

# O oráculo de Mountain View: o Google e sua cartografia do ciberespaço

**Fernanda Bruno<sup>1</sup>**  
**Antônio Leitão**  
**Arthur Lobo**  
**Bruno Boghossian**  
**Luciana Guilhon Albuquerque**  
**Natália Guimarães**  
**Priscila Biancovilli**

**Resumo:** *Este artigo identifica e analisa a constituição de uma nova cartografia do ciberespaço, cujo diferencial consiste no fato de ela não ser apenas um mapeamento dos lugares e dados do espaço informacional, mas sobretudo um monitoramento e uma vigilância dos usuários e seus percursos nesse espaço. A partir desse monitoramento, os rastros deixados na Rede são ordenados e classificados em bancos de dados, possibilitando a construção de perfis sobre padrões comportamentais, preferências e desejos de inúmeros indivíduos e/ou grupos. Os diversos serviços do Google são hoje o maior e mais eficiente exemplo dessa cartografia, constituindo o nosso objeto privilegiado de análise. Procura-se mostrar como se constrói uma cartografia do ciberespaço que é ao mesmo tempo uma taxonomia dos usuários e uma simulação de suas identidades, cujo traço central reside na sua dimensão performativa.*

**Palavras-chave:** *cartografia, vigilância, Google.*

---

<sup>1</sup> Fernanda Bruno (fgbruno@matrix.com.br) é professora do Programa Pós-Graduação em Comunicação e Cultura da UFRJ e Coordenadora do CiberIdea – Núcleo de pesquisa em tecnologias da comunicação, cultura e subjetividade/UFRJ. Antônio Leitão, Arthur Lobo (bolsista PIBIC), Bruno Boghossian, Luciana Guilhon Albuquerque, Natália Guimarães (bolsista PIBIC), Priscila Biancovilli são pesquisadores de iniciação científica do CiberIDEA/UFRJ.

**Resumé:** Cet article identifie et analyse la constitution d'une nouvelle cartographie du cyberspace, dont la différence est dans le fait qu'elle n'est pas seulement une cartographie des lieux et des données de l'espace numérisé, mais surtout une surveillance des utilisateurs et de leurs parcours dans cet espace. De cette surveillance, les vestiges laissés dans la Web sont ordonnés et classifiés en bases de données qui rendent possible la construction de profils sur les habitudes, les préférences et les désirs d'innombrables individus et/ou groupes. Les différents services de Google sont aujourd'hui le plus grand et le plus efficace exemple de cette cartographie, et constituent notre objet d'analyse privilégié. Nous essayons de montrer comment une cartographie du cyberspace est construite, étant en même tant une taxonomie des utilisateurs et une simulation de ses identités, dont le trait central se caractérise par sa dimension performative.

**Mots-clés :** cartographie; surveillance; Google.

**Abstract:** This paper identifies and analyses the constitution of a new cyberspace cartography, whose differential consists of not being just a mapping of the places and data of the informational space, but especially a monitoring and a surveillance of the users and their routes in this space. Starting from this monitoring, the tracks left on the Web are ordered and classified in databases, making possible the construction of profiles about behavioral standards, preferences and desires of countless individuals and/or groups. The several Google services are today the biggest and more efficient example of this cartography, constituting our privileged object of analysis. We intent to show how this cartography of the cyberspace is at the same time a taxonomy of the users and a simulation of their identities, whose central feature resides in its performative dimension.

**Key-words:** cartography; surveillance; Google.

**Resumen:** Este artículo identifica y analiza la construcción de una nueva cartografía del ciberespacio, cuyo diferencial consiste en el hecho de ella no ser solamente un mapeamiento de los lugares y datos del espacio informacional, pero sobretodo una vigilancia de los usuarios y sus recorridos en este espacio. A partir de esta vigilancia, los rastros dejados en la Red son ordenados y clasificados en bases de datos, posibilitando la construcción de perfis sobre modelos comportamentales, preferencias y deseos de innumerables individuos y/o grupos. Los diversos servicios del Google son hoy el mayor y más eficiente modelo de esta cartografía, construyendo el nuestro objeto privilegiado de análisis. Intentamos mostrar como si construye una cartografía del ciberespacio que es al mismo tiempo una taxonomía de los usuarios y una simulación de sus identidades, cuyo carácter principal reside en su dimensión performativa.

**Palabras-llave:** cartografía; vigilancia, Google.

*Naquele Império, a Arte da Cartografia alcançou tal Perfeição que o mapa de uma única província ocupava toda uma Cidade, e o mapa do império, toda uma Província. Com o tempo, esses Mapas Desmesurados não foram satisfatórios e os Colégios de Cartógrafos levantaram um Mapa do Império, que tinha o tamanho do Império e coincidia pontualmente com ele. Menos Afeitas ao Estudo da Cartografia, as Gerações Seguintes entenderam que esse dilatado Mapa era Inútil e não sem Impiedade o entregaram às Inclemências do Sol e dos Invernos. Nos desertos do Oeste perduram despedaçadas Ruínas do Mapa, habitadas por Animais e por Mendigos; em todo o País não há outra relíquia das Disciplinas Geográficas.*  
(Borges, J.L., 1999)

Em Mountain View, Califórnia, encontra-se situada a sede do *Google*, a empresa de buscas e serviços na Internet que, conforme procura mostrar este artigo, efetua hoje a maior cartografia do ciberespaço – especificamente, da Internet. Esta cartografia, como veremos, tem um efeito oracular e incide massivamente sobre os indivíduos que exploram o espaço informacional através do *Google*.

### **Mapeando o ciberespaço**

O surgimento de uma nova modalidade de ‘espaço’ em nossa cultura, o espaço informacional ou o ciberespaço, suscitou a demanda por novas estratégias cartográficas. A consolidação da Internet, em particular, levantou uma série de questões sobre as possibilidades de representação e exploração deste novo espaço nem geométrico nem geográfico; um espaço fluido que não cessa de ser alterado e do qual não se pode traçar os contornos ou apreender a totalidade. Como mapear um território movente e sem bordas? Os primeiros desafios cartográficos da rede lidavam com as dificuldades de se representar e de se construir guias, caminhos nesse ambiente que não coincidia com os conhecidos espaços físico-geométricos (Cf. Dodge & Kitchin, 2001).

Com o crescimento da Internet na década de 1990, um outro problema vai suscitar novos esforços cartográficos: o excesso de informação<sup>2</sup>. A dificuldade de se encontrar em tempo hábil a informação desejada passa a ser um limite nesse espaço ilimitado de informações (Vaz, 2004). Por conta disso, antes do surgimento dos mecanismos de busca mais sofisticados, foram criados os *portais*: sítios que oferecem uma variedade de *links* considerados “úteis e de bom conteúdo” para seus usuários. Torna-se então necessário filtrar e classificar a informação de acordo com os interesses de determinado público. O *portal* surgiu exatamente para mediar o caminho entre a informação e o usuário, através de ambientes virtuais que encurtam a busca e facilitam o percurso pelo grande emaranhado de páginas da rede.

No entanto, essa primeira forma de garantia da qualidade do conteúdo se revelou insuficiente para atender aos interesses dos usuários, pois restringia demasiadamente o seu campo de escolha. Assim, um outro tipo de serviço, mais interativo, mostrou-se indispensável na tentativa de encontrar a agulha no palheiro da *Web*: os sistemas de busca. Com o advento de uma ferramenta capaz de superar a fixidez dos portais e ampliar o acesso à rede, conferiu-se maior abrangência e multiplicidade às pesquisas virtuais, já que os próprios internautas passaram a solicitar, através de palavras-chave, a informação de que necessitavam.

Os primeiros mecanismos de busca da *Web* – até meados da década de 1990 – surgem como sistemas que encontram as palavras-chave indicadas pelos usuários, que indexavam as páginas por meio de diretórios construídos manualmente e mantidos quase que exclusivamente pelo interesse dos internautas. A iniciativa de ter sua página cadastrada partia do próprio usuário, que digitava as palavras-chave e uma descrição do *site* que desejava cadastrar no sistema. Esse empreendimento, no entanto, provou-se falho, pois, segundo Berners-Lee, “os mecanismos de busca perdem-se em meio à massa indiferenciada de documentos que oscilam muito em

---

<sup>2</sup> Estudando a história da Internet, podemos observar que sua popularização foi incentivada pela criação de *softwares* e navegadores que facilitaram o acesso à informação, desenvolvendo uma interface simples o bastante para ser manipulada por qualquer usuário, o que contribuiu ainda mais para a infinita produção de informação na rede. Desde os seus primórdios até o fim da década de 1980, quando a Internet ainda interligava apenas laboratórios e universidades nos EUA, toda a informação corria em formato de texto. Apenas em 1993 foi desenvolvido o primeiro navegador, o *Mosaic*. O *software* foi fundamental para o crescimento da *Web* por dois motivos: por ter conseguido acrescentar sons e imagens ao que antes eram apenas textos e, também, por ter facilitado o acesso à informação.

termos de qualidade, temporalidade e relevância da informação” (1998). Devido à expansão cada vez mais intensa da rede, torna-se necessário o aprimoramento dos mecanismos automáticos de busca que trabalham através dos *crawlers*.

Os primeiros *crawlers* indexavam apenas os títulos das páginas da *Web*; as versões atuais são *softwares* que rastreiam diariamente informações novas em inúmeros *sites*<sup>3</sup>, catalogando o conteúdo de páginas inteiras, bem como de diversos tipos de arquivos e documentos. “O *crawler* envia seus dados a uma enorme base de dados denominada índice. (...) Para cada *site*, o índice lista todas as páginas e todas as informações pertinentes a respeito delas: as palavras em cada página, as conexões, o texto-âncora (texto em torno e dentro de uma conexão), e assim por diante” (John Battelle, 2005). Por funcionar de forma automatizada, o *crawler* possui um banco de páginas exponencialmente maior que os mecanismos de busca antigos, que dependiam da inclusão manual de *sites*.

Através do desenvolvimento de mecanismos de busca por palavras, criação de filtros e dispositivos de seleção, o excesso de informação começou a ser contornado. Todos sabemos que hoje é praticamente impensável navegar na Internet sem o auxílio dos sistemas de busca; estes se tornaram “mapas” indispensáveis à exploração do espaço informacional da Internet. Recentemente, o mecanismo de busca *Google*<sup>4</sup> mostrou-se o mais eficiente da *Web*, uma vez que põe em prática os princípios capazes de atender às demandas dos usuários na busca por informação relevante em tempo hábil: rastreia constantemente as páginas da Internet e mantém-se atualizado, além de filtrar e classificar os dados encontrados, criando informação que o auxilia a organizar a própria informação.

---

<sup>3</sup> A Internet comporta aproximadamente 2,5 bilhões de documentos. A taxa de crescimento de *sites* é de 7,3 milhões de páginas diárias, o que significa uma média 100 bilhões de bytes a mais a cada dia na rede. Se for considerado tudo o que a Internet acessa, entre páginas da *web*, banco de dados e intranets, há 550 bilhões de documentos conectados na rede. ([http://agenciact.mct.gov.br/index.php?action=/content/view&cod\\_objeto=15023](http://agenciact.mct.gov.br/index.php?action=/content/view&cod_objeto=15023)). Em 2005, o *Google* alcança a incrível marca estimada de 60 bilhões de páginas indexadas. [<http://pt.wikipedia.org/wiki/Google>].

<sup>4</sup> Para saber mais sobre o *Google* cf. <http://pt.wikipedia.org/wiki/Google> ; [www.google.com/intl/en/corporate/history.html](http://www.google.com/intl/en/corporate/history.html) .

O *Google* se destaca graças ao seu *crawler* – o *Googlebot*<sup>5</sup> – que rastreia documentos na rede e cria bancos de dados para serem utilizados nas buscas. Outro fator determinante para o seu sucesso é o sistema *PageRank*, um algoritmo que ordena e classifica cerca de 25 bilhões de páginas da *Web* indexadas por *crawlers* de acordo com a sua pontuação em diversos parâmetros: examina as conexões dentro da página, o texto-âncora em torno delas e a popularidade das páginas que se conectam a outras páginas; todos esses fatores são considerados para conferir a relevância final de uma determinada página para a busca efetuada (Cf. Batelle, op. cit., p. 17-18). Deste modo, diferente dos outros mecanismos, o *Google* classifica as páginas em notas de zero a 10, a partir de critérios que pressuponham “credibilidade”: quanto mais *links* apontarem para determinada página, maior será sua *nota* e, assim, sua relevância na página de resultados. Vale, contudo, ressaltar que tais parâmetros e fatores são apenas a parte relativamente explícita dos dispositivos postos em obra pelo *Google*. Na realidade, seus algoritmos analisam “mais de 100 fatores para determinar a relevância de um *site* para as suas palavras-chave” (Idem, p. 18).

Deste modo, o *Google* ilustra de forma exemplar o que chamamos de uma “nova cartografia”, a qual está em pleno desenvolvimento e passa, a nosso ver, por uma transição decisiva – de uma cartografia do espaço informacional para uma cartografia dos usuários ou habitantes deste espaço. Expliquemos: dado que o espaço informacional é constituído e modificado pelos próprios percursos que os indivíduos ali efetuam, a “melhor” maneira de mapear este espaço é mapear os próprios percursos e, em última instância, os próprios indivíduos. Ou seja, em vez de se mapear simplesmente as informações procuradas, mapeia-se também as próprias buscas, delas extraindo padrões que por sua vez otimizarão outras buscas. Nesta cartografia, o foco não é tanto o território, mas os trajetos nele efetuados e os indivíduos que nele transitam – seus hábitos, suas preferências, seus interesses. Na análise deste empreendimento cartográfico - do qual o *Google* é o exemplo privilegiado - pretendemos mostrar como se dá esta transição, quais são as suas características centrais e algumas de suas implicações para um novo regime de constituição e simulação de identidades.

---

<sup>5</sup> Para maiores informações sobre *crawler* e sobre o *Googlebot*, Cf. <http://www.google.com/webmasters/bot.html>; [www.idgnow.uol.com.br/AdPortalv5/InternetInterna\\_290605.html](http://www.idgnow.uol.com.br/AdPortalv5/InternetInterna_290605.html).

## **Cartografar é classificar**

Toda cartografia envolve procedimentos de classificação, seleção e ordenação do que se deseja mapear segundo certos princípios, valores e propósitos. O caso do espaço informacional e do *Google* não é diferente. Sobretudo neste caso, cartografar é classificar, é criar informação sobre a informação, produzindo uma “representação” ordenada da rede que certamente não é o retrato fiel da sua totalidade, mas uma imagem que traduz a própria Internet segundo certos critérios, operando classificações que selecionam regiões dentro desse espaço indeterminado. Notaremos mais tarde que esse mapa, apesar de ser apenas um recorte parcial, acaba por se fazer valer como a própria rede na medida em que, no limite, se efetiva como seu território aquilo que é indexado pelos serviços de busca, uma vez que não se vê quase nada além deste mapa.

Este mecanismo classificatório é também uma forma de produzir conhecimento. Ao disponibilizar a informação segundo uma ordem e um formato distintos daqueles apresentados quando estamos “imersos” nos dados, o mapa permite que se apreenda ou se infira padrões, características, regularidades, tendências que de outro modo – isto é, sem a mediação do mapa – não seriam visualizados. Todo mapa é, portanto, uma máquina classificatória e também uma máquina epistemológica. O mapa desenhado pelo mecanismo de busca *Google* e seus outros serviços constitui um sistema classificatório e epistemológico acerca dos indivíduos e seus trajetos na Internet, aumentando significativamente o poder social da informação coletada e gerada nesse processo.

Vejamos agora como funcionam os mecanismos de classificação vigentes no *Google*, levando-se em conta que eles não se aplicam apenas aos dados e “lugares” da rede, mas envolvem o monitoramento dos próprios indivíduos que nela transitam, produzindo um conhecimento sobre seus hábitos, comportamentos, preferências, desejos etc. Começemos por compreender os dispositivos técnicos que possibilitam tal classificação.

Um dos dispositivos essenciais é a utilização de *cookies*, pequenos arquivos que permitem a identificação do navegador e que recolhem informações sobre as preferências do usuário e as páginas visitadas, rastreando as tendências e padrões



utilizados nas pesquisas feitas através de mecanismos de busca, como o *Google*. Outro dispositivo utilizado pelo *Google* (e por outros mecanismos de busca), são os chamados *server logs*, capazes de gravar as páginas visitadas e coletar informações sobre o número do IP (*Internet Protocol*)<sup>6</sup>, o tipo e a língua do navegador, a data e a hora da pesquisa. Ou seja, quando qualquer indivíduo entra no *site* de buscas do *Google*, fornece, sem perceber, algumas informações sobre seu comportamento, suas preferências, seu modo de realizar pesquisas na Internet. Apesar de não serem informações que necessariamente identifiquem pessoalmente o usuário – como e-mail, nome etc – são informações sobre comportamentos individuais, que tornam possível ao *Google* construir perfis de seus usuários. Há, contudo, alguns serviços, como o *Gmail* (mail.google.com) e o *Orkut* (www.orkut.com), que exigem informações pessoais, como nome e endereço eletrônico<sup>7</sup>.

Os *cookies* e os *server logs* são exemplos de dispositivos que facilitam a coleta de informação sem a necessidade do consentimento explícito dos indivíduos, pois a explicitação de seu funcionamento está no texto sobre a política de privacidade da empresa, o qual muitos usuários não lêem<sup>8</sup>. Além disso, caso o usuário possa e queira bloquear os *cookies*, fica impossibilitado de utilizar a maioria dos serviços e recursos disponibilizados no *Google*. Voluntária ou involuntariamente, a informação acaba funcionando como moeda de troca em um mundo onde os consumidores já estão habituados a fornecer informações pessoais em troca de serviços (Cf. Dholakia; Zwick, 2001). Em suma, ao usar qualquer um dos serviços *Google*, o usuário fornece dados como: o que está buscando, que marca de televisor prefere comprar, que tipo de propaganda chama sua atenção, ou até mesmo seu endereço e o tipo de cartão de crédito que utiliza. O *Google* armazena, organiza e classifica toda a informação em bancos de dados, o que o possibilita apreender padrões e construir perfis computacionais sobre comportamentos, hábitos e preferências de um enorme número de usuários de seus serviços.

---

<sup>6</sup>Protocolo da Internet. Este protocolo identifica, localiza e estabelece conexão entre computadores ligados à Internet (<http://www.aisa.com.br/diciona.html>).

<sup>7</sup>Contudo, ao que parece, não há uma checagem destes dados, o que em princípio deixa aberta a possibilidade de o usuário fornecer informações falsas.

<sup>8</sup>Cf. [www.google.com/privacy.html](http://www.google.com/privacy.html).



Esse processo pode ser observado com clareza nos serviços de publicidade do *Google*. O *Google Adwords* ([adwords.google.com](http://adwords.google.com)), contratado por anunciantes, viabiliza a criação de anúncios direcionados ao público-alvo no momento em que se efetua uma busca de palavra-chave relacionada ao produto anunciado. Essa interação entre *Google* e usuário/potencial consumidor, constitui uma vigilância eletrônica que atua em mão dupla, ou seja, tanto coletando dados quando o usuário expõe uma demanda, quanto utilizando esses dados posteriormente para a melhor adequação de seus serviços ao mercado e às demandas dos usuários ou outros fins.

O *Google Adsense* ([google.com/adsense](http://google.com/adsense)), por sua vez, está voltado para os editores de *sites*, fornecendo automaticamente anúncios gráficos e de texto que são relacionados com precisão ao seu *site* e ao seu conteúdo. Pretende-se que os anúncios sejam tão adaptados ao contexto, que os leitores realmente os considerem úteis. Dessa forma, o serviço permite que, de uma forma rápida e fácil, *sites* de quaisquer dimensões exibam anúncios *Google* relevantes ao conteúdo de cada página, capitalizando-as.

Nota-se, a partir desses exemplos, o quanto vigora hoje no ciberespaço um dispositivo de vigilância eletrônica (Cf. Bruno, 2006) que funciona de forma automatizada, ininterrupta e muitas vezes imperceptível, monitorando as atividades *online*, com o objetivo de identificar, classificar e acessar usuários. Assim, o *Google* garante seu sucesso diferenciado por lançar mão de uma estratégia de rastreamento que ultrapassa a mera cartografia das informações disponibilizadas na rede e se dirige para os indivíduos. Este foi o “pulo do gato” do *Google*: a empresa percebeu que a maneira mais eficiente de se cartografar o espaço informacional era através da cartografia dos seus usuários e dos aspectos que se pode extrair dela, como interesses, hábitos, padrões de comportamento e de consumo.

Essa nova estratégia cartográfica centrada nos usuários se mostrou tão eficiente que passou a ter cada vez mais importância nos serviços do *Google*, que não cessam de se expandir, extrapolando a mera função de facilitar a busca de informações na rede. Hoje, além do mecanismo de busca, a empresa oferece desde o sítio de relacionamentos *Orkut* ([www.orkut.com](http://www.orkut.com)), até um correio eletrônico gratuito – o *Gmail* ([mail.google.com](http://mail.google.com)) - e serviços de informação personalizada com o *Google Notícias* ([news.google.com](http://news.google.com)), que fornece uma espécie de mapa das notícias mais lidas

nos principais jornais on-line, e o *Google Zeitgeist* ([www.google.com/zeitgeist](http://www.google.com/zeitgeist)), que mostra os termos e *sites* mais procurados no mundo. Há também *Google Earth* ([earth.google.com](http://earth.google.com)) e o *Local Search* ([local.google.com](http://local.google.com)), que dispõem de imagens de satélites de várias partes do mundo, bem como mapas locais, com nomes de ruas, restaurantes, cinemas, bibliotecas, estações de metrô e ônibus etc. E ainda o *Froogle* ([froogle.google.com](http://froogle.google.com)), o *Blogger* ([googleblog.blogspot.com](http://googleblog.blogspot.com)), o *Google Talk* ([www.google.com/talk](http://www.google.com/talk)), o *Book Search* ([books.google.com](http://books.google.com)), o *Google Groups* etc. Recentemente, foi criado o *Google Trends* ([www.google.com/trends](http://www.google.com/trends)), que fornece ao usuário o histórico de buscas de determinada(s) palavra(s), bem como o cruzamento desse histórico com o volume de notícias captado pelo *Google News* no período.

Deste modo, o *Google* aprimora seu mecanismo de busca e amplia seus serviços, produzindo bancos de dados e perfis de usuários cada vez mais sofisticados, realizando uma cartografia ampla e multifacetada sobre padrões de comportamento, sociabilidade, consumo etc. Vale lembrar que essa classificação não funciona agrupando os usuários em categorias inteiramente pré-estabelecidas. Boa parte destas categorias são definidas e construídas *a posteriori*. Ainda que numa etapa inicial os indivíduos sejam classificados em bancos de dados segundo categorias prévias – como gênero, faixa etária, nacionalidade etc – a intenção é sempre construir posteriormente classificações mais específicas que sejam um “efeito” das ações, percursos e hábitos dos usuários na rede (Cf. Arvidsson, 2004).

Vimos que todo o conhecimento aí produzido e encarnado nos perfis não está necessariamente fundado na coleta de informação pessoal em num conhecimento profundo das particularidades daqueles que são monitorados. Entretanto, veremos que os perfis pretendem ser extremamente personalizados, pois visam atender a demandas, características e interesses precisos de indivíduos e grupos. Esta estranha combinação de um conhecimento ao mesmo tempo impessoal e ultra-personalizado, infra-individual e individualizado será melhor compreendida se nos detivermos um pouco mais nos elementos e procedimentos envolvidos nessa cartografia dos indivíduos.

## **Uma taxonomia dos usuários: preferências, desejos, tendências**

Com a criação dessa inovadora ferramenta de classificação baseada no cruzamento de dados coletados a partir da movimentação dos indivíduos na *web*, o *Google* constituiu sua própria “taxonomia” dos usuários e um dispositivo que classifica, conhece e intervém no mundo onde está inserido. Essa cartografia ganha incomparável eficiência pelo fato de este conhecimento ser produzido por uma rede distribuída entre humanos e máquinas, delegando boa parte do trabalho aos próprios usuários e aos inúmeros bancos de dados e algoritmos, tornando-o um processo automatizado e ultra atualizado.

Os principais elementos dessa cartografia dos indivíduos são, portanto, os bancos de dados, seus algoritmos e os perfis computacionais gerados a partir deles. No contexto que estamos tratando, o que é propriamente um banco de dados? Uma primeira característica a ressaltar é a de que os bancos de dados arquivam informações que se situam num nível infra-individual. O próprio *Google* nomeia esse tipo de informação de “*aggregated non-personal information*”, a qual se refere a um comportamento ou perfil grupal e não a indivíduos identificáveis, podendo assim ser compartilhada com terceiros ([www.google.com/intl/en/privacy\\_faq.html](http://www.google.com/intl/en/privacy_faq.html)). Ou seja, os bancos de dados não concernem, primeiramente, a pessoas particulares, mas a grupos e populações organizados segundo categorias financeiras, comportamentais, profissionais, educacionais, geográficas etc. Eles não têm apenas a função de arquivo, mas uma função conjugada de registro, classificação, predição e intervenção (Bruno, 2006). A sua lógica, contudo, é menos a da exatidão e da verdade, do que a da agilidade e eficiência na recuperação e utilização das informações armazenadas (Cf. Poster, 1995). E o modo de ação dos bancos de dados envolve simultaneamente previsão, simulação e performatividade. O cruzamento de dados organizados em categorias amplas irá projetar e simular perfis que correspondam a indivíduos e corpos “reais” a serem pessoalmente monitorados, informados, acessados por ofertas de consumo ou serviço, incluídos ou excluídos em listas de mensagens publicitárias, marketing direto, campanhas de prevenção a algum tipo de risco etc.

Desta forma, os serviços oferecidos pelo *Google* constroem, quase que automaticamente, diversos bancos de dados que são incessantemente alimentados à medida que os próprios usuários se servem de tais serviços. Estes bancos de dados

contêm um tesouro de informações sobre os internautas. Esse é o primeiro plano do que chamamos de uma cartografia dos indivíduos e ela gera um conhecimento capaz de antecipar preferências, tendências e padrões comportamentais atuais e potenciais de inúmeros indivíduos, intervindo no seu campo de ações e escolhas. A passagem desse conhecimento num plano infra-individual para uma projeção e uma intervenção no plano individual é garantida pelo perfil. De fato, o próprio ato de classificação dos indivíduos nestes segmentos que constituem os bancos de dados já tem um potencial estimativo a partir do qual serão compostos perfis de indivíduos e grupos. Através da análise desses dados e da correlação entre múltiplas variáveis, compõem-se, por exemplo, perfis do ‘típico’ consumidor de música *pop*, do estudioso de línguas clássicas, do amante de instrumentos óticos, do leitor de romances ou de livros de auto-ajuda, do freqüentador de *chats* de namoro ou de *sites* pornográficos etc.

Vejamos alguns exemplos. O *Gmail*, serviço de e-mail do *Google*, destaca-se por oferecer um espaço de armazenamento de 2 *gigabytes*, gratuito, mas não isento de publicidade. Esta, no entanto, não se dá através de inconvenientes janelas *pop-up*, mas sim por meio de anúncios diretamente relacionados ao conteúdo das mensagens enviadas e recebidas. Por exemplo, se o assunto ou corpo da mensagem contém a palavra “passagem”, pode aparecer no canto direito da página os seguintes anúncios: “Ecoturismo by Freeway”; “Hotéis – passes de trem”; “Excursão a China 2006”. Apesar de muitas pessoas se assustarem com a familiaridade dos anúncios em relação ao conteúdo das mensagens, não há intervenção humana neste processo<sup>9</sup>. O sistema captura palavras-chave e as cruza com o banco de dados de anúncios publicitários, veiculando aqueles de maior relevância. Ainda que estes critérios de relevância e a própria eficácia do *Google Adsense* – mecanismo responsável por esses anúncios – sejam relativamente baixos<sup>10</sup>, esse serviço revela como o cruzamento e a classificação

---

<sup>9</sup> Conforme as informações oferecidas pelo *Gmail*: “Em um processo completamente automatizado, computadores processam o texto em uma mensagem e relacionam-no a anúncios ou informações relacionadas no amplo banco de dados do *Google*. Nenhuma pessoa lê sua correspondência para direcionar anúncios ou outras informações sem seu consentimento” (<http://mail.google.com/mail>).

<sup>10</sup> Isso acontece, principalmente, devido a dois fatores: a falta de análise semântica do conteúdo das mensagens e o número limitado de anunciantes relacionados à determinada palavra-chave. Se, por exemplo, aparecer em uma mensagem a palavra *colher*, o mecanismo não saberá diferenciar o verbo do substantivo, resultando numa publicidade nem sempre

de informações em bancos de dados produzem perfis que pretendem ser adequados, e muitas vezes o são, a desejos e interesses potenciais de usuários. Mas essa adequação, vale notar, encontra a sua eficácia não tanto na verdade ou lealdade do perfil em relação ao usuário, mas no seu caráter performativo. Isto é, ao antecipar um interesse na forma de oferta de um produto ou serviço que o indivíduo pode nem ter pensado, pretende-se não tanto atender a um interesse já dado, mas torná-lo efetivo, a partir da sua própria antecipação, incitando a ação de aceitar a oferta sugerida. E se o indivíduo acessa o anúncio, ele reforça o perfil, aumentando a sua probabilidade de eficácia, mas não tanto a sua veracidade, pois ele pode muito bem falhar com outros indivíduos ou com o mesmo indivíduo em outras ocasiões, sem com isso se tornar falso, mas sim menos eficaz.

O próprio mecanismo de buscas também ilustra esse processo de composição de bancos de dados e perfis de usuários. Ao introduzirmos no campo de busca do *Google* uma palavra qualquer, ela é entendida, em primeiro lugar, como uma intenção, desejo ou interesse. Para atender a essa demanda, o *Google* não procede fornecendo uma informação específica, mas um leque de opções e caminhos de busca onde o usuário poderá achar o que lhe interessa. Com vistas a esse objetivo, o método é o seguinte: primeiro, classificar as páginas da Rede segundo critérios de credibilidade através do método de *PageRank*, conforme já vimos. Deste modo, transforma-se uma parte da rede num banco de dados pronto para cruzamento. Paralelamente, toda uma cartografia da movimentação de dados na rede – isto é, dos procedimentos dos usuários, efetuada das mais diversas formas (feedback da utilização da procura, *cookies* e diversos outros meios de obtenção e de informação presentes nos serviços *Google*) – serve de cruzamento adicional para a composição de uma resposta individualizada à demanda do usuário: a página de resultados da busca é uma atualização do perfil daquele que procura determinada palavra. Ao acessar um dos documentos que compõem tal perfil, o usuário o torna efetivo, atualizando uma virtualidade e tornando “real” o que era apenas uma possibilidade. Aí reside o seu potencial performativo. E tudo se dá *como se* o resultado da busca fosse a resposta precisa e adequada a um desejo, uma intenção, *como se* fosse um

---

condizente com o conteúdo. Ou se uma pessoa de sobrenome japonês enviar um e-mail, pode ser que apareça algo nos anúncios relacionado a rodízios de *sushi* ou esportes nipônicos, ainda que o interlocutor sequer toque no assunto.

espelho desses últimos. No entanto, esse espelho não é o reflexo fiel de um referente estável e determinado, mas uma simulação de preferências, interesses, desejos, intenções que, como tal, encontra a sua efetividade não no seu compromisso com a verdade ou uma referencialidade qualquer, mas sim nas ações, escolhas e experiências que ele incita.

Notamos como uma simples busca está atrelada à produção de bancos de dados com potencial estimativo e à produção automatizada de perfis que funcionam como máquinas performativas. Cabe ainda ressaltar que esse processo é extremamente recursivo e que as respostas do usuário ao resultado obtido na busca ou aos *links* de produtos disponibilizados pelo *Adwords* ou *Adsense*, vistos nos exemplos acima, retroagem sobre os bancos de dados e perfis, tornando mais eficientes os serviços em jogo.

Esta cartografia dos usuários realizada pelo *Google* conta, ainda, em um de seus serviços, com a exposição deliberada dos indivíduos sobre seus traços de personalidade, interesses sociais, profissionais, culturais etc. Trata-se da rede *Orkut*, um *software* de formação de 'redes sociais' ([www.orkut.com](http://www.orkut.com)), que reúne pessoas e comunidades. As pessoas aí se apresentam na forma de perfis (constituídos de fotos, gostos e preferências pessoais, listas de amigos etc) e participam de comunidades. Além de ser uma rede social, o *Orkut* constitui um tesouro informacional a partir das trocas sociais e do que as pessoas declaram acerca de si mesmas (em seus perfis pessoais) ou acerca dos outros<sup>11</sup>. A classificação, o cruzamento e o processamento dessas informações permitem construir valiosos bancos de dados e perfis que expressem uma série de relações entre padrões comportamentais, traços de personalidade, campos de interesses, padrões de consumo, modos de interação social etc.

Em todos os exemplos mencionados, percebe-se o quanto os bancos de dados e as técnicas de composição de perfis procedem por uma antecipação de

---

<sup>11</sup> O *Orkut* tem sido objeto de diversos estudos no campo da comunicação, sendo a maior parte deles dedicada à análise de comunidades e redes sociais na Internet. Neste artigo, pretende-se ressaltar esta segunda face, menos visível da rede *Orkut*, em que o próprio espaço de comunicação e trocas sociais é também o espaço de coleta e classificação de informações, espaço de vigilância, que irá ser transformado em conhecimentos sobre personalidades, preferências e padrões sociais de indivíduos e/ou grupos. Cf. Adamic; Buyukkokten & Adar, 2003).

desejos, interesses e preferências na forma de oferta de produtos, *links* e serviços que os indivíduos podem sequer ter pensado ou saber que existiam. Trata-se de uma “oferta” que pretende ser superindividualizada, mas o perfil é menos um retrato fiel que representa neste caso a verdade dos desejos inerentes ao usuário da Internet, do que uma simulação deste desejo, que, ao se anunciar, tem uma efetividade performativa e proativa<sup>12</sup>, fazendo passar à realidade o que era apenas uma possibilidade, uma potencialidade.

Eis porque todo esse processo de simulação e antecipação comporta uma função quase oracular: os efeitos aí produzidos não se devem tanto à previsão de um futuro certo e necessário, mas ao próprio processo de antecipação, que acaba por intervir nas escolhas, comportamentos e ações presentes, tornando efetivo o que se antecipou. Nessa medida, a cartografia realizada pelos serviços *Google* encontra sua eficácia não tanto na representação de uma realidade prévia ou subjacente, mas nessa função oracular, que efetua uma ‘realidade’ na medida mesma em que a prevê, projeta ou antecipa.

Deste modo, as identidades projetadas nos perfis computacionais são simulações de identidades cuja efetividade não depende de vínculos profundos com os indivíduos a que correspondem, nem do espelhamento de uma personalidade ou caráter subjacentes. Vale notar que tal fenômeno só poderia se sedimentar numa sociedade já acostumada à precedência da imagem sobre o real. Assim, ao produzir perfis – identidades simuladas em que o modelo prevalece sobre o original – a partir do conhecimento acerca de tendências individuais e/ou coletivas, essa “nova vigilância” (Marx, 2002) consegue intervir nas escolhas dos indivíduos, agir sobre a ação possível. Nesse sentido, o perfil é uma construção probabilística relativamente descompromissada com o referente. O fenômeno da simulação, em que se rompe o compromisso com o referente, é um dos princípios básicos dos perfis e os caracteriza tanto numa dimensão temporal de antecipação do futuro como na relação entre imagem e original, onde vige o privilégio da hipermediação sobre o modelo representacional. O perfil é uma simulação dos desejos e preferências potenciais de indivíduos e grupos, cujo efeito de verdade se dá quando os indivíduos se identificam ou se reconhecem de algum modo no que lhes é antecipado, engajando desde então

---

<sup>12</sup> A noção de personalização proativa ganha força no marketing. Cf. Dholakia & Zwick, 2001.



algum tipo de comportamento ou escolha. Quando por exemplo aceito uma oferta personalizada de um produto que eu nem mesmo sabia existir ou que não havia desejado previamente, torno efetivo o perfil ou identidade que me foi antecipado e ao mesmo tempo o reforço para futuras previsões tanto a meu respeito quanto a respeito de outros indivíduos que habitam bancos de dados similares.

Cabe perceber que o principal objetivo em jogo não é tanto produzir um saber sobre um indivíduo especificamente identificável, mas usar um conjunto de informações pessoais para agir sobre outros indivíduos. Os perfis compostos pelos rastros deixados no ciberespaço, em vez de constituírem um todo unificado que corresponderia a um indivíduo real, compõem um mosaico de tendências coletivas. O perfil atua, portanto, como uma forma de organização e categorização do comportamento coletivo a partir do passado e do presente, visando à simulação de comportamentos futuros. Como um dos seus principais objetivos é o aprimoramento de serviços cada vez mais ajustados ao público, ele colabora no delineamento de percursos na rede, produzindo uma relação circular, onde ao mesmo tempo é produto do comportamento dos usuários e ajuda a produzir e reforçar padrões de atuação na Internet.

O caráter circular e recursivo desta cartografia, onde o mapa se sobrepõe ao território, será o nosso último ponto de análise.

### **Do rigor cartográfico e da arbitrariedade enciclopédica**

É provável que Jorge Luis Borges já tenha nos ensinado o essencial da cartografia e, não por acaso, o autor figura em nossa epígrafe. Neste curto e célebre texto intitulado “Do rigor na ciência”, Borges (1999a) nos fala de um mapa tão perfeito que chega a coincidir ponto a ponto com o próprio Império, tornando-se ele mesmo (o mapa) aquele que deveria “representar” (o território). A perfeição da cartografia, assim concebida, é também o seu malogro absoluto, a sua mais alta inutilidade, pois não compreende que todo mapa, para cumprir sua função, deve ser não um decalque fiel do seu território, mas um recorte arbitrário, seletivo, infiel.

A cartografia digital que identificamos nos serviços *Google* nos recoloca o problema das relações do mapa com o seu território. Sem os constrangimentos físico-geográficos enfrentados pelo Colégio dos Cartógrafos de Borges, o *Google* lida com

outros limites, próprios a um espaço sem bordas, movente e excessivo do ponto de vista informacional. Vimos que estes limites são contornados por um sutil deslocamento do foco da cartografia, o qual se volta para os percursos dos próprios indivíduos que navegam no ciberespaço. Tendo como carro-chefe seu mecanismo de busca, o *Google* consegue ser extremamente eficiente na orientação do usuário neste espaço, mesmo sem sequer chegar a cobrir 15% da “totalidade” da rede.

Qual é, pois, a natureza do problema da relação mapa-território aqui colocado? Trata-se do problema da circularidade ou recursividade do mapa constituído. Já apontamos esse problema ao mostrarmos o quanto a eficiência desta cartografia reside, de um lado, num contínuo monitoramento, classificação e ordenação dos caminhos e ações realizados pelos indivíduos e, de outro, no caráter performativo desse mapa que, ao antecipar para o usuário a ordem dos *sites*, documentos ou páginas mais ‘relevantes’ segundo o tema buscado, acaba por aumentar a probabilidade de que esses “caminhos preferidos” se tornem ainda mais pregnantes, a ponto de, no limite, acabarem valendo como os únicos caminhos efetivamente percorridos, os únicos lugares efetivamente visitados, enquanto tantos outros são e seriam possíveis. Aqui o mapa também é levado, no limite, a valer pelo território, não em termos físico-geográficos, mas sim em termos algorítmicos e performativos. Sim, pode-se advertir que esse é um risco inerente a qualquer mapa. Basta pensar nos mapas turísticos que traçam nas cidades percursos que se tornam praticamente necessários e obrigatórios. Mas, nesta cartografia de que estamos tratando, há algumas particularidades que tornam o problema mais complexo: primeiro, o fato de a navegação no espaço informacional ser praticamente impossível sem o auxílio dos mecanismos de busca, tornando o mapa, neste caso, quase que a única via de acesso ao território, a ponto de os navegantes não verem muito além do mapa - o que é um modo de aquele passar a se sobrepor a este, embora jamais haja coincidência entre os dois.

A segunda particularidade, decorrente da anterior, consiste numa maior performatividade desses mapas, na medida em que eles acabam por interferir no território (que no caso são os trajetos e ações efetuados pelos indivíduos) a ponto de torná-lo à semelhança do mapa. Para esclarecer, imaginem que se consiga burlar o sistema de buscas do *Google* na pesquisa da palavra “binóculo” e que apareça no topo

da lista um *site* de vendas de binóculo que não seja de fato o mais procurado do gênero<sup>13</sup>. Ora, não é certo, mas é bastante provável que ele se torne efetivamente o *site* mais procurado da busca de binóculos. Por fim, nesta cartografia, o teste ou a prova de realidade são mais difíceis ou mesmo sem sentido, pois como já vimos, eles não atendem tanto a critérios de verdade ou falsidade, mas sim a efeitos proativos ou performativos. Se um mapa geográfico convencional não for de fato adequado ao seu território, não é difícil detectar o “erro”, pois sempre podemos confrontá-lo com o seu referente, que se mantém relativamente estável para além do mapa. Mas num território em constante movimento e poroso ao seu próprio mapa, temos uma recursividade que torna quase indiscerníveis esses dois termos.

Consideremos um último exemplo, que é uma espécie de mapa de todos os mapas do mecanismo de buscas *Google*. Trata-se do *Zeitgeist*, um *site* que pretende revelar o “espírito da época”<sup>14</sup> na Internet, fornecendo uma espécie de mapa-ranking no qual se visualiza os temas e *sites* mais procurados na Rede em determinada semana, mês ou ano, em diversos países do mundo. Se consultarmos o *Zeitgeist* hoje, no dia em escrevemos este texto, dia 23 de maio de 2006, descobrimos que a palavra mais procurada no Brasil no mês de março foi “rebelde”, fruto do sucesso de uma novela mexicana adolescente emitida pelo SBT, e que esta supera em muito a palavra “carnaval”, nona colocada de uma lista de quinze. A página relativa aos Estados Unidos é mais completa e atualizada, e no mapa desta semana Maggie Q, atriz de *Missão Impossível 3* figura no topo do *ranking*, vindo o filme em segundo lugar. Também podemos saber que o país mais procurado pelos americanos em março de 2006 foi o Irã, enquanto no *Zeitgeist* da Alemanha vemos que a *Wikipédia* ([www.wikipedia.org](http://www.wikipedia.org)) figura entre os *sites* mais procurados no mesmo mês. O *Google Zeitgeist* pretende revelar, assim, o “clima intelectual, moral e cultural de uma época” ([www.google.com/press/zeitgeist](http://www.google.com/press/zeitgeist)). Nas palavras de John Batelle, “o *Zeitgeist* revelou ... que a *Google* tinha mais do que um dedo no pulso da nossa cultura; ela estava

---

<sup>13</sup> Circulam na Internet uma série de estratégias e truques para que um *site* seja listado numa boa colocação pelo mecanismo de buscas do *Google*. Estratégias para ampliar o número de conexões de uma página e para tratar o conteúdo do texto-âncora de um *site*, são alguns exemplos da atenção dada à importância do *Google* na ordenação e “popularidade” das informações disponíveis na Internet, constituindo até um novo conteúdo de disciplina de marketing – o *search marketing*.

<sup>14</sup> Tradução do próprio termo alemão *Zeitgeist*.

conectada diretamente com o seu sistema nervoso ... Meu Deus, pensei, a *Google* sabe o que a nossa cultura quer!” (op. cit., p. 2).

O autor tem, o nosso ver, parcialmente razão. É certo que o *Google* vem construindo um tesouro de dados com diversos usos possíveis e ainda indefinidos. Não é difícil imaginar os tantos sonhos do marketing, da publicidade e do comércio que podem ser vislumbrados – agências de viagens que “saibam” para onde os clientes de diversos países desejam ir, editoras que “saibam” o que os leitores desejam saber e conhecer, indústrias de entretenimento que “saibam” como as pessoas preferem se distrair e se divertir. Contudo, é impreciso afirmar que o *Google*, ou qualquer uma dessas empresas imaginárias, “saiba” de fato o que a nossa cultura quer. E aqui retomamos a questão nos termos que nos interessam – esta cartografia certamente produz e contém um saber sobre interesses e preferências de inúmeros usuários que navegam na Rede, fornecendo, sem dúvida, um mapa da nossa cultura ou, ao menos, da cibercultura. Mas esse “saber”, é preciso advertir, não é da ordem da verdade, como já afirmamos tantas vezes, e sim um conhecimento de cunho performativo e oracular: este mesmo tempo expressa uma certa realidade e intervém sobre ela, reforçando o que revela, antecipa ou faz ver dessa realidade. É bastante plausível supor que os visitantes do *Zeitgeist* acabem por tornar esse mapa ainda mais efetivo, clicando nos *sites* e buscando os temas apontados como o de maior interesse comum. Mapa e território mais uma vez se confundem.

Ora, todo esse problema da relação ente o mapa e o território é atravessado por uma outra questão não menos crucial – a questão do estabelecimento de todo um sistema classificatório ou de um regime de ordenação do mundo ou de uma dada realidade. Ainda uma vez, Borges (1999b) nos fala de uma enciclopédia chinesa intitulada *Empório Celestial de Conhecimentos Benévolos*, cuja engenhosa classificação dos animais consistia em: a) pertencentes ao Imperador, b) embalsamados, c) amestrados, d) leitões, e) sereias, f) fabulosos, g) cães soltos, h) incluídos nessa classificação, i) que se agitam como loucos, j) inumeráveis, (...) n) que de longe parecem moscas”. Essa classificação nos provoca o “riso que perturba todas as familiaridades do pensamento” (Foucault, 1990, p. 12), pois ela nos indica o limite do nosso próprio pensamento, a nossa impossibilidade de pensá-la. A inquietação provocada por esta ordem impensável faz vir à tona a suspeita tácita de que os

critérios com que ordenamos as coisas não lhes pertencem (Vaz, 1997), assim como, e principalmente, nos força a pensar na multiplicidade de formas possíveis de ordenação (Cf. Foucault, op. cit.).

Ora, o *Google* e suas cartografias põem em obra um grande regime de classificação e ordenação, talvez o mais significativo da última década, da “nossa idade e nossa geografia” (Idem, p. 13). Não olhamos seus mapas e indexações sem algum embaraço ou mal-estar, mas aqui estes provêm do perigo de esquecermos a arbitrariedade e contingência das ordens montadas nos bancos de dados e perfis computacionais. Como nos lembra Manovich (s/d), o banco de dados deve ser pensado não simplesmente como invento técnico, mas como forma cultural, como modelo de mundo. Sempre podemos pensar que cada cultura tem a taxonomia que merece, mas também podemos forçar os limites de nossas cartografias, em favor de outras ordens possíveis. Ainda Borges, sobre a Enciclopédia: “notoriamente, não há classificação do universo que não seja arbitrária e conjectural. A razão é muito simples: não sabemos o que é o universo” (1999b, p. 76).

## Referências Bibliográficas

- Arvidsson, A. “On the ‘Pre-History of the panoptic sort’: mobility in market research” in *Surveillance and society*, 1 (4), 2004.
- Battelle, J. *A Busca*. Rio de Janeiro: Ed.Campus/Elsevier, 2005
- Berners-Lee, T. “Realising the full potential of the Web”. Acessado em 28/01/2006  
<http://www.w3.org/1998/02/Potential.html>.
- Borges, J.L. “Do rigor da ciência” in *Obras Completas* Porto Alegre:Globo,1999a  
\_\_\_\_\_ “O idioma analítico de John Wilkins”. In: *Obras Completas* Porto Alegre:Globo,1999b
- Bruno, F. “Dispositivos de vigilância no ciberespaço: Duplos Digitais e identidades simuladas”. In: *Anais da XV Compós*, 2006

- Dodge, M. & Kitchin, R. *Mapping Cyberspace*. Londres: Roudledge, 2001.
- Dholakia, N. e Zwick, D. “Privacy and Consumer Agency in the Information Age: Between Prying Profilers and Preening Webcams” in *JRConsumers.com*, 1, 2001.
- Foucault, M. *As Palavras e as coisas: Uma Arqueologia das Ciências Humanas*. São Paulo: Martins Fontes, 1990
- Manovich, L. “Database as a symbolic form”. Acessado em 16/09/2005  
<http://hipercubo.uniandes.edu.co/redes03/pdf/manovich/database.pdf>
- Marx, G. T. “What’s new about the ‘new surveillance’? Classifying for change and continuity” in *Surveillance & Society* 1 (1), 2002
- Poster, M. *The second media age*. Cambridge: Polity Press, 1995
- Vaz, P.. Esperança e Excesso. In: André Parente (Org.) *Tramas da Rede*. 1ed.. Porto Alegre: Sulina
- \_\_\_\_\_ *O inconsciente Artificial*. São Paulo: Unimarco, 1997