



**A ORGANIZAÇÃO DO CONSUMO DE MÚSICA NA INTERNET
ATRAVÉS DA CLASSIFICAÇÃO DO GOSTO CULTURAL:
ESTUDO DE CASO DO SISTEMA DE RECOMENDAÇÃO
LAST.FM**

Rose Marie Santini

Instituto Brasileiro de Informação Ciência e Tecnologia - IBICT /Universidade Federal Fluminense - UFF, Brasil, mariesantini@gmail.com

Rosali Fernandez de Souza

Instituto Brasileiro de Informação Ciência e Tecnologia - IBICT, Brasil, rosali@ibict.br

Juan C. Calvi

Facultad de Comunicación, Universidad Rey Juan Carlos, España, juancarlosjamon.calvi@urjc.es

RESUMO:

A presente comunicação analisa a organização do consumo musical na Internet a través da classificação dos gostos culturais. Este processo se está intensificando por meio da implementação dos chamados Sistemas de Recomendação (SR), os quais classificam e organizam o consumo cultural recomendando produtos culturais com critérios comerciais baseados nos gostos e preferências culturais dos seus usuários. A principal estratégia dos SR é difundir música sem mensagens comerciais; e para legitimar a recomendação de seus produtos musicais utilizam técnicas que implicam a colaboração dos próprios usuários destes sistemas. Tomamos como caso de estudo o sistema Last.fm, atualmente um dos SR mais populares entre os usuários da Internet.

ABSTRACT:

This paper analyses the organization of music consumption in Internet through the classification of cultural tastes. This process is taking place through the so called Recommender Systems (RS), which classify and organize the cultural consumption by recommending products applying commercial criteria based on users cultural preferences. The main strategy of the RS is to play music without commercial advertisement. In order to legitimate their recommendation of music products they use techniques that involve the collaboration among their users. We take as an example the web site LastFm, one of the most popular RS between the Internet users.

PALABRAS CLAVES:

Internet, Sistemas de Recomendação, Classificação do gosto, Consumo de Música.



INTRODUÇÃO

Os Sistemas de Recomendação (SR) são sistemas informáticos de classificação, organização e recomendação de produtos culturais nas redes digitais, baseados nas preferências dos usuários. A recomendação na web, a partir dos algoritmos dos SR, demonstra potencialidades de uso em situações onde a decisão do usuário é uma questão de gosto e está diretamente relacionada com a recomendação de produtos culturais

A escolha destes sistemas como objeto de estudo se deve às polêmicas sobre os processos de mediação que se configuram no entorno digital. Nos bastidores do discurso sobre a “desintermediação” geral dos mercados através da relação direta com os usuários, encontramos novos intermediadores que concentram e direcionam tanto o tráfico como os usos da Internet, assumindo novas funções na orientação do consumo de produtos culturais e também na organização dos mercados de bens simbólicos. Neste contexto, o surgimento dos SR destaca-se como uma tendência que pode tornar-se dominante entre os modelos mediação de bens simbólicos que se configuram na web, através das técnicas para Recuperação da Informação da Música (RIM).

A Recuperação da Informação da Música (*Music Information Retrieval*) em Sistemas de Recomendação (*Recommender Systems*) na Internet configuram um tema novo para a Ciência da Informação. Os artigos, periódicos, livros-texto, relatórios de pesquisa e bibliografia em geral encontrados tanto na área de Recuperação da Informação da Música como sobre os Sistemas de Recomendação (SR) estão inseridos principalmente no domínio da Computação e a grande maioria em inglês. É escassa a bibliografia sobre o tema no campo das Ciências Sociais, principalmente em português e espanhol.

As considerações acima demonstram, por um lado, que o tema é relevante e inovador para área da Ciência da Informação, e por outro, aponta para grandes desafios de pesquisa.

O SURGIMENTO DOS SISTEMAS DE RECOMENDAÇÃO NA INTERNET E SUAS CARACTERÍSTICAS

Os Sistemas de Recomendação (SR) caracterizam-se como sistemas informáticos de classificação, organização e recomendação de produtos culturais, baseados nas preferências e gostos dos usuários de Internet. Os SR são aplicações de software que pretende dar suporte ao usuário em sua tomada de decisão quando este interage com espaços que oferecem grande quantidade de informação. Os SR recomendam itens baseados no interesse que o usuário expressa explicita ou implicitamente em sua navegação na web.



Os desenvolvedores do sistema de recomendação Tapestry em 1992 (GOLBERG, NICHOLS, & OKI, 1992) criaram a expressão “Filtragem Colaborativa” (*Collaborative Filtering*) para definir estes sistemas visando designar um tipo de software específico no qual a filtragem de informação é realizada com o auxílio humano, ou seja, pela colaboração entre os grupos de interessados e usuários. Posteriormente, esta terminologia passou a ser adotada por muitos para referir-se a estes sistemas.

A Filtragem Colaborativa (FC) não nasce com os SR, mas resulta dos primeiros sistemas de filtragem de informação (*information filtering*) onde a informação relevante é trazida para a atenção do usuário através de padrões de observação de comportamentos prévios para construção de seus perfis. Inicialmente, estes sistemas foram incapazes de ajudar a exploração da web e apresentaram problemas iniciais com os novos usuários para alavancar tendências antes que a filtragem se efetivasse.

A demanda por tecnologias de Filtragem de Informação não é algo novo. Peter Denning (apud LOEB & TERRY, 1992) escreveu em 1982 sobre a preocupação dos engenheiros de computação no que se refere à quantidade de informação que estava sendo gerada pelos diversos tipos de sistema e recebida pelos usuários. O autor defendia que toda a atenção deveria estar concentrada na geração da informação dos usuários para suprir as suas necessidades. Por tanto, diversos autores referem-se a estes sistemas como “sistemas de filtragem colaborativa”.

Utilizaremos o termo mais geral “Sistemas de Recomendação” (RESNICK & VARIAN, 1997) por três razões. Primeiro, porque verificamos que o funcionamento e o desenvolvimento dos SR não necessariamente incluem a colaboração explícita dos recomendadores com os receptores, que podem ser desconhecidos uns para os outros. Segundo, as recomendações podem sugerir itens de interesse muito particulares, ademais de indicar aqueles que poderiam ser removidos ou desconsiderados. E por último, veremos que estes sistemas passaram a incorporar outras técnicas e foram desenvolvidas tecnologias para a recomendação baseada na classificação de itens por conteúdo e filtragem por associação de perfis de usuários, o que amplia o próprio conceito de SR.

Segundo Belvin e CROFT (1992), a Filtragem de Informação é o nome utilizado para descrever uma variedade de processos que envolvem a entrega de informação para as pessoas que dela necessitam. Com o tempo este termo tornou-se popular e artigos técnicos foram escritos descrevendo o emprego de filtragem em diversas aplicações, como e-mail, documentos eletrônicos em escritórios, entre outros. No entanto, a distinção entre a Filtragem de Informação e processos relacionados com a Recuperação de Informações nem sempre é clara, pois ambas as expressões, “Recuperação de Informação” (RI) e “Filtragem de Informação” (FI) descrevem soluções que procuram auxiliar na resolução de problemas referentes à sobrecarga de informações.



O conceito de Recuperação de Informação tradicionalmente envolve representação, armazenamento, pesquisa, índices, tecnologias específicas e busca de informação relevante, requerida a partir da necessidade de informação de um usuário humano para recuperação de documentos textuais (INGWERSEN, 1992).

Para que a recuperação realmente ocorra, o usuário deve descrever sua necessidade de informação. A partir deste momento, o sistema busca associar a necessidade explicitada com os documentos armazenados. Este tipo de abordagem tende a manter uma base com característica estática no armazenamento das informações, e o início da interação é provocado pelo usuário. A eficácia do sistema está relacionada com a obtenção de um ou mais documentos úteis em relação à necessidade de informação do usuário (REATEGUI & CAZELLA, 2005; LANCASTER, 1968; VICKERY, 1970).

A Recuperação de Informação inclui a noção de precisão (*precision*) e revocação (*recall*). Diante dos critérios de desempenho dos sistemas de Recuperação da Informação, o poder de revocação é a habilidade do sistema de descobrir documentos relevantes, sendo este o atributo mais importante do sistema para o usuário. Porém, sempre é possível obter um máximo de revocação recuperando todos os documentos indexados. Assim, os parâmetros de revocação estão relacionados com a performance de precisão, que corresponde ao número de documentos relevantes recuperados dividido pelo número total dos documentos recuperados (LANCASTER, 1968).

Ao comparar com a Recuperação de Informação (RI), a Filtragem de Informação (FI) possui semelhanças e diferenças. Esta abordagem aqui não fica claro a qual delas você se refere – a RI ou a FI - é só uma questão de redação.]geralmente cria ou mantém um perfil de interesses do usuário. A noção de FI [Filtragem de Informação] tem como característica distintiva um maior tempo de duração no que tange à associação ou combinação de interesses (FOLTZ & DUMAIS, 1992). Ou seja, a FI não se refere a uma necessidade específica do usuário, mas sim às suas preferências, enquanto a RI baseia-se na percepção de uma necessidade de informação momentânea. A filtragem deve ser aplicada a cada novo item adicionado, procurando verificar se este atende ou não ao perfil de interesse do usuário (REATEGUI & CAZELLA, 2005).

Cabe destacar que os SR utilizam ambas as abordagens: tanto a RI como a FI, e inclui os conceitos de precisão e revocação. Pode-se dizer que o método da Filtragem de Informação pretende personalizar os Sistemas de Recuperação da Informação, e de alguma forma, os SR representam uma extensão lógica da busca pelo desenvolvimento de máquinas pessoais e de buscadores personalizados. A FI trabalha com o conceito de descoberta de informação a partir do perfil específico de cada usuário, o que também está relacionado com o conceito de Disseminação Seletiva da Informação (*Selective Dissemination of Information*) originalmente empregado no âmbito de sistemas de recuperação de itens em acervos de documentos.



A Disseminação Seletiva da Informação (DSI) se aproxima dos SR na medida em que propõe um serviço personalizado que oferece diretamente ao usuário listas de conteúdos sobre a sua área de interesse e de acordo com seu perfil. É um serviço que permite a recuperação de informação personalizada conforme as perspectivas e necessidades do usuário e que explora a priorização de interesses para um grupo/indivíduo. O processo da DSI envolve uma série de fatores que contribuem para que o usuário gaste menos tempo com o exame e a seleção de conteúdos, característica importante dos serviços oferecidos pelos SR (CHENEY, 1971; GAFFNEY, 1973; LEGGATE, 1975 apud LONGO, 1978).

Os SR são mecanismos para automatizar um método fundamental na perspectiva do usuário, que é a descoberta de informação (*Information Discovery*), e é neste aspecto que apresentam inovações relevantes. O interesse desta pesquisa é menos pelos algoritmos usados para construir os SR e sim por questões mais conceituais, como por exemplo: 1) indagar sobre a forma como se está modelando a arquitetura dos sistemas dentro de uma larga escala de descoberta de informação, particularmente em um ambiente altamente distribuído (que se baseia em padrões e protocolos para intercâmbio de dados); 2) como os SR podem interagir e complementar outros métodos para a descoberta de informação de acordo com interesses individuais, dentro do contexto social destes modelos; e 3) como os SR podem prover o discernimento e a legitimação dos gostos dentro do processo social de descoberta e disseminação informação.

Descoberta de Informação nos Sistemas de Recomendação

Os modelos clássicos de descoberta ou recuperação de informação são baseados na busca (*information seeking*). A representação da informação ou do item procurado pelo usuário é comparada com a indexação dos itens armazenados na base de dados - recuperação baseada em conteúdo (*retrieval based-context*) – e/ou metadados que descrevem os itens nas bases de dados – recuperação de estilo bibliográfica (*bibliographic style retrieval*).

Os objetos semelhantes são retornados, freqüentemente, em algum tipo de ordem baseado em um ranking de proximidade e similitude entre os itens. Uma característica importante do processo de RI é que a resposta a uma dada consulta (*query*) ocorre apenas entre o usuário e os itens pesquisados, e de forma indireta incluir os indivíduos envolvidos na indexação e descrição dos metadados dos itens disponíveis nos sistemas. A descoberta de informação em sistemas de RI não demanda outros tipos de informação sobre os usuários além da sua necessidade pontual de informação (LYNCH, 2001).

Os SR tentam a automatizar aspectos do modelo de descoberta de informação, porém a partir dos perfis dos usuários, e se baseiam em três conceitos para criar os métodos de recomendação. São eles:



1. Modelo baseado no gosto
2. Modelo baseado na reputação
3. Modelo baseado na popularidade

O modelo de descoberta de informação baseado no gosto corresponde aos SR que funcionam baseados na semelhança ou proximidade de gosto a fim de permitir que o sistema organize o encontro de pessoas com gostos similares e assim possam pedir sugestões uma às outras sobre novidades dentro de um domínio e gosto específico. Este é o modelo constituinte dos primeiros SR. Atualmente, muitos destes sistemas operam de uma forma ligeiramente diferente: eles não revelam as identidades das pessoas (com o objetivo de preservar a privacidade e o anonimato do usuário) e apenas passam a recomendação derivada das opiniões ou ações de outros com gostos similares, sem identificação. As outras duas abordagens começaram a emergir no funcionamento dos SR na última década, baseadas no gerenciamento da reputação e na popularidade dos itens disponíveis, que constitui o segundo modelo mencionado, que é baseado na reputação ou no ranking.

Os sistemas baseados na reputação (segundo modelo) focam na identificação de pessoas que o usuário considera como referência e utiliza a opinião ou ação destes indivíduos selecionados para fundamentar ou estruturar as recomendações. O modelo baseado na reputação captura um aspecto diferente da sociabilidade baseada na descoberta de informação, onde os usuários fazem julgamentos ou avaliações de uma forma mais geral, ao invés de comparar o comportamento de uns diretamente com o dos outros. Este modelo de recomendação se origina em um antigo processo social que envolve a mediação de revisores ou “líderes de opinião” na recomendação.

A terceira abordagem dos SR lida com a questão da popularidade. Neste caso, os usuários utilizam o sistema para saber sobre os itens mais populares de uma determinada comunidade e encontrar o que vale a sua atenção. Este método tem precedentes na lista de “best-sellers”, e possui aplicações interessantes no site da Amazon, que descreveremos na seção seguinte deste capítulo. Uma variação também desta ideia pode ser encontrada nas ferramentas de busca do Google, que usa uma abordagem de busca baseada em conteúdo. A ordem ou ranking dos resultados de busca no Google dá um peso maior aos sites que possuem links com muitos outros sites, baseados na ideia de que estes sites possuem maior “autoridade” para os usuários.

Os três modelos de recomendação citados acima são válidos e relevantes, porém limitados frente aos problemas referentes à descoberta da informação em contextos de excesso de informação. Parece claro que os SR podem ser utilizados para encontrar pessoas que possam estar interessadas em sistemas baseados na reputação; que as ferramentas de buscas podem servir para complementar o uso dos SR e que a identificação de itens baseados na popularidade pode aproximar relações com os SR. Porém, o processo de descoberta de informação pode empobrecer diante da



automatização da busca através de uma mediação “invisível” - reduzindo assim as surpresas, a criatividade e a diversidade de combinações (*serendipity*) entre pessoas e conteúdos, em detrimento do controle e sistematização dos sistemas.

Estrutura e Modelos de Funcionamento dos SR

Os estudos sobre os sistemas de recomendação emergiram como uma área independente em meados dos anos noventa - apesar de se tratar de um campo de investigação interdisciplinar - quando os pesquisadores começaram a focar nos problemas de recomendação relacionados explicitamente com a organização da informação baseada na estrutura de ranking, onde a avaliação sobre a relevância de um determinado item ou documento é dada pelo próprio usuário. A formulação mais comum em relação ao problema da recomendação pode ser abordada da seguinte forma: como criar uma estimativa de ranking para os itens que não foram vistos e/ou avaliados pelos usuários e quais são os critérios de relevância nestes casos?

Intuitivamente, esta estimativa é normalmente baseada no ranking dado a outros itens pelos próprios usuários do sistema, entre outras informações que serão descritas a seguir. Na medida em que é possível estimar o ranking dos itens que ainda não foram avaliados ou conhecidos por um determinado usuário, pode-se recomendá-lo os itens com a maior “probabilidade de relevância”. Nos sistemas de informação em geral, e em particular nos SR, a proposta básica é recomendar informações “relevantes” a um usuário ou grupo de usuários.

Entretanto, o conceito de relevância é um dos mais controvertidos na Ciência da Informação (CI). Este conceito ganhou muitos nomes e definições e continua sendo uma das propriedades básicas da informação e dos processos de comunicação. De acordo Saracevic (1970), as relações de relevância e distribuição de vários processos de comunicação estão intimamente ligadas, e as respostas relevantes de um dado sistema dentro de um processo de comunicação normalmente são distribuídas de acordo com a lei de Bradford e Zipf – ou seja, a relevância aumenta proporcionalmente ao uso de um determinado item ou documento.

A relevância não se limita unicamente a uma propriedade do mecanismo interno do sistema ou ao simples resultado de uma combinação entre a representação pelos usuários de sua necessidade de informação com os termos atribuídos aos itens pesquisados. Esta visão implicaria que todos os conteúdos recuperados “corretamente” pelo sistema sejam, por definição, relevante a um indivíduo. Esta abordagem ignora virtualmente o papel do usuário. Na aplicação real dos sistemas, entretanto, é o usuário que decide finalmente se o item ou documento recuperado é útil e se de alguma maneira satisfaz às suas necessidades (SARACEVIC, 1970; BARRY, 1994).



A partir da revisão de literatura sobre o tema, verifica-se que a formulação mais aceita sobre a classificação dos diferentes tipos de SR foi primeiramente publicada por Resnick et. al. (1994); HILL et al. (1995) e Maes & Shardanand (1995). Utilizaremos estes autores como referência para apresentar as categorias de classificação dos SR:

1. Recomendações baseadas em conteúdo: sistemas que recomendam itens similares aos que o mesmo usuário demonstrou preferência no passado;
2. Recomendações colaborativas: sistemas que recomendam ao usuário itens avaliados no passado por pessoas com gostos e preferências similares;
3. Abordagem híbrida: sistemas que combinam o método colaborativo com o método baseado em conteúdo.

Além destes três tipos de SR, que prevêm em valores absolutos as possíveis avaliações dos itens ainda desconhecidos por usuários individuais para calcular as recomendações relevantes, alguns pesquisadores concebem outras categorias para classificar os SR de acordo com os critérios de recomendação. Várias tecnologias têm sido desenvolvidas com o objetivo de identificar padrões de comportamento (consumo, pesquisa, tomada de decisão, etc.) a fim de utilizá-los na personalização do relacionamento com os usuários.

O CASO LAST.FM

A idéia de aplicar técnicas de recuperação automatizada de informação na música data dos anos 1960 (Kassler, 1966). Entretanto, as pesquisas sobre a Recuperação da Informação da Música (RIM) crescem recentemente com a explosão das coleções de obras musicais na forma digital - precipitadas pelo desenvolvimento de tecnologias e redes digitais citadas anteriormente, tais como a compressão de áudio como o MP3; os serviços on-line de compartilhamento P2P (*peer-to-peer system*); os sistemas óptico-musical (OMR) e custos decrescentes do armazenamento digital e da conexão em alta velocidade (FUTRELLE e DOWNIE, 2002).

A motivação das pesquisas em RIM é o grande e crescente volume de música digital disponível atualmente na Internet (DUREY et al., 2001; HOOS et al., 2001; KORNSTÄDT, 2001; YANG, 2001). Porém, os pesquisadores de RIM entendem que a crescente disponibilidade de música na Internet é somente um fator agravante da questão mais importante: existem poucas técnicas efetivas de organização e recuperação para as coleções de música digital. O problema existe desde que a música é codificada digitalmente, mas obtém grande pressão apenas recentemente, quando o custo de armazenar grandes coleções de música digital cai a quase zero e o número dessas coleções conseqüentemente explode.

A nova agenda de pesquisa na área de Recuperação da Informação (*Information Retrieval*) aponta para as discussões sobre a organização da cultura e a promoção da



diversidade cultural no contexto da globalização e das tecnologias digitais. Segundo Michael Lesk (1996, p.11), em seu artigo “*Seven Ages of Information Retrieval*” as pesquisas na área de Recuperação da Informação para as próximas décadas devem contemplar a manipulação de sons e imagens nos contextos de grandes volumes de informação. A análise da Recuperação da Música na Internet pode representar uma maneira importante de discutir as aplicações das comunidades reais de usuários de música e estudar suas necessidades, culturas de uso, modelos de mediação e suas consequências sociais no entorno digital.

A música é um produto social e simbólico de grande importância nas diferentes formações culturais. A música pode usar diferentes linguagens e expressão, sendo produto cultural de características muito especiais: nenhum produto cultural tem mostrado tamanha capacidade de adaptação aos diferentes meios de comunicação.

A indústria da música constitui particularmente um setor chave dentro do conjunto das IC. Primeiramente por ser o setor irrigador dos demais setores, a música é componente fundamental de outros produtos culturais, tais como filmes, programas de rádio e televisão, videogames, etc. Segundo porque a música é um bem simbólico que transcende as barreiras da língua, e assim demonstra ser o produto cultural como maior potencial de disseminação e capilarização na Internet. Dentro do setor da música, a Internet pode potencialmente operar como espaço de mediação, de exploração de novos talentos musicais, de inovação de formas criativas, descobrimento de tendências e novos sons, disseminação de novos gêneros, preservação de catálogos de repertório nacional e local e, em definitiva, demonstra potencial para constituir-se tanto como um canal de mediação como também um acervo essencial para o enriquecimento e preservação da diversidade cultural (CALVI, 2008).

Destacamos os principais argumentos e critérios utilizados que justificam a escolha do SR da Last.fm como objeto específico de análise deste trabalho. O primeiro argumento tem como fundamento o fato de que o SR da Last.fm em relação aos demais SR de Música apresenta de formas mais avançada e transparente determinadas características nos níveis tecnológico, econômico e regulatório. Estas inovações apontam para a identificação de uma nova lógica de mediação das Indústrias Culturais que se configura na Internet para a orientação dos usos sociais da cultura, tanto através dos serviços de informação como das estratégias de disseminação dos bens culturais e simbólicos na Sociedade da Informação.

O SR da Last.fm articula inovações relevantes na lógica de difusão e consumo da música na Internet através da classificação colaborativa da música, recomendações baseada na folksonomia (através das *social tags*), mapas de gostos musicais dos usuários e das redes sociais de ouvintes e dessa forma aponta para as tendências dos usos sociais da música nestes novos contextos.



Em segundo lugar, dentre os SR atualmente disponíveis para uso gratuito na web, o SR da Last.fm é o único sistema que articula e disponibiliza publicamente dados estatísticos pertinentes para a investigação dos usos sociais da música nos novos contextos de recomendação e consumo nas redes e dispositivos digitais.

O SR da Last.fm publica estatísticas de uso da música através dos dados gerados pelos computadores que registram todos os arquivos de música executados por sua rede de usuários. Dessa forma, gera informações sobre as preferências musicais tanto no nível individual como coletivo (comunidades de gosto).

Os dados são complementados através do registro das operações realizadas pelos ouvintes-usuários do sistema em seus serviços através das estratégias de busca, descoberta, acesso, recuperação e uso da música (palavras-chave, formas de classificação e organização das coleções, playlist e etc). Estas estatísticas são cruzadas e articuladas para produzir as recomendações de música de acordo com os critérios de relevância e interesse dos usuários.

Por último, o SR da Last.fm posiciona-se atualmente como a principal plataforma social de música na Internet, conectando aproximadamente 21 milhões de usuários diretos⁴ e 19 milhões de usuários indiretos⁵ a um acervo de mais de 150 milhões de faixas de músicas referentes a 16 milhões de artistas⁶. Nas listas de mensuração de uso, a Last.fm está entre os SR de Música mais populares na Internet, situando-se em primeiro lugar no ranking de número de usuários entre os sistemas de recomendação analisados⁷ para busca e descoberta de informação.

Dessa forma, o SR da Last.fm apresenta características e condições especiais de observação. As condições de observação e análise dos usos sociais da música na Internet fornecidas pelo sistema não são apenas relevantes para esta pesquisa, mas também apresentam dados inéditos que revelam os usos domésticos e privados da música em escala global, que superam as estatísticas de venda de música e as limitações que incluem as entrevistas sociológicas para os estudos culturais relacionados com os gostos.

⁴ Financial Times de maio de 2008 publicou que o número de usuários diretos da Last.fm correspondiam a 26 milhões de usuários. Porém, como não foi possível o acesso ao relatório oficial da empresa, mantivemos o número de 21 milhões de usuários. Para mais detalhes ver: Financial Times Press Digest. CBS targets Last.fm for advertising revenue. 27 de maio de 2008. Disponível em: <http://search.uk.reuters.com/>. Recuperado em 25 de julho de 2008.

⁵ F. Miller, M. Stiksel, and R. Jones. Last.fm in numbers. *Last.fm press material*, February 2008. A ambiguidade sobre o método utilizado para a mensuração da quantidade de usuários do Sistema de Recomendação da Last.fm e de sua audiência será discutida em outro trabalho mais extenso.

⁶ Em fevereiro de 2008 a Last.fm Ltd. divulgou através da mídia um relatório oficial sobre o rápido crescimento de sua base de dados, que neste momento apresentavam estes números de usuário e de acervo musical. Para mais detalhes ver: F. Miller, M. Stiksel, and R. Jones. Last.fm in numbers. *Last.fm press material*, February 2008.

⁷ Alexa. http://www.alexa.com/data/details/traffic_details/last.fm. Estatística de outubro de 2008. Ver também



A Last.fm (www.last.fm) oferece *streaming* (transmissão) de música através do Sistema de Recomendação chamado “Audioscrobbler”, disponível para download e uso gratuito na Internet. O Audioscrobbler oferece diversas funções, serviços e características que articuladas correspondem a uma geração de serviços de música online, que sob os parâmetros de O’REILLY (2005) é chamada de web 2.0. Por outro lado, alguns autores defendem que as inovações destes sistemas correspondem à aurora da websemântica ou web 3.0.

O site Last.fm pode ser definido como uma plataforma do Sistema de Recomendação de Música (SRM) “Audioscrobbler”, mas ao mesmo tempo integra características e funções importantes de uma rádio online, como também de uma rede social que agrega outras comunidades virtuais existentes na Internet – funções que se integram e são constitutivas das estratégias para o serviço de recomendação de música.

As duas últimas funções (rádio online e redes sociais) estão articuladas com a primeira função (Sistema de Recomendação) para gerar informações detalhadas sobre os usos e preferências dos usuários, cujos dados alimentam a base de dados do sistema Audioscrobbler. Por outro lado, as recomendações são oferecidas aos usuários através de um fluxo musical contínuo, que corresponde à função de rádio online segmentada e orientada a partir dos gostos e/ou das demandas momentâneas dos usuários.

A partir do cruzamento das estatísticas sobre os usos destas funções e características mencionadas, o SR da Last.fm constrói perfis e mapas dos gostos musicais de sua rede de usuários registrados, baseando-se nos dados registrados através dos computadores e aplicativos utilizados. A organização dos arquivos de música em seu acervo é baseada no gosto dos usuários que direcionam os processos de busca e recomendação de acordo com os critérios de relevância e precisão de cada usuário.

O SR Audioscrobbler não somente registra as músicas que os usuários escutam através dos aplicativos da Last.fm como também armazena em sua memória os metadados de todos os arquivos de música executados através dos computadores utilizados pelo usuário, independentemente do IP, da fonte ou da plataforma utilizada (como por exemplo, os tocadores de MP3 alimentados através do computador).

"Scrobbling" significa que quando o usuário ouve uma música, o nome da música e do artista é enviado para a Last.fm e adicionados ao seu perfil musical. Assim que o registro e download do software da Last.fm forem feitos, o scrobbling das músicas pode ser feito no computador do usuário, em seu iPod ou iPhone. As músicas que estão sendo ouvidas também aparecem em seu perfil para que outras pessoas tenham acesso ao que um usuário específico está ouvindo. O "scrobbling" é feito mais 600 milhões de vezes ao dia⁸. Estes dados ajudam a Last.fm a organizar e recomendar músicas aos seus

⁸ Last.fm. *Perguntas frequentes Last.fm*. Disponível em: www.lastfm.com



usuários, elaborar rádios personalizadas e sofisticar os critérios de relevância para os processos de busca e descoberta de informação.

O uso dos Sistemas de Recomendação da Last.fm tem se popularizado entre as redes de usuário de Internet por reduzir as dificuldades de recuperação da informação da música, no caótico mar de informações disponíveis na rede, através da sugestão de produtos culturais ou conteúdos digitais que tem mais valor e relevância para cada usuário. Este sistema apresenta-se como um novo mediador entre os grandes produtores culturais e a massa de consumidores na Internet e se constituem na grande vitrine onde se exibem e se recomendam os produtos culturais mais consumidos, tais como a música, os filmes, os livros e os videogames.

Os SR em geral, e especialmente a Last.fm, estão constituindo-se atualmente no modelo de disseminação e comercialização por excelência de produtos culturais na Internet, na principal estratégia comercial dos Grupos Multimídia (GM). Assim, as grandes empresas de tecnologia e de Internet (principalmente os GM) competem com o objetivo de encontrar a forma mais eficaz e eficiente para utilizar esta tecnologia, que orienta o encontro dos usuários com obras musicais recomendadas segundo os padrões de consumo e gostos dos próprios receptores.

A expansão constante e complexa do volume de informação na web tem tornado estes sistemas ferramentas muito atraentes para os usuários em um ambiente muito amplo e variado de busca de informação, recuperação de informação, pesquisas ou em atividades de e-commerce, como é o caso da Internet.

O SR da Last.fm promete ajudar ao usuário administrar o problema do excesso de informação recomendando-o itens que correspondam, ao mesmo tempo, à exatidão e precisão de sua busca, ou seja, aos seus critérios pessoais de relevância. Os SR Last.fm também é utilizado para que o usuário conheça novos itens, de acordo com o seu critério de relevância, interesses e gostos. Assim, a tecnologia dos SR representa o novo paradigma da classificação, organização, busca e uma nova ferramenta de pesquisa em ambientes virtuais através da filtragem colaborativa dos usuários.

CONCLUSÕES

O surgimento dos Sistemas de Recomendação e a história da Internet são paralelos e simultâneos no tempo e espaço. Portanto, torna-se crucial avaliar os modelos de mediação que se implicam na Internet através dos SR. A premissa central desta contribuição é que os SR são um tipo de mediação particular que pode ser uma tendência dominante na Internet e no futuro da digitalização e convergência dos meios de comunicação com a própria web. Estes mediadores começam a definir os métodos e modelos de produção, disseminação e consumo dos produtos culturais e impactam



diretamente na construção e reprodução dos usos sociais dos conteúdos audiovisuais no entorno digital.

Por outro lado, os SR se constituem como a última estratégia dos GM para organizar e orientar o consumo cultural na Internet utilizando técnicas muito refinadas que consideram os gostos e preferências culturais dos usuários em base a uma medição exaustiva dos seus perfis do consumo realizados na Rede.

Assim, os SR se apresentam como uma forma de coordenação muito precisa do ciclo de produção, disseminação, difusão, consumo e uso dos produtos culturais, reduzindo ainda mais a incerteza inerente ao consumo destes bens simbólicos tão particulares. Também, os SR utilizam a informação derivada dos usos feitos pelos grupos de usuários, traçando perfis não só individuais mas dos coletivos de usuários com interesses similares.

Neste sentido, os SR se apresentam como uma nova forma de organização da informação e orientação do consumo cultural na Era Digital, que colocam novos problemas aos estudos clássicos tanto sobre os usos e práticas de consumo cultural como aos debates em torno das novas alternativas à concentração ou à diversidade cultural.

Vemos que os novos fenômenos de intermediação cultural na Internet colocam em questionamento as políticas culturais clássicas, dado que os usos sociais das novas tecnologias desenvolvidos pelos usuários muitas vezes estão à frente das políticas mais progressistas, e que a lógica das redes de usuários tem muito a contribuir para o debate em torno das problemáticas sobre a organização e a classificação da informação digital e também para a formulação de Políticas Públicas que visem à promoção da diversidade cultural nos novos ambientes digitais.

REFERÊNCIAS

- ADOMAVICIUS, G., & TUZHILIN, A. (2005). Toward the next generation of Recommender Systems: A Survey of the State-of-Art and Possible Extension. *IEEE Transaction on Knowledge and Data Engeneering*, junho, vol. 17, n.6 , pp. 734-748.
- ANSARI, A., ESSEGAIER, S., & KOHLI, R. (2000). Internet Recommendation Systems. *Journal of Marketing Research*, Vol. XXXVII , pp. 363-375, agosto.
- BALABANOVIC, M., & SHOHAM, Y. (1997). Fab: Content-Based, Collaborative Recommendation. *Communication of ACM*, v.40, n°3 , pp. 66-72, março.



- BARRY, C. (1994). User-Defined Relevance Criteria: An Exploratory Study. *Journal of American Society for Information Science*, abril.
- BELVIN, N., & CROFT, W. (1992). Information Filtering and Information Retrieval: two sides of the same coin? *Communication of the ACM* , dezembro, p. 29.
- BOURDIEU, P. [1979] (2007). *A distinção: crítica social do julgamento*. São Paulo: Edusp; Porto Alegre: Zouk.
- CALVI, J. (2008). *Reproducción de la cultura o cultura de la reproducción?* Analisis econômico, político y social de La distribución y el consumo de productos audiovisuales em Internet. Madrid: Dykinson.
- COHEN, W., SCHAPIRE, R., & SINGER, Y. (1999). Learning to Order Things. *Artificial Intelligente Research*, vol. 10 , 243-270.
- FARCHY, J. (1999). *La fin de l'exception culturelle?* Paris: CNRS.
- FOLTZ, P., & DUMAIS, S. (1992). Personalized Information Delivery: an analysis of Information Filtering Methods. *Communication of ACM*, v. 35, n° 12 , dezembro, pp. 51-60.
- FOSTER, A., & FORD, N. (2003). Serendipity and information seeking: an empirical study. *Journal od Documentation*, v. 59, n.3, pp.321-340.
- FREUND, Y., IYER, R., SCHAPIRE, R., & SINGER, Y. (1998). An efficient boosting algorithm for combining preference. *Proc. 15th Int'l Conf. Machine Learning*
- GOLBERG, D., NICHOLS, D., & OKI, B. e. (1992). Using collaborative filtering to wave an information tapestry. *Commun, ACM* 35, v.12, dezembro, pp. 61-70.
- HERLOCKER, J. (2000). *Understanding and Improving Automated Collaborative Filtering Systems*. Minnesota: Tese de Doutorado, University of Minnesota.
- HILL, W., STEAD, L., ROSENSTEIN, M., & FURNAS, G. (1995). Recommending and Evaluating Choices in a Virtual Community of Use. *Proc. Conf. Human Factors in Computing Systems*.



- INGWERSEN, P. (1992). *Information Retrieval Interaction*. Londres: Inglaterra.
- JIN, R., SI, L., & ZHAI, C. (2003a). Preference-based graphic models for collaborative filtering. *Proc. 19th Conf. Uncertainty in Artificial Intelligence (UAI)* , 1-8.
- JIN, R., SI, C., ZHAI, C., & CALLAN, J. (2003b). Collaborative Filtering with decoupled models for preferences and rating. *Proc. 12th Int'l Conf. Information and Knowledge Management* .
- KAY, A. (1984). Computer Software. *Scientific American*, 251, setembro, pp. 53-59.
- LACROIX, J.-G. ; TREMBLAY, G. (1997). *The 'information society' and cultural industries theory*. Toronto: Sage.
- LANCASTER, F. (1968). *Information Retrieval Systems: characteristics, testing and evaluation*. New York: John Wiley & Sons, Inc.
- LOEB, S., & TERRY, D. (1992). Information Filtering. *Communication of ACM*, v.35, n°12 , dezembro, pp. 26.
- LYNCH, C. (2001). Personalization and Recommender Systems in the Larger Context: New Directions and Research Questions. *Coalition for Networked Information* .
- MAES, P., & SHARDANAND, U. (1995). Social Information filtering: algorithms for automating "word of mouth". *Human Factors in Computing Systems*, pp.210-217.
- NEGROPONTE, N. (1970). *The Architecture Machine*. Boston: MIT Press.
- O'REILLY, T. (2005). What Is Web 2.0?: design patterns and business models for the next generation of software. p. Disponível em: <<http://www.oreillynet.com/1pt/a/6228>>. Acesso em 01 jun. 2007.
- PINE, J. B. (1993). *Mass Customization: the new frontier in business competition*. Boston, Mass.: Harvard Business School Press.
- REATEGUI, E. B., & CAZELLA, S. (2005). Sistemas de Recomendação. *XXV Congresso da Sociedade Brasileira de Computação* , 306-348.



- RESNICK, P. (1994). GroupLens: an open architecture for collaborative filtering of Netnews. *ACM - Conference of Computer Supported Cooperative Work* , pp. 175-186.
- _____. Recommender Systems - Interview. *University of Michigan School of Information* , fevereiro. Disponível em: www.umich.edu.
- RESNICK, P., & VARIAN, H. R. (1997). Recommender Systems. *Communication of the ACM* , pp. 56-58, vol. 40, nº 3, março.
- RESNICK, P., IACOVU, N., SUCHAK, M., BERGSTROM, P., & RIEDL, J. (1994). GroupLens: an open architecture for collaborative filtering of Netnews. *ACM - Conference of Computer Supported Cooperative Work* , pp. 175-186.
- RICH, E. (1979). User Modeling via Stereotypes. *Cognitive Science*, vol. 3, nº 4 , pp. 329-354.
- RIEDL, J., HERLOCKER, J., GOOD, N., SCHAFER, J., KONSTAN, J., BORCHERSS, A. et al. (1999). Combining Collaborative Filtering with Personal Agents for Better Recommendation. *Proc. of AAAI*, vol.35 , pp. 439-446.
- RIEDL, J., KONSTAN, J. A., MILLER, B. N., MALTZ, D., HERLOCKER, J. L., & GORDON, L. R. (1997). GroupLens: applying collaborative filtering to usenet news. *Comm ACM*, vol.40, nº3 , pp.77-87.
- RIEDL, J., KONSTAN, J., & SCHAFER, J. (1999). Recommender Systems in E-commerce. *Proc. of the 1st ACM Conference on Eletronic Commerce* , pp. 158-166.
- ROBERTS, R. (1989). *Serendipity: Accidental discoveries in science*. Nova York: John Wiley and Sons.
- SARACEVIC, T. (1970). The concept of “relevance” in information science: a historical review. In: T. SARACEVIC, *Introduction to information science*. Nova York: R.R. Bowker, pp. 111-151.
- VICKERY, B. (1970). *Techniques of Information Retrieval*. Londres: Butterworths.
- WILLIAMS, R. (1981). *Culture*. London: Fontana.