

# 1 Introdução

O termo comércio eletrônico refere-se a todos os tipos de transação on-line na *Internet* visando à efetivação de compra e venda de mercadorias ou serviços. Podemos ser mais específicos categorizando essas transações conforme o seu agente (comprador, vendedor). As principais categorias dentro dessa perspectiva são:

- *Consumer to consumer* (C2C) – Esse é o tipo de transação em que um indivíduo negocia bens com outros indivíduos. A plataforma típica para a realização desse tipo de transação é o leilão virtual. Alguns exemplos são: ([www.ebay.com](http://www.ebay.com)) ; ([www.ibazar.com](http://www.ibazar.com)).
- *Consumer to business* (C2B) – Esse é o tipo de transação em que os indivíduos (compradores) definem o preço e todas as suas preferências relacionadas à intenção de consumo e, após a publicação destas informações, vendedores competem e jogam com suas margens de lucro para conquistar a venda. Alguns exemplos são: ([www.priceline.com](http://www.priceline.com)); ([www.valeu.com.br](http://www.valeu.com.br)).
- *Business to business* (B2B) – Esse é o tipo de transação que é feito entre duas empresas. Neste segmento não existe a figura do consumidor pessoa física e, geralmente, a dinâmica de transações caracteriza-se pelo baixo volume de transações de alto valor.
- *Business to consumer* (B2C) – Esse é o tipo de transação em que uma companhia ou organização vende seus produtos ou serviços para as pessoas que navegam pela *Internet*. Diferentemente do segmento C2B, a definição de preço de venda é feita pela companhia ou organização. Inclui-se nesse segmento uma vasta gama de *sites*, tais como varejistas *on-line* e bancos. Alguns exemplos são: ([www.amazon.com](http://www.amazon.com)); ([www.paodeacucar.com.br](http://www.paodeacucar.com.br)); ([www.bb.com.br](http://www.bb.com.br)).

Como este trabalho se enquadra no segmento B2C, daqui para frente falaremos mais especificamente das vantagens e desvantagens específicas desse segmento.

Diferentemente dos demais segmentos do comércio eletrônico, no segmento B2C não há nenhuma quebra de paradigma em relação à forma de compra e venda de produtos, ou seja, nesse segmento o que se tem basicamente são versões digitais de catálogos tradicionais. Os benefícios, tanto para o vendedor como para o consumidor, advêm de o novo meio utilizado permitir a negociação com um alto grau de interatividade.

Entre os benefícios para o vendedor, podemos ressaltar: poder expor seus produtos a qualquer usuário da *Internet* independente de onde ele esteja; capacidade de modificar a forma de exposição dos produtos dentro do *site* conforme as preferências pessoais dos consumidores a um baixo custo (personalização); fazer marketing de forma mais personalizada; etc.

Para o consumidor, a possibilidade de ter acesso a um grande volume de informações sobre vários fornecedores de um produto permite que se tenha uma capacidade mais apurada de decisão sobre uma compra.

Atualmente, apesar dessas vantagens, a principal dificuldade dentro do segmento está em aumentar a utilidade da WWW para o comércio eletrônico através da melhoria de sua interface com usuários.

Uma das formas de aumentar essa usabilidade é através do uso de agentes inteligentes de busca de informação, que seriam utilizados como agentes de compra, visando auxiliar os compradores a encontrar produtos de seu interesse. Isso porém esbarra em uma dificuldade inerente à própria WWW: A mistura da linguagem natural, imagens e informação de *layout* de HTML é uma das maiores barreiras para a automatização do comércio eletrônico, pois a semântica da informação é somente compreensível por seres humanos (Fensel, 2001).

Isto limita a capacidade dos agentes de compra. Apesar de existirem heurísticas para extrair alguma informação da linguagem natural (ex.: *Extreme semantics* (Invention Machine Corp., 2000)), essa não pode ser totalmente compreendida, acarretando um grande esforço para desenvolver e manter esses agentes.

A linguagem XML (Martin et al., 2001; Holzner, 2001) vem para melhorar essa situação. Enquanto HTML é uma linguagem de *layout* de informação para

apresentação de documentos textuais visando à leitura por seres humanos, a XML é uma linguagem para definição de estrutura da informação, voltada principalmente para facilitar a troca de dados entre aplicações.

A XML melhora consideravelmente a sintaxe padrão para troca de informação, porém é necessário definir-se uma semântica precisa que possa ser entendida não só por um humano, mas principalmente por um agente de software. Isto pode ser alcançado com o uso de ontologias (Fensel, 2001).

Além do problema de extrair significado da informação, os agentes de compra se deparam com outro grande problema: a falta de padronização dos diversos *sites* para descrever um mesmo tipo de informação (problema de interoperabilidade semântica). Dois *sites* podem vender o mesmo tipo de produto tendo descrições diferentes para o mesmo. Somada a isto, a forma como um *site* descreve o produto que vende não está acessível em uma linguagem formal, sendo esta subentendida na descrição do produto feita em HTML.

Exemplificando o problema, enquanto um livro em um determinado *site* é descrito pelo seu título, autor e preço, em outro *site* o mesmo livro pode ser descrito pelo nome, autor, gênero, número de páginas e preço. Como não há uma descrição formal de como o livro está descrito em cada *site* (metadescrição), a tarefa de integração das informações dos diversos *sites* pelo agente de compra não pode ser feita de forma automática, exigindo um grande esforço para a implementação e manutenção do agente de compra, pois esse deve considerar as peculiaridades de cada *site*. Dessa forma podemos ver que essa metadescrição é de suma importância para permitir que agentes possam manipular conteúdo de diversas fontes, sem a preocupação com as peculiaridades das mesmas. **Essa metadescrição pode ser organizada como uma ontologia.**

Uma ontologia pode ser entendida como a especificação de uma conceitualização. Ela é a descrição formal de conceitos e relacionamentos entre esses conceitos. O objetivo das ontologias é prover uma semântica, processável por máquina, de fontes de informação. As principais linguagens de representação de ontologias são SHOE (Luke & Heflin, 2000), RDF Schema (Dodds et al., 2001; Brickley & Guha, 2000) e DAML+OIL (Connolly et al., 2001a; Connolly et al., 2001b) sendo as duas últimas baseadas em XML.

Uma vez que temos uma ontologia, podemos associá-la a uma página que siga essa ontologia. Essa página deve ser descrita seguindo as definições dessa

ontologia. A principal linguagem de descrição de recursos é a RDF (Swick & Lassila, 1999), linguagem essa definida pelo consórcio W3C.

O uso de ontologia para descrever informações permite que agentes façam certas inferências a partir das mesmas, pois eles podem conhecer a descrição formal da ontologia.

Esse trabalho pretende utilizar ontologias para, além de padronizar, permitir que agentes de compra façam inferências sobre recursos na WWW. Especificamente os recursos utilizados serão instâncias de lojas geradas pelo *framework* 2BuyNet, um gerador de lojas virtuais. Dessa forma, espera-se fazer com que o processo de compra de um produto por um consumidor seja facilitado com o auxílio de um agente de busca para compra.

Esse trabalho faz algumas contribuições interessantes. Primeiro é feito uma introdução do conceito de ontologia, mostrando-se também algumas das metodologias existentes para o seu desenvolvimento. Depois é feito um estudo comparativo das principais tecnologias que estão por trás da utilização de ontologias em Ciência da Computação (XML, RDF, RDF Schema, DAML+OIL, SHOE). Também é implementado um estudo de caso que se encontra dentro do domínio de *Business to Consumer*, utilizando essas tecnologias. Através do estudo de caso é possível tirar algumas conclusões interessantes a respeito da utilização de ontologias dentro desse domínio.

Os capítulos restantes da dissertação estão assim organizados:

- O capítulo 2 apresenta uma visão geral do segmento de *Business to Consumer*.
- O capítulo 3 fala sobre os vários tipos de ferramentas de busca existentes e suas características.
- O capítulo 4 trata de ontologia.
- O capítulo 5 aborda as principais tecnologias para o desenvolvimento de ontologias na WWW.
- O capítulo 6 apresenta o estudo de caso desenvolvido que utiliza as tecnologias abordadas no capítulo 5.
- O capítulo 7 discute as conclusões e possíveis trabalhos futuros.