

O surgimento da nova camada complexa da Web e a apropriação doméstica das tecnologias digitais conectadas

The emergence of Web new complex layer and domestic appropriation of connected digital technologies

Walter Teixeira Lima Junior

Docente do Programa de Pós-graduação em Comunicação Social da Universidade Metodista de São Paulo (UMESP) – São Bernardo do Campo (SP), Brasil. E-mail: digital@walterlima.jor.br

Resumo

O atual estágio tecnológico da Web permite todos os tipos de apropriações, desde as mais domésticas, como abertura de uma conta em uma rede social, até as mais sofisticadas, como a procura de rastros e dados. Desde a introdução da dinamicidade na Web, quando os componentes das páginas começaram a ser armazenados em bancos de dados, a coevolução da rede permitiu, a cada nível de apropriação pelo ator, determinar qual o patamar de obtenção de seu capital social cognitivo. Para diminuir a assimetria informativa e contribuir para o aumento informativo da esfera pública interconectada, o ator deve apropriar-se com mais profundidade das técnicas e tecnologias da Web.

Palavras-chave: web; tecnologia; capital social.

Abstract

The current technological state of the Web allows all types of appropriations, from the more domestic, as opening an account on a social network, to the most sophisticated one, such as looking for traces and data. Since the introduction of Web dynamics, when the components of the pages began to be stored in databases, the co-evolution of network enabled, at each level of ownership by the actor, to determine his/her level of cognitive social capital acquisition. In order to reduce information asymmetry and contribute to the increase of information of interconnected public sphere, the actor must appropriate more deeply the techniques and technologies of the Web.

Keywords: web; technology; social capital.

O surgimento da nova camada complexa da Web e a apropriação doméstica das tecnologias digitais conectadas

1. Introdução

O poder de conexão entre as pessoas por intermédio da Internet tem aumentado ao longo do tempo. Nos seus primórdios, a rede foi projetada para estabelecer transferência perfeita de dados entre os múltiplos pontos de acesso. Na atualidade, ao ultrapassar a barreira dos bilhões de usuários¹, a Internet revela que atende todos os tipos de requisições, apropriações e desejos dentro das configurações tecnológicas embutidas no modelo camada denominado *Open Systems Interconnection (OSI)*² embutido no protocolo *Transmission Control Protocol/Internet Protocol (TCP/IP)*³

O seu poder também advém do aumento do tráfego de dados, tanto na melhora da velocidade de transmissão, como no barateamento dos serviços que acessam a rede e por causa do desenvolvimento de tecnologias importantes que ampliaram as formas de apropriação dos conteúdos e dos sistemas disponíveis na Internet.

Dentre elas, a tecnologia de acesso a páginas na Web de forma dinâmica⁴, considerada a mais importante para

a mudança da velocidade no fluxo informativo na rede. Foi a dinamicidade das páginas que tornou a inserção e o consumo de informação mais velozes na Web.

Quando a primeira Web começou, havia somente páginas estáticas em HTML. A Internet já funcionava há algum tempo, mas foi somente com a introdução do HTML (com isso, surgiram os navegadores) que aquilo que chamamos de *World-Wide Web* começou. Muito tem acontecido desde então.

Coloca-se o nascimento da Web dinâmica quando a tecnologia *Common Gateway Interface (CGI)* foi introduzida em 1993. CGI foi a forma de permitir que um website “rodasse” *scripts* (usualmente *scripts* na linguagem Perl). Agora a Web tem se desenvolvido tremendamente e está poderosa devido aos serviços configurados com tecnologias e linguagens tais como PHP, Java, APS.NET, Python, Ruby (com Ruby on Rails) e muitos outros.⁵

Na Figura 1, nota-se a evolução das páginas estáticas para dinâmicas em uma linha do tempo. É bom ressaltar que a Internet, na atualidade, é composta tanto por páginas estáticas como por dinâmicas.

A dinamicidade permite também que conteúdos sejam apagados ou escondidos dos sistemas de busca na Web com a mesma velocidade em que eles são inseridos.

2. De volta para o futuro

Se, no seu início, a Internet só era acessada e manipulada por especialistas, pois era necessário conhecer comandos digitados através de teclados para serem observados em

diretamente de um arquivo já existente no servidor, ou dinâmica, se for criada dinamicamente por outro programa, *script* ou *Application Programming Interface (API)* chamado pelo servidor. No caso de uma página dinâmica, o pedido, depois de recebido, é processado pelo servidor web que vai criar dinamicamente o conteúdo que depois será enviado para o cliente. As páginas dinâmicas têm a vantagem de poderem ser programadas, ou seja, usando alguma linguagem de programação (que, dependendo do servidor web, pode ser php, Java, Perl, Visual Basic, NET, C# etc.); podemos criar programas que correm no servidor web eventualmente acessando bases de dados e cujo resultado é enviado para o *browser*. Disponível em: <http://pt.wikipedia.org/wiki/Servidor_web>. Acesso em: 03 nov 2012.

⁵ *A history of the dynamic web*. Pingdom. Disponível em: <<http://royal.pingdom.com/2007/12/07/a-history-of-the-dynamic-web/>>. Acesso em: 7 nov 2012.

¹ *Internet atinge 2,1 bilhões de usuários no mundo em 2011, aponta consultoria*. UOL Tecnologia, 18 de janeiro de 2012. Disponível em: <<http://tecnologia.uol.com.br/ultimas-noticias/redacao/2012/01/18/internet-atinge-21-bilhoes-de-usuarios-no-mundo-em-2011-apon-ta-consultoria.jhtm>>. Acesso em: 7 de novembro de 2012.

² O Modelo OSI permite comunicação entre máquinas heterogêneas e define diretivas genéricas para a construção de redes de computadores (seja de curta, média ou longa distância) independentemente da tecnologia utilizada. Essa arquitetura é um modelo que divide as redes de computadores em sete camadas, de forma a se obterem camadas de abstração. Cada protocolo implementa uma funcionalidade assinalada a uma determinada camada. Disponível em: <http://pt.wikipedia.org/wiki/Modelo_OSI>. Acesso em: 03 nov 2012.

³ O TCP/IP é um conjunto de protocolos de comunicação entre computadores em rede (também chamado de pilha de protocolos TCP/IP). Seu nome vem de dois protocolos: o TCP (*Transmission Control Protocol – Protocolo de Controle de Transmissão*) e o IP (*Internet Protocol – Protocolo de Interconexão*). O conjunto de protocolos pode ser visto como um modelo de camadas, onde cada camada é responsável por um grupo de tarefas, fornecendo um conjunto de serviços bem definidos para o protocolo da camada superior. As camadas mais altas estão logicamente mais perto do usuário (chamada camada de aplicação) e lidam com dados mais abstratos, confiando em protocolos de camadas mais baixas para tarefas de menor nível de abstração. Disponível em: <<http://pt.wikipedia.org/wiki/TCP/IP>>. Acesso em: 03 nov 2012

⁴ A origem do conteúdo enviado pelo servidor web numa resposta a um pedido HTTP pode ser de forma estática, se vier

O surgimento da nova camada complexa da Web e a apropriação doméstica das tecnologias digitais conectadas



Fonte: PINGDOM (2013).

Figura 1: A história da Web dinâmica.

monitores monocromáticos de fósforo, a rede evoluiu e passou para o modo gráfico, por meio do WWW, criado pelo cientista britânico Tim Berners-Lee. Desde então, o cenário de acesso modificou-se bastante para o ator. O conceito de ator foi escolhido, neste trabalho, para designar o ser humano que realiza a interação com a informação digital conectada, através de redes telemáticas.

Segundo Fidel (2012), o termo “ator” tem várias vantagens na pesquisa sobre interação homem e informação digital conectada e é diferente do conceito de usuário:

Centrado mais no participante do que no sistema, quando os atores possuem uma existência independente dos sistemas de informação; coloca interação com participação como parte das atividades de uma pessoa, ao invés de um processo isolado; não cria divisão entre usuários atuais de um sistema e não-usuários, portanto concentra a investigação em todos os participantes que interagem com a informação e podem se beneficiar de um sistema de informação. Essa aproximação é particularmente importante nos casos onde não-usuários evitam o sistema porque não se encaixam nas

necessidades deles. Além disso, para identificar tais casos requer um estudo dos usuários e não-usuários também, isto é, o estudo dos atores (FIDEL, 2012, p. 4).

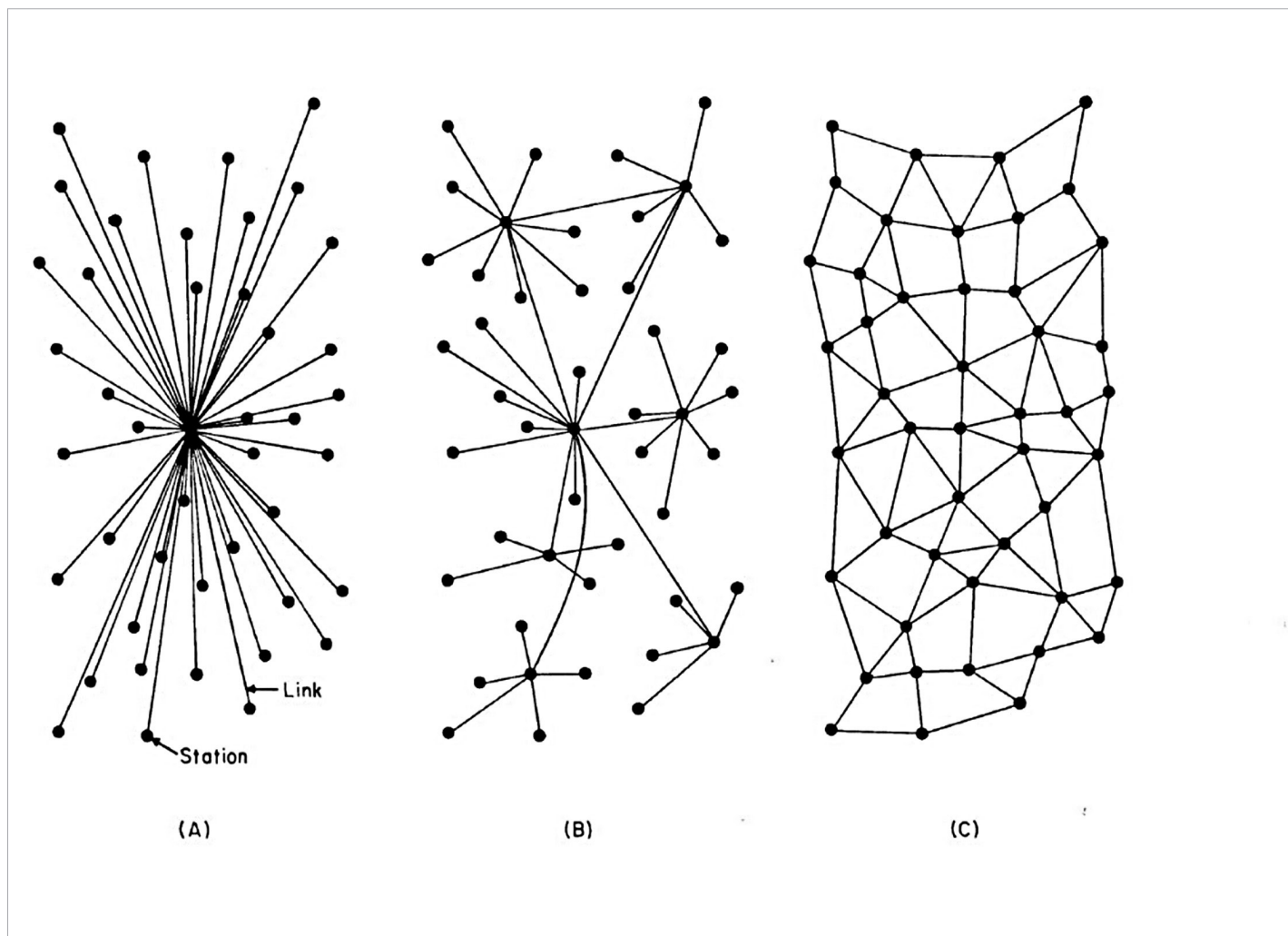
Assim, o ator se apropria das tecnologias digitais conectadas, que fornecem crescente variedade de acesso a serviços de informação em diversas categorias, seja na área de conteúdo de relevância social ou de entretenimento. O ator, dependendo do seu poder aquisitivo, atualmente, tem à disposição uma gama de dispositivos computacionais fixos/portáteis (*PC, smartphones, tablets, e-readers, console de games*) e sistemas (*web e software*). Os dispositivos de acesso combinados com infinitos tipos de *software* (proprietários ou livres) permitem diferentes níveis de apropriação tecnológica da Web, pois a arquitetura da rede idealizada por Paul Baran (Figura 2), em 1964, tornou a Internet uma rede descentralizada, distribuída e de baixa hierarquia.

Para que a topologia funcionasse, Vinton Cerf e Bob Kahn criam o protocolo TCP/IP e o implementaram com sucesso em 1969. Em 1989, com a chegada da camada amigável do WWW, de Berners-Lee, a Internet entra na fase mais amigável, apesar de os conteúdos ainda serem transferidos e manipulados nos servidores por especialistas. Mas, nessa fase, a compreensão tecnológica do funcionamento da Web no seu modo gráfico facilitou a inserção de conteúdos na Web por não engenheiros. Bastava conhecer o código HTML e subir os arquivos para o servidor através do protocolo *File Transfer Protocol* (FTP). Assim, os objetivos propostos originariamente por Tim Berners-Lee foram alcançados:

Quando eu propus a Web em 1989, o pensamento que tinha era a comunicação através do conhecimento compartilhado e a direção de “mercado”, para que isso acontecesse, a colaboração entre pessoas no trabalho e em casa. Para construir uma Web de hipertexto, um grupo de pessoas, seja de que tamanho for, poderia se expressar facilmente, rapidamente adquirir e transmitir conhecimento, superando as incompreensões e reduzindo a duplicação de esforços. Isso daria às pessoas, num grupo, um novo poder para construir alguma coisa juntos (BERNERS-LEE & FISCHETTI, 1999, p. 162).

Portanto, a facilidade de acesso e inserção de conteúdos, além do principal, a colaboração entre os atores, foram os claros objetivos de Berners-Lee, que coloca em xeque os conceitos de Web 2.0⁶ ou Web colaborativa. A Web já nasceu colaborativa.

⁶ C.f. Web 2.0. Disponível em: <http://pt.wikipedia.org/wiki/Web_2.0>. Acesso em: 07 nov 2012.

**O surgimento da nova camada complexa da Web e a
apropriação doméstica das tecnologias digitais conectadas**

Legenda: (A): Centralizado; (B): Descentralizado; (C): Distribuído.

Fonte: Baran (1964).

Figura 2: Diagrama de Paul Baran.

A arquitetura de rede descentralizada e distribuída de Baran, o protocolo de Cerf e Kahn e a Web de Berners-Lee pertencem a um processo de coevolução que permite que a rede, na atualidade, ofereça sistemas extremamente amigáveis, que permitem ao ator, em três ou quatro cliques, subir qualquer tipo de conteúdo e disponibilizá-lo na Web. *Blogs*, *microblogs*, redes sociais, *Content Manager System* (CMS), entre outros sistemas, são utilizados pela grande parte dos atores que usufruem das conexões da Internet.

Assim, existem atores que ainda sobem os seus conteúdos em HTML e via FTP; outros, via Web (*blogs*, *microblogs*, redes sociais, CMS). Entretanto, há outras técnicas

de extração, armazenamento e visualização de dados e informações que não fazem parte das duas formas citadas acima.

Depois da introdução da tecnologia de banco de dados na estrutura do gerenciamento de páginas dinâmicas na Web, percebeu-se o potencial da tecnologia para que a rede armazenasse nos seus servidores conteúdos que poderiam ser manuseados através da Web, sem a necessidade de o ator instalar um *client*, ou seja, armazenar conteúdos na sua própria máquina, para depois subi-los para a rede. Esse serviço de acesso à Web e manipulação de dados sem utilizar um *client*, hoje, é conhecido como Computação em Nuvem (*Cloud Computing*). Também deriva de tal possibilidade

O surgimento da nova camada complexa da Web e a apropriação doméstica das tecnologias digitais conectadas

tecnológica, proporcionada pela utilização de banco de dados e linguagens como XHTML⁷, por exemplo, a extração e relacionamento de dados através de vários conjuntos de dados (*datasets*) e a visualização de informações não triviais via infográficos.

Assim, temos ao mesmo tempo, na Web, diversos tipos de apropriação tecnológica pelos atores. A Web permite, em poucos minutos, que um analfabeto digital crie uma conta em uma rede social em minutos e, também, que *hackers* vasculhem as “profundezas” da Web procurando rastros e dados. São níveis diferentes de apropriação tecnológica. Segundo Kent Norman (2008, p. 234):

Geralmente, para os propósitos de *design* de interface e pesquisa, tendemos a categorizar usuários dentro das seguintes classes de experiência em computador:

- Novato: usuário novo, não familiarizado com a interface, programa e serviços;
- Usuário intermitente: um pouco familiar com a interface, programa e serviços, mas usa com pouca frequência;
- Usuário casual: familiar com a interface, mas não tanto como um usuário frequente, *full-time*;
- Usuário expert (nível 1): usuário treinado com muita experiência; e
- Usuário expert (nível 2): extremamente bem treinado, com conhecimento interno da interface, programa e serviços.

Somente testes científicos controlados podem classificar as categorias em que grupos de atores se encaixam. Entretanto, é notório, pelos números divulgados, que grande parte dos dois bilhões de usuários da Internet se apropriam da rede para fins domésticos⁸ (JORNAL DA GLOBO, 2012). O termo “doméstico” é aqui definido como acesso aos sistemas informativos da Web sem o aprofundamento sobre as possibilidades que eles proporcionam para a obtenção de informação estruturada, ou seja, que

possua valor agregado em auxílio à tomada de decisão, tanto no campo pessoal como no profissional.

Assim, sistemas que possuem configurações de redes sociais têm forte apelo e audiência (medida em frequência de acesso e tempo de permanência). Contudo, são sistemas que estão na Internet através da Web, mas não permitem a utilização de todas as características da rede, como ser descentralizado e distribuído. São sistemas que aprisionam o ator e o induzem, de diversas formas, para que interaja com o sistema conforme os objetivos propostos pelo modelo de negócio que molda a rede social.

Um *software* expressa ideias sobre tudo, desde a natureza de uma nota musical até o que constitui uma pessoa. Um *software* também está sujeito a um processo excepcionalmente rigoroso de “aprisionamento tecnológico”. Dessa forma, as ideias (na era atual, quando as questões humanas são cada vez mais orientadas por *software*) se tornaram mais um objeto de aprisionamento tecnológico do que em eras anteriores. A maioria das ideias que foram sujeitas ao aprisionamento tecnológico até agora não foram tão ruins, mas algumas ideias da chamada Web 2.0 são desprezíveis, de forma que precisamos rejeitá-las enquanto ainda podemos (LANIER, 2010, p. 17).

Jaron Lanier é pioneiro, nos anos 1970, na área de Realidade Virtual. Conhecedor das tecnologias que constituem a Internet, Lanier, em seu livro *Gadget: você não é um aplicativo*, não poupa críticas aos que acreditam na noosfera e aos sistemas que formam silos de informação. Aliás, o termo “silo” foi mencionado por Tim Berners-Lee para designar sites de relacionamento que aprisionam conteúdo:

Várias ameaças à universalidade da web surgiram recentemente. As empresas de TV a cabo, que também vendem conexão à internet, estão pensando em restringir os downloads apenas ao seu próprio mix de entretenimento. As redes sociais apresentam outro tipo de problema: *Facebook*, *LinkedIn*, *Friendster* e outros sites de relacionamento em geral adicionam valor via informações capturadas quando você entra nos sites — o dia do seu aniversário, seu endereço de e-mail, seus gostos e seus amigos (revelados por links e fotos). Essas redes reúnem os dados em brilhantes *databases* e utilizam as informações para fornecer serviços de valor agregado — mas apenas dentro dos próprios sites. Se você entrar com seus dados em um desses serviços, não poderá usá-los facilmente em outro site. Cada um é um silo isolado

⁷ *Extensible Hypertext Markup Language*.

⁸ *Facebook atinge marca de um bilhão de contas ativas. Jornal da Globo*, 4 de outubro de 2012. Disponível em: <<http://g1.globo.com/jornal-da-globo/noticia/2012/10/facebook-atinge-marca-de-um-bilhao-de-contas-ativas.html>>. Acesso em: 07 de novembro de 2012.

O surgimento da nova camada complexa da Web e a apropriação doméstica das tecnologias digitais conectadas

dos outros. Sim, as páginas do seu site estão na web, mas não os seus dados. Você pode acessar uma página com uma lista de pessoas que criou em um site, mas não pode enviar essa lista, ou itens dela, para outro site.⁹

Para Berners-Lee, o isolamento proposital escrito no código do sistema acontece porque cada fragmento da informação não possui um Identificador Uniforme de Recursos (URI). Um URI pode ser classificado como um localizador (URL) ou nome (*Uniform Resource Name* – URN), ou ainda como ambos. Ou seja, segundo o cientista britânico, as conexões entre dados existem apenas dentro do site de relacionamento e não podem ser “encontradas” pelos sistemas de busca, como o Google.

Esse isolamento acontece porque cada fragmento de informação não tem um URI. Conexões entre dados existem apenas dentro de um site; quanto mais dados você fornece, mais fica “trancado” dentro dele. Sua rede social torna-se assim uma plataforma central — um silo fechado de conteúdo, que não lhe dá pleno controle sobre suas próprias informações. À medida que esse tipo de arquitetura se generaliza, a web torna-se cada vez mais fragmentada.¹⁰

Nesse panorama tecnológico, Lanier (2010) faz um alerta aos atores: que eles resistam aos caminhos fáceis que os orientam. “Se você adorar um veículo feito de *software*, corre o perigo de ficar aprisionado nos recentes pensamentos descuidados de alguém. Lute contra isso!” (LANIER, 2010, p. 39).

O sistema que é engendrado para atrair milhares de centenas de atores, ao mesmo tempo, os aprisiona, fornecendo a sensação aos atores de que estão inseridos completamente nas conexões para obtenção de conteúdo estruturado e que estão participando de um sistema com livre trânsito informativo: a denominada “liberdade na Internet”. Lanier (2010) afirma que

a cultura on-line está cheia até a borda de uma retórica sobre qual deveria ser o caminho para um mundo melhor e, nos dias de hoje, esse discurso tende intensamente na direção de uma forma anti-humana de pensar (LANIER, 2010, p. 39).

Outro pesquisador, Eli Pariser (2011), aponta outras questões importantes devido à apropriação doméstica dos sistemas da Internet e a utilização de sites de relacionamento de forma compulsiva. No seu livro intitulado *Filter Bubble*, Pariser aponta que as informações que realmente importam não estão nesses espaços virtuais.

É um desafio: como mais e mais pessoas descobrem notícias e conteúdos através do abastecimento personalizado como o Facebook, as informações que realmente importam ficam de fora. Em um ambiente darwiniano de abastecimento de notícias hiper-relevantes, conteúdos sobre questões como os sem casa ou a mudança no clima não competem com vídeos virais bobos, notícias sobre celebridades e gatinhos. A esfera pública sai de cena. E isto importa, porque enquanto nós podemos perder a visão dos nossos problemas comuns, eles não perdem a visão de nós.¹¹

Mas Pariser (2011) não faz somente críticas; ele aponta caminhos no sentido de que o ator adquira mais conhecimentos sobre a rede e que possa, com isso, “furar a bolha”, que, na visão do pesquisador, filtra as informações relevantes para a discussão de temas importantes para a sociedade. Portanto, temas também fundamentais para o ator que convive em sociedade.

Em outras palavras, é tornar-se mais importante, desenvolvendo conhecimento básico de algoritmos. Progressivamente, os cidadãos poderão julgar os sistemas programados que afetam a nossa vida pública e a vida nacional. E mesmo se não são fluentes o bastante para ler através de milhares de linhas de códigos, loops e memória, esse conhecimento pode colocar luz sobre como estes sistemas funcionam e onde eles podem conter erros (PARISER, 2011, p. 228).

Portanto, Berners-Lee, Lanier e Pariser indicam que os sistemas que estruturam tecnologias de aprisionamento de conteúdo, principalmente redes sociais como o Facebook, são perniciosos tanto à Web quanto aos atores que utilizam esses sistemas de forma doméstica. Acrescenta-se que o ambiente digital conectado fornecido é um atrativo poderoso devido à facilidade de uso e identificação imediata com os desejos dos atores por conversação.

Assim, o ambiente é o contexto digital ideal para que os atores permaneçam (audiência) de forma amigável na sua relação com a tecnologia, sem necessitarem de grandes esforços de

⁹ BERNERS-LEE, Tim. *Longa Vida à Web. Scientific American Brasil*. Disponível em: <http://www2.uol.com.br/sciam/reportagens/vida_longa_a_web.html>. Acesso em: 07 de novembro de 2011.

¹⁰ BERNERS-LEE, Tim. *Longa Vida à Web. Scientific American Brasil*. Disponível em: <http://www2.uol.com.br/sciam/reportagens/vida_longa_a_web.html>. Acesso em: 07 de novembro de 2011.

¹¹ PARISER, Eli. *Filter bubbles, meet Upworthy*, março de 2012. Disponível em: <<http://www.thefilterbubble.com/filter-bubbles-meet-upworthy/>>. Acesso em: 03 de novembro de 2012.

O surgimento da nova camada complexa da Web e a apropriação doméstica das tecnologias digitais conectadas

aprendizagem. A mudança de comportamento só se dará pela mudança do ambiente digital, que deverá possuir serviços e/ou tecnologias que permitam ao ator construir o seu próprio caminho, produzir e compartilhar informação da forma que desejar. Ou seja, o contexto (ambiente digital) é importante para a motivação pela pesquisa de informação relevante.

Pesquisadores têm salientado que a noção de contexto no qual o ator é localizado contribui para a moldagem de uma informação necessária, porque a motivação pela pesquisa de informação não é somente cognitiva, mas também contextual (FIDEL, 2012, p. 85).

Outro fator importante é que o ator, para se apropriar potencialmente da rede, deve possuir habilidades não requisitadas quando atua de forma doméstica. Somente parcela ínfima dos atores na rede possuem tais requisitos, mas que são importantes para a melhora de desempenho na apropriação mais profunda da rede.

Citando as múltiplas inteligências de Gardner (1893, 1999), quando lista oito tipos de inteligência, destaca-se uma primordial para o aprofundamento das lógicas computacionais, a inteligência lógico-matemática, que é a habilidade de trabalhar com números, abstrações e lógica. Essa inteligência favorece matemáticos e programadores de computadores. Para alcançar a lógica que a interface requer, a programando e a computando, contará para a performance (NORMAN, 2008, p. 230).

3. Web de dados

A relação do ator com a informação possui diversos níveis. Uma das áreas da ciência que estudam esse fenômeno é a *Human Information Interaction* (HII), que investiga como os humanos interagem com a informação. Para Fidel (2012, p. 51), a HII é “um fenômeno complexo e multifacetado”, que é influenciado em graus variados por questões tais como estados cognitivo e afetivo; organizacional, cultural e ambiental; aspectos sociais, tecnológicos e da natureza e estrutura das fontes de informação.

Assim, como discutido anteriormente, a Internet e a sua camada amigável Web permitem diversos tipos de apropriações tecnológicas. As formas e níveis de uso determinam como os atores constroem o seu capital social cognitivo através das formas com que acessam e usam tecnologia na rede. Metodologicamente, é desafiador tentar mensurar o grau desse tipo de aquisição de informação e habilidades. Entretanto, a procura por “valor”, seja por informação estruturada, por informações construídas através de cruzamentos

de dados ou por processos de mineração dos mesmos, é uma das áreas que mais se desenvolve no campo da Ciência da Computação, tendo como setor a tecnologia da Internet.

Uma das formas de apropriação tecnológica, de alto nível pela complexidade inserida nos seus procedimentos técnicos, é a manipulação da Web de dados. Sem definição precisa, a Web de dados é um conjunto de tecnologias que são configuradas em função de bancos de dados (*datasets*), disponíveis na Internet, possibilitando cruzamento, relacionamento e a utilização de dados estruturados¹². Na Ciência da Computação, dado estruturado é uma particular forma de armazenar e organizar dados num servidor de Web e que pode ser alcançado e entendido por sistemas computacionais, sendo, portanto, usado eficientemente.

A análise de duas empresas de *Common Crawl*¹³ demonstra que a porcentagem de páginas na Web que contêm dados estruturados aumentou de 6% em 2010 para 12% em 2012 (MÜHLEISEN & BIZER, 2012). A quantidade de dados na Internet também tem aumentado. O relatório *McKinsey Global Institute* demonstra que o crescimento é exponencial:

Vários pesquisadores estudam a quantidade de dados gerados, guardados e consumidos no mundo. As estimativas deles variam muito, mas concordam que o crescimento é exponencial. MGI estima que empresas armazenam globalmente mais do que 7 *exabytes* de novos dados em disk drives em 2010, enquanto consumidores guardam mais do que 6 *exabytes* de novos dados em *PC's* e *notebooks*. Um *exabyte* de dado é equivalente a mais do que 4 mil vezes a informação guardada na Livraria do Congresso dos EUA (235 *terabytes*) (2011, p. 3)¹⁴.

Para encontrar informações relevantes, que não podem ser encontradas relacionando milhões de dados de forma manual, utilizam-se técnicas como *Data mining* e *Web data mining*. Liu (2011) conceitua *Data mining* como técnica para Descoberta de Conhecimento em Banco de Dados (DCB), que

¹² *Definição de dado estruturado*. Disponível em: <<http://xlinux.nist.gov/dads//HTML/datastructur.html>>. Acesso em: 08 de novembro de 2012.

¹³ *Common Crawl* é uma fundação sem fins lucrativos que coleta dados de páginas da Web usando programa com a tecnologia *Crawler* e publica esses dados.

¹⁴ *Bigdata: the next frontier for innovation, competition and productivity*. McKinsey Global Institute Report, May, 2001.

O surgimento da nova camada complexa da Web e a apropriação doméstica das tecnologias digitais conectadas

é comumente definida como processo de descoberta de padrões úteis ou Conhecimento de fonte de dados, como base de dados, textos, imagens, a Web etc. (LIU, 2011, p. 6).

Um atributo da técnica de *Data mining* é ser também utilizada via conexões da Internet, extraído de um conjunto de banco de dados (*datasets*) informações de “valor” agregado, não triviais. Essa técnica é denominada *Web mining*, aproveitando a grande quantidade de dados disponíveis, diversos e em amplo crescimento.

Web mining é a descoberta de informação útil ou conhecimento da estrutura de hyperlinks da Web, conteúdos de páginas e dados. Embora Web mining use muitas técnicas de data mining, web mining não é puramente uma aplicação das tradicionais técnicas de data mining devido à heterogeneidade e natureza semi-estruturada ou desestruturada da Web Data (LIU, 2011, p. 7).

A operação de tecnologias e técnicas no campo da Web de dados permite aos atores, com conhecimento avançado sobre a Internet, Web, estruturas de funcionamento de conexões e banco de dados obterem “valor” da informação ou conjunto de informações contidas na rede. Para Castells,

valor é o que é processado em toda rede dominante, em todo tempo, em todo espaço, de acordo com a hierarquia programada na rede pelos atores agindo sobre a rede (CASTELLS, 2009, p. 29).

Entretanto, tal “valor”, quando obtido por tecnologias e técnicas que envolvem a Web de dados, criam informação com “valor” agregado, construindo nova informação, portanto, estabelecendo assimetria da informação, que gera valor de monetização sobre ela. Castells (2009) insere a produção assimétrica da informação digital como uma forma de poder. Ele afirma que “assimetricamente significa que a influência durante um relacionamento é sempre recíproca, há sempre um grau maior de influência de um ator sobre outro” (CASTELLS, 2009, p. 11). Assim, se um ator possui informação de “valor” agregado, ele, no relacionamento informativo com outro ator que utiliza a Web de forma doméstica, constrói um patamar informativo diferenciado.

Mas por que a Internet e a Web permitem vários tipos de apropriação tecnológica e técnica? Porque a Internet, onde roda a Web, é considerada pelos pesquisadores da área um sistema complexo. Podemos considerar como sistema complexo sintético, pois foi criado pelo homem e não pela natureza. Ou seja, pelos processos evolutivos, a

natureza criou e aperfeiçoa sistemas complexos como a Floresta Amazônica, colônias de insetos, o cérebro, o sistema imunológico, entre outros. Podemos dizer também que o sistema econômico é um sistema complexo sintético. Desta forma, podemos categorizar a Internet/Web como um sistema complexo, pois é um sistema social auto-organizado, uma das principais características de um sistema complexo.

O tópico de sistemas complexos, um campo interdisciplinar de pesquisa que procura explicar como grande número de simples entidades se organiza, sem o benefício de qualquer controle central, para uma inteira coleção que cria padrões, usa informações, além de alguns casos, evolui e aprende (MITCHELL, 2009, p. 4).

Ainda para Mitchell (2009), entre as propriedades comuns de um sistema complexo estão o comportamento complexo coletivo, processamento de sinais e informações e capacidade de adaptação (mudança de comportamento). Ainda para a pesquisadora, o conceito abrange

um sistema no qual grandes redes de componentes sem controle central e com regras simples de operação dão origem a comportamento coletivo complexo, processamento de informação sofisticada e da adaptação via aprendizagem ou adaptação (MITCHELL, 2009, p. 13).

Para o pesquisador na área da Filosofia da Tecnologia, Mario Bunge, um sistema complexo

possui o conceito de ser um objeto complexo do qual toda parte ou todo componente está relacionado no mínimo com outro componente (BUNGE, 2006, p. 358).

Entretanto, no sistema complexo sintético que estrutura a Web, a grande parte dos atores se conecta e o utiliza de forma doméstica. Apenas os atores com o domínio mais aprofundado de *softwares* e/ou das ferramentas da Ciência da Computação conseguem extrair “valor”, em função da apropriação profunda tecnológica e técnica dos atributos pertencentes ao sistema complexo que estabelece a rede.

4. Considerações finais

Com o advento da Web de dados, mais um patamar de complexidade é anexado aos dispositivos e sistemas, fazendo surgir outro tipo de ator, o que é especializado em cruzar dados e informações utilizando diferentes sistemas

O surgimento da nova camada complexa da Web e a apropriação doméstica das tecnologias digitais conectadas

e plataformas. Assim, apesar de viverem à contemporaneidade tecnológica digital conectada, cada nível de apropriação da rede pelo ator determina qual o patamar de obtenção de capital social cognitivo.

A construção do capital social cognitivo do ator que está na Internet também se dá pela qualidade na apropriação dele da tecnologia da rede, em função de obter informação estruturada para tomada de decisão ótima (a melhor possível), seja no aspecto pessoal da vida, seja no campo profissional.

Assim, os atores devem fugir dos silos de informação e furar a bolha informativa e seus filtros. Portanto, não abastecendo com informações sistemas que aprisionam conteúdo. Devem aprender a manipular técnicas e tecnologias da Internet para extrair valor dos dados, diminuindo a

assimetria informativa e contribuindo para o aumento do nível informativo da esfera pública.

A forma de apropriação mais aprofundada das técnicas e tecnologias da Web, no início do WWW, foi decisiva para a formação da visão de que a rede é instrumento importante para aquisição de conhecimento (informação estruturada). Talvez, o atual estágio da interface gráfica amigável da Internet, com a Web de dados, forneça o ambiente digital conectado para que os atores mais qualificados tecnologicamente recolorem a Web no patamar informativo dos seus primórdios. Mas não se deve ter a perspectiva de que se voltará aos tempos quando os acadêmicos disponibilizavam os seus conteúdos. Por se configurar em um sistema complexo, vários tipos de apropriação conviverão, mas novas opções surgem para obtenção de informação relevante para os atores mais qualificados tecnicamente e tecnologicamente e, conseqüentemente, para a sociedade.

Referências

BARAN, Paul. *On Distributed Communications*: I. Introduction to distributed communication networks. Memorandum prepared for the United States Air Force Project RAND, August, 1964, p. 2.

BERNERS-LEE, Tim & FISCHETTI, Mark. *Weaving the Web: the original design and ultimate destiny of World Wide Web by its inventor*. New York: Haper Collins, 1999.

BUNGE, Mario. *Dicionário de Filosofia*. São Paulo: Perspectiva, 2006.

CASTELLS, Manuel. *Communication Power*. New York: Oxford Press University, 2009.

FIDEL, Raya. *Human Information Interaction: an ecological approach to information behavior*. Cambridge: The MIT Press, 2012.

LANIER, Jaron. *Gadget: você não é um aplicativo*. São Paulo: Saraiva, 2010.

LIU, Bing. *Web Data Mining: exploring hyperlinks, contents, and usage data*. New York: Springer, 2011.

MITCHELL, Melanie. *Complexity: a guide tour*. New York: Oxford, 2009.

MÜHLEISEN, Hannes & BIZER, Christian. *Web Data Commons: Extracting Structured Data from Two Large Web Corpora*. Paper apresentado no Workshop *Linked Data on the Web (LDOW2012)*, 2012. Disponível em: <<http://events.linkedata.org/ldow2012/papers/ldow2012-inv-paper-2.pdf>>. Acesso em: 8 nov 2012.

NORMAN, Kent L. *Cyberpsychology: an introduction to human-computer interaction*. New York: Cambridge University Press, 2008.

PARISER, Eli. *The Filter Bubble*. New York: The Pinguim Press, 2011.

PINGDOM. *A history of the dynamic web*. Disponível em: <<http://royal.pingdom.com/2007/12/07/a-history-of-the-dynamic-web/>>. Acesso em: 04 set 2013.