



A Máquina em Aberto: As ferramentas de meta-busca e estratégias de comunicação na rede¹

André Vouga
Mestre em Comunicação e Tecnologias da Informação
Universidade Federal Fluminense

Proposta

Parte da tradição dos estudos da comunicação tem estado orientada para a análise das influências mútuas entre as tecnologias de comunicação e os contextos culturais onde elas são utilizadas. Por ter definido sua utilização para fins comunicacionais há relativamente pouco tempo o ambiente computacional ainda apresenta campos inexplorados por esse tipo de estudo. Este trabalho tem por objetivo a discussão de certas especificidades do ambiente computacional, quando pensado como suporte para meios de comunicação. Recorreremos à referência do padrão de desenvolvimento de *software* em código aberto – que, como veremos, também pode ser associado às linguagens mais utilizadas para a expressão na Internet - para tentar demarcar algumas diferenças trazidas por este ambiente em relação ao padrão histórico das ferramentas de comunicação, discutindo como ele pode estar trazendo algumas novas possibilidades nas formas de construir e se relacionar com essas tecnologias.

Concentraremos a análise na sinergia entre dois aspectos que parecem ser marcantes no universo computacional: uma maior facilidade de disseminação da informação e uma capacidade mais intensa de fixar, de organizar, de traduzir e de permitir o reaproveitamento de certos tipos de memória cultural na constituição de suas próprias ferramentas. Procuraremos discutir como estas características podem estar trazendo novas possibilidades estratégicas em relação aos processos de construção das ferramentas de comunicação. Estaremos tentando pensar como estes fatores podem estar indicando novos papéis para aqueles que ocupavam a posição de emissores nos meios tradicionais de comunicação.

É preciso lembrar que a comunicação, em si, também começou a ser isolada como objeto de estudo num tempo relativamente recente. Apenas a partir do primeiro impacto de meios audiovisuais de massa como o cinema e o rádio. Possivelmente em função dessa origem a teoria da comunicação tem estado orientada para a avaliação das relações nesse tipo de meio, onde poucos emissores têm o privilégio de comunicar a muitos. O ambiente computacional tem colocado novos problemas para a teoria da comunicação tradicional. Um deles seria uma presença marcante de formações que, em termos da produção de conteúdo, não se enquadram bem num padrão "um para muitos" e sim numa articulação de "muitos para muitos". Isso pode ser percebido em casos de sucesso dentro deste ambiente como o das ferramentas de busca, o dos provedores de webmail, o dos serviços de suporte de mailists e o dos sítios especializados no auxílio à publicação de páginas por leigos, entre outros. Neles uma questão fundamental parece ser a de disponibilizar ferramentas que potencializem a produção de conteúdos a partir da integração entre seus usuários.

Em função dessas constatações, este trabalho pretende analisar possíveis papéis para o profissional de comunicação – pensado como o emissor do processo comunicacional - para além da seleção dos conteúdos produzidos pelo corpo social visando sua retransmissão. Uma das hipóteses, com que trabalharemos, é a de que se delineiam para ele novas possibilidades como a de promotor e organizador de uma produção mais generalizada de conteúdos, através da proposição de novos contextos de comunicação e sua contínua transformação.

Virtualização

Existe uma noção fundamental para entender como esta possibilidade estaria dada. É importante demarcar que a habilidade básica do ambiente computacional está na realização de operações lógicas simplificadas com números. A utilização desses números como marcadores simbólicos vai permitir que informações em muitos formatos (como texto, imagem etc) sejam matematicamente convertidas, armazenadas e processadas. Mas nos importa de maneira mais específica observar que este mesmo

1 Trabalho apresentado no NP08 – Núcleo de Pesquisa Tecnologias da Informação e da Comunicação, XXV Congresso Anual em Ciência da Comunicação, Salvador/BA, 04 e 05. setembro.2002.



mecanismo, de fundo simbólico, é a base do próprio seqüenciamento das instruções de processamento. Como veremos mais adiante, é através desse seqüenciamento que o ambiente computacional se transforma, por simulação, numa série de diferentes máquinas capazes de manusear informação. Máquinas constituídas por símbolos que, assim como toda a informação digitalizada no ambiente computacional, contam com seu grande potencial de processamento, tradução e distribuição. Haveria, nesse caso, uma espécie de virtualização do ferramental técnico perceptível nessa codificação. Como analisamos (Vouga, 2002) ela teria como consequência uma aceleração e potencialização dos processos de fixação e transmissão de cultura portados pelos próprios objetos técnicos, como detectados pela antropologia (Leroy-Gourhan, 1964), trazendo novas possibilidades para estas relações.

Steve Johnson (1997) tentou analisar esse processo, através do conceito de interface, procurando caracterizá-la como uma nova instância cultural, que não é nem forma, nem conteúdo. Buscou acompanhar a progressão da adição de dispositivos de relação entre o homem e o aparato computacional. Mas se deteve mais especificamente nas interfaces gráficas que, ao seu ver, permitiram estabelecer representações mais intuitivas dos dados digitais e, também, a ilusão da sua manipulação direta pelo usuário. Seu desenvolvimento teria efetivamente permitido que o computador se popularizasse como um meio de comunicação.

Estamos convivendo, hoje, com a transferência progressiva de muitos aspectos de nossa cultura para o universo digital das redes de computação. Elas têm uma capacidade específica de aglomerar enormes volumes de informação e potencializar grandemente sua distribuição. Em função desse fato vem se estabelecendo um problema de excesso de informações disponíveis e este, por sua vez, definiria uma demanda por mapas, por filtros, ou seja, por interfaces.

Esse fenômeno demarca a importância do acesso à constituição das interfaces, por vários aspectos. Antes de tudo porque estão se definindo novas formas de publicação onde se disponibilizam grandes bases de dados. Para que estas sejam eficientes em termos de comunicação é preciso dotá-las com ferramentas adequadas para interação que, normalmente, precisam portar soluções particulares para cada caso.

Um segundo aspecto é a tendência para a personalização do conteúdo que, por sua vez, embute possibilidades estratégicas de conhecimento dos públicos presentes na rede. Para isso são necessárias interfaces dinâmicas, capazes de registrar demandas e se adaptar àquelas anteriormente registradas.

Um terceiro aspecto possui uma certa ligação com o anterior. Para Johnson (2001, p. 204-225) ocorre um outro fato fundamental com o aparecimento de ferramentas capazes de registrar hábitos, interesses, padrões de operação e, também, de capturar a sinalização entre os usuários. Elas possibilitam reconhecer padrões, classificar e comunicar constantemente o resultado deste processamento entre estes mesmos usuários estabelecendo ciclos de retroalimentação. Esses aspectos seriam fundamentais para o entendimento de um tipo específico de relação de mediação considerado promissor na rede (Vaz, 6/9/2001). Ele aparece em ferramentas que permitem a interação de grupos em larga escala, produzindo conteúdo a partir da articulação entre diversos atos e manifestações de seus usuários. Como exemplos são citados o site de grupos de discussão Slashdot e o de leilões E-bay.

Um outro aspecto seria um pouco mais complexo. Como já vimos as interfaces permitem que o computador seja uma máquina de uso mais fácil e direto. Porém elas também teriam um certo caráter de obstáculo, na medida em que se interpõem entre os usuários e as informações operadas por ele. O interfaceamento é um processo de tradução com todos seus problemas inerentes. Quando delegamos o tratamento prévio de informações a um algoritmo, por um lado ganhamos uma capacidade maior para enfrentar o problema do excesso, mas por outro haverá uma série de perdas em termos da visualização do detalhe, do recorte, da avaliação dos aspectos sutis. Processar freqüentemente inclui selecionar, classificar, ordenar, o que embute, necessariamente, uma série de juízos de valor.

Opera-se, então, no manuseio das interfaces, um paradoxo fundamental: o aumento de nosso poder de operação intuitivo, que estabelece uma sensação de liberdade, se faz ao preço de nossa autonomia, uma vez que os pacotes tecnológicos subjacentes realizam operações que desconhecemos. Dessa forma se estabelece a importância de se interrogar em que medida se pode considerar que as interfaces orientam, recortam, disciplinam nosso olhar (Albuquerque, 2001). E, também, até que ponto é desejável que estes processos sejam dominados por poucos, que continuem opacos para a maioria de seus usuários. Enfim, em que medida seria interessante pesquisar caminhos para que possamos lidar com esse problema.



Abertura

Uma possibilidade promissora viria do fato, já citado, de que a capacidade que o ambiente computacional tem de tradução e conversão simbólica pode ser aplicada para nos traduzir suas próprias instruções de comando. Isso permitiria que um maior contingente de usuários possa compreendê-las, intervir em suas diretrizes ou, pelo menos, usá-las com maior consciência das perdas e ganhos subjacentes.

Esse tipo de perspectiva pode ser encontrado no ideário de grupos como o Movimento pelo *Software* Livre. Surgido sob a influência da revolução contracultural norte-americana ele busca preservar os padrões acadêmicos de produção do conhecimento no universo da programação, contra a tendência, estabelecida a partir do fim dos anos 70, de adoção do modelo baseado na propriedade intelectual.

Este movimento tem como um de seus alicerces o conceito de código aberto (*open source*), noção que surgiu na área da engenharia de *software* para demarcar uma forma específica de geração de aplicativos. Ele se estabeleceu por ser possível, hoje, produzir programas nas chamadas linguagens de alto nível, próximas da compreensão humana. Ocorre que, para se tornar operacional no ambiente computacional, o *software* deve ser traduzido para a codificação da chamada linguagem de máquina (as instruções diretas para o computador) e, assim, sua constituição se torna obscura aos que não estiveram envolvidos em sua produção. O conceito de código aberto está ligado à publicação do código de alto nível. O objetivo seria franquear à comunidade interessada o entendimento da estrutura dos programas, a possibilidade de interferência em sua constituição, de adicionar extensões segundo necessidades específicas, e, mesmo, o reaproveitamento de suas partes componentes.

A nosso ver essas propostas estariam ligadas às possibilidades trazidas pelo fato do ambiente computacional ser uma máquina, simbolicamente definida, de simulação de máquinas de informação. Ou seja, elas tem haver com a virtualização das ferramentas técnicas e da explicitação dos conhecimentos contidos nelas por sua constituição simbólica. Estariam relacionadas ao fato de que estes processos acelerariam os ciclos de inscrição e transmissão do conhecimento, através das técnicas, estabelecendo novos patamares na capacidade de reorganizar e reaproveitar este conhecimento em processos de cópia, recorte e colagem.

Alguns efeitos desses processos estariam numa reexpansão da quantidade de atores envolvidos na criação de ferramentas tecnológicas em função da redução das exigências técnicas e econômicas envolvidas. Entre outras coisas, a virtualização do ferramental lhe traz reprodutibilidade o que reduziu os riscos presentes na bricolagem com as máquinas (Pode-se, por exemplo trabalhar sobre cópias dos programas, preservando as etapas bem sucedidas do processo). Os custos envolvem pouca coisa além de acesso a terminais de computação e treinamento (Ao contrário do que tem acontecido com a pesquisa de tecnologias “físicas”, que cada vez envolvem processos materiais mais caros e complexos). Além disso a própria essência simbólica dessas ferramentas torna intrinsecamente explícitos os conhecimentos fixados nelas (Diferentemente dos padrões mais implícitos historicamente observados nos aparatos técnicos (Leroy-Gourhan, 1964)). Assim se estabelecem as bases para que se possa aprender sobre sua estrutura pela observação direta, permitindo que se desenvolvam saberes mais estruturados acerca delas. Dessa forma estes aspectos parecem estar permitindo uma relação mais rápida e direta de pesquisa e proposição de ferramentas, inclusive as de comunicação. E, também, que este processo comece a se abrir a um público mais amplo do que o de engenheiros e técnicos especializados ligados a grandes corporações.

A conjunção com a facilidade de distribuição e estocagem dos símbolos de constituição dessas ferramentas, propiciada pela Internet, teria sido bastante feliz para estas propostas cooperativas de produção de ferramentas. Teria havido uma tendência à diluição de tarefas amplas em pequenas etapas cumpridas por comunidades de atores pontuais distribuídos nas mais diferentes culturas. Isso pode ser acompanhado em diversos casos recentes de produção de software, um exemplo especificamente visível veio da produção completa de um sistema operacional, o Linux, num nível competitivo com o software comercial.

As novas facilidades para a produção técnica comunitária em pequenas contribuições, com a avaliação por pares e reutilização dos conhecimentos produzidos, ajudaram a que se fortalecesse no movimento pelo *software* livre a perspectiva de resistência de uma visão acadêmica de produção da técnica. Elas aparecem em oposição aos modelos capitalistas de produção de software, altamente dependentes dos



sistemas de propriedade intelectual¹. Isso pode ser fortemente percebido nos mais diferentes textos produzidos por articuladores desse movimento (Stalman, 7/9/00). Há uma premissa de fundo, presente nesses discursos, de que o software livre é uma das frentes onde forças com preocupações sociais devem agir buscando orientar os frutos dos recentes avanços tecnológicos num sentido desalienante. Essas propostas parecem ter grande influência do ideal democrático norte-americano ligado fundamentalmente à liberdade de informação. (Albuquerque, Janeiro de 2000, p. 34)

Na Rede

Em trabalhos anteriores (Vouga, 2002) procuramos demonstrar a presença de alguns padrões semelhantes aos do universo do *software* livre nos processos de criação de ferramentas de comunicação dentro do material publicado na Internet. Questão que estaria ligada ao fato de o principal código utilizado para expressão na rede, o HTML, atender ao modelo do código aberto. Esse fato teria surgido da necessidade de exibir os mesmos dados, com as mesmas formatações, em diferentes sistemas ligados a rede. Se estabeleceu um padrão onde eles não são previamente convertidos para linguagens de máquina. Eles são distribuídos em sua linguagem de concepção e interpretados em cada plataforma. Num certo sentido essa linguagem tem funcionado como uma espécie de esperanto entre as máquinas e programas interligados na Internet. Por isso mesmo as aplicações mais elaboradas, presentes na rede, se utilizam dessa linguagem para suas entradas e saídas de dados.

Assim talvez os padrões detectados nos demais casos ligados à abertura de códigos na programação poderiam ter algum valor para se avaliar estratégias de abordagem para a produção de ferramentas para a rede, mesmo no nível mais básico. No caso da Internet isto parece estar ligado ao reaproveitamento de conteúdos, interfaces e até a requisição remota de processamento.

Requisições

Um caso prático onde se pode perceber a existência desses fatores de reaproveitamento na rede são os processos de configuração das ferramentas meta-busca (Aqueles que interpelam outras ferramentas de busca e condensam seus resultados). Conforme levantamos (Vouga, 2002) muitas das ferramentas de busca presentes na rede disponibilizam, em seus sites, os códigos HTML de pequenos formulários, que acionam sua pesquisa, para serem inseridos nas páginas construídas por qualquer pessoa². Essa estratégia seria utilizada principalmente pela possibilidade de capitalizar os visitantes das páginas que venham a disponibilizá-los.

Com o tempo, surgiram páginas que aproveitavam vários destes formulários ao mesmo tempo, num trabalho de colagem feito, muitas vezes, por pessoas com um conhecimento mínimo de programação e dos processos que estão sendo acionados³. Esse tipo de página aparece inclusive como agregado de sites institucionais de orientação pública, como na Biblioteca Virtual da USP⁴. Posteriormente diversos designers passaram a retirar desses formulários apenas a caixa de entrada de texto visando disponibilizar a consulta com adequação à programação visual da páginas por eles criadas⁵. Em outro nível aparecem casos onde se conseguiu condensar o acionamento de múltiplas ferramentas de busca a partir de uma única entrada de texto⁶ (Dreiliger e Howe, 1997).

¹ De um lado temos a característica essencialmente simbólica do *software*, de outro, o fato de que, atualmente, ele praticamente já não depende mais de suportes físicos para sua distribuição. Assim a possibilidade de exploração comercial do *software* passa a necessitar, de modo fundamental, da instituição do direito autoral.

² No caso do Altavista, por exemplo, ele pode ser encontrado em http://www.altavista.com/sites/search/bf_freeseachbox).

³ Alguns exemplos podem ser encontrados em: <http://www.geocities.com/Broadway/Booth/2992/busca.htm>, <http://www.geocities.com/Area51/Zone/6197/search.html> e <http://www.multserv.com.br/links.htm>.

⁴ <http://www.bibvirt.futuro.usp.br/busca/busca.html>

⁵ Um caso desse tipo pode ser encontrado em <http://www.cosmo.com.br/busca/>.

⁶ Alguns dos mais antigos casos desse tipo seriam o All in One e o Meta search, de 1995, e o Pro Fusion de 1996.



Foi relativamente fácil dar esse último passo já que as formas de acionamento da maioria das ferramentas de busca ficam bastante explícitas, tanto nos códigos dos formulários disponibilizados por elas, quanto no seu próprio uso direto. De maneira geral elas usam os métodos típicos de retorno de formulários do padrão HTML, fazem o navegador requisitar uma página cujo endereço mescla a localização do software de busca com os parâmetros da própria busca. Assim basta conhecer os padrões desse endereço para saber como acionar a busca em cada caso. Seria preciso apenas realizar algumas requisições de busca mudando os parâmetros para conhecer esses padrões. Em resumo, pelo fato desta forma de acionamento ser aberta, disponível para análise, se tornou fácil criar novas ferramentas que conseguem consultar esses sites de busca, mesmo quando eles não disponibilizam o código de formulários para facilitar esse procedimento. E, em princípio, o software da ferramenta de busca consultada não consegue diferenciar se está recebendo um acionamento direto de seu formulário ou de outro instalado numa página alheia.

Com alguns adendos simples é possível evitar que a página de resultados simplesmente substitua a página usada no acionamento. O Blue Book⁷ as coloca em novas janelas do navegador. Um formulário disponibilizado pelo site hacker Astalavista as coloca em partes de um *frameset*⁸. Com isso quem disponibilizou o acionamento pode continuar oferecendo seu conteúdo em paralelo com os resultados. Pode-se, assim, contar com o interesse gerado pelas ferramentas de busca para promover conjuntamente outros conteúdos.

Agentes

Esse tipo de proposta de acionamento remoto de ferramentas de busca foi a base fundamental de alguns projetos mais amplos que surgiram no meio universitário. Por volta do ano de 1995 o ambiente acadêmico americano acompanhava a primeira onda de grande crescimento da Web. O fenômeno chamava a atenção para a necessidade de ferramentas que auxiliassem a dar conta do enorme volume de informação que começava a se apresentar. Algumas das propostas que apareciam com frequência envolviam aplicativos chamados de “agentes”⁹ para auxílio à navegação na rede.

Uma pesquisa conduzida nessa época por Erik Selberg, estudante de graduação da Universidade de Washington, revelou as ferramentas de busca disponíveis retornavam resultados diferentes para uma mesma consulta. Isso se dava porque suas bases de dados e os métodos de pesquisa aplicados sobre elas não coincidiam. Uma parte fundamental dessa questão vinha do fato de que nenhuma delas cobria toda rede (Na verdade, elas cobriam um máximo de 15%). E cada uma cobria grupos diferentes de arquivos, retratados em momentos diferentes. Além disso cada uma atribuía relevância e organizava os resultados de diferentes formas. (Etzioni e Selberg,1996)

Para equacionar em parte esses problemas Selberg propôs a disponibilização na Internet de uma interface para simplificar o acesso dos usuários às ferramentas de busca. Assim na mesma Universidade de Washington onde a primeira ferramenta de busca bem sucedida havia nascido (o Webcrawler) foi disponibilizado, em 1995, o Metacrawler. Ele era primeiro exemplo efetivo das chamadas ferramentas de Meta-busca, hoje citadas também como agentes (Dreiliger e Howe,1997) ou mediadores (Tzitzikas,2000).

O Metacrawler opera com a mesma proposta do acionamento de múltiplas ferramentas de busca a partir de um único formulário já vista anteriormente. Mas nele foram adicionados diversos outros elementos. Seu software requisita, para si, as respostas das diferentes ferramentas, separa as informações das partes do código relativas a formatação, condensa os resultados, elimina a redundância (resultados apontando para os mesmos links). Ele verifica se as páginas indicadas ainda estão disponíveis e ordena a lista que obtém com o somatório das relevâncias declaradas por cada ferramenta ao retornar cada resultado. Por fim entrega a seus usuários páginas de resposta formatadas por ele contendo essa compilação.

Como foi dito uma das principais proposições do Metacrawler seria oferecer uma interface mais simples para o conjunto das principais ferramentas de busca. Ele procura diminuir, por exemplo, a dificuldade de se entender como determinar a interrelação que se deseja entre cada termo de pesquisa em

⁷ <http://www.indexbrasil.net/bluebook/>

⁸ Frameset é um padrão de código para que se exiba múltiplas páginas numa única janela de *browser*, um programa de navegação na Internet. Foi criado para permitir manter a amostragem de determinados conteúdos ou controles enquanto outros vão sendo trocados.

⁹ Para uma introdução ao conceito de agente para a Internet sugerimos Vaz,1999



cada ferramenta (se desejamos que todas as palavras que inserimos apareçam juntas, se uma ou outra, se uma excluindo outra, se queremos pesquisar a ocorrência de uma frase etc). No Metacrawler a maioria dessas opções aparecia numa simples caixa de seleção. Outra intenção era evitar que o usuário fosse obrigado a saber selecionar as melhores ferramentas em função do que buscava. Era dada a ele uma cobertura mais completa da rede num único procedimento. (Etzioni e Selberg,1996)

Posteriormente os endereços oferecidos nas respostas passaram a ser redirecionados pela própria ferramenta: Quando se requisita um endereço listado nos resultados oferecidos é feita uma chamada ao servidor do Metacrawler que, por sua vez, requisita o endereço final. A vantagem deste procedimento seria a capacidade de registrar que respostas foram úteis em função de cada busca permitindo refinar o sistema e, ao mesmo tempo, pesquisar os hábitos dos usuários.

O importante para esse estudo é que o fato do código HTML ser intrinsecamente aberto propiciou essas possibilidades que estivemos retratando acima. Como em qualquer página construída nessa linguagem, é fácil extrair a informação essencial (o texto sem formatação) também das páginas de resposta das ferramentas de busca. Com isso é possível aplicar a essa informação qualquer outra formatação e, principalmente, condensar e reprocessar os resultados de diferentes destas ferramentas.

É interessante lembrar que a reapropriação operada pelas ferramentas de meta-busca não ocorreu sem crises. O Metacrawler teve problemas com seus “fornecedores” involuntários. Apesar dele citar suas fontes em suas respostas aos usuários, os administradores das ferramentas de origem argumentaram que obtinham seus recursos da publicidade, e que ele se utilizava de suas bases de dados e de seu processamento sem permitir que eles exibissem seus patrocinadores. A questão foi resolvida através de acordos para que o Metacrawler exibisse anúncios gerados pelas ferramentas de buscas consultadas. (Sonnenreich,1997)

Pouco depois do surgimento do Metacrawler apareceram outras propostas de meta-busca valendo citar o Savvy Search¹⁰, da Universidade do Colorado, que tinha como diferencial o fato de selecionar automaticamente ferramentas de busca de diferentes protocolos (Além da Web, Usenet, newsgroups, notícias, e-mails e outros) de acordo com as palavras consultadas. Ele foi criado para “aprender” dinamicamente a definir os melhores serviços para achar cada informação e restringir as buscas que realiza ao pequeno grupo de ferramentas que eleger para cada caso (Dreiliger e Howe,1997). Já que sua proposta era trabalhar com o conceito de agente, a idéia de adaptabilidade, de capacidade de “aprendizado” e “memória” seria um aspecto fundamental. No Brasil o Minner¹¹, também surgido da pesquisa universitária, seria nosso principal representante no contexto das meta-buscas.

Reaproveitamento

Se analisarmos mais detidamente todo o processo formador das relações encontradas nas ferramentas de meta-busca poderemos ver uma presença ainda mais extensiva das noções de reaproveitamento e recondução de conteúdo, interfaces e processamento: Para tal precisaremos caracterizar melhor as ferramentas de buscas que servem como suas “fontes”.

Existiriam dois tipos principais de ferramentas de busca: De um lado os diretórios, grupo que teve como um de seus pioneiros o Yahoo, e inclui o Magelan, o Open Directory e o brasileiro Cadê. Eles se caracterizam por funcionar basicamente como ferramentas de comunicação onde quem quer anunciar a existência de algum material publicado na Internet faz uma descrição sucinta, esta é anexada a um sistema categorizado disponível para a pesquisa. Assim os diretórios são mediadores de anúncios da publicação de conteúdos.

Importa para esta análise que haveria neles a acumulação e organização do trabalho de resumo e definição de conteúdos publicados, feito pelos próprios usuários. Há também o registro do esforço de categorização desses resumos. E, ainda, o da produção das próprias interfaces de captura dessas contribuições e do sistema de processamento das consultas.

Um segundo grupo de ferramentas de busca seria o das chamadas *spiders*. O primeiro exemplo foi o Webcrawler, mas outros como Lycos, Altavista, Hotbot e Google, entre diversos, adquiriram importância posteriormente. Essas ferramentas se baseiam em *softwares* que se mantêm fazendo um levantamento automatizado dos documentos publicados na Internet e anexando seu conteúdo textual a bases de dados

¹⁰ Incorporado pelo grupo de mídia CNET em <http://www.search.com>

¹¹ Para conhecer melhor o projeto Minner recomendamos a tese em que seu criador descreve o processo de sua concepção (Ribeiro, 1998)



pesquisáveis. Elas já tem, a priori, uma grande dependência do fato do HTML ser aberto. Este é um dos fatores primordiais para a extração, independente de qualquer autorização, do conteúdo depositado por todos que publicaram na rede, para ser anexado e organizado em suas bases de dados. Nesse sentido elas condensam, organizam, acumulam a informação depositada na rede, e mediam a nossa dificuldade de localização definida pela imensa quantidade de documentos. Nelas também fica registrado o esforço de criação de sistemas de consulta, de processamento e organização dos resultados.

As ferramentas de meta-busca somariam mais uma camada unificando, reorganizando e reaproveitando tudo que citamos acima. Nelas vemos, portanto, a acumulação da memória inscrita em milhões de pequenos atos: Começando pelo de publicar documentos na rede e de indicar neles outros documentos, através de *links*. Passando, também, pelo de resumir seu conteúdo para os diretórios, pelo de categorizar esses resumos e pelo de criar as interfaces de captação dessas inscrições. Estariam incluídos, ainda, o esforço de criar, nas *spiders*, sistemas automatizados de recuperação desses conteúdos para a formação de bases de dados, e o de disponibilizar ferramentas para pesquisá-las com métodos de atribuição de relevância para os resultados. E, por fim, vem o trabalho somar, organizar, esses esforços anteriores, procurando disponibilizar formas simples de requisitá-los, todos de uma única vez, nas ferramentas de meta-busca. Sendo que a criação destas embutiu um progressivo reaproveitamento dos formulários disponibilizados pelas primeiras.

Lembramos, novamente, que boa parte de toda essa cadeia foi uma conseqüência direta de um fator externo, abertura do código intermediário envolvido, o HTML. Ela ajudou a estabelecer a possibilidade de se reaproveitar informação, interface e processo, em tempo real, ou seja, de pensar a rede um pouco como um grande computador.

Nos interessa de modo especial o fato de que a maior parte desses processos de reaproveitamento foi feita de modo não coordenado. Nas últimas etapas da cadeia vista acima vão sendo criadas ferramentas, que na verdade englobam as ferramentas anteriores. Há uma espécie de invenção técnica contínua por somatório, que vai se dando pelos movimentos de diversos diferentes atores, sem que haja um planejamento inicial e uma coordenação desses gestos.

Esse contexto parece atender às noções encontradas nos estudos da complexidade, como nas análises de sociedades de insetos capazes de comportamentos globais elaborados oriundos da associação de comportamentos simples. A nosso ver, nos processos que viemos acompanhando, também encontramos padrões associados a sistemas complexos como: Comportamentos elaborados que emergem diretamente da articulação coletiva de ações sem uma coordenação central. Elementos desempenhando tarefas complexas cooperativamente pela conjunção de movimentos especializados, que remetem aos feitos por seus pares, e estabelecem maior sofisticação quando observados conjuntamente (Johnson, 2001, p. 18). E, também, interação de ações de escopo local produzindo um padrão ordenado num nível mais amplo, estabelecendo novas soluções sem que haja coordenação central (Johnson, 2001, p. 93).

Nas sociedades de insetos esses padrões dependeriam da adaptação ativa dos comportamentos locais através da comunicação por ferormônios. No caso dos processos que acompanhamos estas estruturas parecem estar ligadas a troca de criações culturais ocorrida em meio aos múltiplos reaproveitamentos que descrevemos, que teriam como principal sistema de sinalização o código de programação. O código aberto seria um fator de coesão que propiciou o surgimento de um sistema social complexo, capaz de engendrar novas relações de comunicação pela articulação de diversas ações.

Em princípio, isso indicaria que, assim como nos insetos, para estudar o desenvolvimento desse tipo de sistema é preciso estudar os padrões de ação local, ou seja, os processos singulares de reaproveitamento. Também nos parece interessante analisar o quanto esses tipos de formações, constituídas “de baixo para cima”, indicam possibilidades de contraposição às tendências de concentração tipicamente retratadas no universo da comunicação. Seria interessante interrogar quais de seus aspectos apareceriam como possibilidades socializantes e democratizantes consistentes, que poderiam ser potencializadas e reproduzidas. Procurando, também, evitar cair em frágeis utopias positivistas.

Facilitações

De momento parece que já podemos demarcar que estes processos não são necessariamente esotéricos, não estão acessíveis somente aos engenheiros de software. Na Internet, de forma semelhante a do *software* livre, estes mesmos mecanismos, originados no aspecto simbólico essencial dessas ferramentas, também estabelecem espaço para que atores sem muitos recursos e nem conhecimentos técnicos tenham possibilidades de participação.



Eles têm ganho, inclusive, diversas formas de suporte, utilizando a própria capacidade de tradução e organização do ambiente computacional. Hoje, vemos na Internet uma série de repositórios, de uso gratuito, que procuram catalogar e disponibilizar a maioria dos elementos que são costumeiramente encontrados em páginas da rede, e ensinam a utilizá-los. Encontramos uma série de tutoriais sobre programação HTML, que frequentemente disponibilizam uma grande quantidade de código pronto para recortar e colar. Neles podem ser localizados modelos pré-fabricados, alteráveis, de quase todos elementos tipicamente constituidores de páginas. Inclusive aplicações mais rebuscadas, como as que operam com acionamento de bases de dados. Nem mesmo o conteúdo em si escapa de ser franqueado para a montagem. Existem serviços que disponibilizam códigos para a exibição de material requisitado de fontes remotas através da rede. A forma mais comum é a exibição contínua de manchetes geradas por serviços noticiosos. (Vouga, 2002)

Essa questão da utilização dos próprios recursos da rede para franquear a autoria em seu meio expressivo foi aprofundada com o estabelecimento de grandes sites comerciais dedicados. Como exemplos significativos podemos citar o Geocities (atualmente pertencente ao Yahoo), o Homestead, o Angelfire e o Tripod (Esses dois últimos encampados pelo grupo Terra Lycos). Neles se buscou automatizar e simplificar os processos de construção de páginas para leigos através de ferramentas visuais intuitivas, com grande preocupação didática, às custas de uma certa padronização. (Machado, 12/3/2001, p. 1-4) (N.A., 18/03/2001)

Essa perspectiva de tradução da construção de interfaces em metáforas visuais intuitivas também apareceria em aplicativos como os editores HTML gráficos (como o Microsoft Front Page e o Macromedia Dreamweaver). Eles procuram tirar proveito das interfaces gráficas dos sistemas operacionais atuais para facilitar a criação das interfaces para a rede.

Escolha

Então, se a disponibilização de ferramentas de comunicação tem se mostrado uma forma estratégica de operação na rede, entre os aspectos relevantes estariam incidindo sobre ela podemos listar a capacidade de tradução, de circulação e de reaproveitamento do próprio ambiente computacional. Estes aspectos permitiriam pensar esse processo como uma bricolagem daquilo que já se encontra atualmente disponível na rede, facilitada por interfaces de criação de interfaces.

Lembramos que a concentração do poder de comunicação foi tomada, em diversos momentos, como um aspecto problemático nas mídias de massa. Nos interessa demarcar que a conscientização das potencialidades contidas no fato do HTML ter uma estrutura aberta pode ajudar no sentido de que este jogo estratégico, a proposição de novas ferramentas, conte com a participação de um maior volume e variedade de atores sociais, evitando sua concentração em poucas mãos e seus possíveis efeitos.

No entanto lembramos que nada disso está previamente assegurado. O padrão aberto já não é mais exclusivo na rede e não se pode saber se permanecerá. Nikolai Bezroukov (1999) lembra que a Internet, assim como o movimento pelo *software* livre, teve sua origem ligada ao meio acadêmico onde o conhecimento é produzido de forma a facilitar sua acumulação e reaproveitamento. Na medida em que a rede vai sendo tomada por interesses comerciais cabe pensar se seus padrões de código aberto não serão progressivamente modificados.

Já começam a haver casos de sucesso de tecnologias fechadas. Por exemplo, o *software* de animação vetorial Flash da Macromedia, que gera interfaces que não permitem acesso a seu código constitutivo, tem um plug-in para navegadores Internet que já estaria sendo usado por mais de 250 milhões de usuários. Ele vem se tornando uma referência em termos do design na rede. Ao mesmo tempo a ferramenta de buscas Altavista implementou um método, por leitura de senha em imagens, para evitar que o seu sistema aceite sugestões de novos sites que não venham de seu próprio formulário.

Também a simples presença de interfaces de auxílio à exploração do padrão aberto não define automaticamente possibilidades desalienantes. Se, como Lévy (1993) colocou, a produção de interfaces vem ganhando interfaces, se estabelecerá um novo problema. Como vimos, se de um lado as interfaces nos oferecem maior capacidade de operação intuitiva, é importante lembrar que elas também estabelecem limites, recortes, causam perdas na visão em detalhe etc. As ferramentas de auxílio a produção e publicação de material para rede, se abrem potencialidades de incorporação de uma maior número de participantes, sempre impõem, também, limitações. Seria preciso ter consciência dos ganhos e perdas, em cada caso, e fazer opções.



INTERCOM – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação
XXV Congresso Brasileiro de Ciências da Comunicação – Salvador/BA – 1 a 5 Set 2002

E mesmo o aparecimento de padrões auto-coordenados, que citamos, não é necessariamente socializante. Um lembrete interessante é que entre os exemplos citados, dentro das teorias sociais que se utilizam de conceitos de auto-coordenação, estaria o pensamento econômico de Adam Smith. (Johnson, 2001, p. 18)

Nos anos 60, o teórico tcheco Vilém Flusser foi um dos primeiros a detectar que, em função do advento das novas tecnologias da informação, haveria uma tendência a problematização das formas modernas de relação com as ferramentas técnicas. Uma de suas principais exortações era a de que devemos nos engajar numa relação dialógica com a técnica, que não nos sujeitemos (*unterwerfen*) mas que projetemos (*entwerfen*) nossas próprias máquinas (Röller, 16/12/01, p. 12-13). Como vimos parece ter se estabelecido um interstício, sob tensão mas ainda aberto, que talvez possibilite exatamente o aumento dessa relação dialógica com o ferramental. Parece que cabe a nós tentar estabelecer os mecanismos para explorá-lo. Algumas armas parecem já estar disponíveis. Talvez tenhamos que construir outras. Quem sabe, coletivamente.



INTERCOM – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação
XXV Congresso Brasileiro de Ciências da Comunicação – Salvador/BA – 1 a 5 Set 2002

Bibliografia

ALBUQUERQUE, Afonso - Comunicação Política: o desafio das tecnologias da comunicação e da informação - Anpocs 2001, Niterói, 2001

ALBUQUERQUE, Afonso - Um Outro "Quarto Poder": Imprensa e Compromisso Político no Brasil. Publicado em: Revista Contracampo - Mestrado em Comunicação - UFF, Niterói, Janeiro de 2000

BEZROUKOV, Nikolai - Open Source Software Development as a Special Type of Academic Research - site da Organização Softpanorama, <http://www.softpanorama.org>, 1999

DREILIGER, Daniel e HOWE, Adele - SavvySearch: A Meta-Search Engine that Learns which Search Engine to Query - site do Centro de Tecnologia da NEC Corp., <http://citeseer.nj.nec.com>, 1997

ETZIONI, Oren e SELBERG, Erik - The MetaCrawler Architecture for Resource Aggregation on the Web - site do Centro de Tecnologia da NEC Corp., <http://citeseer.nj.nec.com>, 1996

JOHNSON, Steven - Emergence: The Connected Lives of Ants, Brains, Cities and Software - Scribner, New York, 2001

JOHNSON, Steven - Interface Culture - Basic Books, São Francisco, 1997

LEROY-GOURHAN, André - O Gesto e a Palavra - Edições 70, Lisboa, 1964

LÉVY, Pierre - As Tecnologias da Inteligência - O futuro do pensamento na era da informática - Editora 34, Rio de Janeiro, 1993

MACHADO, André - Monte sua C@sa. publicado em: Jornal "O Globo" - Caderno Informática Etc - , Rio de Janeiro, 12/3/2001

N.A. - Yahoo Geocities Help - site do Yahoo-Geocities, <http://help.yahoo.com>, 18/03/2001

RIBEIRO, Victor Ribeiro - A família Miner de agentes para o world-wide web - Belo Horizonte: Curso de Pós Graduação do Departamento de Ciência da Computação da UFMG , 1998. (Dissertação de mestrado).

RÖLLER, Nils. - Um Platão da Era dos Computadores - Publicado em: Caderno Mais - Folha de São Paulo. São Paulo, 16/12/2001. P.12-13

SONNENREICH, Wes - A History of Search Engines - site da editora John Wiley & Sons, <http://www.wiley.com/legacy/compbooks/sonnenreich/history.html>, 1997

STALMAN, Richard - The GNU Manifesto - Site da Free Software Foundation, <http://www.gnu.org/gnu/manifesto.html>, 7/9/00

TZITZIKAS, Yannis. - Democratic Data Fusion for Information Retrieval Mediators - site do Centro de Tecnologia da NEC Corp., <http://citeseer.nj.nec.com>, 2000

VAZ, Paulo - Mediação e tecnologia - Trabalho apresentado no GT - Comunicação e Tecnologias da Informação XXIV INTERCOM, Campo Grande, 6/9/2001

VOUGA, André - As meta-buscas e o código aberto : Autoria e reaproveitamento nas ferramentas de comunicação da rede. Niterói. Mestrado em Comunicação e Tecnologia da Informação IACS UFF, 2002. (dissertação de mestrado)