

Um estudo sobre google: questões para uma leitura micropolítica das tecnologias da informação

Cecilia Diaz-Isenrath*

Resumo Este artigo descreve um motor de busca, não como ferramenta informacional, mas como dispositivo socio-técnico. Para mapear o meio no qual esse dispositivo toma forma, analisamos um artigo científico sobre o Google e a patente onde o método conhecido como “PageRank” é documentado. Problematizamos um tipo de relação com a tecnologia, a partir de questões tais como: quais os pressupostos que essas técnicas envolvem? Como é pensado o “governo” de pessoas e coisas? Como incide o mercado no desenho desses sistemas? Este texto procura contribuir com as pesquisas sobre as relações humano-máquina, investigando, particularmente, reconfigurações da técnica e da linguagem.

Palavras-chave motores de busca, redes socio-técnicas, técnica e linguagem, governamentalidade, mercado

Abstract This article provides a description of a search engine, not as an informational tool, but as a socio-technical device. In order to map the milieu where that device takes form, we analyze a scientific paper about Google and the patent where the method known as PageRank is documented. We problematize a particular kind of relation with technology, by posing the following questions: what is being taken for granted in those techniques? How is the “government” of persons and things thought? How does the market drive the design of those systems? This text aims to contribute to the researches about human-machine relations, tracing, in particular, some re-configurations of technique and language.

Keywords search engines, socio-technical networks, technique and language, governmentality, market

Das associações lógicas às sociotecnológicas

Dois documentos servirão como fio condutor da nossa análise: um artigo científico e uma patente. O primeiro foi um trabalho apresentado na Sétima Conferência Internacional da World Wide Web, em abril de 1998. Esse *paper*, intitulado “The anatomy of a large-scale hypertextual web search engine”, era assinado por dois engenheiros em computação, naquele momento doutorandos na Universidade de Stanford, Lawrence Page e Sergey Brin. Eles expunham, ali, as dificuldades de aplicar as técnicas tradicionais de recuperação de dados a repositórios de grande escala como a *web*, e apresentavam o protótipo de uma máquina ou motor de busca, cujo objetivo principal era prover resultados de alta qualidade às pesquisas dos usuários, numa *world*

* Cecília Diaz-Isenrath é Doutoranda em Ciências Sociais IFCH/UNICAMP.

wide web em rápido crescimento. No trabalho, Page e Brin, descrevem a anatomia do sistema, os mecanismos automatizados para rastrear a *web*, as técnicas para armazenar e indexar os documentos e aquelas para melhorar a qualidade das buscas e o ordenamento dos resultados. Nesse momento, o buscador, disponível através do *site* da Universidade de Stanford, dispunha de um banco de dados de textos na íntegra interligados mediante *hyperlinks*, de 24 milhões de páginas (Cf. PAGE; BRIN, 1998a; PAGE; BRIN, 1998b).

A patente que consideraremos aqui, depositada alguns meses antes junto ao escritório de patentes americano, reivindica a invenção de um método para classificar e ordenar nós num banco de dados em hipertexto. Esse método será conhecido como PageRank – por Page, a quem se atribuiu a invenção (BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY, 1998).

Se fizermos esta retrospectiva, retrocedendo, no entanto, a uma história bem recente (menos de uma década), não se trata de traçar a trajetória dos inventores do motor de busca Google ou da empresa com o mesmo nome.¹ Longe de nos centrarmos nos indivíduos ou numa organização, importa aqui analisar a evolução de um artefato, um objeto baseado em código, e explorar em que sentido podem ser estudadas, desde uma perspectiva molecular, micropolítica, as chamadas tecnologias da informação. Para isso, e nos apoiando nos documentos acima citados, faremos o exercício de mapear algumas das relações sociotécnicas nas quais esse objeto, o motor de busca, participava em 1998.

Tanto a seção dos agradecimentos no *paper* quanto os créditos na interface para recuperação de dados do protótipo do buscador, remetem a uma multiplicidade de materiais e atores acadêmicos, técnicos, administrativos, econômicos. Seguindo essas referências e créditos, podemos fazer um primeiro traçado da rede na qual o protótipo se inseria: dos colegas, professores e grupos do Departamento de Ciências da Computação em Stanford às empresas doadoras de equipamento (IBM, Intel e Sun), passando pelas instituições financiadoras (NSF, DARPA, NASA) e parceiras do Projeto de Biblioteca Digital daquela universidade, projeto do qual o Google era um desmembramento.² Essas relações formam uma trama de elementos que, nesta análise, não procuraremos separar de acordo com demarcações institucionais ou disciplinares.³ Antes que separar *a priori* setores (público e privado) ou mesmo áreas do conhecimento (matemática, ciências da computação, sistemas de informação, biblioteconomia, gestão da C&T, etc.), interessa-nos pensar deslocamentos, operações de empréstimo e refuncionalização de problemas, conceitos e técnicas. A noção de “rede sociotécnica”, que usamos aqui no sentido de Latour e Callon, refere-se a esse tipo de deslocamentos, alianças, transformações, traduções. Enfatiza, antes de tudo, o aspecto coletivo da construção de fatos e máquinas, isto é, a mobilização e engajamento de elementos heterogêneos que essa construção envolve.⁴

¹ Pode-se encontrar uma história do Google centrada na figura dos “inventores” e “empreendedores”, o americano Lawrence Page e o moscovita naturalizado norte-americano Sergey Brin, no próprio site da empresa (<http://www.google.com/corporate/index.html>).

² Entre os grupos de pesquisa da universidade são mencionados, entre outros, The Database Group e The MIDAS Data Mining Group (para maiores informações sobre esses grupos, podem-se consultar as páginas na Universidade de Stanford, via Internet Archives). Entre as organizações e empresas parceiras do Projeto de Biblioteca Digital, encontramos: Interval Research, Association for Computing Machinery, Bell Communications Research, Hewlett-Packard Laboratories, Hitachi, Hughes Research Laboratory, The MIT Press, NASA, Ames Artificial Intelligence Group, Xerox Palo Alto Research Center.

³ Nesse sentido, Latour propõe que “convém deixar de lado todos os preconceitos sobre as distinções entre o contexto em que está inserido o saber e o próprio saber”. Sobre essa inseparabilidade do contexto e do conteúdo, cf. Latour (2000, p. 266 et seq).

⁴ Os estudos da ciência e a tecnologia realizados por Bruno Latour, Michel Callon e outros foram conhecidos como “teoria de redes de atores” (actor-network theory), rótulo que o próprio Latour reconheceu ter sido bastante infeliz

Lendo o *paper*, entretanto, resulta claro que os interlocutores diretos da apresentação de Page e Brin eram outros pesquisadores, especialistas no campo das ciências da computação e da inteligência artificial, presentes na conferência (www7). Na sessão “Pesquisa acadêmica sobre motores de busca“, Page e Brin declaravam que, apesar da importância dos motores de busca de grande porte na *web*, pouca pesquisa acadêmica teria sido feita sobre eles. Trabalhos em áreas específicas como a de recuperação de informação teriam sido desenvolvidos desde há muitos anos. Entretanto, a maior parte dessas pesquisas teriam sido conduzidas em coleções “bem controladas” que abrangiam um universo “reduzido” e “homogêneo” (como uma coleção de *papers* científicos ou de notícias sobre tópicos particulares). Coisas que funcionavam bem nessas coleções “freqüentemente deixam de produzir bons resultados na *web*”. Nesse sentido, afirmam Page e Brin, as técnicas usadas para lidar com “ambientes controlados” e “sistemas fechados de informação” precisavam ser revistas, com vistas a lidar com coleções de larga escala, onde há pouco ou nenhum controle sobre os conteúdos publicados e onde os documentos apresentam uma extrema variação, “tanto interna aos documentos como nas meta-informações externas que possam estar disponíveis”.

Os autores do paper apontavam outro aspecto característico dessas coleções de larga escala e “menos controladas”. Acompanhando o que seria uma tendência geral da *web*,

os motores de busca migraram do campo acadêmico para o comercial. Até agora, a maior parte do desenvolvimento de motores de busca foi realizado em empresas, com pouca publicação de informações técnicas. Isso faz com que a tecnologia de máquinas de busca permaneça em grande medida uma arte obscura e seja orientada à publicidade.

Para apoiar esse argumento, Page e Brin citam a Mauldin, cientista-chefe da Lycos Inc., que afirmava que “os vários serviços (incluindo Lycos) guardam com cuidado os detalhes de seus bancos de dados”. O projeto do Google tinha o firme propósito de mudar essa situação. O paper fornecia, argumentavam seus autores, “a primeira descrição pública e detalhada desse tipo de ferramenta”.

É a partir dessa descrição “pública” e propriamente “técnica” que podemos estender a nossa rede, a nossa própria descrição da ferramenta de busca. Assim, encontramos toda uma série de “actantes”, de entidades humanas e não humanas, que de uma ou outra maneira estavam sendo conectadas, “alinhas” e postas a funcionar e colaborar com o projeto.⁵ Isto é, além dos colegas,

(LATOURE, 2003). O uso que esses autores fazem da noção de “rede”, certamente muito usada e abusada nos últimos anos, diferencia-se de outras perspectivas sociológicas que excluem qualquer componente não humano. A problemática colocada por eles é, em todo caso, a de como efetuar análises empíricas dos processos sociotécnicos descentradas do humano.

⁵ A noção de “actantes” refere-se a objetos discursivos, às figuras num relato, capazes de designar uma pessoa, uma máquina, um animal, um organismo biológico ou outro tipo de entidade. Essa noção foi tomada de empréstimo da semiótica e empregada por Latour e seus colaboradores como categoria para lidar de forma “simétrica” entre atores humanos e não humanos, no contexto dos estudos sociais da ciência e da tecnologia. Os actantes, segundo Latour, são qualquer coisa ou pessoa capaz de ser representada. Em se tratando de multidões (de coisas e pessoas) cuja comunicação e representação direta é impossível ou, no mínimo, pouco praticável, os “porta-vozes” tornam-se indispensáveis. É claro que os representantes podem trair os interesses dos representados (e, nesse sentido, toda tradução envolve “traição”, transformação) e que, por sua vez, esses últimos também podem resistir ou trair seus porta-vozes. Para exemplos de processos de “tradução” e dos “actantes” mobilizados nos laboratórios e centros de pesquisa, cf. Latour (2000, p. 119, 138). A discussão sobre o papel dos “actantes” na teoria semiótica pode se

cientistas e técnicos, poderíamos reconhecer também o “trabalho” de toda uma série de pessoas e coisas:

- os servidores, seus discos de armazenamento e outros dispositivos de *hardware* e mídia de telecomunicações; as empresas doadoras desses equipamentos que participam de “consórcios” e “projetos em parceria” com a universidade, os representantes dessas partes;
- os sistemas operacionais em que os servidores de Google rodam: o Solaris, sistema UNIX desenvolvido pela Sun Microsystems, e o GNU/Linux; as pessoas e organizações que representam os softwares proprietários ou os de código aberto;
- os programas que, “trabalhando” em paralelo, “rastream” automaticamente a *web* para recuperar e indexar páginas; os milhares de outros servidores que disponibilizam documentos na *web* e com os quais essas aplicações de rastreamento interagem;
- as instruções de C e C++, linguagens de programação com que foi desenvolvido o software do Google, os programas para desenvolvimento de aplicações nessa linguagem, os desenvolvedores e usuários dessa linguagem;
- os programas utilizados para gerenciar bancos de dados, que Google utiliza para organizar e manter atualizado seu repositório de documentos, o banco de *links*, o lexicon e outras compilações; as pessoas e empresas que desenvolvem, utilizam e “representam” os programas para gerenciamento de bancos de dados;
- os programas visualizadores de documentos escritos em hipertexto (*browsers*); os pesquisadores, centros de pesquisa e empresas que os desenvolvem;
- os *hyperlinks* e rótulos ou etiquetas (*tags*) usados nas linguagens de marcação que descrevem a visualização e estrutura dos documentos na *web*; os pesquisadores, centros de pesquisa, grupos de trabalho e organismos que especificam, negociam e representam esses padrões (ANSI, ISO, W3C, etc.);
- os 24 milhões de documentos (páginas *web*) indexados pelo Google em 1998; os que escrevem essas páginas e administram *web sites*, as ferramentas para criação e edição de páginas *web*; os espaços e serviços para troca de recursos de programação e edição;
- os diretórios de documentos mantidos por humanos (Yahoo) e as ferramentas de busca que, em 1998, empregavam “robôs” (www, Lycos, Altavista); os usuários (humanos e *softwares*) que fazem consultas nessas compilações;
- o corpus de palavras e a sintaxe da linguagem “natural” (o inglês) que o protótipo do Google e as outras máquinas de busca podiam processar em 1998; os estudiosos e os falantes dessa língua;
- o TCP/IP e outros protocolos para intercâmbio de informação em redes digitais privadas e públicas; outras tecnologias e serviços da internet (*e-mail, mailing lists, telnet, ftp; etc.*), seus usuários e representantes; as organizações que gerenciam o funcionamento da internet (distribuição de endereços IP e sistema de nomes de domínio); os provedores de acesso; os *backbones* e gestores de pontos de tráfego da internet; etc.

encontrar, especialmente, na obra de A.J. Greimas. Para uma análise que retoma a pragmática de C. S. Peirce e trata de regimes de signos que remetem a agenciamentos não necessariamente lingüísticos, cf. Deleuze e Guattari (1995-1997, v. 2, p. 61-107).

Até que ponto faz sentido essa extensa enumeração? Palavras e linguagens humanas, instruções e linguagens de codificação, documentos, servidores, diretórios, bancos de dados, etiquetas, *links*... e todas as pessoas e coletivos que representam e/ou são representados por esses objetos e produções. De um ponto de vista estritamente “técnico” essa enumeração seria, sem dúvida, bastante discutível. Por outro lado, fica evidente que poderíamos continuá-la, estendê-la ainda mais, com o risco de que “o tamanho do mapa seja maior que o do território”. Contudo, nossa “rede” não visa a ser exaustiva. Procuramos mapear um sistema de associações, o agrupamento dos “recursos” que estavam sendo mobilizados na construção de um sistema de busca. Interessa sublinhar que o processo de construção de máquinas é, sempre, coletivo. A realidade do objeto técnico, a sua realização progressiva, depende do número de elementos amarrados, elementos que produzirão o instrumento e, ao mesmo tempo, serão por ele incorporados.⁶

Nossa própria descrição, podemos pensar, é um instrumento para acompanhar o processo de invenção coletivo no qual um objeto técnico toma forma. Na medida em que entendemos que não existe uma origem ou um sujeito do conhecimento, que nenhum objeto é inventado *ex nihilo*, devemos necessariamente “começar pelo meio”, para usar uma conhecida expressão de Gilles Deleuze. E, em se tratando do trabalho tecno-científico, esse “meio” se estende a um universo dilatado de elementos, reunindo matérias, coisas e pessoas, tanto cientistas e técnicos quanto “leigos”. Cabe notar que, no nosso exemplo, esse meio seria, inseparavelmente, o entorno material, institucional e *milieu* intelectual dos fundadores da Google, assim como o meio tecno-econômico dos dispositivos de hardware e software.

O nosso mapa de relações é composto, portanto, de séries de elementos heterogêneos. Não restringindo a análise às representações e ações humanas, podemos tratar de “associações”, nas quais certas pessoas atuam – nos termos de Latour – como “porta-vozes”, falando no lugar não só de outras pessoas, mas também de máquinas e artefatos. Nesse sentido, na listagem acima, distinguimos uma série de pessoas e coisas (“actantes”) que estavam sendo representados por determinados “porta-vozes”: usuários, pesquisadores, *webmasters*, documentos, servidores, discos de armazenamento, rótulos, *links*, etc. As próprias qualidades dessas pessoas e coisas são objeto de alianças e negociações: alguns requerem dos discos maior “eficiência” no uso do espaço de armazenamento e economia de tempo no processamento, outros desenvolvem e vendem equipamentos com essas qualidades; alguns requerem “robustez” e “performance” dos sistemas operacionais, outros quebram a cabeça para inscrever nos objetos esses critérios de desempenho, etc. As qualidades das coisas, antes de serem propriedades inerentes ou essenciais, são um efeito, resultado de processos complexos de negociações, traduções, deslocamentos e transformações. Mobilizar e interessar outras pessoas e coisas, construir alianças, produzir provas cada vez mais “duras” são estratégias utilizadas para construir fatos e máquinas, para construir fatos a partir das ficções criadas nos laboratórios e centros de pesquisa.⁷ É por essa participação e confrontação em redes de interações socio-técnicas que se define relacionamente a forma de um artefato tecno-científico. A sua própria existência como máquina ou artefato resulta, na verdade, da estabilização desse sistema de interações.

⁶ Latour (1996, 2000) contrapõe, nesse sentido, o modelo de tradução ao modelo de difusão que comumente serve para dar conta de como máquinas e artefatos se propagam e adquirem durabilidade e estabilidade no tempo e no espaço. Uma invenção - e neste ponto Latour tem reconhecido como seu antecessor a Gabriel Tarde - não simplesmente se transmite por força própria: ela produz e depende de inúmeras e microscópicas associações, que modificam, fortalecem ou incorporam o objeto em novos contextos.

⁷ Para uma análise detalhada dessas estratégias de tradução cf. Latour (2000, p. 178). Sobre as “traduções-transformações” de regimes de signos, levando em conta formas de expressão não necessariamente verbais, cf. Deleuze e Guattari (1995-1997, v. 2., p. 91 et seq.).

Pela tradução dos interesses e necessidades dos “usuários” que fazem os desenvolvedores do Google, sabemos que o que aqueles querem é encontrar informação de uma maneira “fácil” e “rápida”, que para isso usam diretórios construídos por humanos e máquinas de busca automatizadas. O número de documentos cobertos por esses diretórios e buscadores teria aumentado de forma considerável, enquanto isso não ocorreria com a “capacidade dos usuários” de consultar esses documentos: as pessoas estão dispostas “a consultar apenas as primeiras dezenas de resultados (PAGE; BRIN, 1998a, p. 108). Por sua vez, os buscadores, nos dizem seus porta-vozes, se diferenciam dos diretórios mantidos por editores humanos, já que não organizam hierarquicamente as páginas que colecionam. Esses motores de busca, geralmente, preocupam-se menos com a seletividade que com a abrangência de suas bases de dados. Para isso os *softwares* que seus “porta-vozes” chamam de robôs (*webbots*), aranhas (*spiders*), agentes (*user-agents*), viajantes (*wanderers*) ou rastejadores (*crawlers*), vão seguindo *links*, de um lugar a outro na internet, passando pelos diferentes documentos espalhados na *web* para integrá-los à base de dados do sistema de busca. Para sua “locomoção”, esses programas utilizam diferentes estratégias, iniciando sua exploração geralmente a partir de *sites* conhecidos e seguindo os *links* encontrados nas páginas desses *sites*.⁸

O Google, em particular, explicam-nos seus porta-vozes, procura não só rastrear e indexar uma vasta quantidade de documentos, mas também retornar resultados “relevantes”. A primeira é tarefa do “Googlebot”, a segunda faz-se graças ao “PageRank”, o método patenteado por Page. Em geral, para indexar documentos e retornar resultados relevantes, os motores de busca requerem que os arquivos tenham rótulos e descritores, que possam ser “lidos” e analisados por eles de forma automática. É preciso, ainda, que os documentos assim escritos e estruturados se conectem uns aos outros, e que apresentem marcas claras e reconhecíveis desses vínculos (os “robôs” ignoram documentos da *web* que não contém *hyperlinks*). Isso, certamente, para que os usuários humanos possam, ao fazer suas “pesquisas”, localizar e se apropriar de maneira “rápida” e “fácil” de qualquer tipo de recursos.

Como vemos, a construção de um sistema que permita a concentração e acumulação de uma vasta quantidade de “recursos” espalhados, assim como o desenvolvimento de mecanismos para localizar e manipular de forma eficiente esses materiais, não envolve apenas mecanismos técnicos.⁹ Podemos dizer, também, que a construção desse artefato supõe uma complexa composição de forças (da percepção e a memória humanas, à probabilidade de ocorrência de palavras numa linguagem “natural”, o custo de digitalizar e indexar textos, o tráfego de *gigabytes* de dados, as verbas investidas nos projetos de desenvolvimento, as regulações dos direitos autorais, etc.). Assim, não se trata apenas de uma questão de mobilização de recursos, de alistamento de aliados, mas de sua atuação como um todo unificado. Como observa Latour,

⁸ Para uma exposição mais detalhada das peculiaridades das diferentes ferramentas de busca, orientada à capacitação dos usuários e visando ao “aumento de eficiência na procura de recursos informacionais”, cf. Cendon (2001). Para uma listagem dos “robôs” da *web*, *user agents* e dos protocolos para exclusão de robôs, pode se consultar Wikipedia (2005a) e Robotstxt.org (2005).

⁹ Cabe notar que o poder de concentração e acumulação de “recursos”, de mobilização numa escala inusitada de tudo o que pode ser inscrito, é uma das características da techno-ciência contemporânea e das sociedades que se autodenominam modernas (Cf. LATOUR, 2000, 1994). Aqui, estudamos as tecnologias da informação e, em particular, aquelas que envolvem bancos de dados, como dispositivos de inscrição que fazem possível tal aumento de escala e como meios de acelerar ainda mais a mobilidade, estabilidade e combinabilidade dos materiais tratados. Por outro lado, cabe pensar que se trata de instrumentos que cristalizam e acentuam o *modus operandi* da techno-ciência: eles incorporam na sua materialidade a rede na qual são construídos e utilizados (incorporam o seu meio associado).

o meio mais simples de transformar o conjunto justaposto de aliados num todo que atue com unicidade é atar as forças reunidas uma à outra e construir uma máquina. Máquina, como o nome indica, é antes de tudo, maquinação, estratagema, um tipo de esperteza em que as forças usadas mantêm-se mutuamente sob controle, de tal modo que nenhuma delas possa escapar do grupo (LATOURE, 2000, p. 213).¹⁰

Tecnologias de governo: código e materialidade

O que se entendia por “Google” em 1998? Tratava-se de um objeto (um objeto baseado em código) ou um projeto de pesquisa? É interessante notar que a primeira “interface com o usuário” de Google não faz uma disjunção entre as duas coisas. O protótipo e o projeto de pesquisa são dificilmente separáveis. Só seria possível separar essa trama de elementos, “técnicos” e “sociais”, humanos e não humanos, após um trabalho de “depuração” (Cf. LATOUR, 1994). Podemos pensar, portanto, que “Google” remetia a um híbrido código-projeto, resultante da reunião e agenciamento de materiais e forças diversas, concentradas naquele momento num centro de pesquisa na Universidade de Stanford. Fora desse centro, “Google” era uma maneira de pronunciar a palavra “googol”, usada por matemáticos para designar 10 elevado à centésima potência.

Mesmo no *paper* aqui analisado, “Google” dá lugar a sentidos diversos (máquina, motor de busca, sistema, arquitetura, projeto de pesquisa, etc.). Agora, além de levar em conta essas diferentes interpretações, cabe perguntar o que o objeto podia fazer e o que podia se fazer ao interagir com ele. Não se trata apenas de ver como as representações e intenções humanas podem introduzir uma espécie de “flexibilidade interpretativa” num mundo de objetos supostamente inertes e “inflexíveis”. Isto é, aquelas diferentes leituras que o *paper* permite, são leituras do quê? Entendemos que se trata de interpretações das possibilidades de ação que o objeto proporciona.

Para não reduzir os artefatos técnicos a propriedades inerentes ou “essenciais”, resulta útil pensar a sua organização material em termos de *affordances*, isto é, as possibilidades de ação que o objeto constringe ou facilita.¹¹ A materialidade dos objetos, assim como a dos ambientes, pode

¹⁰ Não nos deteremos aqui nesse ponto, mas cabe notar, de passagem, a relação disto com a noção de “agenciamento maquínico” (DELEUZE; GUATTARI, 1995-1997) e com a filosofia das técnicas de Gilbert Simondon (1969). Longe de apreender um objeto como algo acabado (como um indivíduo), a ênfase no processo de “tomada de forma” (na “individuação” dos objetos técnicos) supõe reconhecer que há um processo no qual múltiplas forças devem ser continuamente ajustadas, compostas, agenciadas. No nosso exemplo, essas forças vão do tráfego de *gigabytes* de dados às variações das linguagens “naturais” e de codificação, passando pelos fluxos monetários (o investimento das empresas do setor de tecnologia, etc.). Entendemos que, num sentido similar, Latour propõe estudar as associações em vez daquilo que comumente chamamos de “social”: “As forças são heterogêneas no seu caráter: podem incluir átomos, palavras, lianas ou *tattoos*. Elas próprias estão amarradas para criar as máquinas e maquinações que nos mantêm a todos no nosso lugar” (LATOURE, 1996, p. 277). A respeito dessa “maquinação de forças”, na edição francesa de *Ciência em ação*, Latour (1989) remetia à antropologia da tecnologia de André Leroi-Gourham e à obra de Gilbert Simondon (embora este último fosse mal citado na bibliografia). Pretendemos analisar, num outro trabalho, a relação material-forças e a noção de “agenciamento maquínico”, no sentido de Deleuze e Guattari, noção que retoma tanto a contribuição de Leroi-Gourham quanto a de Simondon para pensar a tecnologia. Para uma análise mais aprofundada do alcance e importância dos conceitos de máquina social e de *phylum* maquínico no pensamento de Deleuze e Guattari, cf. Ferreira (2004).

¹¹ Para uma excelente análise do conceito de *affordances*, inscrita na proposta de “leitura micropolítica dos artefatos” que inspirou em grande medida o presente trabalho, cf. Kasper (2004).

“ativar” ou facilitar certas ações, não admitir ou, de alguma forma, constranger as possibilidades para a realização de outras. A noção de *affordances* refere-se a esses aspectos funcionais, assim como a aspectos relacionais, que condicionam, “emolduram”, as possibilidades de ação de quem interage com um objeto, em torno ou através dele (Cf. KASPER, 2004; HUTCHBY, 2001).

Deste ponto de vista, estudamos objetos que são tanto discursivos e semióticos quanto materiais. E isto vale tanto para os artefatos físicos, quanto para as “aplicações”, “soluções”, “ambientes” e “plataformas” de *software*. A primazia do cálculo e da linguagem, e o fato de esses objetos não serem tangíveis têm levado a uma ênfase exagerada na “imaterialidade” e na “espiritualidade” associadas a tecnologias da informação e às redes digitais. Na nossa perspectiva, os objetos baseados em código, como outras produções baseadas na escrita, envolvem tanto dimensões relativas à expressão quanto ao conteúdo, isto é, tanto à organização das funções quanto das matérias.¹²

Uma interface de recuperação de dados, um formulário *on-line* ou um portal, por exemplo, podem induzir ou bloquear ações, afetando as possibilidades de visualização e a seqüência da leitura. Os elementos no seu desenho (visualizados como menus, *banners*, caixas, *combo boxes*, botões, etc.) executam comandos previamente definidos (direcionar, registrar, deixar ver, ocultar, formatar, repetir, etc.) que condicionam a relação do usuário com o objeto, facilitando ou limitando a visibilidade, acelerando ou retardando certos eventos. As *affordances* dos objetos digitais podem ser, sem dúvida, muito complexas: elas não derivam de propriedades “naturais” ou “essenciais” do objeto, mas estão “embutidas” no desenho do artefato (Cf. AKRICH, 1992).

Cabe notar que, desse ponto de vista, pode ser considerada a “estrutura lógica” de um banco de dados. Essa “estrutura” ou “modelo” de dados não apenas envolve dimensões lógicas e epistemológicas, mas também ontológicas. A classificação de entidades, a divisão de campos (os “atributos” dessas entidades), a definição dos tipos de relacionamentos possíveis, etc., não tanto representam o mundo quanto têm um caráter produtivo. Uma “estrutura de dados” preestabelece certas formas de inscrição e de combinação, de recombinação e divisibilidade dos registros. É, portanto, no desenho do banco de dados e das interfaces para sua manipulação que são inscritas as ações possíveis daqueles que cadastram ou consultam registros. Assim, o desenho de um banco de dados condiciona - se bem não determina - as possibilidades de memorização e registro, assim como as próprias formas de descrição e enunciação. Sobretudo, esse “modelo” ou desenho dos dados pauta maneiras específicas de encadear os acontecimentos.

Matemática e política: o “algoritmo democrático”

Deixando de lado as demarcações entre o social e o técnico, e levando em consideração tanto dimensões semióticas quanto materiais, podemos pensar que o que estava tomando forma, o que estava sendo construído sob o nome de Google, em 1998, era um híbrido: uma mistura de algoritmos, linguagens formalizadas e projeto de pesquisa, uma trama de mecanismos de cálculo, codificação e gestão. Ao analisarmos este tipo de híbridos produzidos pela tecno-ciência, resulta irrelevante atribuir a sua invenção a um punhado de indivíduos, restringi-la a um único centro de desenvolvimento ou mesmo a um setor ou esfera de atividade. A demarcação tradicional das áreas de conhecimento resulta também de pouca utilidade.

¹² A escrita pode ser entendida como uma formalização da expressão própria ao conteúdo. Nos apoiamos, aqui, na distinção entre formas de conteúdo e formas de expressão proposta por Deleuze e Guattari (1995-1997).

Graduados em matemática e engenharia de computação, Lawrence Page e Sergey Brin mencionam entre suas linhas de pesquisa o “tratamento de linguagens naturais”, a “extração de informação de fontes não-estruturadas”, a “mineração de dados”, a “exploração de padrões de grandes coleções de textos e de dados científicos”, o “acesso à informação” (PAGE; BRIN, 1998a, p. 117; PAGE; BRIN, 1998b¹³). Os cruzamentos entre diferentes domínios de saber se multiplicam (transdisciplinariedade que, como sabido, é característica das teorias e técnicas em torno da informação).

Isso é claro no caso do “PageRank”, o algoritmo que ainda hoje seria o “coração” do Google. A patente de 1998, como vimos, atribui a Lawrence Page a invenção de

um método para atribuir medidas de importância aos diferentes nós de um banco de dados relacional, como um banco de dados de citações, a rede mundial de computadores ou qualquer outro banco de dados em hipermídia (BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY, 1998).

Para fundamentar esse método, antigos e novos problemas são retomados, assim como diferentes conceitos e técnicas são refuncionalizados. Como determinar a “relevância”, como decidir sobre a qualidade de “importante” ou “interessante” de alguma coisa? E, ainda, o que significa delegarmos a outro (neste caso um *software*) essa decisão?

Se as listagens mantidas por humanos, explicam Page e Brin, podem ser “subjetivas, caras para se construir e se manter”, “de lento desenvolvimento”, etc., as máquinas de busca podem, no entanto, recuperar documentos de “baixa qualidade” e serem “parciais”. Os pesquisadores nos explicam em que consistiria essa parcialidade no caso das máquinas. Num apêndice do *paper*, afirmam que

o modelo de negócios predominante para máquinas de busca comerciais é a publicidade. Os objetivos desse modelo de negócios nem sempre coincidem com os de fornecer aos usuários buscas de qualidade [...]. Do ponto de vista do consumidor, quanto melhor o buscador for, menos anúncios o consumidor encontrará antes de chegar ao que procura.

Com vistas a isso, os autores argumentavam que era preciso construir buscadores que não se afastem das “necessidades dos consumidores”. “A questão da publicidade”, sublinhavam, “traz bastantes incentivos misturados (*mixed motives*) e é crucial haver uma *search engine* competitiva (sic) que seja transparente e se mantenha no domínio acadêmico” (PAGE; BRIN, 1998b).¹⁴

¹³ Item “Research interests”.

¹⁴ Appendix A: “Advertising and mixed motives”. Cabe notar que, no mesmo *paper*, sobrepõem-se várias representações daqueles que fazem “pesquisa” (*search*) com ou no Google. Se aqui lemos que os usuários são “consumidores”, nas primeiras páginas os usuários eram considerados como “pesquisadores” no sentido estrito.

No entanto, mesmo se fosse “restrita a esse domínio”, a questão da qualidade e do ordenamento por relevância continuam sendo problemáticas. Adotar critérios de relevância supõe, obviamente, que numa determinada multiplicidade (digamos, uma “população” de documentos) nem todos os indivíduos serão considerados iguais. Com base em quais critérios selecionar uns e deixar outros de lado? Na medida em que uma amostra ou um banco de dados (digamos, a *web*) aumenta, essa questão torna-se a grande questão. A preocupação já não é a abrangência da amostra, mas sim a questão da seleção.

Grandes bancos de dados de documentos como a web contém muitos documentos de baixa qualidade. Como resultado disso, as buscas recuperam tipicamente centenas de documentos irrelevantes ou não procurados que escondem os poucos que são relevantes (BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY, 1998).

Na medida que o tamanho de um banco de dados aumenta, nos explicam Page e Brin,

necessitamos de ferramentas que tenham uma precisão muito alta (número de documentos relevantes recuperados, por exemplo, entre as primeiras dezenas de resultados). De fato, queremos que de acordo com nossa noção de relevância sejam incluídos apenas os melhores documentos, já que podem existir dezenas de milhares de documentos de pouca importância.

Para tal efeito, o Google, além das palavras-chave, sua proximidade e localização no conteúdo da página, utiliza a estrutura de *links* e o texto sobre os *links* (texto âncora), que “provêm muitas informações tanto para avaliação de relevância como para filtragem da qualidade” (PAGE; BRIN, 1998a, p. 108).¹⁵ Nesses procedimentos de avaliação e seleção, o pressuposto é que uma página que é referenciada por muitas outras tem maior importância que uma referenciada uma única vez.

Esse tipo de cálculo baseado na conectividade não seria, no entanto, inteiramente “novo”. Page e Brin explicam que “a literatura sobre citações acadêmicas tem sido aplicada (sic) à web, principalmente por meio da contagem de citações ou *backlinks* a uma determinada página”. Isso provê uma estimativa da “importância e qualidade da página” (PAGE; BRIN, 1998a, p. 109).

Aplicando aqui a mesma técnica, nós poderíamos fazer o exercício de examinar a literatura que, na patente, serve como “estado da arte” da inovação. Ao fazê-lo, podemos constatar que entre as publicações ali citadas, encontramos algumas que remetem à pesquisa quantitativa nas ciências humanas e a técnicas associadas à gestão da ciência e da tecnologia. A referência com data de publicação menos recente é um artigo de 1953, “A new status index derived from sociometric analysis”, publicado na revista *Psychometrika*. Há também referências a vários artigos sobre

¹⁵ No texto da patente lemos que “a idéia de associar o texto âncora com a página a qual ele aponta, foi inicialmente implementada pelo World Wide Web Worm”. Page faz referência a um artigo publicado, em 1994, na Primeira Conferência Internacional da World Wide Web, no CERN (o Laboratório Europeu de Física de Partículas).

avaliação de *journals* científicos e sobre técnicas de análise de citações, assim como um trabalho publicado na revista *Social Methodology* sobre técnicas de análise quantitativa de redes sociais (Cf. BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY, 1998).

Se esses métodos de contagem de citações e referências são “bem conhecidos”, eles se aplicariam no entanto apenas “a bancos de dados, cujo conteúdo é de uma qualidade e importância relativamente uniformes”. Muitos bancos de dados, explica-se no texto da patente, têm variações extremas na qualidade e importância dos documentos. Nesses casos, as técnicas tradicionais seriam insuficientes e dariam, por exemplo, “o mesmo peso a um documento que é citado uma vez por uma página obscura e a um documento similar que é citado uma vez por uma página bem conhecida e altamente respeitada.” Considerando que as técnicas já conhecidas seriam “excessivamente simplistas”, a patente alega, a invenção de “um ordenamento objetivo baseado na relação entre documentos”, método que é “escalável e que pode ser aplicado a bancos de dados extremamente grandes como a rede mundial de computadores.” Outro aspecto dessa invenção visa

tirar proveito da estrutura relacional de um banco de dados para asignar um peso a cada documento do banco de dados, sendo o peso do documento a medida da sua importância. Antes que determinar a relevância pelo conteúdo intrínseco do documento ou pelo texto âncora dos *backlinks* que apontam a ele, o método inventado determina a importância a partir de relações extrínsecas entre documentos (BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY, 1998).

As métricas de ordenamento de coletivos (documentos, computadores, cientistas, usuários), baseadas na conectividade, aparecem geralmente como sendo razoáveis e costumam proporcionar bons resultados práticos. Cabe, no entanto, examinar mais de perto alguns de seus pressupostos. No caso do algoritmo do motor de busca, um dos pressupostos é que um *link* equivaleria a um “voto”. Se uma página tem um *link* com uma outra seria como se a primeira concedesse um voto à segunda. O fator que aumenta o *PageRank* de uma página é o número de *links* que apontam para ela. Mas nem todas as vozes têm o mesmo peso nessa “eleição”. A importância conferida a um documento será calculada a partir das posições ou “pesos” dos documentos que o citam:

uma página pode ter alto *PageRank* se houver muitas páginas apontando para ela, ou se houver algumas páginas de alto *PageRank* apontando para ela. Intuitivamente, [é fácil aceitar que] páginas que são frequentemente citadas em muitos outros pontos da *web* são merecedoras de uma visita (PAGE; BRIN, 1998a, p. 110).

Essa definição recursiva da relevância poderia ser comparada, na verdade, a uma “capacidade de voto”: quanto maior o PR de uma página, mais PR ela pode passar para outras. Para isso, o método requer um procedimento de cálculo iterativo, que envolve uma normalização prévia dos *links* em cada página, a atribuição de pesos a cada *link* e uma propagação recursiva de pesos

atribuídos às páginas através da estrutura de *links* da *web*. Na patente, lemos que “dado que as citações ou *links* são modos de direcionar a atenção, os documentos importantes correspondem àqueles documentos para os quais é direcionada mais atenção” (BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY, 1998).

Entre a engenharia e as ciências humanas aplicadas, entre os problemas da representação estatística e a representação política, qual o tipo de relação entre humanos e máquinas que está sendo construído? Na seção sobre a escalabilidade (possibilidades de crescimento) do sistema, no último parágrafo do *paper*, podemos ler alguns pressupostos a respeito dessa relação:

dado que os humanos somente podem digitar ou falar uma *quantidade finita* [de informações] e visto que os computadores continuam melhorando, a indexação de textos *vai se ampliar ainda mais [vai escalar ainda melhor]* do que hoje. Certamente, pode haver um volume infinito de conteúdo gerado por máquinas, mas indexar enormes volumes de conteúdo gerado por humanos parece extraordinariamente útil. Por isso, somos otimistas a respeito de que nossa arquitetura centralizada de motor de busca na *web* melhore na sua capacidade de cobrir as *informações de texto pertinentes* e que haja um futuro brilhante para a busca (PAGE; BRIN, 1998b, grifo nosso).¹⁶

Enfrento aqui certos problemas de tradução... O otimismo dos autores, com respeito ao futuro dos motores de busca e à evolução de máquinas e humanos, associa-se aparentemente à quantidade, ao aumento de escala. De que forma as possibilidades de crescimento, isto é, a “escalabilidade”, podem ser “melhores”? Indexando e registrando volumes cada vez maiores? Mais é sempre melhor? Por outro lado, quem decide o que é melhor e quanto é preciso? As informações “pertinentes” são, em qualquer contexto, as mais populares? Todos os assuntos teriam cânones e medidas-padrão igualmente definidos? E, mesmo que os tivessem, cabe sempre subordinar as variações às constantes? Cabe fazer valer em todo e qualquer âmbito apenas princípios majoritários? Em todo caso, cabe notar a enorme responsabilidade conferida aos tradutores.

A expansão da rede: Links e ações

Passemos agora a um formulário, um pedido de registro relativo a Google apresentado, em 2004, perante a Comissão de Valores e Bolsas dos Estados Unidos (Cf. UNITED STATES OF AMERICA, 2004). Diferentemente do artigo científico analisado no início deste trabalho, que alegava trazer a primeira “descrição pública” de um motor de busca, e diferentemente do texto da patente, que conferia direitos de propriedade intelectual ao inventor, este documento dirige-se aos investidores. Ele traz uma extensa descrição de Google, entendido não já como projeto de pesquisa ou ferramenta, mas como empresa. A descrição que se faz “pública” em 2004 (a empresa lista-se agora como uma companhia aberta) não diz respeito à “anatomia do sistema” ou ao algoritmo, mas à “missão da empresa”, ao “modelo de negócio” e ao “plano de ação”.

¹⁶ Apêndice B, item 9.2: “Scalability of centralized indexing architectures”.

Esse documento relativo à oferta pública de ações da Google traz informações sobre os fatores de risco relacionados aos produtos, serviços e ao mercado em que a Google atua. Explica como irá utilizar-se o capital levantado e identificam-se os principais acionistas. Vemos, ali, como a rede de “aliados” e tradutores tem se expandido. Constatamos, também, que na rede traçada acima havia associações mais “fracas” e associações mais “fortes”: o documento é assinado pelos fundadores e também pelo CEO (*chief executive officer*) da Google, um ex-executivo da Sun Microsystems. Na descrição da estrutura organizacional são mencionados outros três diretores da empresa, um deles ex-executivo da Intel e outro executivo da Genentech. Também estão envolvidos corretores de ações, agentes e os auditores (Morgan Stanley & Co., Credit Suisse First Boston LLC) que assinam o balanço financeiro da empresa. Caberia incluir na “rede”, as empresas adquiridas pela Google como, por exemplo, a Applied Semantics Inc. (Cf. UNITED STATES OF AMERICA, 2004).

O documento submetido à SEC inclui, ainda, uma carta dos fundadores da empresa aos futuros investidores (*Letter from the founders: ‘An owner’s manual’ for Google’s shareholders*). Page e Brin manifestam, ali, seu compromisso com os usuários e sua repulsa a práticas abusivas (como, por exemplo, vender colocações nos *rankings* ou manipular os algoritmos dos motores de busca). Essa carta se dirige aos futuros proprietários de ações de uma empresa, cuja ênfase num “ambiente de criatividade e desafio” teria “ajudado a fornecer um acesso à informação não tendencioso, preciso e livre para aqueles que confiam em nós ao redor do mundo.” A carta continua nesse tom: “Sergey e eu fundamos a Google porque acreditávamos que podíamos fornecer um grande serviço ao mundo – distribuir instantaneamente informação sobre qualquer assunto.” Lendo a carta-manual do proprietário, é evidente que a propriedade em questão é sobre as ações. Tudo parece como se se tratasse de uma Enciclopédia-Mundo, que, no entanto, deixaria de lado uma parte da ambição universalista dos enciclopedistas. O conhecimento sobre o algoritmo e toda uma série de mecanismos e desenvolvimentos técnicos associados são, agora, zelosamente guardados por uma corporação.

Se a máquina de busca em 1998 nos remitia a uma pesquisa em andamento e a um punhado de colaboradores (“*Google is a research in progress and there are only a few of us so expect some downtimes and malfunctions*”), Google Inc. é, em 2004, uma companhia com quase 2000 funcionários. A rede na qual a ferramenta estava inscrita, em 1998, restringia-se aos Estados Unidos (“*This new index contains only a very limited number of international pages because we do not want to congest busy international links*”) e mesmo ao âmbito local (no rodapé da primeira interface, encontrávamos um único *link*, “Weather for Stanford, California”). Em 2004, além da sede na Califórnia, a Google tem escritórios de suporte técnico, vendas e centros de P&D em diferentes locais dos EUA e em outras cidades.¹⁷ Possui 95 nomes de domínios de internet registrados e interfaces com o usuário disponíveis em 97 idiomas. E todas as palavras deste mundo parecem não lhe bastar: o buscador apresenta interfaces em duas línguas “extraterrestres” (Klingon e Leet), tomadas de uma série de ficção científica.

A rede dilatada envolve, obviamente, os “usuários” do motor de busca. É importante salientar que o papel daqueles é inseparável da construção de Google como máquina: cada vez que

¹⁷ Amsterdã, Atlanta, Bangalore, Boston, Chicago, Dallas, Denver, Detroit, Dublin, Hamburgo, Hyderabad, Londres, Los Angeles, Madri, Milão, Mountain View, Nova York, Paris, Santa Monica, Seattle, Sydney, Tokyo, Toronto e Zurique (UNITED STATES OF AMERICA, 2004). A organização espacial do sistema de busca é sugestiva com respeito à importância econômica das “cidades globais” em detrimento do papel dos Estados-Nação. Ao considerar a arquitetura centralizada do sistema de busca constata-se, também, a hipótese de que a crescente dispersão territorial dos intercâmbios estaria criando novas necessidades de administração e controle centralizado (SASSEN, 1991). O que resulta interessante, justamente, é que aquele movimento de desterritorialização é relativo na medida que aparece conjugado a uma reterritorialização: o Google opera de forma “distribuída” para aumentar o desempenho no “rastreamento” das páginas *web* e no processamento das consultas, mas tanto do ponto de vista da organização da indexação quanto da estrutura de dados e da propriedade, trata-se de um sistema centralizado.

utilizamos, sem questioná-lo, esse ou qualquer outro artefato, o fortalecemos, colaboramos na sua construção. Se no *paper* de 1998 Page e Brin explicavam ter projetado o sistema de busca especialmente para que outros pesquisadores e estudantes pudessem “entrar rapidamente, processar grandes porções da *web* e produzir resultados interessantes” que “difícilmente teriam sido obtidos de outra maneira”, em 2004 é claro que a busca e a experimentação não se restringem às atividades dos pesquisadores em sentido estrito. No formulário submetido à Comissão de Valores e Bolsas superpõem-se diversas representações dos “usuários” ou “consumidores” que fazem buscas e “pesquisas” no ou por meio do Google. Por outro lado, se procurássemos identificar os “usuários” do buscador, antes que classificar pessoas segundo nacionalidade, por exemplo, caberia pensar na constituição do “dentro” e do “fora” como processos sempre relativos: como enunciado na revista *Wired*, mais da metade do *tráfego* de usuários do Google tinha se originado “fora” (*abroad*), isto é, fora dos Estados Unidos (GLASNER, 2004).

No *paper* de 1998, Page e Brin justificavam o uso da palavra “google” (10 elevado à centésima potência) porque combinava com o objetivo do projeto de construir uma máquina de busca de larga escala. Em 2004, essa palavra “reflete a missão da companhia de organizar o imenso, aparentemente ilimitado volume de informação disponível na *web*”. Trata-se de uma marca da empresa, registrada nos EUA e em outros países. Outras marcas e serviços da Google são PageRank, Froogle, AdSense, AdWords, Blogger e Gmail.¹⁸ Também são marcas registradas “*I’m Feeling Lucky*” (o texto sobre um dos botões de comando na interface do motor de busca) e “*It’s all about results*” (UNITED STATES OF AMERICA, 2004). A ampla difusão de “Google” como marca tem trazido alguns efeitos não previstos pela empresa: muitos internautas, sobretudo falantes de inglês, usariam “*to google*” como um verbo, sinônimo de “procurar alguma coisa no Google” ou, em geral, de “fazer uma busca na *web*”. Os representantes do Google, lemos no documento submetido à SEC, preocupam-se com que o nome se torne uma marca genérica e com o conseqüente risco de que a forma verbal possa diluir a imagem da empresa.

O índice da *web* do Google tinha, segundo esse documento, aproximadamente dois terabytes de tamanho, abrangendo 3 bilhões de páginas da *web*. Em fevereiro de 2005, segundo a Wikipedia e a própria empresa, o Google teria indexado mais de 8 bilhões de páginas, mais de um bilhão de imagens e oitocentos milhões de mensagens *usetnet*. O sistema de busca armazena todos os dados que indexa e vasculha nessa gigante coleção em menos de meio minuto (GOOGLE INC., 2005). Para fornecer resultados instantâneos e manter um banco de dados dessas dimensões, emprega uma vasta *intranet* de mais de 3.500 servidores, que roda em Linux. O índice de pesquisa é distribuído entre milhares de discos rígidos, permitindo que um *software* de balanceamento de carga aponte as consultas para os servidores e subsistemas de disco mais disponíveis.

Se na nossa “rede sociotécnica” de 1998, desenhada a partir da descrição do *paper*, contávamos entre os “actantes” os cinco servidores e vários *crawlers* trabalhando de forma distribuída, no documento submetido à SEC de 2004 a empresa alega deter uma das redes privadas mais poderosas do mundo. Um analista da Intel, empresa fornecedora de servidores da Google, estima que a sua cliente precisaria “adicionar aproximadamente 30 novos servidores por dia à rede (sic), somente para acompanhar o aumento da demanda” (!?). O técnico da Intel se pergunta:

¹⁸ Outros serviços prestados pela empresa são Google Maps, Google Scholar e o Diretório Google que, como outros buscadores e portais, utiliza as categorias e o conteúdo do Open Directory Project. A Google é também proprietária do código do software do Orkut, um conhecido serviço destinado às chamadas “redes sociais” *on line* ou “comunidades virtuais” (Cf. TERDIMAN, 2004). Orkut é marca registrada da empresa Orkut.com LLC.

com mais de 3.500 servidores [...] zumbindo nas instalações de *co-location* da Google, a questão é: quantos mais a empresa instalará? Nesse ponto, o céu parece ser o limite. A Google diz que tem crescido a uma taxa de 25 % por mês e, à medida que mais e mais *web sites* empregam o serviço gratuito e pago de pesquisa na *web* da Google, essa taxa pode se acelerar (UNITED STATES OF AMERICA, 2004).

Como vimos, o algoritmo do motor de busca (o método para selecionar documentos “relevantes”) tratava as variações subordinando-as às constantes. O desenho da infra-estrutura tecnológica do sistema apoia-se também numa padronização em grandíssima escala. Aqui o que está em jogo não é só a extensão de uma rede tecno-científica numa escala até agora inconcebível. Pode se pensar que, nesta rede dilatada, toda uma série de mecanismos de cálculo e seleção são inscritos nos objetos e nas práticas cotidianas, o que deriva de e, ao mesmo tempo, contribui para sustentar um processo de “aceleração total” dos procedimentos de acumulação (Cf. SANTOS, 2004).¹⁹

Links patrocinados. A otimização da visibilidade

Palavras como “missão”, “sorte”, “confiança”, “satisfação”, “maldade”, “valor” não pareceriam muito adequadas para a descrição de um objeto técnico. No entanto, como lemos no documento submetido à Comissão de Valores e Bolsas, elas são inseparáveis da realidade de Google em 2004. A maioria dos serviços fornecidos pela empresa é gratuita para o usuário final (como é o caso da maioria dos motores de busca). A principal fonte de renda provém, entretanto, dos diferentes serviços que a empresa presta. Esses serviços consistem em licenciar a sua tecnologia a grandes clientes como portais ou grandes jornais eletrônicos e em exibir publicidade paga em páginas de resultado de pesquisas.²⁰

Os serviços para anunciantes da Google não apresentam, contudo, publicidade “intrusiva” ou que incomode o usuário, mas sim anúncios discretos e bem diferenciados dos resultados das pesquisas. Preocupados com o impacto positivo dos produtos da empresa, na carta aos investidores, os fundadores advertem que a sua prioridade será manter a aliança com os usuários. Larry Page e Sergey Brin asseguram que o código de conduta da empresa, cujo princípio é “*don't be evil*”, será respeitado, mantendo a “confiança” dos usuários e não aceitando pagamentos por resultados de busca. Essa precaução é explicada, mais adiante, como visando a “fornecer aos usuários informação comercial de alto valor” e pela necessidade de conciliar “a relevância”, “a satisfação dos usuários”, a “satisfação dos anunciantes” e “as necessidades dos parceiros” (UNITED STATES OF AMERICA, 2004).

De novo, cabe pensar que, nessas complexas alianças, há associações mais fortes e mais fracas, algumas que resistem melhor ou pior que outras. E, certamente, a enorme audiência

¹⁹ Especialmente: “Müller e o ritmo dos tempos”; e “Tecnologia e seleção”.

²⁰ Pode se constatar a importância crescente dada a esses serviços nas mudanças no próprio desenho das interfaces de recuperação de dados do buscador e na descrição que se faz dele. Se em 2001, na seção “About” não se ofereciam serviços de marketing, em 2002 eles passam a ser um dos principais tópicos na seção “Tudo sobre Google”. Além disso, a partir de agosto de 2004, nessa descrição do buscador encontramos em destaque os anúncios dos próprios serviços de marketing da Google. Cf. o site da Google, via Internet Archive.

proporcionada pelas buscas chama a atenção dos anunciantes. Como são então traduzidas/traídas as “necessidades” e “interesses” dos usuários e dos anunciantes? Como conciliar aquilo que seria a “satisfação dos usuários” e a daqueles que têm mais capacidade de pagar para ter uma visibilidade “otimizada”?

“Personalização” vs. “privacidade”

A possibilidade de dispor de um repositório de documentos cobrindo “quase a totalidade da *web*”, assim como de um registro de dados sobre as ações de milhões de usuários do motor de busca e outros *sites*, tem levantado sérias questões concernentes à privacidade. Algumas organizações têm chamado a atenção para o fato de que essa centralização poderia criar um antecedente problemático para usos de dados pessoais pelos governos ou por outras corporações.²¹

O Google, como outros motores de busca e serviços na *web*, utiliza *cookies* com a função de facilitar a “personalização da ferramenta”.²² Os *cookies* e outros arquivos de *log* de acesso são utilizados, em muitos *sites* (*sites* de comércio eletrônico, portais, bibliotecas digitais, etc.), para armazenar informações pessoais e, em alguns casos, compor perfis dos visitantes e propagandear ou configurar de forma “personalizada” os produtos e serviços oferecidos. Muitos analistas argumentam que o uso desses arquivos, sempre que não armazenem informações sigilosas, não representa um perigo à privacidade dos internautas. As técnicas de personalização, pelo contrário, tornariam possíveis serviços de informação “customizados” e serviços de publicidade direcionada.

Para outros, entretanto, a relação entre “personalização” e “privacidade” é problemática. Assim, nas controvérsias suscitadas em torno do Google tem se sublinhado o fato de que os *cookies* que o sistema utiliza para registrar o histórico de buscas dos usuários, são armazenados durante muito tempo, tendo “prazos de validade” não comuns a outros serviços da *web*. A empresa, por sua vez, se compromete a manter sua “política de privacidade” e não vender ou transmitir, sem o consentimento do usuário, dados pessoais coletados no seus sites a outros indivíduos ou empresas. Contudo, muitos apontam a possibilidade e os riscos de a empresa utilizar um único *cookie* ou *cookies* que permitam fazer um “cruzamento” de informações nos diversos serviços oferecidos pela empresa – históricos das buscas, registro de ações nas aplicações de e-mail, nos grupos on-line, etc. (Cf WIKIPEDIA, 2005b). Para ter acesso a muitos desses serviços, como é conhecido, o próprio usuário cadastra informações pessoais que, em alguns casos, podem servir para identificá-lo. Cabe pensar que, nesse sentido, os “usuários” são tanto a fonte como os que inscrevem a informação (ao cadastrar dados pessoais e gerar registros de suas ações no navegador). A própria atividade de “uso” produz inscrições e alimenta bancos de dados com informações úteis para a modificação futura dos produtos e serviços informacionais.

²¹ São as posições defendidas, por exemplo, por Google Watch (<http://www.google-watch.org>) e Google Fight (<http://www.googlefight.com>).

²² São chamados de “HTTP cookies” os arquivos ou pacotes de informação transmitidos por um servidor a um *browser* e depois re-transmitidos pelo *browser* cada vez que acessa esse servidor. Esses arquivos comumente não permitem identificar os usuários pelo nome: eles registram ações executadas no *browser* (é o *software* que, nos termos de muitos especialistas em redes digitais, é o “user-agent” e o “cliente” do servidor).

Com relação ao Gmail, o serviço de correio eletrônico gratuito da Google, as controvérsias envolvem sobretudo preocupações em torno do uso de técnicas de leitura automática do conteúdo, conhecidas também como técnicas de “mineração de textos”, utilizadas para incluir anúncios direcionados nas mensagens.²³ A principal questão colocada é que a correspondência de pessoas que não cadastraram contas no Gmail (isto é, que não acordaram necessariamente com os “termos do serviço”), mas que enviam e-mails a usuários de Gmail pode ser lida e analisada sem a sua permissão (PRIVACY RIGHTS CLEARINGHOUSE et al., 2004). Os representantes da empresa alegam que esses métodos para análise das mensagens já estariam sendo usados por alguns programas *anti-spam* instalados nos provedores de internet ou em outros serviços de correio eletrônico. Explicam, também, que os textos não são lidos por humanos, exceto pelo dono da conta de correio, e que essas técnicas seriam usadas para melhorar a relevância dos anúncios. Isto é, nos simpáticos termos do prospecto, tratar-se-ia de prover ao usuário de e-mail “informações comerciais relevantes e de alto valor”.

Nas situações e contextos em que são vistas como produções controvertíveis, e menos como “matérias de fato”, as tecnologias de informação aparecem fortemente conectadas com outros debates (como o da redefinição do público e do privado). Esse é o caso das técnicas de registro, extração e análise de dados, em particular, quando se trata de dados que permitem identificação de pessoas: registros de identidade, números de cartões de crédito, contas bancárias, licenças de conduzir, cartões de cliente, históricos médicos, registros de pagamento de serviços públicos, etc. Quando questionadas, essas técnicas têm levantado problemas associados, principalmente, à erosão da privacidade, abusos por parte das empresas e ao aumento da vigilância e do controle estatal sobre os cidadãos.²⁴

²³ Em geral, o “*data mining*”, nas diferentes definições que circulam entre os profissionais da informação, refere-se ao processo de extração e análise de dados, realizado com vistas a descobrir padrões ou relações previamente não conhecidos e potencialmente úteis a determinados fins. Essa “descoberta” de relações significativas ou tendências “implícitas” nos dados faz-se a partir do tratamento e análise, geralmente (mas não necessariamente), de grandes volumes de dados em formato eletrônico. Para isso, utilizam-se técnicas de inferência da estatística tradicional e técnicas desenvolvidas no âmbito do aprendizado máquina e da inteligência artificial (redes neurais, árvores de decisão, etc.), que envolvem métodos semi-automáticos de modelização dos dados. A denominação desse processo como *data mining* difundiu-se particularmente no âmbito empresarial, junto com expressões como “*data surfing*”, “descoberta de conhecimento nos dados” e outras usadas sobretudo pelos vendedores de software como “tecnologias analíticas”, “inteligência competitiva” ou “inteligência de negócio”. Alguns especialistas notam, no entanto, que o *data mining* não se refere a uma técnica particular ou a um algoritmo específico, mas é antes um método ou processo que, tendo como ponto de partida a exploração de um banco de dados, permite realizar predições a partir da generalização de um padrão para um certo tipo ou grupo de dados. As aplicações mais comuns desses processos de extração e análise de dados incluem a segmentação de mercado, a criação de “perfis” de clientes, a detecção de fraudes, a análise e monitoramento de risco de crédito, de históricos médicos, etc.

²⁴ Nos EUA, os bancos de dados, em geral, e as técnicas de *data mining*, em particular, tornaram-se um assunto controvertido quando, em 2003, o Congresso desse país cortara o financiamento para um projeto conhecido como “Total Information Awareness” (TIA) da Defense Advanced Research Projects Agency (DARPA). Essa pesquisa envolvia o uso de técnicas de reconhecimento de padrões em grandes bancos de dados, com o intuito de descobrir seqüências de atividades que poderiam revelar “planos de conspirações terroristas”. Como em outras iniciativas de construção de sistemas de identificação nessa escala, para esse projeto foi precisa a integração e o tratamento de enormes quantidades de dados, fornecidos por diferentes fontes de governo e empresas comerciais. Organizações de defesa das liberdades civis, sobretudo dos EUA, manifestaram preocupações com respeito ao uso de informações pessoais e apontaram o fato de que seria infrutífero e ineficiente, na prevenção do terrorismo, o uso de métodos para descobrir nesses enormes repositórios padrões de ações e relações de pessoas “suspeitas” (“é como procurar uma agulha no palheiro”). Por sua vez, o escritório geral de contabilidade do governo americano apresentou um relatório frente ao congresso, mostrando quão estendido estava o uso de técnicas de *data mining* no âmbito das agências do governo federal. Esse relatório, entretanto, chamara a atenção de que nem todos os projetos envolvendo técnicas de *data mining* seriam invasivos ou violariam as liberdades civis. Muitos desses projetos, advertiam os representantes do governo, estavam destinados a prevenir fraudes, abusos ou dispêndio inecessário de recursos, a gerenciar

Muitas dessas controvérsias, centradas na questão da privacidade, têm dado lugar a propostas visando reforçar as medidas de segurança, tanto do ponto de vista legal quanto técnico. Nesse último sentido, tem se fomentado o desenvolvimento de sistemas de acesso a bancos de dados com controles mais estritos: “*We can use technology to protect privacy (...) Let's build privacy into the design of these systems from the start*”, propunha um defensor das liberdades civis.²⁵ Como em outras situações nas quais se colocam em questão os efeitos de desenvolvimentos tecno-científicos, as soluções propostas procuram paliar os efeitos da tecnologia desenvolvendo mais tecnologia. Neste caso, procurou-se desenvolver medidas de segurança que incluem, por exemplo, mecanismos de “anonimização”, envolvendo técnicas de encriptação de dados como as usadas para proteger dados confidenciais no correio eletrônico, de forma que os projetos de *data mining* comerciais e das autoridades de governo possam compartilhar seus bancos de dados, sem revelar informações sobre pessoas “não suspeitas”. Além disso, tem sido estimulado o desenvolvimento de procedimentos de autenticação para controlar o acesso às informações disponibilizadas nos bancos de dados, assim como a criação e uso de arquivos de *log* (registros de transações e arquivos de auditoria) para monitorar o acesso e julgar eventuais abusos ou usos inapropriados da informação.

Outras controvérsias dizem menos respeito à privacidade quanto a problemas ligados à “representatividade”. Recentemente, a Google anunciou o projeto de indexar e digitalizar milhões de livros e publicações de algumas das maiores bibliotecas do mundo, entre elas as das universidades de Stanford, Harvard, Oxford, Michigan e da Biblioteca Pública de Nova York. Muitos afirmaram que a digitalização desses textos significaria dar a um grande número de indivíduos e populações acesso ao “conhecimento global”. Outros, entretanto, preocuparam-se com o fato de que o critério de seleção fosse guiado pelo olhar anglo-saxão e os efeitos que isso acarretaria a respeito de especificidades lingüísticas e culturais ou em termos de identidades nacionais. Nesse sentido, o projeto foi especialmente controvertido na França.²⁶

É interessante notar que se trata, ao mesmo tempo, de um problema de representação política e estatística. Se bem é formulada em termos de identidades lingüísticas ou nacionais, a questão não parece ser a da justa representação dos indivíduos pelos seus dirigentes ou de diferentes línguas e dialetos por uma língua oficial, mas, antes, a da representação dos diferentes países e línguas em bancos de dados e coleções digitalizadas. Na medida em que se pressupõe que todos os “conteúdos” (desde os dados pessoais às histórias dos Estados-Nação, desde as anotações e diários íntimos à memória coletiva) são “codificados” e “publicados” em formato digital, o que importa é, sobretudo, a inclusão das diferentes populações nessas coleções (Cf. FOUCAULT, 1999).²⁷ Nesse sentido, caberia pensar que os artefatos e mecanismos aqui estudados dizem

recursos humanos ou melhorar os serviços prestados pelas agências governamentais (Cf. OLSEN, 2004a; OLSEN, 2004b; ZETTER, 2004).

²⁵ A expressão é do diretor do Center for Democracy and Technology, citada em Zetter (2004).

²⁶ Foi o diretor da Biblioteca Nacional da França, um renomeado historiador, quem alertou sobre o que ele entendia como a “ameaça” do Google: “Não gostaria que se falasse sobre a Revolução Francesa somente em livros selecionados pelos Estados Unidos. É possível que nossa idéia sobre o tema seja pior, mas será nossa.” Como resposta ao projeto da Google e às reações que ele provocou, o presidente da França teria se reunido com autoridades francesas e dos governos vizinhos para avaliar as condições para que as coleções das grandes bibliotecas francesas e européias possam ser colocadas rapidamente na internet (Cf. FRANÇA CONTRA-ATACA..., 2005; FRANÇA PROTESTA..., 2005)

²⁷ Foucault analisa, no curso dos anos 1977-1978, a relação histórica entre três movimentos: aquele que coloca em questão a noção de soberania e introduz o problema do governo (o governo dos outros e de si), o movimento que faz aparecer a população como um “dado” e, em terceiro lugar, aquele que faz da economia um campo específico de realidade e da economia política, uma técnica de intervenção sobre esse campo. Na quarta aula desse curso, analisou como, a partir do século XVIII, a noção de “economia” foi tomando seu sentido moderno, ligado à administração das pessoas e as coisas. Essa noção, que designava até então a administração de um conjunto reduzido de indivíduos

menos respeito a instrumentos para legitimar a posse de um território ou a proteção de uma propriedade, quanto a técnicas de governo das pessoas e das coisas, que são imanentes ao campo da população.

Referências

AKRICH, M. The de-scription of technical objects. In: BIJKER, W. E., LAW, John (Ed.) *Shaping technology / building society: studies in sociotechnical change*. Cambridge: The MIT Press, 1992.

ARAYA DUJISIN, R. El factor sociotecnológico. *Nueva Sociedad*, Caracas, v. 117, p. 154-160, ene./feb. 2002.

BARRY, A.; OSBORNE, T.; ROSE, N. (Ed.) *Foucault and political reason: liberalism, neo-liberalism and rationalities of government*. Chicago: The University of Chicago Press, 1996.

BECK, U.; BONSS, W.; LAU, C. The theory of reflexive modernization: problematic, hypotheses and research programme. *Theory, Culture & Society*, Thousand Oaks, CA: Sage Publications, v. 20, n. 2, p.1-33, Apr. 2003.

BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY. Lawrence Page. Method for node ranking in a linked database. Int CI G06F 017/30, Jan. 9, 1998. Disponível: <<http://www.uspto.gov/patft/>>. Acesso em: 14 fev. 2005.

BRUNO, C. *The Google adwords happening*. Disponível em: <<http://www.iterature.com/adwords/>>. Acesso em: 6 jul. 2004.

CALLON, M.. El proceso de construcción de la sociedad: el estudio de la tecnología como herramienta para el análisis sociológico. In: DOMÉNECH, Miguel; TIRADO, Francisco J. (Ed.) *Sociologia simétrica*. Barcelona: Gedisa, 1998. p. 143-170.

CENDON, B. V. Ferramentas de busca na web. *Ciência da Informação*, Brasília: Instituto Brasileiro de Ciência e Tecnologia, v. 30, n. 1, p. 39-49, jan./abr. 2001. Disponível em: <<http://www.scielo.br/>>. Acesso em: 22 mar. 2005.

DELEUZE, G. *Foucault*. Buenos Aires: Paidós, 2003.

e riquezas (ligados à casa e à família), é introduzida no exercício político. Trata-se de um processo não de “estatização da sociedade” mas de “governamentalização” que é, ao mesmo tempo, interior e exterior ao Estado. Nesse processo, a estatística foi um dos elementos que permitiu o “desbloqueio da arte de governar”. Ao permitir agregar e quantificar os fenômenos da população, os procedimentos da estatística os mostram como acontecimentos irredutíveis ao marco reduzido da família e os tomam como objeto em função seus efeitos econômicos específicos. Foucault propôs, nesse sentido, uma história da “governamentalidade”, entendida como o conjunto de instituições, procedimentos e técnicas que tem como objeto a gestão das populações, como forma primordial de saber a economia política, e cujos mecanismos técnicos essenciais são os dispositivos de segurança. Retomando em muitos pontos essa problemática, e se inspirando provavelmente no curso de Foucault sobre o nascimento da biopolítica, Deleuze analisa como, em meios abertos, o poder deixa de ser massificante e individuante (deixa de moldar a individualidade de cada membro de um corpo coletivo), e passa a operar pelo exercício de um controle contínuo sobre “dividuais” e sobre massas, que se tornaram “amostras, mercados ou bancos” (DELEUZE, 1990, p. 222).

_____. Post-criptum sobre as sociedades de controle. In: _____. *Conversações*. Rio de Janeiro: Ed. 34, 1990.

DELEUZE, G; GUATTARI, F. *Mil platôs: capitalismo e esquizofrenia*. Rio de Janeiro: Ed. 34, 1997-1999.

FERREIRA, P. Máquinas sociais: o filo maquínico e a sociologia da tecnologia. In: ARTE, TECNOLOGIA E POLÍTICA, 2004, Campinas. *Anais eletrônicos...* Campinas: IFCH/Unicamp, 2004. Disponível em: <<http://www.ifch.unicamp.br/cteme>>. Acesso em: 1 fev. 2005.

FOUCAULT, M. La gubernamentalidad. In: _____. *Estética, ética y hermenéutica*. Barcelona: Paidós, 1999. p. 175-197. (Obras Esenciales, 3). Quarta aula do curso do Collège de France, 1977-1978: Segurança, território e população.

_____. *Vigilar y castigar: el nacimiento de la prisión*. México: Siglo Veintiuno, 1987.

FRANÇA CONTRA-ATACA projeto do Google. *Observatório da Imprensa*, São Paulo, ano 11, n. 322, 29 mar. 2005. Disponível em: <<http://observatorio.ultimosegundo.ig.com.br/artigos.asp?cod=322ENO003>>. Acesso em: jun. 2005.

FRANÇA PROTESTA contra biblioteca do Google. *O Estado de S. Paulo*, São Paulo, 19 fev. 2005.

GLASNER, Joanna. More reasons to love Google. *Wired News*, New York: CondéNet, Apr. 29, 2004. Disponível em: <<http://www.wired.com/news/business/0,1367,63286,00.html>>. Acesso em: 20 mar. 2005.

GOOGLE INC. *Google fun facts*. Disponível em: <<http://www.google.com/press/funfacts.html>>. Acesso em: 20 mar. 2005.

HUTCHBY, I. Technologies, texts and affordances. *Sociology*, Thousand Oaks, CA: Sage Publications, v. 35, n. 2, p. 441-456, May 2001.

KASPER, C. P. Aspectos do desvio de função. In: ARTE, TECNOLOGIA E POLÍTICA, 2004, Campinas. *Anais eletrônicos...* Campinas: IFCH/Unicamp, 2004. Disponível em: <<http://www.ifch.unicamp.br/cteme>>. Acesso em: 1 fev. 2005.

LATOUR, B. *Aramis, or the love of technology*. Cambridge: Harvard University Press, 1996.

_____. *Ciência em ação: como seguir cientistas e engenheiros sociedade afora*. São Paulo: Ed. Unesp, 2000.

_____. Is re-modernization occurring – and if so, how to prove it?: a commentary on Ulrich Beck. *Theory, Culture & Society*, Thousand Oaks, CA: Sage Publications, v. 20, n. 2, p. 35-57, Apr. 2003.

_____. *Jamais fomos modernos: ensaio de antropologia simétrica*. Rio de Janeiro: Ed. 34, 1994.

_____. *La science en action*. Paris: Éditions la Découverte, 1989.

_____. The powers of association. In: LAW, John (Ed.). *Power, action and belief: a new sociology of knowledge?*. London: Routledge & Kegan Paul, 1996. p. 264-280.

_____. When things strike back: a possible contribution of “science studies” to the social sciences. *The British Journal Of Sociology*, London: Blackwell Publishing, v. 51, n. 1, p. 107-123, Jan./Mar. 2000.

LAW, J. (Ed.). *Power, action and belief: a new sociology of knowledge?* London: Routledge & Kegan Paul, 1996, p. 1-19.

LAW, J. Networks, relations, cyborgs: on the social study of technology. London: 2000. On-line paper. Disponível em: <<http://www.comp.lancs.ac.uk/sociology/soc042jl.html>>. Acesso em: 10 ago 2003.

MACKENZIE, A. Open source software: when is a tool? What is a commodity? *Science As Culture*, London: Routledge, v. 10, n. 4, p. 541-552, Dec. 2001.

_____. The potential of the attractor: non-metaphorical transport between sciences and social theory. Jan. 2004. Paper on-line disponível em: <<http://www.lancs.ac.uk/staff/mackenza/papers/non-metaphorical-complexity.pdf>>. Acesso em: 20 jul. 2004.

_____. *Transductions: bodies and machines at speed*. London: Continuum, 2002.

MANOVICH, L. Database as a genre of new media. *Ai & Society*, London: Springer, v. 14, n. 2, p. 176-183, June 2000.

OLSEN, Florence. GAO finds widespread fed data mining. *Federal Computer Week*, Falls Church, VA: FCW Media Group, May 28, 2004a. Disponível em: <<http://www.fcw.com/article83039-05-28-04-Web>>. Acesso em: jul. 2005.

_____. Where oh where does all the data go? *Federal Computer Week*, Falls Church, VA: FCW Media Group, v. 18, n. 22, p. 40, May 7, 2004b. Disponível em: <<http://www.fcw.com/fcw/articles/2004/0705/pol-gao-07-05-04.asp>>. Acesso em: jul. 2005.

PAGE, L.; BRIN, S. The anatomy of a large-scale hypertextual web search engine. *Computer Networks And ISDN Systems*, [S.l.]: Elsevier, v. 30, n. 1-7, 1998a. Disponível em: <<http://www.sciencedirect.com/science/journal/01697552>>. Acesso em: 1 fev. 2005. Texto integrante dos anais da 7ª International World Wide Web Conference (WWW7), em Brisbane, 1998.

PAGE, L.; BRIN, S. The anatomy of a large-scale hypertextual web search engine. Stanford: Computer Science Department/Stanford University, 1998b. Disponível em:<<http://www-db.stanford.edu/~backrub/google.html>>. Acesso em: 1 fev. 2005.

POSTER, Mark. *The mode of information: poststructuralism and social context*. Cambridge: Polity Press, 1990.

PRIVACY RIGHTS CLEARINGHOUSE et al. *Thirty-one privacy and civil liberties organizations urge Google to suspend Gmail: an open letter to Google regarding its proposed Gmail service*. San Diego, 2004. Disponível em: <<http://www.privacyrights.org/index.htm>>. Acesso em: 10 mar. 2005.

ROBOTSTXT.ORG. *The web robot pages*. Disponível em: <<http://www.robotstxt.org/wc/robots.html>>. Acesso em: ago. 2005.

SANTOS, L. Garcia dos. *Politizar as novas tecnologias: o impacto sóciotécnico da informação digital e genética*. São Paulo: Editora 34, 2003.

SASSEN, S. *The global city*: New York, London, Tokyo. Princeton, NJ: Princeton University Press, 1991.

SIMONDON, G. *Du mode de existence des objets techniques*. Paris: Aubier-Montaigne, 1969.

STRATHERN, M. Cutting the network. *Journal of The Royal Anthropological Institute*, London: Royal Anthropological Institute of Great Britain and Ireland v. 2, n. 3, p. 517-535, Sept. 1996.

_____. *The relation: issues in complexity and scale*. Cambridge: Prickly Pear Press, 1995.

TERDIMAN, Daniel. Lawsuit: Google stole Orkut code. *Wired News*, New York: CondéNet, Jun. 30, 2004. Disponível em: <<http://www.wired.com/news/business/0,1367,64046,00.html>>. Acesso em: jul. 2005.

UNITED STATES OF AMERICA. Securities and Exchange Commission. *Registration Statement of Google Inc.:* form S-1 under the Securities Act. Washington, DC: SEC, 2004. Disponível em: <<http://www.sec.gov/>>. Acesso em: 22 fev. 2005.

WIKIPEDIA. *Google*. Disponível em: <<http://en.wikipedia.org/wiki/google>>. Acesso em: Acesso em: 10 mar. 2005. (a)

_____. *User agent*. Disponível em: <http://en.wikipedia.org/wiki/user_agent>. Acesso em: ago. 2005. (b)

ZETTER, Kim. GAO: fed data mining extensive. *Wired News*, New York: CondéNet, May 27, 2004. Disponível em: < <http://www.wired.com/news/privacy/0,1848,63623,00.html>>. Acesso em: ago. 2005.