

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E INFORMAÇÃO

GILBERTO BALBELA CONSONI

RECUPERAÇÃO DE INFORMAÇÃO EM SISTEMAS DE RECOMENDAÇÃO:

Análise da interação mediada por computador e dos efeitos da filtragem colaborativa na seleção de itens no website da Amazon.com

Porto Alegre
2014

GILBERTO BALBELA CONSONI

RECUPERAÇÃO DE INFORMAÇÃO EM SISTEMAS DE RECOMENDAÇÃO:

Análise da interação mediada por computador e dos efeitos da filtragem colaborativa na seleção de itens no website da Amazon.com

Tese apresentada ao Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul, como requisito parcial à obtenção do título de Doutor em Comunicação e Informação.

Orientador: Prof. Dr. Alex Fernando Teixeira Primo

Porto Alegre
2014

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Carlos Alexandre Netto

Vice-reitor: Prof. Dr. Rui Vicente Oppermann

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretora: Ana Maria Mielniczuk de Moura

Vice-diretor: André Iribure Rodrigues

PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E INFORMAÇÃO

Coordenadora: Profa. Dra. Miriam de Souza Rossini

Coordenadora substituta: Profa. Dra. Nísia Martins do Rosário

Consoni, Gilberto Balbela

Recuperação de informação em sistemas de recomendação: Análise da interação mediada por computador e dos efeitos da filtragem colaborativa na seleção de itens no website da Amazon.com / Gilberto Balbela Consoni. – 2014.

196 p.

Orientador: Alex Fernando Teixeira Primo.

Tese (Doutorado) – Universidade Federal do Rio Grande do Sul, Faculdade de Biblioteconomia e Comunicação, Programa de Pós-Graduação em Comunicação e Informação, Porto Alegre, BR-RS, 2014.

1. Tecnologias da informação e comunicação. 2. Sistemas de recomendação. 3. Recuperação de informação. 4. Interação mediada por computador. 5. Amazon. I. Primo, Alex Fernando Teixeira, orient. II. Título.

GILBERTO BALBELA CONSONI

RECUPERAÇÃO DE INFORMAÇÃO EM SISTEMAS DE RECOMENDAÇÃO:

Análise da interação mediada por computador e dos efeitos da filtragem colaborativa na seleção de itens no website da Amazon.com

Esta tese foi analisada e julgada adequada para a obtenção do título de Doutor em Comunicação e Informação e aprovada em sua forma final pelo Orientador e pela Banca Examinadora designada pelo Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul.

Prof. Dr. Alex Fernando Teixeira Primo
Orientador

Profa. Dra. Miriam de Souza Rossini
Coordenadora

Aprovado em: ____ / ____ / ____.

BANCA EXAMINADORA

Prof. Dr. Eduardo Campos Pellanda - PUCRS _____.

Profa. Dra. Samile Andrea de Souza Vanz - UFRGS _____.

Profa. Dra. Sandra Portella Montardo - Feevale _____.

Profa. Dra. Sônia Elisa Caregnato - UFRGS _____.

*A minha amada esposa,
Leticia Angheben Consoni.*

AGRADECIMENTOS

O número de pessoas a agradecer no final da tese após esses quatro anos de aprimoramento no doutorado provavelmente excede o espaço apropriado para esta seção. É praticamente impossível recordar no momento da finalização do trabalho todas as pessoas que merecem serem lembradas. Certamente um sistema de recomendação para recuperar as pessoas da memória seria útil neste momento, mas felizmente não é de forma objetiva que se recorda dos entes queridos. Ainda assim, duas pessoas precisam ser citadas, visto à importância delas para a realização dessa pesquisa.

Agradeço ao meu orientador pela oportunidade de realizar meu doutoramento em uma instituição de tanto prestígio quanto a UFRGS e principalmente por compartilhar seu conhecimento comigo, essenciais para a execução desta tese. Agradeço também pela importante contribuição em toda a minha carreira docente e por dividir turmas de alunos na UFRGS para que desenvolvesse a atividade de professor. Agradeço os conhecimentos passados na véspera do concurso para professor da UFPel, quando deixou de brincar com sua filha na chuva para passar instruções primordiais para minha aprovação. Agradeço por esses quase seis anos de confiança que iniciou em 2008 com o mestrado e que seguiu no doutorado.

Agradeço à Leticia Angheben Consoni, minha eterna namorada que aceitou se casar comigo durante o doutorado. A participação dela nesta tese está além do compartilhamento de sua linda vida comigo e dos sacrifícios necessários no início de nossa vida de casados para a realização do meu doutorado. Sua formação na área da biblioteconomia possibilitou sua ajuda com indicações de bibliografias e também ser uma ouvinte profissional das minhas ideias a respeito desta tese.

Gostaria de agradecer também aos profissionais da UFRGS, especialmente os professores e funcionários do PPGCOM e da FABICO. Agradeço a essas pessoas que me modificaram ao minimizar as verdades que eu tinha do mundo e maximizarem as minhas incertezas, mas o genial dessa mudança é que agora compreendo melhor a vida.

Agradeço a UFRGS pela qualidade de seu ensino e a CAPES pelos anos de bolsa ao longo do doutorado, a qual permitiu dedicação no aprimoramento acadêmico.

Muito obrigado a todos vocês!

Um registro, se é para ser útil para a ciência, precisa ser continuamente ampliado, deve ser armazenado, e acima de tudo deve ser consultado.¹

Vannevar Bush (1945, p. 102)

Um sistema de recuperação de informação tenderá a não ser usado sempre que for mais penoso e incômodo para o cliente ter a informação por ele do que não tê-la.²

Calvin Mooers (1996, p. 22)

¹ Tradução nossa para “A record, if it is to be useful to science, must be continuously extended, it must be stored, and above all it must be consulted” (BUSH, 1945, p. 102).

² Tradução nossa para “An information retrieval system will tend not to be used whenever it is more painful and troublesome for a customer to have information than for him not to have it” (MOOERS, 1996, p. 22).

RESUMO

Como os interagentes selecionam conteúdo sob a influência dos sistemas de recomendação digital é o problema de pesquisa apresentado nesta tese. A abundância de dados nos repositórios digitais exige sistemas de recuperação de informação eficazes para auxiliar o usuário na gestão e na seleção de itens de informação. Desta forma, o objetivo geral deste trabalho pretende investigar o comportamento dos interagentes na seleção de itens frente ao sistema de recomendação digital do website da loja virtual Amazon. O sistema de recomendação da Amazon foi investigado com a intenção de se compreender como o usuário utiliza um sistema automatizado de listas de referências em forma de recomendação de conteúdo. O funcionamento dos sistemas de recomendação é fundamentado com a proposta de conhecer suas características e funcionalidades. Como o problema de pesquisa tem em sua temática a recomendação de itens de informação, tornou-se necessário compreender como os usuários interagem com os sistemas para perceber como as recomendações são feitas. O aporte teórico desta pesquisa aproxima os estudos dos campos da Informação e da Comunicação. As técnicas de pesquisas aplicadas envolvem métodos de pesquisa qualitativa. Ao distinguir as recomendações a partir das interações reativas e mútuas dos usuários, propõe-se nesta tese a Matriz de Recomendações de Itens de Informação constituída pelos seguintes quadrantes: Recomendações Objetivas Digitais; Recomendações Subjetivas Digitais; Recomendações Objetivas Analógicas e Recomendações Subjetivas Analógicas. Para analisar o comportamento dos interagentes no uso dessas recomendações, a estratégia metodológica aplicou entrevista em profundidade e observação direta. Os resultados desta pesquisa consideram que o internauta recorre a mais de um tipo de recomendação quando a seleção envolve conteúdo significativo, enquanto segue passivamente sistemas de recomendações automatizados quando o custo pessoal diretamente aplicado é baixo ou inexistente.

Palavras-chave: Tecnologias da informação e comunicação. Sistemas de recomendação. Recuperação de informação. Interação mediada por computador. Amazon.

ABSTRACT

As the interacting select content under the influence of digital recommender systems is the research problem presented in this thesis. The abundance of data in digital repositories recovery requires effective information systems to assist the user in the management and selection of information items. Thus, the objective of this study was to investigate the behavior of the interacting in the selection of digital items across the recommendation of the Amazon's bookshop website. The Amazon's recommendation system was investigated with the aim of understanding how the user uses an automated reference lists in the form of content recommendation. The performance of recommender systems is founded with the purpose of knowing their characteristics and functionalities. As the research problem is in your subject to the recommendation of information items, it became necessary to understand how users interact with this system to understand how the recommendations are made. The theoretical contribution of this research approaches the fields of Information and Communication. The technique applied involves qualitative research methods. By distinguishing the recommendations from reactive and mutual interactions of users, is propos in this research the Model of Recommendation Information Items consist of the following quadrants: Digital Objective Recommendations; Digital Subjective Recommendations; Analog Subjective Recommendations and Analog Objective Recommendations. To analyze the behavior of interactors in the use of these recommendations, the methodological strategy applied in-depth interviews and direct observation. The results of this research consider that the Internet uses more than one type of recommendation when the selection involves significant content, while passively follows recommendations of automated systems when applied directly to the personal cost is low or nonexistent.

Keywords: Information and communication technologies. Information retrieval. Recommendation systems. Computer-mediated interaction. Amazon.

LISTA DE ILUSTRAÇÕES

Figura 1 – O que outros clientes estão vendo agora mesmo.	19
Figura 2 – Mais itens para considerar.	20
Figura 3 – Relacionado aos itens vistos por você.	20
Figura 4 – Inspirado pelo seu histórico de navegação.	20
Figura 5 – Itens adicionais a explorar.	21
Figura 6 – Recomendado baseado no seu histórico de navegação.	21
Figura 7 – Seu histórico de navegação.	22
Figura 8 – Gerencie o seu histórico de navegação.	22
Figura 9 – <i>Star Rating</i> – sistema de classificação por estrelas.	23
Figura 10 – Página de resenhas de clientes.	24
Figura 11 – Your Amazon.com: Página inicial do visitante que efetuou <i>login</i>	25
Figura 12 – Recomendação para você na loja Kindle.	25
Figura 13 – Recomendações inspiradas na lista de desejos.	26
Figura 14 – Itens frequentemente comprados juntos.	26
Figura 15 – Votos em resenhas e comentários.	27
Figura 16 – Comentários das resenhas.	27
Figura 17 – Avaliação simplificada por estrelas.	28
Figura 18 – Página de criação de resenhas.	28
Figura 19 – Reportagem A tortura da seca.	39
Figura 20 – Four aspects of information.	42
Figura 21 – Fluxo de documentos através do Tapestry.	66
Figura 22 – Arquitetura do Sistema de Recuperação de Imagens para o iPad-2.	69
Figura 23 – Resultados de pesquisa para a palavra nutrição na Amazon Brasil.	79
Figura 24 – Tela de incentivo ao preenchimento no eHarmony.	81
Figura 25 – Tela para preenchimento de preferências do usuário no Netflix.	83
Figura 26 – Exemplo de recomendações no Netflix.	84
Figura 27 – Sala de aula preparada para entrevistas em profundidade.	121
Figura 28 – Cartão com escala de classificação por estrelas.	122
Figura 29 – Interface do software de gravação da observação direta.	124
Figura 30 – Tabulação das recomendações na Amazon.com.	125
Figura 31 – Tabulação e cruzamento das respostas das entrevistas em profundidade.	126
Figura 32 – Tabulação e cruzamento dos dados coletados nas observações diretas.	126
Figura 33 – Tipos de Recomendações de Itens de Informação.	127
Figura 34 – Improve Your Recommendantios.	130
Figura 35 – Slider na Amazon.	131
Figura 36 – Seção na Amazon.	131
Figura 37 – Relacionado aos itens que você viu.	133
Figura 38 – Consumidores que compraram itens em seu histórico de navegação.	134
Figura 39 – Recomendações baseadas em compras dos usuários.	134
Figura 40 – Computers & Technology Best Sellers.	136
Figura 41 – The New York Times® Best Sellers.	136
Figura 42 – Internet & Web Culture Best Sellers.	137
Figura 43 – Campo de busca.	138
Figura 44 – Filtro do Campo de Busca.	139
Figura 45 – Sugestões de busca.	140
Figura 46 – Star Rating.	142
Figura 47 – Classificação pelas avaliações dos clientes.	143

Figura 48 – Votos nas resenhas.	143
Figura 49 – Descrição da obra.	146
Figura 50 – Resenhas dos editores da Amazon e da PW.....	148
Figura 51 – 20 principais escolhas dos editores para os melhores livros de 2013.	149
Figura 52 – Página dos editores da Amazon.	150
Figura 53 – Comentário em postagem da editora Sara Nelson.	150
Figura 54 – Seleções de celebridades: o que os autores estão lendo.	151
Figura 55 – Melhores do ano para o autor George Saunders.	152
Figura 56 – Comentário na postagem 2013 Best Books of the Year: Celebrity Picks.	152
Figura 57 – Seções de livros premiados.	154
Figura 58 – Grau de confiança em formas de publicidade online.	156
Figura 59 – Link para resenhas de consumidores.....	157
Figura 60 – Resenhas de consumidores diretamente na página do item.	157
Figura 61 – Página das resenhas no perfil do consumidor.	158
Figura 62 – Comentário oculto da lista de comentários.	160
Figura 63 – Conteúdo do comentário oculto da lista de comentários.	160
Figura 64 – Página das resenhas no perfil do consumidor.	161

LISTA DE TABELAS

Tabela 1 – Quadro teórico.	35
Tabela 2 – Sistemas de recomendação, algoritmos e técnicas.	77
Tabela 3 – Taxonomia dos Agentes de Recomendação.	80
Tabela 4 – Matriz de Recomendações de Itens de Informação.	118
Tabela 5 – Roteiro da entrevista em profundidade semiaberta.	122
Tabela 6 – Roteiro das atividades da observação direta.	124
Tabela 7 – Matriz e Tipos de Recomendações de Itens de Informação.	181

SUMÁRIO

1	INTRODUÇÃO	14
1.1	JUSTIFICATIVA	15
1.2	OBJETO	16
1.2.1	Amazon.com	16
<i>1.2.1.1</i>	<i>Recomendações e avaliações para visitantes que não efetuaram login</i>	<i>19</i>
<i>1.2.1.2</i>	<i>Recomendações e avaliações para visitantes que efetuaram login</i>	<i>24</i>
1.3	PROBLEMA DE PESQUISA.....	29
1.4	OBJETIVOS.....	34
1.4.1	Objetivo Geral	34
1.4.2	Objetivos Específicos	34
1.5	QUADRO TEÓRICO.....	35
2	SISTEMAS DE INFORMAÇÃO	37
2.1	A INFORMAÇÃO-COMO-COISA.....	37
2.2	COMO PENSAR A RECUPERAÇÃO DE INFORMAÇÃO	45
2.3	A DEFINIÇÃO DE RECUPERAÇÃO DE INFORMAÇÃO.....	53
2.4	SISTEMAS DE RECOMENDAÇÃO.....	62
2.4.1	Da Filtragem Colaborativa aos Sistemas Implícitos de Recomendação	63
2.4.2	Classificação e Anatomia dos Sistemas de Recomendação	73
3	INTERAÇÃO MEDIADA POR COMPUTADOR	86
3.1	RECIPROCIDADE ENTRE OS INTERAGENTES	87
3.2	RELAÇÃO E INTERAÇÃO.....	89
3.3	INTERAÇÃO MEDIADA	97
3.4	CONSUMO MEDIADO POR COMPUTADOR	104
4	QUADRO METODOLÓGICO	116
4.1	CLASSIFICAÇÃO DAS RECOMENDAÇÕES	117
4.2	ENTREVISTA EM PROFUNDIDADE E OBSERVAÇÃO DIRETA	119
4.2.1	Entrevista em profundidade	120
4.2.2	Observação direta	123
4.3	ANÁLISE DOS DADOS	125
5	RECOMENDAÇÃO DE ITENS DE INFORMAÇÃO NA AMAZON	127
5.1	CLASSIFICAÇÃO DAS RECOMENDAÇÕES DE ITENS DE INFORMAÇÃO	128
5.1.1	Recomendações do tipo Objetiva Digital na Amazon	129
<i>5.1.1.1</i>	<i>Interface Gráfica Dinâmica</i>	<i>129</i>
<i>5.1.1.2</i>	<i>Recomendações Personalizadas</i>	<i>132</i>
<i>5.1.1.3</i>	<i>Recomendações por Popularidade</i>	<i>135</i>
<i>5.1.1.4</i>	<i>Recomendações Sistêmicas</i>	<i>138</i>
<i>5.1.1.5</i>	<i>Recomendações Sociais Reativas</i>	<i>142</i>
5.1.2	Recomendações do tipo Subjetiva Digital na Amazon	143
<i>5.1.2.1</i>	<i>Recomendações Editoriais</i>	<i>144</i>
<i>5.1.2.2</i>	<i>Recomendações de Obras Premiadas</i>	<i>153</i>
<i>5.1.2.3</i>	<i>Recomendações Sociais Mútuas</i>	<i>155</i>
5.2	EXPERIÊNCIA DE SELEÇÃO DE ITENS NA AMAZON.COM	162
5.2.1	Busca e Recuperação de Itens de Informação na Amazon	163
5.2.2	Aspectos Externos da Seleção de Itens de Informação na Amazon	167
5.2.3	Motivações e Uso do Sistema de Recomendações da Amazon	173
5.2.4	Avaliação e Uso das Recomendações Editoriais e Sociais Mútuas na Amazon	177
6	CONSIDERAÇÕES FINAIS	182

REFERÊNCIAS.....	189
ANEXO A – MODELO DO TERMO DE CONSENTIMENTO	194
GLOSSÁRIO	195

1 INTRODUÇÃO

A presente tese investiga o comportamento dos internautas ao selecionarem itens em websites com sistemas de recomendação digital. Ao estudar os sistemas de recuperação de informação na Ciência da Informação, nos quais os sistemas de recomendação digital como abordados nesta tese estão inseridos, pesquisa-se as formas de indicar itens para seleção a partir da interação mediada por computador no viés da Comunicação Social.

Nesta tese, entende-se por recomendação a ação de indicar algo no sentido de aconselhamento, ou seja, trata-se do ato de indicar informações para apresentar, advertir ou aconselhar algo frente à qualidade do que é recomendado. Nesse sentido, a recomendação está no ato de apresentar uma opção qualificada para efeito de seleção. A recomendação diferencia-se da propaganda, pois esta possui o cunho comercial frente à qualidade da indicação a quem se sugere algo, logo, na recomendação a qualidade deveria estar à frente do interesse de venda. Embora se observe propagandas travestidas de recomendações, as quais também serão analisadas nesta tese, para efeito de análise será considerado apenas aqueles itens de informação que não estão visivelmente marcados como propaganda. Portanto, um editorial que é pago para promover um produto será analisado no website da Amazon, mas os banners digitais publicitários serão deliberadamente excluídos dos dados da análise.

Antes de abordar e apresentar o problema e os objetivos desta pesquisa nos subcapítulos 1.3 e 1.4, e fundamentar a bibliografia publicada nos capítulos 2 e 3, apresenta-se no subcapítulo 1.2 o objeto empírico a ser pesquisado que é o website da Amazon.com³. Opta-se por esta estruturação do texto para que as abordagens relacionadas aos sistemas de recomendações contemporâneos, feitas ao longo da tese, aproximem-se desde o início do objeto de pesquisa. Como websites com sistemas de recomendação digital estão intrínsecos à vida cotidiana dos leitores deste texto, esta descrição técnica inicial do objeto não deve gerar dúvidas quanto ao seu funcionamento, porém deve despertar inquietações que tocam o tema de pesquisa pelas peculiaridades que serão evidenciadas no próprio uso do website da Amazon. Conhecida a performance técnica permitida pelo objeto empírico pesquisado, passa-se à fundamentação do objeto teórico que são os sistemas de recuperação de informação.

³ Disponível em <<http://amazon.com>>

1.1 JUSTIFICATIVA

O montante de informações que precisa ser gerenciada pelos usuários de sistemas de informação exige, necessariamente, o uso de algum tipo de sistema de recuperação de informação, para que se obtenham dados relevantes quando se pesquisa por determinado item. Os sistemas de recuperação de informação não são mais exclusividade daqueles habituados a visitarem bibliotecas e a utilizarem bases de dados científicas. Praticamente todos os grandes websites com bases de dados centralizadas operam hoje sob influência de algum sistema de recuperação de informação para ajudar seus visitantes a encontrarem o que desejam. Dentre estes sistemas, os sistemas de recomendação de conteúdo são aqueles que os desenvolvedores mais buscam implementar, para aproveitar os dados de navegação dos usuários, para indicar conteúdo a eles próprios. O acentuado e cada vez mais frequente uso desses sistemas de recomendação pelos internautas é a justificativa mais evidente para que estes sejam investigados, porém há outras justificativas que merecem ser destacadas para sustentar a necessidade de melhor investigação e análise desses sistemas.

O uso frequente dos sistemas de recomendação pode influenciar a seleção de itens pelos internautas. Frente à quantidade de informação que precisa gerenciar, o internauta pode utilizar um sistema de recomendação para ajudá-lo a filtrar os dados relevantes frente a esse montante de dados. Este aspecto justifica investigar o funcionamento destas recomendações e o comportamento do internauta sob influência desses sistemas, visto que pode ficar sujeito à máquina decidir quais itens são relevantes a serem acessados. O que também justifica este estudo está não apenas nos dados que o sistema decide mostrar, mas àqueles que decide esconder do usuário ao calcular de forma matemática, por meio de algoritmos, os dados que não são relevantes para o usuário do sistema.

A justificativa acadêmica para realização desta pesquisa é que apesar de os sistemas de recomendação serem tratados como objeto de pesquisa nas áreas da Ciência da Informação, Matemática, Informática e Computação, defende-se que merecem também um melhor olhar do viés da Comunicação Social a partir do estudo da interação mediada por computador. Estudar os sistemas de recomendação pelo viés da Comunicação e da Informação permite analisar as formas de indicação de itens de informação, tanto a partir de um tratamento objetivo de dados reais, quanto subjetivo das interações entre as pessoas conectadas ao sistema. Desse modo, a presente tese justifica-se por associar essas duas áreas a fim de aproveitar o que cada uma delas potencializa em seus estudos: a investigação da informação e das interações humanas.

Do ponto de vista da aplicabilidade desta pesquisa, os resultados alcançados podem ser utilizados para o desenvolvimento de melhores sistemas de recomendação e não apenas para o avanço dos algoritmos, mas também das interações entre seus usuários. Esta preocupação com a aplicação justifica-se pela entrada da Amazon no Brasil, por ser atualmente esta a maior livraria do mundo e por ter como diferencial seu sistema de recomendação. A loja virtual pode vir a tornar-se líder no mercado brasileiro, como o fez nos EUA e causar rupturas no mercado editorial de publicação de livros, bem como na própria seleção de itens desse mercado. Investigar este objeto pode apontar comportamentos a serem tomados pelos consumidores de informação e pelas próprias livrarias brasileiras, principalmente ao considerar que no mesmo dia 6 de dezembro de 2012, quando a Amazon lançou a versão brasileira do seu site, a livraria Cultura lançou produto semelhante ao da Amazon para a leitura de livros eletrônicos. Portanto, esta pesquisa justifica-se também por sua aplicabilidade. Esta aplicação pode ser tanto para o mercado, ao melhorar as indicações em websites de lojas virtuais, quanto na academia, ao contribuir para a recuperação de informação no que toca o desenvolvimento do conhecimento.

1.2 OBJETO

A maioria dos websites da Internet com grandes bases de dados possuem algum tipo de sistema de recuperação de informação. Estes sistemas partem de simples mecanismos de busca para a procura de itens conhecidos e chegam a avançados sistemas de recomendação de conteúdo que, a partir de seus algoritmos, podem indicar itens de interesse do usuário desconhecidos dele próprio. Os sistemas de recomendação de conteúdo são, por vezes, a garantia de se conseguir recuperar informações relevantes em meio ao montante de dados armazenados nos servidores dos websites da Internet. Dentre esses, o website com sistema de recomendação digital em funcionamento abordado nesta tese é o da Amazon.

1.2.1 Amazon.com

A Amazon foi escolhida como objeto por ser a maior loja virtual do mundo (BRANDT, 2011) que tem em sua natureza a oferta de recomendações aos seus usuários. O sistema da Amazon indica itens que variam desde o mercado editorial de livros e de assinaturas de revistas até produtos de varejo como relógios e roupas. A escolha dá-se

também pelo motivo de ser o website de loja virtual pioneiro em sistemas de recomendação, com o maior número de avaliações de itens por parte dos próprios usuários.

A empresa foi fundada em julho de 1994 por Jeffrey Preston Bezos e tornou-se a maior livraria online do mundo. Antes de fundar a Amazon, Bezos trabalhou por alguns anos com o desenvolvimento de softwares para corretoras de valores de Wall Street, o que foi essencial para aplicar seus conhecimentos em engenharia da computação e aprimorar-se no mundo dos negócios. Segundo biografia autorizada, Bezos desejava empreender em seu próprio negócio desde que deixou a Princeton University, porém ele próprio reconhecia que precisava de um período no mercado de trabalho para encontrar seu negócio original para explorar na Internet. O empreendedor “procurava algo que [...] só conseguiria fazer on-line, algo que não pudesse ser replicado no mundo físico” (BRANDT, 2011, p. 48). A natureza inovadora da Amazon parte desta filosofia empregada desde seu primeiro website em 1995, em que o design de interação projetado por Bezos já era centrado no usuário: “O foco deve ser no cliente, não no site. É bastante óbvio que um site simples é mais fácil de usar que um outro com um monte de coisas extravagantes. Ele foi inflexível quanto a isso” (BRANDT, 2011, p. 16). Brandt explica que até chegar ao modelo pensado para a livraria online, Bezos definiu alguns critérios, dos quais dois deles são essencialmente importantes nesta tese:

Criação de um banco de dados dos livros à venda: [...] os atacadistas já estavam preparando o terreno para a era digital. Bezos foi a um encontro da American Booksellers Association (ABA, Associação dos Livreiros Americanos) e descobriu que todos os livros à venda recebiam um ISBN, que é o International Standard Book Number, um padrão internacional para o número do livro, em português. Foi criado na década de 1960 e depois padronizado pela International Organization for Standardization (ISO, Organização Internacional para Padronização), em 1970. Isso tornou mais fácil criar um banco de dados de livros e depois pesquisá-los pelo ISBN. No final de 1980, a Ingram and Baker & Taylor tinha atualizado sua lista de inventário de microfichas, que um indivíduo precisava examinar manualmente, para o formato digital, que os funcionários podiam pesquisar pelo computador. Para Bezos, parecia que todas as informações sobre livros “havia sido tão meticulosamente organizadas que podiam ser colocadas on-line”.

[...]

Potencial on-line: [...] os programas de software podiam classificar, pesquisar e organizar os títulos e as categorias dos livros, tornando-os mais fáceis de serem encontrados e comprados on-line. A maior livraria do mundo estocava só 175 mil títulos, mas Bezos sabia que os software podia classificar 1 milhão de títulos em um banco de dados, desde que tivesse alguns computadores razoavelmente poderosos. (BRANDT, 2011, p.50)

Portanto, o que Bezos observou no mercado de livros da época era que os itens comercializados já possuíam todos os dados necessários para serem classificados em uma

base de dados e rapidamente recuperados. Como as lojas virtuais não possuem limitações físicas dos itens para oferecer produtos, como as presenciais, a livraria de Bezos poderia rapidamente oferecer maior variedade de itens. Porém, não bastava apenas colocar muitos itens em um único local, era necessário indicar aqueles que realmente interessavam aos clientes. Foi a partir disso que Bezos passou a implementar a ideia de que um avançado sistema de recomendação era necessário para que fossem encontrados os itens que realmente interessavam a cada um dos seus clientes.

No final da década de 1990, “a Amazon comprou um pequeno serviço online chamado PlanetAll, para fazer o cruzamento das vendas de livros com o código postal e os endereços de e-mail das pessoas” (BRANDT, 2011, p. 14). Desse modo, a livraria passou a desenvolver seus algoritmos para aprimorar as recomendações de itens aos seus clientes. A Amazon cruza os dados de compra e as avaliações de um cliente com o histórico de avaliações e compras dele próprio e de outros usuários para recomendar itens. Os dados que alimentam o algoritmo da Amazon para recomendações são gravados a partir da experiência interativa e os rastros de navegação que cada usuário tem no site. “Ao recolher dados sobre os livros que eles haviam comprado e compará-los com os livros adquiridos por clientes com gostos similares, os computadores da Amazon conseguem recomendar títulos que outras pessoas talvez não encontrariam de outra forma” (BRANDT, 2011, p. 18).

Uma das mais inovadoras propostas da Amazon no processo de avaliação de itens foi a de permitir que os próprios clientes publicassem junto a cada item, resenhas críticas dos mesmos. Com isso, a livraria esperava avaliações mais espontâneas do que aquelas pagas a críticos profissionais para promoverem os livros: “como as críticas ruins eram permitidas da mesma forma que as boas, os clientes entendiam que a indicação de livros da Amazon não iria desapontá-los” (BRANDT, 2011, p. 17). Porém, esta é apenas uma das inúmeras formas de recomendação e avaliação de conteúdo no site da Amazon. Na sequência, com o intuito de conhecer-se o objeto de estudo em questão, descrevem-se as formas de avaliação e recomendação de conteúdo na Amazon. Esta descrição é dividida em dois grupos: visitantes que não efetuaram *login*⁴ e visitantes que efetuaram *login*.

⁴ No final desta tese, pode-se encontrar o glossário com termos específicos ao tema pesquisado. Para facilitar a localização desses termos, optou-se por utilizar notas finais indicadas no texto com números romanos.

1.2.1.1 Recomendações e avaliações para visitantes que não efetuaram login

No momento em que um visitante acessa a página da Amazon pela primeira vez, ele recebe ofertas de itens a partir dos destaques previamente definidos pela empresa, porém já há desde a primeira página recomendações baseadas nos históricos de navegação de outros clientes, como pode ser observado em uma parte da página inicial na Figura 1.



Fonte: Amazon.com

A partir do momento em que o visitante passa a fazer buscas no site e a clicar em itens, sua experiência de navegação individual passa a definir os próximos itens a serem ofertados. A Amazon utiliza *cookies*ⁱⁱ para gravar no computador de seus visitantes os dados de navegação deles e reconhecer suas preferências de navegação. Um exemplo comum de *cookie* é quando o site grava dados frequentes utilizados em formulários para que o usuário não necessite digitá-los novamente. Esses dados são gravados no próprio computador do visitante em um arquivo texto denominado *log*ⁱⁱⁱ. Com essa tecnologia em uso, a página inicial da Amazon oferece recomendações dirigidas aos seus visitantes após a primeira busca, mesmo que eles não possuam um cadastro de perfil ou não tenham efetuado *login*. Observa-se nas Figuras 2, 3, 4, 5 e 6 a seguir, recortes da página inicial da Amazon de um visitante que não efetuou *login*, mas que já havia feito algumas buscas no website.

Figura 2 – Mais itens para considerar.

More Items to Consider





You viewed	Customers who viewed this also viewed		
 <p>Linked: How Everything is Connected... > Albert-László Barabási Paperback ★★★★★☆ (132) €16.00 \$11.31</p>	 <p>Connected: The Surprising Power of... > Nicholas A. Christakis, James H. Fowler Paperback ★★★★★☆ (65) €15.99 \$11.40</p>	 <p>Sync: How Order Emerges From Chaos In... > Steven H. Strogatz Paperback ★★★★★☆ (82) €17.99 \$12.77</p>	 <p>Six Degrees: The Science of a... > Duncan J. Watts Paperback ★★★★★☆ (39) €17.95 \$11.76</p>

> [View or edit your browsing history](#)

Fonte: Amazon.com

Figura 3 – Relacionado aos itens vistos por você.

Related to Items You've Viewed





You viewed	Customers who viewed this also viewed		
 <p>Fortis Men's 669.10.31M B-47... \$3,950.00</p>	 <p>Armand Nicolet Men's 9670A-AG-P670NR1... ★★★★★☆ (1) €4,400.00 \$3,960.00</p>	 <p>Fortis Men's 659.27.92 Andora... ★★★★★☆ (3) \$7,236.36</p>	 <p>Fortis Men's 627.22.11 M Spacematic... €9,925.00 \$8,457.26</p>

> [View or edit your browsing history](#)

Fonte: Amazon.com

Figura 4 – Inspirado pelo seu histórico de navegação.

Inspired by Your Browsing History

You viewed	Customers who viewed this also viewed		
 <p>Versace Men's 23C80D008 S009 V-Race... \$1,975.00</p>	 <p>VERSACE 2021 color 100213 Sunglasses ★★★★★☆ (13) €230.00 \$128.50</p>	 <p>Versace Men's M8C80D008 S080... \$1,995.00</p>	 <p>Gold Medusa Heads with a 13mm 36 Inch... JOTW ★★★★★☆ (2) \$39.95</p>

> [View or edit your browsing history](#)

Fonte: Amazon.com

Figura 5 – Itens adicionais a explorar.

Additional Items to Explore

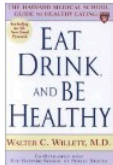
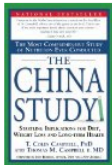
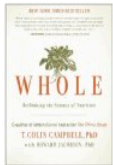
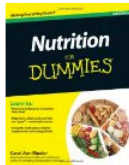
You viewed	Customers who viewed this also viewed		
 <p>Madden NFL 25 Anniversary Edition... Electronic Arts Xbox 360 \$99.99</p>	 <p>NCAA Football 14 Electronic Arts PlayStation 3 ★★★★☆ (45) \$59.99 \$50.17</p>	 <p>NBA 2K14 2K Sports Xbox 360 \$59.96</p>	 <p>Madden NFL 13 Electronic Arts Xbox 360 ★★★★☆ (912) \$29.99 \$26.05</p>

[View or edit your browsing history](#)

Fonte: Amazon.com

Figura 6 – Recomendado baseado no seu histórico de navegação.

Recommended Based on Your Browsing History

 <p>Eat, Drink, and Be Healthy: The... Walter Willett, Edward L. Giovannucci, P. J. Skerrett Hardcover ★★★★☆ (201)</p>	 <p>The China Study: The Most... T. Colin Campbell, Thomas M... Paperback ★★★★☆ (1,691) \$16.95 \$12.66</p>	 <p>Whole: Rethinking the Science of... T. Colin Campbell, Howard Jacobson Hardcover ★★★★☆ (159) \$26.95 \$18.66</p>	 <p>Nutrition For Dummies Carol Ann Rinzler Paperback ★★★★☆ (19) \$24.99 \$15.68</p>
--	---	--	--

[View or edit your browsing history](#)

Fonte: Amazon.com

Observa-se nestas cinco figuras anteriores, que todos os itens que são ofertados na página inicial aproveitam-se dos rastros de navegação do visitante para definir quais itens serão ofertados a partir dos itens já vistos. Desse modo, seria difícil até mesmo prever uma página inicial estática, ou voltar-se a mesma página anterior, visto que cada cliente tem sua própria página que se molda a partir da sua primeira experiência de navegação. A Amazon permite que seus visitantes editem seus históricos de navegação para aprimorar as buscas. Ao clicar em *View or edit your browsing history*, visita-se a seguinte seção do site apresentada na Figura 7:

Figura 7 – Seu histórico de navegação.

[Your Amazon.com](#) > [Your Browsing History](#)
(Click here to [sign in.](#))

Narrow by Category
 ▶ **All Categories** (1)
[Kindle Store](#) (1)

Manage Your Browsing History

Click here to [turn off your browsing history.](#)

Need more help?
 Visit the [Help](#) area.

Your recently viewed items



Kindle Paperwhite, 6" High Resolution Display with Built-in Light, Wi-Fi - Includes Special Offers
 by Kindle (October 1, 2012)
 ★★★★★ (15,150)

Price: \$119.00
[9 used & new from \\$93.99](#) [\[Delete this item\]](#)

Fonte: Amazon.com

Na seção *Seu Histórico de Navegação*, o internauta pode apagar itens recentemente visualizados que não interessam, deletar todos os itens do histórico de navegação ou desabilitar o uso do histórico de navegação para receber itens. Ao clicar em desabilitar histórico de navegação, o visitante para em *Gerencie seu Histórico de Navegação*.

Figura 8 – Gerencie o seu histórico de navegação.

[Your Amazon.com](#) > [Your Browsing History](#) > [Manage Your Browsing History](#)
(Click here to [sign in.](#))

You're in Control

Your Browsing History is meant to help you keep track of some of the items you've recently viewed, searches you've recently made, and product categories you've recently visited.

If you do not enjoy this feature and want to turn it off, simply click Turn Off Browsing History in the box to the left. Turning this feature off will turn it off for anyone who uses your Web browser. If you'd like to re-enable Your Browsing History for your browser, return to this page and click the Turn On Browsing History button that will appear. Because your selection is managed through HTTP cookies, if you delete these cookies or use a different browser, you will have to make this same selection again.

Visit our [help](#) area to learn more.

For information about your browsing history on Amazon Associate Websites, click [here](#).

Fonte: Amazon.com

A empresa ressalta acima do botão observado na Figura 8 que desabilita o histórico de navegação que “*You're in Control*” e no texto ao lado, explica que o histórico ajuda o visitante a receber melhores indicações a partir dos itens vistos recentemente.

A partir da página inicial e na medida em que o cliente alimenta o seu histórico de navegação, as páginas seguintes da Amazon, divididas por departamentos, moldam-se e indicam conteúdo a partir da experiência de navegação individual do visitante em questão e também da experiência de navegação de outros visitantes.

Os visitantes da Amazon que não efetuaram *login* não podem avaliar produtos pelo *Star Rating*^{iv}, escrever resenhas ou votar se as resenhas foram úteis ou não, porém eles podem ler as resenhas e visualizar os votos dos outros internautas. Ao passar o mouse por

cima das estrelas de classificação, o visitante pode observar a quantidade de votos avaliativos feitos de 1 a 5 no item visualizado, exemplificado na Figura 9.

Figura 9 – *Star Rating* – sistema de classificação por estrelas.

The screenshot displays the Amazon product page for the Kindle Fire HD Tablet. The product title is "Kindle Fire HD Tablet" with specifications: "7" HD Display, Dolby Audio, Dual-Band Dual-Antenna Wi-Fi, 16GB or 32GB". It has a star rating of 4.2 out of 5 stars based on 16,109 customer reviews. A pop-up window shows a detailed view of the star rating system, listing the number of reviews for each star level: 5 stars (9,645), 4 stars (2,960), 3 stars (1,500), 2 stars (854), and 1 star (1,150). The pop-up also displays three customer reviews with their star ratings and text. The main product page includes a "Product Tour" button, a "See all 35 customer images" link, and a "Add to Cart" button. The price is listed as \$159.00.

Star Rating	Number of Reviews
5 star	9,645
4 star	2,960
3 star	1,500
2 star	854
1 star	1,150

4.2 out of 5 stars

5 star: (9,645)
4 star: (2,960)
3 star: (1,500)
2 star: (854)
1 star: (1,150)

See all 16,109 reviews

“Picture quality is amazing and speakers are so clear and sound great!”
Maureen R. | 1,836 reviewers made a similar statement

“There are plenty of free apps and with a Prime Membership they have access to many more apps, movies, and books.”
Jared | 943 reviewers made a similar statement

“The size and weight are perfect... comfortable to hold, light, and easy to carry.”
JDeldios | 1,479 reviewers made a similar statement

technology for rich color and deep contrast from any viewing angle

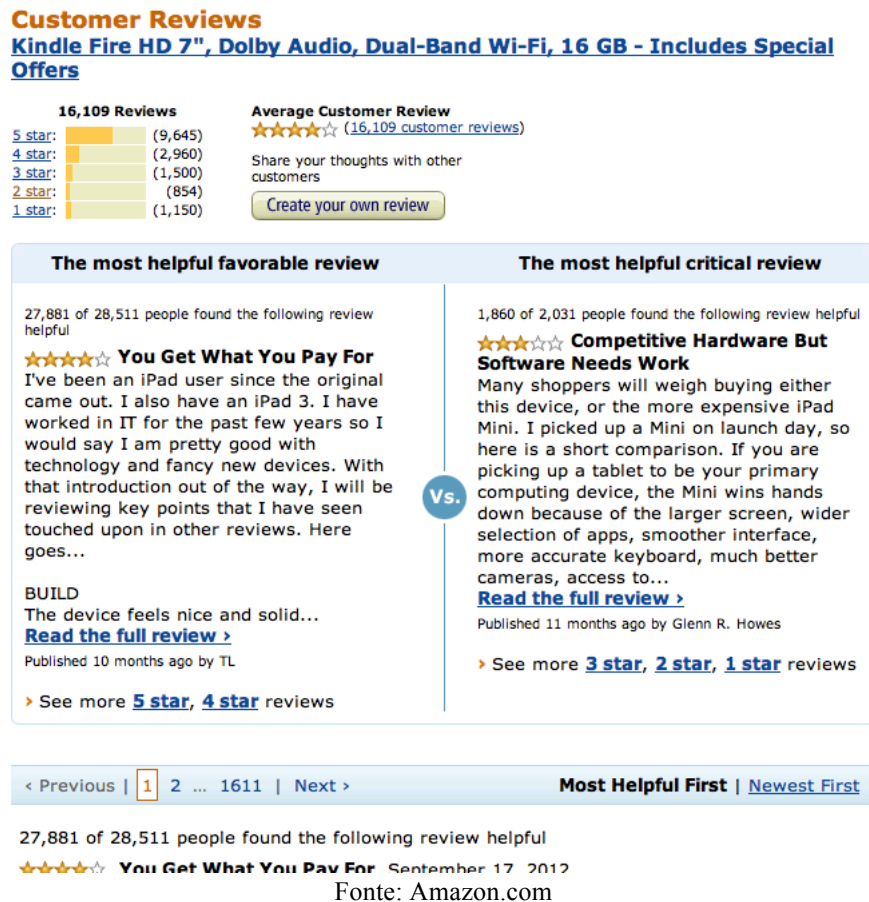
Quantity: 1
Add to Cart
Fire HD, 16GB, Special Offers: \$159.00
to Cart as Shown
or
on 1-Click ordering.
File Accessories
PowerFast for
ed Charging (not
):
14.99
Standing Case:
Black \$44.99 \$29.99
OtterBox Case: Black
\$69.95

Fonte: Amazon.com

O visitante pode clicar em uma das estrelas da escala para ler todas as resenhas dos internautas que avaliaram o produto com aquele índice. Por exemplo, na figura anterior, observa-se que o índice *2 star* está marcado como já visitado. Ao lado do *Star Rating* estão extratos de sentenças dos textos com o número de resenhas que fizeram sentenças semelhantes.

Ao clicar para visualizar as resenhas, é apresentada a página com as resenhas dos internautas, onde é destacada a resenha favorável mais útil ao lado da resenha crítica (desfavorável) mais útil. As resenhas marcadas com estrelas de 4 a 5 são consideradas pela Amazon como favoráveis e aquelas com estrelas de 1 a 3 como desfavoráveis. Na Figura 10, observam-se essas resenhas lado a lado. Abaixo delas, são primeiramente apresentadas as resenhas classificadas como mais úteis pelos votos de outros internautas.

Figura 10 – Página de resenhas de clientes.



O visitante pode clicar no botão *Newest First* para visualizar primeiramente as resenhas mais recentes. Toda a resenha possui título, seguido da data em que foi publicada, da classificação por estrelas definida, do nome do autor, da verificação se a resenha foi escrita por alguém que comprou o item na Amazon, do número de pessoas do total de votantes na resenha que a acharam útil, do *hyperlink*^v para reportar abuso, do *permalink*^{vi}, dos botões para votar se a resenha foi útil ou não e do espaço para comentários. Algumas dessas funções estão ativas apenas para aqueles visitantes que efetuaram *login* e serão descritas no próximo subcapítulo.

1.2.1.2 Recomendações e avaliações para visitantes que efetuaram login

O visitante que acessa a Amazon e efetua *login* tem a sua primeira página denominada como *Your Amazon.com* e é direcionado até esta seção que divide por departamentos as recomendações para o cliente que está conectado com seu perfil cadastrado anteriormente, como observado na Figura 11.

Figura 11 – Your Amazon.com: Página inicial do visitante que efetuou *login*.

The screenshot shows the Amazon.com homepage for a user named Gilberto. At the top, there are navigation links for 'Gilberto's Amazon.com', 'Today's Deals', 'Gift Cards', 'Sell', and 'Help'. A 'Try Prime' button is visible. The main navigation bar includes 'Shop by Department', a search bar with 'Books' selected, and buttons for 'Go', 'Hello, Gilberto Your Account', 'Try Prime', 'Cart', and 'Wish List'. Below the navigation bar, there are links for 'Your Amazon.com', 'Your Browsing History', 'Recommended For You', 'Amazon Betterizer', 'Improve Your Recommendations', 'Your Profile', and 'Learn More'. The main content area is titled 'Recommended for You > Books'. It features a section 'Just For Today' with a 'Browse Recommended' link. Below this is a list of book recommendations. The first recommendation is 'Emergence: The Connected Lives of Ants, Brains, Cities, and Software' by Steven Johnson, with a list price of \$46.00 and a current price of \$12.09. The second recommendation is 'Future Perfect: The Case For Progress In A Networked Age' by Steven Johnson, with a list price of \$26.95 and a current price of \$18.66. The third recommendation is 'Everything Bad is Good for You: How Today's Popular Culture is Actually Making Us'.

Fonte: Amazon.com

Ao clicar para ir até a página inicial da Amazon, o visitante conectado tem as mesmas formas de indicações de quem não efetuou *login* baseado no histórico, porém agora os dados de compra também passam a ser considerados para indicar conteúdos, como é observado na Figura 12 a seguir, em que são ofertados livros baseados na loja Kindle do usuário:

Figura 12 – Recomendação para você na loja Kindle.

Recommendations for You in Kindle Store

The screenshot shows the 'Recommendations for You in Kindle Store' section. It features four book recommendations in a grid layout. Each recommendation includes the book cover, title, author, edition type, customer review (stars and count), price, and a 'Why recommended?' link. The first recommendation is 'The Wealth of Networks: How Social... How Social...' by Yoichi Benkler, a Kindle Edition with a 4.5-star rating (15 reviews) and a price of \$12.07. The second is 'The Shallows: What the Internet Is... What the Internet Is Doing to Our Brains' by Nicholas Carr, a Kindle Edition with a 4.5-star rating (232 reviews) and a price of \$11.46. The third is 'iDisorder: Understanding Our... Understanding Our...' by Larry D. Rosen, L. Mark Carrier, and Nancy A. Cheever, a Kindle Edition with a 4.5-star rating (18 reviews) and a price of \$9.99. The fourth is 'Networked: The New Social... Networked' by Harrison Rainie and Barry Wellman, a Kindle Edition with a 4.5-star rating (12 reviews) and a price of \$16.30.





Fonte: Amazon.com

O visitante cadastrado também recebe indicações em sua página inicial baseadas em sua lista de desejos, porém estas recomendações aparecem de forma intermitente, como observado na Figura 13.

Figura 13 – Recomendações inspiradas na lista de desejos.

Inspired by Your Wish List

You wished for Customers who viewed this also viewed

			
Xbox LIVE 12 Month Gold Membership Card Microsoft Xbox 360 ★★★★☆ (4,912) \$59.99 \$46.88	Xbox 360 Wireless Controller - Glossy... Microsoft Xbox 360 ★★★★☆ (1,250) \$49.99 \$36.37	Xbox Live 1 Month Gold Membership Card Microsoft Xbox 360 ★★★★☆ (99) \$9.95	iPearl mCover Hard Shell Cover Case... mCover® ★★★★☆ (118) \$49.95 \$19.99



[See more in your Wish List](#)

Fonte: Amazon.com

Ao selecionar um item, o usuário recebe recomendações de outros itens baseados nas compras de outros usuários. São indicados itens comprados frequentemente juntos e também itens comprados pelos usuários que também compraram o item em visualização, visto na Figura 14.

Figura 14 – Itens frequentemente comprados juntos.

Frequently Bought Together

 +  **Price for both: \$23.82**

[Add both to Cart](#) [Add both to Wish List](#)

[Show availability and shipping details](#)

- This item:** Alone Together: Why We Expect More from Technology and Less from Each Other by Sherry Turkle Paperback **\$11.76**
- The Shallows: What the Internet Is Doing to Our Brains** by Nicholas Carr Paperback **\$12.06**

Customers Who Bought This Item Also Bought

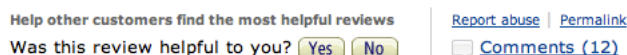
Page 1 of 34

		
The Shallows: What the Internet Is Doing to ... Nicholas Carr ★★★★☆ (232) Paperback \$12.06	Personal Connections in the Digital Age Nancy Baym ★★★★☆ (7) Paperback \$14.46	Dirty Electricity: Electrification and ... MD Samuel Milham ★★★★☆ (5) Paperback \$10.96

Fonte: Amazon.com

O internauta que efetuou *login* tem acesso às mesmas resenhas e avaliações daqueles que não efetuaram *login* no sistema, porém a diferença em conectar-se com seu perfil é que pode-se agora votar nas resenhas que julga útil ou não, reportar abuso e escrever comentários, exemplificado na Figura 15.

Figura 15 – Votos em resenhas e comentários.



Fonte: Amazon.com

Os comentários podem ser feitos para responder a resenha ou para responder os próprios comentários. Pode-se também votar nos comentários para avaliar se eles contribuem ou não para o debate. Como discussões acirradas e até mesmo ofensivas podem emergir nos espaços de comentários, é possível ignorar um cliente específico para ocultar seus comentários e resenhas. A Figura 16 apresenta comentários em uma resenha.

Figura 16 – Comentários das resenhas.

Comments
 Track comments by e-mail
 Tracked by 3 customers

Showing 1-10 of 12 posts in this discussion Sort: **Oldest first** | Newest first

EFK says: Initial post: May 5, 2011 11:52:28 AM PDT
 Hello,
 You mention "taking into account how better integrated minds react to digital technology".
 But what many people are wondering is how digital technology will affect future generations' ability to be better intergrated.
[Reply to this post](#) [Permalink](#) | [Report abuse](#) | [Ignore this customer](#)
 7 of 7 people think this post adds to the discussion. Do you?

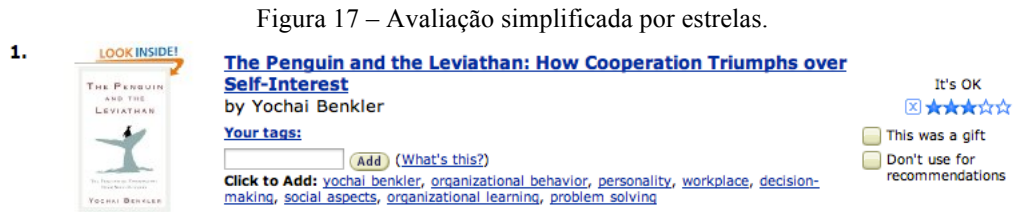
Martin Zook says: In reply to an earlier post on May 5, 2011 3:16:54 PM PDT
 Hey Emmet.
 It's a wide open question, isn't it? I think Turkle is more alarmed about the prospects than I am, but the concerns she points to are valid.
[Reply to this post](#) [Permalink](#) | [Report abuse](#) | [Ignore this customer](#)
 4 of 4 people think this post adds to the discussion. Do you?

IRG says: In reply to an earlier post on May 22, 2011 11:29:11 AM PDT
 Enjoyed your overview. However, I, for one, do remain concerned about the future for those growing up in and being educated in the digital age. Although the concerns apply to all ages since you've older folks displaying the same bad habits as younger ones. (Forty something husband on the computer with porn while loving and attractive wife is alone in bedroom.)

Fonte: Amazon.com

O internauta que efetuou *login* pode fazer resenhas e avaliações dos itens. O sistema permite escrever resenhas e avaliar produtos que o internauta não comprou na Amazon, porém se o item foi comprado na Amazon aparecerá o rótulo de compra verificada na avaliação que o cliente publicar.

A primeira forma de avaliação é simples, basta clicar na estrela desejada para classificar o item como na Figura 17, porém esta forma simplificada é disponibilizada apenas para itens comprados na Amazon.



Fonte: Amazon.com

A forma completa para avaliação de um item na Amazon é a partir da produção de uma resenha. As publicações podem ser feitas a partir de textos ou vídeos e seguem basicamente três etapas: primeiro, o item deve ser classificado pelo *Star Rating*; segundo, deve-se escrever um título para a resenha; e por fim, escreve-se a resenha na própria interface ou sobe-se o arquivo do vídeo com o conteúdo da avaliação gravada, como visto na Figura 18.

Figura 18 – Página de criação de resenhas.

Fonte: Amazon.com

Antes da publicação, o resenhista pode ainda marcar se permite que os outros internautas saibam que ele comprou o item na Amazon, se deseja lembrar esta opção para outras resenhas, e se quer receber por e-mail os comentários feitos em suas resenhas por outros usuários.

A partir dessa descrição, observam-se as formas de recomendação e de avaliação de itens no website da Amazon para posterior problematização e estudo dos sistemas de

recomendações. A descrição do sistema apresentada mostra que tanto os usuários cadastrados na Amazon, quanto aqueles não cadastrados estão por default sob recomendações do sistema baseadas no histórico de navegação, nas avaliações e nas compras do usuário visitante e de todos os outros visitantes do website da Amazon.

1.3 PROBLEMA DE PESQUISA

A busca de informação na Internet é cada vez mais árdua para o internauta que precisa garimpar o que deseja em meio à abundância de dados que ora potencializa a recuperação de resultados inimagináveis no meio físico, ora suscita ruído na busca desejada frente à miscelânea de dados no meio virtual. O paradoxo do acesso à informação na vida contemporânea está diretamente relacionado às bases de dados – abarrotadas de todo tipo de informação – ao passo que, ao mesmo tempo em que se tem acesso a muitos dados, mais difícil torna-se saber quais são verdadeiramente relevantes. A própria relevância desses dados no ambiente digital é colocada em xeque quando uma informação de muita importância para adolescentes pouco importa para adultos e vice-versa, visto à concomitância de públicos presenciada no meio virtual. Ademais, esses dados apresentam alternância de relevância entre questões profissionais e pessoais do internauta. A simples busca por um título editorial feita por um pesquisador exemplifica essa intercalação do interesse, pois em determinado momento ele pode buscar por uma obra literária para seu lazer e em outro por um título acadêmico para desenvolver sua tese. O próprio título literário pode ser de interesse acadêmico, caso sua tese apresente tema linguístico. A diversidade de papéis desempenhados no cotidiano pelo internauta alterna seus interesses em diferentes momentos. Ao pensar nas respostas do sistema de busca de conteúdo on-line, antes mesmo de relevar a qualidade da informação que será oferecida, apresentam-se essas três limitações: da abundância de dados, da concomitância de públicos e da diversidade de papéis.

Os arquivos que compõem as bases de dados no meio virtual são acompanhados de dados que os descrevem, denominados metadados. São dados sobre os dados que permitem a classificação desse conteúdo. Os metadados hoje são, por vezes, gerados no mesmo momento da criação do arquivo, como é o caso das câmeras fotográficas com geolocalização^{vii}. Nestes equipamentos no momento em que um arquivo de imagem é criado grava-se consigo metadados que classificam a imagem pelas técnicas pela qual foi gerada, como a abertura e velocidade do obturador, bem como registra dados potencialmente informativos como a localização geográfica de origem. Esses metadados servem para

organizar a fotografia tanto pela técnica utilizada quanto pela localização em que foi clicada. Os metadados também podem ser gerados manualmente, como no caso da catalogação de livros pelo formato MARC^{viii} nas bibliotecas, pelo qual características das obras são explicitadas em campos específicos que compõem bancos de dados legíveis pelo computador.

A organização do conteúdo no meio virtual supera as limitações físicas dos átomos, já que através dos metadados carregados em cada arquivo é possível classificar um dado específico em diversas categorias e colocá-lo em diferentes locais. O ambiente digital rompe o sistema de classificação de árvore, à medida que a hierarquia na disposição dos conteúdos não precisa seguir uma única categorização de cima para baixo. Por exemplo, para dar conta da limitação da diversidade de papéis que cada usuário desempenha, o livro procurado pelo cientista pode estar distribuído igualmente em diferentes categorias. O mesmo livro de culinária pode estar disponível para quem procura por nutrição ou por dieta. Para atender o problema da concomitância de públicos, criam-se categorias como *Infantis* e *Infanto-juvenis*, *Autoajuda*, etc. Categorizados os conteúdos, precisa-se doravante de um sistema de filtragem para que dados apropriados sejam recuperados, para superar, desse modo, as fraquezas apresentadas pela abundância de dados. Neste cenário, pode-se recuperar respostas relevantes para qualquer conteúdo devidamente categorizado. A assertiva estaria correta, não fosse a conhecida adversidade para categorizar corretamente os conteúdos para posterior recuperação. O obstáculo de respostas relevantes prepondera.

O problema de categorizar os conteúdos no espaço virtual da mesma forma como se faz no espaço físico é que se repetem as mesmas dificuldades ao tentar pensar quais consumidores se interessarão pelo conteúdo – mais penoso que isso – identificar quais perguntas o usuário fará para encontrá-lo. Em uma loja física, resolve-se a questão facilmente com atendentes bem treinados e, mesmo que estes não consigam dar conta de toda a base de dados que a empresa disponibiliza, a capacidade de raciocínio associativo do homem ajuda a encontrar algo em uma miscelânea de dados. No entanto, no virtual – infelizmente ou não – os computadores não contam com essa qualidade humana. A categorização imprecisa de um dado pode impedir que as perguntas dos consumidores o encontrem na base de dados ou – ainda mais ineficiente – apresentar respostas erradas.

Conta-se com especialistas da área da informação e com editores de texto para classificar em bibliotecas e livrarias as obras com o intuito de oferecer respostas relevantes aos consumidores. Os críticos literários e os acadêmicos encarregam-se de diferenciar o bom do mau conteúdo. As editoras e livrarias tomam as providências para expor nas vitrines os títulos mais rentáveis, os quais também estarão em lugares de destaque nas prateleiras das

lojas. O sistema no meio físico, repetido por vezes no virtual, para se obter repostas relevantes, ainda que para o resultado financeiro de empresas, define a forma como os produtos serão apresentados aos consumidores nas prateleiras das lojas físicas e virtuais. Porém, na tentativa de se organizar o conteúdo a ser oferecido para todo o tipo de consumidor, corre-se o risco de não atender bem nenhum deles, ao tratá-los de forma generalista. O Mercado de Massa segue essa lógica de distribuição de produtos.

O mesmo fenômeno pode ser observado na Comunicação de Massa quando os especialistas dos veículos da imprensa selecionam o conteúdo que será oferecido a sua audiência através de seus canais no meio físico e que também se repete por vezes no virtual. Igualmente, na tentativa de atingir a massa, o Marketing de Massa torna as campanhas publicitárias de seus clientes cada vez mais custosas. Ainda que especialistas categorizem os consumidores em faixa etária, classe social, etc., na tentativa de dividir em parte a massa, a adversidade de se oferecer respostas relevantes – neste caso ao consumo de bens e serviços – repete-se, pois o consumidor é pensado de forma generalista.

A forma como a informação é categorizada e distribuída por especialistas para posterior recuperação nas áreas do conhecimento e da comunicação trilhou o caminho para a gestão das primeiras bases de dados em rede. O potencial de expansão dos metadados sob a informação no meio virtual facilita a recuperação de dados a partir de classificações facetadas⁵. Ao procurar um livro sobre TV em livrarias online, em vez de percorrer caminhos da árvore da biblioteca ou livraria como Ciências Sociais Aplicadas>Comunicação>Rádio e Televisão>Videodifusão ou Engenharia>Telecomunicação, digita-se a palavra TV no campo de busca e encontra-se rapidamente todos os livros categorizados nesta divisão. Os especialistas podem romper a classificação por árvore e expandir para múltiplas facetas nos sistemas que gerenciam as bases de dados online. O sistema vai além e permite que o internauta crie suas próprias categorias e organize os dados como melhor lhe convier, ao criar

⁵ Na classificação em árvore os dados são dispostos em categorias hierarquicamente organizadas. Por exemplo, em uma biblioteca um livro pertence apenas a uma das grandes áreas do conhecimento e segue catalogado em subcategorias previamente definidas. Para se encontrar um livro teórico sobre programas de TV, procura-se em Ciências Sociais Aplicadas>Comunicação>Rádio e Televisão>Videodifusão. Isto apontará facilmente seu local na estante da biblioteca para os comunicadores. Porém, sabe-se que a mesma obra pode fazer parte de diferentes áreas e neste caso a classificação facetada poderá ajudar, ao inclui-la também nos catálogos da Antropologia com suas respectivas subcategorias. No meio físico, a classificação facetada tem um limite, pois há a limitação física para recuperar a informação, pois se todos os livros aparecem em todos os catálogos, é como se eles não fizessem parte de nenhum e de nada adiantaria a categorização para a recuperação de informação. No digital, com auxílio da informática, livros podem estar em múltiplas categorias, diferentes facetas, como se a árvore se moldasse no momento de cada busca.

etiquetas (*tags*). O sistema de gerenciamento de base de dados online possibilita a customização do conteúdo para atender os interesses individuais de cada consumidor.

O sistema da Amazon é provavelmente o mais avançado nos dias de hoje na organização de conteúdo, principalmente no mercado editorial de livros que é o principal mercado atendido pela empresa e cerne desta tese. Assim como sistemas de busca, como a plataforma Google, a Amazon grava os dados de navegação de seus consumidores para oferecer respostas mais relevantes aos seus desejos. À medida que o consumidor efetua buscas e faz suas primeiras compras, o sistema oferece respostas customizadas com base na sua experiência individual de navegação e também na dos outros internautas que buscam por itens semelhantes. A experiência coletiva de navegação expande a possibilidade de melhores resultados. Se muitos consumidores procurarem por um assunto e escolherem o mesmo livro para compra, este subirá nos resultados classificados por popularidade. É possível saber qual é a obra mais popular para a busca efetuada, conhecendo exatamente a colocação do livro no ranking de vendas da Amazon.

A Amazon conta ainda com o engajamento de seus consumidores para aprimorar cada vez mais sua base de dados. A partir de resenhas de consumidores, escritas de forma espontânea, e votos de 1 a 5 no sistema de classificação por estrelas, a empresa oferece aos seus usuários produtos que não foram pré-definidos como bom ou mau conteúdo por especialistas profissionais, mas sim baseados na experiência coletiva. Os livros da Amazon não deixam de receber as críticas de especialistas como no Mercado de Massa, mas contam ainda com as recomendações do consumidor e do próprio sistema que oferece dados estatísticos relevantes baseados na experiência da massa.

Ao considerar que os livros da Amazon são itens de informação armazenados em dígitos com seus respectivos metadados, e que o sucesso da empresa está diretamente relacionado à forma como ela engaja os consumidores ao oferecer melhores respostas às perguntas destes, pondera-se que o produto no meio digital é a gestão da informação, visto à abundância de dados (itens de informação) que são cadastrados nas bases de dados.

O sistema da Amazon rompe em parte a filtragem exclusiva dos editores, críticos e livrarias, ao potencializar a filtragem inclusiva no momento em que disponibiliza muitas obras para a venda, indiferente das críticas dos especialistas ou da rentabilidade do produto. A vitrine da Amazon é definida a partir da filtragem colaborativa, que possibilita a inclusão do primeiro livro escrito por um autor desconhecido ao lado do *best-seller* de autores de sucesso, desde que os consumidores assim tenham experimentado. Deve-se aqui diferenciar filtragem participativa de filtragem colaborativa. A primeira diz respeito ao potencial das pessoas

participarem do processo de filtragem com contribuições explícitas, enquanto a segunda diz respeito ao processo automatizado de algoritmos que operam sobre os dados. Na filtragem colaborativa as contribuições são a partir de interações explícitas (votos) e/ou também implícitas (cliques) para alimentar a base de anotações para calcular correlações e indicar conteúdo. Em suma, o internauta pode colaborar com a filtragem sem saber disso.

Observam-se, neste cenário, novos obstáculos a serem pensados, visto que a Amazon potencializa limitações em relação à abundância de dados, concomitância de públicos e diversidade de papéis. Ao procurar por livros sobre TV para sua pesquisa, o cientista pode obter como resposta obras bem avaliadas pela experiência coletiva e mal avaliadas pelos especialistas da área. A decisão a respeito de quem se deve confiar em grandes bases de dados é um desafio custoso não apenas no sistema da Amazon, mas em toda a Internet, em consequência da concomitância de públicos. A escolha do livro certo pode ser decidida a partir de trocas de mensagens com um dos espontâneos resenhistas, mas ainda assim pode ser um trabalho penoso ao consumidor. Ainda que o pesquisador optasse por confiar exclusivamente nas recomendações da Amazon, quanto mais a sua experiência aprimorar as recomendações do sistema para livros acadêmicos, menos este servirá para encontrar leituras de lazer. Enfrenta-se, atualmente, adversidade semelhante no sistema de busca da Google. O internauta poderia desconectar sua conta e navegar off-line ou desligar o histórico de navegação, porém perderia todas as vantagens do sistema baseado em sua experiência individual. São paradoxos que o meio virtual apresenta ao tratar todas as informações como dados, já que por muitas vezes se está frente ao sistema binário, de respostas exatas sim, não. O tratamento subjetivo dos dados hoje, provavelmente, só poderá ser realizado por um humano e talvez por esse motivo, busca-se cada vez mais nas lojas virtuais, nas páginas de sites de redes sociais, nos sistemas de informação e gestão do conhecimento, o engajamento entre consumidores, ora de informação, ora de produtos. A subjetividade humana pode ser a resposta e o engajamento entre estes a solução.

Esta tese pretende investigar se o internauta recorre a mais de um sistema de recomendação quando a busca envolve significativo dispêndio a ele, e se segue exclusivamente sistemas de recomendações automatizados quando o dispêndio diretamente aplicado é baixo. Faz-se necessário observar que o custo aqui referido não envolve necessariamente montante financeiro. Ainda que o investimento financeiro possa servir como parâmetro no processo decisório de compra, o custo aqui referido diz respeito ao quão significativa e dispendiosa é a escolha para o internauta. Por exemplo, o custo que envolve a seleção de um livro para uso acadêmico na tese de um pesquisador será significativamente

mais elevado do que a escolha de um livro para sua leitura de lazer. Nessa linha de raciocínio, por um lado, os sistemas de recomendações objetivos aprimoram a busca por informação que envolve trocas dispendiosas, por outro, diminuem a participação e influência dos sistemas de recomendações subjetivos no processo de seleção de informações cotidianas e de lazer a serem consumidas.

O principal paradigma do meio virtual a ser considerado é que se vive em uma era de espaços de fluxos (CASTELLS, 2009), onde culturas, políticas, mercados e interações sociais dividem o mesmo espaço, como se estivessem em uma miscelânea sócio-digital. Nesse ambiente, vive-se uma Experiência de Massa, onde a experiência individual de cada internauta nos diferentes sistemas contribui para alimentar e – como na Sabedoria das Multidões (SUROWIECKI, 2006) – aprimorar a própria base de dados por onde ele navega. No entanto, apenas a gravação e o cruzamento estatístico probabilístico dos rastros de navegação do internauta nem sempre serão suficientes para se obterem respostas relevantes ao que se deseja. A partir dessa abordagem, o problema de pesquisa que delimita esta tese será: **Como os interagentes selecionam conteúdo sob a influência dos sistemas de recomendação digital?**

1.4 OBJETIVOS

Os objetivos desta tese dizem respeito ao comportamento do usuário na seleção de itens no website da Amazon e estão divididos em geral e específicos.

1.4.1 Objetivo Geral

Investigar o comportamento dos interagentes na seleção de itens frente ao sistema de recomendação digital do website da Amazon.

1.4.2 Objetivos Específicos

- a) Distinguir as indicações de itens nos sistemas de recomendação digital a partir das interações dos usuários da Amazon;
- b) Investigar como os interagentes utilizam o sistema de recomendação digital da Amazon;
- c) Apurar os parâmetros conhecidos que destacam conteúdos no sistema de recomendação digital da Amazon.

1.5 QUADRO TEÓRICO

O quadro teórico desta tese debruça-se em dois principais campos: da Informação, ao estudar os sistemas de recomendações inseridos na recuperação de informação; e, da Comunicação, ao estudar a Interação Mediada por Computador no viés desta área. As principais teorias e autores adotados nesta tese serão abordados nos capítulos a seguir conforme o quadro teórico da Tabela 1:

Tabela 1 – Quadro teórico.

	Sistemas de Informação	Interação Mediada por Computador
Conceitos	Informação; Recuperação de informação; Filtragem Colaborativa; Taxonomia e anatomia dos sistemas de recomendação.	Reciprocidade; Orientações da interação mediada; Interação Social; Interação Reativa; Interação Mútua; Consumo mediado por computador.
Autores	ADOMAVICIUS e TUZHILIN, 2005; BAEZA-YATES e RIBEIRO-NETO, 2013; BALABANOVIĆ e SHOHAM, 1997; BARABÁSI, 2003; BRANDT, 2011; BREESE, HECKERMAN e KADIE, 1998; BUCKLAND, 1991; BUSH, 1945; CASTELLS, 2009; DENNING, 1982; FAIRTHORNE, 1961; FELINTO, 2001; GLEICK, 2013; GOLDBERG <i>et al.</i> , 1992; GOLDBERG, 1992; HU, SMEATON e SUN, 2012; LATOUR, 2005; MANNING, RAGHAVAN e SCHÜTZE, 2008; MONTANER, LÓPEZ e ROSA, 2003; MOOERS, 1950a;1950b;1951;1996; O'REILLY, 2005; OTLET, 1934; PARISER, 2012; RALPH, 1949; RESNICK e VARIAN, 1997; RICH, 1998; SHANNON e WEAVER, 1949; SUROWIECKI, 2006; WAL, 2007; WEINBERGER, 2007; WHEELER, 1994; WISE e PERRY, 1950.	ANDACHT, 2004; BLUMER, 1986; BRAGA, 2008; BRECHT, 2005; CONVILLE e ROGERS, 1998; ENZENSBERGER, 1978; GILLMOR, 2006; GOFFMAN, 1985; HINDE, 1979; LEMOS, 1998;2002; MELO, 1994; NIELSEN, 2009; PRIMO, 2007; RECUERO, 2008; THOMPSON, 1998; WHITE, 1950; ROGERS, SHARP e PREECE, 2013; ROYO, 2008; CAPLAN, 1990; DUARTE, J., 2005; DUARTE, M. Y. M., 2005; FLICK, 2009; LIEBSCHER, 1998; PATTON, 1990; BAUDRILLARD, 1998; BLACKWELL, MINIARD e ENGEL, 2009; FEATHERSTONE, 1995; LEVINE <i>et al.</i> , 2000; LIN, 2007; MONTARDO e ARAÚJO, 2013; NETO e FLORIDIA, 2008.

Fonte: Próprio autor

O aporte teórico inicia o estudo dos sistemas de recomendação a partir dos sistemas de informação no capítulo 2. Parte-se da definição de informação, onde esta é abordada a partir da sua representação como coisa ao adquirir materialidade e permitir o seu manuseio enquanto bit^{ix} de forma automatizada pelos sistemas de recuperação de informação. Segue-se a fundamentação com a abordagem e definição dos sistemas de recuperação de

informação e encerra-se este capítulo com o estudo específico dos sistemas de recomendação, quando estes tem seu funcionamento interpelado acuradamente no que toca a filtragem colaborativa e seus algoritmos na indicação de itens.

No capítulo 3 é estudada a interação mediada por computador. Os sistemas de recomendação digital como observados nesta pesquisa seguem duas linhas para recomendar itens. A primeira é a abordagem objetiva dos dados, quando estes recebem um tratamento automatizado nos sistemas de recomendação e as indicações são resultantes de interações reativas. O exemplo mais usual neste aspecto é a correlação de itens: quem acessou este item, também acessou estes outros. Na segunda linha, os dados recebem um tratamento subjetivo, quando as indicações são feitas a partir de interações mútuas entre os próprios usuários, onde os dados não podem receber um tratamento matemático para análise, visto que dependem do ato cognitivo exclusivo do homem. No objeto sob investigação, as resenhas feitas pelos usuários da Amazon ilustram essa forma de recomendação em sistemas digitais. Portanto, observa-se nos sistemas de recomendação digital um viés da Comunicação Social para sua análise diretamente relacionado com a interação entre pessoas de forma mediada. Desse modo, ao abordar a reciprocidade entre os interagentes nos meios de comunicação e a construção de seus relacionamentos, estudar-se-á como pessoas interagem pelo computador com o objetivo de observar quais interações potencializam indicações ora objetivas, ora subjetivas. No final deste capítulo, o aporte teórico é encerrado com o estudo do consumo mediado por computador, o qual é definido na conjuntura pós-moderna para abordar o processo de decisão do consumidor digital. A forma como os usuários interagem com o sistema pode diferenciar significativamente o tratamento dos dados para a indicação, a avaliação e a seleção de itens em websites sob influência de sistemas de recomendação, como abordados nesta tese.

No capítulo 4 é apresentado o quadro metodológico, para responder ao problema de pesquisa desta tese. Para atender o caráter subjetivo, que envolve as próprias formas de interações entre as pessoas, utilizou-se técnicas de pesquisa qualitativa. As técnicas aplicadas envolveram entrevista em profundidade e observação direta.

No capítulo 5, as recomendações e o comportamento de seleção de itens de informação dos usuários no website da Amazon foram classificados a partir da Matriz de Recomendações de Itens de Informação resultante desta tese que se inicia com o estudo dos sistemas de informação no capítulo subsequente.

2 SISTEMAS DE INFORMAÇÃO

A abordagem adotada nesta tese para a definição, a classificação e a anatomia dos sistemas de recomendação parte da definição de *informação-como-coisa* (BUCKLAND, 1991), pois estes sistemas lidam com informações enquanto conceito e que neste caso está representado pelo bit com potencial informativo. Definida a informação, parte-se para o estudo da recuperação de informação quando primeiro é discutida a forma de pensar o conhecimento em uma era em que a produção supera a capacidade de absorção (BUSH, 1945). Posteriormente, a recuperação de informação é fundamentada enquanto base aos sistemas de recomendação (MOOERS, 1950b). Por fim, os sistemas de recomendação são definidos a partir do estudo de filtragem colaborativa (GOLDBERG *et al.*, 1992), das implicações dos algoritmos implícitos (HU, SMEATON e SUN, 2012; PARISER, 2012), da sua taxonomia (MONTANER, LÓPEZ e ROSA, 2003) e da sua anatomia (RESNICK e VARIAN, 1997).

2.1 A INFORMAÇÃO-COMO-COISA

A definição para informação adotada nesta tese é a do autor Buckland (1991) que, ao aplicá-la em sistemas de informação, a observa como coisa. A definição do autor parte da *informação-como-processo*, quando a pessoa é informada e o que ela sabe até então é modificado. Esse processo segue o mesmo sentido do uso da palavra comunicação adotado por Shannon e Weaver (1949, p. 3): “um sentido muito amplo para incluir todos os procedimentos pelos quais uma mente afeta outra”⁶. A partir dessa perspectiva, informação é o ato de informar, é a “comunicação do conhecimento ou ‘notícias’ de algum fato ou ocorrência; a ação de dizer ou fato de ser informado de alguma coisa”⁷ (BUCKLAND, 1991, p. 351). Então, a transmissão de uma mensagem no momento da leitura da notícia em um jornal – o processo informativo – é o que o autor observa na *informação-como-processo*.

Nesse processo, ao ser modificada, a informação recebe o segundo significado adotado pelo autor, da *Informação-como-conhecimento*:

⁶ Tradução nossa para “a very broad sense to include all of the procedures by which one mind may affect another” (SHANNON e WEAVER, 1949, p. 3).

⁷ Tradução nossa para “communication of the knowledge or ‘news’ of some fact or occurrence; the action of telling or fact of being told of something” (BUCKLAND, 1991, p. 351).

Informação é também usada para denotar aquilo que é percebido na “informação-como-processo”: o “conhecimento comunicado consiste algum fato particular, assunto, ou evento; em que um dos quais é informado ou dito; inteligência, notícia”. A noção de informação como aquilo que reduz incerteza poderia ser visto como um caso especial de “informação-como-conhecimento.” Algumas vezes, a informação aumenta a incerteza.⁸ (BUCKLAND, 1991, p. 351)

Desse modo, a informação quando absorvida é tomada como conhecimento, visto que modifica a referência anteriormente conhecida, ou seja, pode tanto esclarecer sobre alguma noção ao reduzir a incerteza, quanto pode favorecer ainda mais a dúvida frente ao que se sabe. Nesse sentido, a quantidade de dados disponíveis na Internet pode tanto esclarecer a noção do indivíduo, quanto pode confundi-lo ainda mais em relação ao que se busca saber.

Deve-se aqui observar que informação é o que o receptor toma para si enquanto algo que ele reconhece e o modifica, ou seja, se o receptor não é ciente do sinal do que lhe é passado não pode ser considerado informação, como esclarece Fairthorne (1961, p. 255): “informação é um atributo do conhecimento do receptor e interpretação do sinal, e não do emissor, nem de algum observador onisciente, tão pouco do seu sinal em si mesmo”⁹. Desse modo, a informação é a evidência que leva o receptor a interpretar algo. Porém, para ser evidente ele precisa reconhecer aquilo como prova.

Nessa linha, a fim de definir mais adiante a *informação-como-coisa*, o autor Buckland (1991) observa então a informação como evidência. “Se algo não pode ser visto como tendo as características de evidência, então é difícil ver como isto poderia ser considerado como informação. Se tem valor como informação sobre alguma coisa, então parece ter valor como evidência de alguma coisa”¹⁰ (BUCKLAND, 1991, p. 353).

O *Star Rating* é comumente utilizado por sites que desejam permitir que os internautas votem no conteúdo, ao atribuírem de uma a cinco estrelas para o item avaliado. Uma estrela significa que o votante não gostou do conteúdo, enquanto cinco estrelas indicam que julgou excelente. Estas estrelas servem como informação, são a prova da satisfação dos internautas frente àquele conteúdo. Porém, se alguém acessar o sistema e não reconhecer este

⁸ Tradução nossa para “Information-as-knowledge: ‘Information’ is also used to denote that which is perceived in ‘information-as-process’: the ‘knowledge communicated concerning some particular fact, subject, or event; that of which one is apprised or told; intelligence, news’ The notion of information as that which reduces uncertainty could be viewed as a special case of ‘information-as-knowledge.’ Sometimes information increases uncertainty” (BUCKLAND, 1991, p. 351).

⁹ Tradução nossa para “information is an attribute of the receiver's knowledge and interpretation of the signal, not of the sender's, nor some omniscient observer's nor of the signal itself” FAIRTHORNE (1961, p. 255).

¹⁰ Tradução nossa para “If something cannot be viewed as having the characteristics of evidence, then it is difficult to see how it could be regarded as information. If it has value as information concerning something, then it would appear to have value as evidence of something” (BUCKLAND, 1991, p. 353).

sinal, por mais intuitivo que ele possa parecer, ele não será uma informação, por não evidenciar nada para o receptor. Por outro lado, aquele receptor que reconhece o sinal, tem a informação da qualidade do conteúdo baseado na experiência de outros internautas, o que pode levá-lo a consumir ou não o conteúdo. Inclusive, pode modificar a sua opinião frente àquele conteúdo. “A essência da evidência é justamente essa percepção daquilo que pode levar a mudanças no que as pessoas acreditam saber”¹¹ (BUCKLAND, 1991, p. 353).

Desse modo, para algo ser informação é necessário evidenciar algo. O chão petrificado enquanto sinal de que houve uma extensa seca naquele local traz informação quando quem o vê reconhece este sinal. Então, uma fotografia de uma paisagem com essa característica é uma mensagem que transmite a informação da seca ao receptor que a reconhece.

Figura 19 – Reportagem A tortura da seca.



Fonte: (VEJA, 1983)

Porém, se o receptor não compreende como fica o chão quando existe seca excessiva, observará apenas uma paisagem e pouco representará a imagem para ele.

Podemos dizer que aqueles objetos dos quais ninguém está ciente não podem ser informação, embora apressemo-nos a acrescentar que eles bem podem se transformar em informação quando alguém torna-se ciente deles. Não é incomum inferir que algum tipo de prova da qual não estamos cientes deveria ou poderia existir. Se encontrada, seria de particular importância enquanto evidência, como quando detetives investigam mais ou menos sistematicamente por pistas.¹² (BUCKLAND, 1991, p. 356)

¹¹ Tradução nossa para “The essence of evidence is precisely that perception of it can lead to changes in what people believe that they know” (BUCKLAND, 1991, p. 353).

¹² Tradução nossa para “We might say that objects of which nobody is aware cannot be information, while hastening to add that they might well become so when someone does become aware of them. It is not uncommon to infer that some sort of evidence, of which we are not aware, ought to or might exist and, if found, would be of

Em uma investigação policial, vários objetos são anotados, porém só serão considerados provas aqueles que evidenciam algo que possa esclarecer o acontecimento, ou seja, só aqueles que trazem consigo alguma informação. As provas de um crime são os objetos e informações das evidências da materialidade do delito. O objeto quando se torna prova de algo, evidencia alguma coisa, representa então uma informação ao receptor que a reconhece. A materialidade também é observada por Felinto (2001, p. 3) na comunicação:

Em primeira instância, falar em “materialidades da comunicação” significa ter em mente que todo ato de comunicação exige a presença de um suporte material para efetivar-se. Que os atos comunicacionais envolvam necessariamente a intervenção de materialidades, significantes ou meios pode parecer-nos uma idéia já tão assentada e natural que indigna de menção. Mas é precisamente essa naturalidade que acaba por ocultar diversos aspectos e conseqüências importantes das materialidades na comunicação – tais como a idéia de que a materialidade do meio de transmissão influencia e até certo ponto determina a estruturação da mensagem comunicacional.

Ao olhar uma sequência de dígitos na tela de um computador, provavelmente não se compreende nada, pois é necessária uma interface material para a visualização dos dados potencialmente informativos. Assim como a comunicação materializa as mensagens a partir dos seus meios, a informática materializa informação a partir da decodificação de dados visualizados nas interfaces de seus dispositivos. No momento em que um programa decodifica a sequência de bits para sinais reconhecidos, como o fazem as interfaces digitais de computadores, *tablets*^x e *smartfones*^{xi}, passa-se a obter informações a partir desses sinais projetados nas telas desses dispositivos.

A interpretação do sinal é um importante aspecto da *informação-come-conhecimento* a ser observado, visto que o mesmo sinal pode oferecer diferentes significados, como é o caso da figura de uma caveira que em alguns países como no Brasil representa perigo, tristeza, enquanto que em outros, como no México, representa a celebração, alegria dos festejos do Dia dos Mortos. As referências para a associação são diferentes. Esse processo de interpretação do sinal é o que coloca a *informação-come-conhecimento* como algo intangível.

particular importance as evidence, as when detectives search, more or less systematically, for clues” (BUCKLAND, 1991, p. 356).

Uma característica chave da “informação-como-conhecimento” é que ela é intangível: ela não pode ser tocada ou medida de qualquer maneira direta. Conhecimento, crença e opinião são pessoais, subjetivas e conceituais. Portanto, para comunicá-los, eles tem de ser expressados, descritos ou representados de alguma maneira física, como um sinal, um texto ou uma comunicação. Qualquer expressão, descrição ou representação seria “*informação-como-coisa*”.¹³ (BUCKLAND, 1991, p. 351)

Desse modo, antes da informação ser transmitida (*informação-como-processo*) e modificar o que se sabe ao ser tomada como conhecimento (*informação-como-conhecimento*) é necessária sua representação. “Qualquer representação é necessariamente de forma tangível (sinal, sinalização, dados, texto, filme, etc.) e então *representações* de conhecimento (e de eventos) são necessariamente ‘*informação-como-coisa*’”¹⁴ (BUCKLAND, 1991, p. 352). Nesta constituição é que se define a *informação-como-coisa*, enquanto algo tangível, observada em livros, jornais, sons, imagens, interfaces digitais, etc.

Informação-como-coisa: O termo “informação” é também usado atributivamente para objetos, tais como dados e documentos, que são referidos como “Informação” porque eles são considerados como sendo informativos, como “tendo a qualidade de transmitir conhecimento ou de comunicar informação; instrutivo”.¹⁵ (BUCKLAND, 1991, p. 351)

A essência da definição de informação nesta tese é que ela é conceitual. Os elementos conceituais do desenho de ponto, linha, plano e volume ajudam a clarear essa essência conceitual da informação. Da mesma forma que o conceito de ponto enquanto elemento conceitual do desenho não existe em tamanho exato, mas delimita conceitualmente um local específico que ora pode medir um pequeno pixel da tela de um computador, ora pode ser muito maior quando observado de perto em um outdoor, a informação é sempre conceitual ao representar uma ideia ou um objeto. Porém, assim como o conceito de ponto toma forma visual ao ser definido formato, tamanho, cor e textura, a informação é também representada por sinais. O próprio ponto é uma informação em determinado momento. Desse

¹³ Tradução nossa para “A key characteristic of ‘information-as-knowledge’ is that it is intangible: one cannot touch it or measure it in any direct way. Knowledge, belief, and opinion are personal, subjective, and conceptual. Therefore, to communicate them, they have to be expressed, described, or represented in some physical way, as a signal, text, or communication. Any such expression, description, or representation would be ‘information-as-thing’” (BUCKLAND, 1991, p. 351).

¹⁴ Tradução nossa para “Any such representation is necessarily in tangible form (sign, signal, data, text, film, etc.) and so representations of knowledge (and of events) are necessarily ‘information-as-thing’” (BUCKLAND, 1991, p. 352).

¹⁵ Tradução nossa para “Information-as-thing: The term ‘information’ is also used attributively for objects, such as data and documents, that are referred to as ‘information’ because they are regarded as being informative, as ‘having the quality of imparting knowledge or communicating information; instructive’” (BUCKLAND, 1991, p. 351).

modo, tanto ponto, quanto informação tornam-se tangíveis quando representados espacialmente. O bit que é considerado como unidade mínima de informação é tangível, pois mesmo que de forma microscópica ele ocupa algum espaço. Os dados são gravados tanto nos discos rígidos quanto nas memórias flash de forma magnética, ou seja, ocupam algum espaço físico.

A *informação-como-coisa* é a definição que permite armazenar, transportar e transmitir o conhecimento: “pode não haver tal coisa como um sistema especializado ‘baseado-no-conhecimento’ ou um sistema de ‘acesso-ao-conhecimento’, mas apenas sistemas baseados nas representações físicas do conhecimento”¹⁶ (BUCKLAND, 1991, p. 352).

A discussão resume-se em duas distinções: (1) entidade e processo; e (2) intangíveis e tangíveis. Tomadas em conjunto, essas duas distinções produzem quatro aspectos de informação e de sistemas de informação representados na Figura 20.

Figura 20 – Four aspects of information.

	INTANGÍVEL	TANGÍVEL
ENTIDADE	2. <i>Informação-como-conhecimento</i> Conhecimento	3. <i>Informação-como-coisa</i> Dados, documentos
PROCESSO	1. <i>Informação-como-processo</i> Informar-se	4. <i>Processamento de informação</i> Processamento de dados

Fonte: (BUCKLAND, 1991, p. 352)

A informação colocada nesta configuração é de especial interesse para estudar os sistemas de informação, pois é dessa forma que eles a utilizam para armazenar e recuperar a informação, ou seja, para garantir o transporte do conhecimento.

É com a informação neste sentido que os sistemas de informação lidam diretamente. Bibliotecas lidam com livros, sistemas de informação informatizados manipulam dados na forma de bits e bytes físicos; museus lidam diretamente com objetos. A intenção pode ser que usuários tornem-se informados (*informação-como-processo*) e que haverá uma comunicação do conhecimento (*informação-como-conhecimento*). Mas os meios fornecidos, o que é manuseado e operado, o que é armazenado e recuperado, é informação física (*informação-como-coisa*).¹⁷ (BUCKLAND, 1991, p. 352)

¹⁶ Tradução nossa para “On these definitions, there can be no such thing as a ‘knowledge-based’ expert system or a ‘knowledge access’ system, only systems based on physical representations of knowledge” (BUCKLAND, 1991, p. 352).

¹⁷ Tradução nossa para “It is with information in this sense that information systems deal directly. Libraries deal with books; computer-based information systems handle data in the form of physical bits and bytes; museums deal directly with objects. The intention may be that users will become informed (information-as-process) and that there will be an imparting of knowledge (information-as-knowledge). But the means provided, what is handled and operated upon, what is stored and retrieved, is physical information (information-as-thing)” (BUCKLAND, 1991, p. 352).

Os sistemas de recomendação digital que são estudados nesta tese armazenam, transformam e transmitem a informação a partir de bits, ou seja, utilizam a *informação-como-coisa*. Esta forma de uso da informação, ao representar conhecimento, é o que permitirá colocar os diferentes conteúdos da Internet lado a lado, visto que sempre representam algo a partir de sequências de bits. “Os sistemas de informação somente podem lidar *diretamente* com ‘*informação-como-coisa*’ [...]. qualquer coisa *poderia ser informação-como-coisa*”¹⁸ (BUCKLAND, 1991, p. 358). Essa definição da *informação-como-coisa* é o que possibilita a comparação na Internet de sistemas de recomendação de notícias com sistemas de recomendação de produtos comerciais. A semelhança nestes sistemas de recomendação – notícias e produtos – é comumente observada nas interfaces digitais. Cada vez mais a *informação-como-coisa* é difundida pelo desenvolvimento dos dispositivos e sistemas digitais.

Avanços na tecnologia da informação aumentam as possibilidades para criar e usar *informação-como-coisa*. Muito da informação nos sistemas de informação tem sido processada sendo codificada, interpretada, resumida ou senão transformada. Livros são um ótimo exemplo. Virtualmente todos os livros nas coleções são baseados, ao menos em parte, em evidências anteriores, tanto quanto textos e outras formas de informação. A academia está repleta de descrições e resumos, ou, como preferimos chamá-los, representações.¹⁹ (BUCKLAND, 1991, p. 358)

Notícias, produtos comerciais, vídeos, postagens de blogs, relacionamentos por interações mediadas por computador compartilham na Internet a mesma técnica de representação da *informação-como-coisa* a partir dos bits. Como observado por Gleick (2013, p. 163), os próprios campos da ciência e seus teóricos estudam seus objetos por essa partícula comum, o bit:

Cada vez mais, os físicos e os teóricos da informação são iguais uns aos outros, tornam-se uma mesma coisa. O bit é uma partícula fundamental de outro tipo: não apenas minúsculo, mas também abstrato — um dígito binário, um circuito flip-flop, um sim-ou-não. Trata-se de algo sem substância, mas, à medida que os cientistas enfim começam a compreender a informação, eles se perguntam se esta não seria a questão principal: mais fundamental do que a própria matéria.

¹⁸ Tradução nossa para “information systems can deal *directly* only with ‘information-as-thing’ [...] anything might be information-as- thing” (BUCKLAND, 1991, p. 358).

¹⁹ Tradução nossa para “Progress in information technology increases the scope for creating and using information-as-thing. Much of the information in information systems has been processed by being coded, interpreted, summarized, or otherwise transformed. Books are a good example. Virtually all of the books in the collections are based, at least in part, on earlier evidence, both texts and other forms of information. Scholarship is permeated with descriptions and summaries, or, as we prefer to call them, representations” (BUCKLAND, 1991, p. 358).

Gleick (2013), ao citar John Archibald Wheeler (1994), colaborador de Einstein, coloca que é a informação que permite a própria existência. No capítulo *It from bit*, Wheeler (1994, p. 296) diz que a existência de “cada ser — cada partícula, cada campo de força, até o próprio continuum espaço-tempo” é possibilitada pela informação.

Deve-se observar que a intenção, nessa linha de raciocínio proposta da *informação-como-coisa*, apresentada até o momento, não é simplificar todo o conhecimento, as relações, a existência, o saber a um bit, mas entender que essas entidades enquanto informação são representadas como alguma coisa, na abordagem desta tese em bit.

Por tratar-se de representação, ressalta-se que não é uma apresentação fiel do original, pois toda a representação é incompleta. Trata-se de codificação²⁰ e decodificação: “A ação do transmissor em modificar a mensagem para o sinal frequentemente envolve um processo de codificação”²¹ (SHANNON e WEAVER, 1949, p. 8). Portanto, as seguintes características da representação devem ser consideradas na leitura deste texto ao se utilizar a *informação-como-coisa*.

(1) Toda representação pode ser esperada para ser mais ou menos incompleta em alguma consideração. [...] A narrativa escrita refletirá o ponto de vista do escritor e as limitações da linguagem. Filmes e fotografias normalmente mostram apenas uma perspectiva. Alguma coisa do original é sempre perdido. Há sempre alguma distorção, mesmo que apenas através da incompletude. (2) Representações são feitas por conveniência, as quais neste contexto tendem significar o que for mais fácil para armazenar, entender, e/ou pesquisar. (3) Devido a busca por conveniência, representações são normalmente uma troca de evento ou objeto para texto, de um texto para outro texto, ou de objetos e textos para dados. [...] (4) Detalhes adicionais relacionados ao objeto, mas não evidentes nele podem ser adicionados a representação, tanto para informar quanto para desinformar. (5) Representações podem continuar indefinidamente. Pode haver representações das representações das representações. (6) Por razões práticas, representações são comumente (mas não necessariamente) breves e menores do que está sendo representado, concentrando-se nas características que se espera serem as mais significativas. Um resumo, por definição, é uma descrição incompleta.²² (BUCKLAND, 1991, p. 358)

²⁰ O problema da codificação aplicada à recuperação de informação é abordado no artigo *Coding, Information Retrieval, And The Rapid Selector* de Mooers (1950a), onde o pesquisador critica a proposta de codificação de Wise e Perry (1950) para uso do *Rapid Selector* publicada no texto *Multiple coding and the rapid selector*.

²¹ Tradução nossa para “The action of the transmitter in changing the message into the signal often involves a coding process” (SHANNON e WEAVER, 1949, p. 8).

²² Tradução nossa para “Representations have important characteristics: (1) Every representation can be expected to be more or less incomplete in some regard. [...] A written narrative will reflect the viewpoint of the writer and the limitations of the language. Films and photographs usually show only one perspective. Something of the original is always lost. There is always some distortion, even if only through incompleteness. (2) Representations are made for convenience, which in this context tends to mean easier to store, to understand, and/or to search. (3) Because of the quest for convenience, representations are normally a shift from event or object to text, from one text to another text, or from objects and texts to data. [...] (4) Additional details related to the object but not

Como é observada nessas características, a representação da informação oferece problemas limitadores à transmissão fiel do saber original. A percepção do receptor e sua interpretação estão sujeitas a essas limitações no momento em que avalia fotografias, sinais, textos, etc. Essas limitações também foram observadas nos problemas comunicacionais na Teoria Matemática da Comunicação de Shannon e Weaver (1949, p. 4) em três níveis: o problema técnico, “Nível A. Como os símbolos da comunicação podem ser transmitidos precisamente?”; o problema semântico, “Nível B. Como os símbolos transmitidos levam precisamente o significado desejado?”; e o problema da eficácia, “Nível C. Como o significado recebido afeta efetivamente o comportamento no sentido desejado?”²³.

Essas limitações de representação da *informação-corno-coisa* em sistemas de recomendação corroboram ainda a justificativa de investigar como são escolhidos os conteúdos e produtos em sites com sistemas que permitem a recuperação de informação.

2.2 COMO PENSAR A RECUPERAÇÃO DE INFORMAÇÃO

As dificuldades que envolvem a recuperação de informação estão diretamente associadas ao desenvolvimento das técnicas de produção e armazenamento do conhecimento. A partir do momento em que a humanidade passa a produzir e armazenar muito mais informações do que é capaz de absorver, surge a necessidade de instrumentos para recuperar informações relevantes às necessidades demandadas. Por esse motivo, torna-se difícil apontar na história um marco para a gênese do problema da recuperação de informação. Porém, certamente há pontos marcantes na história relacionados ao armazenamento e recuperação de informação que favorecem o seu estudo, como o surgimento de tecnologias de impressão.

A impressão por tipos de Gutenberg disseminou a aprendizagem em massa e contribuiu para a instauração de uma era com economia baseada no conhecimento, e também marcou a publicação de livros em grande escala, o que aumentou ainda mais o problema da recuperação de informação. A impressão por tipos permitiu reduzir significativamente o custo dos livros, a ponto de facilitar sua produção e conseqüentemente consumo. Porém, tornou-se

evident from it might be added to the representation, either to inform or to misinform. (5) Representation can continue indefinitely. There can be representations of representations of representations. (6) For practical reasons representations are commonly (but not necessarily) briefer or smaller than whatever is being represented, concentrating on the features expected to be most significant. A summary, almost by definition, is an incomplete description” (BUCKLAND, 1991, p. 358).

²³ Tradução nossa para “LEVEL A. How accurately can the symbols of communication be transmitted? [...] LEVEL B. How precisely do the transmitted symbols convey the desired meaning? [...] LEVEL C. How effectively does the received meaning affect conduct in the desired way?” (SHANNON e WEAVER, 1949, p. 4)

cada vez mais difícil consumir tudo o que era publicado, mesmo ao se tratar de um único tema. Observa-se neste cenário que a dificuldade de se recuperar informação é um problema secular, mas que só passa a ser tratado como área de pesquisa na Ciência da Informação a partir da metade do século XX, quando a informação recebe tratamento computacional, tanto para armazenamento quanto para recuperação.

Por esses motivos, mesmo que se compreenda os problemas estudados por Otlet (1934) no movimento da documentação, ao criar o sistema facetado de Classificação Decimal Universal, que posteriormente permitiu o desenvolvimento de sistemas de organização de bibliotecas, bem como sua importante definição de documentação ao incluir “objetos potencialmente informativos” (OTLET, 1934, p. 217), serão adotados como objeto teórico e empírico nesta tese os estudos e sistemas que envolvem a recuperação de informação a partir de meados do século passado.

Após a Segunda Guerra Mundial, Bush (1945) publicou o ensaio *As We May Think* no qual problematiza o uso que os cientistas fazem dos arquivos do conhecimento. O texto mostra que a ciência implantou a comunicação que provê o registro do conhecimento além da vida de um indivíduo, porém destaca que o obstáculo não está em publicar ideias, mas sim no seu uso e gerenciamento. O ensaio enfatiza que, com a crescente especialização necessária ao progresso, cresce constantemente a quantidade de pesquisas e que esse crescimento passa a impor barreiras ao acesso de seus resultados. Bush (1945, p. 1) explica que “o pesquisador é surpreendido pelas descobertas e conclusões de milhares de outros colegas – conclusões as quais ele não encontra tempo para absorver, muito menos lembrar – à medida que aparecem”²⁴. O problema está na habilidade do pesquisador em processar a quantidade de resultados de pesquisa publicados. Bush (1945) denuncia que importantes resultados de pesquisa podem se perder no montante de publicações, como exemplifica com o caso das leis da genética de Mendel que passaram despercebidas por uma geração inteira. Administrar esse montante de informação é um dos desafios ao investigador contemporâneo.

²⁴ Tradução nossa para “The investigator is staggered by the findings and conclusions of thousands of other workers - conclusions which he cannot find time to grasp, much less to remember, as they appear” BUSH (1945, p. 1).

A dificuldade parece ser, não tanto por publicarmos excessivamente no ponto de vista de extensão e variedade nos interesses dos dias de hoje, mas porque a publicação tem se estendido além da nossa habilidade presente de fazer uso real do registro. O somatório da experiência humana tem sido expandida a uma taxa prodigiosa, e os meios que usamos para acharmos o caminho pelo consequente labirinto para um importante item momentaneamente é o mesmo que era usado nos dias de navios de velas quadradas.²⁵ (BUSH, 1945, p. 101)

A crítica do autor à época mostra-se cada vez mais pertinente na contemporaneidade, à medida que se desenvolvem vertiginosamente técnicas de produção e publicação de conteúdo, porém aquelas destinadas a recuperar a informação avançam de forma mais lenta ou, em busca da automatização do processo, repetem erros em pleno século XXI antecipados há mais de 60 anos.

Ao abordar as barreiras que a ciência já enfrentava em 1945 e adiantar problemas para consultar os arquivos do conhecimento que ainda estavam por se expandir nesse inflame de produção científica especializada, o autor foi além e sugeriu que o pensamento lógico fosse aplicado mecanicamente à seleção de informações a quem as demandavam. Bush (1945, p. 102) escreveu que “um registro, se é para ser útil para a ciência, precisa ser continuamente ampliado, deve ser armazenado, e acima de tudo deve ser consultado”²⁶.

A ironia da afirmação do autor é que quanto mais se desenvolve a capacidade de expansão e armazenamento de registros científicos, mais barreiras são criadas para consulta desse arquivo do conhecimento. “Mera compressão, claro, não é suficiente; há a necessidade de não apenas produzir e guardar um registro, mas também de consultá-lo, e esse aspecto do problema vem depois. Mesmo a grande biblioteca moderna não é consultada geralmente; é vista aos poucos”²⁷ (BUSH, 1945, p. 102). Para enfrentar tal paradoxo e poder selecionar resultados relevantes frente a esse montante de informação, o autor propõe que se utilize de uma máquina para automatizar a aplicação do pensamento lógico na seleção dos arquivos.

²⁵ Tradução nossa para “The difficulty seems to be, not so much that we publish unduly in view of the extent and variety of present-day interests, but rather that publication has been extended far beyond our present ability to make real use of the record. The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships” (BUSH, 1945, p. 101).

²⁶ Tradução nossa para “A record, if it is to be useful to science, must be continuously extended, it must be stored, and above all it must be consulted” (BUSH, 1945, p. 102).

²⁷ Tradução nossa para “Mere compression, of course, is not enough; one needs not only to make and store a record but also to be able to consult it, and this aspect of the matter comes later. Even the modern great library is not generally consulted; it is nibbled by a few” (BUSH, 1945, p. 102).

Sempre que os processos de pensamento lógico são empregados – ou seja, sempre que o pensamento por um tempo corra sobre um leito aceitável – há uma oportunidade para a máquina. [...] Há a possibilidade imediata de existir uma máquina que manipule as premissas de acordo com a lógica formal, simplesmente por um uso inteligente de circuitos de relé. Ponha um conjunto de premissas em tal dispositivo e vire a alavanca, e ela emitirá prontamente uma conclusão atrás da outra, tudo de acordo com a lei da lógica, com a mesma quantidade de erros que se espera de uma máquina de adição com teclado.²⁸ (BUSH, 1945, p. 105)

O uso de circuitos de relé para a seleção a que Bush (1945) se refere foi o mesmo aplicado posteriormente por Ralph (1949) para desenvolver o *Rapid Selector*. Tratava-se de uma máquina eletrônica para busca de dados registrados em rolos de microfilmes, a partir de uma indexação previamente estabelecida. Segundo Goldberg (1992), o modelo de Ralph (1949) foi baseado em um protótipo desenvolvido pelo próprio Bush entre 1938 e 1940, mas que não chegou a ser operacionalizado. A limitação evidenciada no *Rapid Selector* era o modelo de indexação, o qual seguia uma hierarquia para recuperar o conteúdo gravado.

A máquina certamente acelerava o processo de busca quando se conhecia exatamente o documento, relatório ou obra que se procurava. No entanto, como será abordado mais adiante, na definição de recuperação de informação de Mooers (1950b), como também é afirmado por Bush (1945, p. 105): “A ação primordial de uso é seleção, e isso é um ponto de dificuldade que nos detém”²⁹. Esse é inclusive o ponto principal que difere um sistema de busca de dados de um de recuperação de informação, pois no primeiro busca-se algo conhecido como o título de um livro, obras de um autor específico, assuntos previamente catalogados em palavras-chaves, enquanto no segundo espera-se a indicação de conteúdo baseada em problemas, questionamentos e necessidades que se deseja atender.

A recuperação de dados, no contexto de um sistema de RI, consiste na identificação de quais documentos da coleção contêm as palavras-chaves da consulta do usuário. O que, com frequência, não é suficiente para satisfazer a necessidade de informação do usuário. De fato, o usuário de um sistema de RI está mais interessado em recuperar *informações* sobre um assunto do que em recuperar dados que satisfaçam uma dada consulta. (BAEZA-YATES e RIBEIRO-NETO, 2013, p. 7)

²⁸ Tradução nossa para “Whenever logical processes of thought are employed - that is, whenever thought for a time runs along an accepted groove - there is an opportunity for the machine. [...] It is readily possible to construct a machine which will manipulate premises in accordance with formal logic, simply by the clever use of relay circuits. Put a set of premises into such a device and turn the crank, and it will readily pass out conclusion after conclusion, all in accordance with logical law, and with no more slips than would be expected of a keyboard adding machine” (BUSH, 1945, p. 105)

²⁹ Tradução nossa para “The prime action of use is selection, and here we are halting indeed” (BUSH, 1945, p. 105).

Ao visitar-se um website jornalístico, o sistema deve responder a seguinte questão: Quais são as notícias que devem ser lidas para o leitor manter-se informado? Em um site de produtos comerciais, ao considerar-se a proximidade da data comemorativa do dia das mães, a pergunta que o sistema deverá responder é: Qual é o presente que deve ser comprado para a mãe do cliente? Ao considerar-se que os leitores e clientes possuem peculiaridades diversas, os sistemas precisam estar prontos para fornecerem diferentes respostas para as mesmas perguntas. Os websites de redes sociais também precisam responder perguntas semelhantes, como a que diz respeito sobre com quem se deve manter interações a ponto de ser adicionado como amigo na rede do internauta. Por exemplo, ao basear-se no gosto musical do internauta, a pergunta ao sistema seria: Quais são as pessoas e seus respectivos perfis que o internauta pode interagir para compartilhar músicas? No caso de relacionamentos: Quais amigos o internauta deve adicionar ao seu perfil para interagir, compartilhar e conversar sobre temas comuns? Portanto, a recuperação de informação trata os conteúdos de forma abrangente em detrimento às indexações e às classificações pré-estabelecidas.

O cerne real do problema da seleção, porém, é mais profundo que a demora na adoção dos mecanismos por bibliotecas, ou uma falta de desenvolvimento de dispositivos para seu uso. Nossa inaptidão de chegarmos ao registro é causada, em grande parte, pela **artificialidade dos sistemas de indexação**. Quando os dados de qualquer tipo são colocados no armazenamento, esses são catalogados alfabética ou numericamente, e a informação é encontrada (quando o é) descendo na hierarquia de subclasses.³⁰ (BUSH, 1945, p. 106, grifo nosso)

Aplicar a lógica a uma máquina para obter respostas relevantes não é uma tarefa fácil, pois o que mais se observa nos sistemas criados para consulta dos registros são classificações taxonômicas que ordenam o conteúdo de forma sistemática, ou seja, geralmente percorrem por vocabulários fechados e categorias formais no enquadramento do conteúdo. Catalogar os dados alfabética e numericamente dessa forma hierárquica pode ser uma excelente solução para a organização das prateleiras, porém se mostra ineficaz para gerir informação enquanto conhecimento. Bush (1945, p. 106) alerta que: “a mente humana não funciona dessa forma. Ela opera por associação. Com um item em tela, a mente instantaneamente o relaciona ao próximo sugerido pela associação de pensamentos, de acordo

³⁰ Tradução nossa para “The real heart of the matter of selection, however, goes deeper than a lag in the adoption of mechanisms by libraries, or a lack of development of devices for their use. Our ineptitude in getting at the record is largely caused by the **artificiality of systems of indexing**. When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass.” (BUSH, 1945, p. 106, grifo nosso).

com alguns traços intrínsecos na teia de trilhas feita pelas células do cérebro”³¹. A aplicação do conceito de seleção por associação do autor em uma máquina, ao materializar uma espécie de metáfora do pensamento associativo humano, é realizada em sua proposta de um dispositivo de microfilme ajustável que anteciparia em décadas o conceito de hipertexto, de teia, aplicado posteriormente por Tim Berners-Lee ao implementar a *World Wide Web*.

Seleção por associação, em vez de por indexação, também pode ser mecanizada. Considere um dispositivo futuro de uso individual, que é um tipo de arquivo e biblioteca mecanizados. Isso precisa de um nome e, para cunhar um a esmo, usaremos “memex”. Um memex é um dispositivo no qual um indivíduo armazena todos seus livros, registros e comunicações, e que é mecanizado para que tudo isso seja consultado com máxima velocidade e flexibilidade. É um suplemento íntimo de sua memória aumentada.³² (BUSH, 1945, p. 106).

Por um lado, no primeiro momento em que toda a biblioteca de seu proprietário está armazenada, a memex é como qualquer biblioteca anterior em que todos os registros possuem uma classificação. Por outro lado, a máquina permite que anotações, comentários e novas ligações possam ser adicionadas aos registros pré-existentes. Pode-se dessa forma fazer-se associações entre os registros e registrá-las através de anotações e comentários, ou seja, a informação primeiramente armazenada recebe metadados que podem transformar-se em metainformação nas pesquisas futuras.

Qualquer livro da sua biblioteca pode então ser chamado e consultado com mais facilidade do que se fosse tirado de uma estante. Como o usuário tem diversas posições de projeção, pode deixar um item em posição enquanto ele chama outro. Ele pode adicionar anotações e comentários [...] Esta é a característica essencial do memex. O processo de ligar dois itens juntos é o mais importante.³³ (BUSH, 1945, p. 6-7)

Ao pensar-se em metadado e metainformação, faz-se aqui a mesma analogia referida no subitem anterior ao definir-se *informação-como-coisa* em que o metadado é um

³¹ Tradução nossa para “The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain” (BUSH, 1945, p. 106).

³² Tradução nossa para “Selection by association, rather than by indexing, may yet be mechanized. [...] Consider a future device for individual use, which is a sort of mechanized private file and library. It needs a name, and to coin one at random, ‘memex’ will do. A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory” (BUSH, 1945, p. 106).

³³ Tradução nossa para “Any given book of his library can thus be called up and consulted with far greater facility than if it were taken from a shelf. As he has several projection positions, he can leave one item in position while he calls up another. He can add marginal notes and comments [...] This is the essential feature of the memex. The process of tying two items together is the important thing” (BUSH, 1945, p. 6-7).

registro potencialmente informativo, mas que só poderá ser considerado metainformação ao ser consultado e ao modificar o conhecimento anterior de quem o acessa. A proposta de Bush (1945), ao projetar a memex, evidencia o potencial informativo dos metadados por possibilitar a gravação de ligações associativas entre os registros, o que ele denomina como trilhas associativas, e potencializar seu acesso futuro a partir de caminhos construídos anteriormente e transformar o metadado em metainformação.

Para esclarecer essa indexação associativa, o autor escreve sobre uma hipotética pesquisa na memex sobre arco e flecha, mais especificamente o porquê do arco turco ser superior ao inglês nas batalhas das Cruzadas. A partir das dezenas de livros e artigos que possui em sua memex, o pesquisador inicia em uma obra de referência – sua enciclopédia – e encontra um artigo embrionário do tema. Deixa o microfilme do artigo armazenado na memex projetado na tela da máquina e – na sequência – encontra no histórico um outro item que também fica projetado e é feita a ligação entre os dois registros.

Assim ele segue, construindo uma trilha de muitos itens. Ocasionalmente ele insere um comentário dele próprio, ou conecta-o à trilha principal ou liga-o a uma trilha lateral a um item particular. Quando torna-se evidente que as propriedades elásticas dos materiais disponíveis têm muito a ver com o arco, o usuário faz uma bifurcação numa trilha paralela, a qual o leva através de livros sobre elasticidade e tabelas de constantes físicas. Ele insere uma página de suas análises escritas a mão. Assim ele constrói uma trilha de seu interesse através do labirinto de materiais disponíveis a ele. E essa trilha não desaparece.³⁴ (BUSH, 1945, p. 7)

A trilha construída pelo pesquisador para responder a pergunta que originou a pesquisa apresenta o caminho que ele percorreu de forma associativa entre os itens que possui em sua memex. Essas trilhas não se apagam e podem ser acessadas futuramente para responder outras perguntas. Como também em um exemplo hipotético de Bush (1945), no caso do mesmo pesquisador, em uma conversa com um amigo, anos depois, desejar argumentar sobre a resistência de povos às inovações tecnológicas utilizar a mesma trilha para mostrar a resistência dos europeus aos arcos turcos. Ele vai além, o amigo poderia levar a trilha desse pesquisador para casa e colocar em sua própria memex. Pode assim começar sua própria pesquisa a partir da trilha anteriormente gravada.

³⁴ Tradução nossa para “Thus he goes, building a trail of many items. Occasionally he inserts a comment of his own, either linking it into the main trail or joining it by a side trail to a particular item. When it becomes evident that the elastic properties of available materials had a great deal to do with the bow, he branches off on a side trail which takes him through textbooks on elasticity and tables of physical constants. He inserts a page of longhand analysis of his own. Thus he builds a trail of his interest through the maze of materials available to him. And his trails do not fade” (BUSH, 1945, p. 7)

O projeto da memex de Bush (1945) ficou no papel, mas a forma de indexação associativa pensada mecanicamente pelo autor é observada nos sistemas de recuperação de informação de ponta que estão na Internet. A forma de pensar o uso do conhecimento, de forma automatizada, implementada pelo autor, direcionou as pesquisas e o desenvolvimento desses sistemas. O texto de Bush (1945) a respeito do desenvolvimento da Ciência da Informação na época foi visionário:

Formas totalmente novas de enciclopédias surgirão, pré-fabricadas com uma malha de trilhas associativas correndo por elas, prontos para serem inseridas no memex e ali amplificadas. O advogado tem ao seu dispor todas opiniões e decisões associadas de sua própria experiência, e da experiência de seus amigos e de autoridades. O advogado de patentes tem disponíveis milhões de trilhas concedidas, com trilhas familiares para todos pontos de interesse de seu cliente. O médico, intrigado com as reações de seus pacientes, bate-se com a trilha estabelecida anteriormente em estudos anteriores de casos similares, e corre rapidamente através de histórias de casos análogos, com referências laterais para os clássicos pertinentes a anatomia e histologia. O químico, lutando com a síntese de um composto orgânico, tem toda a literatura de química anterior a ele em seu laboratório, com trilhas seguindo as analogias dos compostos, e trilhas paralelas de seus comportamentos físicos e químicos.³⁵ (BUSH, 1945, p. 8)

O autor além de prever a forma de uso associativo dos registros do conhecimento, antecipa sistemas colaborativos como a Wikipedia. A enciclopédia online é construída a partir da contribuição colaborativa dos internautas que a utilizam. Ao fazer analogia à memex, cada página da Wikipedia é um registro da máquina de Bush (1945), e os links da enciclopédia são as trilhas. Quando procura-se por arco e flecha na Wikipedia, tem-se atalhos, ligações para outros itens deixados por internautas anteriores. As trilhas de Bush (1945) são evidenciadas nos verbetes que se tornam cada vez mais precisos quanto mais pessoas contribuem. Não está escrito em *As We May Think*, mas o projeto da memex evidencia 50 anos antes um dos princípios do conceito de Web 2.0, cunhado por O'Reilly (2005, *online*): “quanto mais pessoas utilizam o serviço, automaticamente melhor ele se torna”³⁶.

³⁵ Tradução nossa para “Wholly new forms of encyclopedias will appear, ready-made with a mesh of associative trails running through them, ready to be dropped into the memex and there amplified. The lawyer has at his touch the associated opinions and decisions of his whole experience, and of the experience of friends and authorities. The patent attorney has on call the millions of issued patents, with familiar trails to every point of his client's interest. The physician, puzzled by its patient's reactions, strikes the trail established in studying an earlier similar case, and runs rapidly through analogous case histories, with side references to the classics for the pertinent anatomy and histology. The chemist, struggling with the synthesis of an organic compound, has all the chemical literature before him in his laboratory, with trails following the analogies of compounds, and side trails to their physical and chemical behavior” (BUSH, 1945, p. 8)

³⁶ Tradução nossa para “the service automatically gets better the more people use it” (O'REILLY, 2005, *online*).

Porém, Bush (1945) não problematizou o montante de informações e conexões que as máquinas armazenariam. Ao valorizar a relevância de um conteúdo por suas conexões, pelo uso constante de suas trilhas, surgem problemas observados em sistemas que mensuram o valor das conexões quantitativamente. O fenômeno é abordado por Barabási (2003) no livro *Linked* onde afirma que “Ricos ficam mais ricos”, ao se referir aos nós que se tornam progressivamente mais visíveis na rede quanto mais conexões recebem. Hubs que já possuem muitas conexões recebem progressivamente mais conexões. Quanto mais visível é um nó no mecanismo de busca, mais cliques e conexões ele recebe e, num círculo vicioso, mais e mais cliques e conexões ele recebe sucessivamente.

Potencialidades apontadas por Bush (1945) e observadas nos sistemas contemporâneos da Web 2.0, bem como problemas muitas vezes gerados por eles próprios no que diz respeito ao peso das conexões em relação aos conteúdos serão abordados mais à frente neste texto. Antes, torna-se necessário definir recuperação de informação, pois como já adiantado existe grande diferença entre a busca de dados conhecidos e de recuperação de informação, pois “o objetivo maior do sistema de RI é recuperar informações que seja úteis e relevantes para o usuário. A ênfase está na recuperação de *informação*, não na recuperação de *dados*” (BAEZA-YATES e RIBEIRO-NETO, 2013, p. 4). Como será exposto na sequência, ao pensar a informação-como-conhecimento, representada como coisa através do bit, a busca dos dados está aquém da recuperação de informação, visto à dificuldade de se manusear digitalmente informação não-numérica.

2.3 A DEFINIÇÃO DE RECUPERAÇÃO DE INFORMAÇÃO

O termo recuperação de informação foi cunhado pelo cientista da computação Calvin Mooers (1950b) em sua pesquisa para desenvolver um mecanismo automatizado para acessar informações armazenadas em grande escala. A intenção era aplicar análises objetivas e científicas na questão de como recuperar informação. “O problema considerado é a revocação do armazenamento de itens de informações não numéricas”³⁷ (MOOERS, 1950b, p. 2). A teoria do autor está associada à própria comunicação humana, que por ser verbal é considerada não numérica e que, como ideias e conceitos, não pode ser mapeada de forma euclidiana.

³⁷ Tradução nossa para “The problem considered is the recall from storage of items of non-numerical information” (MOOERS, 1950b, p. 2).

O problema da recuperação é um problema digital, pois de fato toda a comunicação humana é digital. Recuperação de Informação é um problema não numérico em parte porque a maior parte da comunicação humana é verbal, mas mais importante porque a maioria das ideias ou dos conceitos não podem ser mapeados em um espaço euclidiano tridimensional, ou espaço superior. Enquanto há valores de escalas para a representação de alguma informação, esses são relativamente poucos e sem importância. Conceitos espaciais e métricos não são aplicáveis à maioria das informações, ao menos não de simples formas. Ainda, embora o problema da recuperação de informação seja não numérico, não parece haver qualquer alternativa para o uso de técnicas digitais para a sua solução.³⁸ (MOOERS, 1950b, p. 3)

A informação possivelmente não ocupa espaço para que seja calculada matematicamente como Euclides fez na geometria sólida, técnicas digitais para sua recuperação parecem não solucionar o problema, porém como colocado pelo próprio Mooers (1950b), tais técnicas digitais de recuperação de informação – empregando máquinas – já eram operacionais antes de 1950 e o sucesso em suas aplicações indicavam o direcionamento do progresso na área. Os sistemas digitais contemporâneos aplicados à recuperação de informação evidenciam esse progresso.

Informação não ocupa espaço, não possui tamanho, porém sua representação por métodos de codificação permite que estas sejam recuperadas a partir de sistemas digitais. Porém, “[...] o sucesso comparativo e a complexidade mecânica de cada máquina digital está intimamente conectado com o princípio de codificação digital empregado. Cada item de informação deve ser caracterizado para seleção por um conjunto descritivo de termos ou de ‘descritores’”³⁹ (MOOERS, 1950b, p. 2).

Mooers (1950b, p. 2) observa que os métodos de codificação digital são distinguidos pela maneira que encaram ou se esquivam do problema da descrição da informação. Os métodos citados pelo autor e utilizados na época são:

³⁸ Tradução nossa para “The retrieval problem is a digital problem, for in fact all human communication is digital. Information retrieval is a non-numerical problem in part because the most of human communication is verbal, but more important because most ideas or concepts cannot be mapped into a Euclidean 3-space, or higher space. While there are scale readings for the representation of some information, these are relatively few and unimportant. Spatial and metrical concepts do not apply to most information, at least not at the simpler levels. Yet, though the information retrieval problem is non-numerical, there does not seem to be any alternative to the use of digital techniques for its solution” (MOOERS, 1950b, p. 3).

³⁹ Tradução nossa para “[...] the comparative success and machine complexity of each [digital machine] is intimately connected with the principle of digital coding employed. Each information item must be characterized for selection by a set of descriptive terms or ‘descriptor’” (MOOERS, 1950b, p. 2).

(1) classificação alfabética, (2) classificação com códigos numéricos, (3) classificação decimal de Dewey, (4) método de subcampos exclusivos, (5) sistema único de cartão, (6) codificação por microfilme no *Rapid Selector* (Sistema do Departamento de Agricultura), (7) codificação por microfilme no *Rapid Selector* (revisão da AEC), e (8) codificação aleatória por sobreposição (Zatocoding)⁴⁰.

As especificidades de cada uma das formas de descrição da informação determinam o sucesso na operacionalidade a que se propõe. A classificação alfabética é provavelmente uma das mais eficientes formas de descrever a informação quando o objetivo é a simples ordenação hierárquica dos dados, como a ordenação de listas de livros, arquivos, nomes, etc. Porém, à medida que a intenção está além do nome que identifica o referido item, quando o objetivo é conhecer o conteúdo do item e seu potencial informativo, a simples ordenação alfabética pouco servirá para o propósito da recuperação de informação. O sistema de classificação por códigos numéricos acelerava o resultado de pesquisa, principalmente quando era realizada a partir do *Rapid Selector*, mas carregava consigo o mesmo problema de hierarquização e de ordenação dos itens apresentados na classificação alfabética.

O problema está na forma como os itens são descritos e, como já observado anteriormente ao se definir *informação-como-coisa*, a representação da informação em seu simples processo de materialidade já carrega consigo limitações de codificação e decodificação, o qual é potencializado no tratamento automático dos itens armazenados.

A pergunta que surge é: como descrever itens de forma precisa a serviço da recuperação de informação? O primeiro tratamento à informação apontado por Mooers (1950b) para atingir esse objetivo deve seguir a seguinte linha de raciocínio:

Um item de informação será considerado um único relatório, artigo científico, ou um pedaço unitário de dados. Em um momento veremos que é impossível alcançar a recuperação por qualquer sistema de ordenação dos próprios itens físicos; e a observação lógica deste fato é que a recuperação requer ou alguma forma de varredura na totalidade dos itens de informação na coleção, ou alguma maneira de lidar com abreviações simbólicas dos itens do conteúdo. Varredura na totalidade é humanamente impraticável e tecnicamente indesejável. Portanto, a atenção deve ser direcionada aos métodos de descrição simbólica.⁴¹ (MOOERS, 1950b, p. 4)

⁴⁰ Tradução nossa para “(1) alphabetical sorting, (2) numerical code with sorting, (3) Dewey decimal coding, (4) method of exclusive subfields, (5) unit card system, (6) Microfilm Rapid Selector coding (Department of Agriculture system), (7) Microfilm Rapid Selector coding (AEC revision), and (8) superimposed random coding (Zatocoding)” (MOOERS, 1950b, p. 2).

⁴¹ Tradução nossa para “An item of information will be considered to be a single report, scientific paper, or unitary piece of data. In a moment we will see that it is impossible to achieve retrieval by any system of ordering of the physical items themselves; and the logical observe of this fact is that retrieval requires either some form of scanning in entirety of all information items in the collection, or of some manner of dealing with symbolic

Ao observar cada um dos conteúdos como um único item de informação na base de dados, determina-se que cada relatório, artigo, ou pedaço de informação pode tanto ser apontado de forma precisa quanto pode ser desfragmentado. Porém, como bem colocado por Mooers (1950b), é humanamente impossível e tecnicamente oneroso percorrer todos os dados, mesmo que de forma mecanicamente automatizada, para recuperar tais itens.

A proposta do pesquisador para solucionar o problema é um método de descrição simbólica dos itens a partir de assuntos aleatórios, que poderão ser tratados unicamente pela máquina seletora, o qual denomina como Zatocoding. “O assunto importante de cada documento ou outra unidade de informação é caracterizado ou descrito por meio de um conjunto de ‘descritores’ tomados de um vocabulário formal de tais termos”⁴² (MOOERS, 1951, p. 25).

O método de criar descritores para chamar itens previamente identificados por esses marcadores pode ser relacionado à forma de indexação associativa da memex, visto que pedaços de itens de informação podem ser vinculados entre si. Ao definirem-se atributos específicos a partir de um vocabulário fechado, torna-se possível recuperar informações de forma algébrica.

Consideremos que haja uma linguagem sem sinônimos cujo vocabulário é uma lista fixa de declarações ou ideias, a qual chamaremos alternativamente de “atributos” ou “descritores”. No caso mais simples, esta linguagem é fornecida nada mais algébrica do que o produto lógico dos descritores. Álgebras mais complicadas têm sido utilizadas com muito sucesso em química, mas é possível ir muito além nesta álgebra simples. Ao olhar-se para uma maçã, alguém poderia aplicar os descritores “fruto” e “vermelho”. Cada um é aplicado separadamente, e a maçã, sendo um fruto vermelho, é caracterizada por este produto lógico que descreve aquelas coisas que são ambas vermelho e um fruto.⁴³ (MOOERS, 1950b, p. 4)

A aplicação da álgebra à tabela periódica supera o problema da classificação alfabética e agrega informação aos elementos: “[...] o layout da tabela periódica parece ser o

abbreviations of the content of the items. Scanning in entirety is humanly impracticable and technically undesirable. Therefore, attention must be directed to methods of symbolic description” (MOOERS, 1950b, p. 4)

⁴² Tradução nossa para “The unit of information is a single document, paper, or report. [...] The subject matter of each document or other unit of information is characterized or described by means of a set of “descriptors” taken from a formal vocabulary of such terms” (MOOERS, 1951, p. 25).

⁴³ Tradução nossa para “Let there be an artificial language without synonyms whose vocabulary is a fixed list of statements or ideas, which we shall call alternatively “attributes” or “descriptors”. In the simplest case, this language is given no further algebra than the logical product of the descriptors. More complicated algebras have been used quite successfully in chemistry, but it is possible to go very far this simple algebra. Upon looking at an apple, one could apply the descriptors “fruit” and “red”. Each is applied separately, and the apple, being a red fruit, is characterized by this logical product which describes those things that are both red and a fruit” (MOOERS, 1950b, p. 4).

oposto de uma ordem arbitrária como a alfabética, que não acrescenta informação alguma aos itens por ela organizados. A tabela revelou algumas verdadeiras articulações da natureza” (WEINBERGER, 2007, p. 41). Ainda que não se possa pensar a informação como elementos de forma precisa como na tabela periódica, descritores podem ser precisamente definidos enquanto variáveis a serem tratadas matematicamente e revelarem articulações entre itens a ponto de se tornarem sentenças informativas desconhecidas no momento inicial da pesquisa.

Qualquer item de informação pode ter um conjunto de sentenças feitas sobre ele em termos de descritores escolhidos de um vocabulário. Assim, o item pode ser representado por um complexo desses descritores. Pode haver apenas um número finito de descritores nesse complexo. Onde o conjunto $((a_i))$ é o vocabulário, então o item de informação i^{th} pode ser simbolicamente representado por $C_i(a_1, a_2, a_3...a_n)$. Os requisitos da recuperação de informação, de encontrar informação cuja localização ou existência é a-priori desconhecida, agora exige que seja possível por alguma técnica eficiente especificar uma seleção de complexo C_j por meio de qualquer conjunto ou combinação de descritores escolhidos de qualquer maneira de um vocabulário $((a_i))$. Deve haver total independência na escolha e no uso de descritores.⁴⁴ (MOOERS, 1950b, p. 5)

A partir desse cenário, Mooers (1950b) restringe o campo de exigências para sistemas de recuperação de informação. O sistema precisa ter capacidade de processar informações em coleções com itens na casa superior a 10^6 . Os itens para serem descritos adequadamente devem receber ao menos cinco descritores. No momento da recuperação do item serão necessários no mínimo operações que envolvem três descritores e, por fim, o sistema será dependente de uma máquina. No estudo do pesquisador, foram adotados cartões perfurados que carregavam consigo a localização real do item e seus respectivos descritores. A designação simbólica dada à seleção foi a digital para permitir o tratamento mecânico dos cartões.

Após a designação de descritores aos itens da coleção, com o devido respeito ao vocabulário fechado, basta cruzarem-se os assuntos para recuperar os itens desejados. Como o processamento é feito a partir de uma máquina, quanto mais itens a coleção oferecer, maior será o número de opções e de sequências de itens a serem recuperados. Ao aplicarem-se

⁴⁴ Tradução nossa para “Any information item can have a set of assertions made about it in terms of descriptors chosen from the vocabulary. Thus, the item can be represented by a complex of these descriptors. There can be only a finite number of descriptors in this complex. Where the set $((a_i))$ is the vocabulary, then the j^{th} information item can be symbolically represented by $C_j(a_1, a_2, a_3...a_n)$. The requirements of information retrieval, of finding information whose location or very existence is a-priori unknown, now requires that it be possible by some efficient technique to specify a selection of complexes C_j by means of any set or combination of descriptors chosen in any way from the vocabulary $((a_i))$. There must be complete independence in the choice and use of descriptors” (MOOERS, 1950b, p. 5).

técnicas booleanas nas pesquisas, atingem-se resultados significativos no que toca a recuperação de informação da coleção em questão.

O modelo Booleano é um modelo de recuperação simples baseado na teoria de conjuntos e na álgebra Booleana. Como consequência, o modelo é bastante intuitivo e possui uma semântica precisa. [...] O modelo Booleano considera que os termos de indexação estão presentes ou ausentes nos documentos, ou seja, as frequências na matriz de termos por documentos são todas binárias. Uma consulta *q* é composta por termos de indexação ligados por três conectivos Booleanos: *not*, *and* e *or* (note que alguns sistemas podem optar por implementar um operador “menos” em vez de “not” por razões de eficiência, pois computar o complemento é uma operação cara). (BAEZA-YATES e RIBEIRO-NETO, 2013, p. 29)

Como os descritores são tratados de forma aleatória, o conhecimento presente nos itens da coleção não recebe nenhuma organização até o momento em que alguém requerer a informação. “Apenas nesse momento a ‘organização’ é trazida à existência. Em seguida, o pedido é enquadrado por um conjunto de descritores, a máquina escaneia os registros, e prepara a bibliografia demandada”⁴⁵ (MOOERS, 1951, p. 27). Então, sempre que uma recuperação é atendida, a organização dos itens recebe um novo ponto de vista. Vê-se neste aspecto a realização das trilhas da memex em operação.

O método Zatocoding considera uma premissa do sistema de classificação decimal: cada item de informação no universo pode ser mapeado em um único ponto dentro do intervalo da linha real de 0 a 1. Mooers (1950b, p. 7) destaca que: “Este é o ideal do bibliotecário de ‘apontar precisamente’ a informação”⁴⁶.

Acredita-se mais, [...] que este mapeamento pode ser continuamente feito topologicamente contínuo. O que isto significa é que sobre qualquer ponto no intervalo da linha real há uma vizinhança, ou um pequeno segmento de linha, no qual todos pontos estão associados com o item de informação tendo um conceito fechado similarmente entre si. Este mapeamento de ideias sobre os grupos alinham as ideias – ou as classificam – de acordo com as crenças dos defensores do sistema decimal de Dewey. Além disso, eles acreditam que o mapeamento é tal que a vizinhança sobre um determinado ponto (com a vizinhança assumida a ser ligada, e não quebrada em segmentos) deve conter todos os pontos conceitualmente semelhantes, com ênfase no “todo”.⁴⁷ (MOOERS, 1950b, p. 7)

⁴⁵ Tradução nossa para “Only at that time is “organization” brought into being. Then the request is framed by a set of descriptors, the machine scans the tallies, and prepares a demand bibliography” (MOOERS, 1951, p. 27).

⁴⁶ Tradução nossa para “This is the librarian’s ideal of ‘pin-pointing’ the information” (MOOERS, 1950b, p. 7).

⁴⁷ Tradução nossa para “It is further believed, [...] that this mapping can be made topologically continuous. What this means is that about any point on the real line interval there is a neighborhood, or small segment of the line, in which all points are associated with the information item having a close conceptual similarity to each other. This mapping of ideas onto the line ‘groups’ the ideas – or classifies them – according to the beliefs of proponents of the Dewey decimal system. Moreover, they believe that the mapping is such that the neighborhood

Essa premissa pode até ser confirmada, mas o problema é que ela depende da definição de conceitos fechados, como os descritores, porém segundo Weinberger (2007, p. 57): “O mundo é diversificado demais para qualquer sistema único de classificação funcionar para todos, em todas as culturas, o tempo todo”. O conhecimento é mutável em si e está carregado de valores culturais do momento. A criação de categorias arbitrárias, como o caso de enquadrar o homossexualismo em “Desvios Sexuais”, como um “distúrbio sociopático de personalidade” apontada por Weinberger (2007, p. 34) evidenciam este problema. “O construtivismo social tem razão de dizer que às vezes desenhamos linhas de maneira arbitrária, que isso traz conseqüências reais e que as elites usam linhas arbitrárias para se manterem no poder”. Os responsáveis pelo vocabulário padrão do ZatoCoding poderiam manipular ativamente os descritores para atender seus ideais. Mas o mais grave é que essa designação pode ocorrer de forma passiva, quando os descritores são criados baseados em crenças e conceitos tomados como verdades e baseados em preconceitos. A tentativa de enquadramento de um item de informação a um único ponto, para posterior recuperação, perdura e pior do que não responder as perguntas de pesquisa é oferecer respostas erradas e preconceituosas.

Feita essa importante ressalva em relação a conceitos fechados, o que se faz pertinente nesta tese, é a necessidade de mais de um sistema objetivo para respostas relevantes. Considera-se que o método ZatoCoding, ao usar atributos para descrever itens, pode facilitar a recuperação de informação. Ademais, ao aplicar o método a um vocabulário aberto de atributos, potencializar-se-ia sistemas de recuperação associados à Folksonomia, que “é o resultado da livre etiquetagem pessoal de informação e objetos (qualquer coisa com um URL) para a própria recuperação. A etiquetagem é feita em um ambiente social (geralmente compartilhado e aberto a outros)”⁴⁸ (WAL, 2007, *online*).

A aplicação mais recente a qual se pode associar à Folksonomia é o uso de etiquetas, denominadas hashtags^{xii}, utilizadas em Sites de Redes Sociais como Facebook e Twitter, para compartilhamentos criados pelos próprios internautas. Porém, por tratar-se de recuperação, somente os hashtags utilizados com o fim de posterior acesso, poderão ser associados à Folksonomia. Os websites que melhor exemplificam a Folksonomia são aqueles

about a given point (with the neighborhood assumed to be connected, and not broken into segments) must contain all the conceptually similar points, with emphasis on the ‘all’” (MOOERS, 1950b, p. 7).

⁴⁸ Tradução nossa para “Folksonomy is the result of personal free tagging of information and objects (anything with a URL) for one's own retrieval. The tagging is done in a social environment (usually shared and open to others)” (WAL, 2007, *online*).

destinados para o compartilhamento de links, como é o caso do Delicious⁴⁹ que é uma rede de compartilhamento de links que permite etiquetar endereços da web e ainda acompanhar as etiquetas de outros internautas. O movimento da Folksonomia pode ser evidenciado com grande potencial para a recuperação de informação também na blogosfera. Uma etiqueta pode ser criada para classificar uma postagem em um blog e automaticamente associar o conteúdo a milhares de outras postagens, como ocorre na plataforma Wordpress⁵⁰.

Se Bush (1945) foi um visionário ao pensar a aplicação da indexação associativa na memex e potencializar a Web, Mooers (1951, p. 25) também o foi ao desenvolver o Zatocoding e definir a recuperação de informação a partir de descritores:

Recuperação de informação é o nome dado para o processo ou método por meio de que um esperado usuário de informação está apto a converter sua necessidade de informação em uma lista atual de citações para documentos armazenados contendo informação útil a ele. É o processo de procura ou de descoberta que diz respeito para informações armazenadas. É outro, mais geral, nome para a produção de uma bibliografia demandada. Recuperação de informação envolve os aspectos intelectuais da descrição da informação e suas especificações para busca, e também qualquer sistema, técnica ou máquina que são utilizadas para efetuar a operação. Recuperação de informação é crucial para documentação e organização do conhecimento.⁵¹

Ao aplicar as características e a definição da recuperação de informação a sistemas contemporâneos de recomendação de conteúdo, observa-se que tais aspectos também se aplicam a outras bases de armazenamento de dados, que não exclusivamente àquelas destinadas a conteúdo bibliográfico. Definições de recuperação de informação mais contemporâneas seguem a mesma linha de Mooers:

A Recuperação de Informação trata da representação, armazenamento, organização e acesso a itens de informação, como documentos, páginas Web, catálogos online, registros estruturados e semiestruturados, objetos multimídia, etc. A representação e a organização dos itens de informação devem fornecer aos usuários facilidade de acesso às informações de seu interesse. (BAEZA-YATES e RIBEIRO-NETO, 2013, p. 1)

⁴⁹ Disponível em <<https://delicious.com>>

⁵⁰ Disponível em <<http://wordpress.com/tags/>>

⁵¹ Tradução nossa para “Information retrieval is the name for the process or method whereby a prospective user of information is able to convert his need for information into an actual list of citations to documents in storage containing information useful to him. It is the finding or discovery process with respect to stored information. It is another, more general, name for the production of a demand bibliography. Information retrieval embraces the intellectual aspects of the description of information and its specification for search, and also whatever systems, techniques, or machines that are employed to carry out the operation. Information retrieval is crucial to documentation and organization of knowledge” (MOOERS, 1951, p. 25).

Respeitadas as devidas proporções, o consumidor do sistema de recomendação da Amazon espera converter a sua necessidade de informação em uma lista de itens a serem consumidos da mesma forma que o pesquisador em um sistema de uma biblioteca. Os aspectos que envolvem a descrição de informação acadêmica em uma base de dados científica, também operam sobre itens disponíveis em bases de conteúdos não acadêmicos.

A definição de Mooers (1950b) apresentada aborda essencialmente a recuperação de informação a partir de dois aspectos: codificação da informação a partir de descritores e decodificação a partir de um mecanismo automatizado por cartões lidos pelo *Rapid Selector*, ou seja, resumidamente armazenamento e leitura das informações. Os sistemas de recomendação, como serão tratados a partir do próximo subcapítulo desta tese, lidam com parâmetros além do armazenamento e da leitura. Aproveitam-se do manuseio dos itens para recuperar a informação.

Um novo aspecto pode ser agregado ao problema da recuperação de informação: a experiência de navegação do internauta. Segundo Baeza-Yates e Ribeiro-Neto (2013), a área de recuperação de informação pode ser estudada em dois vieses: centrada no computador, que se ocupa da construção de índices, processamento de consultas e desenvolvimentos de algoritmos; e centrada no usuário, em que a recuperação de informação “consiste principalmente em estudar o comportamento do usuário, entender suas principais necessidades e determinar como esse entendimento afeta a organização e a operação do sistema de recuperação” (BAEZA-YATES e RIBEIRO-NETO, 2013, p. 1). Nesta tese, mesmo que se estude o funcionamento do sistema e seus algoritmos, o objetivo está no comportamento, logo, no ponto de vista centrado no usuário. Como será abordado nos sistemas estudados, o rastro deixado pelo perfil do internauta ao acessar dados interfere tanto nos seus resultados de pesquisa quanto nos de outros internautas.

Ainda que o desenvolvimento de tecnologias e de sistemas exija atualizações no tratamento da informação, o que parece prevalecer é a lei de Mooers (1996, p. 22): “Um sistema de recuperação de informação tenderá por não ser usado sempre que for mais penoso e incômodo para o cliente ter informação por ele do que não tê-la”⁵².

Porém, se o problema era codificar as representações materiais da informação para recuperá-las posteriormente – catalogar livros para encontrá-los nas prateleiras – o

⁵² Tradução nossa para “An information retrieval system will tend not to be used whenever it is more painful and troublesome for a customer to have information than for him not to have it” (MOOERS, 1996, p. 22).

problema agora é materializar os resultados de cruzamentos de dígitos que podem oferecer dados informativos – cruzar dados de navegação dos internautas para recomendar conteúdo.

2.4 SISTEMAS DE RECOMENDAÇÃO

Os sistemas de recomendação presentes na Internet, como abordados neste texto, são parte dos sistemas de recuperação de informação. A classificação destes sistemas de recuperação pode ser feita a partir da distinção da escala de dados em que eles operam (MANNING, RAGHAVAN e SCHÜTZE, 2008). Basicamente, o tamanho e armazenamento da base de dados delimitam as diferenças que dividem os sistemas de recuperação de informação em três grandes grupos: **sistemas de busca da web** (Google⁵³, Bing⁵⁴, Yahoo⁵⁵), que operam com consultas que lidam com bilhões de documentos armazenados em milhões de servidores conectados através do hipertexto; **sistemas de recuperação de informação pessoal** (Mac OS X Spotlight^{xiii}, Windows 8 File Explore^{xiv}, Mail^{xv}, Outlook^{xvi}), os quais respondem às consultas e classificações dos documentos pessoais geralmente – não necessariamente⁵⁶ – armazenados no próprio computador do usuário; e os **sistemas de busca empresariais, institucionais e de domínio específico** (Amazon, Last.fm⁵⁷, SABi⁵⁸), que são aqueles em que a base de dados é tipicamente centralizada e objetivam recuperar informações de um domínio específico. Neste último grupo, estão inseridos os sistemas de recomendação presentes na Internet similares ao do objeto de estudo desta tese que é o sistema de recomendação da Amazon.

Ainda que o foco para esta tese esteja nos sistemas de domínio específico, características dos outros dois grupos serão abordadas ao longo do texto que segue, visto que os três grupos de sistemas de recuperação podem ser utilizados em conjunto pelo usuário que busca por informação. Por exemplo, o uso do sistema de buscas Google Scholar^{xvii} pode ser associado ao sistema de buscas da Amazon quando o usuário busca por livros acadêmicos. Estes aspectos serão tratados na sequência, que inicia-se com a origem dos sistemas de recomendação na Internet.

⁵³ Disponível em <<https://google.com.br>>

⁵⁴ Disponível em <<http://bing.com>>

⁵⁵ Disponível em <<http://yahoo.com>>

⁵⁶ A ressalva é necessária devido à tendência de armazenamento na nuvem ser cada vez mais frequente.

⁵⁷ Site de Rede Social destinado à recomendação de músicas. Disponível em <<http://www.lastfm.com.br>>

⁵⁸ SABi – Sistema de Automação de Bibliotecas da UFRGS. Disponível em <<http://sabi.ufrgs.br>>

2.4.1 Da Filtragem Colaborativa aos Sistemas Implícitos de Recomendação

A abordagem fundadora em direção aos sistemas de recomendação de conteúdo na Internet é do cientista da computação Peter Denning (1982) que cunhou o termo *Lixo Eletrônico*⁵⁹ em seu ensaio de mesmo nome. Ao apresentar a potencialidade de ferramentas informatizadas para a produção e difusão de documentos, o escritor denuncia problemas em relação ao uso dessas técnicas automatizadas e em relação à recepção de conteúdo.

Não há dúvida que essas ferramentas, [sistemas de e-mail, editores de texto, planilhas eletrônicas, etc.] especialmente quando todas são disponibilizadas juntas, aperfeiçoam significativamente a produtividade daqueles que precisam escrever e distribuir documentos. Mas tais ferramentas também podem aumentar o palavreado sem aumentar o número de ideias: elas tornam fácil combinar uma série de parágrafos a partir de vários arquivos e chamar o resultado de um documento. [...] Quem salvará os receptores de afogar-se numa maré crescente de informação assim gerada?⁶⁰ (DENNING, 1982, p. 163-164)

O problema causado pelo lixo eletrônico nas caixas de entrada de e-mail na era contemporânea dispensa explicação, visto que mesmo com avançados sistemas AntiSpam^{xviii} ainda perdura o problema evidenciado desde 1982, quando “mensagens urgentes e pessoais são confundidas com mensagens de rotina e de negócios”⁶¹ (DENNING, 1982, p. 164). Os programas de gerenciamento de e-mail estão na ordem dos sistemas de recuperação de informação pessoal, mas mesmo ao considerar a base de dados do usuário em bem menor escala do que a de sistemas de busca da web ou de domínio específico, padecem os gerenciadores de mensagens eletrônicas sob o montante de documentos que processam para adequada recuperação de informação.

Para minimizar o problema do lixo eletrônico, o autor propõe oito assistências para a recepção de informação via mídia eletrônica: “(1) organização hierárquica da caixa de entrada, (2) caixas de entradas privadas⁶², (3) formas especiais de recebimento, (4) filtros de conteúdo, (5) números de importância, (6) batente de recepção, (7) desfragmentador do

⁵⁹ Tradução nossa para “Electronic Junk” (DENNING, 1982).

⁶⁰ Tradução nossa para “There is no question that these tools, [sistemas de e-mail, editores de texto, planilhas eletrônicas, etc] especially when all are available together, greatly improve the productivity of those who must write and distribute documents. But such tools may also increase the verbiage without increasing the number of ideas: they make it easy to combine an assortment of paragraphs from various files and call the result a document. [...] Who will save the receivers from drowning in the rising tide of information so generated?” (DENNING, 1982, p. 163-164).

⁶¹ Tradução nossa para “Urgent and personal mail is mingled with routine and business mail” (DENNING, 1982, p. 164).

⁶² Deve-se considerar que na época era comum as caixas de e-mail serem públicas.

documento, e (8) certificação de qualidade”⁶³ (DENNING, 1982, p. 164-165). Dentre esses oito aspectos de ajuda propostos, dois merecem mais atenção:

FILTROS DE CONTEÚDO. A pessoa pode anexar um processo a uma caixa de entrada para varrer o assunto e os campos das palavras-chaves, e até mesmo o texto, das mensagens recebidas. Este processo pode descartar mensagens indesejadas e encaminhar outras para caixas de entrada de pessoas que normalmente lidam com estes tipos de mensagens.

NÚMEROS DE IMPORTÂNCIA. Cada mensagem vem com um número de importância selecionado pelo remetente. As mensagens são armazenadas em ordem decrescente de importância nas caixas de entrada. Para prevenir inchaço, uma pessoa indicará sua própria opinião sobre a importância de cada mensagem recebida, e um número de “tendência” será atualizado para cada remetente. Mensagens recebidas serão subidas ou rebaixadas de acordo com a tendência gravada por cada remetente. Remetentes serão notificados das tendências gravadas por eles.⁶⁴ (DENNING, 1982, p. 164)

O ato de marcar mensagens indesejadas por um usuário que contribui para filtrar conteúdo é uma maneira de usar um descritor de forma colaborativa. Ao definir números de importância para cada mensagem enviada – sinalizar o conteúdo para ordená-lo de forma prioritária na caixa do destinatário – um usuário colabora com a hierarquização do conteúdo do outro. Embora os programas de gerenciamento de e-mail estejam no grupo de sistemas de recuperação pessoais, observa-se especificamente nessas duas assistências propostas por Denning (1982) interações mediadas por computador entre os usuários, que afetam seus conteúdos entre si. Há interferência humana no processo de filtragem.

Ainda que o pesquisador não tenha criado seu próprio programa de gerenciamento de e-mails para aplicar sua proposta de técnicas para filtrar mensagens nas caixas de entrada, estas assistências serviram de ponto de partida para sistemas de recomendação, como é o caso do Tapestry, que é considerado o primeiro sistema de recomendação informatizado.

O Tapestry foi um sistema experimental de e-mail desenvolvido por Goldberg *et al.* (1992) na Xerox Palo Research. A motivação dos autores em desenvolver o sistema parte

⁶³ Tradução nossa para “(1) Hierarchical Organization of Mailboxes, (2) Separate Private Mailboxes, (3) Special Forms of Delivery, (4) Content Filters, (5) Importance Numbers, (6) Threshold Reception, (7) Documents Skimmer, (8) quality Certification” (DENNING, 1982, p. 164-165).

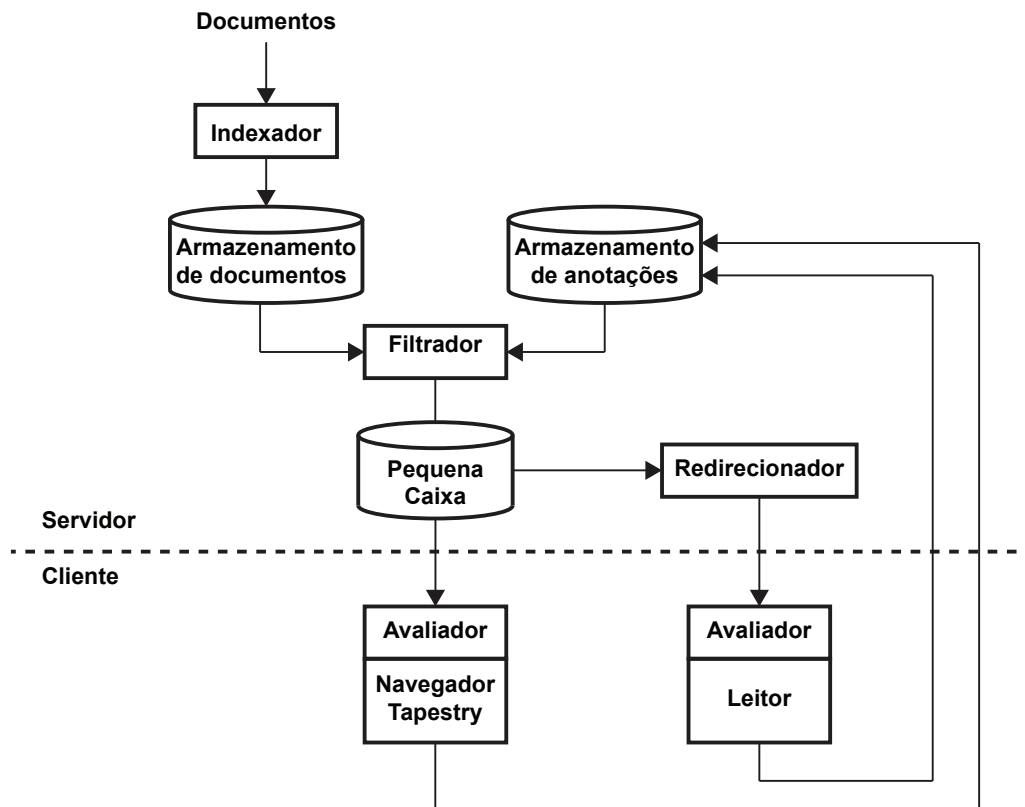
⁶⁴ Tradução nossa para “CONTENT FILTERS. A person can attach a process to a mailbox to scan the subject and keyword fields, and even the text, of incoming messages. This process can discard unwanted messages and forward others to mailboxes of persons who normally handle those types of messages. IMPORTANCE NUMBERS. Each message comes with an importance number selected by the sender. Messages are stored in order of decreasing importance numbers in mailboxes. To prevent inflation, a person will state his own opinion of the importance of each received message, and a "bias" number will be updated for each sender. Incoming messages will be upgraded or downgraded according to the recorded bias for each sender. Senders will be notified of the biases recorded for them.” (DENNING, 1982, p. 164).

do trabalho de Denning (1982) para solucionar o problema de inundação de mensagens consideradas lixo eletrônico nas caixas de entrada. A premissa do sistema é que a participação de humanos melhora o processo de filtragem: “Um princípio básico do trabalho do Tapestry é que uma filtragem mais eficaz pode ser feita através do envolvimento de humanos no processo de filtragem”⁶⁵ (GOLDBERG *et al.*, 1992, p. 61).

O Tapestry permite que seus usuários criem anotações nos documentos. Essas anotações passam a operar nas futuras consultas e entregas de documentos nas caixas de entradas. O fluxo dos documentos no sistema está sob a seguinte arquitetura apresentada na Figura 21, na qual: o indexador é responsável pela leitura e por particionar em campos de indexação os documentos de fontes externas, como de e-mail, da NetNews ou de agências de notícias; o **armazenamento de documentos** é mantido por longo prazo para formar a base de dados e aprimorar as consultas; no **armazenamento de anotações** são mantidas as anotações associadas aos documentos; o **filtrador** opera o lote de consultas fornecidas pelo usuário sobre o conjunto de documentos e os entrega na caixa pequena; cada proprietário possui uma **pequena caixa**, onde documentos de seu interesse podem ser lidos e apagados através do seu leitor de documentos; o **redirecionador** é o servidor responsável por entregar periodicamente os documentos àqueles que desejam utilizar seu próprio leitor de documentos; no **avaliador** aplica-se a classificação personalizada aos documentos do usuário; e, por fim, no **navegador** ou no **leitor** de mensagens é provida a interface para acesso aos serviços do Tapestry, onde existem facilidades, como adicionar, deletar e editar filtros para recuperar documentos, organizar documentos em pastas, fornecer anotações e realizar consultas *ad hoc*.

⁶⁵ Tradução nossa para “A basic tenet of the Tapestry work is that more effective filtering can be done by involving humans in the filtering process” (GOLDBERG *et al.*, 1992, p. 61).

Figura 21 – Fluxo de documentos através do Tapestry.



Fonte: (GOLDBERG *et al.*, 1992, p. 63)

Ao aplicar no Tapestry a participação do usuário, o sistema passa a contar com a subjetividade humana no processo de classificação dos documentos e da recuperação de informação. Como as anotações do sistema são armazenadas no servidor é possível compartilhar a base com todos os usuários do Tapestry, ou seja, pode-se aproveitar as anotações dos usuários entre si. Ao possibilitarem essa interação de forma participativa, Goldberg *et al.* (1992) cunharam o conceito de filtragem colaborativa, que definem como:

Filtragem colaborativa significa simplesmente que pessoas colaboram umas com as outras para ajudar a realizar filtragem através da gravação de suas reações nos documentos que leem. Tais reações podem ser de um documento que foi particularmente interessante (ou particularmente desinteressante). Estas reações, geralmente chamadas de anotações, podem ser acessadas pelos filtros dos outros.⁶⁶ (GOLDBERG *et al.*, 1992, p. 61)

A filtragem colaborativa está presente hoje na maioria dos sistemas de recomendação de conteúdo da Internet. No momento em que se reporta uma mensagem como

⁶⁶ Tradução nossa para “Collaborative filtering simply means that people collaborate to help one another perform filtering by recording their reactions to documents they read. Such reactions may be that a document was particularly interesting (or particularly uninteresting). These reactions, more generally called annotations, can be accessed by others’ filters” (GOLDBERG *et al.*, 1992, p. 61).

Spam ou Phishing^{xix} no aplicativo web de e-mail Gmail, colabora-se com os demais usuários do sistema que inclusive podem ser protegidos de fraudes que podem causar prejuízos financeiros. Como a filtragem colaborativa é feita a partir da gravação das reações dos usuários em relação aos documentos da base de dados, no simples ato de clicar no link de uma notícia do site Folha de S. Paulo⁶⁷, por exemplo, se contribui para a filtragem dos conteúdos mais populares. “Pela introdução de um objeto de anotação, o qual sempre tem um campo de mensagem que o conecta a um documento, os tipos de consultas que suportam filtros colaborativos se tornam simples”⁶⁸ (GOLDBERG *et al.*, 1992, p. 65). No momento em que se visita dois produtos na loja da Amazon, cria-se uma trilha associativa entre eles que servirá de recomendação para outros usuários no modelo de *Clientes que visualizaram este item também visualizaram aquele outro*.

Como seus próprios desenvolvedores escreveram, o design do Tapestry potencializou a aplicação da filtragem colaborativa além dos programas gerenciadores de correio eletrônico:

Tapestry é mais que um sistema de e-mail, porque ele é projetado para manusear qualquer fluxo de entrada de documentos eletrônicos. Mensagens eletrônicas são apenas um exemplo de tamanho fluxo: outros são notícias de agências e artigos da NetNews. Além disso, Tapestry não é apenas um mecanismo de filtragem de correio, ele é também um repositório de mensagens enviadas no passado. Tapestry unifica consultas *ad hoc* sobre este repositório com a filtragem dos dados de entrada.⁶⁹ (GOLDBERG *et al.*, 1992, p. 63)

A manutenção da base de dados a longo prazo, tanto dos documentos quanto das anotações sobre eles, deixa o sistema cada vez mais sujeito às ações dos usuários quando opera a filtragem colaborativa. “A novidade do Tapestry reside em seu suporte para a filtragem colaborativa. Os usuários são convidados a criar anotações nos documentos, e estas anotações podem então ser usadas para filtragem”⁷⁰ (GOLDBERG *et al.*, 1992, p. 69). Este convite à participação faz-se cada vez mais presente na Internet. Websites jornalísticos,

⁶⁷ Disponível em <<http://www.folha.uol.com.br>>

⁶⁸ Tradução nossa para “By introducing an annotation object, which always has a field msg that links it to a document, the kind of queries that support collaborative filters become simpler.” (GOLDBERG *et al.*, 1992, p. 65).

⁶⁹ Tradução nossa para “Tapestry is more than a mail system, because it is designed to handle any incoming stream of electronic documents. Electronic mail is only one example of such a stream: others are newswire stories and NetNews articles. Moreover, Tapestry is not only a mechanism of filtering mail, it is also a repository of mail sent in the past. Tapestry unifies *ad hoc* queries over this repository with the filtering of incoming data.” (GOLDBERG *et al.*, 1992, p. 63).

⁷⁰ Tradução nossa para “The novelty of Tapestry lies in its support for collaborative filtering. Users are encouraged to annotate documents, and these annotations can then be used for filtering.” (GOLDBERG *et al.*, 1992, p. 69).

institucionais, empresariais e até mesmo weblogs e páginas pessoais convidam seus visitantes a comentar conteúdo, criar resenhas, Curtir^{xx} no Facebook, marcar no g+^{xxi} e criar anotações sobre os documentos que acessam para servirem de filtragem a outros internautas. Quando respondem ao convite, os internautas contribuem explicitamente com o sistema. Porém, observa-se que não são todas as anotações sobre os documentos que se dão de forma explícita.

Os desenvolvedores do primeiro sistema de recomendação, Tapestry, cunharam o termo “filtragem colaborativa” e muitos outros o tem adotado. Nós preferimos o termo mais geral “sistemas de recomendação” por duas razões. Primeiro, recomendadores podem colaborar não explicitamente com receptores, que talvez sejam desconhecidos uns dos outros. Segundo, recomendações podem sugerir itens particularmente interessantes, além de indicarem aqueles que precisam ser filtrados para fora.⁷¹ (RESNICK e VARIAN, 1997, p. 56)

Como é criticado por Resnick e Varian (1997), o uso do termo colaborativo na filtragem pode ser capcioso em participações implícitas. Como os autores escrevem, nem sempre os usuários envolvidos colaboram explicitamente uns com os outros. A palavra colaboração sugere trabalhar em conjunto, mas na maior parte das vezes os envolvidos nem se conhecem. Filtrar sugere colocar algo fora, limpar o lixo eletrônico, tirar o ruído, enquanto que uma base de dados ideal só armazenaria documentos que não são lixo, pois se o fossem nem deveriam estar armazenados. Por esses motivos, preterem o uso do termo filtragem colaborativa e preferem sistemas de recomendação.

Esta tese não deseja descobrir qual é a melhor nomenclatura para estes sistemas, mas a reflexão se torna necessária devido aos motivos que originaram a crítica relacionada à participação implícita do usuário. Por tratar-se de indicações baseadas em ações de desconhecidos, como se garante a fidedignidade das recomendações? Outro questionamento a ser levantado é que por serem as respostas muitas vezes baseadas em reações implícitas, como garantir que o sistema foi automaticamente (implicitamente) alimentado com as informações verdadeiramente relevantes?

A alimentação do sistema de armazenamento de anotações nos sistemas contemporâneos se sobressai aos aspectos das ações implícitas do usuário. Hu, Smeaton e Sun (2012) mostram que dispositivos como *tablets* utilizam tecnologias sensoriais para

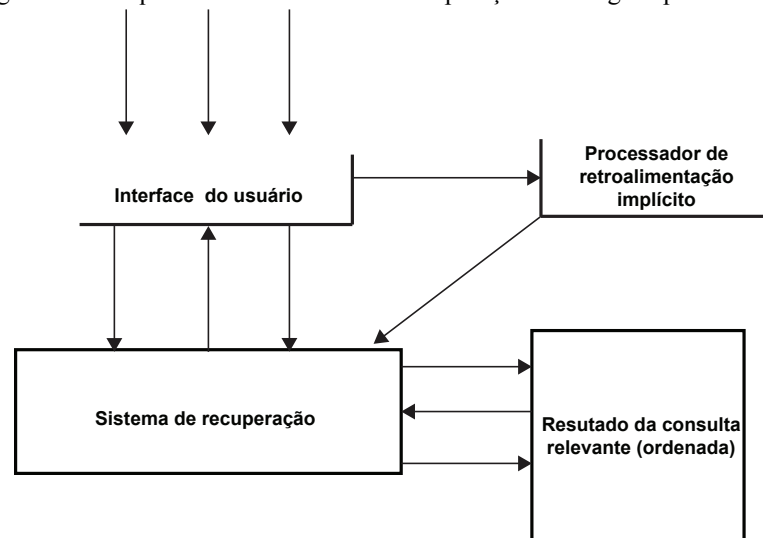
⁷¹ Tradução nossa para “The developers of the first recommender system, Tapestry, coined the phrase “collaborative filtering” and several others have adopted it. We prefer the more general term “recommender system” for two reasons. First, recommenders may not explicitly collaborate with recipients, who may be unknown to each other. Second, recommendations may suggest particularly interesting items, in addition to indicating those that should be filtered out” (RESNICK e VARIAN, 1997, p. 56).

recomendar conteúdo. Na arquitetura do sistema de recuperação de imagens para o iPad-2 proposta pelos autores apresentada na Figura 22, existe um processador de retroalimentação implícito que utiliza os dados de uso para aprimorar o sistema de recuperação, como o tempo que um determinado documento é exibido na tela do dispositivo, o que pode sugerir maior ou menor interesse por ele.

Retroalimentação implícita está relacionada ao comportamento, ação ou reação do usuário, visível ou invisível ou ambas. Por exemplo, se um assistente clica em um determinado documento, isto implica o seu interesse no documento clicado. Se a duração de exibição de alguns documentos é mais longa do que de outros, isto significa que ele pode ser mais interessante no conteúdo.⁷² (RESNICK e VARIAN, 1997, p. 21)

As ações explícitas e visíveis ao usuário como o toque na tela do *tablet* para a abertura de uma imagem, que é o caso do estudo dos autores em questão, e as implícitas e invisíveis aos olhos dele, como tempo de exibição, alimentam a base de armazenamento de anotações no processador de retroalimentação implícito.

Figura 22 – Arquitetura do Sistema de Recuperação de Imagens para o iPad-2.



Fonte: (HU, SMEATON e SUN, 2012, p. 21)

O processador implícito pode aproveitar ainda a tecnologia de geolocalização do dispositivo para oferecer conteúdo da região a qual ele está. Esse uso implícito é comum nos

⁷² Tradução nossa para “Implicit feedback is related to the behavior, action or reaction of the user, either visible or invisible or both. For example if an assessor clicks a certain document, this implies their interest in the clicked document. If the duration of viewing some documents is longer than others, this means that he might be more interested in the content” (RESNICK e VARIAN, 1997, p. 21).

aplicativos de notícias e no sistema de buscas Google, que utilizam a localização do usuário para alimentar seu sistema de recuperação.

Há duas importantes diferenças entre a retroalimentação implícita e a explícita apresentadas por Hu, Smeaton e Sun (2012, p. 21). Primeiro, “o usuário não está avaliando a relevância para o sistema de recuperação de informação, mas apenas satisfazendo suas próprias necessidades”⁷³. Segundo, “O usuário não é necessariamente informado que seu comportamento será utilizado como retroalimentação relevante”⁷⁴. Portanto, um sistema de notícias que utiliza implicitamente a geolocalização para recomendar conteúdo no Feed de Notícias^{xxii} pode destacar manchetes que não necessariamente são as mais relevantes ao usuário, pois apenas a localização é um critério limitado para filtrar conteúdo. Quando vários critérios implícitos são combinados para recomendar conteúdo, o processador de retroalimentação implícito pode aprimorar os resultados, mas novos problemas também podem surgir a partir do momento em que o usuário fica sob um expressivo número de ações implícitas.

O problema foi evidenciado por Eli Pariser (2012) no livro *O filtro invisível*, onde o autor problematiza o uso excessivo de filtragem implícita no Facebook que esconde dele mensagens políticas de seus amigos com visões contrárias às suas em uma época de eleições. Para ele, se está em uma era na qual a personalização cria filtros invisíveis de conteúdos para os usuários sujeitos aos sistemas de recomendação implícitos em sistemas como o de buscas Google e de sites de redes sociais como o Facebook. Por detrás da página do Facebook opera o algoritmo EdgeRank que é responsável pela personalização do conteúdo que aparecerá no Feed de Notícias do usuário:

O EdgeRank classifica todas as interações ocorridas no site. A matemática é complicada, mas a ideia básica é bastante simples, baseando-se em três fatores. O primeiro é a **afinidade**: quanto mais próxima a nossa amizade com alguém – o que é determinado pelo tempo que passamos interagindo com a pessoa e investigando seu perfil –, maior será a probabilidade de que o Facebook nos mostre suas atualizações. O segundo é o **peso relativo de cada tipo de conteúdo**: atualizações sobre relacionamento, por exemplo, têm peso grande; todos gostam de saber quem está namorando quem (muitos observadores suspeitam que esse peso também seja personalizado: pessoas diferentes dão mais ou menos importância a cada tipo de conteúdo). O terceiro é o **tempo**: itens mais recentes têm mais peso do que postagens mais antigas. (PARISER, 2012, p. 39, grifo nosso)

⁷³ Tradução nossa para “The user is not assessing relevance for the benefit IR system, but only satisfying their own needs” (HU, SMEATON e SUN, 2012, p. 21).

⁷⁴ Tradução nossa para “The user is not necessarily informed that their behavior will be used as relevance feedback” (HU, SMEATON e SUN, 2012, p. 21).

O problema do filtro invisível agrava-se quando os internautas passam a utilizar principalmente esses sistemas para informar-se. Ao retomar-se o exemplo das mensagens relacionadas à política em épocas de eleições, evidencia-se esta situação, pois o usuário recebe cada vez mais conteúdo pela afinidade que tem com os outros amigos da rede. Como a tendência natural das pessoas é relacionar-se mais com aqueles que compartilham de ideias semelhantes, as mensagens recebidas fortalecerão cada vez mais o discurso já conhecido. “Uma vez estabelecido o perfil de interesses, é fácil pegar-se carona nas avaliações de consumo fornecidas pelos outros”⁷⁵ (RESNICK e VARIAN, 1997, p. 57). Ainda mais perigoso é estar sob um sistema que implicitamente relativiza o conteúdo por um algoritmo pré-estabelecido, visto que ele pode ser manipulado para atender os interesses políticos e econômicos dos seus desenvolvedores. “Se qualquer um pode fornecer recomendações, proprietários de conteúdo podem gerar montanhas de recomendações positivas para seus próprios materiais e recomendações negativas para seus competidores”⁷⁶ (RESNICK e VARIAN, 1997, p. 57).

A Amazon, considerada por muitos internautas como a livraria online que prioriza as avaliações e recomendações de seus clientes, foi alvo em 1998 da denúncia de repórteres que descobriram que a empresa cobrava 10 mil dólares de editores que desejassem destacar livros na página principal com etiquetas como “O novo e Notável” e “Destinado à fama”. Os títulos patrocinados recebiam “forte suporte editorial da Amazon, incluindo perfil ou entrevista com o autor” (BRANDT, 2011, p. 13). Em entrevista concedida ao jornalista David Streitfeld do Washington Post, Bezos afirmou que:

Se a Editora X representa para nós uma negociação melhor [de um novo título] do que a Editora Y”, dizia Bezos, “e se prevemos que o cliente vai gostar de ambos os livros igualmente, mas só existe espaço para mostrar um, nós vamos mostrar aquele com o qual ganhamos mais dinheiro. (*apud* BRANDT, 2011, p. 13-14)

A Amazon é uma empresa comercial e destacar produtos mais rentáveis é certamente a escolha certa a fazer quando se visa o lucro. O problema é passar ao cliente a impressão de que o sistema é isento e de que se baseia apenas na experiência de consumo e de recomendações do próprio usuário.

⁷⁵ Tradução nossa para “once one has established a profile of interests, it is easy to free ride by consuming evaluations provided by others.” (RESNICK e VARIAN, 1997, p. 57).

⁷⁶ Tradução nossa para “if anyone can provide recommendations, content owners may generate mountains of positive recommendations for their own materials and negative recommendations for their competitors” (RESNICK e VARIAN, 1997, p. 57).

A passividade do internauta aos sistemas de recomendação o coloca em situações em que interesses comerciais podem prevalecer sobre os resultados dos algoritmos. Mesmo quando opera sem manipulação, a própria arquitetura dos sistemas de personalização indica conteúdo que pode deixar o internauta passivo à escolha de informações, se não buscar em outras fontes para ter direito ao benefício da dúvida.

A melhor maneira de evitar o sobreajuste é tentar provar que o modelo está errado, construindo algoritmos que gerem o benefício da dúvida. Se o Netflix me mostra uma comédia romântica e eu gostar do filme, vai me mostrar outra e começar a pensar em mim como um fã de comédias românticas. Mas se quiser ter uma boa imagem de quem eu realmente sou, deveria testar constantemente essa hipótese, mostrando-me *Blade Runner* na tentativa de provar que a hipótese estava errada. De outra forma, eu poderia acabar preso num máximo local povoado por Hugh Grant e Julia Roberts. (PARISER, 2012, p. 121)

Como em todo sistema automatizado, deve-se sempre lembrar que as respostas, por mais que se mostrem interessantes – e realmente são, pois se baseiam no perfil do usuário – não necessariamente serão as melhores escolhas. O conteúdo à margem pode ser o que realmente interessa e a exceção, excluída pelos padrões do sistema, pode ser o item a ser recuperado. É preciso lembrar que: “Os modelos estatísticos que geram a bolha dos filtros descartam as exceções. Entretanto, entre seres humanos, são as exceções que tornam a vida interessante e nos dão inspiração. E as exceções são os primeiros sinais de mudança” (PARISER, 2012, p. 121). Como já citado: a filtragem torna-se melhor quando envolve humanos. Essa sentença precisa ser expandida e incluir a palavra diversidade à frente de humanos. Se as recomendações do Facebook são baseadas na afinidade entre os usuários, aconselha-se que o usuário interaja muito mais com quem não concorda para abrir espaço a conteúdos que provavelmente ele detestaria, mas que podem modificá-lo. Da mesma forma, aconselha-se que nos sites de recomendações de músicas o usuário marque trilhas que ele provavelmente nunca ouviria, nos sites de livros, leia autores que detesta, nos sites de filmes, assista o vencedor do Oscar, da Framboesa de Ouro e os alternativos franceses e alemães. Pariser (2012, p. 207) faz uma proposta interessante neste sentido:

[...] Google ou Facebook poderiam incluir um botão deslizante que corresse de “só coisas que eu curto” a “coisas que outras pessoas curtem e que eu provavelmente vou detestar” no topo dos resultados de buscas e do Feed de Notícias, permitindo assim que os próprios usuários determinassem seu próprio equilíbrio entre a personalização estrita e um fluxo de informações mais variado. Esse método teria duas vantagens: deixaria claro que está sendo utilizado um processo de personalização e daria ao usuário maior controle sobre ele.

Ressalta-se que a informação-como-conhecimento é alcançada apenas quando o receptor do sinal é modificado. Aqueles conteúdos que nada modificam o receptor são vazios de informação. Talvez essa observação explique o fenômeno de internautas passarem horas em sites como o Facebook e após o uso, ao serem interpelados sobre o que leram, não saberem informar nada de concreto. A partir deste óbvio estereótipo, deseja-se apenas salientar os problemas que sistemas sujeitos a filtros invisíveis podem causar. A intenção está longe de criticar o uso desses sistemas, apenas defende-se que a diversidade de fontes humanas é necessária no processo de escolha de conteúdo. No próximo item, ao classificar e esboçar a anatomia dos sistemas de recomendação, a participação humana, será retomado o processo de escolha de conteúdo.

2.4.2 Classificação e Anatomia dos Sistemas de Recomendação

Os sistemas de recomendação são geralmente classificados em três grandes grupos: (a) *recomendações baseadas no conteúdo*, o usuário é recomendado a partir dos itens que ele próprio preferiu no passado; (b) *recomendações colaborativas*, o usuário é recomendado com itens que pessoas com gostos semelhantes preferiram no passado; e, (c) *abordagens híbridas*, quando métodos de recomendações colaborativas e baseadas no conteúdo são combinados para recomendar itens (ADOMAVICIUS e TUZHILIN, 2005; BALABANOVIĆ e SHOHAM, 1997). Visto que este terceiro é um agrupamento das mesmas funcionalidades dos dois primeiros, apenas as técnicas aplicadas nas recomendações baseadas no conteúdo e colaborativas serão aprofundadas.

As recomendações baseadas no conteúdo varrem os textos e os metadados dos documentos para indicar itens. Os sistemas de recuperação de informação, como os geralmente presentes nas bibliotecas, respondem às consultas exclusivamente baseado na representação do conteúdo de seus documentos.

Nesta forma de tratamento dos dados, algumas fraquezas são observadas por Adomavicius e Tuzhilin (2005). A análise de conteúdo é limitada, visto que está sujeita aos recursos explicitamente associados aos dados que os documentos oferecem. O conteúdo do documento deve estar todo presente digitalmente, em formato de texto na base para processamento, e a análise baseia-se apenas nestes dados ou nos atributos inseridos manualmente, como no processo de catalogação dos bibliotecários. Neste segundo aspecto, encontra-se a segunda limitação da análise exclusiva no conteúdo, da sobreposição de descritores que pode impossibilitar a distinção entre os documentos. A terceira fraqueza está

associada aos problemas já discutidos do filtro invisível, quando há nas recomendações o excesso de especialização. O usuário passa a receber recomendações progressivamente semelhantes quando o seu perfil é cada vez mais associado sob a mesma base de dados.

Neste aspecto, observa-se o sistema de recomendação como um agente não humano que influencia no processo de indicação de conteúdo e conseqüentemente na seleção de itens que serão consumidos. Para Latour (2005), agentes humanos e não-humanos estão ligados a uma rede social de elementos que constituem a sociedade. Estes agentes, os quais Latour (2005) denomina de actantes, constituem-se no momento da interação, ou seja, os sistemas de recomendação transformam-se em atuantes ao longo do processo interativo de seleção de itens. Portanto, o sistema de recomendação da Amazon pode ser considerado um mediador – actante – no processo de seleção de livros.

Extrapolando-se hipoteticamente o problema na seguinte sentença: Quanto maior é a busca por livros de culinária francesa na Amazon, menor é a chance de receber indicações de mecânica, informática ou física. Neste exemplo, o sistema torna-se um actante que se auto constitui a partir das interações no processo de seleção de itens. A concomitância de papéis que as pessoas desempenham na vida cotidiana também agrava este tipo de situação, como no seguinte questionamento: Como recomendar literatura acadêmica sobre nutrição para a investigação de um nutricionista que frequentemente pesquisa sobre culinária, e não receber inúmeros itens de autoajuda e dietas milagrosas? O periódico cruzamento da experiência de navegação com o perfil do usuário baseado apenas no conteúdo da base pode criar um ciclo vicioso e causar a excessiva personalização. A última vulnerabilidade está relacionada com a chegada de novos usuários ao sistema. O sistema precisa aprender com as interações do usuário para conhecer suas preferências. Nas primeiras interações estas anotações no perfil de preferências do usuário não existem. Desse modo, o sistema não é capaz de fazer recomendações precisas em relação ao perfil do usuário.

Os métodos de *recomendações colaborativas* procuram prever a utilidade de um item para um usuário particular baseado nas avaliações de outros usuários. Ele não está mais limitado apenas ao conteúdo da base de dados e às interações do usuário em questão, mas é alimentado periodicamente pelas anotações feitas por todos os usuários do sistema ao avaliarem conteúdo. “Em um aplicativo de recomendação de filmes, no intuito de recomendar filmes a um usuário *c*, o sistema de recomendação colaborativo tenta encontrar os ‘pares’ do

usuário c , ou seja, outros usuários com gostos semelhantes em filmes”⁷⁷ (ADOMAVICIUS e TUZHILIN, 2005, p. 735-736). Definido os pares, apenas os filmes melhores avaliados por eles passam a ser recomendados ao usuário em questão. O primeiro sistema de recomendação a propor o uso de estereótipos para definir padrões de usuários foi o Grundy (RICH, 1998). No sistema foram aproveitados os estereótipos dos usuários para agrupá-los e depois recomendar romances para eles. Os algoritmos de filtragem colaborativa estão presentes na maior parte dos sistemas de recomendação de lojas virtuais como a Amazon.

Existem dois grupos de métodos de algoritmos para a filtragem colaborativa. O primeiro método é o *algoritmo baseado na memória* onde: “a tarefa na filtragem colaborativa é prever os votos de um usuário particular (vamos nos referir a ele como o usuário ativo) a partir da base de dados dos votos de usuários de uma amostra ou população de outros usuários”⁷⁸ (BREESE, HECKERMAN e KADIE, 1998, p. 44). O valor da classificação desconhecida $r_{c,s}$ para um usuário c e item s é normalmente calculada a partir da soma das avaliações de algum outro usuário (geralmente o N mais similar) para o mesmo item s . Na fórmula a seguir, \hat{C} denota um conjunto de usuários N que é o mais similar ao usuário c e que avaliaram o item s .

$$r_{c,s} = \text{aggr}_{c' \in \hat{C}} r_{c',s}$$

Resumidamente, a ordenação dos dados apresentados ao usuário ativo respeita a memória de avaliações de outros usuários similares. Para a tarefa de calcular a similaridade são utilizadas duas técnicas: *correlação* e *vetor de similaridade* (ADOMAVICIUS e TUZHILIN, 2005; BREESE, HECKERMAN e KADIE, 1998). Na primeira, o coeficiente de correlação de Pearson é utilizado para medir o grau de correlação entre duas variáveis. A correlação entre os usuários a e i será conhecido na fórmula a seguir, onde as somas sobre j representam todos os itens em que os usuários a e i gravaram votos.

$$(a, i) = \frac{\sum_j (v_{a,j} - \bar{v}_a)(v_{i,j} - \bar{v}_i)}{\sqrt{\sum_j (v_{a,j} - \bar{v}_a)^2 \sum_j (v_{i,j} - \bar{v}_i)^2}}$$

⁷⁷ Tradução nossa para “in a movie recommendation application, in order to recommend movies to user c , the collaborative recommender system tries to find the ‘peers’ of user c , i.e., other users that have similar tastes in movies” (ADOMAVICIUS e TUZHILIN, 2005, p. 735-736).

⁷⁸ Tradução nossa para “the task in collaborative filtering is to predict the votes of a particular user (we will refer to this user as the active user) from a database of user votes from a sample or population of other users” (BREESE, HECKERMAN e KADIE, 1998, p. 44).

O resultado da estatística descritiva nesta fórmula será um valor entre 1 e -1, em que 1 significa perfeita correlação positiva, -1 correlação totalmente negativa e 0 que as duas variáveis são independentes linearmente. Limitações e potencialidades serão abordadas em seguida à explicação do segundo modelo para cálculo da similaridade, conheça-se antes a segunda técnica algorítmica para definição da similaridade ainda no modelo baseado na memória.

A técnica de *vetor de similaridade* é a mesma utilizada no campo da recuperação de informação, onde a similaridade pode ser medida tratando cada documento como um vetor de frequências de palavras e calculando o cosseno do ângulo formado pelos dois vetores de frequência. Esta técnica, segundo Breese, Heckerman e Kadie (1998), pode ser adotada para os filtros colaborativos, quando os usuários tomam o papel dos documentos, títulos tomam o lugar das palavras, e os votos fazem o papel das frequências de palavras. Os pesos de relevância são tomados a partir da seguinte fórmula:

$$w(a, i) = \sum_j \frac{v_{a,j}}{\sqrt{\sum_{k \in I_a} v_{a,k}^2}} \frac{v_{i,j}}{\sqrt{\sum_{k \in I_{ai}} v_{i,k}^2}}$$

Neste caso, os votos observados indicam uma preferência positiva neste algoritmo, não há lugar para votos negativos, e os itens não avaliados recebem o voto zero. Os denominadores ao quadrado servem para normalizar votos.

O segundo método de algoritmo para filtragem colaborativa é o *baseado no modelo*, onde em contraste ao modelo anterior não há um intervalo pré-definido para os valores dos votos. Neste método o algoritmo aprende com a coleção de anotações de votos dos usuários. Assume-se que os votos tomam um valor inteiro no intervalo de 0 a m e chegue-se a seguinte fórmula:

$$P_{a,j} = \sum_{i=0}^m \Pr\{v_{a,j} = i | v_{a,k}, k \in I_a\} i$$

Estima-se neste algoritmo a probabilidade do voto que um usuário c vai dar para um item s . Para estimar esta probabilidade são propostos dois modelos: *modelo de clusters* e *redes bayesianas*. No primeiro, os usuários que possuem pensamentos iguais são agrupados em grupos. A partir disso é feita a associação das classes de usuários e suas classificações são consideradas independentes. A partir disso, o número das classes e os parâmetros são aprendidos a partir dos dados. O segundo modelo adota cada item como um nó em uma rede bayesiana, em que o estado de cada nó para a rede corresponde aos valores possíveis para a

classificação de cada item. “Tanto a estrutura de rede quanto as probabilidades condicionais são aprendidas a partir dos dados”⁷⁹ (ADOMAVICIUS e TUZHILIN, 2005).

Na Tabela 2 pode-se visualizar de forma esquemática cada um dos grupos de sistemas de recomendação e seus respectivos algoritmos e técnicas aplicadas para a indicação de itens.

Tabela 2 – Sistemas de recomendação, algoritmos e técnicas.

Grupos de sistemas	Algoritmo	Técnicas
Recomendações baseadas no conteúdo	Análise de conteúdo	Similaridade do conteúdo dos itens
Recomendações colaborativas	Baseado na memória	Correlação de Pearson Vetor de similaridade
	Baseado no modelo	Escala variável de avaliações (modelo de clusters e redes bayesianas)
Abordagens híbridas	Análise de conteúdo Algoritmo baseado na memória Baseado no modelo	

Fonte: Próprio autor

Mesmo que a explicação matemática pareça complicada, descrever esses modelos e suas técnicas permite observar características nos algoritmos que podem comprometer os resultados de busca.

Um dos problemas da votação é que diferentes usuários podem usar a escala de avaliação diferentemente para avaliar os dados. Por mais que o algoritmo possa ser escrito para atualizar a escala de votos, cada usuário seguirá seus próprios critérios para escolher qual voto fará. Por exemplo, em um sistema que disponibiliza a classificação por estrelas de 1 a 5 para o usuário x a estrela 2 pode representar ruim e para o usuário y pode representar regular. Na outra ponta da escala, o usuário x pode escolher a estrela 5 como bom e o y como muito bom. A própria classificação ruim, regular e muito bom, citada neste exemplo, já é subjetiva. Ainda relacionada a essa limitação, o mesmo usuário pode seguir sua escala subjetiva de forma intermitente, para votar em momentos diferentes. Hoje, ele pode considerar a estrela 4

⁷⁹ Tradução nossa para “Both the structure of the network and the conditional probabilities are learned from the data” (ADOMAVICIUS e TUZHILIN, 2005).

como muito bom, amanhã pode defini-la como regular. Ao tratar os votos matematicamente, essa subjetividade na essência do voto pode oferecer resultados até mesmo controversos.

Outra limitação é que o sistema depende de um número de avaliações sobre um determinado item para começar a calcular a similaridade dele com os demais itens para fazer as indicações. Por um lado, quanto mais avaliações o item recebe, mais variáveis para calcular a similaridade o algoritmo poderá tratar para aprimorar a recomendação. Por outro lado, o elevado número de avaliações torna o resultado sujeito, cada vez mais, a variáveis subjetivas em que cada usuário define a sua própria escala de raciocínio, como abordado no parágrafo anterior.

Por buscar formar um perfil estereotipado do usuário, ainda que matematicamente e respeitando avançados algoritmos da estatística descritiva, pode-se direcionar à excessiva especialização dos resultados e causar a bolha dos filtros. Esse é também o problema evidenciado na abordagem por *cluster*, onde cada usuário pode ser inserido em um único grupo. Então, o problema da diversidade de papéis desempenhados pelos usuários nos diferentes momentos em que usa o sistema e a concomitância de públicos, limitam os resultados de indicações de itens.

No grupo recomendações colaborativas, o problema de novos usuários se repete, como no baseado no conteúdo, visto que sem um perfil do usuário conhecido para alimentar o algoritmo, não haverá dados para calcular e prover indicações. No caso de novos itens, o problema se agrava ainda mais, visto que, pelo fato dos itens estarem dispostos em rede, aqueles mais antigos têm maior potencial de receber votos e aumentar sua popularidade. Como já citado, o problema de links ricos em conexões ficarem cada vez mais ricos foi evidenciado por Barabási (2003). Por fim, os novos itens aumentam a dispersão das avaliações, visto que os sistemas possuem muitos mais itens não avaliados do que aqueles que receberam votos. Por mais que se desenvolvam os algoritmos para superarem essas fraquezas, todas essas limitações são evidenciadas nos sistemas de recomendação de conteúdo contemporâneos. Experimente buscar por *nutrição* no site da Amazon Brasileira⁸⁰ e em seguida por *nutrition* na Amazon Internacional⁸¹. Na loja do Brasil, ainda com uma limitada base de dados de documentos e de anotações, devido à recente chegada da empresa no país, o resultado vai de *Inteligência Alimentar*, como primeira indicação, ao clássico *O Príncipe* como quarta posição.

⁸⁰ Disponível em <<http://amazon.com.br>>

⁸¹ Disponível em <<http://amazon.com>>

O que diria Maquiavel ao ver sua obra do lado do *Guia da Dieta para Aplicações em Excel*? As diferentes técnicas baseadas na filtragem colaborativa e no conteúdo podem ser combinadas para superar essas fraquezas, mas estas se repetirão, visto que o sucesso dos sistemas de recomendações depende, a partir desses algoritmos, da crítica expressiva de uma massa de usuários. O resultado da pesquisa na loja brasileira da Amazon pode ser observado na Figura 23.

Figura 23 – Resultados de pesquisa para a palavra nutrição na Amazon Brasil⁸².

The screenshot displays the Amazon.com.br search results for the keyword "nutrição". The page layout includes a top navigation bar with the Amazon logo, user account information, and a search bar. Below the search bar, there are filters for language (Português), department (Loja Kindle), and price range. The main content area shows a grid of six eBooks with their covers, titles, authors, and prices. The results are ordered by relevance.

Book Title	Author	Price (Kindle Edition)
INTELIGÊNCIA ALIMENTAR: A Nova Refeição e os Ingredientes do Equilíbrio	E. AL. ROOPER	R\$8,90
Um Brinde à Saúde. Sem glúten e nem lactose, POR FAVOR!	Marcelo Facini, Gislene Ribeiro, Angela Annibal e Simone Demolinari	R\$9,99
Nutrição aplicada ao esporte	Monise Viana Abranches	R\$50,94
O Príncipe	Nicolau Maquiavel	R\$0,00
Nutrição	Steven G Carley	R\$1,99
Guia da Dieta para Aplicações em Excel, Volumes 1 e 2	Flavio Morgado	R\$9,00

Fonte: Amazon.com

A participação dos usuários, explícita por votos ou implícita por cliques, é subjetiva para ser apropriada à aplicação em algoritmos que oferecem resultados matemáticos. Ainda assim, observa-se cada vez mais aprimoramento nas indicações que, quando associadas a outras fontes de recomendações, como a simples indicação direta de um professor, pode melhorar o processo de escolha por meio desses algoritmos. O que se destaca nesses modelos estudados até aqui é a importância da geração do perfil do usuário para ser usado pelo algoritmo no mecanismo de recomendação, o qual passa a ser abordado na sequência.

Montaner, López e Rosa (2003, p. 286) classificam os sistemas de recomendação a partir de duas abordagens: (a) *espacial*, “uma classificação dos agentes de acordo com o

⁸² A busca foi realizada em junho de 2013 em <http://www.amazon.com.br>

domínio de aplicação”⁸³; e, (b) *funcional*, “uma classificação dos agentes baseada nas diferentes técnicas para completar tarefas usadas no sistema”⁸⁴. Na abordagem espacial a classificação é feita a partir do domínio do sistema, que podem ser: filtro de notícias, comércio eletrônico, recomendações da web, jornal personalizado, compartilhamento de notícias, recomendações de filmes, recomendações de músicas, recomendações de documentos, recomendações de informação, etc. Na abordagem funcional, a classificação se dá a partir das características do sistema disponibilizadas para a realização das recomendações, que envolvem dois grupos específicos: *geração e manutenção do perfil do usuário*, e *métodos de filtragem da informação*. A abordagem funcional é a que interessa para esta tese, pois vai além da simples classificação do domínio do sistema – os próprios autores não chegam a problematizá-la – e descreve a anatomia do ponto chave de todos os sistemas de recomendação: as técnicas para geração do perfil do usuário e as técnicas para realizar as tarefas de indicação de conteúdo para o usuário quando personificado.

Nesta anatomia, a partir do olhar funcional dos sistemas de recomendação, oito dimensões são observadas: cinco relacionadas à *geração e manutenção do perfil do usuário* que são, na sequência, (1) técnica de representação do perfil, (2) técnica de criação do perfil inicial, (3) fonte de retroalimentação relevante, (4) técnica de aprendizagem do perfil, e (5) técnica de adaptação do perfil. As outras três dizem respeito aos *métodos de filtragem de informação* utilizados pelo sistema, que são (6) filtragem demográfica, (7) filtragem baseada no conteúdo, e (8) filtragem colaborativa.

A Tabela 3 esquematiza essas oito dimensões da taxonomia dos agentes de recomendação investigadas por Montaner, López e Rosa (2003, p. 290):

Tabela 3 – Taxonomia dos Agentes de Recomendação.

TAXONOMIA DOS AGENTES DE RECOMENDAÇÃO	
<i>Geração e manutenção do perfil do usuário</i>	<i>Métodos de filtragem de informação</i>
Técnica de representação do perfil	Filtragem demográfica
Técnica de criação do perfil inicial	Filtragem baseada no conteúdo
Fonte de retroalimentação relevante	Filtragem colaborativa
Técnica de aprendizagem do perfil	
Técnica de adaptação do perfil	

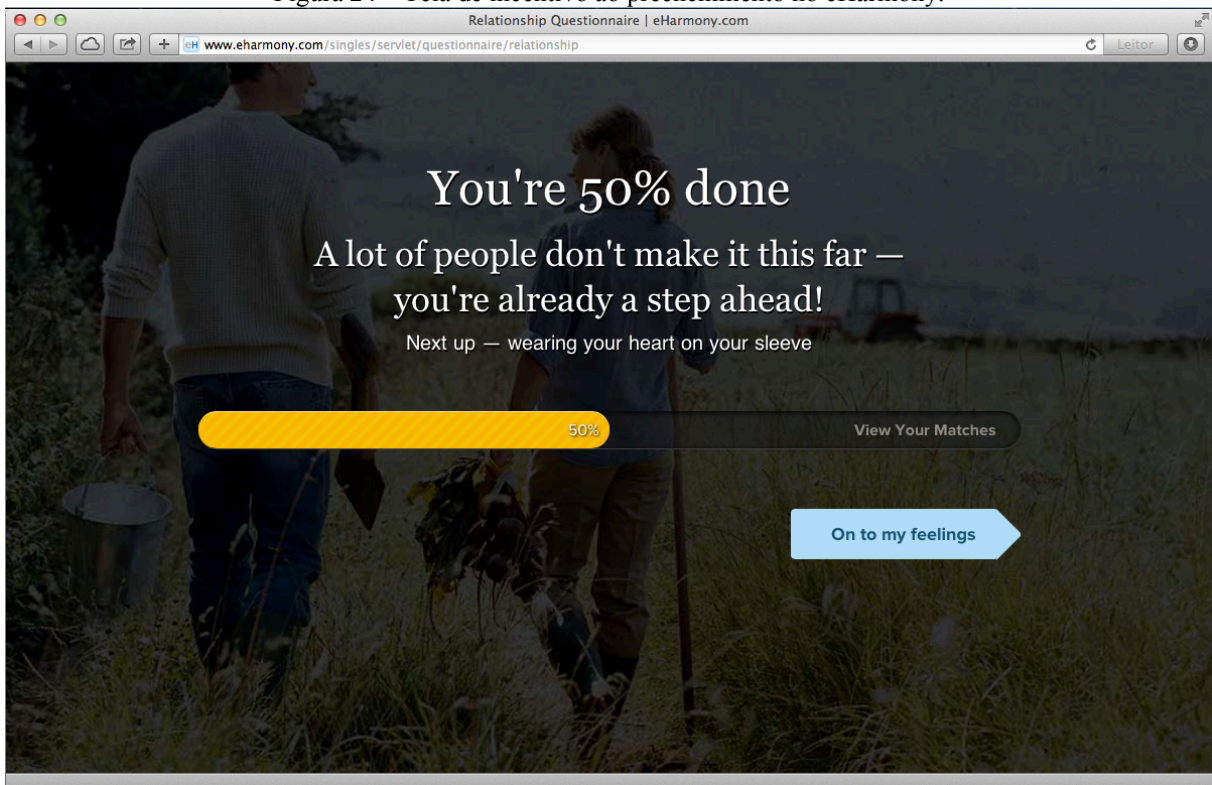
Fonte: (MONTANER, LÓPEZ e ROSA, 2003, p. 290)

⁸³ Tradução nossa para “a classification of agents according to the application domain” (MONTANER, LÓPEZ e ROSA, 2003, p. 286).

⁸⁴ Tradução nossa para “The functional approach produces a classification based on the different task-achievement techniques used in the system” (MONTANER, LÓPEZ e ROSA, 2003, p. 286).

A técnica de representação é o primeiro passo a ser tomado para iniciar a geração do perfil do usuário. O sistema precisa saber o máximo do usuário para dar início a esta geração. Esta técnica – quando bem aplicada – possibilita o sistema oferecer indicações relevantes desde a finalização desta etapa. A partir desta representação, será criado o perfil inicial do usuário que definirá as técnicas seguintes. O usuário interage com o sistema e fornece dados com indícios para a geração do seu perfil desde o início. Esta técnica pode se dar a partir do simples preenchimento de um cadastro com poucas informações até formulários que podem durar vários minutos. A escolha de um formulário mais curto ou extenso dependerá dos objetivos que o usuário espera atingir com o sistema. No sistema de encontros eHarmony⁸⁵, o usuário precisa preencher um extenso questionário que pode levar mais de uma hora para formar seu perfil. O questionário é tão extenso a ponto de aparecerem telas de incentivo ao longo de seu preenchimento para que o usuário não desista da tarefa. A tela apresentada na Figura 24 a seguir apareceu aproximadamente 30 minutos após o início do preenchimento.

Figura 24 – Tela de incentivo ao preenchimento no eHarmony.



Fonte: eHarmony.com

⁸⁵ Disponível em <<http://www.eharmony.com>>

A partir do estudo de sua personalidade, baseado nos dados fornecidos pelo usuário, o sistema passa a explorar pessoas que combinam com ele. Essa correlação pode atingir expressivos resultados. O eHarmony foi responsável por 5%⁸⁶ dos casamentos nos EUA no ano de 2009.

A partir dessas interações com o usuário, o sistema armazena dados que servem como fonte de retroalimentação relevante. Esta é, certamente, a principal dimensão dos sistemas de recomendação. Como já observado, essa alimentação das anotações na base de dados ocorrem de duas formas: explícita, quando fica evidente ao usuário que ele informa o sistema sobre seus gostos; e implícita, quando o sistema utiliza os rastros de navegação do usuário ao interagir com o sistema. A partir do armazenamento do perfil inicial do usuário e dos dados que determinam as respostas relevantes para ele, o sistema necessita de uma técnica para aprender com os dados de navegação explícitos e implícitos gerados nas interações entre usuário e sistema. Os gostos dos usuários mudam com maior frequência que os programadores dos algoritmos gostariam e, muitas vezes, são inclusive contrários ao que o próprio usuário acreditava. As interações implícitas com o sistema podem oferecer um perfil de preferências para o usuário que ele próprio desconhece. Esse é inclusive um dos grandes potenciais dos sistemas de recomendação – um dos objetivos dos sistemas de recuperação de informação – oferecer respostas mais relevantes para perguntas mais generalistas das quais se desconhece o resultado. A busca em um sistema de recomendação, por tratar de recuperação de informação, é diferente da simples consulta de um item conhecido em uma base de dados, como a busca por uma obra pela qual se conhece o título.

Por esses motivos, os sistemas de recomendação necessitam a partir de todas as interações explícitas e implícitas presentes, atualizar a fonte de retroalimentação relevante. Precisam ainda cruzar os dados dessas interações explícitas e implícitas com os dados do perfil inicial do usuário para adaptar a todo instante este perfil, ou seja, é necessária uma técnica de adaptação do perfil para que o sistema continue a indicar itens continuamente com relevância ao usuário.

Definido o perfil do usuário, torna-se essencial a existência de métodos de filtragem da informação. Como amplamente discutido nesta tese, o montante de dados nas bases de armazenamento é cada vez maior e constantemente alimentadas com mais e mais dados a partir das novas interações. Portanto, técnicas de filtragem precisam ser aplicadas para recomendar conteúdo.

⁸⁶ Fonte: instituto Harris Interactive (2009)

A filtragem demográfica usa descrições das pessoas para aprender a relação entre um único item e o tipo de pessoa que gosta dele. Essa técnica utiliza o mesmo método de estereotipar o usuário de Rich (1998) para formar *clusters* de usuários. Os dados das interações dos usuários são utilizados para fazer uma classificação demográfica a partir das preferências. Geralmente, esses dados são obtidos no formulário de cadastro como apresentado na Figura 25.

Figura 25 – Tela para preenchimento de preferências do usuário no Netflix.

How often do you watch	Never	Sometimes	Often	
Moods				
Absurd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Need some examples?
Adrenaline Rush	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Need some examples?
Chilling	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Need some examples?
Controversial	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Need some examples?
Cynical	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Need some examples?
Dark	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Need some examples?
Deadpan	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Need some examples?
Emotional	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Need some examples?
Exciting	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Need some examples?

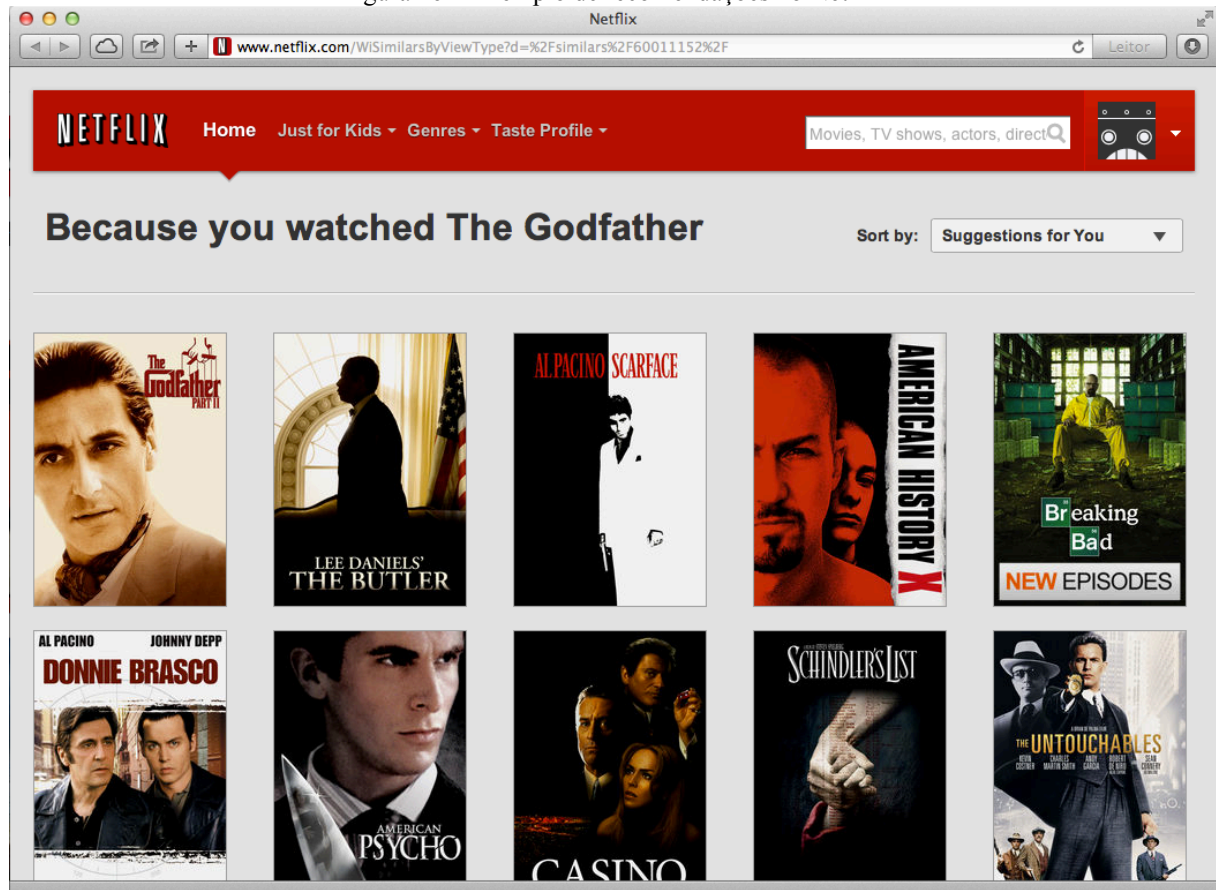
Fonte: Netflix.com

As principais limitações, como já abordadas anteriormente, são a generalização das preferências dos usuários e a falta de adaptação ao perfil individual do usuário.

A filtragem baseada no conteúdo observa as avaliações anteriores dos itens, feita pelo usuário para indicar novos itens. Resumidamente, os itens são indicados pela semelhança de itens preferidos no passado. O perfil do usuário é baseado nos dados extraídos destas interações. Simplificadamente, se um usuário assiste a muitos filmes do gênero Comédia Pastelão no Netflix, a chance dele receber indicações de filmes do diretor e ator Woody Allen é grande. Os dados fornecidos à base de anotações são obtidos por meio da gravação dos rastros de navegação do usuário. Quanto mais filmes de comédia ele assistir, mais indicações

de filmes com atores como Ashton Kutcher e Jennifer Aniston ele receberá. O problema é que, por mais que os sistemas façam boas indicações, é difícil imaginar no momento esses dois atores em um filme de Woody Allen. Meryl Streep e Alec Baldwin são atores que se enquadram no perfil do diretor, porém estes fazem muitos outros filmes que não são nada próximos ao gênero da Comédia.

Figura 26 – Exemplo de recomendações no Netflix



Fonte: Netflix.com

Basear-se apenas nos cruzamentos das preferências, a partir dos rastros de navegação, pode associar itens de forma tão cômica quanto as comédias feitas pelos citados atores. Basear-se apenas em informações objetivas para fazer indicações pode ser um problema, visto que a maior parte das seleções feitas pelos usuários são subjetivas. As escolhas feitas quando se lê notícias e se compra produtos são, muitas vezes, muito mais baseadas em preferências subjetivas do que objetivas. Por fim, outro problema desta abordagem é a já citada excessiva especialização das recomendações. Quanto mais conteúdo do gênero Comédia o usuário assistir, menor será a chance de ele receber indicações de filmes do gênero Drama ou Policial. As preferências não são assim objetivas para indicar conteúdo desta forma.

Por fim, a filtragem colaborativa é aquela, também já discutida neste texto, que cruza as preferências entre perfis semelhantes. As indicações são feitas a partir do padrão das analogias entre as votações de pessoas de perfil similares. Os dados armazenados na base de anotações podem ser fornecidos dessa forma explícita ou implicitamente, a partir do monitoramento do comportamento dos usuários. As fraquezas são: itens e usuários recentes, pois ainda não foram avaliados; dispersão de itens, visto que os itens já avaliados tendem a superar os nunca avaliados; variedade de preferências, já que há usuários com gostos muito dispersos que receberão indicações pobres e, por fim, a dependência da participação, já que se ninguém coopera ou utiliza o sistema, não existirá uma base de armazenamento de anotações.

A forma de tratamento da informação pelos sistemas de recomendação, por mais que ofereçam indicações que podem aprimorar o processo de escolha de conteúdo na Internet, utiliza de algoritmos que tratam a informação de forma objetiva. O contrário nem seria possível, visto que a subjetividade é uma característica, neste aspecto, exclusivamente humana. Mesmo quando as votações e cliques que alimentam os sistemas partem de uma pessoa e que a subjetividade opere nestas ações, o sistema passa a tratá-las objetivamente para indicar conteúdo. A estatística e probabilidade contribuem para indicar conteúdo, porém o seu uso exclusivo pode não ser suficiente para recuperar informação do montante de dados nas bases de armazenamento de conteúdo.

A resposta para os problemas levantados pode partir da lei de Mooers (1996), que afirmou que o usuário desiste de buscar informação quando o sistema tornar a busca muito custosa. Certamente o usuário pode desistir do sistema, mas não necessariamente da informação. Pode ele buscar em uma conversa com outro humano, a partir de interações mútuas, mesmo de forma mediada. A interação mediada por computador será então abordada no capítulo a seguir.

3 INTERAÇÃO MEDIADA POR COMPUTADOR

As interações humanas passam nas últimas décadas pela influência cotidiana de tecnologias digitais, que possibilitam às interações mediadas – sem coincidência espacial/temporal – se aproximarem das interações face a face, ao permitirem que os seres humanos conversem sociavelmente à distância. Como se dá a interação entre os seres humanos de forma mediada por computador é o tema estudado neste capítulo. A intenção neste momento é estudar as interações entre humanos, salientando-se àquelas que possibilitam interações mútuas. É, contudo, importante salientar que nem toda interação entre homens é dessa forma, como no caso de uma troca de links entre dois internautas que nem sempre é considerada como mútua, ou quando se interage com uma atendente de um *call center* que segue um roteiro fechado de respostas para as perguntas de quem fez a ligação. Mesmo que apresente questões relacionadas à interação homem-máquina, explicita-se, desde já, que o interesse está na mutualidade entre as partes, uma atividade cognitiva exclusiva do homem, que recebe extensões quando as tecnologias de comunicação digitais oferecem a possibilidade de se interagir à distância e que é merecedora de maior investigação teórica.

A perspectiva que se segue para a compreensão de como os seres humanos interagem através de sistemas informáticos é a “sistêmico-relacional” de Primo (2007) que divide a interação mediada por computador em “reativa” e “mútua”. Compreende-se que o enfoque dado pelo autor em sua categoria denominada interação mútua, articula-se com a natureza do “interacionismo simbólico” de Blumer (1986), pelo caráter “recursivo sobre a relação e sobre o comportamento dos interagentes” (PRIMO, 2007, p. 228). A interação mútua é a que permite a conversação entre os interagentes, pois, como será apresentado, essa forma de interação possibilita o diálogo necessário para que uma conversa se estabeleça em um ambiente virtual e possíveis recomendações sejam feitas entre pessoas.

No texto que segue, o leitor observará que se busca trazer para a discussão outros conceitos da interação mediada por computador, como aqueles cunhados por Thompson (1998) e Lemos (1998), e outros com enfoques transmissionista ou informacional criticados por Primo (2007). A intenção nesta fundamentação não é a escolha de uma taxonomia da interação, porém é de fazer apontamentos nas teorias já publicadas, a fim de compreender como se insere o diálogo na interação mediada por computador, questão essa importante para se considerar recomendações de conteúdo de forma mediada entre os interagentes.

Na construção teórica em que se apresenta o estabelecimento das interações entre os homens em ambientes virtuais, parte-se do estudo de como as relações sociais se formam e

da importância das interações entre os homens para a constituição dessas relações. Para que a recomendação online se estabeleça de forma mútua, exige-se reciprocidade entre as partes envolvidas. Porém, não necessariamente os interagentes formam relações sociais bem constituídas neste tipo de interação, para isso é necessário historicidade. Faz-se desde já essa ressalva, pois percebe-se que nem toda interação na Internet apresentará diálogo ou tornará o ambiente em um espaço de discussão. Conforme será apresentado, são vários os fatores que incidem em uma interação para que se estabeleça a interação mútua, ao permitir o diálogo entre os interagentes e a formação de relações sociais de forma mediada.

Na interação mútua, as trocas entre os interagentes são recursivas, mas não necessariamente há diálogo. A ação de um professor frente a sua turma de alunos pode apresentar diversas situações de interações mútuas onde não há conversa, como no caso hipotético da recomendação de leituras, em que os alunos ficam descontentes com as indicações, mas que o professor ignora e sai da sala de aula sem dialogar. As interações desse tipo apresentam ações e reações em que as características da interação mútua aparecem, mas não há uma conversa, já que o professor não abre espaço para a fala dos seus interlocutores, mas os modifica ao ignorá-los. As questões de recursividade e mutualidade serão retomadas ao longo desse capítulo, mas antes, se torna necessário atentar-se para como a reciprocidade se apresenta entre os interagentes nos meios de comunicação.

3.1 RECIPROCIDADE ENTRE OS INTERAGENTES

As mídias digitais apresentam diversas formas de interação mediada que aproximam os produtores das informações dos seus receptores e, até mesmo, invertem esses papéis. Porém, a interação mediada não é privilégio apenas da era digital. No tempo do início do rádio já era possível manter contato com os ouvintes através do telefone, como nos próprios jornais impressos que se podia interagir por cartas, mas a interação não ocorria pela mesma mídia. As formas de interação mediada são discutidas na última parte deste capítulo, apenas adiantam-se esses dois exemplos para remeter ao aspecto de reciprocidade entre os meios de comunicação e o público, o que é criticado por Enzensberger (1978) e Brecht (2005)⁸⁷ que, apesar das possibilidades, mostram que não havia reciprocidade entre as partes nesses meios.

⁸⁷ O texto original de Bertold Brecht é do início do século passado e desde aquela época já era percebida a possibilidade de se aproximar o público receptor dos produtores através de uma interação mediada.

Enzensberger (1978, p. 45) alerta que os meios de comunicação não permitiam “qualquer influência recíproca entre o emissor e o receptor; do ponto de vista técnico, reduzem o *feedback* ao nível mínimo possível”. O autor diz que essa relação é refletida pela mesma divisão do trabalho em uma sociedade capitalista, em que o emissor é o dominante e o receptor o dominado. Exemplifica ao mencionar a participação dos eleitores em uma eleição:

A participação pessoal dos eleitores/espectadores é mínima: como nas eleições parlamentares, num sistema bipartidário, o *feedback* fica reduzido às cifras-índices. A manifestação popular reduz-se a um processo simples de três alternativas: escolher o Primeiro Programa, escolher o Segundo Programa, desligar o aparelho (abstenção). (ENZENSBERGER, 1978, p. 45-46)

O autor vê na radiodifusão uma alternativa que vai contra esses distanciamentos entre quem produz e quem recebe a informação. Baseia-se no texto de Brecht que diz que: “é preciso transformar o rádio, convertê-lo de aparelho de distribuição em aparelho de comunicação” (BRECHT, 2005, p. 42).

A radiodifusão poderia ser o mais fantástico meio de comunicação imaginável na vida pública, um imenso sistema de canalização. Quer dizer: isto se não somente fosse capaz de emitir, como também de receber; em outras palavras, se conseguisse que o ouvinte não se limitasse a escutar, mas também falasse, não ficasse isolado, mas relacionado. (BRECHT *apud* ENZENSBERGER, 1978, p. 50)

A possibilidade mencionada por Brecht (2005) no rádio também é vista do mesmo modo, desde o início dos anos 1990, com o surgimento da web. No entanto, é importante perceber que tais possibilidades existem – como já existiam – mesmo com as várias mudanças vistas nos meios digitais, e que a participação do público na construção do conteúdo em sites jornalísticos, por exemplo, é ainda pequena. Mesmo partindo-se da ideia de que, com os elementos de colaboração proporcionados pela Web 2.0, surgem mídias mais democráticas como os weblogs e os sites de avaliação de conteúdo como o Yelp, é necessário fazer essa crítica para não cair em utopias exacerbadas.

As formas de interação apresentadas mais adiante servem para estreitar esse relacionamento do público com a informação. Porém, seria precipitado afirmar que a web é um meio sempre democrático, ao considerar a questão da exclusão digital, cujo problema aumenta ainda mais a barreira entre a coletividade e os meios de comunicação. A partir do momento que importantes informações e processos comunicativos mais avançados se apresentam na web, pessoas que não possuem acesso à rede ficam excluídas por mais essa barreira social que se estabelece.

Mesmo para aqueles que possuem acesso à rede, a crítica de Brecht (2005) e a problematização da reciprocidade são levantadas para que se observe que, mesmo com as possibilidades de interações permitidas pelos ambientes participativos, é necessário mais do que meios técnicos para que ocorram interações dialógicas. Os aplicativos web que se apresentam no cotidiano permitem a produção e publicação de conteúdo coletivo do público de forma mais participativa, nos quais através das interações online se torna possível construir uma mídia mais democrática. No entanto, para que isso ocorra é necessário que as pessoas efetuem trocas que transformem uns aos outros. As recomendações entre os seres humanos no ambiente presencial permitem que essas trocas ocorram a ponto de modificar suas escolhas. Nesse sentido, a interação mútua na web também permitirá que interações mediadas por computador transformem as pessoas e estendam os diálogos e recomendações já existentes para um ambiente mediado, ao oferecer ainda modificações na forma como dialoga-se na contemporaneidade.

A participação e a cooperação são necessárias para a web se tornar mais sociável e as interações podem permitir isso, mas quem fará o meio democrático são as pessoas e não o conteúdo. Por isso, compreender a forma como se interage na web é importante para definir o próprio meio e observar como os seres se relacionam em ambientes virtuais. Desse modo, deve-se pensar desde já que para essas interações ocorrerem é necessária a reciprocidade entre os interagentes envolvidos.

3.2 RELAÇÃO E INTERAÇÃO

A forma como os homens se relacionam possui importante papel na compreensão de suas interações, pois é através das trocas entre os seres humanos que as relações sociais se estabelecem. A comunicação está intimamente ligada às relações sociais. Conville e Rogers (1998) dizem já ser isto um consenso entre os estudiosos nas últimas décadas. Baseando-se em Berger e Kellner os autores afirmam que: “[...] essa ideia uma vez radical de ver os relacionamentos como ‘conversações continuadas’ ganhou aceitação mais ampla, com estudos contemporâneos”⁸⁸ (CONVILLE e ROGERS, 1998, p. 69). Nas conversas se estabelecem as relações sociais.

⁸⁸ Tradução nossa para “[...] this once radical ideal of viewing relationships as ‘ongoing conversations’ has gained wider acceptance, with contemporary studies” (CONVILLE e ROGERS, 1998, p. 69).

Os relacionamentos se dão através de formas de comportamentos entre seus participantes. Os “relacionamentos são vistos como se movendo [...] criativamente formados pelo comportamento interativo dos participantes”⁸⁹ (CONVILLE e ROGERS, 1998, p. 70). A interação determinará como os homens se relacionam. Cada homem tem uma forma de interagir com o mundo e consigo mesmo. No entanto, os relacionamentos constroem-se quando ele interage com os outros. Para Gottman, “[...] um relacionamento consiste de formas temporais que são criadas quando duas pessoas estão juntas”⁹⁰ (*apud* CONVILLE e ROGERS, 1998, p. 74). Após a inserção das tecnologias digitais de comunicação no cotidiano social, o estar junto não é mais visto exclusivamente nas interações presenciais. Precisa-se pensar esse aspecto no momento em que aplicativos web disponibilizam espaços que permitem as pessoas se encontrarem em ambientes virtuais. No entanto, esse aspecto temporal ainda é relevante para o estabelecimento de uma interação dialógica. As formas como interagimos determinam como nos relacionamos com os outros. Deve-se considerar que essas relações agora também ocorrem de forma mediada.

Hinde (1979) considera essa questão temporal e a amplia ao mostrar que somente poucas interações entre os homens não são suficientes para os relacionamentos serem construídos. Para estabelecermos uma relação é necessário que haja interação ao longo do tempo. Será no conjunto de interações e trocas entre os interlocutores que as relações serão moldadas. O autor diz que para haver um relacionamento é necessário interação entre mais de uma pessoa, reciprocidade e continuidade.

Ao considerar as recomendações, observa-se que aquelas entre dois interagentes que mantém um relacionamento podem ser mais relevantes daquelas em que eles sequer se conhecem. Portanto, este aspecto da reciprocidade passa a ser importante no tocante às recomendações feitas a partir de interações em que há reciprocidade entre as partes envolvidas. O mínimo de dois participantes na interação, reciprocidade e continuidade na interação também são exigências para que um diálogo se estabeleça. Desse modo, ao que já se percebe, a conversa é uma condição para a constituição de um relacionamento. A conversa – online ou face a face – permite com que os seres humanos se relacionem.

Como o objeto estudado nesta tese são os sistemas de recomendações digitais, e que estas recomendações podem ocorrer a partir de trocas dialógicas, conversas, entre os

⁸⁹ Tradução nossa para “relationships are viewed as moving [...] creatively shaped by the interactive behaviors of the participants” (CONVILLE e ROGERS, 1998, p. 70).

⁹⁰ Tradução nossa para “[...] a relationship consists of the temporal forms that are created when two people are together” (*apud* CONVILLE e ROGERS, 1998, p. 74).

interagentes, estes três aspectos levantados por Hinde (1979) tornam-se importantes. Ao considerar uma interação através de resenhas da Amazon, constata-se que, para haver a constituição de uma conversa e a construção de um relacionamento, é necessário que dois ou mais interlocutores troquem comentários entre eles de forma recíproca e contínua. O simples comentário em uma resenha, sem o retorno e a continuidade de comentários do autor ou de outros comentaristas, não garante o estabelecimento de conversas e de relacionamentos. Conversas e relacionamentos não se estabelecem de forma individual, “residem, não nas participações individuais, mas no coletivo, em formas padronizadas de se relacionar”⁹¹ (CONVILLE e ROGERS, 1998, p. 71).

Da mesma forma que os relacionamentos não residem nas participações individuais, ao discutir o relacionamento para conceituar a interação mútua, Primo (2007, p. 104) diz que “o relacionamento [...] não pode ser visto como aglutinação de traços pessoais, pois trata-se de um processo emergente não-somativo, ou seja, o relacionamento entre os personagens [interagentes] [...] se dá através da inter-relação das ações recíprocas de cada participante”.

O que o autor apresenta nesse aspecto não-somativo⁹² é que uma interação mútua não é a soma de ações individuais de um interagente, ou seja, a interação mútua se estabelece nas ações recíprocas dos dois interagentes envolvidos. Nesse sentido, a ação de um interagente *A* pode gerar uma reação a um interagente *B* e a mesma ação de *A* pode gerar infinitas reações a outros interagentes.

Mesmo ao considerar essa infinidade de respostas a uma ação, as interações entre os homens podem oferecer padrões na coletividade, ou seja, as pessoas tendem a seguir padrões formais, estabelecidos na sociedade. Quando se deseja bom dia a um colega ao chegar ao trabalho, espera-se receber um bom dia como resposta. Esta é uma convenção social estabelecida e adquirida na experiência cotidiana. Nesse mesmo sentido, porém, não mais em tom amigável, um funcionário que chega constantemente atrasado em seu trabalho, mesmo após seu superior ter chamado sua atenção no dia anterior, espera uma reclamação e

⁹¹ Tradução nossa para “[...] resides not in the individual participants but in the collective, patterned forms of interrelating” (CONVILLE e ROGERS, 1998, p. 71).

⁹² A não-somatividade de ações em interações mútuas é uma característica apontada por Primo importante para a compreensão de como as conversas se estabelecem visto que, ao conversar, o homem faz perguntas em que as respostas são inesperadas. Por mais que padrões surjam e se espere por uma resposta para uma pergunta, pode-se sempre surpreender com a reação para uma ação em interações com esse processo não-somativo presente, não é uma questão como nos sistemas informáticos como se/então/senão (if/so/else).

não um bom dia sorridente do seu chefe. Respostas diferentes nessas situações do cotidiano certamente causariam estranhamento por se afastarem dos padrões sociais.

Ao tomar Bateson como referência, Conville e Rogers (1998, p. 71) consideram que esses padrões formais, presentes nas interações, são contínuos ao longo do tempo. “O intercâmbio de nível relacional do significado da mensagem oferece uma descrição interativa caso o conjunto construído de padrões, ao longo do tempo e em combinação, caracterize a natureza do relacionamento”⁹³.

Padrões são formados pela coletividade para se relacionar. Esses padrões apontam o conjunto de comportamentos com características comuns, que oferecem pistas que possibilitam estudar as interações a partir da natureza de suas relações. A natureza das relações, tomada por base para compreender a forma como os seres humanos interagem, está na perspectiva interacionista de que o homem, ao interagir, interpreta a ação do interagente com quem se relaciona baseado nos significados observados em seu mundo, conforme verificada por Blumer (1986) ao cunhar o termo “interacionismo simbólico”.

O interacionismo simbólico é uma perspectiva que estuda a vida social a partir da observação da descrição do comportamento humano baseado no ato social. Blumer (1986) aponta que os homens, em vez de apenas reagirem às ações dos outros, interpretam as ações uns dos outros. Para delinear a natureza do interacionismo simbólico parte de três premissas:

A primeira premissa é que **os seres humanos agem em relação às coisas com base nos significados que o ser humano pode observar em seu mundo** – objetos físicos, tais como árvores ou cadeiras; outros seres humanos, tais como uma mãe ou uma balconista; categorias de humanos, tais como amigos ou inimigos; instituições, como uma escola ou um governo; ideais guias, tais como independência individual ou honestidade; atividades de outros, tais como seus comandos ou pedidos; e tais situações como um encontro individual em sua vida cotidiana. A segunda premissa é que **os significados de tais coisas são derivados de, surgem da, interação social que alguém tem com um companheiro**. A terceira premissa é que **esses significados são manejados, e modificados através de um processo interpretativo usado pelas pessoas ao lidar com as coisas que elas encontram**.⁹⁴ (BLUMER, 1986, p. 2)

⁹³ Tradução nossa para “The interchange of relational level message meaning offers an interactive description if the jointly constructed patterns that, over time and in combination, characterize the nature of the relationship” (CONVILLE e ROGERS (1998, p. 71).

⁹⁴ Tradução nossa para “The first premise is that human beings act toward things on the basis of meanings that the human being may note in his world—physical objects, such as trees or chairs; other human beings, such as a mother or a store clerk; categories of human beings, such as friends or enemies; institutions, as a school or a government; guiding ideals, such as their commands or requests; and such situations as an individual encounters in his daily life. The second premise is that the meaning of such things is derived from, or arises out of, the social interaction that one has with one’s fellows. The third premise is that these meanings are handled in, and

Desse modo, a ação dos homens em suas interações está condicionada à sua experiência observada em seu mundo. A forma como o homem reagirá em uma conversa, como ele responderá aos outros, e a sua conduta, esta condicionada a significados estabelecidos em sua interação social, significados que se modificam para o homem em sua interpretação das ações do outro. “Interacionismo simbólico [...] vê o significado como decorrente do processo de interação entre as pessoas. O significado de uma coisa para uma pessoa se desenvolve a partir dos modos como outras pessoas que agem em direção à pessoa no que diz respeito à coisa”⁹⁵ (BLUMER, 1986, p. 4). O significado é concebido no processo de interação social, na interpretação que os homens tem das ações uns dos outros.

Os significados são vistos por Blumer (1986) como produtos sociais. Eles são constituídos na atividade cognitiva de interação humana. O contexto da interação social formará o significado que o homem percebe dos objetos do mundo, não apenas no sentido lato da palavra, como coisas materiais, mas como processos cognitivos de interpretação. Tais processos que delineiam demarcam o significado das ações nas interações humanas e nas interpretações do ato social, tratando-se de: “[...] criações que se formam na e através da definição das atividades das pessoas como elas interagem”⁹⁶ (BLUMER, 1986, p. 5).

Nesse sentido, a experiência interativa entre os interagentes em ambientes mediados por sistemas informáticos, como os espaços de troca nos sistemas de recomendação digital, merece atenção, visto que observar a interação social nesses espaços e a interpretação que os interagentes possuem uns das ações dos outros, nortearão os significados que se estabelecem como produtos sociais em ambientes virtuais.

Como apresentado até o momento, o contexto da interação e a interpretação da pessoa envolvida na interação geram o significado, mas é importante salientar que esse processo não necessariamente deriva significados posteriores similares.

modified through, an interpretative process used by the person in dealing with the things he encounters” (BLUMER, 1986, p. 2).

⁹⁵ Tradução nossa para “Symbolic interactionism [...] sees meaning as arising in the process of interaction between people. The meaning of a thing for a person grows out of the ways in which other persons act toward the person with regard to the thing” (BLUMER, 1986, p. 4).

⁹⁶ Tradução nossa para “[...] creations that formed in and through the defining activities of people as they interact” (BLUMER, 1986, p. 5).

Embora o significado das coisas é formado no contexto de interação social e é derivado pela pessoa a partir dessa interação, é um erro pensar que o uso do significado de uma pessoa é uma aplicação do conceito de modo derivado. [...] a utilização dos sentidos por uma pessoa em sua ação envolve um processo interpretativo.⁹⁷ (BLUMER, 1986, p. 5)

Ao correlacionar esse aspecto com a perspectiva sistêmico-relacional de Primo (2007), no momento em que a interpretação de um sujeito é baseada em sua experiência e em sua observação do mundo, seria falho afirmar que as interpretações são meramente aglutinação de traços sociais. Assim como um relacionamento é um processo emergente, não somativo, a interpretação do sujeito também emerge dessa forma: “[...] interpretação não deve ser considerada como uma mera aplicação automática de significados estabelecidos, mas como um processo formativo em que os significados são usados e revisados como instrumentos para a orientação e formação de ação”⁹⁸ (BLUMER, 1986, p. 5). Desse modo, o significado não é uma aplicação automática, mas orienta as ações dos interagentes, ou seja, como Conville e Rogers (1998) também consideram, há padrões formados na coletividade que apontam um conjunto de comportamentos.

Na maioria das situações em que as pessoas agem em direção um ao outro, elas tendem em avançar uma compreensão sólida de como agir e de como a outra pessoa age. Eles compartilham significados comuns e pré-estabelecidos do que é esperado na ação dos participantes e, conseqüentemente, cada participante é capaz de guiar seu próprio comportamento por tais significados.⁹⁹ (BLUMER, 1986, p. 17)

As conversas nos sistemas de recomendações podem apontar interpretações que formam significados e orientam as ações dos interagentes. A experiência conversacional dos interagentes nesses ambientes guiará, dessa forma, o comportamento desempenhado por eles. Deve-se verificar que os significados formados nessas interações se dão nas interpretações dos interagentes do grupo em questão, que podem se modificar para outros grupos. No entanto, quando formados, os produtos sociais (significados e sentidos) orientam o comportamento de grupos em processos interativos convencionados, como no exemplo citado

⁹⁷ Tradução nossa para “While the meaning of the things is formed in the context of social interaction and is derived by the person from that interaction, it is a mistake to think that the use of meaning by a person is an application of the meaning so derived. [...] the use of meanings by a person in his action involves an interpretative process” (BLUMER, 1986, p. 5).

⁹⁸ Tradução nossa para “[...] interpretation should be not be regarded as a mere automatic application of established meanings but as a formative process in which meanings are used and revised as instruments for the guidance and formation of action” (BLUMER, 1986, p. 5).

⁹⁹ Tradução nossa para “In most situations in which people act toward one another they have in advance a firm understanding of how to act and of how other people act. They share common and pre-established meanings of what is expected in the action of the participants, and accordingly each participant is able to guide his own behavior by such meanings” (BLUMER, 1986, p. 17).

anteriormente da relação de cumprimento entre duas pessoas (bom dia/bom dia). Mas é importante perceber que é o grupo quem determina as regras e não essas que estabelecem os sentidos do grupo. “É o processo social no grupo que cria e mantêm as regras, não as regras que criam e sustentam a vida do grupo”¹⁰⁰ (BLUMER, 1986, p. 19).

Há outro aspecto na perspectiva interacionista que também pode ser correlacionado à continuidade da interação, anteriormente citada, que diz respeito à experiência interativa que o grupo adquire ao longo do tempo ao estabelecer seus significados.

[...] qualquer instância de ação conjunta, seja recém-formada ou estabelecida em longo prazo, surge necessariamente de um fundo de ações anteriores dos participantes [...] além de que, a nova forma de ação conjunta sempre emerge e está relacionada com um contexto de ação conjunta anterior.¹⁰¹ (BLUMER, 1986, p. 20).

O histórico de interações anteriores é importante para o contexto de ações conjuntas. O comportamento do autor de uma resenha no website da Amazon em relação à provocação de um comentarista pode ser tomado de forma negativa se não houver uma historicidade entre os dois interagentes, por outro lado, se o interagente for um comentarista ativo nas críticas deste resenhista com um histórico de ação conjunta, a provocação pode ser vista até mesmo como brincadeira. Desse modo, o histórico das ações conjuntas é importante para a observação dos significados formados nas ações dos interagentes. A interpretação do interagente está sujeita ao histórico de ações conjuntas que ele mantém com os outros. As formas como as pessoas se comportam e se representam em suas interações estão sujeitas ao histórico de ações conjuntas, como ao ambiente em que as interações ocorrem. Nesse sentido, a representação dos sujeitos nas interações carece de um olhar mais aprofundado para a continuidade dessa fundamentação. Porém, como se dá a representação do *self*?¹⁰²

¹⁰⁰ Tradução nossa para “It is the social process in the group that creates and upholds the rules, not the rules that create and uphold group life” (BLUMER, 1986, p. 19).

¹⁰¹ Tradução nossa para “[...] any instance of joint action, whether newly formed or long established, has necessarily arisen out of a background of previous actions of the participants. [...] thus, the new form of joint action always emerges out of and is connected with a context of previous joint action” (BLUMER, 1986, p. 20).

¹⁰² O objetivo dessa tese não é apresentar como os sujeitos se representam na web, mas a discussão teórica de temas como *self* e outros é necessária para a construção teórica para compreensão da forma de como os internautas se comportam na web. Por esse motivo, o leitor perceberá no decorrer do texto que teorias são estudadas nesse sentido.

O ser humano é visto como um organismo que não apenas responde aos outros no nível não-simbólico, mas como alguém que faz indicações para os outros e interpreta a sua indicação. Ele pode fazer isso, como Mead mostrou de forma tão enfática, apenas em virtude de possuir um "self"¹⁰³. (BLUMER, 1986, p. 12)

Desse modo, o sujeito possui uma imagem de si para os outros e para si próprio, trata-se do *self* privado e do *self* público que também são formados baseados na perspectiva interacionista, ou seja, o *self* público que será percebido do sujeito pelas pessoas com quem ele interage, será instituído em significados constituídos na interação social. Nesse mesmo sentido, o sujeito constituirá o seu *self* privado nos significados formados na interpretação da interação consigo próprio. “Essa noção de si mesmo como um objeto se encaixa na discussão anterior sobre os objetos. Tal como outros objetos, o objeto de si mesmo emerge do processo de interação social em que outras pessoas estão definindo uma pessoa para si mesmo”¹⁰⁴ (BLUMER, 1986, p. 12).

A forma como o eu é representado na vida real foi estudada também por Goffman (1985) na obra *A Representação do Eu na Vida Cotidiana*. O autor sugere que os indivíduos que representam papéis na sociedade aparecem em um *self oficial* e em um *self dramático*. O *self oficial* é a visão que os outros têm do indivíduo a partir do lugar que ele ocupa: “[...] se supõe que um juiz seja seguro e sóbrio, que um piloto na sua cabine tenha sangue frio; que um contador seja preciso e cuidadoso na sua tarefa.” (GOFFMAN, 1985, p. 97). O *self dramático* “[...] corresponde ao ator social (*performer*¹⁰⁵)” (ANDACHT, 2004, p. 140), ou seja, a forma como o indivíduo representa o melhor papel frente aos seus deslizes, analisados por Goffman (1985) a partir da reação de radialistas ao cometerem erros no ar – os *bloopers*¹⁰⁶ – ao avaliar como agiam após seus deslizes. O sujeito reage para manter o papel que está representando e superar o erro frente sua audiência. “Como personagens encenados para uma audiência, não obstante, não podemos estar sujeitos a altos e baixos” (GOFFMAN, 1985, p. 58). Mas a forma como a audiência perceberá o sujeito que representa será

¹⁰³ Tradução nossa para “The human being is seen as an organism that not only responds to others on the non-symbolic level but as one that makes indications to others and interprets their indication. He can do this, as Mead has shown so emphatically, only by virtue of possessing a ‘self’” (BLUMER, 1986, p. 12).

¹⁰⁴ Tradução nossa para “This notion of oneself as an object fits into the earlier discussion of objects. Like other objects, the self-object emerges from the process of social interaction in which other people are defining a person to himself” (BLUMER, 1986, p. 12).

¹⁰⁵ O ator social possui um personagem pré-existente em seu *self oficial*. O *performer* é a entidade humana que executa esse personagem da melhor maneira possível.

¹⁰⁶ *Bloopers* são erros inusitados ocorridos quando indivíduos atuam em um papel social. O termo é usado para erros em filmes, quando o ator comete deslizes ao atuar seu personagem. Deslizes de políticos em pronunciamentos oficiais que geram carga de humor é um exemplo de *blooper* que pode ocorrer no cotidiano a partir desse olhar de Goffman.

significativo na sua interpretação da interação estabelecida. O *self* como é colocado por Goffman (1985) toma também a interação social como meio de estabelecê-lo, seguindo essa perspectiva interacionista.

As resenhas da Amazon são mantidas por pessoas que desejam comentar e criticar produtos, mas também podem ser para defender suas ideias e representar sua imagem no ambiente virtual. Os votos nas resenhas podem servir para que seus autores recebam avaliações dos outros interagentes. As críticas e elogios recebidos nos comentários das resenhas são formas de apresentar o *self* oficial do autor da crítica a ele próprio, mostrar a visão que os outros interagentes possuem dele. Os comentários e votantes servem como retribuição ao autor para se manter engajado em sua produção e publicação de resenhas no ambiente, ao verificar a avaliação que as pessoas possuem do seu eu e do seu posicionamento frente aos itens que aborda. Braga (2008, p. 267) identificou fenômeno semelhante na blogosfera: “O sistema de retribuição de visitas parece ser a força motriz que constitui o circuito interativo em torno dos blogs”. Como a interação com os outros é o que define o *self* do sujeito, as conversas nos comentários não só avaliam a produção do resenhista, como também definem a imagem do próprio sujeito. Recuero (2008, p. 4) observou também na blogosfera as motivações dos autores para manterem seus espaços ativos que são a “*criação de autoridade*” e a “*criação de popularidade*”¹⁰⁷. Nas resenhas da Amazon e em seus respectivos comentários, fenômeno semelhante pode ocorrer e implicar relevantemente nas recomendações feitas nas interações entre os interagentes. Por isso, torna-se importante estudar as formas de interação mediada que influenciam nas indicações de itens nos sistemas de recomendação digital.

3.3 INTERAÇÃO MEDIADA

A interação entre homens é categorizada por Thompson (1998, p. 78) em três formas: face a face; interação mediada e quase mediada (ou quase interação). A face a face é a que os interagentes estão submetidos a uma mesma lógica de espaço e tempo, em que há uma multiplicidade na possibilidade de deixas simbólicas, como expressões faciais e gesticulações. Na interação mediada, os atores não estão mais copresentes, há uma separação dos contextos, em que há a disponibilidade estendida no tempo e no espaço, como no caso de uma conversa por telefone. Nesta forma de interação, as possibilidades de deixas simbólicas são limitadas,

¹⁰⁷ Tradução nossa para “*Creating Authority e Creating Popularity*” RECUERO (2008, p. 4).

já que os interlocutores não se veem, mas ainda aparecem deixas. Nessas duas primeiras formas de interação de Thompson (1998, p. 78) há uma importante característica interativa para a recomendação entre pessoas, que é o diálogo. O autor considera as duas dialógicas, mesmo que na interação mediada sejam mais limitadas. As duas possuem a necessidade de manter a comunicação fluente entre as partes para garantir o diálogo.

Se, por exemplo, a pessoa com quem se fala deixa de manifestar sinais indicativos de que está acompanhando o que se está dizendo (a falta de um sinal dos olhos, a ausência de um afirmativo “sim” ou “um-hum”, etc.), é sinal de que se deve interromper a narrativa e investigar explicitamente (“Está me ouvindo?”) para provocar uma resposta, ou de alguma outra forma subentendida se certificar de que o outro está atento, sem deixar transparecer a dúvida. Na medida que a interação mediada (como uma conversa telefônica) é dialógica, ela também implica a monitorização reflexiva das respostas alheias, embora as deixas e mecanismos simbólicos acessíveis aos participantes sejam geralmente mais restritos do que na interação face a face. (Indicações verbais, como “sim” e “um-hum”, são muito importantes em conversas telefônicas, precisamente por causa da ausência de deixas visuais. (THOMPSON, 1998, p. 89)

Na interação quase mediada os interagentes estão em contextos diferentes e há uma limitação ainda maior de possibilidades das deixas simbólicas. Mas, a principal diferença entre ela e as duas formas anteriores é a questão da sua orientação. Nas formas de interações face a face e mediada, os interlocutores se dirigem um diretamente ao outro, enquanto na quase mediada um interlocutor dirige-se a um número indefinido de receptores, como ocorre nos meios de comunicação de massa. Portanto, essa forma de interação é monológica, não há um retorno direto do receptor ao emissor e vice-versa. Por exemplo, na televisão, em que os apresentadores dos telejornais não sabem a reação das pessoas frente a um fato anunciado no mesmo momento da transmissão, assim como o receptor não tem a possibilidade de interferir naquele mesmo momento na fala do apresentador. Não existe diálogo na relação entre emissor e receptor.

O aspecto do caráter dialógico é o que se deseja apontar como possibilidade nas interações mediadas, nas mídias digitais, para não se cair na terceira forma de interação de Thompson que é unilateral, pois a bidirecionalidade entre as partes é o que permite o diálogo, já que exige trocas entre os dois interagentes. Não trocas de ação-reação, mas sim de diálogo nas duas vias também apontado por Braga (2008, p. 43):

A comunicação humana não é um processo de mão única. A Internet é um meio interativo que possibilita comunicação e *feedback* em dois sentidos, permitindo trocas de mensagens de um-para-um, como no caso dos e-mails; um-para-muitos, como é o caso das *webpages*; muitos-para-muitos, como no caso das listas de discussão, e todos relacionados a uma base de dados comum.

No sentido das novas mídias, Thompson (1998, p. 81) já dizia na primeira década da web que “[...] outras formas de interação podem ser criadas, por exemplo, pelo desenvolvimento de novas tecnologias de comunicação que permitem um maior grau de receptividade”. O autor percebeu a necessidade de outras tipologias para denominar as formas de interação que surgem com as novas tecnologias devido ao maior grau de receptividade apresentado. Nessas novas formas de interação, o autor fala de uma interação de muitos para muitos. Primo (2007, p. 19) faz a mesma observação baseado em Thompson (1998, p. 235) ao citar que as “[...] redes de computadores possibilitam a comunicação de ida-e-volta que não se orienta para outros específicos, mas que é de ‘muitos para muitos’”.

Nessa questão das interações um-um, um-todos, todos-todos, Primo (2007, p. 99) mostra que “[...] ao se diferenciar a interação mediada por computador da comunicação de massa, por sua capacidade de mediar intercâmbios ‘um-um’ e ‘todos-todos’, aproxima-se a primeira da comunicação interpessoal e a segunda da interação em grupos”.

Nos espaços de resenhas, a comunicação interpessoal (um-um) ocorre quando o resenhista interage diretamente com os comentaristas ou um comentarista interage com outro comentarista, como em diálogos diretos entre eles. A interação grupal (todos-todos) surge nas interações em que resenhista e/ou comentaristas se engajam em um processo interativo em que efetuam trocas, o que pode ocorrer na discussão sobre um tema, em que resenhista e comentaristas publicam suas ideias através de comentários. Primo (2007, p. 99) explica que o estudo dessas interações um-um e todos-todos deve levar em conta o aspecto relacional entre as partes e não somente os aspectos homem-máquina. Com isso em mente, vê-se que as duas formas de interação de Primo (2007), mútua e reativa, avançam a tipologia de Thompson (1998) e satisfazem melhor as interações mediadas por computador, destacando-se as dialógicas presentes em ambientes da Web 2.0, como nos espaços de resenhas da Amazon e em seus respectivos comentários.

Como já foi adiantada, a perspectiva sistêmico-relacional de Primo (2007) afirma que uma interação mútua é não-somativa. Assim como no interacionismo simbólico, para se compreender os significados da interação mútua, devem-se considerar as ações de um interagente sobre o outro. Uma interação mútua não pode ser vista como uma soma de ações

individuais. “Entende-se pelo princípio sistêmico de *não-somatividade* que esse tipo de interação é diferente de mera soma de ações ou das características individuais de cada interagente” (PRIMO, 2007, p. 101-102). Esse caráter de soma de ações é previsto pelo autor na interação reativa¹⁰⁸, em que há necessariamente a condição de estímulo-resposta:

[...] as interações mútuas distanciam-se da lógica de causa e efeito – onde a condição antecedente A é suficiente para causar a condição conseqüente B, isto é, “se A, então B” – presentes em sistemas reativos e que sublinha as perspectivas transmissionista e behaviorista (estímulo-resposta). (PRIMO, 2007, p. 106)

A recursividade no processo interativo caracteriza a interação mútua, enquanto a soma de ações define a interação reativa, quando uma ação prevê necessariamente a mesma reação, presentes nos sistemas informáticos em que as respostas estão prontas, hospedadas em uma base de dados previamente programada.

As interações mútuas são decorrentes de relações entre os seres humanos que efetuam trocas recursivas, onde os interagentes definem seu comportamento e suas respostas na experiência do processo interativo: “[...] os participantes em interação mútua, mediados por redes informáticas, vão se transformando em cada interação que se engajam” (PRIMO, 2007, p. 112). Nesse sentido, essa transformação do sujeito, ou seja, os significados que se formam nesse processo interativo, segue a mesma perspectiva interacionista, onde os sentidos são derivados da interação social que uns têm com os outros.

Na interação reativa as respostas estão prontas e podem ser corrigidas se algum erro interromper a interação, como no caso de uma falha em uma linha de um código do sistema, por exemplo, um link quebrado na web que não encontra a página que deveria apontar. No momento em que o programador corrigir o endereço URL no código, a interação pode ter continuidade: “[...] em sistemas reativos, basta apresentar as mesmas variáveis, nas mesmas condições, que elas apresentarão consistentemente os mesmos resultados” (PRIMO, 2007, p. 116). Os cálculos estatísticos e probabilísticos dos algoritmos dos sistemas de recomendações exemplificam essas interações reativas.

Por outro lado, uma ação em uma interação mútua pode gerar reações inesperadas que marcam definitivamente a experiência interativa, como apresenta Primo (2007, p. 115):

¹⁰⁸ Interação mútua e interação reativa não são excludentes, o leitor deste texto deve ter em mente que: “[...] esses dois tipos interativos (mútuos e reativos) não se estabelecem de forma exclusiva. Pode-se pensar em algo como uma *multi-interação*, no sentido que várias serão as interações simultâneas” (PRIMO, 2007, p. 229).

Em interações mútuas [...] um comportamento não pode ser apagado ou retirado. Uma ofensa através de um *e-mail*, por exemplo, é um evento no tempo que não pode ser retirado da evolução da interação. O conflito gerado por aquele texto será trabalhado no curso de novos eventos comunicativos.

A interação mútua, como na perspectiva interacionista, se dá na interpretação que um interagente faz das ações do outro ao estabelecer seus significados, pode-se pensar que uma ação que inferiu o processo interativo seja contornada mutuamente, mas não se espera que seja esquecida. O mesmo fenômeno deve ser pensado no histórico da ação conjunta. Como já foi apresentada, a interação se dá em uma negociação entre as partes em que as ações anteriores agem significativamente nas ações futuras. Nesse sentido, as recomendações podem inclusive ocasionar desequilíbrios na interação.

Enquanto se comunicam, os interagentes promovem uns nos outros constantes desequilíbrios. A própria definição de seu relacionamento, estando em constata negociação, exige contínua reelaboração em vista das desestabilizações. Assim sendo, pode-se dizer que as interações mútuas se complexificam e se desenvolvem diante do próprio desequilíbrio, sendo este um propulsor de novas atualizações. (PRIMO, 2007, p. 121)

Os desequilíbrios entre os interagentes não devem ser percebidos como fator negativo à interação. Uma discussão prevê necessariamente essa desestabilização e pode tomar uma proporção significativa em sua continuidade, já que atualizações serão necessárias, ao estabelecer um verdadeiro debate. Ao pensar em recomendações, os desequilíbrios podem aprimorar as indicações de itens de forma mútua.

Na *interação mútua*, os interagentes reúnem-se em torno de contínuas problematizações. As soluções inventadas são apenas momentâneas, podendo participar de futuras problematizações. A própria relação entre os interagentes é um problema que motiva uma constante negociação. Cada ação expressa tem um impacto recursivo sobre a relação e sobre o comportamento dos interagentes. [...] a interação mútua é um constante vir a ser, que se atualiza através das ações de um interagente em relação à(s) do(s) outro(s), ou seja, não é mera somatória de ações individuais. (PRIMO, 2007, p. 228)

A interação mútua, ao envolver dois (um-um) ou mais participantes (todos-todos), em ambientes virtuais, significa a possibilidade de se estabelecer espaços destinados ao debate público mediado por sistemas informáticos, como o espaço de resenhas nos sites. As interações de Thompson não satisfazem todas as características interativas encontradas na web na contemporaneidade, mas suas observações em relação a perceber a orientação no processo comunicativo, como um-para-um, um-para-muitos e muitos-para-muitos, são importantes para se compreender as possibilidades dos tipos interativos de Primo (2007). Nas

interações mútuas, uma conversa por Skype¹⁰⁹ ou entre dois comentaristas é do tipo um-um, a postagem de um internauta que publica textos em um blog ou em um jornal que não abre caminho para interação com a audiência de forma recíproca, caracteriza o tipo de interação um-para-muitos, enquanto as conversas nos fóruns da web ou nos comentários de resenhas entre vários interagentes são do tipo muitos-para-muitos.

Desse modo, o produtor se aproxima do público através de interações um-um e todos-todos a partir do momento em que o leitor tem a possibilidade de deixar comentários na resenha, interagindo com o produtor e com os outros leitores, em um espaço de interação mútua, que possibilita a recomendação através da conversação.

Essas formas de interação mediada um-um e todos-todos podem ser vistas em uma das formas de interação de Lemos (2002, p. 113-114). O autor identifica três formas: a Mecânico-analógico, com a máquina; a Eletrônico-digital, com o conteúdo; e, a Social, em que os homens passam a interagir de forma mediada nas formas um-um e todos-todos.

Com a entrada das tecnologias digitais, novas formas de circulação da informação surgem. Ao modelo “Um-Todos” dos media tradicionais, opõem-se o modelo “Todos-Todos”, ou seja, uma forma descentralizada e universal (tudo pode ser convertido em “bits” - sons, imagens, textos, vídeo...) de circulação das informações. (LEMOS, 1998, p. 3)

Percebe-se até aqui a possibilidade que o receptor tem, por meio das mídias digitais, de se relacionar com o conteúdo, o que vem a proporcionar a interação social de Lemos (2002) na perspectiva sistêmico-relacional de Primo (2007). Essa interação com o conteúdo está presente no potencial que o internauta possui de colaborar na construção do seu conteúdo da Internet. Por exemplo, na medida em que um veículo de comunicação publica uma notícia e abre espaço para comentários, as pessoas colaboram com suas opiniões ou apresentam fatos novos, que podem ser aproveitados pelo próprio redator ou serem discutidas com outros interagentes. As opiniões dos internautas podem relacionar-se entre si, o que dá origem ao caráter dialógico, onde cada um apresenta seu ponto de vista independente, ao possibilitar até mesmo o surgimento de novos temas diferentes da matéria apresentada pelo repórter. Esses novos fatos apresentados seriam os temas que os interlocutores pensam ser de relevância para serem informados, suprimindo assim o espaço de discussão reivindicado pela sociedade e salientado por Melo (1994). A ligação entre o receptor e o emissor se dá no conjunto de fatos do que está acontecendo. Melo (1994, p. 15) ressalta que o ponto de tensão

¹⁰⁹ Aplicativo para conversas online da Microsoft <<http://www.skype.com/>>

entre os dois está no que: “[...] a coletividade gostaria de saber e o que a instituição jornalística quer fazer saber”. Nesse sentido, surge a participação que os receptores possuem na produção do material publicado, alertada e criticada pela sua falta por Brecht (2005) já no rádio.

Os espaços de interação mútua, vistos desse modo, podem estimular o diálogo, mas é importante notar a questão sistêmico-relacional de Primo (2007), onde percebe-se que apenas existirá diálogo se os interagentes se relacionam entre si, caso contrário não passarão de opiniões soltas frente a um item, o que também pode ser válido para recomendações de itens, mas não ocorre de forma mútua.

[...] além da interatividade técnica de tipo “analógico-mecânica” e da interação social, podemos dizer que os novos mídia digitais vão proporcionar uma nova “qualidade” de interação, ou o que chamamos hoje de “interatividade”: uma interação técnica de tipo “eletrônico-digital”, correspondendo à superação do paradigma “analógico-mecânico”. (LEMOS, 1998, p. 2-3)

Deve-se diferenciar as interações entre homem-máquina e homem-homem. A interação que interessa neste momento é a entre os homens, pois “[...] tanto um clique em um ícone na interface quanto uma conversa na janela de comentários de um *blog* são interações. Portanto, é preciso diferenciá-las qualitativamente” (PRIMO, 2007, p. 113). O diálogo entre as partes é relevante para identificar comportamentos em recomendações em ambientes mediados por computador. Os graus de interatividade sugeridos por Lemos (2002), que originam essas três formas de interação: Mecânico-analógico, com a máquina; a Eletrônico-digital, com o conteúdo; e, a Social, em que os homens interagem de forma mediada nas formas um-um e todos-todos, são úteis no sentido de considerar os aspectos que envolvem a interação Social, que pode ser encontrada nas contribuições em sites de avaliação de conteúdo, como àqueles que destinam espaço para resenhas críticas dos itens.

A partir dos conceitos abordados a respeito da interação mediada por computador, observa-se que a interação em sistemas de recomendação digital pode ocorrer de forma reativa, como no caso das respostas dos algoritmos às perguntas do usuário, bem como de forma mútua quando dois usuários conversam entre si para fazer recomendações. Observa-se que nem toda resenha, mesmo que escrita por humano apresenta-se como uma interação mútua. Pelo contrário, na maior parte do consumo de resenhas, a simples leitura de uma crítica referente a um conteúdo deverá aparecer como reativa, já que não há necessariamente uma modificação entre as partes. Ainda assim, destaca-se a importância dessas interações,

quando mútuas, na recomendação de um conteúdo e o potencial apresentado por essas formas de interação para aprimorar as indicações.

Nesta conjuntura, a forma como as pessoas interagem reativamente com os conteúdos disponíveis na Internet, bem como mantêm interações mútuas umas com as outras, torna-se essencial no processo de recomendação e de consumo de conteúdo na rede. Ademais, “[...] a interação entre consumidores torna-se um elemento crítico na formação da experiência de consumo” (NETO e FLORIDIA, 2008, . 124). Portanto, o potencial interativo das pessoas com o conteúdo e umas com as outras na Internet pode interferir no processo de consumo. Desse modo, torna-se necessário aprofundar como se dá o consumo mediado por computador que é tema do próximo e último tópico deste capítulo.

3.4 CONSUMO MEDIADO POR COMPUTADOR

A definição moderna de consumo remete ao consumo de bens e serviços, cuja “a cultura de consumo tem como premissa a expansão da produção capitalista de mercadorias, que deu origem a uma vasta acumulação de cultura material na forma de bens e locais de compra e consumo” (FEATHERSTONE, 1995, p. 31). Essa perspectiva de definir consumo, ao pensar na produção de bens e serviços no viés mercadológico, está mais ligada ao consumo destinado para satisfazer as necessidades básicas de sobrevivência. Neste consumo, para satisfazer as necessidades básicas, incluem-se as roupas, os alimentos, as moradias, entre outros bens e serviços básicos voltados à sobrevivência das pessoas.

No entanto, no mundo desenvolvido, apenas uma pequena fração de consumo é para a sobrevivência básica. Decisões de consumo frequentemente refletem uma determinação de que o item consumido será de uso de alguma outra forma que não da meramente para permitir a sobrevivência.¹¹⁰ (LIN, 2007, p. 16)

Na perspectiva pós-moderna do consumo, observa-se que “é importante focalizar a questão da proeminência cada vez maior da ‘cultura’ de consumo, e não simplesmente considerar que o consumo deriva inequivocamente da produção” (FEATHERSTONE, 1995, p. 31). A visão exclusivamente mercadológica e funcionalista da definição de consumo foi bem representada pelo fordismo, pois corresponde principalmente ao setor primário de

¹¹⁰ Tradução nossa para “In the developed world, however, only a small fraction of consumption is for basic survival. Decisions to consume often reflect a determination that the item being consumed will be of use in some way other than merely enabling survival” (LIN, 2007, p. 16).

produção. Porém, existem outras formas de consumo que tocam nos bens culturais da sociedade além de suas necessidades básicas. Baseado em Douglas e Isherwood (1980), Featherstone (1995, p. 35) coloca três classes de consumo:

[...] em relação ao consumo de três conjuntos de bens: um conjunto de artigos de consumo geral, correspondente ao setor primário de produção (comida, por exemplo); um conjunto tecnológico, correspondente ao setor secundário de produção (percurso e equipamento capital do consumidor); e um conjunto de informação, correspondente à produção terciária (bens de informação, educação, artes, atividades culturais e de lazer).

O fordismo, representante da administração científica, certamente fortaleceu a perspectiva da economia clássica, na qual o objetivo da produção é o consumo, porém também desenvolveu as oportunidades de consumo das outras classes. “Afirma-se que a expansão da produção capitalista, especialmente depois da gerência científica e do ‘fordismo’, por volta da virada do século, necessitou da construção de novos mercados e da ‘educação’ de novos públicos consumidores por meio da publicidade e da mídia” (FEATHERSTONE, 1995, p. 32). Logo, a mesma perspectiva que impulsionou o conjunto de bens de consumo do setor primário também expandiu e fortaleceu o conjunto de bens do setor terciário, que envolve o consumo da educação e da informação. Na abordagem dos frankfurtianos, nessa perspectiva o consumo não é mais destinado exclusivamente para atender a produção, a mercadoria está “livre para assumir um valor de uso secundário, uma vez que a dominância do valor de troca conseguiu suprimir a memória do valor de uso original dos bens. Assim, as mercadorias ficam livres para adquirir uma ampla variedade de associações e ilusões culturais” (FEATHERSTONE, 1995, p. 32). Tratam-se dos bens culturais, como aqueles da alta cultura, os quais muitas vezes servem para posicionar as pessoas na sociedade, bem como dividi-las.

A publicidade do mercado vestuário não deseja vender roupas para atender demandas das estações climáticas, mas transferir a promessa de beleza, estilo contemporâneo e todos os signos que envolvem o mercado da moda. Ao observar as mercadorias associadas a ilusões culturais fora do seu uso original, em que o objetivo não é satisfazer uma necessidade básica, mas diferenciar as pessoas umas das outras, chega-se ao que Baudrillard (1998) definiu como mercadoria-signo. O consumo de educação e de informação pode ser visto como prática de consumo que remete status social ao indivíduo.

As pessoas informam-se e educam-se tanto com o objetivo de desempenhar o trabalho para sobreviver quanto para se colocar em posições de status social. O consumo de entretenimento evidencia esse caráter do valor de troca de bens culturais que está além do uso. O consumo dos livros da série Harry Potter pelas crianças na contemporaneidade não é

exclusivamente destinado à leitura infantil, mas é também uma forma de situar o leitor infantil em seu grupo. “[...] o consumo é considerado como dado, sendo empregado como sinônimo de qualquer tipo de ação, como fator de mediação das interações” (MONTARDO e ARAÚJO, 2013, p. 473). O consumo da obra é mais um meio de situar a criança frente aos amigos da escola, cujo conhecimento dos personagens pode ser através da leitura dos livros, bem como a partir dos filmes e dos canais de televisão pelo auxílio do marketing.

Essa experiência de consumo evidencia o comportamento pós-moderno frente aos bens culturais de consumo. As experiências contemporâneas de consumo estão frequentemente sujeitas aos significados que as mercadorias materiais ou culturais carregam consigo. “A lógica de consumo pode ser definida como a manipulação de signos”¹¹¹ (BAUDRILLARD, 1998, p. 114).

[...] a forma e o conteúdo do ato de consumir assumiram características antes mais observadas nas relações de consumo de bens simbólicos. A principal implicação dessa constatação é que a formação do valor dos produtos e serviços passou a depender mais do esforço de construção, negociação e troca de modelos de interpretação e significados. E, portanto, os vetores de satisfação das necessidades e desejos dos consumidores se deslocaram para o interior desse processo. (NETO e FLORIDIA, 2008, p. 114)

Nesta abordagem pós-modernista de definir o consumo, a publicidade está a serviço do marketing que utiliza os meios de comunicação de massa como instrumento para manipular as mercadorias-signos. O consumidor utiliza esses meios tanto para se informar dos itens que deseja consumir, quanto para tomar conhecimento de novos significados que esses itens podem assumir. “A lógica do consumo sempre se ampara em arranjos de comunicação social. Os consumidores precisam ter acesso a um nível mínimo de informação para tomar suas decisões” (NETO e FLORIDIA, 2008, p. 114).

Não se deseja aqui discutir a influência dos meios de comunicação de massa nas decisões das pessoas, mas sim apresentar a lógica de consumo contemporânea que tem na informação e em seus meios a manipulação de signos que definem o consumo pós-moderno. Até o surgimento da Internet e das tecnologias de informação e comunicação potencializadas pela Web¹¹² a partir dos anos 1990, os meios de comunicação de massa foram importantes representantes no tratamento e na comunicação dos significados dos bens de consumo. Antes

¹¹¹ Tradução nossa para “The logic of consumption [...] can be defined as a manipulation of signs” (BAUDRILLARD, 1998, p. 114).

¹¹² Para saber mais sobre as implicações sociais e tecnológicas consequentes da história da Internet que teve início nos anos 1960 e da Web a partir dos anos 1990, aconselha-se a leitura do livro *The Internet galaxy* de Manuel CASTELLS (2001).

da Web, a indústria fonográfica era o principal meio de produzir novas mercadorias-signos musicais. As rádios, representantes dos meios de comunicação de massa na lógica de consumo moderna, eram as responsáveis por comunicar as novas músicas aos consumidores. Com o surgimento da Web e de sistemas de compartilhamento de músicas, o mercado da indústria fonográfica foi abalado e precisou se reestruturar para superar os problemas que passou a enfrentar, os quais estão além da pirataria.

Desde a popularização da gravação de áudio, por meio de fitas magnéticas nos anos 1960, a pirataria era possível por meio da gravação de músicas transmitidas pelas rádios. A gravação dessas músicas foi ainda facilitada com as tecnologias digitais contemporâneas. Porém, a ruptura cultural que abalou a indústria fonográfica não está apenas na facilidade de piratear arquivos musicais, mas sim na possibilidade de compartilhá-los na rede. Agora, os consumidores também têm domínio sobre as tecnologias de informação e comunicação (TICs) das novas mercadorias-signos. Essas “[...] TICs modificam as possibilidades de comunicação entre consumidores” (NETO e FLORIDIA, 2008, p. 115). Ademais, os próprios significados dos bens culturais são por vezes determinados por meio da comunicação mediada por computador, em que “atores vinculados a diferentes esferas sociais estão consolidando novas formas de articulação. Essas experiências sugerem que é possível atenuar o isolamento e diminuir o hiato informacional entre os diversos agentes envolvidos nos processos de consumo” (NETO e FLORIDIA, 2008, p. 116). Além de copiar e distribuir músicas, o consumidor pode criticar e recomendar as trilhas consumidas a partir de sites de compartilhamento, como ocorre no site de rede social destinado a recomendação de músicas Last.fm. “Por meio da Internet, as pessoas que fazem parte do seu mercado estão descobrindo e inventando novas formas de conversar. Elas estão falando de sua empresa. Elas estão dizendo umas às outras a verdade, com vozes muito humanas.” (LEVINE *et al.*, 2000, p. 19)

A mesma ruptura causada ao mercado da indústria fonográfica por consequência da comunicação mediada por computador também afetou outros mercados de bens e serviços. As interações mantidas em Websites como o Reclame Aqui¹¹³ e o TripAdvisor¹¹⁴ evidenciam como as TICs modificaram a comunicação entre os consumidores.

¹¹³ Disponível em <<http://www.reclameaqui.com.br>>

¹¹⁴ Disponível em <<http://www.tripadvisor.com.br>>

Evidências mostram que fontes de informação na internet exercem influência sobre as decisões de consumo daqueles que são expostos a elas. [...] As conseqüências que podem advir desse contexto em construção ainda não são inteiramente conhecidas. Entretanto, alguns estudos sugerem tendências com boas chances de consolidação. Entre elas, a aceleração e intensificação dos ciclos de boca-a-boca, o deslocamento de poder em direção aos consumidores, o surgimento de novos focos de inovação com valor competitivo e a interferência das redes de relacionamento entre consumidores no processo de criação e gestão das marcas. (NETO e FLORIDIA, 2008, p. 118-119)

Nessa perspectiva contemporânea de definição do consumo em que a informação tem importante influência no processo de decisão do consumidor, a Internet tem o papel de tornar os consumidores mais ativos na avaliação e na recomendação dos bens que consomem: “com a ajuda das redes online, as expectativas dos consumidores são mais afetadas por informações/recomendações provenientes de fontes impessoais” (NETO e FLORIDIA, 2008, p. 124).

No website da Amazon, as recomendações pessoais providas de interações mútuas (PRIMO, 2007) através das resenhas dos próprios consumidores, bem como as avaliações feitas a partir de interações reativas (PRIMO, 2007) por meio dos votos no Star Rating, apresentam o potencial da comunicação mediada por computador influenciar no processo de decisão do consumidor. Observa-se que: “Além dos aspectos concretos, como critérios técnicos, os produtos e serviços são avaliados por suas dimensões simbólicas” (NETO e FLORIDIA, 2008, p. 123). Neste cenário, torna-se necessário compreender esse processo que leva o consumidor a decidir os itens que compra.

O processo de decisão do consumidor que influencia nas compras das pessoas apresentado nesta tese segue o clássico modelo de Blackwell, Miniard e Engel (2009), o qual apontam sete estágios decisórios. Ao decidirem a compra de um produto, “[...] os consumidores normalmente passam por sete estágios maiores de tomadas de decisão: reconhecimento da necessidade, busca de informações, avaliação de alternativas pré-compra, compra, consumo, avaliação pós-consumo e descarte” (BLACKWELL, MINIARD e ENGEL, 2009, p. 73). À medida que se avança na explanação de cada estágio, serão aprofundados os mais pertinentes a esta pesquisa que intenta observar como os interagentes selecionam conteúdo sob a influência do sistema de recomendação digital da Amazon.

O primeiro estágio diz respeito ao reconhecimento das necessidades ou, como é frequentemente denominado nos estudos de comportamento do consumidor, reconhecimento do problema. Os autores sustentam que: “Os consumidores compram coisas quando acreditam que a habilidade do produto em solucionar problemas vale mais que o custo de comprá-lo,

fazendo assim do reconhecimento de uma necessidade não satisfeita o primeiro passo da venda de um produto” (BLACKWELL, MINIARD e ENGEL, 2009, p. 74). Observa-se neste estágio que a dimensão do problema e a emergência em saná-lo influenciará diretamente no investimento a ser feito. Investiga-se nesta tese este comportamento ao considerar que leituras acadêmicas apresentam problemas mais dispendiosos ao consumidor pós-graduando do que as leituras de lazer. O custo envolvido na seleção de bibliografia para pesquisas está além do investimento financeiro, pois não trata apenas de gastar menos dinheiro, mas sim das consequências que a seleção errada de livros pode causar aos resultados esperados em suas pesquisas. A forma como o consumidor percebe o problema e suas soluções pode sofrer influências que envolvem: “*família, valores, saúde, idade, renda e grupos de referência*” (BLACKWELL, MINIARD e ENGEL, 2009, p. 75). Destaca-se que as pessoas influentes nesse processo, bem como os profissionais de marketing e suas campanhas publicitárias, não podem criar necessidade, “mas podem mostrar como um produto atende a necessidades ou problemas não percebidos anteriormente pelos consumidores” (BLACKWELL, MINIARD e ENGEL, 2009, p. 76). Apesar de não criarem necessidades, percebe-se a considerável influência desses aspectos no que toca o consumo para a solução dos problemas do consumidor. Logo, com a expansão dos processos de comunicação entre os consumidores, potencializada pelas interações mútuas (PRIMO, 2007) na comunicação mediada por computador, observa-se emergentes canais para a recomendação de itens de informação fora dos meios tradicionais de comunicação, como os meios de comunicação de massa. Ademais, as próprias recomendações de amigos e de grupos de referências apresentam expansão nesse processo de decisão do consumidor na condição em que as interações mútuas e reativas multiplicam-se na rede. “Cada vez mais, a cultura do consumidor online é absolutamente pessoal, tendo-se em vista as múltiplas possibilidades de customização dos espaços ocupados e das formas de atuação em sites de redes sociais” (MONTARDO e ARAÚJO, 2013, p. 479).

O próximo estágio do processo de decisão é o mais importante para esta tese, pois trata da busca de informações no viés do consumo para a satisfação das necessidades do consumidor. Reconhecido o problema, o consumidor precisa de informações para avaliar as alternativas no estágio posterior. “A busca pode ser **interna**, recuperando o conhecimento na memória ou, talvez, nas tendências genéticas, ou ela pode ser **externa**, coletando informações entre os pares, familiares e no mercado [...]” (BLACKWELL, MINIARD e ENGEL, 2009, p. 76). O consumidor carrega consigo uma experiência de consumo que influencia nas suas decisões. No mercado editorial de livros, essa busca interna considera a memória de editoras e autores considerados relevantes pelo consumidor. Com relação a busca externa, o consumidor

considerará as informações buscadas com familiares, amigos, especialistas, entre outros meios que servirão como recomendações para a decisão de compra. Neste aspecto, no que diz respeito ao mercado editorial de livros, deve-se investigar a diferença das recomendações entre amigos e especialistas, visto que o consumidor poderá assumir uma hierarquia nas indicações dessas pessoas. À vista disso, a própria busca externa considerará aspectos internos que tocam na memória da qualidade de indicações de diferentes pessoas em experiências anteriores. “A busca [...] representa a ativação motivada de conhecimento armazenado na memória ou na aquisição de informação do ambiente relacionado à satisfação potencial de necessidades” (BLACKWELL, MINIARD e ENGEL, 2009, p. 110).

As fontes de informações nas buscas externas apontadas pelos autores também merecem atenção, visto que consideram questões mercadológicas que podem influenciar na própria produção da informação.

[...] os consumidores buscam uma variedade de fontes para a obtenção das informações de que precisam a fim de tomar decisões de produto com as quais se sintam confortáveis. Essas fontes podem ser categorizadas como (1) dominadas pelo profissional de marketing, ou (2) dominadas por outros profissionais, não mercadológicos [...] Fontes dominadas por outros profissionais não mercadológicos incluem amigos, família, líderes de opinião e mídia. Muitas dessas influências vêm por meio do boca a boca. (BLACKWELL, MINIARD e ENGEL, 2009, p. 77-78)

As recomendações feitas pelos profissionais de marketing naturalmente objetivam incrementar a venda dos produtos que oferecem, enquanto que as não mercadológicas não carregam necessariamente consigo essa intenção. Apesar disso, ao observar esse aspecto mercadológico por detrás das informações do marketing, atenta-se à necessidade de se expandir essas categorias de fontes de informação no processo decisório, visto que mesmo as informações dominadas por profissionais não mercadológicas não estão necessariamente livres de interesses. Indiferente de quem tem mais importância, não se pode considerar no mesmo grupo as informações obtidas por meio de amigos e as obtidas por meio da mídia. Deve-se ainda salientar as informações obtidas em websites da Internet, cujo o próprio uso do site implicará na qualidade da fonte de informação.

Há cada vez mais a busca de informações na Internet. Apesar de algumas buscas na rede poderem levar um longo tempo, outras são muito mais rápidas dependendo do projeto do site. Alguns pesquisadores indicam que se o varejo on-line diminuir o custo de busca de informações de preço, os consumidores se tornarão mais sensíveis ao preço. Outros estudos têm mostrado que alterando o projeto do site, tornando mais fácil a busca de informações e a comparação de qualidade, os consumidores se tornarão menos sensíveis a preço e tenderão mais a comprar produtos de qualidade. **É o uso do site que influencia como os consumidores vão aplicá-lo no processo de decisão de consumo.** (BLACKWELL, MINIARD e ENGEL, 2009, p. 79, grifo nosso).

Conhecer o comportamento de seleção de itens de informação no website da Amazon pode apresentar relevantes indícios que afetam o processo de decisão do consumidor. Ainda em relação à busca de informação, deve-se salientar o empenho assumido pelo consumidor neste estágio que se aproxima da lei de recuperação de informação de Mooers (1996, p. 22): “De acordo com uma perspectiva de **custo versus benefício**, *as pessoas buscam informações relevantes para a tomada de decisão quando os benefícios percebidos da nova informação são maiores que os custos percebidos para obter a informação*” (BLACKWELL, MINIARD e ENGEL, 2009, p. 112).

Após buscar as informações necessárias, o consumidor passa a avaliar as alternativas que melhor responderão ao problema apresentado.

[...] os consumidores buscam respostas para questões como “Quais são as minhas opções?” e “Qual é a melhor entre elas?” [...] Os consumidores comparam o que conhecem sobre diferentes produtos e marcas com o que consideram mais importante e começam a estreitar o campo de alternativas antes de finalmente resolver comprar uma delas. (BLACKWELL, MINIARD e ENGEL, 2009, p. 79)

Com relação a seleção de livros na loja da Amazon, os consumidores poderão utilizar dos filtros disponíveis no sistema do website para ordenar os livros com as informações de interesse, como a data de edição, o preço, os formatos disponíveis, etc. A partir de interações reativas com o conteúdo, o consumidor pode avaliar quais itens são mais relevantes. Caso ainda esteja em dúvida, o consumidor poderá aproveitar das recomendações feitas pelo site baseadas no histórico de navegação dele próprio ou de outros consumidores. Na dúvida da compra de um livro específico, saber que outros consumidores compraram este livro em conjunto a outros considerados relevantes pode oferecer indícios importantes na avaliação. Logo, além dos aspectos individuais que levam à decisão de compra, os aspectos coletivos também são considerados. “A forma pela qual os indivíduos avaliam suas escolhas é

influenciada tanto por diferenças individuais quanto por variáveis ambientais” (BLACKWELL, MINIARD e ENGEL, 2009, p. 79).

Para avaliar as alternativas, o consumidor considera os atributos salientes (preço, confiabilidade e fatores pouco variáveis entre produtos similares) e os atributos determinantes (detalhes como estilo, acabamento e características específicas). Os atributos salientes são os que o consumidor geralmente julga mais importantes no processo de decisão, enquanto que os determinantes são aqueles mais apontados quando os salientes são equivalentes. São os atributos determinantes que indicarão quais marcas os consumidores comprarão. Enquanto os salientes estão mais focados na racionalidade de uso do produto como na perspectiva clássica de consumo, os atributos determinantes têm por base o estilo do bem a ser consumido.

Para o consumo de bens como filmes e livros, ainda que o preço seja relevante, serão detalhes como autor, estilo, editora, personagens, entre outras variáveis subjetivas que poderão ser consideradas mais importantes pelo consumidor. A decisão de compra desses bens é difícil de ser tomada visto que geralmente não são experimentados antes da efetivação da compra.

Os filmes servem como uma boa categoria de pesquisa, uma vez que são bens experienciais difíceis de se avaliar antes de assistir a eles. Apesar de algumas variáveis, como a boca a boca e resenhas críticas, serem identificadas como influenciadores críticos na escolha dos filmes por consumidores, outras variáveis psicológicas como expectativas emocionais e interesse latente do produto também desempenham um importante papel na decisão. (BLACKWELL, MINIARD e ENGEL, 2009, p. 81)

As interações no website da Amazon podem contribuir para salientar os atributos determinantes dos livros oferecidos pela loja. A análise do uso do website permitirá observar a relevância dessas informações no processo de seleção de itens de informação.

O próximo estágio diz respeito a efetivação da compra, que envolve os consumidores em duas fases:

Na primeira, eles escolhem um vendedor em vez de outro (ou qualquer outra forma de varejo como catálogos, vendas eletrônicas anunciadas na TV ou PC, ou vendas diretas). A segunda fase envolve escolhas dentro da loja influenciadas pelo vendedor, pelas vitrines, meios eletrônicos e propagandas no ponto-de-venda. (BLACKWELL, MINIARD e ENGEL, 2009, p. 82)

A primeira fase evidencia um importante aspecto das lojas virtuais, que diz respeito ao fato do consumidor geralmente não contar no meio digital com o auxílio de vendedores para efetuar as suas compras. O consumidor online está sujeito ao que as autoras Rogers, Sharp e Preece (2013) denominam como “ônus da interação”. No virtual, a

responsabilidade de selecionar os itens corretos no sistema, de cadastrar completamente o endereço de entrega e informar a forma e condições do pagamento é do próprio usuário do sistema. Ainda que os sites possibilitem cancelamentos ou devoluções de compras: “Pressionar o botão errado pode resultar em uma experiência frustrante [...] especialmente para usuários de primeira viagem” (ROGERS, SHARP e PREECE, 2013). Este aspecto pode inclusive inibir a compra em situações em que o consumidor abandona o processo com receio de comprar itens que não atendam suas expectativas. Com relação a segunda fase da compra, as recomendações feitas pelo site da Amazon podem ser comparadas àqueles itens que são estrategicamente posicionados nas lojas físicas próximo às filas dos caixas. As recomendações de compras conjuntas de outros consumidores, bem como de itens populares, podem influenciar outras escolhas e levar o consumidor a selecionar outros itens que não tinha intenção de adquirir naquele momento.

Os primeiros quatro estágios do processo de decisão de compra abordados até o momento são os que mais interessam para os objetivos desta tese por envolverem diretamente a seleção dos itens para consumo. Ainda assim, os outros três estágios serão abordados com o intuito de se conhecer e observar todo o processo decisório no que toca o consumo mediado por computador.

O quinto estágio do processo é denominado pelos autores do modelo como consumo. Como abordado anteriormente ao definir-se consumo, o título dado a este estágio é problemático, visto que se refere mais a questão de uso específico do item do que ao consumo que envolveria todo o processo. A própria explicação dos autores para o estágio evidencia que o foco está no uso ao afirmarem que: “O consumo pode ocorrer tanto imediatamente quanto em um momento posterior” (BLACKWELL, MINIARD e ENGEL, 2009, p.83). Nesta citação, observa-se que o consumo se concretizaria apenas após o início do uso. Logo, ao seguir essa lógica, afirmar-se-ia que se não houver uso não há consumo. Ainda que este não seja o foco desta pesquisa, observa-se neste aspecto a necessidade de atualização do título do estágio para atender a definição pós-moderna de consumo. Por esse motivo, prefere-se o termo uso para denominação deste estágio.

O uso de itens comprados na loja da Amazon será naturalmente sempre posterior à compra quando se tratar de livros impressos, porém também poderá se dar imediatamente quando o consumidor selecionar a compra de e-books para entrega no seu Kindle. O uso imediato dos livros apresenta um potencial diferente das lojas físicas que é da leitura iniciar instantaneamente posterior a compra. A entrega e leitura imediatamente a compra podem influenciar na avaliação do próprio livro no site, visto que o leitor pode rapidamente se

frustrar com o item e retornar para criticar por meio das resenhas que a Amazon permite que os consumidores escrevam para avaliar os produtos. Logo, o processo de avaliação pós-consumo, que pela mesma contraposição entre uso e consumo será melhor denominado se chamado pós-uso, também apresenta diferenças significativas na loja da Amazon.

No estágio de avaliação pós-uso o consumidor define o grau de satisfação do bem ao relacioná-lo com o seu desempenho na solução do problema anteriormente reconhecido.

A satisfação acontece quando a performance percebida confirma as expectativas dos consumidores; quando experiências e performance frustram expectativas, a *insatisfação* ocorre. Os resultados são significantes porque os consumidores guardam suas avaliações na memória e se referem a elas em decisões futuras [...] Se o consumidor tiver alto grau de satisfação, as decisões subseqüentes de compra se tornarão muito mais simples e rápidas. (BLACKWELL, MINIARD e ENGEL, 2009, p. 83)

A avaliação pós-uso alimentará a memória do consumidor como fonte para as buscas internas em processos de decisões futuros. A avaliação pode se dar de forma ainda mais sistemática quando o consumidor se propõe a produzir uma resenha para avaliar o produto no próprio site que o adquiriu. As avaliações pós-consumo da Amazon por meio das resenhas são úteis também para as buscas externas dos outros consumidores. Portanto, no estágio de avaliação, observa-se um potencial ambiente externo que poderá ser considerado nas buscas de informações para o processo de decisão do consumidor.

Por fim, no último estágio, após o uso do item é feito o seu descarte. “Os consumidores têm diversas opções, incluindo descarte completo, reciclagem ou revenda” (BLACKWELL, MINIARD e ENGEL, 2009, p. 86). Destaca-se neste ponto a possibilidade de revenda de livros no próprio website da Amazon. Desse modo, ao mesmo tempo que o consumidor tem a possibilidade de descartar seus itens com retorno financeiro, aumenta a base de dados e quantidade de livros disponíveis na Amazon. Com isso, a Amazon passa a contar com a possibilidade de oferecer livros raros, visto que o consumidor pode vender qualquer livro usado que desejar no site, mesmo que o tenha comprado em outras livrarias.

O aporte teórico desta tese iniciou com a definição de informação, a qual foi definida como coisa para permitir seu manuseio digital por meio de dispositivos informáticos. Abordou-se então nos sistemas de informação a recuperação de informação, cujo destaque foi dado aos sistemas de recomendação ao estudar seus aspectos como a filtragem colaborativa e as implicações da retroalimentação explícita e implícita. Passou-se a estudar a interação mediada por computador, quando se abordou os aspectos da reciprocidade entre os interagentes até classificar-se a interação mediada. Por fim, foi abordado o consumo mediado

por computador, em que se partiu da definição pós-moderna de consumo até chegar-se ao processo de decisão do consumidor no viés digital. Com essa fundamentação em mente, passa-se ao quadro metodológico seguido pela presente tese para responder como os interagentes selecionam conteúdo sob a influência dos sistemas de recomendação digital.

4 QUADRO METODOLÓGICO

A estratégia metodológica para responder o questionamento feito nesta tese, que objetiva responder como os interagentes selecionam conteúdo sob a influência dos sistemas de recomendação digital, envolveu métodos de pesquisa qualitativa. “A principal ênfase em métodos qualitativos reside em observar, gravar, e analisar interações reais entre pessoas e entre pessoas e sistemas”¹¹⁵ (LIEBSCHER, 1998, *online*). O problema em questão, no que toca a seleção de itens como abordado nesta tese, quando visto mais acuradamente, remete-se a interações entre usuários de sistemas de recomendação digital, com o próprio sistema, como também entre os próprios usuários. Portanto, o objeto em questão toca nas interações mediadas por computador entre as pessoas e delas com os sistemas. Por esse motivo, as interações humanas, por serem naturalmente subjetivas, para que sejam compreendidas mais amplamente e por lidarem com informações subjetivas são melhores analisadas por técnicas qualitativas de pesquisa. Para Flick (2009, p. 8), a pesquisa qualitativa:

[...] visa a abordar o mundo “lá fora”(e não em contextos especializados de pesquisa, como os laboratórios) e entender, descrever e, às vezes, explicar os fenômenos sociais “de dentro”, de diversas maneiras diferentes: Analisando experiências de indivíduos ou grupos. [...] Examinando interações e comunicações que estejam se desenvolvendo. [...] Investigando documentos (textos, imagens, filmes ou música) ou traços semelhantes de experiências ou interações.

Ainda que a pesquisa qualitativa também investigue documentos, diferentemente da quantitativa que pode lidar com dados reais que podem ser acessados sem uma observação direta do objeto, envolve a observação do campo a ser pesquisado.

Enquanto o pesquisador quantitativo pode coletar dados sem sair da sua instituição, o pesquisador qualitativo frequentemente coleta dados através de observações “no campo”. De fato, uma imersão no ambiente natural dos fenômenos sob observação é muitas vezes um elemento essencial da pesquisa qualitativa. Como Cavan (1966) apontou, se você deseja saber como pessoas socializam em bares, você tem que visitar bares para observá-las.¹¹⁶ (LIEBSCHER, 1998, *online*)

¹¹⁵ Tradução nossa para “The principal emphasis in [...] qualitative methods lies in [...] to observe, record, and analyse real interactions between people and between people and systems” (LIEBSCHER, 1998, *online*).

¹¹⁶ Tradução nossa para “While the quantitative researcher can gather data without leaving her institution, the qualitative researcher most often gathers data through observations “in the field.” Indeed, an immersion in the natural setting of the phenomena under observation is often an essential element of qualitative research. As Cavan (1966) pointed out, if you want to know how people socialize in bars, you have to visit bars to observe them” (LIEBSCHER, 1998, *online*).

Para compreender o comportamento dos usuários de sistemas de recomendação, tornou-se necessário observar de perto o seu uso a fim de coletar dados para investigar quais comportamentos individuais emergiam desse processo interativo. Para observar o comportamento individual, foram coletados dados da experiência de seleção de conteúdo de indivíduos isolados. As técnicas qualitativas de coletas de dados utilizadas nesta pesquisa observaram a experiência individual a partir de observação direta e de entrevista em profundidade. “A entrevista em profundidade é um recurso metodológico que busca, com base em teorias e pressupostos definidos pelo investigador, recolher respostas a partir da experiência subjetiva de uma fonte, selecionada por deter informações que se deseja conhecer” (DUARTE, J., 2005, p. 64). A observação direta “[...] serve para fornecer dados adicionais sobre o tema em análise. Compreende atividades formais como desenvolver protocolos de observação” (DUARTE, M. Y. M., 2005, p. 230).

Desse modo, aplicaram-se as técnicas de pesquisa qualitativa de entrevista em profundidade junto à observação direta do sistema da Amazon, em uso pelos entrevistados no que toca a interação individual destes, com vistas a conhecer o comportamento de cada usuário analisado frente aos tipos de recomendações. Antes dessas duas etapas, fez-se uma leitura flutuante das formas de recomendação e avaliação de itens do website analisado, no intuito de classificar os tipos de recomendações a partir da Matriz de Recomendações de Itens de Informação que será proposta nesta tese, a qual será apresentada na sequência. Essa classificação permitiu tomar-se conhecimento das funcionalidades do sistema que contribui para o conhecimento empírico do mesmo e para as etapas de coleta de dados posteriores.

Após coletados os dados, estes são analisados a partir das categorias da matriz proposta, quando correlaciona-se os resultados das entrevistas em profundidade com as observações diretas para esboçar como se dá a seleção de itens em websites sob influência de sistemas de recomendação digital.

4.1 CLASSIFICAÇÃO DAS RECOMENDAÇÕES

Nesta primeira etapa, as formas de recomendações foram classificadas a partir de categorias definidas a partir do quadro teórico no intuito de reconhecer quais tipos de recomendações envolvem o processo de seleção de itens pelos usuários. Nesta fase, foi feita a análise do site da Amazon para identificar os tipos de recomendações possíveis, de forma mediada. Porém, ressalta-se que também aparecem nas categorias da matriz aquelas recomendações oriundas de interações analógicas, visto que um dos pressupostos desta tese

considera que para a escolha de determinados itens são necessárias consultas além dos sistemas de recomendação digital. As recomendações analógicas foram investigadas na etapa das entrevistas em profundidade.

A matriz a seguir apresenta as categorias que enquadram, como elementos, os tipos de recomendações de itens de informação.

Tabela 4 – Matriz de Recomendações de Itens de Informação.

Recomendações	Objetivas	Subjetivas
Digital	Recomendações estatísticas matemáticas informatizadas baseadas na experiência do usuário.	Recomendações de terceiros baseadas em suas experiências pessoais.
Analógica	Recomendações impressas baseadas em estatísticas, pesquisas e procedimentos.	Recomendações de terceiros baseadas em suas experiências pessoais
Interação	Reativa	Mútua

Fonte: Próprio autor

Na orientação vertical da matriz, observa-se as recomendações a partir da forma como os dados potencialmente informativos são tratados, onde os elementos são dispostos em duas colunas: *Recomendações Objetivas*, as quais ofertam conteúdo a partir de um tratamento matemático dos dados resultante de interações reativas, ou seja, objetivas; e, *Recomendações Subjetivas*, quando os dados são avaliados e indicados a partir de interações mútuas, logo subjetivas.

Na orientação horizontal, percebe-se os sistemas a partir do ambiente em que as interações ocorrem e do seu tipo, em que seus elementos são apresentados em duas linhas: *Recomendações Digitais*, quando as recomendações são feitas a partir de interações digitais com um dispositivo conectado a algum tipo de sistema de recomendação informatizado; e, *Recomendações Analógicas*, quando as indicações de itens de informação ocorrem a partir de interações analógicas ou por dispositivos não conectados a sistemas de recomendação informatizados.

O cruzamento das colunas com as linhas da Matriz de Recomendações resulta em quatro tipos: *Recomendações Objetivas Digitais*, que são aquelas presentes nos sistemas de recomendações, as quais utilizam de algoritmos e cálculos matemáticos para tratar dados e recomendar conteúdos a partir da interface de sistemas informatizados, como as recomendações feitas a partir do histórico de navegação do usuário no website da Amazon onde operam as interações reativas; *Recomendações Objetivas Analógicas*, as quais são resultado de recomendações geralmente impressas e baseadas em informações matemáticas,

estatísticas, técnicas ou previamente definidas, como ocorre nas recomendações a partir dos *Blurb Index* das centrais de atendimento (*call centers*) e também aquelas oriundas de avaliações técnicas de revistas impressas; *Recomendações Subjetivas Digitais*, em que as indicações são feitas a partir de críticas e avaliações de itens geralmente no formato de textos de resenhas ou também por indicações providas de interações entre internautas mantidas em sites de redes sociais ou programas de bate papo; e por fim, *Recomendações Subjetivas Analógicas*, que são todas as indicações oriundas de conversas e discussões com o objetivo de indicar e avaliar itens, as quais podem ser feitas por especialistas, professores, familiares, amigos e até mesmo vendedores ou consultores de lojas e instituições.

Como se observa nos tipos de recomendações, em relação ao website da Amazon, nesta primeira etapa da pesquisa, será possível classificar apenas os elementos que dizem respeito às *Recomendações Digitais*. Os elementos das *Recomendações Analógicas* serão considerados nas etapas posteriores da pesquisa, onde os usuários serão interpelados sobre estes aspectos, porém os elementos das *Recomendações Digitais* também serão exaustivamente abordados nas entrevistas e nas observações diretas, visto que é no comportamento dos usuários que estão os principais objetivos desta tese.

4.2 ENTREVISTA EM PROFUNDIDADE E OBSERVAÇÃO DIRETA

Os internautas selecionados para amostra da pesquisa deveriam, necessariamente, ser usuários do website da Amazon.com para a busca de itens, mas não precisavam ter efetuado qualquer compra no website, pois objetivava-se também avaliar o comportamento daqueles recém chegados ao sistema e ainda dos que o utilizavam apenas para fazer uso da recomendação digital. Não foi feita predefinição de gênero ou escala de idade na seleção dos entrevistados, visto que essa distinção poderia limitar a avaliação do comportamento entre usuários nascidos em uma era de dispositivos digitais e aqueles anteriores a essa geração. “É importante obter informações que possam dar visões e relatos diversificados sobre os mesmos gatos. Pessoas em papéis sociais diferentes, recém-chegadas ou que tenham deixado a função recentemente, podem dar perspectivas e informações bastante úteis” (DUARTE, J., 2005, p. 69). A única limitação de idade feita foi que os entrevistados deveriam necessariamente ser maiores de 18 anos, pois estariam legalmente aptos a participar de pesquisas sem a necessidade da autorização de responsáveis legais. Visto que esta etapa da pesquisa exigia participação presencial para execução da mesma, os usuários deveriam residir

necessariamente na região metropolitana da cidade de Porto Alegre para viabilizar a aplicação das entrevistas junto à observação direta.

A seleção dos entrevistados foi não probabilística e por conveniência. “A seleção dos entrevistados em estudos qualitativos tende a ser não probabilística, ou seja, sua definição depende do julgamento do pesquisador e não de sorteio a partir do universo, que garante igual chance a todos (característica das amostras probabilísticas)” (DUARTE, J., 2005, p. 69). Ainda que houvesse intencionalidade na seleção dos entrevistados, pois deviam fazer parte do grupo de usuários do website da Amazon, não se fez juízo particular para selecionar pessoas com conhecimentos específicos do assunto pesquisado. A conveniência na viabilidade de realização das entrevistas foi por serem presenciais e necessitarem de proximidade geográfica. Os informantes enquadravam-se na categoria padrão, por ser um tipo de “fonte envolvida com o tema de pesquisa, mas que poder substituída por outra sem que se espere prejuízo na qualidade de informações obtidas” (DUARTE, J., 2005, p. 70).

Definidos os primeiros entrevistados, o recorte da amostra para seleção dos entrevistados subsequentes considerou a técnica de “bola de neve” (PATTON, 1990). Esta técnica permite identificar pessoas em meio a uma amostra de difícil acesso ou identificação.

A técnica de seleção por bola de neve consiste em identificar primeiramente alguns entrevistados elegíveis que aceitem¹¹⁷ participar da pesquisa. Conforme cada um dos participantes é entrevistado, solicita-se que seja indicada outra pessoa com perfil semelhante ao dele, adequado à amostra da pesquisa. Especificamente para esta pesquisa, a indicação foi de pessoas maiores de 18 anos que utilizavam o website da Amazon e que eram colegas do entrevistado. Essa prática seguiu até que os dados passassem a se repetir e fossem suficientes para atender aos objetivos da pesquisa. Por esse motivo, não havia um número de entrevistados previamente definidos para participarem da pesquisa. Foram entrevistados 10 pessoas de quatro grupos acadêmicos diferentes, especificamente 4 pesquisadores da Comunicação, 4 pesquisadores da Biologia, 1 pesquisador da Sociologia e 1 pesquisador da Biblioteconomia. Os entrevistados eram 4 homens e 6 mulheres com idade entre 22 e 35 anos.

4.2.1 Entrevista em profundidade

A entrevista seguiu um roteiro composto de perguntas semiestruturadas e cada uma durou aproximadamente 40 minutos. As respostas foram gravadas em áudio e foram

¹¹⁷ A permissão do entrevistado foi formalizada por termo de consentimento conforme o Anexo A.

tomadas notas das deixas simbólicas dos entrevistados que podiam oferecer importantes indícios a análise posterior.

O ambiente para a realização das entrevistas em profundidade e das observações diretas foi previamente preparado para evitar interferências externas que pudessem interromper a dinâmica dos questionamentos e das atividades desempenhadas. O ambiente foi preparado em salas universitárias, como a apresentada na Figura 27.

Figura 27 – Sala de aula preparada para entrevistas em profundidade.



Fonte: Próprio autor

No intuito de deixar o entrevistado à vontade para a realização da entrevista, antes de passar ao roteiro de perguntas, o entrevistador fez uma breve apresentação pessoal e, ao agradecer pela participação espontânea, apresentou o termo de consentimento, falou do tema da pesquisa e disponibilizou-se a enviar os resultados do trabalho quando finalizado. Neste momento, o entrevistado foi informado que os dados oriundos da pesquisa seriam guardados por 3 anos e que posteriormente a esse período os mesmos seriam incinerados. O entrevistado também foi informado que poderia abandonar a entrevista em qualquer momento do processo sem qualquer tipo de prejuízo a ele, bem como poderia omitir-se a responder qualquer uma das questões. Todos os entrevistados responderam todas as perguntas e, até mesmo pelo conteúdo das questões, não se percebeu situações em que os entrevistados hesitariam responder os questionamentos.

Para não induzir respostas dos entrevistados, a apresentação foi breve, e dúvidas em relação à participação das etapas da pesquisa foram respondidas ao participante. Após este

momento, passou-se a gravar o áudio e o roteiro semiestruturado da Tabela 5 a seguir foi aplicado:

Tabela 5 – Roteiro da entrevista em profundidade semiaberta.

Questões semiestruturadas	Tópicos abordados
1) Como você acessa o site da Amazon e qual é a sua primeira ação ao visitar?	(caminho de entrada, primeiras ações, <i>login</i>)
2) Por que você acessa o site da Amazon?	(motivações, frequência de uso, tempo de uso em anos, fazer uso do sistema de recomendação)
3) Como você seleciona itens no Website da Amazon?	(experiência de navegação, uso de sistemas de busca, uso de recomendações)
4) Como você faz compras na loja da Amazon?	(formas de pagamento, cuidados com segurança, é comprador?)
5) Como você escolhe os itens que consome?	(motivações principais e secundárias para seleção, compra de itens de baixo custo e alto custo, tipos de indicações)
6) Como você vê as recomendações feitas na loja da Amazon?	(utilidade, segue recomendações, quais sistemas reconhece, tipos e custo de itens selecionados, sistemas objetivos)
7) O que você pensa das resenhas escritas pelos próprios clientes da Amazon?	(faz uso, utilidade, influência na seleção, interage com resenhistas/comentaristas, vota nas resenhas)
8) Como você faz avaliações de itens no website da Amazon?	(escreve resenhas, vota em resenhas, responde comentários, comenta, vota no star rating, importância das interações)
9) A partir deste cartão, você poderia escrever do lado de cada uma das escalas o que cada nível representa para você?	(O cartão da Figura 28 é entregue ao entrevistado)
10) Por fim, o quanto a Amazon influencia nas suas compras?	

Fonte: Próprio autor



Fonte: Próprio autor

Releva-se que o roteiro de entrevista serviu como guia e que não necessariamente foram aplicadas todas as questões a todos os entrevistados, bem como novas perguntas

emergiram ao longo dos questionamentos, conforme os tópicos listados em cada questão do roteiro. Essa flexibilidade foi necessária para que o pesquisador estivesse livre para explorar melhor o direcionamento que a entrevista tomaria.

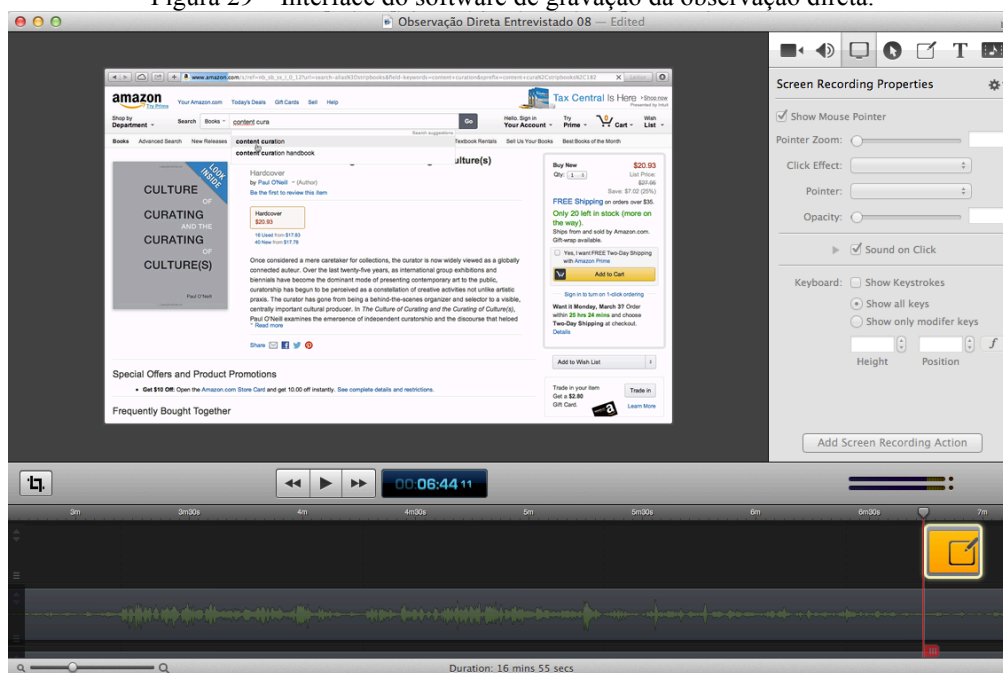
As entrevistas foram transcritas e as respostas foram tabuladas junto às anotações feitas pelo pesquisador ao longo das mesmas. O objetivo dessa tabulação era estruturar os dados de forma a encontrar indícios de padrões entre os usuários que oferecessem resultados para esboçar-se o comportamento dos usuários durante a seleção de itens de informação sob influência de sistemas de recomendação digital. As respostas obtidas nesta etapa também foram aproveitadas na etapa posterior de observação direta.

4.2.2 Observação direta

Nesta etapa, os entrevistados utilizaram o website da Amazon sob observação do pesquisador. Algumas das interações mantidas no website da Amazon pelos usuários, por serem muitas vezes implícitas tornam praticamente impossível de serem descritas pelos próprios usuários. Por isso, optou-se por uma técnica projetiva em que atividades foram induzidas ao usuário com o objetivo de fazer-se a leitura dessas interações e dos comportamentos por vezes desconhecidos pelo próprio usuário. Esta observação seguiu um roteiro de tarefas que foram cumpridas pelos entrevistados, as quais ora induziam a forma de uso do sistema, ora deixava-os livres para utilizar o site como de hábito. A partir do software¹¹⁸ apresentado na Figura 29, o qual permite gravar a tela do computador enquanto é utilizado, toda a experiência interativa ao longo da observação direta foi gravada.

¹¹⁸ O software utilizado foi o ScreenFlow Versão 4.3 para Mac OS

Figura 29 – Interface do software de gravação da observação direta.



Fonte: Próprio autor

Ao longo desta observação, enquanto o entrevistado executava as tarefas passadas, o pesquisador anotava informações relevantes para posterior análise. Toda a experiência captada dizia respeito exclusivamente à busca de itens de livros para delimitar melhor os dados analisados posteriormente.

A seguir, observa-se o roteiro de tarefas executadas no site da Amazon, que foram divididas em dois grupos de atividades: espontâneas e induzidas.

Tabela 6 – Roteiro das atividades da observação direta.

Atividades espontâneas

- 1) Você poderia selecionar um livro que deseja no site da Amazon? (o navegador estará fechado)
- 2) Agora, poderia explicar como navega pelo site e quais são as principais seções de livros e funcionalidades do site que você utiliza?

Atividades induzidas

- 1) Você poderia procurar e selecionar um livro para a sua vida acadêmica ou profissional?
- 2) Agora, você poderia procurar e selecionar um livro para leitura de lazer?
- 3) Você poderia selecionar os livros da sua página de entrada que você compraria?
- 4) Você poderia procurar, avaliar e selecionar um livro de mais de 100 reais que você compraria?
- 5) Você poderia selecionar um livro de seu interesse e escolher mais um item a partir das recomendações correlatas?

Fonte: Próprio autor

As atividades descritas neste roteiro consideram as informações passadas anteriormente na entrevista individual e as atividades induzidas que foram desempenhadas

anteriormente, na época da observação, pelo próprio pesquisador com o objetivo de atualizar as atividades em razão da sazonalidade e intermitência do destaque de alguns itens.

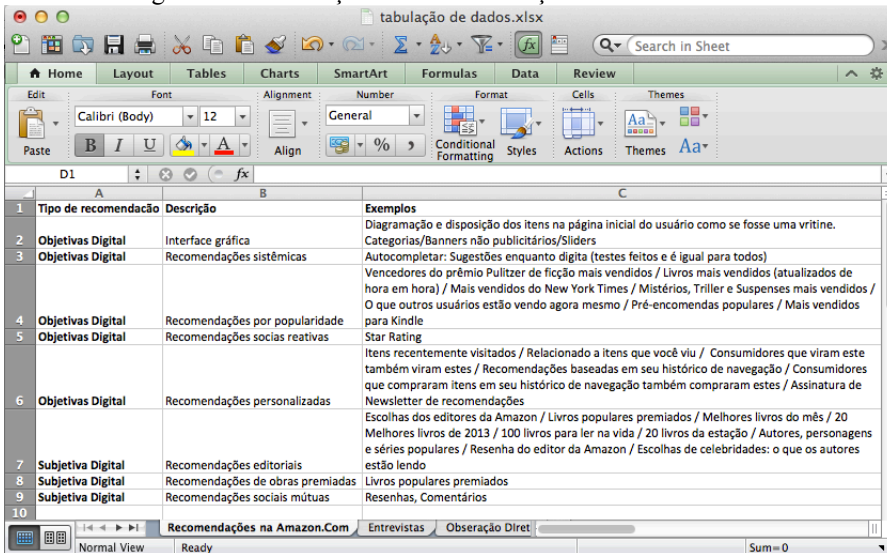
Os resultados da observação direta foram tabulados juntos aos dados já tabulados na entrevista em profundidade, a fim de avaliar as informações passadas pelo próprio entrevistado, bem como coletar dados diretamente da experiência interativa que poderiam oferecer melhores indícios do que aqueles resultantes da interpelação oral no momento da entrevista.

4.3 ANÁLISE DOS DADOS

Os dados coletados na pesquisa foram analisados em correlação e confrontados com a fundamentação teórica. As experiências individuais foram correlacionadas umas às outras para posterior descrição e análise dos dados.

A análise inicia com a classificação das recomendações no website da Amazon, no qual as formas de recomendação no site são agrupadas em elementos na Matriz de Recomendações de Itens de Informação. A tabulação dos dados dos tipos de recomendações coletados do website da Amazon foi feita na planilha representada pela Figura 30.

Figura 30 – Tabulação das recomendações na Amazon.com.



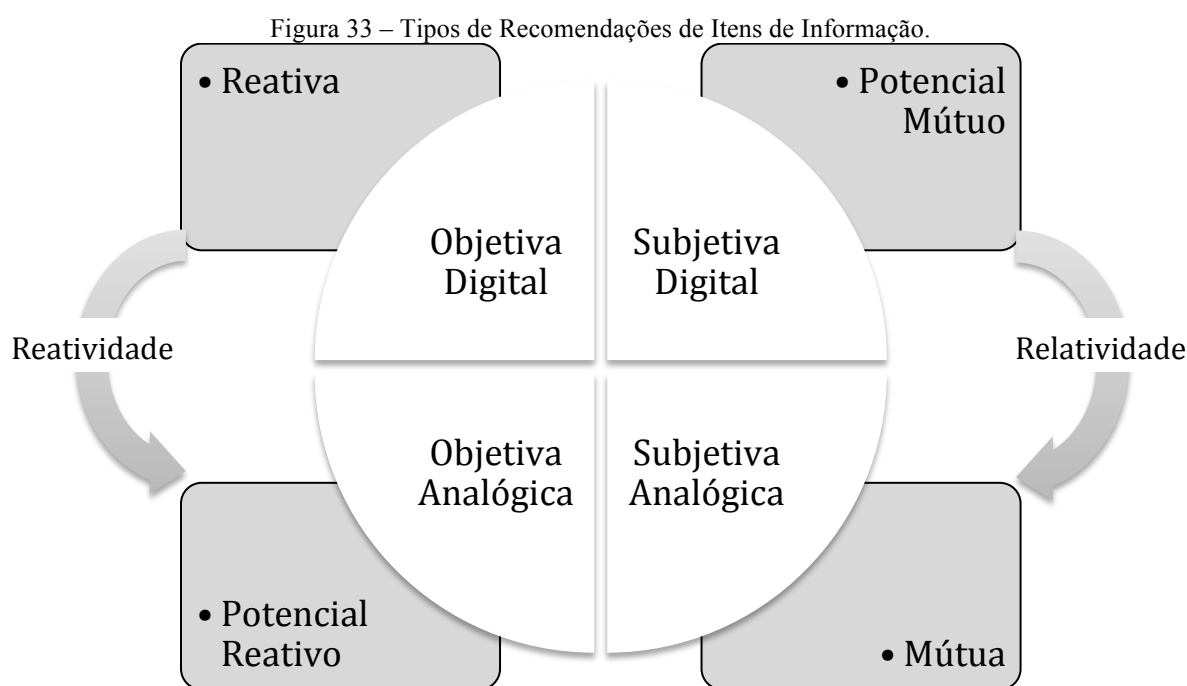
A	B	C	
1	Tipo de recomendação	Descrição	Exemplos
2	Objetivos Digital	Interface gráfica	Diagramação e disposição dos itens na página inicial do usuário como se fosse uma vitrine. Categorias/Banners não publicitários/Sliders
3	Objetivos Digital	Recomendações sistêmicas	Autocompletar: Sugestões enquanto digita (testes feitos e é igual para todos)
4	Objetivos Digital	Recomendações por popularidade	Vencedores do prêmio Pulitzer de ficção mais vendidos / Livros mais vendidos (atualizados de hora em hora) / Mais vendidos do New York Times / Mistérios, Thriller e Suspenses mais vendidos / O que outros usuários estão vendo agora mesmo / Pré-encomendas populares / Mais vendidos para Kindle
5	Objetivos Digital	Recomendações sociais reativas	Star Rating
6	Objetivos Digital	Recomendações personalizadas	Itens recentemente visitados / Relacionado a Itens que você viu / Consumidores que viram este também viram estes / Recomendações baseadas em seu histórico de navegação / Consumidores que compraram itens em seu histórico de navegação também compraram estes / Assinatura de Newsletter de recomendações
7	Subjetiva Digital	Recomendações editoriais	Escolhas dos editores da Amazon / Livros populares premiados / Melhores livros do mês / 20 Melhores livros de 2013 / 100 livros para ler na vida / 20 livros da estação / Autores, personagens e séries populares / Resenha do editor da Amazon / Escolhas de celebridades: o que os autores estão lendo
8	Subjetiva Digital	Recomendações de obras premiadas	Livros populares premiados
9	Subjetiva Digital	Recomendações sociais mútuas	Resenhas, Comentários
10			

Fonte: Próprio autor

Posteriormente à classificação dos tipos de recomendações, é analisado a partir das entrevistas em profundidade e das observações diretas, o comportamento dos interagentes na experiência de seleção de itens de informação, onde são abordadas além das recomendações digitais também aquelas oriundas de interações analógicas.

5 RECOMENDAÇÃO DE ITENS DE INFORMAÇÃO NA AMAZON

As recomendações presentes no website da Amazon podem ser classificadas nas categorias de Recomendações Objetivas Digitais ou Recomendações Subjetivas Digitais. Como será apresentado neste capítulo, as Recomendações Objetivas pesquisadas são aquelas que apresentam as interações reativas e com potencial reativo e as Recomendações Subjetivas encontradas são caracterizadas pela interação com potencial mútuo e mútuas. Analisou-se esse potencial interativo visto que, como Primo (2007) coloca que a interação mútua exige recursividade inter-relacional e Hinde (1979) destaca que para haver relação é necessária interação ao longo do tempo, nem toda Recomendação Subjetiva será necessariamente mútua, bem como nem toda Recomendação Objetiva será totalmente reativa. Logo, as interações que resultam nas Recomendações Objetivas e Subjetivas, como serão analisadas na sequência, apresentam-se, ora potencialmente reativas, ora potencialmente mútuas. Por esse motivo, a linha que diz respeito ao tipo de interação envolvida nos tipos de recomendações na Matriz de Recomendações de Itens de Informação proposta na Tabela 4 anteriormente nesta tese, deve ser observada como um espectro, onde as interações podem ser alocadas em seus extremos desde reativas até mútuas, ou entre elas onde são dispostas desde aquelas com potencial reativo até as com potencial mútuo. O que determinará a classificação será a reatividade ou a relatividade presente no processo interativo, como observado na Figura 33 a seguir que apresenta graficamente os Tipos de Recomendações de Itens de Informação.



Fonte: Próprio autor

No texto que segue, primeiramente, as recomendações presentes no website da Amazon, ao serem analisadas em 5.1, são classificadas a partir da Matriz de Recomendações de Itens de Informação, na qual diferentes formas de recomendações são agrupadas em elementos em cada um dos tipos de recomendações da matriz. Essa primeira etapa de análise ainda não considera as Recomendações Analógicas, pois não dizem respeito ao meio virtual, mas estas receberão merecida atenção no tópico subsequente, onde os dados oriundos das entrevistas em profundidade e das observações diretas permitirão analisar o comportamento dos interagentes na seleção de itens de informação em um website com um sistema de recomendação que abriga tanto Recomendações Objetivas quanto Recomendações Subjetivas, analisados em 5.2.

5.1 CLASSIFICAÇÃO DAS RECOMENDAÇÕES DE ITENS DE INFORMAÇÃO

A primeira análise feita em relação às informações presentes no Website da Amazon diz respeito ao reconhecimento dos dados enquanto informação por parte do receptor, que deve decodificar (SHANNON e WEAVER, 1949) os dados apresentados e atribuir algum conhecimento sobre eles. Visto que Fairthorne (1961) coloca que a informação é um atributo do receptor, os dados apresentados à primeira vista do internauta somente serão reconhecidos como informação, e doravante recomendações, se assim forem atribuídos. A ressalva torna-se importante para efeito de análise, visto que este primeiro tópico tratará de classificar todos os tipos de recomendações no site da Amazon, mas que nem todos esses tipos aqui analisados necessariamente aparecerão nas análises subsequentes das entrevistas e das observações diretas. Logo, nem todos os tipos de recomendações feitas pelos editores, consumidores e pelo próprio sistema da Amazon de forma automatizada são atribuídas como informação pelos internautas, pois nem todos dados tem a sua materialidade (FELINTO, 2001) reconhecida enquanto informação atuante como actante (LATOUR, 2005), neste caso como um agente não humano, no processo de seleção de itens de informação. Embora reconheça-se essa necessidade de atribuição de informação sobre os dados para tornarem-se recomendações, a primeira análise será feita a partir do reconhecimento de um potencial informativo que esses dados carregam consigo e por isso todos os tipos de indicações presentes serão analisadas e classificadas na Matriz de Recomendações de Itens de Informação. Nos tópicos a seguir, as recomendações do tipo Objetiva Digital serão analisadas em 5.1.1 e as do tipo Subjetiva Digital em 5.1.2.

5.1.1 Recomendações do tipo Objetiva Digital na Amazon

As recomendações do tipo Objetiva Digital são aquelas que indicam itens de informação a partir de um tratamento matemático dos dados resultantes de interações reativas no meio virtual. Nesse tipo de recomendação, o sistema utiliza algoritmos para selecionar os itens que serão apresentados na interface digital do usuário de forma automatizada. As recomendações desse tipo encontradas no website da Amazon estão agrupadas nos cinco elementos seguintes: Interface Gráfica Dinâmica, Recomendações Personalizadas, Recomendações por Popularidade, Recomendações Sistemáticas e Recomendações Sociais Reativas.

5.1.1.1 Interface Gráfica Dinâmica

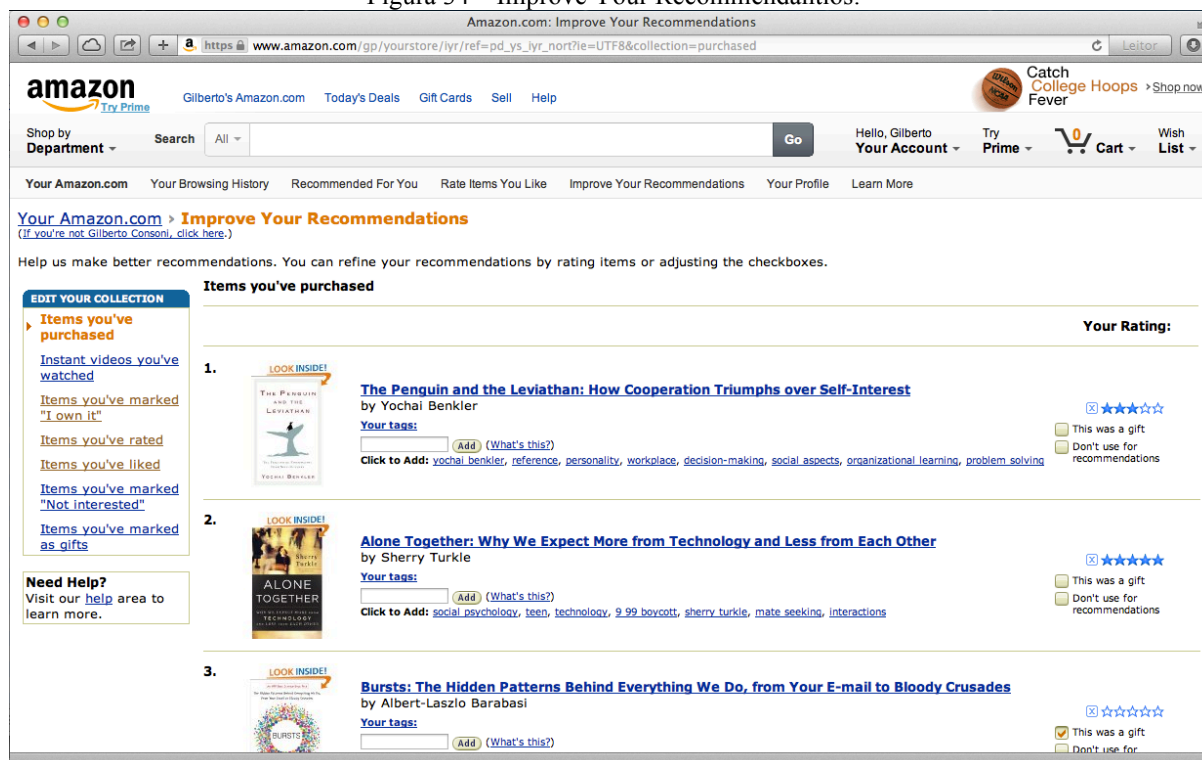
As páginas do website da Amazon moldam-se a partir de informações obtidas pelo sistema no perfil do usuário para dispor os itens de informação. Como a disposição dos itens podem ocupar mais ou menos visibilidade na diagramação da página, a própria interface gráfica das páginas da Amazon torna-se uma espécie de vitrine personalizada para os internautas que navegam pelo site. A diagramação da interface gráfica do website de forma customizada é resultado do sistema de recomendação objetivo digital que se baseia em um perfil de usuário para montar automaticamente as páginas que são apresentadas.

Como estudado em Montaner, López e Rosa (2003), um dos pontos-chaves dos agentes de recomendação é a geração e a manutenção do perfil do usuário. No caso do site da Amazon, as técnicas apresentadas pelos autores de representação, criação, retroalimentação, aprendizagem e adaptação do perfil do usuário baseiam-se em retroalimentações implícitas e explícitas (HU, SMEATON e SUN, 2012). A partir do histórico de navegação e da avaliação de itens do usuário o sistema alimenta implicitamente a sua base de dados para gerir o perfil do usuário e fazer recomendações. Quando os usuários não efetuam *login* o sistema utiliza os *cookies* para recomendar conteúdo, e quando efetuam *login*, o sistema utiliza a alimentação implícita do histórico de navegação.

Há basicamente duas formas de o usuário adaptar explicitamente o seu perfil no site da Amazon. A primeira opção explícita que se aproxima desse formato de adaptação do perfil é a ação do usuário de adicionar itens a sua *wish list*^{xxiii}, mas esta retroalimentação para o sistema de recomendação só poderá ser considerada explícita se o usuário reconhecer essa ação. Ainda que o sistema aproveite dessa interação intencional do usuário de adicionar um

item a *wish list*, não é explícito que essa ação será aproveitada pelo sistema de recomendação. A segunda forma de retroalimentar o sistema explicitamente é visitando a seção “*Improve Your Recommendations*” apresentada na Figura 34, onde o usuário pode avaliar pelo *Star Rating* itens que comprou, itens que assistiu, itens que marcou que tem, itens que gostou, itens que marcou que não interessam e itens que marcou como presentes.

Figura 34 – Improve Your Recommendations.



Fonte: Amazon.com

O site ainda disponibiliza a possibilidade de se adicionar etiquetas a cada um dos itens, que poderia ser aproveitado como uma forma de lista de preferências, o que potencializaria a Folksonomia (WAL, 2007), mas o serviço foi descontinuado e o próprio sistema recomenda que o usuário possa se interessar em utilizar a *wish list*, as Resenhas dos Clientes e as Recomendações.

O que se observa desde já é que a maior parte da interface gráfica da Amazon que é apresentada ao usuário é montada a partir de interações reativas (PRIMO, 2007) mantidas com o sistema, o qual utiliza de retroalimentação implícita (HU, SMEATON e SUN, 2012) para gerenciar e manter o perfil do usuário (MONTANER, LÓPEZ e ROSA, 2003) e dispor a diagramação dos itens de informação que são dinamicamente apresentados ao internauta.

As duas partes que não são apresentadas automaticamente a partir do algoritmo, mas sim colocadas arbitrariamente são os links, das seções e subseções, e os *sliders*^{xxiv}, da página inicial e das páginas internas. Porém, deve-se observar que estas partes do site não

recomendam itens de informação específicos e sim áreas do site que a Amazon deseja destacar. Observe nas figuras seguintes um exemplo de *slider* na Figura 35 e de seção na Figura 36.

Figura 35 – Slider na Amazon.

Fonte: Amazon.com

Figura 36 – Seção na Amazon.

Fonte: Amazon.com

Os links do sistema de navegação (seções e subseções) e os *sliders* que aparecem nas páginas da Amazon são os mesmos para todos os internautas e, indiferente se o usuário efetuou *login* ou não, possuem a mesma estrutura. Porém, ao clicar em uma seção específica, o conteúdo apresentado ao usuário na página passa a considerar o seu perfil e deixa de ser o mesmo para todos os usuários.

Desse modo, ainda que a estrutura de navegação seja a mesma para todos os internautas do site, o conteúdo que é apresentado em cada seção é baseado no perfil do usuário. As seções no site da Amazon são como nas livrarias físicas, iguais para todos, mas a diferença está na disposição dos livros que diferente da loja física, onde os títulos expostos nas prateleiras são os mesmos para todos, nas prateleiras virtuais – páginas do site – os títulos expostos acompanham as preferências e o perfil de cada visitante, e moldam-se a cada visita.

Portanto, a própria interface gráfica dinâmica do site da Amazon é um elemento a ser considerado inserido no tipo de recomendação Objetiva Digital, visto que na intenção de apresentar conteúdo ao visitante da página a interface molda-se a partir de interações reativas que alimentam o perfil do usuário. Mesmo aquelas seções que são arbitrariamente criadas pelos administradores do sistema têm o conteúdo de suas páginas baseado nessas interações que, em sua maioria, adaptam implicitamente a base de dados do perfil do usuário, a partir de interações reativas dele próprio. A interface gráfica do usuário na Amazon é estruturada a partir do seu perfil que é gerado pelo algoritmo do sistema que, de forma objetiva, procura oferecer conteúdo relevante. Dessa forma, a interface gráfica dinâmica torna-se um elemento na Matriz de Recomendações de Itens de Informação do tipo Objetivo Digital.

5.1.1.2 Recomendações Personalizadas

A gestão do perfil do usuário, baseada no histórico de navegação e nas avaliações feitas no sistema para apresentar páginas personalizadas para as preferências do internauta, procura antecipar suas próprias perguntas e atender ao problema da recuperação de informação colocado por Mooers (1951), que intenciona converter a necessidade de informação do usuário em uma lista atual de documentos armazenados contendo informação útil a ele. As páginas da Amazon oferecem de forma personalizada a cada usuário listas de itens de informação baseadas em suas necessidades de informação. A representação e a organização dos itens na Amazon fornecem aos usuários facilidade de acesso às informações de seu interesse (BAEZA-YATES e RIBEIRO-NETO, 2013).

As listas de recomendações personalizadas no site são geradas automaticamente, a partir de um sistema de recomendação de abordagem híbrida, (ADOMAVICIUS e TUZHILIN, 2005; BALABANOVIĆ e SHOHAM, 1997) que utiliza exclusivamente de interações reativas para recuperar documentos. Desse modo, o grupo de recomendações personalizadas é mais um elemento que exemplifica recomendações do tipo Objetiva Digital na Matriz de Itens de Informação, já que utiliza apenas das interações reativas no meio virtual para estruturar as páginas que serão apresentadas aos usuários.

A Amazon considera as características dos itens que o usuário viu anteriormente, para recomendar outros itens que também poderão interessar. Ao apoiar-se nessas características, o sistema faz uma recomendação baseada no conteúdo (ADOMAVICIUS e TUZHILIN, 2005) em que um algoritmo busca correlacionar itens pela similaridade entre eles. Porém, visto que o sistema do site é de abordagem híbrida, associado à análise do conteúdo, também utiliza-se de filtragem colaborativa (GOLDBERG *et al.*, 1992) ao relacionar os itens vistos pelo usuário com os itens vistos por outros usuários, como observado na Figura 37.

Figura 37 – Relacionado aos itens que você viu.

Related to Items You've Viewed

You viewed

Designing the User Interface...
Ben Shneiderman, Catherine Plaisant
Hardcover
★★★★☆ (5)

Customers who viewed this also viewed

Designing the User Interface...
Ben Shneiderman, Catherine Plaisant, ...
Hardcover
★★★★☆ (9)
\$147.00 **\$114.98**

Designing Web Usability
Jakob Nielsen, 篠原 稔和
Paperback
★★★★☆ (241)
\$54.00 **\$36.73**

Prioritizing Web Usability
Jakob Nielsen, Hoa Loranger
Paperback
★★★★☆ (31)
\$55.00 **\$38.01**

Mobile Usability
Jakob Nielsen, Raluca Budiu
Paperback
★★★★☆ (13)
\$34.99 **\$22.83**

Interaction Design: Beyond Human...
Yvonne Rogers, Helen Sharp, Jenny Preece
Paperback
★★★★☆ (16)
\$59.89

Fonte: Amazon.com

O sistema correlaciona, também, as ações de compras dos usuários para recomendar conteúdo. O site apresenta uma lista de livros comprados por outros usuários baseada em itens vistos anteriormente no histórico de navegação como apresentado na Figura 38. Como a ação de comprar um item pode ser observada com mais peso pelo usuário do que a simples visita, por envolver investimento financeiro, as recomendações relacionadas às compras efetivadas podem receber mais atenção por parte do usuário.

Figura 38 – Consumidores que compraram itens em seu histórico de navegação.

Continue Shopping: Customers Who Bought Items in Your Recent History Also Bought Page 1 of 10

The screenshot shows a horizontal carousel of five book covers with their titles, authors, ratings, and prices. From left to right: 'About Face 3: The Essentials of Interaction Design' by Alan Cooper (\$29.48); 'Homepage Usability: 50 Websites...' by Jakob Nielsen (\$34.97); 'The Design of Everyday Things' by Donald A. Norman (\$11.37); 'The UX Book: Process and...' by Pardha S. Pyla (\$77.37); and 'Designing Interfaces' by Jenifer Tidwell (\$32.75). Each item includes a star rating and the number of reviews.

Fonte: Amazon.com

Ao comprar um item, o usuário avança no processo de seleção, pois não está apenas navegando pelos itens do site de forma despreocupada. O usuário visita o item, analisa-o e efetua a compra, ou seja, busca a informação, avalia as alternativas do item e completa a seleção.

Os usuários que considerarem estas recomendações baseadas nas compras dos outros têm em mente que outros usuários fizeram esse caminho anteriormente e passaram por um processo de decisão ao selecionar o item em questão. O usuário pode visitar itens, analisar suas características e rapidamente decidir abandoná-lo para seguir em frente, mas o ato de comprar caracteriza a própria seleção em si.

Diferentemente das recomendações baseadas apenas nos itens vistos, aquelas que consideram o ato de comprar são baseadas na seleção (compra) efetivada. A Amazon observa isso e destaca em sua interface essas recomendações personalizadas, baseadas nas compras dos usuários, como observado na Figura 39, que apresenta compras frequentes e conjuntas na página de um item específico.

Figura 39 – Recomendações baseadas em compras dos usuários.

Frequently Bought Together

The screenshot shows three book covers: 'Prioritizing Web Usability', 'Designing Web Usability', and 'Mobile Usability'. Below them, a price summary states 'Price for all three: \$97.57' and offers buttons for 'Add all three to Cart' and 'Add all three to Wish List'. A list of items with checkboxes shows: 'This item: Prioritizing Web Usability by Jakob Nielsen Paperback \$38.01', 'Designing Web Usability by Jakob Nielsen Paperback \$36.73', and 'Mobile Usability by Jakob Nielsen Paperback \$22.83'. A note mentions shipping details.

Customers Who Bought This Item Also Bought Page 1 of 19

The screenshot shows a horizontal carousel of five book covers with their titles, authors, ratings, and prices. From left to right: 'Don't Make Me Think: A Common Sense Approach' by Steve Krug (\$22.72); 'Designing Web Usability' by Jakob Nielsen (\$36.73); 'E-Commerce User Experience' by Jakob Nielsen (\$56.41); 'Mobile Usability' by Jakob Nielsen (\$22.83); and 'Eyetracking Web Usability' by Jakob Nielsen (\$40.28). Each item includes a star rating and the number of reviews.

Fonte: Amazon.com

A abordagem do sistema da Amazon considera tanto a experiência individual do internauta como também vai além e a correlaciona com a experiência coletiva dos outros consumidores. A partir disso, quanto mais pessoas utilizam o sistema, mais aprimorada será a filtragem colaborativa. Portanto, quanto mais pessoas utilizam o serviço, mais ele se aprimora (O'REILLY, 2005). O uso dessas correlações que são aproveitadas para recomendar conteúdo, ao considerar a experiência dos usuários no sistema de recomendações da Amazon, é uma forma de utilizar a informação como evidência (BUCKLAND, 1991), quando o usuário pode evidenciar nas compras dos outros usuários os itens preferidos para seleção.

O problema das recomendações personalizadas é que elas são feitas a partir de estereótipos (RICH, 1998) gerados para o perfil do usuário e isso pode causar excessiva especialização (ADOMAVICIUS e TUZHILIN, 2005) das recomendações, ou seja, o usuário que busca por literatura francesa passa a receber cada vez mais indicações de títulos do gênero.

O rastreamento da experiência de navegação dos usuários, ora por visualização, ora por seleção e compra de itens, apresenta uma espécie de aproveitamento de trilhas associativas (BUSH, 1945). Ao visitar ou comprar um item, anotações (GOLDBERG *et al.*, 1992) são feitas para futuras filtrações colaborativas. Neste cenário, é como se descritores (MOOERS, 1950b) fossem criados reativamente em um sistema dinâmico constantemente modificado por interações reativas (PRIMO, 2007).


5.1.1.3 Recomendações por Popularidade

Os itens que compõem o grupo que constitui o elemento de recomendações por popularidade na Matriz de Recomendações de Itens de informação são aqueles apresentados em listas, formadas pelas quantidades absolutas de vendas e encomendas, bem como aqueles que são mais vistos no momento.

O site da Amazon oferece várias formas de filtragem dos itens por sua popularidade de vendas e prioriza estas listas ao destacar sempre na primeira página de cada seção os itens mais vendidos, como observado na Figura 40.

Figura 40 – Computers & Technology Best Sellers.

Computers & Technology Best Sellers Page 1 of 4

				
Minecraft: Essential Handbook: An... Scholastic Hardcover ★★★★☆ (22) \$7.99 \$4.91	Minecraft: Redstone Handbook: An... Scholastic Hardcover ★★★★☆ (5) \$7.99 \$6.79	Final Fantasy X-X2 HD Remaster... Joe Epstein Paperback ★★★★☆ (27) \$24.99 \$14.99	Dark Souls II Collector's Edition... Future Press Hardcover ★★★★☆ (52) \$34.99 \$20.99	250+ Best Kindle Fire HDX and HD Apps... Tom Edwards, Jenna Edwards Kindle Edition ★★★★☆ (698) \$1.02

[See all best sellers in Computers & Technology](#)

Fonte: Amazon.com

A Amazon mantém parceria com o The New York Times para usar a conhecida lista *Best Sellers* do jornal. A lista é apresentada na Figura 41. Segundo o website¹¹⁹ do próprio jornal, a metodologia utilizada para a constituição da lista considera as vendas de locais que oferecem ampla variedade de títulos de interesse geral. Os locais de venda de livros impressos incluem varejistas independentes de livro; cadeias norte-americanas nacionais, regionais e locais; on-line e varejistas de entretenimento multimídia; supermercados; universidades; lojas de departamento de presentes e de desconto e bancas de jornais. O ranking E-books reflete as vendas dos principais vendedores online de e-books, com variedade de formatos populares de e-readers.

Figura 41 – The New York Times® Best Sellers.

The New York Times® Best Sellers

Hardcover [Fiction and Nonfiction](#) | [Trade Paperback Fiction](#) | [Advice & How-To](#) | [Children's Picture Books](#)

Hardcover Fiction Best Sellers Page 1 of 4

				
Power Play: A Novel Danielle Steel Hardcover ★★★★☆ (151) \$28.99 \$16.79	Night Broken (Mercy Thompson) Patricia Briggs Hardcover ★★★★☆ (735) \$26.95 \$19.12	Be Careful What You Wish For Jeffrey Archer Hardcover ★★★★☆ (307) \$27.99 \$16.79	Stone Cold (A Joe Pickett Novel) C. J. Box Hardcover ★★★★☆ (214) \$26.95 \$19.86	Words of Radiance Brandon Sanderson Hardcover ★★★★☆ (882) \$28.99 \$17.60

[See the full list](#)

Fonte: Amazon.com

¹¹⁹ Disponível em <<http://www.nytimes.com/best-sellers-books/overview.html#moreOnBestsellers>>

O problema das recomendações por popularidade, que consideram exclusivamente as quantidades de vendas dos itens é que os produtos populares tendem a ficar cada vez mais populares, ou seja, os produtos mais vendidos são cada vez mais vendidos, como Barabási (2003) observou ao estudar redes sem escalas, em que os links ricos em conexões tendem a ficar cada vez mais ricos. Para amenizar esse impacto da popularidade de itens, a Amazon atualiza as recomendações por popularidade de hora em hora. Dessa forma, o recorte mais reduzido da mostra possibilita que novos títulos consigam se destacar frente aos títulos mais populares. Desde que estes novos sejam muito vendidos na última hora.

Por um lado, as listas de livros mais vendidos podem ser observadas negativamente, por considerarem apenas a quantidade de vendas, mas por outro lado pode ser útil saber quais são os itens mais populares em uma específica seção do site, quando o usuário tem a intenção de saber o que seus pares compram. Dessa forma, as recomendações por popularidade da seção Internet & Web Culture da Amazon podem servir como uma forma de atualização para o pesquisador na área, como observado na Figura 42.

Figura 42 – Internet & Web Culture Best Sellers.

Books : Internet & Web Culture Best Sellers Updated hourly

 <p>1. The Second Machine Age: Work, Progress, and Prosperity in a... Erik Brynjolfsson ... Kindle Edition \$12.99</p>	 <p>2. The Second Machine Age: Work, Progress, and Prosperity in a... Erik Brynjolfsson ... Hardcover \$26.95 \$17.04</p>	 <p>3. Start Your Own Business, Fifth Edition: The Only Start-Up... The Staff of Entrepreneur Media Paperback \$24.95 \$14.10</p>	 <p>4. Don't Make Me Think, Revisited: A Common Sense Approach to... Steve Krug Paperback \$45.00 \$30.02</p>
 <p>5. How to Survive the End of the World as We Know It: Tactics,... James Wesley Rawles Paperback \$18.00 \$10.17</p>	 <p>6. How To Make Money Blogging: How I Replaced My Day-Job With... Bob Lotich Kindle Edition \$0.99</p>	 <p>7. Blog, Inc.: Blogging for Passion, Profit, and to Create... Joy Deangdeleert Cho ... Kindle Edition \$1.99</p>	 <p>8. How to Blog for Profit (Without Selling Your Soul) Ruth Soukup Kindle Edition \$4.99</p>
 <p>9. Physics of the Future: How Science Will Shape Human Destiny... Michio Kaku Paperback \$16.95 \$9.04</p>	 <p>10. 10 Ways To Make \$1000 Online A.C. Jones Kindle Edition \$2.99</p>		

Fonte: Amazon.com

O pesquisador pode utilizar o sistema com o exclusivo objetivo de garimpar a bibliografia que tem recebido mais atenção na sua área para posteriormente buscar pelos títulos em sua biblioteca ou até mesmo em outras livrarias virtuais ou físicas. Desse modo, o pesquisador utiliza os itens resultantes do sistema de recomendação da Amazon no aspecto de informação-como-conhecimento (BUCKLAND, 1991) que pode satisfazer a sua necessidade

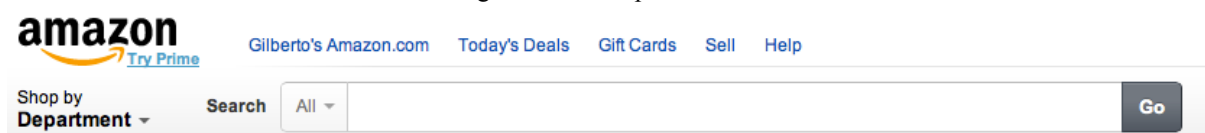
na pesquisa bibliográfica, ou aumentar ainda mais a incerteza frente ao montante de itens que podem ser sugeridos.

As recomendações por popularidade também são consideradas do tipo Objetiva Digital, pois são feitas exclusivamente baseadas na efetivação das compras dos consumidores. As vendas no próprio site da Amazon alimentam reativamente essas quantidades, a partir das interações reativas dos usuários, ao clicarem no botão comprar e finalizarem a encomenda. Por mais que o processo de seleção e escolha para a compra de um livro possa envolver interações mútuas, as listas apresentadas nas recomendações por popularidade são resultantes exclusivamente de somas matemáticas indiferentes se quem comprou o livro gostou ou não da leitura do texto.

5.1.1.4 Recomendações Sistêmicas

O website da Amazon disponibiliza na interface de todas as suas páginas um campo de formulário destinado para a ferramenta de busca por meio de entrada de texto digitado pelo usuário. O campo, que pode ser observado na figura 43, está localizado na parte superior do site e merece atenção, porque é o que melhor representa os sistemas de recuperação de informação tradicionais. Como serão explicados, os resultados para as buscas são resultantes de indexações prévias e anotações sobre os itens.

Figura 43 – Campo de busca.



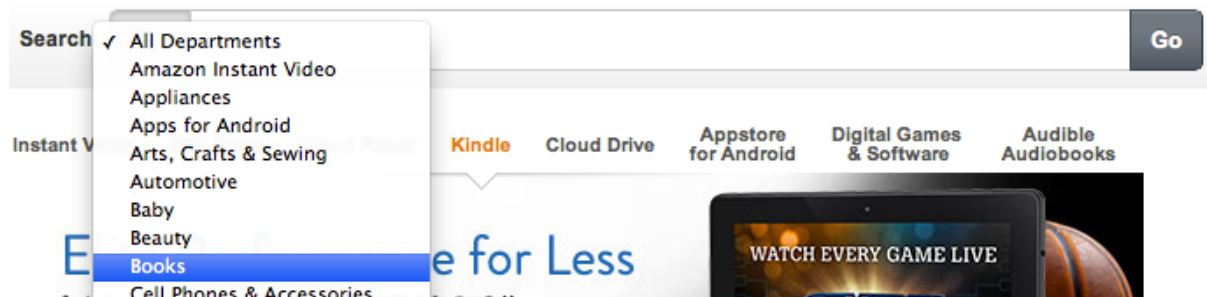
Fonte: Amazon.com

O usuário pode digitar no campo de busca qualquer dado em texto que pode estar associado ao item de informação que deseja recuperar. A busca pode ser feita por qualquer anotação que a obra possa possuir em sua indexação, como pelo título da obra, pelo nome do autor, pelo tema, ano de publicação, etc. Pode-se ainda fazer uma busca por mais de uma anotação em conjunto ou utilizar os conectivos do modelo booleano.

Antes mesmo de iniciar a sua busca, o usuário pode selecionar, a partir da interface gráfica do sistema, o departamento que deseja efetuar sua busca. Na Figura observa-se a seleção do departamento Books, a partir do menu que aparece ao clicar no botão All. Neste caso, é como se a expressão “Books” acompanhasse todos os textos digitados no campo

de busca. Para refinar a busca pode-se utilizar os conectivos booleanos representados por - ou +, por exemplo “information +retrieval -introduction” ou “design +digital -graphic”.

Figura 44 – Filtro do Campo de Busca.



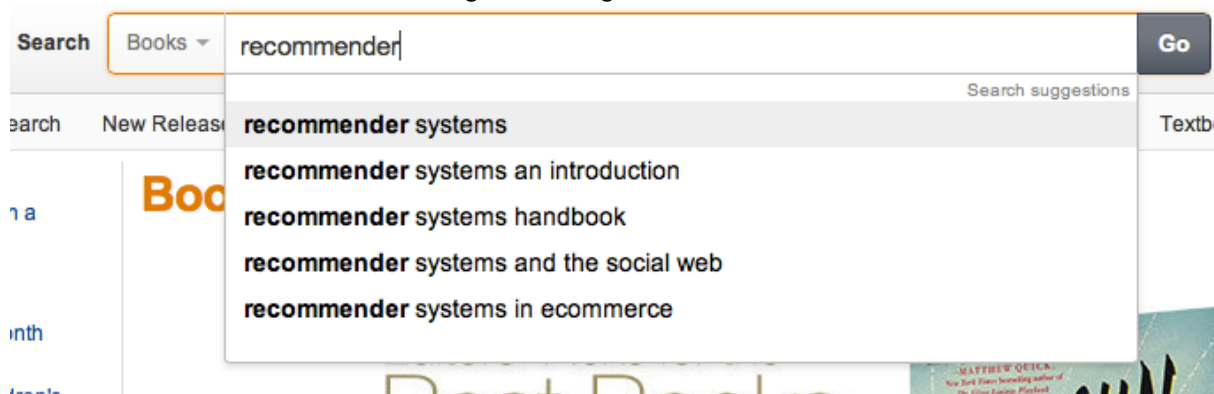
Fonte: Amazon.com

A ordenação dos resultados de busca, que apresenta a lista dos itens recuperados no site da Amazon, se dá a partir de um algoritmo que, como os principais sistemas de buscas conhecidos, é mantido em sigilo pela empresa. Porém, a partir da experiência de uso do site e de informações publicadas pela própria Amazon, sabe-se que os resultados levam em conta a popularidade do item, correspondência dos termos de busca e elementos sociais.

A autora Sansevieri (2013) sugere formas de como os autores podem otimizar a posição dos seus livros no site da Amazon, ao recomendar técnicas de SEO (*Search Engine Optimization*). Essas técnicas merecem análise nesta tese para observar, na medida em que se aprende a otimizar a posição da página de um livro nos resultados de busca do site, como o algoritmo trata os dados indexados para ranquear os itens de informação.

A primeira ação a ser feita trata das palavras-chaves de indexação do item. Primeiro, deve-se procurar qual é a principal palavra-chave que remete a obra. Essa busca, pela principal palavra-chave, não deve ser feita na própria obra, mas sim nas palavras-chaves utilizadas por obras similares na base de dados do próprio site. Deve-se depois, observar se esta palavra-chave reporta o provável termo de busca que os usuários utilizam para procurar pelos itens deste tema. A dica da autora é iniciar a digitação dos termos de busca no Google e na própria Amazon para observar quais são os termos de busca sugeridos como exemplificado na Figura 45.

Figura 45 – Sugestões de busca.



Fonte: Amazon.com

No caso do tema desta tese, ao seguir esse critério, a palavra-chave principal seria “recommender systems”. Esta palavra-chave deve ser cadastrada em primeiro lugar no campo “Search keywords” do cadastro, seguida das outras seis palavras-chaves que devem seguir essa mesma sugestão de otimização e serem inseridas em ordem decrescente de importância. Desse modo, para esta tese as seis palavras-chaves restantes seriam cadastradas na seguinte ordem: “information retrieval, human computer interaction, information technology, social web, social computing, e social communication”. Observe que essas palavras-chaves não aparecem em momento algum para os internautas da Amazon, contudo, essas informações estão no que os designers denominam como *back-end* do site, ou seja, dados que não aparecem para o usuário final, mas apenas para os administradores do site.

Definidas as palavras-chaves, o passo seguinte é otimizar o título da obra. Esse pode ser um ponto de conflito que o autor enfrentará, pois provavelmente pensou em um título original para o seu livro junto ao seu editor. Porém, como nesta simulação a intenção é de fato exclusivamente melhorar a posição da obra nos resultados de buscas do site, o título da obra deverá ser otimizado para iniciar com a principal palavra-chave anteriormente cadastrada e conter o máximo número das outras 6 palavras. Como o subtítulo da obra também deverá seguir essa linha de raciocínio, um possível título e respectivo subtítulo para esta tese ser bem classificada no sistema de busca da Amazon seria: “*Recommender systems and information retrieval for the human computer interaction in the information technology: how the social computing affects the social web and social communication?*”.

O próximo passo é escrever a descrição da obra e, mais uma vez, fazer uso da principal palavra-chave que deve aparecer de 2 a 5 vezes para cada 100 palavras do texto. A descrição da obra pode ter até 700 palavras e não há número mínimo exigido, mas é aconselhado que o texto não possua menos de 500 palavras.

Após esta etapa, passa-se para as categorias. As categorias funcionam como as seções de livros das lojas físicas e a obra deve ser cadastrada no máximo em duas categorias. Neste ponto, deve-se privilegiar categorias menos populares e, como Sansevieri (2013) sugere, até mesmo obscuras. O caso é que, diferentemente da loja física, um livro que não está cadastrado na principal categoria do tema da obra ainda será encontrado pelos termos de indexação no momento da busca. Enquanto a presença de um livro sobre sistemas de recomendação possa parecer estranho na prateleira de jornalismo em uma loja física, na virtual provavelmente passará despercebido. A vantagem de se cadastrar o livro em uma categoria menos popular é que a chance da obra se tornar a mais vendida é muito maior. Após ter se tornado a mais vendida ou ocupar uma boa posição na categoria obscura, a chance da obra ser recomendada com os mais vendidos da categoria que o livro realmente pertence também se torna maior, visto que as duas obras são as mais vendidas em suas categorias e possuem termos de indexação (título, subtítulo, e palavras-chaves) semelhantes. Com isso, a obra passa a ser mais vendida e a melhorar a sua posição nos resultados de busca, já que passa a atender também ao critério de popularidade de vendas considerado pelo algoritmo da Amazon.

Assim como o algoritmo do Google, que ranqueia melhor as páginas dinâmicas em relação às estáticas, a Amazon também privilegia em seus resultados as páginas que possuem algum movimento. Por esse motivo, é altamente aconselhado que atualizações de preços, revisões da descrição e mudanças de categorias sejam feitas em uma periodicidade de seis semanas.

Por fim, deve-se dar atenção aos elementos sociais. Este é o ponto mais difícil de otimizar, visto que depende das interações dos consumidores em avaliar os itens e escrever resenhas, já que a Amazon também dá preferência nos resultados de busca às obras mais e melhores avaliadas. Para impulsionar a participação dos leitores, sugere-se que o autor escreva suas próprias resenhas, responda as resenhas dos consumidores de forma informal e publique informações da obra como resumos, personagens, prêmios, etc, no Shelfari¹²⁰. A outra estratégia é buscar pelos resenhistas mais populares¹²¹ no próprio site, garimpar por aqueles principais que costumam fazer resenhas do tema da obra e oferecer uma cópia para avaliação, mas claramente que o autor não deve em momento algum sugerir a redação de uma

¹²⁰ O Shelfari é uma comunidade virtual desenvolvida pela Amazon para apreciadores de livros, bem renomada nos Estados Unidos. Disponível em <<http://www.shelfari.com>>

¹²¹ A Amazon disponibiliza um ranking para seus resenhistas onde pode-se acompanhar informações dos resenhistas como a sua posição no ranking, nome e link para as resenhas feitas. Disponível em <http://www.amazon.com/reviews/top-reviewers/ref=cm_pdp_top_reviewers>

resenha no site. A obra deverá ser oferecida sem compromisso algum. Este aspecto dos resenhistas será aprofundado nas recomendações do tipo Subjetiva Digital no próximo tópico, no momento apresenta-se apenas as estratégias de SEO da autora Sansevieri (2013).

O conhecimento dessas técnicas torna-se importante para observar que os resultados de buscas podem oferecer obras que receberam otimização nos termos de busca no momento do cadastro.

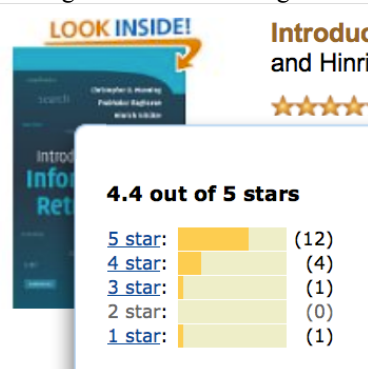
Com essas técnicas de SEO em mente, observa-se os principais parâmetros que o algoritmo da Amazon leva em conta para recuperar as informações por relevância. Verifica-se que a recuperação se dá principalmente a partir dos termos de indexação (descritores) Mooers (1950b), pelas quantidades numéricas de avaliações e de vendas, ou seja, as recomendações sistêmicas também se dão a partir de interações reativas e, por isso, também se enquadram nas recomendações do tipo Objetiva Digital.

5.1.1.5 Recomendações Sociais Reativas

O último elemento a ser classificado nas recomendações do tipo Objetiva Social no site da Amazon apresenta a participação explícita dos usuários na avaliação do conteúdo. As interações consideradas nessas avaliações sociais são aquelas exclusivamente reativas, pois tratam-se de votos reativos nos itens do catálogo, como o aproveitamento do Star Rating para recomendar conteúdo.

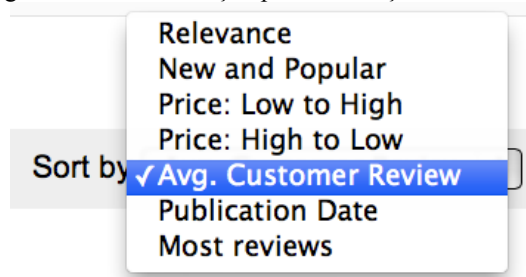
O resultado das votações por estrelas, além de aproveitado pelo algoritmo como elemento social para ranquear os resultados de busca, pode ser utilizado visualmente pelo usuário para saber como está a avaliação do título ou para classificar os itens exclusivamente pelo voto dos clientes exemplificados, respectivamente, na Figuras 46 e Figura 47:

Figura 46 – Star Rating.



Fonte: Amazon.com

Figura 47 – Classificação pelas avaliações dos clientes.

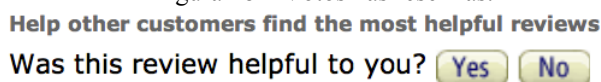


Fonte: Amazon.com

O problema das recomendações por esse tipo de avaliação é que os usuários podem pegar carona nas avaliações dos outros (RESNICK e VARIAN, 1997), tanto para selecionar como também para avaliarem itens. As recomendações sociais reativas representam uma forma de recomendação colaborativa (ADOMAVICIUS e TUZHILIN, 2005) e também são aproveitadas para contribuir com a técnica de aprendizagem do perfil do usuário (MONTANER, LÓPEZ e ROSA, 2003).

A outra forma de participar de forma reativa para avaliação diz respeito a possibilidade do usuário votar se as resenhas foram úteis ou não, apresentada na Figura 48.

Figura 48 – Votos nas resenhas.



Fonte: Amazon.com

Os votos nas resenhas contribuem tanto para destacar as resenhas mais úteis para os consumidores, como também para ranquear os próprios resenhistas. Como serão observadas no próximo tópico, as resenhas aparecem como recomendações subjetivas baseadas em interações com potencial mútuo, mas como é vista aqui, a ação reativa social é o que determina a sua posição mais acima ou mais abaixo nas páginas dos itens.

5.1.2 Recomendações do tipo Subjetiva Digital na Amazon

As recomendações do tipo Subjetiva Digital são aquelas que envolvem, necessariamente, a apreciação subjetiva humana na indicação e ocorrem, exclusivamente, a partir de interações mediadas por computador mútuas ou com potencial mútuo. Neste tipo de recomendação, a interferência do homem na recomendação vai além do manuseio reativo do sistema e considera seu posicionamento subjetivo em relação ao conteúdo apreciado. As recomendações podem provir da avaliação crítica da obra, positiva ou negativa, bem como do resultado de premiações que contam com a criticidade humana para definir seus premiados.

Porém, apenas as premiações que avaliam de forma qualitativa o conteúdo da obra enquadram-se nesse formato de indicação, visto que aquelas que premiam, a partir de um resultado quantitativo, como campeões de vendas, enquadram-se em recomendações objetivas, analisadas no tópico anterior. As avaliações subjetivas, como seu próprio nome sugere, consideram o caráter subjetivo, relativo e qualitativo, da pessoa que avalia a obra. Ocorre a partir de uma experiência sobre o conteúdo adquirido pelo conhecimento da própria leitura prévia da obra ou do reconhecimento do próprio autor. Neste tipo de recomendação, mais importante do que os dados quantitativos que envolvem a obra, está a relação de quem recomenda frente à obra e a quem é recomendado. A recomendação de um reconhecido editor ou renomado resenhista do site da Amazon poderá estar proeminente àquelas dos editores ou dos resenhistas iniciantes, daí o caráter relativo da indicação entre as partes da interação. Logo, nas recomendações subjetivas digitais a origem da recomendação também interfere no processo de seleção. As recomendações que atendem a essas características encontradas no website da Amazon estão agrupadas nos quatro elementos seguintes: Recomendações Editoriais, Recomendações de Obras Premiadas, Recomendações de Personalidades e Recomendações Sociais mútuas.

5.1.2.1 Recomendações Editoriais

As recomendações editoriais no website da Amazon são aquelas provindas das seções em que os próprios editores, da loja virtual ou das editoras, recomendam obras que aparecem no formato de listas de itens de informação ou de resenhas escritas por eles próprios. Nestas recomendações, enquadram-se também as descrições de obras escritas por seus próprios editores/autores, bem como as resenhas de revistas renomadas, como as da *Publisher Weekly*,¹²² que aparecem nas páginas de algumas obras do site da Amazon. Ainda que essas indicações coincidam muitas vezes com os livros mais vendidos, o propósito da recomendação é estritamente qualitativo e a produção deste tipo de recomendação é feita por profissionais do mercado editorial. A característica base dessa recomendação é o caráter subjetivo dado à avaliação qualitativa da obra pelos editores. As recomendações feitas por

¹²² A *Publisher Weekly*, conhecida por PW, é uma revista semanal de notícias focadas no mercado internacional de livros fundada em 1872. A publicação oferece listas de mais vendidos e estatísticas do mercado editoria, mas é principalmente reconhecida pela publicação de cerca de 8 mil resenhas anuais. Disponível em <<http://www.publishersweekly.com/>>

esses profissionais formam o grupo que constitui o elemento de recomendações editoriais da Matriz de Recomendações de Itens de informação.

As indicações feitas pelos editores compõem o tradicional formato de recomendações subjetivas produzidas pelo próprio mercado editorial há décadas e são caracterizadas pelo mesmo modelo de interação dos meios de comunicação de massa, em que a orientação da interação é de um-para-muitos (THOMPSON, 1998). Da mesma forma que a teoria do Gate Keeper (WHITE, 1950) aponta que a produção de notícias no jornalismo é um processo de escolhas que passa por diversos *gates* (portões), onde seus porteiros (assessores, jornalistas, organizações) agem arbitrariamente e subjetivamente para determinar a pauta, as recomendações editoriais tradicionais também seguem um processo semelhante para determinar quais obras merecem atenção a fim de compor a lista de recomendações e serem merecedoras da produção profissional de resenhas.

Mesmo que se reconheça a limitação da teoria do Gate Keeper ao analisar a produção de notícias apenas a partir de quem a produz, já que é sabido que os jornalistas reconhecem a importância e a influência da audiência na produção da pauta, a teoria apresenta a destacada posição do produtor em relação ao receptor da informação. Visto à falta de reciprocidade nos meios de comunicação de massa (BRECHT, 2005), a relação pode se dar de forma em que o emissor é dominante e o receptor dominado (ENZENSBERGER, 1978). Para que essa relação não se estabeleça nessa forma de dominância, torna-se necessário o estabelecimento de confiança na relação entre quem emite o conteúdo e quem o recebe. Essa relação entre os interagentes se dá ao longo do tempo (HINDE, 1979). Dessa forma, o receptor da informação poderá confiar e não se colocar na posição de dominado frente alguns produtores que reconhece há mais tempo. Esse reconhecimento do produtor, no caso os editores, autores e respectivas editoras, para o leitor pode ser ora positivo, quando associa boa qualidade às obras e às recomendações, ora negativo, quando observa a falta de qualidade nas obras e nas recomendações. O mau avaliador (resenhista) será aquele reconhecido por fazer indicações ruins, talvez rasas ou até mesmo enganosas. O bom será aquele que é renomado por construir boas avaliações, tanto no sentido de persuadir quanto de dissuadir, conhecido por ser útil no processo de seleção de itens.

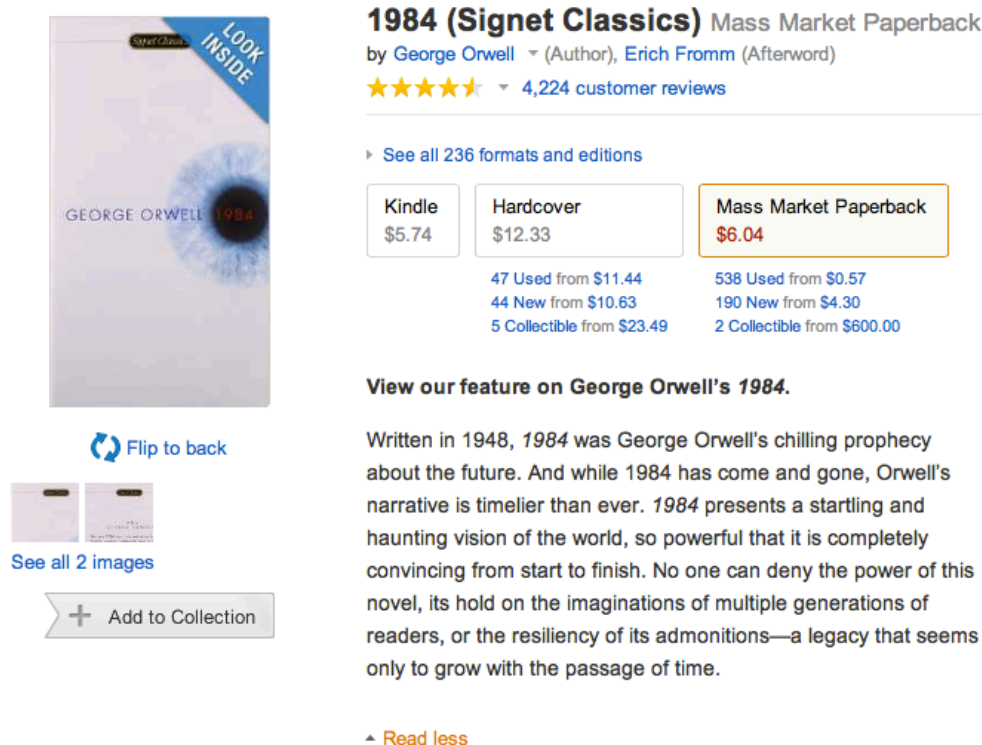
Esse aspecto da relação no processo interativo nas recomendações editoriais é importante porque esse tipo de recomendação, Subjetiva Digital, baseia-se em interações mútuas ou com potencial mútuo. Como a interação mútua é resultado de trocas recursivas não-somativas (PRIMO, 2007), quanto mais se aproxima a recomendação desse tipo de interação, maior será a chance do receptor saber da procedência da recomendação. Ressalta-se

que a relação de causa/consequência não é de quanto mais mútua é a interação, melhor é a recomendação ou o avaliador. No entanto, quanto mais mútua é a interação, maior é a chance de saber a qualidade do avaliador, tanto para o bem quanto para o mal.

A relação de confiança que a revista PW construiu com seus leitores tem mais de um século e é natural que as pessoas confiem na qualidade de suas recomendações. Porém, as resenhas da revista que são publicadas junto à página dos livros na Amazon, por mais que se caracterizem pela subjetividade, a sua simples leitura não pode ser considerada uma interação totalmente mútua, visto que não é comum redatores e leitores interagirem entre si. Normalmente, o canal de comunicação estabelecido na troca não é direto um-para-um, mas existe esse potencial. Por este motivo, as recomendações editoriais estão posicionadas no espectro mais para o lado da interação com potencial mútuo do que mútuas. Trata-se de uma questão de relatividade entre as partes envolvidas, que não necessariamente interfere na qualidade e na confiança da recomendação e sim na forma como ela se estabelece.

Nas páginas dos livros da Amazon, existem basicamente três tipos de recomendações editoriais. A primeira delas diz respeito à própria descrição do livro, que é fruto de texto redigido e informado pela editora da obra. A descrição da obra está posicionada logo ao lado da figura da capa do livro e abaixo das informações de título, autor(es), formatos disponíveis e preços, como observado na Figura 49.

Figura 49 – Descrição da obra.



1984 (Signet Classics) Mass Market Paperback
by [George Orwell](#) (Author), [Erich Fromm](#) (Afterword)
★★★★★ 4,224 customer reviews

▶ See all 236 formats and editions

Kindle \$5.74	Hardcover \$12.33	Mass Market Paperback \$6.04
------------------	----------------------	--

47 Used from \$11.44
44 New from \$10.63
5 Collectible from \$23.49

538 Used from \$0.57
190 New from \$4.30
2 Collectible from \$600.00

View our feature on George Orwell's 1984.

Written in 1948, *1984* was George Orwell's chilling prophecy about the future. And while *1984* has come and gone, Orwell's narrative is timelier than ever. *1984* presents a startling and haunting vision of the world, so powerful that it is completely convincing from start to finish. No one can deny the power of this novel, its hold on the imaginations of multiple generations of readers, or the resiliency of its admonitions—a legacy that seems only to grow with the passage of time.

▲ [Read less](#)

Fonte: Amazon.com

Ao pensar no cenário discutido anteriormente, no que diz respeito ao relacionamento do consumidor com a editora do livro em questão, a descrição pode ser observada positivamente, como um resumo do livro que, apresenta de forma rápida, recomendações úteis para decidir selecionar ou não o item, ou negativamente, quando considerada pelo leitor apenas como uma espécie de propaganda com o objetivo de vender o livro. Portanto, a descrição da obra pode ser observada como um bom meio de recomendação quando a editora em questão é reconhecida e bem conceituada pelo leitor.

As outras duas formas de recomendações editoriais nas páginas dos livros são as dos editores da Amazon e as dos editores da PW. A diferença do caráter subjetivo entre essas duas formas, editores da Amazon e da PW, e a analisada anteriormente, editores das editoras, pode ser observada na relação profissional com a obra e o objetivo em promovê-la. Deve-se relembrar nessas formas de recomendações editoriais expostas que sempre há a questão arbitrária da teoria do Gate Keeper e que, naturalmente, há a presença do objetivo da venda detrás da recomendação. Nessa linha, a editora aparece como maior interessada na venda dos livros, do contrário nem os publicaria, seguida da Amazon que, por mais que tenha o objetivo de atender fielmente seu consumidor, também deseja vender o maior número de livros possíveis para aumentar sua receita, e por último aparece a revista PW que, ainda que não possua participação direta nas vendas das obras, tem interesse em atender as expectativas de consumo dos assinantes e de publicidade dos anunciantes, bem como vender assinaturas da própria publicação. Portanto, a criticidade em relação à avaliação da obra apresenta-se inversamente proporcional à expectativa de venda da mesma.

A descrição é uma recomendação que geralmente não envolve teor crítico, e objetiva apenas resumir o tema, personagem, entre outros aspectos gerais da obra. As recomendações dos editores da Amazon e da PW avaliam a obra criticamente e, como observado na Figura 50, são assinadas pelo resenhista que avaliou o livro.

Figura 50 – Resenhas dos editores da Amazon e da PW.

Editorial Reviews

Amazon.com Review

An Amazon Best Book of the Month, October 2013: It's hard to articulate just how much—and why—*The Goldfinch* held such power for me as a reader. Always a sucker for a good boy-and-his-mom story, I probably was taken in at first by the cruelly beautiful passages in which 13-year-old Theo Decker tells of the accident that killed his beloved mother and set his fate. But even when the scene shifts—first Theo goes to live with his schoolmate's picture-perfect (except it isn't!) family on Park Avenue, then to Las Vegas with his father and his trashy wife, then back to a New York antiques shop—I remained mesmerized. Along with Boris, Theo's Ukrainian high school sidekick, and Hobie, one of the most wonderfully eccentric characters in modern literature, Theo—strange, grieving, effete, alcoholic and often not close to honorable Theo—had taken root in my heart. Still, *The Goldfinch* is more than a 700-plus page turner about a tragic loss: it's also a globe-spanning mystery about a painting that has gone missing, an examination of friendship, and a rumination on the nature of art and appearances. Most of all, it is a sometimes operatic, often unnerving and always moving chronicle of a certain kind of life. "Things would have turned out better if she had lived," Theo said of his mother, fourteen years after she died. An understatement if ever there was one, but one that makes the selfish reader cry out: Oh, but then we wouldn't have had this brilliant book! —Sara Nelson

From Publishers Weekly

Donna Tartt's latest novel clocks in at an unwieldy 784 pages. The story begins with an explosion at the Metropolitan Museum that kills narrator Theo Decker's beloved mother and results in his unlikely possession of a Dutch masterwork called *The Goldfinch*. Shootouts, gangsters, pillowcases, storage lockers, and the black market for art all play parts in the ensuing life of the painting in Theo's care. With the same flair for suspense that made *The Secret History* (1992) such a masterpiece, *The Goldfinch* features the pulp of a typical bildungsroman—Theo's dissolution into teenage delinquency and climb back out, his passionate friendship with the very funny Boris, his obsession with Pippa (a girl he first encounters minutes before the explosion)—but the painting is the novel's secret heart. Theo's fate hinges on the painting, and both take on depth as it steers Theo's life. Some sentences are clunky (suddenly and meanwhile abound), metaphors are repetitive (Theo's mother is compared to birds three times in 10 pages), and plot points are overly coincidental (as if inspired by TV), but there's a bewitching urgency to the narration that's impossible to resist. Theo is magnetic, perhaps because of his well-meaning criminality. *The Goldfinch* is a pleasure to read; with more economy to the brushstrokes, it might have been great. Agent: Amanda Urban, ICM. (Oct. 22)

[See all Editorial Reviews](#)

Fonte: Amazon.com

Naturalmente, pelos interesses comerciais existentes na relação, é esperado que as críticas da Amazon estejam mais sujeitas a tais interesses do que as das PW, mas ainda assim, estão neste aspecto de isenção à frente da descrição da editora. A própria assinatura do autor presente na resenha é uma forma de tentar expressar o posicionamento isento, já que a crítica não é da empresa e sim daquele profissional do ramo editorial que assina. A publicação do nome do autor da resenha junto ao texto é, inclusive, o que aparece como elemento permissionário a interações potencialmente mútuas e que os leitores estabeleçam interação mútua (PRIMO, 2007) com orientação um-para-um (THOMPSON, 1998) com o resenhista. É neste ponto que aparecem as listas de itens de informação, formadas a partir de recomendações do tipo Subjetiva Digital, que também tem potencial mútuo, porém aquém das resenhas assinadas pelos editores.

Diferentemente das resenhas, em que se conhece exatamente a pessoa que a escreveu, as listas de recomendações editoriais da Amazon não são assinadas e estão dispostas em seções do site, como os melhores do ano observados na Figura 51.

Figura 51 – 20 principais escolhas dos editores para os melhores livros de 2013.

Amazon Editors' Top 20 Picks for the Best Books of 2013

Page 1 of 5

We're happy to share with you our picks for the best books of the year.



Book Title	Author	Rating (Stars)	Reviews	Hardcover Price	Kindle Edition Price
The Goldfinch	Donna Tartt	★★★★☆	(6,215)	\$16.95	
And the Mountains Echoed	Khaled Hosseini	★★★★☆	(5,488)	\$16.36	\$8.29
Thank You for Your Service	David Finkel	★★★★☆	(187)	\$14.69	\$11.41
Life After Life: A Novel	Kate Atkinson	★★★★☆	(2,087)	\$27.99	

[See all Best Books of 2013](#)

Fonte: Amazon.com

Outro aspecto a ser observado é que ao visualizar as listas de recomendações editoriais pode-se associá-las às listas de mais vendidos. Deve-se deixar claro que, por mais que as listas recomendadas pelos editores se assemelhem graficamente daquelas de *best sellers*, geradas a partir do algoritmo, as recomendações editoriais se baseiam na avaliação subjetiva e arbitrária do grupo de editores da Amazon.

Certamente que cada editor preocupa-se com a sua idoneidade no momento de sugerir as obras para a lista do grupo, mas também é verdade que eles devem se preocupar ainda mais quando escrevem resenhas assinadas com seus nomes. No grupo, a responsabilidade das indicações é dividida entre os editores.

Este aspecto pode ser tanto benéfico, no sentido de indicarem sem medo de serem pessoalmente associados à recomendação, quanto maléfico ao indicarem um item que não merecia estar citado entre as obras da lista. Felizmente, é provável que a audiência refute uma indicação não merecida e isso poderia colocar em xeque a reputação dos editores. Portanto, visto que a audiência passou a ter mais voz na avaliação das obras, como será visto mais adiante nas recomendações sociais, os editores deverão estar atentos a este aspecto.

A Amazon disponibiliza para seus clientes uma página com a relação de todos os editores da empresa, onde é publicado junto à fotografia e de seus respectivos nomes, o link para as resenhas no site, o link para as postagens no blog Omnivoracious¹²³ dos editores e uma pequena descrição profissional de cada um, observado na Figura 52.

¹²³ Disponível em <<http://www.omnivoracious.com>>

Figura 52 – Página dos editores da Amazon.

Meet the Editors

**Sara Nelson**

- Sara became the Editorial Director of Amazon.com after working as books editor at O, the Oprah magazine, and as editor in chief of Publishers Weekly. In other words, her job for many, many years has been to read a lot of books and talk and write a lot about books. Tough life, huh?
- [Visit Sara's Page](#)
- [Read Sara on Omnivoracious](#)

**Mari Malcolm**

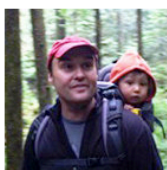
- If someone dropped a match next to Mari's bed, they'd ignite a pyre fueled by piles of novels and guides to being a better gardener, crafter, cook, designer, writer, and person. She rarely reads by candlelight.
- [Visit Mari's Page](#)
- [Read Mari on Omnivoracious](#)

**Jon Foro**

- An avid reader since age six when he ordered *Hardy Boys 53: The Clue of the Hissing Serpent*, Jon has spent over 20 years in the book business, and over 15 years at Amazon.com. He enjoys ancient history, literary fiction, and adventure and nature writing, especially books about bears.
- [Visit Jon's Page](#)
- [Read Jon on Omnivoracious](#)

**Neal Thompson**

- A journalist & [author](#), amateur photographer/ videographer, and compulsive reader-writer, Neal's rampant tastes veer from narrative non-fiction to literary fiction to long-form journalism to memoir/ biography to history, sports, music, food... Favorite way to kill an hour: book, bourbon, Miles Davis.
- [Visit Neal's Page](#)
- [Read Neal on Omnivoracious](#)

**Chris Schlupe**

- Chris spent over a decade editing books in New York before moving to Seattle, where he lives with his wife and two sons and writes about books instead. With a wide, unpredictable reading range, he takes pride in connecting readers with the books they'll love--one of the best jobs in the world.
- [Visit Chris's Page](#)
- [Read Chris on Omnivoracious](#)

**Kevin Nguyen**

- Kevin's a sucker for short story collections, magical realism, novels unrelated to NYC, comics that reinvent superhero mythology, books with colophons, and other literary things that he'll talk about for hours on end at parties. Which might explain why he hasn't been invited to a party since 2006.
- [Visit Kevin's Page](#)
- [Read Kevin on Omnivoracious](#)

**Robin A. Rothman**

- This former rock critic, TV & Radio reporter, and Entertainment & Features editor proudly wears the "Editorial Team Geek" title. She's drawn to quirky fiction, funny Sci-Fi, big fantasy, cult classics, pop culture nonfiction, and anything that will help her survive the zombie apocalypse.
- [Visit Robin's Page](#)
- [Read Robin on Omnivoracious](#)

**Seira Wilson**

- Seira can't remember a time when she wasn't infatuated with reading--17+ years in the book industry later, her house is brimming with shelves upon shelves of memoirs, cookbooks, all manner of fiction, and an insane number of children's books and YA novels. Don't judge, just borrow a book.
- [Visit Seira's Page](#)
- [Read Seira on Omnivoracious](#)

Fonte: Amazon.com

Nas resenhas dos editores nas páginas das obras na Amazon não existe link para entrar em contato com o editor ou forma de postar comentários diretamente na resenha editorial. Porém, no blog Omnivoracious, os internautas podem entrar em contato com os editores e se posicionarem nos espaços destinados aos comentários, como na Figura.

Figura 53 – Comentário em postagem da editora Sara Nelson.

Comments



You can follow this conversation by subscribing to the [comment feed](#) for this post.

RC OLeary on February 04, 2014 at 07:42 AM

Hi Sara--two quick points. Thank you for the great insight into how Oprah selects her books. I bet that information will make her fans feel even more connected to her when they read one of her selections.

Regarding *The Invention of Wings*, my wife is currently reading the book for her book club and has mentioned what a great book it is.

RC

Fonte: Amazon.com

No caso dos editores responderem aos internautas, observa-se nesses espaços destinados a comentários, a possibilidade de se manter interações mútuas entre produtores e

consumidores. Mesmo que não respondam aos comentários, a possibilidade do internauta opinar publicamente sobre as resenhas é um aspecto a ser considerado com a devida atenção pelos editores. Neste cenário, não se trata exclusivamente de uma interação de um-para-muitos (THOMPSON, 1998), mas também de um-um (interação interpessoal) e todos-todos (interação grupal) (PRIMO, 2007). Nos espaços dos comentários, os internautas podem criticar as resenhas dos editores, bem como as próprias listas de obras sugeridas por eles.

Os editores da Amazon prepararam uma lista de melhores do ano de 2013, baseada na seleção de alguns autores. Os editores pediram aos autores selecionados para nomearem o título favorito publicado em 2013 e recomendarem mais dois outros títulos. Na Figura 54 a seguir observa-se a relação de autores que apresentaram suas recomendações.

Figura 54 – Seleções de celebridades: o que os autores estão lendo.

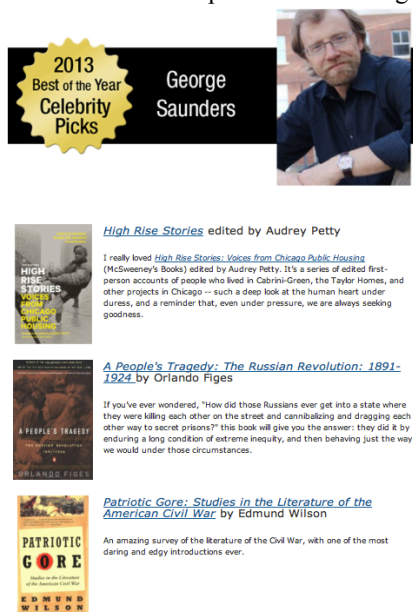
Celebrity Picks: What Writers Are Reading



Fonte: Amazon.com

A ideia dos editores da Amazon agrega às recomendações a autoridade de cada autor selecionado. A lista oferece aos internautas do site a recomendação de obras que inspiram os autores que costumam ler. A lista foi publicada na página da Amazon, conforme a figura anterior e, ao clicar no link de cada autor, o internauta era encaminhado ao blog Omnivoracious, onde era possível ver as três recomendações de cada um dos autores com seus respectivos comentários, como pode ser observado nas indicações feitas pelo autor George Saunders na Figura 55 a seguir.

Figura 55 – Melhores do ano para o autor George Saunders.



Fonte: Amazon.com

O que os editores da Amazon talvez não esperassem, ou talvez esperassem e consideraram o risco-benefício favorável, é que alguns internautas questionassem a composição da lista, como observado nos comentários da Figura 56.

Figura 56 – Comentário na postagem 2013 Best Books of the Year: Celebrity Picks.

Terrible!!! A wasted opportunity. Why would men be over-represented when women buy more romance and mystery and fantasy and Children's books (the best-selling categories!!!).

And why does everyone look alike? Where's the diversity?

An appalling (and uncharacteristic) failure on the part of Amazon.

Dawn on December 15, 2013 at 02:05 PM

I'm also dismayed by the lack of diversity. How about Amy Tan, Lisa See, Jacqueline Woodson, Tavis Smiley, Malcolm Gladwell, Jhumpha Lahiri.....

I am pleased to see authors that write for young adults included. That market is blasting wide open with excellent literature for young adults to read that their parents are enjoying as well!

Sharon McMillan on December 10, 2013 at 07:50 AM

I totally agree.....where are the GREAT African-American Authors?????

Ann Strickland on December 08, 2013 at 02:44 PM

I need some books for someone who does not like to read to keep him busy over the winter months

G. Muzio on December 08, 2013 at 07:25 AM

I did a little experiment, one that I've done before and suggest you try. Show this group of photos to someone and ask them what the people pictured have in common. In my experience, people will say "they're over 40," "they're all sitting," "they all have brown eyes." They will even stretch the truth to come up with an answer. I have tried this sort of "game" with friends and colleagues.

I've used an all-male photo roster from various other displays hoping that someone would notice that women are excluded. I've gotten the same kind of disappointing answers.

I'm flummoxed why no one at Amazon thought it was important to have a more diverse panel of authors. What is most disturbing is that, like the people in my ongoing "game", I suspect that no one at Amazon even noticed.

Fonte: Amazon.com

Os internautas questionaram a diversidade étnica e de gênero na composição da lista de autores selecionados e posicionaram veementemente suas insatisfações com relação aos autores escolhidos pelos editores nos comentários da postagem. Como os editores não informaram transparentemente como foi feita a seleção dos autores, não responderam aos comentários e tampouco esclareceram as dúvidas dos internautas em postagens futuras, a interação se mostrou orientada no sentido de um-para-muitos (THOMPSON, 1998). Porém o potencial mútuo ficou estabelecido nas trocas entre o grupo de comentaristas nesta postagem, os quais inclusive estabeleceram trocas interpessoais (PRIMO, 2007) um-para-um (THOMPSON, 1998). Os editores da Amazon, ao não responderem, perderam a oportunidade de estabelecer o equilíbrio, mas ao que tudo indica consideram melhor manter o silêncio referente à insatisfação dos leitores. Essa situação apresenta o quanto a aparentemente simples seleção de uma lista pode se tornar complexa e colocar em xeque a recomendação.

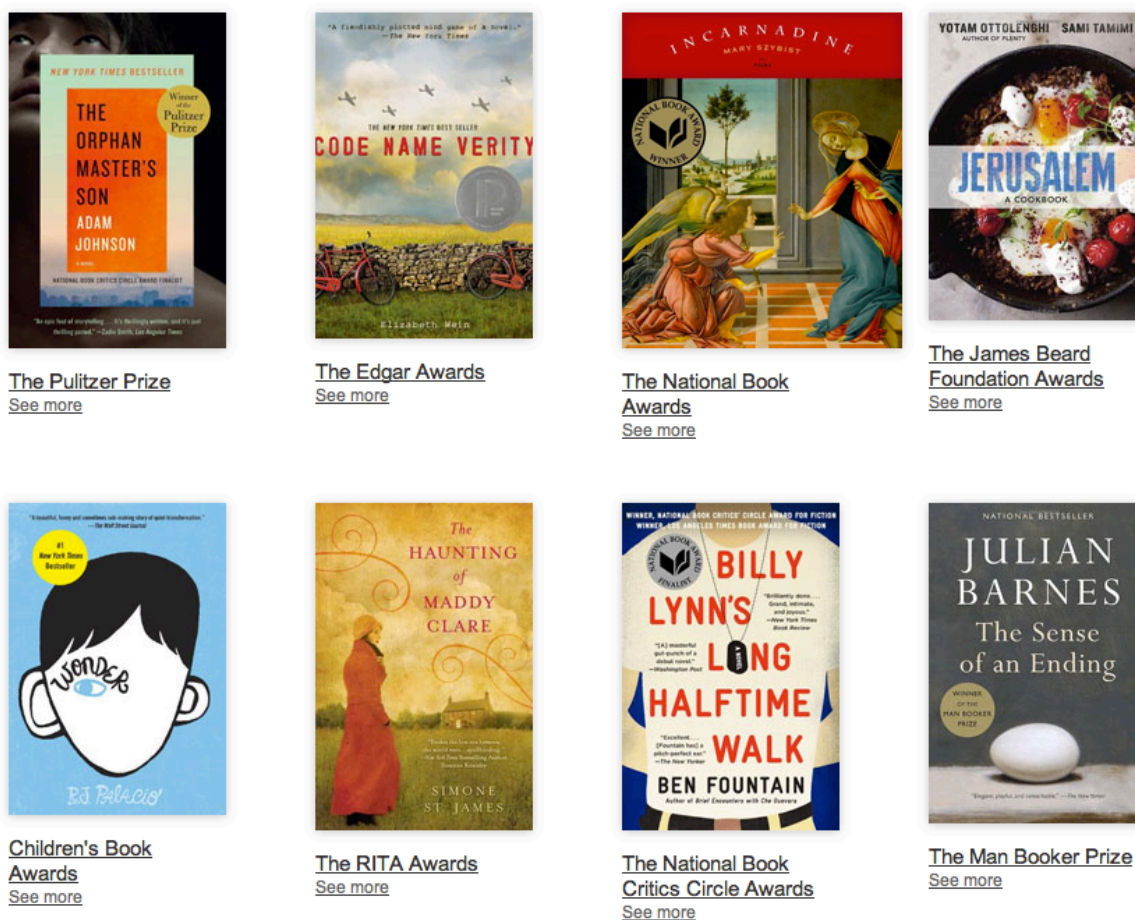
O que se esclarece a partir dessa explanação é que, como as resenhas, as listas de recomendações editoriais de seções como, *Escolhas dos editores da Amazon, Melhores livros do mês, 20 Melhores livros de 2013, 100 livros para ler na vida, 20 livros da estação*, entre outras, carregam consigo um caráter qualitativo e subjetivo na avaliação que caracterizam as formas de indicações que compõem o elemento de recomendações editoriais do tipo Subjetiva Digital na Matriz de Recomendações de Itens de Informação.

5.1.2.2 Recomendações de Obras Premiadas

As recomendações de obras premiadas no site da Amazon possuem características semelhantes às listas de recomendações editoriais, pois também têm a sua representação gráfica disposta no formato de lista e são oriundas da escolha subjetiva de pessoas. A diferença está na especificidade das pessoas que selecionam as obras e o objetivo da seleção. Não se trata de editores que selecionam e avaliam obras no sentido de recomendar, mas sim de pessoas renomadas e especialistas em suas áreas, que compõem um grupo de julgadores para avaliar as obras para os respectivos prêmios. A forma como os premiados são escolhidos não objetiva indicar um item, mas fazer a avaliação do melhor, escolher o premiado.

Porém, a forma como as listas são dispostas no site da Amazon tem o objetivo de recomendar destacadamente títulos renomados, como visto na Figura 57.

Figura 57 – Seções de livros premiados.



Fonte: Amazon.com

A recomendação de obras premiadas no site da Amazon está na seleção da própria lista em si, que é escolhida para aparecer no site. Essas listas não podem ser consideradas objetivas porque não são formadas a partir de ações reativas, mas sim de forma subjetiva e arbitrária por aqueles envolvidos nas premiações. A própria seleção das listas em si, que aparecerão no site, também não pode ser considerada uma recomendação objetiva, já que ela ocorre de forma arbitrária e subjetiva pelos editores da Amazon. O consumidor do site pode inclusive questionar por que a lista dos premiados pelo *Pulitzer* está à frente daqueles premiados pelo *Edgar Awards*. Consumidores brasileiros podem questionar a falta da lista do *Prêmio Esso*.

Nestas listas de recomendações de obras premiadas, aplicam-se as mesmas características das listas de recomendações de editores. Porém, ainda que aqui o potencial de interação mútua seja menor com quem seleciona cada item da lista, o potencial é o mesmo com os editores que escolhem quais prêmios aparecerão nas listas de obras premiadas.

Como a escolha dos itens é feita pelos prêmios, as mesmas questões comerciais anteriormente discutidas não se aplicam. Por esse motivo, a recomendação presente em uma

lista de premiados pode ser observada de forma mais idônea do que aquela que possui itens escolhidos pelos editores do site.

Da mesma forma que nas recomendações editoriais, as de obras premiadas também seguem o cenário de relação entre as partes envolvidas. Porém, em vez da relação ser apenas entre consumidor e editor, entra no processo a instituição promotora do prêmio. Pelas mesmas características de recursividade e relação ao longo do tempo, discutidas anteriormente, a relação do leitor com o prêmio também determinará sua confiança nas obras selecionadas.

Baseado nas relações entre consumidores e editores que aparecem da mesma forma que nas recomendações editoriais e pelo caráter subjetivo e arbitrário pelo qual são escolhidos os premiados, essas listas compõem o elemento de recomendações de obras premiadas do tipo Subjetiva Digital na Matriz de Recomendações de Itens de Informação.

5.1.2.3 Recomendações Sociais Mútuas

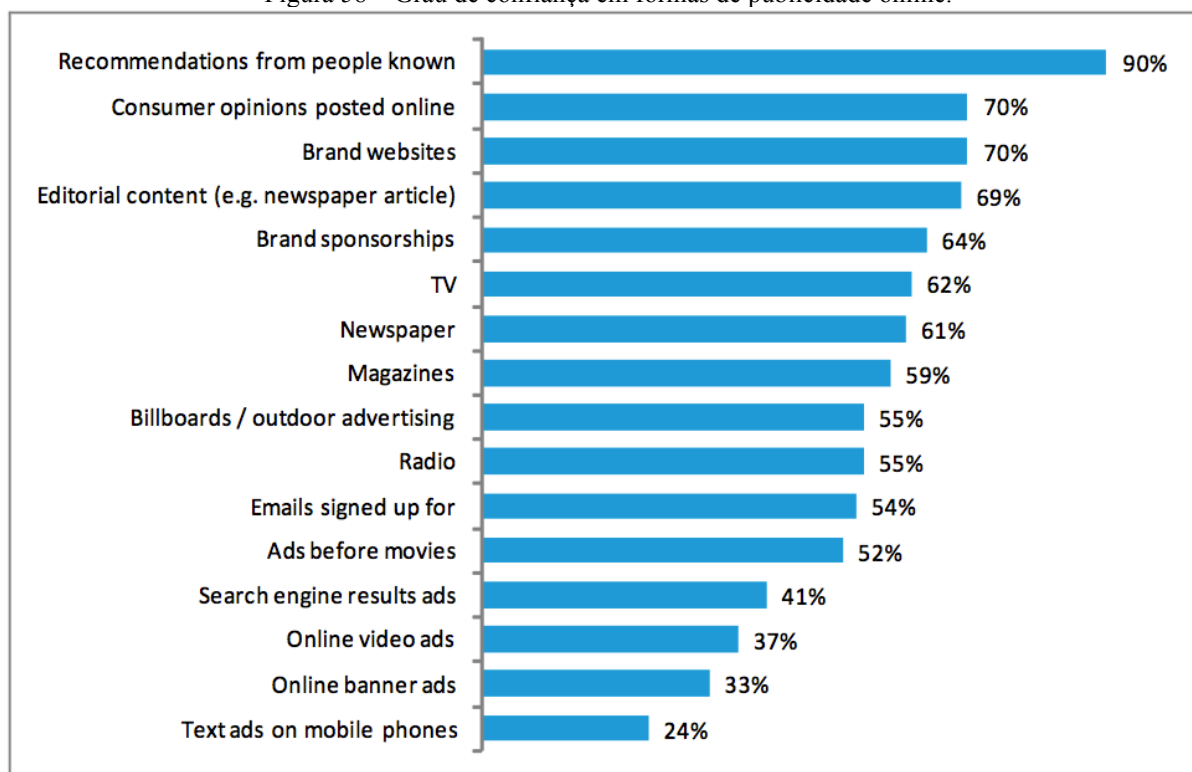
As recomendações sociais mútuas no website da Amazon são as que melhor caracterizam as recomendações do tipo Subjetiva Digital na Matriz de Itens de Informação, pois possuem o maior potencial para as interações mútuas em grupo (PRIMO, 2007), orientadas de muitos-para-muitos (THOMPSON, 1998), visto que os consumidores podem produzir resenhas e interagirem mutuamente uns com os outros nos espaços para comentários dessas avaliações.

Neste formato de recomendação é onde os consumidores podem posicionar mais ativamente suas opiniões de forma qualitativa frente às recomendações editoriais, visto que têm a possibilidade de escrever e publicar resenhas junto às páginas das próprias obras avaliadas. Com isso, as limitações de reciprocidade e dominância (ENZENSBERGER, 1978) entre produtores e receptores têm a chance de serem superadas neste espaço, mas para que isso aconteça é necessário que os consumidores participem e publiquem resenhas críticas das obras.

Na escala de isenção em relação a interesses comerciais, como abordado anteriormente nas recomendações editoriais, as recomendações sociais são as que têm maior chance de serem aceitas pelos consumidores, visto que comumente o resenhista não possui interesse na venda ou em fazer propaganda das obras que avalia. Em relação ao grau de confiança na fonte das recomendações, uma pesquisa da Nielsen Company com 25 mil consumidores de internet de 50 países diferentes revelou que 90% deles confiam nas

recomendações de pessoas conhecidas e 70% nas opiniões de consumidores postadas online. A Figura 58 apresenta o gráfico com os percentuais de grau de confiança nas formas de publicidade consideradas na citada pesquisa.

Figura 58 – Grau de confiança em formas de publicidade online.



Fonte: Nielsen Global Online Consumer Survey (NIELSEN, 2009)

O resultado mostra que os consumidores confiam significativamente em opiniões semelhantes às resenhas da Amazon e que o grau de confiança eleva-se a 90% quando a avaliação é de pessoa conhecida. Desse modo, as resenhas dos conhecidos têm ainda mais valor, e este ponto revela o quanto é importante existir a possibilidade dos consumidores interagirem mutuamente no sistema, para que se conheçam e confiem ainda mais uns nos outros, pois 70% dos pesquisados também confiam nos websites publicitários. Como a mesma confiança dada às opiniões de desconhecidos na web é dada a sites publicitários, a significativa diferença residirá na relação entre pessoas conhecidas, ou seja, está nas interações sociais o potencial de melhorar a confiança na recomendação.

Esse aspecto de construção de confiança a partir de interações mútuas (PRIMO, 2007) pode ainda ser analisado pelo viés do interacionismo simbólico (BLUMER, 1986), em que os significados das coisas são formados no contexto das interações sociais. Quanto mais as pessoas interagirem, maior será a chance de elas constituírem relações sociais interpessoais

(CONVILLE e ROGERS, 1998), tornando-se conhecidas, e conseqüentemente confiarem umas nas outras.

As resenhas dos consumidores no site da Amazon possuem destaque nas páginas dos livros. Podem ser acessadas a partir do link disposto ao lado do Star Rating e abaixo do título da obra como na Figura 59 ou lidas após as informações do autor diretamente na página do livro como na Figura 60.

Figura 59 – Link para resenhas de consumidores.



Fonte: Amazon.com

Como para escrever uma resenha é necessário o voto no Star Rating, o número que aparece ao lado das estrelas diz respeito ao número total de avaliações no Star Rating, mas não necessariamente é o número total de resenhas, visto que se pode votar nas estrelas e não escrever resenhas. No exemplo da Figura 59 anterior, a obra teve 4226 votos no Star Rating e recebeu 423 resenhas.



Figura 60 – Resenhas de consumidores diretamente na página do item.
author of six novels as well as numerous essays and nonfiction works.

Amazon Author Rank ^{beta} ([What's this?](#))

#3 in [Books](#) > [Literature & Fiction](#) > [Classics](#)

#23 in [Books](#) > [Science Fiction & Fantasy](#) > [Science Fiction](#)

#40 in [Kindle eBooks](#) > [Science Fiction](#)

#51 in [Books](#) > [Literature & Fiction](#) > [Contemporary](#)

#66 in [Kindle eBooks](#) > [Literature & Fiction](#) > [Literary Fiction](#)

Customer Reviews

★★★★★ (4,226)

4.5 out of 5 stars

5 star		2,936
4 star		819
3 star		262
2 star		101
1 star		108

[See all 4,226 customer reviews](#)

This book was very the plot of the story very well.

Kristy

George Orwell was a very political writer in books such as '1984' and, book, 'Animal Farm.'

yoshiki

Most Helpful Customer Reviews

625 of 688 people found the following review helpful

★★★★★ **The kind of distressing book you NEED to read...**

By [M. B. Alcat](#) on August 7, 2004

Format: [Mass Market Paperback](#)

Fonte: Amazon.com

As resenhas selecionadas automaticamente pelo sistema para aparecerem diretamente na página, como na Figura 60, são ordenadas pelos votos de utilidade dos próprios consumidores. Os votos dados pelos próprios consumidores, ao responderem se a resenha foi útil em que a resposta deve necessariamente ser *sim* ou *não*, permitem ordenar as avaliações mais úteis e classificar os próprios resenhistas.

Ao basear-se em experiências anteriores, o internauta que considera as resenhas no seu processo de decisão, pode reconhecer o resenhista da avaliação pelo próprio nome para saber da confiança das opiniões ou, no caso de não reconhecer, pode olhar nesses votos de utilidade o que os outros consumidores pensam a respeito daquela avaliação. Ademais, caso deseje investigar ainda mais a utilidade das avaliações do resenhista, pode clicar no nome do avaliador e ir até a página de seu perfil para conferir sua reputação em relação a todas as avaliações que fez, conforme observado na Figura 61.

Figura 61 – Página das resenhas no perfil do consumidor.

Michael Crane's Activity: Reviews

Michael Crane
Oriand Park, IL USA
Reviewer ranking: #59,402
87% helpful
votes received on reviews
(9,351 of 10,762)
www.facebook.com/authormichaelcrane

ABOUT ME
NOTE: I'm still getting a lot of review requests. While I appreciate it, I really don't review much these days unless I do it at random. With that said, I have to decline any review requests.
[See more](#)

INTERESTS
writing, reading and movies

Reviewed
Sidetracked: A Short Story
3 of 4 people found this review helpful
★★★★★ [Heeeerrreee, Kitty-Kitty....](#)
[Heeeerrreee, Kitty-Kitty....](#)
October 23, 2010
"Sidetracked" by John Fitch V is a short story about something we do often—attempt to bring our pets to the vet. It's a dirty job, but somebody's gotta do it. Randy Turner is no different, and he does his best to get his cat, Sweetums into the carrier so they can be on their way.
Sweetums has other plans.
Next thing you know, Sweetums gets [...Read more](#)

Reviewed
33 A.D. (Bachiyr, Book 1)
6 of 9 people found this review helpful
★★★★★ "You have been lied to, vampire..."
September 20, 2010
"33 A.D." by David McAfee is one of the most original and inventive novels I've read in a long time. While I'm not big on vampires, I've been known to enjoy the occasional vampire flick. I was really interested with the premise of "33 A.D." and when I began reading it, I knew there was no turning back. Within only a short amount of time, I was already invested with the story and the characters.
The story takes place in Jerusalem. T...[Read more](#)

Reviewed
Taking Care of Katrina; A Joel Arnold Quickie
3 of 3 people found this review helpful
★★★★★ "This is MY day. Don't ever tell
..."

Reviewed
The Weight of Blood (The Half-Orcs, Book 1)
2 of 4 people found this review helpful
★★★★★ "Be careful. Your power was
..."

Fonte: Amazon.com

Nesta página de perfil, podem-se observar à esquerda informações úteis para avaliar a confiabilidade do resenhista. Além do nome, localização e posição no ranking de resenhistas da Amazon, aparece o percentual de utilidade dos votos dados pelos consumidores às resenhas publicadas pelo resenhista do referido perfil.

Visto que algumas obras podem receber um elevado número de avaliações textuais, como no caso do livro *1984* de George Orwell que na data desta análise apresentava 423 resenhas, essa forma de ordenação das resenhas pelos votos dos próprios consumidores é importante para filtrar socialmente as avaliações consideradas mais úteis no processo de seleção de itens.

O voto na utilidade das resenhas pelos consumidores é caracterizado por uma interação reativa (PRIMO, 2007) e essa indicação da resenha também representa uma forma de recomendação de itens de informação do tipo Objetiva Digital. Os internautas recomendam objetivamente, a partir de interações exclusivamente com o sistema, resenhas uns aos outros. Neste aspecto em que todos podem votar nas resenhas de todos, observa-se uma forma de interação reativa (PRIMO, 2007), orientada de muitos-para-muitos (THOMPSON, 1998), que define quais as resenhas serão mais úteis enquanto recomendação no processo de escolha dos itens. Porém, destaca-se que a resenha propriamente dita não é uma recomendação Objetiva, pois seu conteúdo é escrito a partir do olhar crítico do consumidor, baseado em uma avaliação qualitativa. As resenhas nas páginas dos itens nas obras da Amazon podem ser observadas como postagens de blogs, onde o internauta tem seu próprio espaço para publicar suas opiniões, indiferente da qualidade e abordagem adotadas. Quem determinará a qualidade dessas opiniões e, no caso da Amazon, a posição delas no sistema são os próprios internautas. A ferramenta para a postagem de resenhas na Amazon compõe “tecnologias que fazem *Nós a Mídia* possível”¹²⁴ (GILLMOR, 2006, p. 23).

Além de votar de forma reativa nas resenhas, os consumidores podem postar comentários diretamente nas postagens e interagirem de forma mútua. Agora, a interação não é uma ação-reação (se-então), *se sim, então útil*, mas trata-se de uma interação mútua (PRIMO, 2007), cujos interagentes escrevem comentários subjetivos sobre as avaliações. Ao interagir com o resenhista, o comentarista realiza uma interação mútua (PRIMO, 2007) de um-para-um (THOMPSON, 1998), e, ao interagir com o grupo de comentaristas, caracteriza interações mútuas (PRIMO, 2007) de muitos-para-muitos (THOMPSON, 1998).

Os internautas também podem avaliar os comentários uns dos outros de forma reativa. O objetivo é ponderar se o comentário adiciona elementos úteis à discussão ou não, como observado na Figura 62 a seguir em que 25 internautas votaram que o comentário não adiciona elementos à discussão e apenas 5 votaram que sim. A partir dessa votação, o comentário foi ocultado na sequência de comentários da resenha.

¹²⁴ Tradução nossa para “Technology that makes We the Media possible” (GILLMOR, 2006, p. 23).

Figura 62 – Comentário oculto da lista de comentários.

Showing 1-10 of 15 posts in this discussion Sort: **Oldest first** | Newest first

Initial post: May 3, 2007 3:55:12 PM PDT
Last edited by the author on May 3, 2007 3:57:24 PM PDT

jeffreytg says:

[Customers don't think this post adds to the discussion. [Show post anyway.](#)]

Fonte: Amazon.com

Mesmo que o comentário fique oculto pelo fato de os consumidores pensarem que ele não adicionava elementos à discussão, é possível clicar no link para que seja mostrado de qualquer maneira. Na Figura 63 observa-se o conteúdo do comentário anteriormente ocultado, o qual coloca que o livro analisado “é muito inspirado por eventos reais que vieram abaixo durante a era do comunismo na Rússia”.

Figura 63 – Conteúdo do comentário oculto da lista de comentários.

jeffreytg says:

[Customers don't think this post adds to the discussion. [Hide post again.](#) ([Show all unhelpful posts](#))]

"The book is greatly inspired by real events that went down during the era of communism in Russia"- huh?
The book was written over 50 years before the fall of communism in Russia.

[Reply to this post](#) [Permalink](#) | [Report abuse](#) | [Ignore this customer](#)

5 of 30 people think this post adds to the discussion. Do you?

Fonte: Amazon.com

A interpretação dos internautas, que desconsideraram este comentário como interessante para a discussão, pode ter sido em função de acharem o comentário fantasioso, a ponto da ação do grupo ocultar o comentário para futuros visitantes.

Os novos visitantes até podem clicar para ler o comentário oculto e visitar a página do perfil do comentarista, que teve sua opinião desconsiderada pelo grupo, para conhecer sua reputação. Porém, ao fazer isso no caso desse comentarista, observará que no histórico de resenhas de seu perfil existem apenas duas avaliações. Apesar de essas resenhas terem sido consideradas úteis por 80% dos que nelas votaram, observará que foram apenas 5 votos no total. Portanto, frente às informações garimpadas ao longo da navegação pelo site, provavelmente será considerado que o comentário ocultado realmente não deve ser avaliado. Porém, ressalta-se que esta é uma interpretação do internauta que está selecionando a obra. Reforça-se que a informação é atributo do conhecimento do receptor (FAIRTHORNE, 1961).

Neste cenário, observa-se que o processo de seleção pode se tornar uma navegação complexa que ora considera recomendações Objetivas, resultantes de interações reativas (PRIMO, 2007), ora pondera recomendações Subjetivas, decorrentes de interações mútuas (PRIMO, 2007). A experiência de navegação para a escolha de cada item definirá até que ponto o internauta usará o diferentes tipos de recomendações.

Ao considerar a reputação dos resenhistas nas páginas de seus perfis no site da Amazon, bem como a própria colocação desses avaliadores no ranking de resenhistas do site, deve-se ter em mente que os dados dizem respeito a todos os departamentos, e não apenas a livros.

Ao visitar o ranking de resenhistas, encontra-se em primeiro lugar a resenhistas Joanna Daneman, observada na Figura 64, que escreveu 2770 resenhas, possui um total de 62091 votos em seus textos, que 97% dos votos são favoráveis e que 642 votantes são considerados fãs por apreciarem com frequência as opiniões da resenhistas.

Figura 64 – Página das resenhas no perfil do consumidor.

The screenshot shows the 'Hall of Fame Reviewers' section on Amazon. It features a table with the following data:

Rank	Customer Reviewer	Total Reviews	Helpful Votes	Percent Helpful	Fan Voters
# 1	 Joanna Daneman  See all 2,770 reviews	2,770	62,091	97%	642

Fonte: Amazon.com

Os números apresentados no ranking certamente são úteis para avaliar a qualidade das avaliações e podem ser muito úteis para saber se o avaliador é confiável ou não. Porém, como esta tese trata exclusivamente de livros, observa-se uma limitação deste ranking, visto que ele considera todos os produtos avaliados pelos resenhistas. Ao visitar a página do perfil da primeira colocada, observa-se que suas resenhas são feitas para avaliar desde livros até, medicamentos, luminárias, molhos para saladas, até aspiradores de pó. A respeito deste último item, por exemplo, há um modelo avaliado pela resenhistas cuja resenha recebeu 141 votos favoráveis de um total de 143. Ao navegar pelo perfil da consumidora na Amazon, observa-se que poucas de suas resenhas ultrapassam os 5 votos.

O que se deseja reforçar e alertar com essa análise é que esses dados brutos mostram que o resenhistas pode ser considerado um escritor de resenhas úteis dos itens que consome, mas que não necessariamente de livros. Visto que o fato do sujeito ser útil ao avaliar um aspirador de pó não quer dizer que fará boas resenhas de romances. O problema levantado apresenta o quanto rankings baseados em dados brutos, oriundos exclusivamente de interações reativas, são limitados. Para saber a reputação de um resenhistas de livros na Amazon será necessário mais do que a colocação no ranking, deve-se considerar o tipo de itens e, no caso de livros, de gênero que costuma avaliar. Um bom apreciador e resenhistas de romances não será, necessariamente, um bom avaliador para livros de culinária ou mesmo de

livros acadêmicos. O cenário apresentado não quer negar a possibilidade do resenhista fazer boas resenhas de diferentes itens e gêneros e ser reconhecido por isso, mas sim que será necessário conhecer mais a respeito do resenhista do que apenas dados numéricos calculados genericamente, como no ranking da Amazon, para informar o percentual de utilidade das recomendações de um determinado perfil quando coloca obras literárias ao lado de aspiradores de pó.

As recomendações sociais mútuas, quando exploram o potencial da interação mútua que possuem, certamente oferecem indícios para excelentes recomendações do tipo Subjetiva Digital, mas é importante tomar cuidado ao avaliar o resultado dessas recomendações, se estas decorrem mesmo de interações mútuas ou se por detrás delas estão na verdade interações reativas. Portanto, como analisado neste tópico, as resenhas dos consumidores da Amazon compõem o elemento de recomendações sociais mútuas do tipo Subjetiva Digital na Matriz de Recomendações de Itens de Informação.

Neste capítulo, os tipos de recomendações presentes no site da Amazon foram classificadas, caracterizadas e analisadas em grupos, divididos por elementos, os quais compõem os diferentes tipos de recomendações digitais na Matriz de Recomendações de Itens de Informação.

Os tipos de recomendações analisados até aqui consideram exclusivamente as interações mediadas por computador, logo foram analisadas apenas aquelas recomendações do tipo digital. Como pôde ser observado na análise feita até o momento, há diferentes tipos de recomendações digitais em cada uma das categorias, quando oriundas de interações reativas, ou de potencial reativo, caracterizam elementos do tipo de recomendação objetiva, e quando resultantes de interações mútuas ou de potencial mútuo qualificam elementos do tipo de recomendação subjetiva. Porém, visto que informação é atributo do conhecimento do receptor, será necessário que os usuários do site reconheçam e utilizem os diferentes tipos de recomendações digital para que estes elementos sejam identificados como informação. A forma como os usuários selecionam itens de informação e fazem uso desses elementos será apresentada no próximo tópico, em que a experiência interativa na seleção de itens no website da Amazon será analisada.

5.2 EXPERIÊNCIA DE SELEÇÃO DE ITENS NA AMAZON.COM

As recomendações no website da Amazon foram analisadas até o tópico anterior considerando apenas os dados resultantes da classificação dos tipos de recomendações a partir

da Matriz de Recomendações de Itens de Informação. Neste tópico, os dados das entrevistas em profundidade e das observações diretas desta pesquisa serão analisados para se conheça o comportamento dos internautas entrevistados ao selecionarem itens de informação no website da Amazon, onde opera um sistema de recomendação digital. Deseja-se, a partir da análise da experiência interativa dos internautas, responder como os interagentes selecionam conteúdo sob a influência dos sistemas de recomendação digital. Ainda que o foco desta tese debruçe-se nas recomendações do sistema do website da Amazon, logo aquelas exclusivamente digitais, as recomendações analógicas também serão analisadas, visto que um dos pressupostos desta tese é que o internauta recorre a mais de um tipo de recomendação quando a seleção envolve conteúdo significativo, enquanto segue passivamente sistemas de recomendações automatizados, quando o custo pessoal diretamente aplicado é baixo ou inexistente. Portanto, à medida que se analisa os dados, tanto as recomendações digitais como as analógicas serão descritas, caracterizadas e analisadas neste tópico, para investigar o comportamento dos internautas na seleção de itens de informação.

5.2.1 Busca e Recuperação de Itens de Informação na Amazon

Antes de iniciar a análise da experiência de seleção de itens no website da Amazon, deve-se recordar a diferença entre um sistema de busca e um sistema de recuperação de informação. Como será observado na análise da experiência do usuário, por um lado o sistema é muitas vezes usado para a busca de dados referentes a obras específicas de antemão conhecidas, ou seja, para identificar na base de dados quais documentos correspondem àquela consulta por dados específicos, o que caracteriza o sistema de busca. Por outro lado, quando o usuário não conhece a obra que procura e está mais interessado em recuperar informações sobre um assunto do que em dados para satisfazer determinada consulta, define-se o sistema de recuperação de informação. O interesse do usuário na recuperação de informação vai além de procurar dados que correspondam alguma consulta (BAEZA-YATES e RIBEIRO-NETO, 2013), está em selecionar obras que atendam à necessidade de informação frente a determinado tema.

Neste aspecto, o sistema de recomendação de um website, observado como um sistema de recuperação de informação, deseja não apenas atender às consultas de dados mas sim apresentar listas de documentos em resposta às perguntas do usuário frente a determinados assuntos (MOOERS, 1951). Nesta conjuntura, será a experiência de uso do sistema pelo usuário que o definirá, ora como sistema de busca, em que se procura por obras

conhecidas, ora como sistema de recuperação de informação, quando se deseja recuperar uma lista de obras que atendem a determinado assunto.

Nas entrevistas em profundidade e nas observações diretas com os usuários da loja on-line da Amazon, foi observada que a principal ferramenta utilizada para a busca ou recuperação de itens é a barra de buscas. Quando buscam por itens conhecidos específicos, os usuários digitam na barra de buscas um ou mais descritores (MOOERS, 1950b) correspondentes ao item procurado. Os descritores utilizados são quase sempre o autor ou o título da obra pesquisada. Depois de efetuada a busca, os usuários percorrem pela lista resultante na interface gráfica até encontrarem o item procurado. Neste ponto, ressalta-se o destaque dado à figura ilustrativa da capa do livro procurado que é encontrado mais rápido quando esta se assemelha à capa do item físico impresso. Como observado no relato dos entrevistados, a capa do livro é destacada no momento em que percorrem a lista de resultados: “Eu vou muito pela capa, primeiro, se a capa for meio esquisita eu já nem olho o conteúdo, é um preconceito, mas tudo bem. Já é um meio de eliminação” (informação verbal)¹²⁵.

A seleção do item escolhido nesta prática que caracteriza a ferramenta como um sistema de busca não sofre influência direta pela lista de itens apresentada no site. A decisão de escolha do item foi tomada anteriormente por outros meios, como aqueles a partir de recomendações subjetivas analisadas mais adiante, e não pela influência do sistema da Amazon. Neste aspecto, a experiência do usuário definiu o site como um mero sistema de busca, que serviu para encontrar um livro de forma mais rápida. A experiência é semelhante à busca por um item conhecido em uma loja física, quando se pergunta a um vendedor por um item específico e ele é entregue diretamente ao consumidor ou apontada a sua localização na prateleira. Porém, ainda que parecida com a experiência na loja física, há aspectos diferentes entre a prateleira e a lista on-line, percebidos na observação direta desta pesquisa que merecem atenção.

[...] o que mais me chama a atenção às vezes, é o título e às vezes, a capa. Como eu trabalho com isso sempre me chama a atenção às capas. E aí eu fui direcionado para aquele livro, eu já clico naquele, aparecem novas recomendações e às vezes eu me perco assim nessas recomendações. Eu já acabei comprando [por isso]. (informação verbal)¹²⁶.

Ao percorrerem a lista de resultados na procura do item conhecido, os usuários tendem a explorar outros itens apresentados. Como a lista apresentada considera tanto os

¹²⁵ Informação verbal obtida em entrevista com entrevistado 04.

¹²⁶ Informação verbal obtida em entrevista com entrevistado 09.

descritores (MOOERS, 1950b) digitados como as anotações (GOLDBERG *et al.*, 1992) da obra resultante, os livros apresentados conjuntamente tendem a ser muito relevantes para o usuário frente ao tema em questão da obra conhecida. No entanto, como já explicado, o que determinará se o sistema será caracterizado por um sistema de busca ou de recuperação de informação é a experiência do próprio usuário. Se o usuário procurar por um item e logo selecioná-lo, sem explorar outros itens, o sistema se caracterizará como de busca, mas se explorar os itens vizinhos ao primeiro item procurado constituirá os atributos de recuperação de informação. Ainda que as prateleiras das lojas físicas sejam organizadas com a intenção de aproximar os livros de temas semelhantes, e que o visitante possa explorar os itens vizinhos, elas são desorganizadas pelos consumidores a todo tempo. Naturalmente não existe a mesma facilidade em se percorrer uma prateleira física que em se explorar uma lista na tela do computador, bem como a disposição de itens na prateleira da livraria física não recebe atenção do profissional da recuperação de informação como em uma biblioteca ou sistema que envolve a arquitetura da informação.

Mesmo na circunstância do usuário utilizar o site como um sistema de busca, a forma como as listas são constituídas para serem apresentadas caracteriza recomendações sistêmicas quando os itens vizinhos resultantes são explorados. Este aspecto fica ainda mais evidente quando ocorre o que se evidenciou na observação direta de um dos entrevistados, que não encontrou pelo livro específico do autor que procurava e acabou por selecionar um dos outros itens oferecidos na lista de resultados.

Os visitantes das livrarias físicas geralmente levam consigo o título da obra ou nome do autor quando buscam por um item específico para que os vendedores procurem pelo item na loja. Consequentemente, a informação do vendedor de que o item não se encontra em estoque pode ser suficiente para que o visitante desista da busca naquele ambiente. Porém, quando procuram por um item na loja da Amazon, a experiência observada nos usuários analisados é que raramente buscam diretamente pelo título da obra, mas sim pelo nome do autor ou pelas palavras-chaves que resultem na obra pesquisada.

Esta prática exige necessariamente a exploração de uma lista de livros resultante da pesquisa na busca do item desejado. Como a lista formada segue os atributos da recuperação de informação, existe um potencial na experiência interativa em tornar corriqueiras buscas em complexas recuperações de informações. Ressalva-se que a intenção não é colocar a recuperação de informação à frente dos sistemas de buscas e mensurar o seu grau de importância, mas evidenciar como buscas facilmente transformam-se em recuperação de informação nos sistemas digitais como no website da Amazon. Ao pensar em sistemas de

recomendação, observar que sistemas de buscas se transformam em sistemas de recuperação de informação é importante para destacar a importância dos resultados apresentados, pois opera por detrás do algoritmo que define a disposição dos itens em melhor ou pior posicionamento nas listas de itens recomendados. Logo, a própria apresentação da lista já caracteriza uma recomendação sistêmica, que pode levar a seleção ou não de um item de informação.

A busca por livros acadêmicos evidencia essa preocupação do posicionamento dos itens no que diz respeito à recuperação de informação. Ao serem induzidos pela busca de livros acadêmicos, os usuários participantes da observação direta procuraram diretamente na barra de buscas por palavras-chaves de suas áreas de pesquisa e exploraram os itens resultantes na primeira página dos resultados. A mesma experiência foi relatada na entrevista em profundidade:

Algumas vezes eu procuro alguma coisa relacionada aquele tema: livros de ecologia, Livros de estatística, aí eu vejo os títulos que vão aparecendo, conheço alguns [...]. Como eu já tive várias disciplinas disso, conheço os autores mais importantes, então eventualmente tem um título que me chama mais atenção, por estar relacionado ao meu trabalho. (informação verbal)¹²⁷

[...] eu vou colocando lá ciências humanas, geralmente eu compro livros de sociologia que são mais específicos. Então, vou colocando as grandes categorias até chegar na sociologia e vejo as sugestões que ele dá ali. (informação verbal)¹²⁸

[...] se eu estou procurando jornalismo digital, eu geralmente vou pela busca ali, coloco umas palavras-chaves. Mas, geralmente quando eu entro [no site] eu já sei o livro que eu estou procurando, mas às vezes eu faço isso pra ver o que é que tem de repente, mas sempre por palavras chaves, como jornalismo digital. (informação verbal)¹²⁹

Visto que a obra era desconhecida, e apenas o tema de interesse conhecido, os entrevistados utilizaram o sistema para a recuperação de informação. Nesta atividade, observou-se que, ainda que um deles aproveitasse os filtros para ordenar os resultados por data de publicação, alterando assim a lista da primeira página apresentada, geralmente os itens selecionados já estavam nesta primeira página de resultados. Ao percorrerem a lista e selecionarem os itens escolhidos, os usuários relataram que suas escolhas levam em conta principalmente os autores, o título e o ano da obra.

¹²⁷ Informação verbal obtida em entrevista com entrevistado 06.

¹²⁸ Informação verbal obtida em entrevista com entrevistado 01.

¹²⁹ Informação verbal obtida em entrevista com entrevistado 02.

A análise feita dessa experiência é de que, por um lado, a escolha de itens da primeira página pode ser consequência da eficácia do próprio algoritmo, que almeja oferecer logo no início dos resultados os itens mais relevantes em resposta à busca do internauta. Por outro lado, ao conhecer as técnicas de SEO (SANSEVIERI, 2013), como analisadas nas recomendações sistêmicas anteriormente, a escolha de itens nos primeiros resultados pode ser resultante da boa otimização no cadastro da obra.

Ao observar a prática dos internautas em selecionar os primeiros resultados, a forma como os itens são ordenados nas respostas das pesquisas torna-se relevante para efeito de análise. Anteriormente, ao classificar as recomendações sistêmicas, observou-se que dentre os aspectos levados em conta no site da Amazon para a ordenação dos livros, além dos atributos da recuperação de informação, estão a popularidade de venda dos itens e os elementos sociais. Portanto, a classificação por estrelas que é uma recomendação social reativa também contribui para o melhor ou pior posicionamento da obra na lista.

Indiferente se a ordenação é resultado do algoritmo que leva em conta atributos de recuperação de informação ou de interações sociais reativas, observa-se nestes aspectos analisados a influência e limitações das interações reativas na seleção de itens de informação. Se seguir exclusivamente as recomendações objetivas oriundas de interações reativas (PRIMO, 2007), o usuário poderá ficar sujeito a resultados que não necessariamente atendam sua necessidade de informação, mas talvez às necessidades de vendas da loja. Esses dados mostram o cuidado que o internauta precisa tomar ao seguir exclusivamente um sistema de recomendação para escolher os itens que selecionará para consumo.

5.2.2 Aspectos Externos da Seleção de Itens de Informação na Amazon

Quando questionados como escolhem os itens que consomem, os entrevistados relataram que a seleção leva em conta as citações bibliográficas em leituras prévias de artigos de periódicos científicos e de livros de autores conhecidos, bem como por indicações pessoais de professores e de colegas em sala de aula, nos grupos de pesquisa e nos congressos científicos que participam.

De repente alguma discussão que surgiu no grupo de pesquisa, e daí um colega mencionou esse autor, aí se eu não conheço esse autor eu vou atrás. Se parece que, às vezes a gente recomenda, no caso da academia, por conhecer a pesquisa né, alguém indica que tu podes se interessar por esse. (informação verbal)¹³⁰

[...] geralmente conversando com alguns pesquisadores em eventos, acaba-se tendo algumas indicações, além das bancas de livros dos eventos. [...] Geralmente autores conhecidos, que são referência na área, quando o título me atrai muito. (informação verbal)¹³¹

De contato com o pessoal, ou em workshops, congressos, pelos autores dos artigos que citam livros também. Então, me faz querer buscar aquela informação quando não é disponível, se é disponível em PDF eu uso o PDF. (informação verbal)¹³²

Para as leituras de lazer, aparecem ainda as indicações feitas a partir de compartilhamentos nos sites de redes sociais da Internet e as recomendações de amigos que compartilham de interesses comuns.

Na leitura de lazer é mais indicação de amigos, aí às vezes eu vejo alguém na *timeline* nas redes sociais que a pessoa leu tal livro, aí às vezes eu acho curioso, vou buscar e me interesso pelo livro, li uma resenha do livro, vou buscar e acabo comprando. (informação verbal)¹³³.

As indicações pessoais merecem análise mais atenta, visto que são estas que melhor caracterizam as recomendações do tipo subjetivas, pois são oriundas de interações mútuas (PRIMO, 2007), mas antes é preciso conferir atenção às citações bibliográficas que também podem ser analisadas como uma forma de recomendação, ora objetiva, quando procedente de interações reativas (PRIMO, 2007) ao quantificar o fator de impacto (FI) de um periódico, ora subjetiva, quando constituída de interações com potencial mútuo no momento em que o autor que faz a citação orienta à leitura da obra citada.

Segundo a ABNT, citação é “menção de uma informação extraída de outra fonte” (ABNT, 2002, p. 1). Pelo teor desta tese e pela característica de seus leitores, dispensa-se maior definição para o termo. Contudo, destaca-se que as citações são também utilizadas para mensurar o FI dos periódicos científicos indexados. A partir da média de citações que os artigos recebem, mede-se o FI do periódico onde são publicados. Quanto mais citações receberem os artigos de uma revista científica, maior será seu FI. Ainda que existam críticas à metodologia, este formato é basicamente aceito pela academia para comparar diferentes

¹³⁰ Informação verbal obtida em entrevista com entrevistado 08.

¹³¹ Informação verbal obtida em entrevista com entrevistado 05.

¹³² Informação verbal obtida em entrevista com entrevistado 07.

¹³³ Informação verbal obtida em entrevista com entrevistado 09.

periódicos. O mesmo critério, ainda que não mensurado matematicamente com essa precisão do FI, é aproveitado empiricamente pelos acadêmicos para a pesquisa bibliográfica cotidiana na procura por textos para suas pesquisas. A lógica é provavelmente conhecida e utilizada por qualquer pesquisador, mesmo aqueles iniciantes compartilham dessa prática.

Os entrevistados desta tese, todos acadêmicos de pós-graduação, relataram que dedicam atenção a determinados autores ou textos quando observam muitas citações a eles em outras leituras.

Geralmente [escolho livros], lendo algum outro livro em citações, [...] porque esse cita alguma coisa que não está especificado. Eu vou atrás de citações mesmo. Esse livro aqui [*mostra o livro*], por exemplo, muitos artigos citavam, eu tentei encontrar e não encontrei, entrei na Amazon e comprei. (informação verbal)¹³⁴

Bom, eu vou fazendo meio que uma corrente: eu leio um livro e leio as referências daquele livro, vejo o que me interessa, vejo se ele citou algum autor, e vou fazendo isso. Vou fazendo esse processo assim. (informação verbal)¹³⁵

Normalmente [os livros escolhidos] eram livros dos quais eu já tinha lido em algum lugar, eu já tinha conhecimento prévio. Já tinha visto em outros trabalhos, que na verdade são poucos na área, do sistema ainda, agora que está começando, agora que deu o “bum”. Mas, enfim, eu já pegava dali e vinha com a minha listinha pronta de títulos e autores, e sim o sistema de recomendação também me influenciou um pouco nesse sentido até pra me mostrar. (informação verbal)¹³⁶

Mas quem fala de livros da academia mesmo, são pessoas dos mesmos temas de estudo, conhecem os autores, aí a gente vai indicando, autor também é bom pra ter essa motivação também de ler. A gente vai lendo um, aí gostou vai procurar mais. (informação verbal)¹³⁷

Já tem os autores que eu gosto, às vezes ele dá aquele, quem comprou esse também comprou esse, dou uma lida, mas na verdade geralmente eu já sei o que eu quero comprar, mesmo que eu não saiba o título, eu já sei o autor. (informação verbal)¹³⁸

Assim, pelo pensamento associativo (BUSH, 1945), mentalmente, os pesquisadores carregam consigo uma bagagem de autores e de obras que empiricamente julgam de maior ou menor impacto para suas pesquisas. Ao produzirem seus textos, os acadêmicos também reconhecem a importância ao citar fontes, pois conferem-lhe crédito ao

¹³⁴ Informação verbal obtida em entrevista com entrevistado 01.

¹³⁵ Informação verbal obtida em entrevista com entrevistado 02.

¹³⁶ Informação verbal obtida em entrevista com entrevistado 03.

¹³⁷ Informação verbal obtida em entrevista com entrevistado 04.

¹³⁸ Informação verbal obtida em entrevista com entrevistado 05.

citarem seus trabalhos. Mesmo quando é para criticar um estudo, raramente observa-se um autor citar a obra de outro que não mereça devido crédito. Desse modo, a citação é vista como a moeda dentro da vida acadêmica e, inclusive, também determina o financiamento em pesquisas, visto que indicadores bibliométricos são por vezes utilizados para avaliar resultados de investimentos em pesquisa.

A explanação anterior não deseja discutir a forma como o impacto dos periódicos é mensurado ou como o financiamento de pesquisas é distribuído, mas sim apresentar como os pesquisadores comportam-se ao fazer uso das citações para a seleção de itens.

Com isso, quando utilizada para medir o FI de um periódico, a citação aparece como uma forma de recomendação objetiva, oriunda de interações reativas (PRIMO, 2007) . Se um determinado autor for muito citado, então o determinado autor será considerado como relevante na seleção. Se um livro for muito citado, então a obra é útil para leitura. Este aspecto matemático de quantificar as citações caracteriza-as como recomendações objetivas. Visto que se trata de uma questão de quantidade de citações, ressalva-se que não é uma questão de popularidade da obra, como nas recomendações por popularidade, já que na academia as obras citadas passam pelo crivo científico de quem as leu e as citou. Este aspecto é o que abre precedente para as citações também ocuparem espaço nas recomendações subjetivas.

Quando utilizada para quantificar o fator de impacto, a citação serve como recomendação objetiva, mas quando é vista pelo leitor como forma do autor do texto orientar à leitura de outro texto, pelos aspectos subjetivos que envolvem o ato de citar textos, a citação é observada como recomendação subjetiva. A citação neste aspecto pode inclusive potencializar interações mútuas (PRIMO, 2007), visto que o leitor do texto pode entrar em contato com o autor e questionar a menção feita ou, como é comum na academia, produzir textos científicos posteriormente à leitura para criticar a referência anterior. Com isso, as citações quando utilizadas para quantificar o FI são observadas como recomendações objetivas, mas quando potencializam interações mútuas caracterizam-se como recomendações subjetivas.

Ao percorrerem a lista de resultados para uma determinada pesquisa no site da Amazon, os usuários analisados nesta tese carregam consigo essa relação de autores conhecidos para selecionar ou não a obra de determinado autor. Quando seguem essa prática, os usuários usufruem de dois tipos de recomendações: a objetiva digital, feita no momento da constituição da lista a partir da recomendação sistêmica; e a subjetiva, formada na ocasião da leitura de artigos a partir de recomendações oriundas das citações bibliográficas. O contato

com os autores inclusive pode ocorrer de forma pessoal em congressos científicos ou nas salas de aulas de pós-graduação, o que caracterizaria ainda mais o caráter mútuo das recomendações subjetivas analógicas.

Os entrevistados relataram que as indicações pessoais para a seleção de itens provêm basicamente de amigos, colegas de sala de aula ou de grupos de pesquisa, de professores ou de contatos em congressos científicos. Para efeito de análise, essas recomendações subjetivas analógicas serão enquadradas em três elementos: amigos, conhecidos e especialistas.

Ao estudar como as pessoas se relacionam no capítulo 3, quando em 3.1 se abordou sobre relação e interação, observou-se que para haver relação é necessário que haja interação ao longo do tempo (HINDE, 1979), bem como o significado das coisas, os produtos sociais, são formados no contexto das interações sociais na perspectiva do interacionismo simbólico (BLUMER, 1986). Portanto, deve-se considerar esta perspectiva interacionista ao analisar as indicações pessoais que podem oferecer diferentes relações para os diferentes grupos de pessoas, visto que estas são oriundas de interações mútuas.

Os entrevistados demonstraram que hierarquizam as pessoas que fazem as indicações. Em relação às indicações provindas de amigos, relataram que existem aqueles com quem interagem mais e que possuem gostos semelhantes e que as indicações destas pessoas tem mais peso para recomendações de leituras de lazer do que aquelas dos amigos que não possuem a mesma afinidade.

[Para as leituras de lazer as indicações são de]: Pessoas que leram, gostaram, que sabem mais ou menos o meu estilo e dizem que eu vou gostar. Dai é uma ótima indicação. [...] geralmente quem gosta de ler sabe, a gente tem uma conversa, ai acaba discutindo sobre livros e ai tem essa de que sabe o teu perfil, então, não é toda pessoa que vai indicar. (informação verbal)¹³⁹

O aspecto em questão não está no quanto a pessoa é amiga uma da outra, mas no quanto uma tem afinidade cultural com a outra. Consequentemente, aquelas pessoas que têm mais compatibilidade entre si serão aqueles que farão as melhores indicações umas às outras.

Contudo, a amizade que pode refletir a compatibilidade necessária para as indicações de lazer, não será necessariamente importante para as leituras acadêmicas que são indicadas por conhecidos (colegas, contatos em congressos) e especialistas (professores, autores, editores). Nestas recomendações também existe uma hierarquia entre os conhecidos e

¹³⁹ Informação verbal obtida em entrevista com entrevistado 04.

os especialistas, visto que é dada maior importância às recomendações deste segundo grupo pela autoridade que professores e autores geralmente têm frente aos temas que recomendam. “Tem pessoas mais academicamente confiáveis do que outras, tanto professores como colegas” (informação verbal)¹⁴⁰.

Os entrevistados relatam que a importância das indicações dos colegas está mais nas recomendações das leituras clássicas da área, enquanto que o destaque às indicações de conhecidos está em conhecer novidades na área, já que as indicações dos colegas do grupo de pesquisa e de sala de aula tendem a repetir-se. Será no grupo de conhecidos que os pesquisadores encontrarão novos caminhos e recomendações bibliográficas desconhecidas. As indicações dos especialistas foram igualmente associadas mais às leituras clássicas da área, no caso dos professores, mas estas também são as recomendações que os entrevistados mais confiam para as leituras acadêmicas. Com a exceção de um dos entrevistados que relatou não hierarquizar as indicações de conhecidos com a dos professores, os outros foram veementes na posição de que a recomendação do orientador é a mais segura. “O que o orientador indica é lei” (informação verbal)¹⁴¹. Porém, o entrevistado em exceção ressaltou que isto ocorre porque, segundo ele, seu tema de pesquisa foge da área de especialização do orientador devido à originalidade no tema de pesquisa.

Desse modo, as recomendações pessoais para a seleção de itens acadêmicos constituem-se em interações mútuas (PRIMO, 2007), cujo grau de confiança dado a cada recomendação depende do propósito da indicação e do grupo que a pessoa que indica pertence. Como as pesquisas por obras acadêmicas feitas pelos entrevistados no site da Amazon, partem dessas indicações para posterior exploração das listas em que os itens resultantes são ordenados a partir das recomendações sistêmicas, observa-se o uso de mais de um tipo de recomendação na seleção de itens com fins acadêmicos.

O usuário parte de uma recomendação subjetiva analógica que pode provir de recomendações de amigos, conhecidos ou especialistas e passa a utilizar o sistema em busca dos descritores que, a partir de recomendações objetivas digitais, recupera uma lista de resultados para seleção. Porém, ressalta-se que esse cuidado para a seleção de itens é mais pontual para a busca por livros acadêmicos e que para a busca de livros para a leitura de lazer é naturalmente mais despreocupada. Ainda que relatem as indicações de amigos para este tipo

¹⁴⁰ Informação verbal obtida em entrevista com entrevistado 05.

¹⁴¹ Informação verbal obtida em entrevista com entrevistado 04.

de leitura, os entrevistados informaram que raramente recorrem às recomendações pessoais para a indicação de livros de lazer.

A busca por obras para a leitura de lazer apresenta um comportamento heterogêneo dos entrevistados. O primeiro aspecto a ser destacado é que nem todos utilizam o site para leituras de lazer, visto que alguns preferem visitar exclusivamente as livrarias físicas para explorarem itens para suas leituras. Entre os que utilizam o site também para o consumo de leituras de lazer, há os que buscam pelas histórias, personagens ou pelos autores preferidos e navegam pelos resultados como nas buscas por obras acadêmicas. Há os que exploram as listas de recomendações para o seu perfil de usuário ou aqueles que navegam pela página de entrada e selecionam itens. Neste tipo de seleção aparece ainda o uso das seções do site, como no caso em que o usuário observado visitou a seção ficção e explorou os resultados até passar por um autor conhecido e selecionar um item. Por último, aparece o uso das listas de livros mais vendidos e populares para a seleção de itens.

O que foi observado nesta forma de seleção é que os usuários estavam mais sujeitos às indicações do próprio sistema e que as recomendações por popularidade e personalizadas passaram a ser mais aproveitadas. Por esta prática, torna-se necessário analisar os aspectos comportamentais dos entrevistados que tocam na gestão do perfil do usuário e o gerenciamento da base de dados para aprimorar as recomendações para diferentes perfis frente a uma abundante base de dados.

5.2.3 Motivações e Uso do Sistema de Recomendações da Amazon

Os internautas da loja on-line da Amazon entrevistados para esta pesquisa relatam que a principal motivação para o uso do site é o acesso a publicações estrangeiras e a variedade de itens disponíveis em sua base de dados.

Porque de usar ele? Acho que é porque tem a variação mais completa assim, pra achar vários livros que não se acha em outro lugar. Na realidade eu acesso a Amazon só por livros que eu não acho no Brasil, se tem livros que eu acho no Brasil, até com o preço um pouco mais caro, eu prefiro [comprar] aqui porque a entrega é mais rápida. Mas a Amazon eu acesso porque é uma biblioteca grande, eu acesso acho que desde a metade da minha graduação, assim que eu comecei a ler coisas mais específicas. (informação verbal)¹⁴²

¹⁴² Informação verbal obtida em entrevista com entrevistado 01.

Pela variedade, em primeiro lugar, de encontrar fontes diversas tanto nacionais, agora que a Amazon está no Brasil. Mas, no início, era mais pra encontrar fontes internacionais. Por exemplo, na minha pesquisa que era sobre crowdfunding eu encontrei várias fontes interessantes em função disso, de ter esse acesso direto à biblioteca da Amazon. (informação verbal)¹⁴³

Principalmente pela variedade de livros: tem em inglês, em português, e tudo pelo teu aparelho, então facilita, porque eu não mexi nele até ganhar o Kindle. (informação verbal)¹⁴⁴

Bom, eu acesso em função da variedade de produtos que eles oferecem, a pesquisa é em função de livros, mas se tem algum CD de algum artista, eu acabo indo lá para procurar também. É mais em função da diversidade de produtos que eles oferecem no site americano. (informação verbal)¹⁴⁵

Observa-se neste ponto que uma das principais motivações para acesso ao site está relacionada precisamente a um dos principais paradigmas da gestão de informação contemporâneo, problema da recuperação de informação, que é a abundância de dados e gerenciamento do lixo eletrônico (DENNING, 1982). Em meio a essa variedade de livros disponíveis aos consumidores, torna-se necessário um sistema que vá além da busca de itens conhecidos, visto que é praticamente impossível à memória humana conhecer todas as obras de uma biblioteca. Como é defendido nesta tese, para atender às necessidades de informação do consumidor, a Amazon disponibiliza um avançado sistema de recuperação de informação para que seu usuário garimpe informações relevantes meio a essa diversidade de itens para resolver esse paradoxo, o qual ao mesmo tempo que coloca a variedade de itens como principal aspecto motivacional, tem na abundância desses dados o problema para a recuperação de informação.

A observação direta do uso do site permitiu visualizar que os entrevistados geralmente utilizam o sistema conectados à sua conta de perfil. Essa prática permite com que a Amazon consiga aprimorar o seu algoritmo a cada interação reativa (PRIMO, 2007) que o usuário mantém com o sistema ao alimentar implícita e explicitamente (HU, SMEATON e SUN, 2012) as anotações (GOLDBERG *et al.*, 1992) da base de dados para a geração e manutenção do perfil do usuário (MONTANER, LÓPEZ e ROSA, 2003). Baseado na experiência de navegação dos usuários, o sistema consegue filtrar os itens de informação de uma abundante base de dados e apresentar listas de resultados relevantes, ao aproveitar tanto as pesquisas inseridas no momento da recuperação de informação como as experiências

¹⁴³ Informação verbal obtida em entrevista com entrevistado 03.

¹⁴⁴ Informação verbal obtida em entrevista com entrevistado 04.

¹⁴⁵ Informação verbal obtida em entrevista com entrevistado 10.

anteriores de cada perfil. Como os usuários fazem pesquisas com seus perfis conectados, os resultados oferecidos são customizados e consideram as especificidades dos interesses de cada perfil. A prática dos usuários em utilizar o site conectado ao seu perfil permite que o sistema ofereça Recomendações Personalizadas do tipo Objetiva Digital. Na análise dos dados observou-se que, principalmente para as leituras de lazer, os usuários demonstram utilizar as recomendações da interface gráfica dinâmica, personalizadas e por popularidade, mas este uso também aparece para as leituras acadêmicas em um segundo plano.

Nas seleções por obras acadêmicas, os usuários costumam chegar até a loja da Amazon em busca de títulos e autores específicos, mas como observado anteriormente na análise da seleção de itens no site, os usuários tendem a explorar as listas de resultados e navegar por outros itens que não procuravam no momento. Ao selecionarem a própria página do livro específico que procurava, as recomendações que aparecem nesta página, como as compras conjuntas, tornam-se relevantes para exploração e ainda para o processo de decisão de seleção do usuário. O fato dos entrevistados geralmente estarem condicionados à busca de itens específicos nas seleções acadêmicas coloca as recomendações personalizadas como uma ferramenta de recuperação de informação de segundo plano, mas nem por isso deixa de ser considerado ao longo da experiência de navegação.

A ferramenta se apresenta útil para os usuários, principalmente para a busca de itens desconhecidos, quando a pesquisa por temas de áreas específicas resulta em novidades publicadas em suas áreas de conhecimento. Os entrevistados relatam que a Amazon serve para tomar conhecimento de novos autores e de novas obras dos autores já conhecidos, bem como das principais novidades em suas áreas de pesquisa. Comparam o site com o sistema de bibliotecas das universidades e informam preferir o sistema da Amazon pela variedade e precisão nas referências bibliográficas que acompanham as obras. Os entrevistados consideram a ferramenta inteligente e destacam as indicações resultantes das correlações feitas entre as experiências de navegação dos consumidores, principalmente as recomendações personalizadas do tipo Objetiva Digital no formato de “quem comprou este também comprou este”.

Esse sistema é muito recente, mas eu me lembro que a primeira vez que eu vi eu fiquei meio que chocado, porque é surpreendente, parece que o sistema pensa. Ele dá sugestões muito inteligentes na Amazon. Na maior parte eu acabo me instigando e tem muito a ver com o que tu estas buscando. (informação verbal)¹⁴⁶

¹⁴⁶ Informação verbal obtida em entrevista com entrevistado 09.

Ao serem solicitados para selecionarem uma recomendação a partir da página de uma obra anteriormente selecionada, os usuários percorreram a página e levaram em conta principalmente essas recomendações personalizadas, baseadas nas compras de outros usuários. Informaram que essas indicações servem ainda para reforçar a escolha como mais um elemento no processo decisório no caso de incerteza. Após percorrerem as listas e selecionarem um livro, quando chegam até a página do item, os usuários percorrem e analisam a descrição, as resenhas dos editores da Amazon, os detalhes do produto, as resenhas dos usuários e as recomendações correlatas. O uso das resenhas editoriais e sociais serão analisadas no próximo tópico, antes deve-se dar mais atenção a essas recomendações personalizadas do tipo Objetiva Digital que aparecem nas páginas das obras.

Como é defendido nesta tese, o internauta recorre a mais de um tipo de recomendação para a seleção de itens que envolvam maior custo pessoal. Os livros acadêmicos para os pesquisadores são importantes ferramentas de trabalho e se enquadram neste grupo de itens de alto custo pessoal. Desse modo, quando procura por livros acadêmicos na loja da Amazon, os usuários entrevistados demonstraram recorrer a mais de um tipo de recomendação e foi dado destaque às recomendações personalizadas baseadas no histórico de navegação e também nas compras de outros usuários no momento da seleção de itens.

Os usuários relataram ser comum retornarem ao site e explorarem as listas de recomendações baseadas em seus históricos de navegação quando estão em uma navegação exploratória, mas deram ainda mais destaque para as recomendações baseadas nas compras de outros usuários, pelo motivo de haverem investido dinheiro naquela obra e ainda por relacionarem o conteúdo do livro em análise com obras que já possuíam ou conheciam. Neste momento da seleção, prevalecem também as recomendações sociais reativas na forma de classificação por estrelas no processo decisório. A experiência de seleção de itens dos usuários analisados na observação direta evidencia como as recomendações objetivas digitais passam a influenciar significativamente no processo de seleção.

O potencial dessas interações reativas (PRIMO, 2007) quando aproveitadas pelo sistema de recomendação pode usar a experiência dos usuários para aprimorar cada vez mais as recomendações personalizadas. Anteriormente foi dito que os entrevistados compararam a Amazon com as bibliotecas de suas universidades. A vantagem do sistema da Amazon em relação às bibliotecas pode estar além do tamanho da base de dados. Ao explorar por itens nas bibliotecas universitárias, o usuário do sistema não usufrui desse tipo de recomendação personalizada baseada na experiência coletiva. A experiência dos usuários no site da Amazon

observado nesta tese indica um potencial dado, oriundo de interações reativas (PRIMO, 2007), que ainda não é apresentado pelas bibliotecas para seus usuários.

O sistema da biblioteca poderia correlacionar os empréstimos dos usuários para fazer recomendações de itens. Certamente há diferença entre vendas e empréstimos de livros, mas ainda assim a recomendação baseada na experiência coletiva ofereceria mais um elemento no momento da seleção de itens. A biblioteca, pelo público especialista que possui, poderia oferecer a possibilidade de filtrar as recomendações exclusivamente baseadas na experiência de empréstimos de professores e doutores.

Desse modo, observa-se nas recomendações do tipo Objetiva Digital o potencial que as interações reativas (PRIMO, 2007) podem ser úteis para a recomendação de itens. Como é resultante da interação reativa do próprio usuário com o sistema, o aspecto negativo das recomendações objetivas diz respeito ao ônus de interagir com o sistema. “Pressionar o botão errado pode resultar em uma experiência frustrante [...] especialmente para usuários de primeira viagem” (ROGERS, SHARP e PREECE, 2013). O usuário que segue a recomendação baseada na compra de outros usuários para selecionar uma obra é quem será afetado pela escolha correta ou não da seleção. O ônus da interação é um fardo que os usuários de sistemas digitais carregam consigo no momento de fazer suas escolhas na seleção de itens de informação. Enquanto em uma livraria física presencial, o cliente insatisfeito pode até mesmo devolver um item que foi encomendado erroneamente por outra pessoa, na loja virtual a seleção está exclusivamente ao encargo do próprio internauta.

O ônus de interagir com o sistema evidencia como os usuários precisam recorrer a mais de um tipo de recomendação para selecionarem seus itens e não ficarem exclusivamente sujeitos às recomendações objetivas digitais do sistema. O mesmo cuidado deve ser tomado nas indicações subjetivas digitais, pois mesmo que as recomendações não sejam resultantes de cálculos matemáticos e de algoritmos, como será tratado no tópico a seguir, o prejuízo de seguir recomendações subjetivas digitais erradas, sejam editores ou sociais, será do próprio usuário que segue a indicação ao selecionar os itens que consome.

5.2.4 Avaliação e Uso das Recomendações Editoriais e Sociais Mútuas na Amazon

Os usuários do website da Amazon analisados nesta tese apresentam um comportamento de seleção de itens que parte de recomendações subjetivas analógicas (amigos, conhecidos e especialistas), consideram a própria experiência anterior de leituras acadêmicas ou de lazer, até chegar às pesquisas na barra de buscas do site e à exploração da

lista de resultados recuperados a partir de recomendações sistêmicas. Quando chegam à página do item em avaliação para a seleção, os usuários passaram a considerar as informações da descrição e as resenhas editoriais e sociais, as quais serão analisadas nesta última seção deste capítulo.

As informações dispostas nos detalhes do produto apareceram como fatores importantes na seleção, no que diz respeito ao número de páginas da obra, ano de publicação, edição e editora. Observou-se que obras com um número menor de páginas eram preteridas pelos entrevistados e que a data de publicação e número de edição também serve como fator decisório na seleção dos itens, mas é a editora que publicou a obra que recebe maior atenção, visto que é considerada como recomendação editorial.

Às vezes eu associo com a editora, se as estrelas são baixas e a editora é boa eu vou verificar a informação. (informação verbal)¹⁴⁷

No site dos editores, por exemplo, a Associação Americana de Bibliotecários. É daí que vêm as bibliografias. Então eu vejo que tem um livro que me interessa sobre catalogação que é publicado pela Associação. Dificilmente eu compro direto deles, porque eles não fazem entrega aqui, aí eu vou no site da Amazon e compro. (informação verbal)¹⁴⁸

Os usuários informaram reconhecer as editoras que publicam obras de melhor qualidade, tanto para livros acadêmicos quanto para leituras de lazer, e que este aspecto faz diferença no momento da seleção. A relação construída entre os entrevistados com as editoras de sua preferência, ao longo do tempo, aparece como elemento no momento de escolha do item por um único detalhe da obra.

Após olhar os detalhes do produto, os usuários passam à leitura das recomendações editoriais, especificamente a descrição do livro e as resenhas dos editores. Os usuários observam a diferença entre a descrição do livro e as resenhas editoriais, em que a primeira deseja expor rapidamente o conteúdo da obra, enquanto a segunda faz uma crítica para apreciar seu conteúdo. Contudo, os usuários não diferem a posição do produtor das descrições, da posição do autor das resenhas editoriais, pois consideram que os autores são da própria Amazon ou da editora do livro. Logo, os produtores das recomendações editoriais para os entrevistados desta pesquisa, fazem parte todos da mesma organização.

As descrições e resenhas editoriais são observadas pelos entrevistados com maior peso na seleção de livros acadêmicos do que as resenhas sociais. Para o grupo observado, os

¹⁴⁷ Informação verbal obtida em entrevista com entrevistado 07.

¹⁴⁸ Informação verbal obtida em entrevista com entrevistado 09.

profissionais que produziram as descrições e as resenhas editoriais estão mais tecnicamente preparados para analisar metodicamente a obra.

[...] tem gente que eu não sei qual a procedência, o que o sujeito estuda, porque é que ele gostou ou porque é que ele não gostou, eu vejo e olho, mas não considero muito, como são livros específicos que eu vou buscar, mesmo que tenha uma avaliação baixa eu vou comprar, porque foi uma outra fonte que eu consegui pra comprar. (informação verbal)¹⁴⁹

Os entrevistados relataram desconhecem os resenhistas sociais e respectivas experiências na área apreciada para poderem confiar na análise de livros acadêmicos. Portanto, o know-how de quem recomenda é um fator importante, considerado pelos entrevistados para selecionarem as leituras para suas pesquisas, teses e dissertações.

Neste aspecto, as recomendações subjetivas digitais sociais mútuas aparecem em segundo plano, em relação às recomendações editoriais. A observação direta do uso do sistema pelos entrevistados evidenciou que na pesquisa por livros acadêmicos os usuários raramente percorrem as resenhas sociais, porém para as buscas por literatura de lazer esse comportamento mostrou-se heterogêneo.

Contudo, ainda que as recomendações sociais mútuas sejam mais aproveitadas na seleção de leituras de lazer, existe uma dicotomia na opinião dos entrevistados em relação às resenhas escritas pelos próprios internautas no website da Amazon. Alguns confiam nesse tipo de avaliação por carregar consigo a opinião subjetiva do próprio consumidor comum, mas há quem desconfie da credibilidade dessa análise exatamente por ter sido produzida por este consumidor que, como analisado anteriormente, não estaria apto tecnicamente para produzir a resenha, por não ser um especialista na área em questão.

Porém, ao serem questionados da influência das resenhas sociais no processo de escolha, os entrevistados relatam já terem consumido itens devido à leitura de análises positivas nessas recomendações sociais, mas o contrário não ocorreu, visto que não deixam de consumir livros exclusivamente por uma resenha negativa da obra. Alguns relatam que não deixariam de ler um livro apenas porque algum resenhista desconhecido não gostou, já que o seu gosto pode ser diferente.

¹⁴⁹ Informação verbal obtida em entrevista com entrevistado 01.

Eu penso que [as resenhas] não são muito boas, os de livros de lazer são ok, mas as de áreas específicas são mais coisas soltas sem ter muito que ver. [...] O problema é que geralmente ela é ampla demais, e não mostram se tem exatamente os pontos que tu está buscando. [...] Nunca deixei de comprar, mas já comprei por causa da resenha. Nesse caso era livro de lazer e eu estava procurando literatura nova e a resenha me seduziu e eu comprei. (informação verbal)¹⁵⁰

Essa observação apresenta a importância de se conhecer o autor das resenhas sociais, já que se for produzida por algum amigo ou especialista no tema da obra poderá ser levada em conta para a compra de um produto.

As resenhas aparecem como mais um elemento no processo decisório de escolha, mais expressivas na seleção de leituras de lazer, mas pouco pertinentes para livros acadêmicos, já que os entrevistados relatam serem muito abrangentes as críticas para avaliação científica. Por isso, também buscam por resenhas fora do site da Amazon, especificamente em periódicos científicos, onde os elementos técnicos científicos da área e de credibilidade do periódico estão em questão e foram sujeitas ao parecer dos pares.

As resenhas sociais, classificadas como recomendações Subjetivas Digitais, são fruto de interações com potencial mútuo. O que pode ser observado nessa experiência de uso do site é que as recomendações desse tipo serão mais relevantes se as interações passarem do potencial mútuo para interações mútuas (PRIMO, 2007), quando trocas recursivas tornam os envolvidos na interação conhecidos. O que se observa é que no aspecto das recomendações sociais mútuas os usuários não ficam sujeitos ao ônus da interação, já que tomam cuidado ao considerarem as resenhas sociais ao utilizarem diferentes tipos de recomendações para a seleção de itens.

Por fim, a experiência dos usuários entrevistados na seleção de itens de informação no website da Amazon mostra que o sistema influencia nas escolhas dos entrevistados principalmente quando utilizam o site como um sistema de recuperação de informação e aproveitam as recomendações objetivas digitais oriundas das interações reativas. A pesquisa por bibliografia acadêmica privilegia as recomendações subjetivas analógicas enquanto que a busca por leitura de lazer favorece as recomendações objetivas digitais.

A partir desta análise, chega-se a seguinte Matriz e Tipos de Recomendações de Itens de Informação relevantes na experiência de seleção de itens de informação no website da Amazon, apresentada na Tabela 7.

¹⁵⁰ Informação verbal obtida em entrevista com entrevistado 05.

Tabela 7 – Matriz e Tipos de Recomendações de Itens de Informação.

Recomendações	Objetivas	Subjetivas
Digital	<i>Interface Gráfica Dinâmica; Recomendações Personalizadas; Recomendações por Popularidade; Recomendações Sistêmicas; Recomendações Sociais Reativas.</i>	<i>Recomendações Editoriais; Recomendações de Obras Premiadas; Recomendações Sociais Mútuas. Recomendações mediadas de amigos, conhecidos e especialistas.</i>
Analógica	<i>Recomendações por Popularidade Impressas;</i>	<i>Recomendações de Amigos; Recomendações de Conhecidos; Recomendações de Especialistas. Recomendações Editoriais Impressas;</i>
Interação	Reativa	Mútua

Fonte: Próprio autor

A busca por novos itens acadêmicos considera na maior parte do tempo mais de um tipo de recomendação, e normalmente partem das recomendações subjetivas analógicas para posterior pesquisa na barra de busca do site onde são consideradas as recomendações objetivas digitais. Como observado na análise da experiência de seleção de itens dos usuários entrevistados, as recomendações objetivas digitais sistêmicas e as recomendações subjetivas digitais editoriais presentes no site da Amazon são as mais influentes no processo de seleção, após as recomendações analógicas de especialistas para a seleção de livros acadêmicos. Ainda que também sejam consideradas para as buscas acadêmicas, as recomendações objetivas digitais sociais reativas e recomendações subjetivas digitais sociais mútuas são mais frequentes e relevantes na seleção de livros destinados à leitura de lazer. As pesquisas por novos itens de lazer levam em conta ainda as recomendações objetivas digitais da interface gráfica dinâmica, por popularidade e recomendações objetivas analógicas por popularidade. Ainda que não exista um comportamento homogêneo dos internautas para a seleção de leituras de lazer, percebe-se que ao procurarem por esse tipo de leitura no website da Amazon estão mais suscetíveis exclusivamente ao sistema das recomendações do tipo objetivas digital. Portanto, ao classificar os tipos de recomendações e analisar a experiência de navegação dos internautas no website da Amazon, percebe-se que os usuários recorrem a mais de um tipo de recomendação quando a seleção envolve conteúdo significativo, no caso a busca por livros acadêmicos, enquanto seguem exclusivamente sistemas de recomendações automatizados quando buscam por item em que o custo pessoal diretamente aplicado é baixo ou inexistente.

6 CONSIDERAÇÕES FINAIS

A ascendente produção científica e literária contemporânea demanda o uso de eficientes sistemas de gestão de dados informáticos para oferecer ao usuário listas de referências relevantes com o objetivo de realizar, de forma eficaz, a recuperação de informação. A abordagem adotada nesta tese para informação a definiu como coisa para que dessa forma pudesse ser tratada como dados binários por sistemas informáticos. O tratamento digital da informação possibilita a rápida consulta por documentos em uma base de dados informatizada, quando o usuário conhece os dados específicos que procura. Contudo, como estudado na definição de recuperação de informação, o problema está fundamentalmente na limitação do usuário de possuir conhecimento da totalidade de documentos armazenados na base de dados do sistema. Com as bases de dados cada vez mais abundantes, eleva-se a indispensabilidade de eficazes sistemas de recuperação de informação para oferecer listas de conteúdos relevantes aos usuários, os quais desejam além disso separar os documentos conforme seu interesse. Dentre esses, considera-se o sistema de recomendação digital do website da Amazon como um eficaz sistema de recuperação de informação, cujo auxilia os usuários na tarefa de superar o paradigma da abundância de dados e de garimpar conteúdo relevante, especificamente, na seleção de livros para leituras acadêmicas e de lazer.

À medida que o objetivo geral e os específicos desta pesquisa são considerados neste último capítulo, pondera-se o sistema de recomendação da Amazon como um sistema de recuperação de informação, o qual tem a base de anotações alimentada implícita e explicitamente por interações reativas que, em conjunto a técnicas de filtragem colaborativa, gerencia o perfil do usuário para recomendar conteúdo relevante. O estudo do funcionamento do sistema foi ressaltado com devido espaço neste trabalho, mas o cerne desta pesquisa estava em seu uso. Por isso, a análise das interações mútuas também foi relevante para compreender **como os interagentes selecionam conteúdo sob a influência desses sistemas de recomendação digital**. Segue-se às considerações finais desta tese com os objetivos específicos, visto que estes contribuem para atender ao objetivo geral e responder o problema de pesquisa. Encerra-se o texto com as limitações encontradas e os caminhos abertos para trabalhos futuros.

Com relação ao objetivo de **distinguir as indicações de itens nos sistemas de recomendação digital a partir das interações dos usuários da Amazon**, considera-se que se avançou no estudo dos sistemas de recuperação de informação ao aproximar-se o campo da Comunicação ao da Informação quando se propôs a Matriz de Recomendações de Itens de

Informação. Os estudos da Informática são frequentemente voltados à eficiência do sistema e de seus algoritmos, ao passo que na presente tese o estudo foi centrado nas interações do usuário.

A definição e caracterização das Recomendações Objetivas Digitais, resultantes das interações reativas do usuário com o sistema, deu origem aos elementos da matriz baseados exclusivamente no histórico de navegação, assim denominados: Interface Gráfica Dinâmica, Recomendações Personalizadas, Recomendações por Popularidade, Recomendações Sistêmicas e Recomendações Sociais Reativas.

A análise do website permitiu observar que a apresentação da interface gráfica do usuário na Amazon é baseada na gestão do perfil do usuário. Considera-se essa interface como um tipo de recomendação objetiva digital, visto que o perfil é gerido exclusivamente por interações reativas com o objetivo de ofertar conteúdo relevante ao perfil em questão. A interface gráfica dinâmica da Amazon é considerada uma espécie de vitrine virtual customizada às preferências do usuário.

Nesta interface também estão presentes as ofertas de livros baseadas nas recomendações personalizadas, as quais consideram a experiência de navegação individual do usuário para oferecer conteúdo, como as indicações baseadas nos itens anteriormente visitados ou comprados pelo próprio usuário. A limitação considerada nesse tipo de recomendação é que existe a propensão à superespecialização das recomendações. O usuário passa a receber cada vez mais indicações de obras dentro do seu próprio universo de preferência, o que eleva o risco de causar uma espécie de filtro invisível para os conteúdos que o algoritmo reativamente exclui das listas de resultados. Porém, observa-se que essa limitação é superada quando o algoritmo passa a aproveitar igualmente a experiência coletiva para recomendar livros.

As recomendações personalizadas também ocorrem a partir da experiência coletiva dos usuários, quando são feitas recomendações baseadas nas compras conjuntas de outros usuários. O problema deste tipo de recomendação é que os itens em destaque tendem a receber mais atenção dos usuários e à medida que um livro é destacado ao lado do outro, existe a tendência dos dois serem selecionados cada vez mais e, sucessivamente, tornarem-se cada vez mais e mais escolhidos. Ainda assim, a experiência coletiva na seleção de itens na Amazon que resulta nessas recomendações personalizadas é considerada como uma espécie de curadoria baseada em interações reativas. O cruzamento dos dados baseados nas experiências individuais e coletivas, tratados conjuntamente pelo algoritmo, oferecem recomendações objetivas digitais baseadas em interações reativas relevantes pelos usuários do

site. Como os visitantes do site carregam consigo suas experiências anteriores de autores e livros que julgam ser relevantes, cada vez que selecionam um item em conjunto a outro alimentam a base de anotações do sistema de recomendação, o qual levará essa informação em conta para indicar itens a outros usuários.

Desse modo, visto que os usuários chegam ao site da Amazon com indicações previamente estabelecidas, os outros elementos que compõem a matriz baseados em interações mútuas também merecem consideração.

O estudo baseado nas interações mútuas resultou nos elementos da matriz que compõem as Recomendações Subjetivas Digitais, os quais titulados por: Recomendações Editoriais, Recomendações de Obras Premiadas, Recomendações de Personalidades e Recomendações Sociais Mútuas.

Diferentemente do que eram esperadas, as recomendações editoriais, as quais são produzidas pelos profissionais da Amazon, de revistas especializadas e das editoras, foram consideradas pelos usuários pesquisados mais importantes do que as recomendações sociais mútuas. Apesar de os usuários julgarem relevantes as resenhas da Amazon no processo de seleção de livros, elas aparecem como elemento secundário em relação às recomendações dos editores, especialmente na seleção de livros acadêmicos. Ainda assim, consideram-se as recomendações sociais mútuas significativas quando aproveitadas como elemento social para classificar a disposição dos itens nas listas de resultados.

Mesmo que os usuários relatem que o conteúdo das resenhas é menos relevante na decisão de selecionar itens, a interação com os elementos sociais do site contribui para ranquear melhor os livros. Como será considerado mais à frente, o número de resenhas, de votos no *star rating* e nas próprias resenhas são levados em conta na classificação dos itens. O livro que receber maior número de resenhas e de votos nas estrelas de classificação será melhor ordenado na lista de resultados recuperados.

Nessa conjuntura, consideram-se as resenhas mais importantes para ranquear as obras nas listas recuperadas do que a própria avaliação crítica feita pelos internautas. Releva-se que a consideração não descarta o conteúdo da resenha produzida pelos internautas, visto que os pesquisados inclusive relataram selecionar livros devido à indicação em resenhas, mas destaca a sua importância mais pelo aspecto reativo do que relativo da interação.

As recomendações editoriais, ao prevalecerem frente às recomendações sociais, apontam ainda a permanência e influência das interações do tipo um-para-muitos no processo de seleção de itens. Por mais que as tecnologias digitais contemporâneas ofereçam espaços para tipos alternativos ao modelo de comunicação de massa, considera-se nesse fenômeno

observado na seleção de itens no website da Amazon, a repetição no digital dos modelos antecessores a esse meio.

As interações reativas e mútuas também foram importantes para a análise das recomendações externas ao meio digital. A análise dessas recomendações tornou-se essencial para o estudo devido ao fato de que os usuários chegam ao site com conhecimento prévio de autores, temas, obras e/ou com indicações de outras pessoas. As indicações oriundas dessas experiências resultam nas Recomendações Objetivas Analógicas, provindas de interações reativas, e Subjetivas Digitais, derivadas de interações mútuas.

As Objetivas Analógicas são melhor representadas pelas Recomendações por Popularidade Impressas. A partir das entrevistas em profundidade com os usuários, consideram-se as recomendações impressas nos veículos de comunicação tradicionais mais relevantes àquelas disponíveis no website da Amazon, apesar da matemática para chegar aos itens populares ser a mesma no digital e no impresso. Acredita-se que o fato da recomendação no impresso ocupar espaço físico na diagramação do jornal e de envolver maior custo que no digital seja o responsável por esse fenômeno.

As Recomendações Subjetivas Analógicas são consideradas como as principais indicações na seleção de livros acadêmicos, representadas pelos seguintes elementos da matriz: Recomendações de Amigos, Recomendações de Conhecidos, Recomendações de Especialistas e Recomendações Editoriais Impressas.

Com base na análise das entrevistas, considera-se que as Recomendações de Especialistas são as mais confiáveis para a seleção de leituras acadêmicas, visto que estes possuem conhecimento técnico frente aos temas que indicam livros. As Recomendações de Conhecidos e de Amigos não necessariamente atendem a esse critério. Contudo, as Recomendações de Amigos são consideradas na seleção de livros para a leitura de lazer, já que estes conhecem melhor o perfil e as preferências para quem indicam itens. Porém, julgam-se essas recomendações pouco relevantes no comportamento de seleção de itens de lazer, visto que os usuários relatam raramente recorrerem a recomendações de amigos para esse tipo de leitura.

Considera-se que as Recomendações Subjetivas Analógicas são as que mais se caracterizam por interações mútuas, visto que são oriundas de conversas entre as pessoas no face a face, cujas tem mais chance de modificar o conhecimento dos interagentes envolvidos. Mesmo que as Recomendações Subjetivas Digitais, como as Recomendações Mediadas de Amigos, Conhecidos e Especialistas, também se caracterizem pela interação mútua, trata-se mais de um potencial mútuo do que uma interação realmente recursiva.

Ao distinguir as indicações de itens a partir das interações dos usuários, observa-se que a seleção de itens acadêmicos considerará elementos objetivos e subjetivos da matriz, enquanto a escolha por livros de lazer estará geralmente suscetível às recomendações objetivas. A forma como os usuários utilizam o sistema determinará essa relação dos elementos da matriz.

Ao atender ao objetivo de **investigar como os interagentes utilizam o sistema de recomendação digital da Amazon**, considera-se que as Recomendações Sistêmicas são as que mais influenciam no processo de seleção de itens de informação no website.

Os usuários da Amazon pesquisados utilizam com frequência a barra de buscas do site para recuperar itens com base nas palavras-chaves que representam os assuntos de interesse. Como concluído que os usuários costumam selecionar os primeiros itens apresentados, observa-se nesta prática a influência das listas recuperadas pelo sistema em resposta aos descritores inseridos na barra de buscas. Na observação direta do uso do website, a prática mostrou-se frequente tanto na busca de livros acadêmicos desconhecidos quanto na de leituras de lazer. Visto que para as leituras de lazer os usuários são mais suscetíveis às indicações do sistema, considera-se essas recomendações ainda mais influentes nesta utilização.

Para as leituras acadêmicas, essas recomendações são consideradas significativas no momento em que deixam de apresentar nas primeiras páginas obras que podem ser relevantes para o usuário. Porém, a potencial limitação é amenizada pela experiência de navegação observada no comportamento dos usuários analisados. Visto que existe maior preocupação com a qualidade do conteúdo nesta utilização, os usuários fazem uma navegação exploratória pelo site e garimpam os itens ao avaliarem outras informações. Quando visitam um item para análise, o usuário navega pela página em busca de informações que julgam importantes para a seleção e exploram sucessivamente as recomendações editoriais, as recomendações sociais reativas e as recomendações personalizadas. Considera-se neste processo de seleção as recomendações oriundas de interações reativas bem como as derivadas de interações mútuas. Conclui-se que após as recomendações subjetivas analógicas, seguem-se também as recomendações objetivas e subjetivas digitais para a seleção de itens acadêmicos, enquanto que para as leituras de lazer frequentemente as recomendações digitais são consideradas suficientes no processo de seleção. Desse modo, os parâmetros que ordenam os itens foram essenciais para compreender a influência das recomendações objetivas na seleção de itens de informação na Amazon.

Após a análise do website investigado nesta tese, considera-se que **os parâmetros conhecidos que destacam conteúdos no sistema de recomendação digital da Amazon** ponderam a popularidade dos itens, a correspondência dos termos de busca e os elementos sociais, bem como o gerenciamento do perfil do usuário. Conclui-se que a filtragem colaborativa, que aproveita a experiência coletiva dos usuários do site, contribui para a ordenação dos itens. Contudo, a popularidade dos itens e a correspondência dos termos de busca sobrepõem os elementos sociais. Logo, como apresentadas nesta tese, as técnicas de SEO podem influenciar na seleção de itens, especialmente nas leituras de lazer.

Em síntese, considera-se o sistema de recomendação digital do website da Amazon como uma ferramenta de recuperação de informação que, a partir de interações reativas e mútuas, oferece resultados relevantes aos seus usuários. Destacam-se nestes sistemas as recomendações objetivas oriundas das interações reativas que alimentam a base de anotações com dados a serem tratados pelo algoritmo, que considera, principalmente, a popularidade e a correspondência dos termos de busca, mas que também leva em conta os elementos sociais. A ressalva que se deve fazer com relação aos resultados é que diferentemente do sistema de recuperação de uma biblioteca, o website da Amazon visa lucro e que as obras podem receber destaque no site para incrementar as vendas. Contudo, como observado no **comportamento dos interagentes na seleção de itens frente ao sistema de recomendação digital do website da Amazon**, considera-se que esse cuidado é tomado especialmente para a seleção de itens acadêmicos, visto que os usuários recorrem a mais de um tipo de recomendação quando o custo pessoal envolvido é elevado.

A presente pesquisa analisou o comportamento de seleção de itens em um grupo de entrevistados que é especializado na seleção de bibliografia. O fato de todos os entrevistados serem pós-graduandos, por um lado tornou possível analisar a influência de um sistema de recomendação digital em um grupo especializada na seleção de livros acadêmicos, mas por outro, a pouca diversidade no recorte limita os resultados na seleção de livros em geral. Ainda que esse grupo também ofereça bons indícios para o comportamento neste tipo de seleção, o resultado em um recorte heterogêneo pode apresentar resultados diferentes especialmente para livros destinados à leitura de lazer.

O website da Amazon foi selecionado pelo conhecimento empírico que se possuía de seu sistema de recomendação digital considerado relativamente bom, mas como os próprios usuários pesquisados salientaram há sites em que o sistema de recomendação é precário. O próprio site brasileiro da Amazon ainda não possui uma base de dados suficiente para fazer boas recomendações. Por esse motivo, deve-se ter em mente que os resultados

desta tese consideram a análise de um sistema de recuperação de informação eficaz, enquanto que a análise em um sistema ineficaz pode oferecer comportamentos diferentes na seleção de itens, principalmente nas recomendações objetivas digitais.

As considerações feitas com relação às resenhas sociais nesta tese foram exclusivamente baseadas na experiência de internautas brasileiros. A aplicação da mesma pesquisa a um público de norte-americanos, por exemplo, pode oferecer um resultado mais favorável às recomendações sociais mútuas, visto que este tipo de recomendação é fundamentado principalmente no interacionismo simbólico, teoria em que a cultura é essencial na análise. Como os entrevistados navegaram essencialmente no website da Amazon norte-americana, o idioma e o afastamento da cultura daquele país pode ser um limitador para uso das recomendações sociais mútuas.

As limitações apresentadas nesta pesquisa servem como caminho para trabalhos futuros. A aplicação em um público heterogêneo e com uma mostra mais expressiva poderá oferecer outros indícios para expandir a Matriz de Recomendações de Itens de Informação. A aplicação desta matriz na análise de websites com sistemas de recomendações ineficientes pode apresentar os pontos que necessitam desenvolvimento para superar suas limitações, bem como podem expandir os próprios elementos da matriz.

Por fim, a principal sugestão desta pesquisa para trabalhos futuros diz respeito a aplicação da matriz em sistemas de recomendações de notícias, como do New York Times e do Portal Terra. Os dois websites de notícias oferecem desde o início de 2014 um sistema de recomendação digital de conteúdo baseado na experiência de navegação do usuário. A análise do comportamento dos leitores na seleção de notícias em websites, sob a influência dos sistemas de recomendação digital, pode resultar em rupturas sociais e tecnológicas na curadoria jornalística.

REFERÊNCIAS

- ADOMAVICIUS, G.; TUZHILIN, A. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. **IEEE Transactions on Knowledge and Data Engineering**, v. 17, n. 6, p. 734-749, 2005. Disponível em: < <http://dl.acm.org/citation.cfm?id=1070751> >. Acesso em: 13 ago. 2013.
- ANDACHT, F. A Representação do Self na obra de Goffman: sociosemiótica da identidade. In: GASTALDO, E. **Erving Goffman: Desbravador do Cotidiano**. Porto Alegre: Tomo Editorial, 2004. p.125-146.
- BAEZA-YATES, R.; RIBEIRO-NETO, B. **Recuperação de Informação: Conceitos e tecnologias das máquinas de busca**. 2. ed. Porto Alegre: Bookman, 2013. 590 p.
- BALABANOVIĆ, M.; SHOHAM, Y. Fab: content-based, collaborative recommendation. **Communications of the ACM**, v. 40, n. 3, p. 66-72, 1997. Disponível em: < <http://dl.acm.org/citation.cfm?id=245124> >. Acesso em: 13 ago. 2013.
- BARABÁSI, A.-L. S. **Linked: how everything is connected to everything else and what it means for business, science, and everyday life**. New York: Plume, 2003. 294 p.
- BAUDRILLARD, J. **The consumer society: Myths and structures**. Londres: Sage, 1998. 224 p.
- BLACKWELL, R. D.; MINIARD, P. W.; ENGEL, J. F. **Comportamento do Consumidor**. São Paulo: Cengage Learning, 2009. 606 p.
- BLUMER, H. **Symbolic interactionism: Perspective and Method**. Berkeley: University of California Press, 1986. 208 p.
- BRAGA, A. A. **Personas Materno-Eletrônicas: Feminilidade e Interação no Blog Mothern**. Porto Alegre: Sulina, 2008. 286 p.
- BRANDT, R. L. **Nos bastidores da Amazon: o jeito Jaff Bezos de revolucionar mercados com apenas um clique**. São Paulo: Saraiva, 2011. 176 p.
- BRECHT, B. Teoria do Rádio (1927-1932). In: MEDITSCH, E. **Teorias do Rádio: textos e contextos**. Florianópolis: Insular, v.1, 2005.
- BREESE, J. S.; HECKERMAN, D.; KADIE, C. Empirical analysis of predictive algorithms for collaborative filtering. **Anais da UAI'98 Fourteenth conference on Uncertainty in artificial intelligence** Madison, Julho de 1998. p.43-52.1998. Disponível em: < <http://research.microsoft.com/pubs/69656/tr-98-12.pdf> >. Acesso em: 13 ago. 2013.
- BUCKLAND, M. K. Information as Thing. **Journal of American Society for Information Science**, v. 42, n. 5, p. 351-360, 1991. Disponível em: < <http://skat.ihmc.us/rid=1KR7VC4CQ-SLX5RG-5T39/BUCKLAND%281991%29-informationasthing.pdf> >. Acesso em: 13 ago. 2013.

BUSH, V. As we may think. **Atlantic Monthly**, v. 176, n. 1, p. 101-108, 1945. Disponível em: < <http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/> >. Acesso em: 13 ago. 2013.

CAPLAN, S. Using focus group methodology for ergonomic design. **Ergonomics**, v. 33, n. 5, p. 527-533, 1990/05/01 1990. ISSN 0014-0139. Disponível em: < <http://dx.doi.org/10.1080/00140139008927160> >. Acesso em: 13 ago. 2013.

CASTELLS, M. **The Internet galaxy**: Reflections on the Internet, business, and society. New York: Oxford University Press, 2001.

_____. **A sociedade em rede**: A era da informacao : economia, sociedade e cultura. São Paulo: Paz e Terra, 2009. 698 p.

CONVILLE, R. L.; ROGERS, E. **The Meaning of “Relationship” in Interpersonal Communication**. Westport: Praeger, 1998. 204 p.

DENNING, P. J. Eltronic Junk. **Communications of the ACM**, v. 25, n. 3, p. 162 - 165, mar. 1982. Disponível em: < <http://dl.acm.org/citation.cfm?id=358453.358454&coll=DL&dl=ACM&CFID=214189238&CFTOKEN=95414675> >. Acesso em: 13 ago. 2013.

DUARTE, J. Entrevista em Profundidade. In: DUARTE, J. e BARROS, A. **Métodos e técnicas de pesquisa em comunicação**. São Paulo: Atlas, 2005. p. 62-83.

DUARTE, M. Y. M. Estudo de Caso. In: DUARTE, J. e BARROS, A. **Métodos e técnicas de pesquisa em comunicação**. São Paulo: Atlas, 2005. p. 213-235.

ENZENSBERGER, H. M. **Elementos para uma teoria dos meios de comunicação**. Rio de Janeiro: Tempo Brasileiro, 1978. 120 p.

FAIRTHORNE, R. A. **Towards information retrieval**. Londres: Butterworths, 1961. 211 p.

FEATHERSTONE, M. **Cultura de consumo e pós-modernismo**. São Paulo: Studio Nobel, 1995. 225 p.

FELINTO, E. ‘Materialidades da Comunicação’: Por um Novo Lugar da Matéria na Teoria da Comunicação. **Ciberlegenda**, n. 5, 2001. Disponível em: < <http://www.uff.br/ciberlegenda/ojs/index.php/revista/article/view/308> >. Acesso em: 13 ago. 2013.

FLICK, U. **Desenho da pesquisa qualitativa**. Porto Alegre: Artmed, 2009. 164 p.

GILLMOR, D. **We the media: Grassroots journalism by the people, for the people**. O'Reilly Media, Inc., 2006.

GLEICK, J. **A informação**: Uma história, uma teoria, uma enxurrada [Edição Kindle]. Brasil: Companhia das Letras, 2013. 541 p.

GOFFMAN, E. **A Representação do Eu na vida Cotidiana**. 9 ed. Petrópolis: Vozes, 1985. 236 p.

GOLDBERG, D. *et al.* Using collaborative filtering to weave an information tapestry. **Communications of the ACM**, v. 35, n. 12, p. 61 - 70, dec. 1992. Disponível em: < <http://dl.acm.org/citation.cfm?id=138867> >. Acesso em: 13 ago. 2013.

GOLDBERG, E. Electronic Document Retrieval, And Vannevar Bush's Memex. **Journal of the American Society for Information Science**, v. 43, n. 4, p. 284-294, 1992. Disponível em: < <http://onlinelibrary.wiley.com/doi/10.1002/%28SICI%291097-4571%28199205%2943:4%3C284::AID-ASI3%3E3.0.CO;2-0/abstract> >. Acesso em: 13 ago. 2013.

HINDE, R. A. **Towards understanding relationships**. Londres: Academic Press, 1979. 367 p.

HU, F.; SMEATON, A. F.; SUN, Y. **An image retrieval system based on explicit and implicit feedback on a tablet computer**. iHCI 2012: Irish Human Computer Interaction Conference, 2012 of Conference., Galway, Ireland. Edition, June 2012. p.20-21. Disponível em: < <http://doras.dcu.ie/17081/> >. Acesso em: 13 ago. 2013.

LATOURE, B. **Reassembling the social : an introduction to actor-network-theory**: Clarendon lectures in management studies. Oxford ; New York: Oxford University Press, 2005. x, 301 p.

LEMOES, A. Anjos interativos e retribalização do mundo: sobre interatividade e interfaces digitais. **Signo Revista de Comunicação**, v. 3, n. 5, p. 26-42, 1998. Disponível em: < <http://www.facom.ufba.br/ciberpesquisa/lemos/interativo.pdf> >. Acesso em: 13 ago. 2013.

_____. **Cibercultura: tecnologia e vida social na cultura contemporânea**. 2. ed. Porto Alegre: Sulina, 2002.

LEVINE, R. *et al.* **O manifesto da economia digital : o fim dos negócios como nós conhecemos**. Rio de Janeiro: Campus, 2000. 190 p.

LIEBSCHER, P. Quantity with quality? Teaching quantitative and qualitative methods in an LIS masters program. **Library Trends**, v. 46, n. 4, p. 668, Spring98 1998. ISSN 00242594. Disponível em: < <http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=948106&lang=pt-br&site=ehost-live> >.

LIN, A. C. Virtual Consumption: A Second Life for Earth? **Brigham Young University Law Review**, v. 1, p. 47-114, 2007.

MANNING, C. D.; RAGHAVAN, P.; SCHÜTZE, H. **Introduction to information retrieval**. New York: Cambridge University Press, 2008. xxi, 482 p.

MELO, J. M. D. **A opinião no jornalismo brasileiro**. Petrópolis: Editora Vozes, 1994. 208 p.

MONTANER, M.; LÓPEZ, B. L.; ROSA, J. L. D. L. A Taxonomy of Recommender Agents on the Internet. **Artificial Intelligence Review**, v. 19, n. 4, p. 285-330, 2003. Disponível em: < <http://link.springer.com/article/10.1023/A:1022850703159> >. Acesso em: 13 ago. 2013.

MONTARDO, S. P.; ARAÚJO, W. Performance e práticas de consumo online: ciberativismo em sites de redes sociais. **Revista FAMECOS-Mídia, Cultura e Tecnologia**, v. 20, n. 2, 2013.

MOOERS, C. N. Coding, Information Retrieval, And The Rapid Selector. **American Documentation**, v. 1, n. 4, p. 225-229, 1950a. Disponível em: < <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=16875315&lang=pt-br&site=ehost-live> >. Acesso em: 13 ago. 2013.

_____. **The theory of digital handling of non-numerical information and its implications to machine economics**: (Zator technical bulletin). Boston: Zator Co., 1950b. 17 p.

_____. Zato coding Applied to Mechanical Organization Of Knowledge. **American documentation**, v. 2, n. 1, p. 20-32, 1951. Disponível em: < <http://onlinelibrary.wiley.com/doi/10.1002/asi.5090020107/pdf> >. Acesso em: 13 ago. 2013.

_____. Mooers' Law or why some retrieval systems are used and others are not. **Bulletin of the American Society for Information Science**, v. 23, n. 1, p. 22-23, 1996. Disponível em: < <http://onlinelibrary.wiley.com/doi/10.1002/bult.37/pdf> >. Acesso em: 13 ago. 2013.

NETO, E. C.; FLORIDIA, G. Circuitos online de comunicação relacionada ao consumo. **Comunicação, Mídia e Consumo**, v. 5, n. 12, p. 113-129, 2008.

NIELSEN. **Personal recommendations and consumer opinions posted online are the most trusted forms of advertising globally**. The Nielsen Company. New York: abril. 2009. Disponível em: < http://www.nielsen.com/content/dam/corporate/us/en/newswire/uploads/2009/07/pr_global-study_07709.pdf >. Acesso em: 10 dez. 2009.

O'REILLY, T. What is Web 2.0? **O'Reilly Media**, 2005. Disponível em: < <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html> >. Acesso em: 13 ago. 2013.

OTLET, P. **Traité de documentation**: le livre sur le livre: théorie et pratique. Bruxelas: Mundaneum, 1934.

PARISER, E. **O filtro invisível**: o que a internet está escondendo de você. Rio de Janeiro: Zahar, 2012. 250 p.

PATTON, M. Q. **Qualitative evaluation and research methods**. 2ª ed. Thousand Oaks, CA, US: Sage Publications, Inc, 1990. 532 p.

PRIMO, A. F. T. **Interação Mediada por Computador**: comunicação, cibercultura, cognição. Porto Alegre: Sulina, 2007.

RALPH, S. R. The Rapid Selector. **Journal of Documentation**, v. 5, n. 3, p. 164-171, 1949. Disponível em: < <http://www.emeraldinsight.com/journals.htm?articleid=1649390&show=abstract> >. Acesso em: 13 ago. 2013.

RECUERO, R. D. C. Information Flows and Social Capital in Weblogs: A case study in the Brazilian blogosphere. **Hypertext'08**, Pittsburgh. 2008.

RESNICK, P.; VARIAN, H. R. Recommender Systems. **Communications of the ACM**, v. 40, n. 3, p. 56 - 58, mar. 1997. Disponível em: < <http://dl.acm.org/citation.cfm?doid=245108.245121> >. Acesso em: 13 ago. 2013.

RICH, E. User modeling via stereotypes. In: MAYBURY, M. T. e WAHLSTER, W. **Readings in intelligent user interfaces**. San Francisco: Morgan Kaufmann Publishers Inc., 1998. p. 329-342 Disponível em: < <http://www.sciencedirect.com/science/article/pii/S0364021379800129> >. Acesso em: 13 ago. 2013.

ROGERS, Y.; SHARP, H.; PREECE, J. **Design de Interação: Além da Interação Humano Computador**. Porto Alegre: Bookman, 2013. 585 p.

ROYO, J. **Design Digital**. São Paulo: Rosari, 2008. 169

SANSEVIERI, P. C. **How To Sell Your Books By The Truckload On Amazon.com**. Amazon Digital Services, Inc., 2013.

SHANNON, C. E.; WEAVER, W. **The mathematical theory of communication**. Urbana: University of Illinois Press, 1949. v (i.e. vii), 117 p.

SUROWIECKI, J. **A sabedoria das multidões**. Rio de Janeiro: Record, 2006. 375 p.

THOMPSON, J. B. **A mídia e a modernidade: uma teoria social da mídia**. 6 ed. Petrópolis: Vozes, 1998.

VEJA, edição nº780. São Paulo, p.56, 17 agosto. 1983.

WAL, T. V. Folksonomy Coinage and Definition. **vanderwal.net**, 2007. Disponível em: < <http://vanderwal.net/random/category.php?cat=153> >. Acesso em: 13 ago. 2013.

WEINBERGER, D. **A nova desordem digital**. Rio de Janeiro: Elsevier, 2007. 273 p.

WHEELER, J. A. It from bit. In. **At home in the Universe**. Nova York: American Institute of Physics, 1994. p. 276-368.

WHITE, D. M. The gatekeeper: A case study in the selection of news. **Journalism quarterly**, v. 27, n. 4, p. 383-390, 1950.

WISE, C. S.; PERRY, J. W. Multiple coding and the rapid selector. **American Documentation**, v. 1, n. 2, p. 76-83, 1950. Disponível em: < <http://onlinelibrary.wiley.com/doi/10.1002/asi.5090010203/abstract> >. Acesso em: 13 ago. 2013.

ANEXO A – MODELO DO TERMO DE CONSENTIMENTO

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM COMUNICAÇÃO E INFORMAÇÃO

Termo de Consentimento Livre e Esclarecido

Esta é uma pesquisa que investiga o comportamento do entrevistado durante a seleção de livros em websites.

Você está sendo convidado a participar de uma entrevista pessoal e, assim, integrar o universo da pesquisa. Para isto, é requerida a autorização prévia para gravação das mesmas mediante assinatura deste documento.

Nesta entrevista você será questionado por cerca de 40 minutos sobre o seu uso do site da Amazon. O entrevistador fará perguntas semiestruturadas a partir de um roteiro previamente preparado. As perguntas dizem respeito a sua navegação e seleção de itens no site da Amazon. À medida que você responde às perguntas, o entrevistador anotará informações que se destacam em suas respostas e também gravará o áudio da entrevista para posterior transcrição e análise. Quando necessário, o computador utilizado será do próprio pesquisador e este não terá acesso aos dados do seu computador.

As informações fornecidas pelo entrevistado visam à produção de conhecimento científico, contribuindo para o esclarecimento do tema investigado. As informações serão utilizadas somente sob propósito acadêmico para fins de consulta, análise e argumentação nesta pesquisa.

Sua participação é voluntária, isto é, a qualquer momento você pode recusar-se a responder qualquer pergunta ou desistir de participar e retirar seu consentimento. Sua recusa não trará nenhum prejuízo em sua relação com o pesquisador. Não haverá riscos de natureza física relacionados à sua participação.

O seu nome não será divulgado no texto da análise da pesquisa. Para garantir isso, quando os dados relacionados a você forem citados no texto, o seu nome será substituído por um número, no seguinte formato: entrevistado #.

A entrevista será gravada em suporte digital em áudio para posterior transcrição e será guardada por três (03) anos em uma mídia de CD. Após este período, todo o material resultante da pesquisa será incinerado.

Eu, _____, declaro ter recebido informação sobre esta pesquisa, concordo em participar voluntariamente e autorizo a gravação da entrevista.

Porto Alegre, ____ de _____ de 2014.

Assinatura do (a) participante

Assinatura do pesquisador

Esta pesquisa está registrada no Comitê de Ética em Pesquisa da UFRGS, telefone (51) 3308-3738

GLOSSÁRIO

- ⁱ Login – ato de conectar-se com sua conta/perfil a um computador, base de dados ou sistema.
- ⁱⁱ Cookies: pequenos pedaços de dados enviados por um website e armazenados no navegador do internauta enquanto o internauta navega pelo website visitado.
- ⁱⁱⁱ Log – arquivo que guarda informações de um programa que é utilizado para registrar os eventos relevantes de um sistema computacional.
- ^{iv} Star Rating – sistema de classificação, normalmente com cinco escalas, destinado à avaliação de itens.
- ^v Hyperlink – ligação de um arquivo de hipertexto para um outro local ou arquivo.
- ^{vi} Permalink – hiperlink estático e permanente para uma página web ou para uma postagem de blog.
- ^{vii} Geolocalização – funcionalidade de alguns dispositivos móveis que usa do localizador GPS para identificar a posição exata em que o aparelho se localiza.
- ^{viii} MARC – Do inglês *Machine Readable Cataloging* o format MARC significa catalogação legível por computador.
- ^{ix} Bit – palavra sugerida por John W. Turkey para definir uma unidade de informação condensada em dígito binário, ou seja, cada unidade assume dois valores, geralmente, 0 e 1 que representam falso e verdadeiro.
- ^x Tablet – Dispositivo móvel de computador geralmente com tela sensível ao toque e em uma única peça.
- ^{xi} Smartphone – Dispositivo móvel com funcionalidades de telefone com acesso à Internet de banda larga com funcionalidades avançadas como geolocalização, bússola, giroscópio, entre outras.
- ^{xii} Hashtag – etiqueta primeiramente utilizada com essa nomenclatura por usuários do Twitter com o objetivo de classificar mensagens. O termo surgiu a partir da soma do nome do símbolo sustenido, # hash em inglês, mais a palavra etiqueta, *tag* em inglês.
- ^{xiii} Mac OS X Spotlight – sistema de recuperação de informação presente na plataforma do sistema operacional Mac OS para recuperar dados do próprio armazenamento do computador do usuário.
- ^{xiv} Windows 8 File Explore – sistema de recuperação de informação presente na plataforma do sistema operacional Windows 8 para recuperar dados do próprio armazenamento do computador do usuário.
- ^{xv} Mail – aplicativo de gerenciamento de e-mail do sistema operacional Mac OS.

-
- ^{xvi} Outlook – aplicativo de gerenciamento de e-mail da Microsoft.
- ^{xvii} Google Scholar – sistema de buscas em bases de dados acadêmicas da Google.
- ^{xviii} AntiSpam – programa que aplica técnicas de filtragem colaborativa para evitar entrada de lixo eletrônico na caixa de entrada dos gerenciadores de e-mail.
- ^{xix} Phishing – ato de crackers ao tentar obter informações confidenciais dos usuários de forma ilegal a partir de e-mail com programas de vírus anexados ou links para websites falsos.
- ^{xx} Curtir – botão destinado à avaliação positiva de mensagens compartilhadas no site de rede social Facebook. Este botão pode ser adicionado diretamente na interface de outros sites.
- ^{xxi} g+ – botão destinado à avaliação positiva de mensagens compartilhadas no site de rede social Google Plus. Este botão pode ser adicionado diretamente na interface de outros sites.
- ^{xxii} Feed de Notícias – local destinado às atualizações de notícias de forma automatizada diretamente na interface do aplicativo ou da página de acesso do usuário.
- ^{xxiii} Wish list – lista de desejos do usuário mantida no website e adicionada por botões nos itens.
- ^{xxiv} Slider – Banner dinâmico disponível nos sites em formato de banner.