para trabalhos futuros.

##### **4.1. Conclusões**

A proposta de construir uma aplicação para simular uma consulta em um nutricionista, alimentada pelas informações de sintomas e fatores relevantes resultou no foco das seguintes hierarquias de classes de conceitos: sintoma, doença, causa, nutriente, alimento e fatores humanos. A utilização destas seis hierarquias de classes, resultou na dispensa da definição de um glossário de termos, conforme previsto na metodologia UPONS, descrita na seção 2.1.3.2.

A heurística de construir uma ontologia baseada na elaboração das questões que a ontologia deve responder, proposta na metodologia Methontology descrita na seção 2.1.3.1, foi adaptada para a realização deste trabalho. A adaptação desta heurística resultou na metáfora da consulta a um nutricionista, no qual os sintomas e fatores humanos fornecidos pelo paciente, direcionam a escolha das relações entre conceitos presentes na ontologia, influenciando a resposta dos conceitos relevantes para a consulta.

O resultado deste trabalho foi a geração de uma aplicação na qual as respostas fornecidas são justificadas com base nos conceitos encadeados a partir de relações existentes na ontologia construída para suportar a aplicação. A informação, por parte do usuário, de mais de um sintoma ou de fatores humanos relevantes, resulta na priorização de doenças associadas a mais de um sintoma ou aos fatores humanos informados.

A utilização do texto da Enciclopédia Britânica como base para o trabalho viabilizou a realização da proposta de simular uma consulta no nutricionista do ponto de vista qualitativo, ou seja, informar as doenças, causas, nutrientes e alimentos que possam estar associados aos sintomas e fatores humanos recomendados. No entanto, a informação é meramente qualitativa, dado que o texto raramente provê informações quantitativas que pudessem ser utilizadas para caracterizar a probabilidade de ocorrências de doenças ou das causas de deficiências de nutrientes retornadas como resposta.

Foram adotadas duas abordagens que não são geralmente encontradas nas ontologias; mas que facilitaram a implementação e não acarretaram prejuízo para as respostas geradas na aplicação proposta: (1) a padronização do nome das relações com foco nos conceitos envolvidos na relação, dispensando o nome do texto original; e (2) a utilização de relações ternárias, envolvendo três conceitos, em vez das relações binárias comumente utilizadas nas ontologias reportadas na literatura.

#### **4.2. Dificuldades encontradas**

A construção da ontologia é, visivelmente, a etapa mais importante do trabalho. Contudo, durante o estudo das metodologias para construção de ontologias, somente uma heurística, foi utilizada para a construção deste trabalho, e mesmo assim, adaptada para a metáfora da consulta no nutricionista.

Outra dificuldade relevante é a dificuldade de encontrar informações detalhadas a respeito de trabalhos correlatados na literatura estudada. Embora existam aplicações que abordam o domínio de nutrição, a reprodução desses trabalhos é inviável, em função da superficialidade das informações reportadas nas publicações estudadas.

Durante a extração de relações da fonte de informações estruturada, a proposta inicial do trabalho visava nomear os relacionamentos entre todos os conceitos de forma independente, com o intuito de preservar as informações contidas no texto. Porém, a

vasta quantidade de informações se mostrou um obstáculo, gerando uma grande diversificação nos nomes de relacionamentos. Para contornar tal dificuldade, foi realizada a padronização dos conceitos e o nome de seus relacionamentos, assim como a adição do uso de relações ternárias. Esse conjunto de decisões permitiu a simplificação das relações de forma a conservar as informações consideradas relevantes durante o estudo da fonte de informações.

### **4.3. Trabalhos futuros**

O grande objetivo desta trabalho foi suportar consultas a partir de uma ontologia e, adicionalmente, adquirir experiência na construção de uma pequena ontologia.

No presente trabalho, foi construída uma ontologia manualmente, que foi utilizada para automatizar consultas, a partir de informações fornecidas pelo usuário do sistema. A padronização do nome das associações não acarretou nenhuma restrição para a realização das consultas.

A extração automática de relações para construção de ontologias é um trabalho futuro de grande utilidade para a construção automática de ontologias.

## REFERÊNCIAS

CORCHO, O; FERNÁNDEZ-LÓPEZ, M; GÓMEZ-PÉREZ, A. **Methodologies, tools and languages for building ontologies. Where is their meeting point?** Data & Knowledge Engineering 46, 2003.

FERNÁNDEZ, M; GÓMEZ-PÉREZ, A; JURISTO, N. **From Ontological Art Towards Ontological Engineering.** AAAI Technical Report SS-97-06. 1997.

NOY, N; MCGUINNESS, D. **Ontology Development 101: A Guide to Creating Your First Ontology.** Stanford University, California. 2001.

ÖHRGEN, A; SANDKUHL, K. **Towards a methodology for Ontology development in small and medium-sized enterprises.** IADIS International Conference on Applied Computing. 2005.

JONES, D; BENCH-CAPON, T; VISSER, P. **METHODOLOGIES FOR ONTOLOGY DEVELOPMENT.** Department of Computer Science, University of Liverpool, 1998.

STEVE, G; GANGEMI, A; PISANELLI, D. **Integrating Medical Terminologies with ONIONS Methodology.** Department of Medical Informatics, Institute for Biomedical Technologies.

USCHOLD, M; GRUNINGER, M. **ONTOLOGIES: Principles, Methods and Applications.** Knowledge Engineering Review, Volume 11 Number 2, Junho 1996.

TUMNARK, Piyaporn et al. **Ontology-Based Personalized Dietary Recommendation for Weightliftin.** International Workshop on Computer Science in Sports, 2013.

NICOLA, A; MISSIKOFF, M; NAVIGLI, R. **A software engineering approach to ontology building.** Information System 34, P.258-275. 2009.

SNAE, C; BRÜCKNER, M. **FOODS: A Food-Oriented Ontology-Driven System. International Conference on Digital Ecosystems and Technologies.** Thailand. P.168-176, Fevereiro 2008.

WHITBREAD, D. **Top 10 lists for the most nutritious foods.** Disponível em: <<https://www.healthaliciousness.com/most-nutritious-foods-lists.php>>. Acesso em março de 2016.

WEININGER, J. **Nutritional disease. 2007.** Disponível em: <<http://global.britannica.com/science/nutritional-disease>>. Acesso em abril de 2015.