

Pós-Graduação em
Gestão da Informação
e Marketing
Intelligence

Representação
do
Conhecimento

Grupo:

Susana Ribeiro

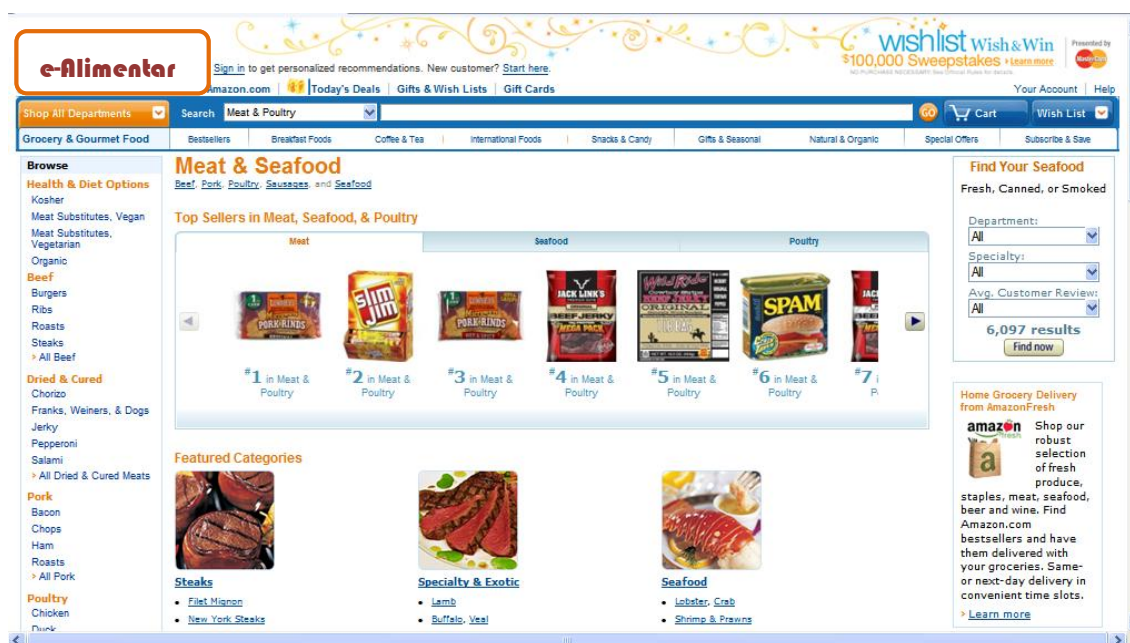
Susana Leal

Ontologia
Loja Online **e-Alimentar**

Outubro 2010

Pós-Graduação em Gestão da Informação e Marketing Intelligence 3ª Edição 2010/2011

“Ontologia Loja Online **e-Alimentar**”



Trabalho Prático realizado no âmbito da disciplina **Representação do Conhecimento**

Docente: Eng. José Carlos Ramalho

Grupo: Susana Ribeiro e Susana Leal

Outubro 2010

II. ÍNDICE

I. APRESENTAÇÃO	1
II. ÍNDICE	2
III. INTRODUÇÃO	3
IV. DOMÍNIO E TUTORIAL	3
V. DESENVOLVIMENTO	4
1. ÂMBITO	4
2. CONSTRUÇÃO DA ONTOLOGIA	4
2.1 CLASSES	4
2.2 INDIVÍDUOS	8
2.3 PROPRIEDADES	8
3. ASPECTOS RELEVANTES	9
VI. CONCLUSÕES	10
VII. BIBLIOGRAFIA	11
VIII. ANEXOS	12

III. Introdução

A ontologia desenvolvida neste trabalho, denominada **Ontologia Loja Online e-Alimentar**, avaliará o perfil do cliente numa loja on-line do sector alimentar.

O processo de adesão/identificação do cliente e de compra numa loja on-line gera um universo de informação que, devidamente tratada e avaliada, torna-se num suporte precioso para acções de marketing direccionado. Assim, a **Ontologia Loja Online e-Alimentar** pretende ser uma ferramenta facilitadora da pesquisa, integração e tratamento de toda essa informação, permitindo a segmentação de clientes.

A ontologia desenvolvida permitirá analisar diversos níveis de informação:

- Informação fornecida pelo cliente no processo de adesão, essencialmente constituída por dados demográficos e geográficos.
- Informação comportamental do cliente no processo de compra, como lealdade e frequência
- Informação psicográficas, como avaliação do estilo de vida através do seu cabaz de compras.

Esta ontologia tem como **destinatários** retalhistas e suas equipas de marketing.

Para escolha deste tema foi decisivo o conhecimento profissional nesta área de um dos elementos do grupo. Com a experiência adquirida como clientes de lojas on-line e, também, pelas perspectivas de crescimento deste tipo de negócio, a escolha deste tema tornou-se mais evidente.

Com este trabalho pretendemos desenvolver uma ontologia simples, mas bem estruturada, e fugir à tentação de detalhar exaustivamente classes, relações e propriedades.

IV. Domínio e Tutorial

O tema que escolhemos para o nosso trabalho, loja online do sector alimentar, tem um enorme potencial de desenvolvimento. Diferentes ontologias poderão ser construídas dependendo do objectivo que se pretende alcançar, como, por exemplo, a “arrumação” dos produtos por secções e subsecções com o objectivo de facilitar a sua pesquisa. No entanto, e porque nos pareceu um desafio mais interessante, decidimos definir como **domínio** da nossa ontologia **a segmentação de clientes de uma loja online**.

Como ferramenta decidimos pelo editor de ontologias **Protégé 4.1 Beta**.

V. Desenvolvimento

1. Âmbito

Para o desenvolvimento desta Ontologia criámos uma loja online à qual chamamos **e-Alimentar**, que é propriedade da empresa **Lojas Online**. Para simplificar, limitamos a oferta da loja a 6 secções de produtos alimentares: Bebidas, Frescos, Lacticínios, Mercearia, Peixaria e Talho. Finalmente enchemos as prateleiras destas secções com alguma variedade produtos.

Povoamos a ontologia com oito clientes para os quais atribuímos um nome, uma idade, um e-mail, um género, um agregado familiar, um número de identificação e uma morada.

Desses oito clientes, sete efectuaram compras na loja **e-Alimentar**. Assim, foram também carregados na ontologia os seus dados de compras – valores totais de compras, artigos comprados e frequência de compras – relativos aos últimos 3 meses.

Como pretendemos obter a **segmentação de clientes**, foi prioritário proceder à definição dos critérios de segmentação que utilizaríamos. Assim optamos por definir a frequência de visita ou compra na loja, e o valor do carrinho de compras como **Critérios Comportamentais**. O agregado familiar com **Critérios Demográfico**. A área de residência como **Critério Geográfico**. E, por fim, o de cabaz de compras com **Critério Psicográfico**.

Outras regras de Segmentação, poderão ser definidas mais tarde, de acordo com a necessidade de cada retalhista ou equipa de Marketing.

2. Construção da Ontologia

Após a definição do âmbito do nosso trabalho e estabelecidos os pressupostos, procedemos à inserção da informação no **Software Protegé 4.1 Beta**.

2.1. Classes

Começamos pela **definição das classes**, e percebemos desde logo que não há soluções universais. Várias hipóteses de diferentes combinações de classes seriam possíveis para atingir o nosso objectivo. Optamos por definir 7 classes, conforme **Figura 1**.

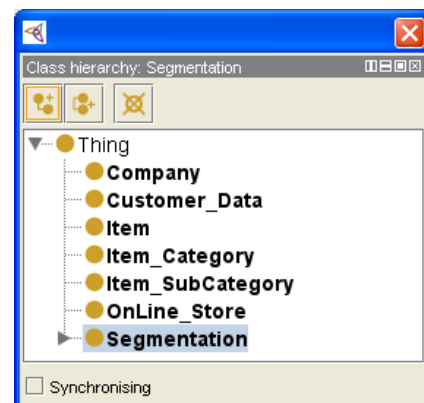


FIGURA 1: DEFINIÇÃO DE CLASSES

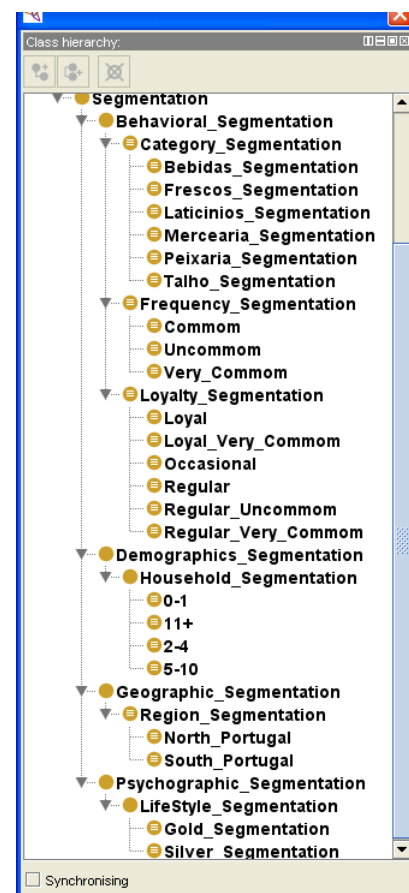
Destas classes apenas a classe *Segmentation* tem subclasses, de vários níveis (ver **Figura 2**).

Esta hierarquia de subclasses de 1º, 2º e 3º nível foi sofrendo diversas alterações ao longo da construção da ontologia. Inicialmente a estrutura de subclasses era mais simples pois pensávamos ser possível definir a segmentação através de expressões matemáticas mais complexas. Mas rapidamente percebemos que o Protégé não nos permitiria enveredar por esse caminho com sucesso. Fiéis ao nosso objectivo primordial, e sem nos desviarmos dos Critérios de Segmentação inicialmente definidos, escolhemos então o caminho da segmentação através de subclasses de vários níveis, da classe *Segmentation*.

Assim a classe *Segmentation* passou a integrar 4 subclasses de primeiro nível: (A)*Behavioral Segmentation*, (B)*Demographics Segmentation*, (C)*Geographic Segmentation* e (D)*Psychographic Segmentation*.

(A) A subclasse *Behavioral Segmentation*, que por sua vez, tem 3 subclasses de segundo nível e respectivas subclasses de terceiro nível, a saber:

FIGURA 2: SUBCLASSES DE SEGMENTATION



→ *Category Segmentation*

Segmentos	Condições
Category Segmentation	buys some Item
Bebidas Segmentation	Category_Segmentation and (buys some (hasAscendant value Bebidas))
Frescos Segmentation	Category_Segmentation and (buys some (hasAscendant value Frescos))
Laticinios Segmentation	Category_Segmentation and (buys some (hasAscendant value Laticinios))
Mercearia Segmentation	Category_Segmentation and (buys some (hasAscendant value Mercearia))
Peixaria Segmentation	Category_Segmentation and (buys some (hasAscendant value Peixaria))
Talho Segmentation	Category_Segmentation and (buys some (hasAscendant value Talho))

→ *Frequency Segmentation*

Segmentos	Condições
Frequency Segmentation	hasFrequency some integer
Commom	(hasFrequency some integer[>= 3]) and (hasFrequency some integer[< 9])
Uncommom	hasFrequency some integer[< 3]
Very Commom	hasFrequency some integer[>= 9]

→ *Loyalty Segmentation*

Segmentos	Condições
Loyalty Segmentation	hasAvgAmount some integer
Loyal	((hasFrequency some integer[>= 3]) and (hasFrequency some integer[< 9])) and (hasAvgAmount some integer[>= 100])
Loyal Very Commom	(hasAvgAmount some integer[>= 50]) and (hasFrequency some integer[>= 9])
Occasional	(hasAvgAmount some integer[< 50]) and (hasFrequency some integer[< 3])
Regular	((hasFrequency some integer[>= 3]) and (hasFrequency some integer[< 9])) and (hasAvgAmount some integer[< 100])
Regular Uncommom	(hasAvgAmount some integer[>= 50]) and (hasFrequency some integer[< 3])
Regular Very Commom	(hasAvgAmount some integer[< 50]) and (hasFrequency some integer[>= 9])

(B) A subclasse *Demographic Segmentation*, que incorpora a segmentação por agregado familiar:

Segmentos	Condições
Household Segmentation	
0-1	(hasHousehold some integer[>= 0]) and (hasHousehold some integer[<= 1])
2-4	(hasHousehold some integer[>= 2]) and (hasHousehold some integer[<= 4])
5-10	(hasHousehold some integer[>= 5]) and (hasHousehold some integer[<= 10])
11+	hasHousehold some integer[>= 11]

(C) A subclasse *Geographic Segmentation*, efectua a segmentação por região:

Segmentos	Condições
Region Segmentation	(hasCountry some string) and (hasRegion some string)
North Portugal	Region_Segmentation and (hasCountry value "Portugal") and (hasRegion value "Norte")
South Portugal	Region_Segmentation and (hasCountry value "Portugal") and (hasRegion value "Sul")

(D) A subclasse *Psychographic Segmentation*, que faz a segmentação por estilo de vida:

Segmentos	Condições
LifeStyle Segmentation	
Gold Segmentation	buys some (hasItemRating value "Gourmet")
Silver Segmentation	buys some (hasItemRating value "Normal")

2.2 Indivíduos

Seguiu-se a criação dos **Indivíduos**.

- Indivíduos da **Classe *Customer Data***, são os clientes da loja, detalhados na **Figura 3**.
- Indivíduos da **Classe *Item Category***, que são as secções da loja: Bebidas, Frescos, Lacticínios, Mercearia, Peixaria e Talho.
- Indivíduos da **Classe *Item SubCategory***, que são as subcategorias de produtos (ver **Anexo C**).
- Indivíduos da **Classe *Item***, que são artigos pertencentes à gama da loja (ver **Anexo C**).



FIGURA 3: INDIVÍDUOS DA
CLASSE CUSTOMER DATA

2.3. Propriedades

Enunciadas as **Classes** e os **Indivíduos**, passamos à introdução das **Propriedades** que nos permitirão definir a relação entre **Classes** ou **Indivíduos**:

- **Object Properties**, relacionam um indivíduo com outro indivíduo.
- **Datatype Properties**, ligam um indivíduo a um valor específico.

Descrevemos as seguintes **Object Properties**:

- *buys*
- *hasAscendent*, transitiva e inversa de *isAscendentOf*
- *isAcendentOf*, transitiva e inversa de *hasAscendent*
- *owns*, inversa de *isOwned*
- *isOwned*, inversa de *owns*
- *isSold*, inversa de *sells*
- *sells*, inversa de *isSold*

Das **Datatype Properties** criadas, e que se podem ver **Figura 4**, as propriedades

- *hasAge*
- *asGender*
- *hasHousehol*
- *hasID*,

são propriedades funcionais.

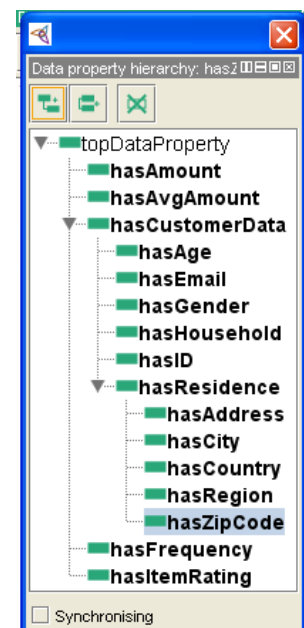


FIGURA 4: DATATYPE PROPERTIES

3. Aspectos Relevantes

O desenvolvimento desta ontologia realizou-se, inicialmente, na versão 4.1.0 (Build 209) do Protégé. Versão esta que, no decorrer do trabalho, se foi demonstrando muito pouco estável, evidenciando comportamentos desacertados da aplicação.

Dada a in experiência da equipa com a aplicação, esta instabilidade conduziu a um trabalho adicional. Primeiramente, no sentido de se perceber se o desenvolvimento estava a decorrer de uma forma incorrecta, e depois, na tentativa de se encontrarem soluções que ultrapassassem essas questões.

Dos casos detectados, evidenciamos os 3 mais críticos:

- Impossibilidade do uso de caracteres exclusivos do alfabeto português. Aqui, e na grande maioria das vezes, os caracteres portugueses eram substituídos por outros caracteres, criando a ideia de desformatação. Assim, e na tentativa de se resolver esta questão, fizeram-se diferentes tentativas de correcção, que passaram pelo seguinte:
 - Reinstalações do próprio Protégé.
 - Instalação do Protégé em versões de Sistemas Operativos (Windows) diferentes.
 - Instalações de diferentes versões do Java VM.
 - Tentativa de alteração do *character set* quer do Sistema Operativo, quer do próprio Protégé.

Como não se obteve sucesso, optámos então por usar classes, subclasses e propriedades em inglês. Os indivíduos já foram populados em português, mas aqui decidimos excluir caracteres portugueses.

- Falta de consistência aquando do apuramento da informação através do Reasoner. Em algumas situações, detectou-se que o Reasoner só era capaz de produzir resultados, se a ontologia fosse fechada e novamente aberta. Nestes casos, a única solução encontrada não foi além disso mesmo: quando havia a certeza que o desenvolvimento estava correcto, fechávamos e abríamos de novo a aplicação.
- Inconsistências de tipos de dados. Em expressões do tipo “hasAmount some integer[> 230]”, o Reasoner disparava um erro de inconsistência de dados. Depois de avaliados os tipos de dados, quer dos indivíduos quer das propriedades, concluímos então que se poderia tratar de novo bug da aplicação. Para o contornar, a cada edição do filtro da classe, passámos a alterar o ficheiro XML directamente.

Já no final do desenvolvimento da ontologia, tomámos conhecimento do lançamento de um patch por parte da Protégé, que resolvia, entre outros, estes 3 casos. Instalámos assim

esta nova versão - 4.1.0 (Build 213) – e foi então com esta que concluímos o nosso desenvolvimento.

Salientamos também o facto de que a aplicação não se demonstrou muito prática para o carregamento de dados massivos. Para contornar essa dificuldade foram criados geradores de código XML, com base em ficheiros Excel e Access, que depois se encaitava no ficheiro OWL.

VI. Conclusões

Durante o desenvolvimento da ontologia no Protégé, fomos nos deparando com algumas dificuldades, que tentamos descrever ao longo deste relatório. Dificuldades essas que passaram não só pela falta de experiência no manuseamento da aplicação, numa fase inicial, como pela instabilidade que a versão 4.1.0 (Build 209) veio a apresentar.

Com o decorrer do tempo, a aquisição de conhecimento evoluiu. E assim, conseguimos progredir não só no desenvolvimento da própria ontologia, como permitiu que superássemos algumas das instabilidades detectadas.

Já no final do desenvolvimento, foi então instalada a nova versão, versão 4.1.0 (Build 213), que se apresentou mais estável. No entanto, não foram aqui feitos testes exaustivos, dada a fase avançada em que nos encontrávamos.

Em algumas situações, teria sido importante que a própria ontologia realizasse operações matemáticas (p.e.: médias, produtos, divisões), no entanto isso não se tornou possível por limitações da própria aplicação. Assim, os dados que popularam a ontologia tiveram que ser previamente tratados, e só depois carregados.

Sobre o desenvolvimento propriamente dito, concluímos que se cumpriram todos os requisitos estabelecidos na fase inicial. Segundo um conjunto de informação relativa a Clientes, a ontologia apurou, e com sucesso, a lista de clientes nos diferentes segmentos, cumprindo as condições por nós estipuladas. Mais ainda, a ontologia foi também capaz de relacionar duas segmentações (*Frequency Segmentation* e *Loyalty Segmentation*) que à primeira vista se mostravam independentes.

A experiência adquirida na construção desta ontologia, coroada pelo resultado final do trabalho, permitiu-nos perceber a verdadeira potencialidade deste tipo de ferramenta, e deixou-nos com a vontade de tentar ir mais além.

VII. Bibliografia

SAIAS, José Miguel. 2003. Uma metodologia para a construção automática de Ontologias e sua aplicação a sistemas de Recuperação de Informação. Mestrado em Engenharia Informática Universidade de Évora. Disponível em: <
<http://host.di.uevora.pt/~jsaias/data/dissertacao-mei-js.pdf>

Natalya F. Noy, Deborah L. McGuinness. Ontology Development 101: A Guide to Creating Your First Ontology. Stanford University

www.linguateca.pt(consulta em Outubro 2010)

www.tecnet.pt (consulta em Outubro 2010)

www.daml.org/ontologies (consulta em Outubro 2010)

VIII. Anexos

- A. Representação Gráfica da Ontologia (grafos)
- B. Ficheiro OWL da Ontologia (RC_Segmentacao_Clientes)
- C. Dados que popularam a Ontologia (Dados para popular a ontologia)
- D. Códigos XML gerados (exemplos de codigo)



Grafos.docx



RC_Segmentacao_Clientes.owl



Dados para popular a ontologia.xlsx



exemplos de codigo.zip