

# A cena cibercultural do jornalismo contemporâneo: *web* semântica, algoritmos, aplicativos e curadoria

*The cybercultural scenario of contemporary journalism: web semantics, algorithms, applicatives and curatorship*

ELIZABETH SAAD CORRÊA \*

DANIELA BERTOCCHI \*\*

## RESUMO

Este trabalho discute os diferentes aspectos ciberculturais que o jornalismo contemporâneo vivencia, sobretudo a partir da emergência da *web* semântica e das funcionalidades técnicas dela decorrentes – como o uso de algoritmos e aplicativos. Explicamos os conceitos básicos deste conjunto técnico e sua correlação aos campos da Comunicação e do Jornalismo em particular. Apresentamos as possibilidades de aproximação de tal cenário com a atividade jornalística a partir de exemplos já em aplicação em marcas jornalísticas internacionais. Propomos a configuração do papel de curadoria para o profissional que irá atuar neste novo contexto cibercultural.

**Palavras-chave:** cibercultura, jornalismo digital, *web* semântica, algoritmos, curadoria

## ABSTRACT

The present article debates the different cybercultural aspects experienced by contemporary journalism, especially after the emergence of web semantics and technical resources along with it – such as the use of algorithms and applicatives. We present the basic notions involved in this technical conjunction and also the relationship they establish to the fields of Communication in general and of Journalism in particular. Considering examples in use by international journalistic brands we indicate possibilities for gathering journalistic activity in such scenario. We then establish a proposal for the role of curatorship to the professional willing to enter this new cybercultural context.

**Keywords:** cyberculture, digital journalism, web semantics, algorithms, curatorship

\* Professora Titular do Departamento de Jornalismo e Editoração da Escola de Comunicações e Artes da Universidade de São Paulo e do PPGCom-USP. Coordenadora do curso de especialização DIGICORP e do grupo de pesquisa COM+. E-mail: bethsaad@gmail.com

\*\* Doutoranda do Programa de Pós-Graduação em Ciências da Comunicação da Universidade de São Paulo. Mestre em Ciberjornalismo pela Universidade do Minho, Portugal. Pesquisadora do grupo de pesquisa COM+. E-mail: bertocchi@usp.br

## INTRODUÇÃO

INTERESSA-NOS AQUI PENSAR qual comunicação temos quando máquinas também comunicam, especialmente no contexto da informação jornalística. Essa inquietação nasce em decorrência dos recentes e cada vez mais presentes debates em torno do desenvolvimento da *Web Semântica* (WS), uma das mais conhecidas previsões de Tim Berners-Lee (Bourenane, Szoniecky & Saleh, 2009) e que agora começa a alcançar seu estágio de consolidação (Adolphs, *et al.*, 2010). A partir da WS, também nos interessa investigar as discussões em torno do uso de algoritmos e aplicativos (*apps*) no processo jornalístico, exploração que se revela como tendência irreversível nas principais redações de destaque na *web* (Rabaino, 2011 e Carr, 2011).

Considerando que a WS, *apps*, algoritmos, bancos de dados, entre outros, são manifestações ciberculturais contemporâneas e alteram formas de sociabilidade, nos vemos, portanto, diante de mais um campo de discussões e reflexões, para não dizer de mudanças prementes. É nesse sentido que iremos desenvolver nossa argumentação.

A primeira parte desse artigo explica o *status* da *Web Semântica* como um processo amplo. Ao defini-la, privilegiamos o ponto de vista da comunicação, da linguagem e da cognição em detrimento da visão técnica sobre o tema. Acreditamos que iniciar o trabalho apresentando a WS com esse enquadramento se faz necessário por pelo menos dois motivos. O primeiro: a produção de uma literatura específica sobre a WS tem sido prolífera nos campos das Ciências da Informação, da Computação e também nas Engenharias, e frequentemente em Língua Inglesa, porém escassa no campo da Comunicação. Consideramos importante, portanto, trazê-la para próximo dos debates relativos à comunicação digital e à cibercultura. O segundo motivo decorre do primeiro: o conceito de WS ainda está nas primeiras discussões entre pesquisadores e profissionais da Comunicação, o que nos obriga a primeiramente explicar a sua lógica de funcionamento antes de nos aprofundarmos em especificidades. Também nessa primeira parte, faremos a correlação da WS com algoritmos, *apps* entre outras formas contemporâneas de comunicação.

Na segunda parte desse texto, por outro lado, discutimos o cenário em que a máquina compreende dados e, num passo seguinte, os correlaciona e os interpreta, apontando especificamente como tal cenário se configura em ambientes jornalísticos.

Na terceira e última parte refletimos, então, sobre o papel do jornalista enquanto comunicador e do próprio campo da Comunicação nesse contexto em que se faz evidente o excesso de dados disponíveis na rede e nas diversas tentativas em organizá-los e conferir-lhes sentido.

O tema WS tem sido discutido de diferentes formas entre os acadêmicos das Ciências da Comunicação e da Informação, além de apropriado por inúmeros representantes da mídia como *o futuro* da rede mundial de computadores. Ao realizarmos uma breve e assistemática pesquisa bibliográfica, tanto em buscadores acadêmicos quanto no próprio Google, a partir do termo *web semântica* encontramos uma miríade de expressões que traduzem, em nossa opinião, visões equivocadas acerca de seus impactos sociais: a ideia de que a rede terá uma *inteligência* autônoma; de que será um repositório do conhecimento do mundo; de que conseguirá reconhecer os padrões individuais de uso e personalizar conteúdos para cada usuário de forma a fortalecer essa inteligência universal; um sistema de alavancagem do *marketing* digital; uma extensão da memória do usuário e, ao mesmo tempo, um arquivo individualizado de seus conhecimentos, entre outros.

O que vemos, em geral, é a mitificação da WS, erroneamente relacionada apenas com a ideia da *Web 3.0*, anunciada como a grande nova etapa da rede mundial de computadores, a entrada para um mundo de facilidades na relação homem-máquina, o mais novo momento transformador da sociedade contemporânea.

Evidentemente que a WS irá promover transformações significativas no uso da *web* – e em muitos casos isso já ocorre – mas pretendemos aqui refletir sobre tais transformações como resultantes de um processo que tem seu cenário instalado no *back office* da rede (algoritmos, *software*, sistemas de bancos de dados, aplicativos, linguagens de máquina etc.), um processo maquínico que só terá uma dimensão socialmente transformadora a partir do exercício do processo comunicativo, ou seja, a partir da atuação do profissional de comunicação como agenciador, municador, mediador, articulador e analista humano comportamental dessa etapa de organização e criação de significados à imensa massa informativa que circula no ciberespaço. A saber, o que chamamos aqui de curador de informação – alguém que coloca perspectiva sobre o dado.

De fato, quando o *information overload* torna-se insuportável, o usuário recorre aos especialistas, delegando a eles a tarefa de organizar e dar sentido aos dados. No caso da informação noticiosa, espera-se que o jornalista execute essa tarefa. Para alguns, no cenário da *Web 3.0*, os algoritmos serão os grandes *experts* capazes de cumprir essa missão, chegando mesmo a substituir a editorialização humana. Na nossa visão, uma WS socialmente relevante não irá falar por si mesma, mas por meio da Comunicação. Ou seja, sem o comunicador como agenciador dessa rede – ao lado da máquina – há somente uma *web* semântica totalitária. Cabe, portanto, repensar o posicionamento do jornalista e das empresas informativas diante deste próximo panorama.

# D

## PARTE 1 – O QUE É E COMO FUNCIONAM A WEB SEMÂNTICA E OS ALGORITMOS CURADORES

1. Importante lembrar que a Comunicação trabalha, em seus diferentes processos narrativos, com o recurso da ambiguidade para persuadir, fazer crer, informar, contextualizar.

Daí a importância do comunicador como mediador e municiador ativo da aplicabilidade da WS.

2. A definição mais comumente citada é a seguinte: “A Web Semântica não é uma rede separada, mas uma extensão da atual, na qual a informação recebe um melhor significado, os computadores são mais bem programados e as pessoas trabalham em colaboração” (Berners-Lee, Hendler & Lassila, 2011: 35). Breslin, Passant & Decker (2009) definem dessa maneira: “A Web Semântica é uma plataforma útil para conectar e realizar operações com uma diversidade de dados vinculados a pessoas-objetos recuperados de *websites* sociais heterogêneos”. Para outras definições, ver Bertocchi (2010).

3. Vale aqui destacar a ideia de *projeto*, ou seja, a WS não ocorre como um salto quântico, como muitos preconizam, como súbito adentramento a uma nova etapa em rede. De fato, a WS nada mais é que a própria *Web* dentro de outra configuração.

4. O termo pesquisado (*Quem escreveu Principia Mathematica?*) foi realizado no buscador Google em 15 de março de 2010.

UMA DAS DEFINIÇÕES mais precisas para *Web Semântica* é a de *unambiguous web* (Siegel, 2010). Essa visão coloca a WS como um espaço digital *desambiguado*<sup>1</sup>, onde máquinas conseguem entender o significado dos dados em diversos contextos e sem interpretações sobrepostas<sup>2</sup>.

Compreendem, por exemplo, que o Esporte Clube Vitória tem sede em Salvador, e não na cidade de Vitória da Conquista, na Bahia, localidade afastada da Vitória capixaba. Sabem que existem milhares de meninas chamadas Vitória, mas que apenas uma cresceu e reinou por mais de sessenta anos na Inglaterra. Conhecem, na literatura estrangeira, uma obra intitulada *El Pensamiento* de Francisco de Vitoria e outra, na literatura brasileira, sobre a lenda da Vitória-régia; que, por sua vez, no contexto da biologia, é uma *Victoria Amazonica*, flor que pouco ou nada tem a ver com o Esporte Clube Vitória.

Para o homem, tais sutilezas em torno do termo *vitória* são facilmente apreendidas consoante o contexto e as condições de enunciação. Para a inteligência artificial da máquina, tal apreensão é mais penosa. Daí a necessidade de um projeto – computacional, linguístico, comunicacional – chamado *Web Semântica*<sup>3</sup>.

A utilidade mais óbvia da WS revela-se ao tentamos recuperar uma informação no ciberespaço. Hoje tal processo está atrelado aos mecanismos de busca baseados em algoritmos tipo Google. Ao perguntarmos à máquina: *Quem escreveu Principia Mathematica?*, ela recupera dois tipos de informação: os três volumes que compõem *Principia Mathematica*, de Alfred North Whitehead e Bertrand Russell, e também *Philosophiæ Naturalis Principia Mathematica*, de Isaac Newton<sup>4</sup>.

Logo, para um sistema de busca computacional tanto as obras que deram origem à matemática moderna como aquelas que apresentaram as principais teorias da física que conhecemos são, *grosso modo*, a mesma coisa; embora saibamos que isso não poderia ser considerado verdadeiro. Falta-lhe, portanto, o olhar mais sensível. Ou, melhor, faltou explicar à máquina que um mesmo termo pode ter diversos significados e apontar quais são esses sentidos em cada tipo de contexto. Falta semântica ou, como afirmariam os programadores, há uma ausência de ontologias.

Na conceitualização do *World Wide Web Consortium* (W3C)<sup>5</sup>, a ontologia

(...) define os termos utilizados para descrever e representar uma área do conhecimento. A *web* semântica necessita de ontologias com diferentes níveis de estruturas. Estas servem para especificar descrições para os seguintes tipos de conceitos:

classes (geralmente coisas) nos diversos domínios de interesse; as relações que podem existir entre as coisas; e as propriedades (ou atributos) que tais coisas possam ter.

King & Reinold (2008: 8) simplificam dizendo que ontologia é uma “coleção de conceitos, organizados em uma hierarquia de categorias, combinados com as relações entre os conceitos, a fim de refletir o vocabulário de uma área de conhecimento”.

Por que a *web* semântica precisa de ontologias? Não bastaria um dicionário ou uma taxonomia bem estruturada para explicar à máquina as nuances entre os termos? O dicionário tem como propósito definir conceitos. O formato permite apresentar mais de uma definição por termo. Há dicionários de sinônimos, antônimos, de termos específicos de uma área, entre outros. O *Tesauros* trabalha na lógica do sinônimo, aproximando conceitos similares. A taxonomia coloca os conceitos em hierarquia. Mas é a ontologia que cria relações semânticas entre os termos – e, de fato, com ontologias, um número indeterminado delas pode ser criado, como um mapa mental.

Emerge aqui uma primeira questão ao centrarmos o foco no campo do jornalismo digital: a necessidade ou não da construção específica de ontologias para o campo. Tal questionamento será discutido mais adiante.

As ontologias podem ser enriquecidas pela folksonomia (*tagging* social) (Angeletou et al., 2007). O *tagueamento* livre de conteúdos é dinamicamente atualizado por uma plêiade de usuários, o que o torna rico em jargões, expressões populares e termos corriqueiros que podem ser incorporados às ontologias formal e previamente estruturadas que passam a ganhar as características da linguagem natural: o multilinguismo, a ambiguidade, a inventividade.

Numa *web* que dispensa a ontologia, geralmente a navegação do usuário é conduzida pelo sistema de busca que, por sua vez, é baseado num algoritmo (um protocolo padrão de varredura do conteúdo dos dados disponíveis na rede). Se considerarmos o onipresente Google, o *spider* faz uma varredura nos dados *tagueados* e organizados no espaço de metadados propiciado pela estrutura da linguagem HTML de uma página; faz o mesmo sucessivamente para todas as páginas endereçadas na rede; organiza estas palavras-chave num índice de dados; quando um usuário alimenta a ferramenta com um termo de busca, o sistema processa tal lista de palavras-chave e rastreia os sites que contêm a mesma, apresentando o resultado por meio de um algoritmo – o *PageRank*, que os ordena pela frequência com que a palavra aparece na página, pelo tempo que a mesma está no ar e pelo número de outras páginas que remetem *links* para a mesma, entre outras alternativas.

5. Entidade global, sediada no EUA, que define os padrões de uso e os protocolos do WWW, e que atualmente possui um grupo específico para a definição dos padrões da WS.

# D

## A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria

Com isso, o papel do comunicador neste processo depende do seu domínio na construção de um conteúdo de página passível de *tagueamento* dentro das regras de programação; e de sua competência em fazer este conteúdo visível para o buscador por meio de técnicas de SEO (*Search Engine Optimization*). É um processo em que a ação comunicativa é determinada pelo sistema de busca e não pelos significados atribuídos pelo público ao produto ou serviço de uma dada marca ou empresa e às relações que estas estabelecem com os públicos.

É neste ponto que folksonomias e ontologias constituem-se como elementos-chave para introduzir a significação do mundo real nos processos de busca e uso da rede. São elementos básicos de aproximação entre gerador/usuário de dados por meio de um processo de troca de significados comunicacionais que são incorporados semanticamente ao processo de busca dos sistemas. A presença do comunicador na construção de folksonomias e ontologias – e na instrução do processo construtivo de páginas com *links* semânticos – será fundamental para a consolidação da WS.

A pesquisadora Silvia Laurentiz (2010) acredita que o maior desafio para a WS reside na questão da criação de ontologias.

Definir ontologias prevê um conjunto de métodos e técnicas automáticas ou semiautomáticas para aquisição de conhecimento utilizando textos, dados estruturados e semiestruturados, esquemas relacionais e outras bases do conhecimento (Laurentiz, 2010: 17).

Quando transportamos as palavras de Laurentiz para a cena cotidiana das redações jornalísticas vemos que a ideia de uma estruturação prévia universalizante – a ontologia – fica deslocada do processo jornalístico que pressupõe, simultaneamente, critérios editoriais de noticiabilidade do conteúdo, e aqueles relativos ao SEO na construção de títulos, menus, seções, *tags* e metadados. Seria um cenário de embates entre o automatismo e a práxis jornalística.

Aqui é oportuno refletir qual o papel do comunicador, do jornalista e da empresa jornalística para manter tal conjunto de valores num processo que visa a universalização destes, e que é construído, por total falta de opções, a partir de um conjunto de subjetividades (aquelas dos construtores da ontologia).

Tem sido observada nas diferentes práxis redacionais com significativa presença na rede a opção por uma espécie de *caminho do meio* para adentrar ao mundo da WS sem, contudo, abdicar da especificidade do campo. Tal caminho não prioriza a ontologia como base de uma atuação semântica das marcas jornalísticas na rede, mas enfatiza a construção proprietária de CMS's (*content management systems*) – os sistemas publicadores. Fala-se, portanto, do uso de

*marcadores semânticos* (não mais ontologias) incorporados no próprio CMS, ou seja, um conjunto de termos e palavras, definidos estrategicamente, segundo o modelo de negócio e a linha editorial, pela empresa informativa proprietária e usuária do CMS, que funcionam como parâmetros para qualquer jornalista durante a produção noticiosa. Tais parâmetros levam em conta características do segmento/audiência da publicação, e os indicadores do SEO. O CMS, por *default*, conduz o jornalista a categorizar o conteúdo produzido dentro desse conjunto de marcadores.

São práticas que viabilizam a associação do CMS semântico às demais práticas emergentes da cena jornalística 3.0 que incluem os algoritmos e os *apps*, conforme veremos a seguir.

Em resumo, seja via CMS, ontologias, algoritmos ou *apps*, a matéria-prima de todo o processo comunicativo/jornalístico permanece centrada na palavra como signifiante do mundo e valor-notícia. É pela palavra que se constroem ontologias, que se efetiva a ação do buscador, que o jornalista marca o conteúdo e o algoritmo personaliza (por meio das *hashtags* #) a oferta informativa. Tal constatação consolida o papel do artífice da palavra.

Num caminho paralelo, na *web* ou através de aplicativos para *tablets* e outros dispositivos móveis, observa-se uma significativa presença de soluções algorítmicas para lidar com a profusão informativa. O buscador Google (e outros mais segmentados<sup>6</sup>) e a rede social Facebook são baseados em algoritmos<sup>7</sup> curadores que decidem qual a informação será disponibilizada. Os sistemas agregadores, baseados na organização de fluxos informativos em RSS<sup>8</sup>, utilizam um algoritmo simplificado, sendo o GoogleReader o melhor exemplo. Os sistemas de recomendação como o da livreria virtual Amazon, ou dos *sites* de música, como LastFm e Spotify, também se utilizam de algoritmos relativamente simples. E há os algoritmos criados para sustentar modelos de negócio e que se transformaram em produtos específicos baseados em *feeds* pessoais e/ou *tags* definidas pelos usuários: é o caso da Zite e da Flipboard, revistas personalizáveis, únicas para cada usuário e exclusivas, por ora, para *tablets*; do Paper. Li que organiza jornais pessoais a partir da navegação dos usuários nas redes sociais; do Scoop.it que cria conteúdo dinâmico a partir de temáticas definidas pelo usuário; o Storify que constrói *narrativas sociais* e viraliza o resultado obtido; o Social Pulse da agência de notícias Reuters apresenta o noticiário mais tuitado e retuído por uma rede de influenciadores; o Washington Post Social Reader, agregador de notícias igualmente personalizado; e as ferramentas de organização da informação por meio de imagens, trazendo a inovação do aprendizado social na construção de narrativas visuais, a exemplo do Pinterest e do WeHeartIt.

6. Incluem-se, entre outros, nessa categoria de buscadores o Addict-o-matic que privilegia a busca do *buzz* momentâneo nas redes sociais; o DuckDuckGo que busca conteúdos em *sites* de *crowdsourcing*; o Icerocket que busca em tempo real na rede.

7. Adotamos o termo *algoritmo-curador* para identificar esse tipo de ação automatizada.

8. Really Simply Syndication.

# D

## A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria

Referenciamos apenas os exemplos mais populares daquilo que denominamos ferramentas de organização da abundância informativa por meio de algoritmos. O que temos a refletir sobre essa disponibilidade é o seu forte papel curatorial, já intrínseco às ferramentas, e a noção de que todo algoritmo é produto de um processo humano, com critérios de escolha previamente definidos com base em algum contexto de oferta da informação, conforme explicado a seguir.

O termo *algoritmo* é uma palavra latinizada, derivada do nome de Al Khowarizmi, matemático árabe do século XIX. Na computação, um algoritmo é um “procedimento criado para cumprir uma tarefa específica” (Skiena, 2008: 3).

Trata-se de um passo a passo computacional, um código de programação, executado numa dada periodicidade e com um esforço definido. O conceito de algoritmo permite pensá-lo como um procedimento que pode ser executado não apenas por máquinas, mas ainda por homens, ampliando seu potencial de acuidade associada à personalização.

Deste modo, desenhar um algoritmo é elaborar uma série de instruções com a finalidade de resolver um problema. No caso da *timeline social* do Facebook, o algoritmo é utilizado especialmente para ordenar elementos (dados sobre outros usuários) por critério de importância (definida pelo programador da empresa). Para o Google, o algoritmo busca resolver o problema da compreensão dos dados e retorno de ocorrências afins. Para Zite ou Flipboard, o objetivo é produzir uma revista digital apenas com conteúdos relevantes para aquele usuário, e assim por diante. No cenário da comunicação digital, a rigor, o algoritmo trabalha com a missão de expurgar informações indesejáveis, oferecendo apenas o que o usuário julgaria eventualmente o mais relevante para si, conforme um modelo de negócio definido ou de acesso às informações também previamente determinado pelo proprietário do algoritmo.

Ao considerarmos o uso de algoritmos para decidir critérios de noticiabilidade no jornalismo, a exemplo do Storify – um algoritmo curador – estaríamos avalizando um processo que prescinde das variáveis *jornalista* e *acontecimento social relevante* típicos do campo.

Podemos depreender do funcionamento desse tipo de algoritmos curadores que a função jornalística nesse contexto reduz-se à produção da informação e sua disponibilização para o imenso cadinho informativo da *web*. Caberá somente à máquina e à intensidade de uso destas plataformas por usuário individual a decisão sobre relevância, recorte editorial, marca jornalística que serão oferecidos ao leitor?

## PARTE 2. A MÁQUINA QUE COMPREENDE E AS REDAÇÕES QUE CONVIVEM COM MÁQUINAS INTELIGENTES

Estamos diante de um cenário cibercultural para o mundo informativo que disponibiliza para as redações contemporâneas uma rede mundial de computadores com um excesso de informações nunca visto; uma proposta de médio prazo de uma outra rede – WS; uma realidade iminente de um processo de construção da opinião coletiva pautado por plataformas baseadas em algoritmos curadores; e uma possibilidade (caracterizada como investimento) de um processo intermediário com a marcação semântica dos CMS's e a folksonomia jornalística.

Tal cenário é, ao mesmo tempo, previsto por diferentes analistas (conforme citados mais adiante) como o futuro do jornalismo e também como uma nova onda de mudanças paradigmáticas.

O chamado *information overload* que a rede mundial de computadores abriga tem sido o *leit motif* para justificar a emergência de sistemas automatizados em detrimento de um processo de modelamento com evidente intervenção humana. Vejamos algumas reflexões autorais sobre o tema.

Excesso informativo, ansiedade de informação, dieta informativa (Turkle, 2011: 242; Pariser, 2011; Johnson, 2012) são expressões que começaram a surgir na década de 70 do século passado<sup>9</sup> e, desde então, passaram a ser incorporadas no discurso de pesquisadores da Comunicação e, cada vez mais, nas conversas de não acadêmicos e da mídia. Reflexo de um momento em que instituições científicas acadêmicas, escolas, bibliotecas, museus, organizações, e corporações de toda espécie – e não somente os meios tradicionais de informação e comunicação – partilham dados digitais na rede, multiplicando a oferta informativa para a sociedade.

Constata-se uma abundância informativa e uma disseminação de múltiplas narrativas sobre os acontecimentos do mundo, criadas a partir de distintas fontes, e apresentadas em formatos para todos os gostos. Contudo, inexistente a garantia de transformação de *dados* em *informação* e tampouco estes em *conhecimento* por parte do homem<sup>10</sup>. O especialista em mídia Neal Gabler (2011) chega a afirmar que a sociedade vive na era da *pós-ideia*, ou seja, os indivíduos se tornaram grandes acumuladores de fatos e informações, mas já não conseguem desenvolver um pensamento crítico e profundo sobre um fato. O comentarista de mídia declara que a era digital nos libertou para a *ignorância bem informada*.

Weinberger (2012) identifica neste cenário de *information overload* uma *crise de conhecimento*. Para o autor, no mundo pré-internet, as informações permaneciam clara e concretamente localizadas (em livros, bibliotecas, jornais) e permitiam a construção do conhecimento por meio de *trabalho duro* de estudiosos que se tornavam, por consequência, especialistas em assuntos. No

9. Alvin Toffler introduziu a ideia de excesso de informação em 1970 com a publicação de seu livro *O Choque do Futuro*.

10. Referenciamos aqui a famosa pirâmide DIKW (dado, informação, conhecimento e sabedoria) criada em 1988 pelo pesquisador Russell Ackoff.

# D

## A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria

contexto digital, vive-se o oposto, segundo o autor: as informações encontram-se espalhadas desordenadamente; são produzidas por amadores, plagiadores e usuários que consideram um bom conteúdo aquele que possui o maior número de polegares indicando *curtir*. Ainda assim, reitera Weinberger, os meios de comunicação, as empresas, os governos e a ciência se beneficiam dos dados disponíveis na rede, sobretudo por conta das contribuições de leitores que, ao colaborar com o processo, tornam os dados muito mais localizáveis e compreensíveis por outros tantos leitores.

Ou seja, a crise do conhecimento nasce numa época de exaltação do conhecimento. O conhecimento em rede, embora menos acurado, torna-se *mais humano*.

Analisando o mesmo cenário sob um recorte da filosofia social, ressurgem o pensador Pierre Lévy que defende – desde pelo menos 1991, quando publicou *Ideografia Dinâmica* – que as escritas-máquinas (programação) se aproximam do funcionamento do pensamento humano. Para ele, contudo, sempre faltou à programação computacional desenvolver o lado representativo e expressivo da linguagem (o simbólico, próprio das escritas do homem). Lévy propõe um olhar de convivência entre máquinas inteligentes, construtores de informações e a sociedade como um todo e, a partir de 2006, começou a se aprofundar no que ele chama de *la langue de l'intelligence collective*, intitulada de IEML (*Information Economy Meta Language*).

A IEML é uma linguagem artificial da inteligência coletiva que é: a) manipulável por computadores e, b) capaz de expressar as nuances semânticas e pragmáticas das línguas naturais. É a WS sendo compreendida de um ponto de vista não somente computacional, mas enquadrada também numa perspectiva filosófica e social. Ou, como afirma o autor, trata-se de um trabalho “*human-centric social computing*” (Lévy, 2009: 32).

Na proposta de Lévy, a marcação semântica e a folksonomia jornalística que hoje começam a ser praticadas em grandes redações, deveriam ser ampliadas para a *web* como um todo, criando-se um verdadeiro *espaço semântico* de construção social.

Na mesma linha filosófico-social temos o pesquisador João Fernandes Teixeira (2010) que vai mais além e trata do mundo pós-evolutivo – ou pós-humano, cuja característica principal é a impossibilidade de continuação da evolução do corpo e cérebro humanos, forçando a espécie a recorrer à *parabiose* (associação do homem com formas de vida seca):

Até a chegada do mundo pós-evolutivo estamos enclacrados numa situação na qual as máquinas nos trazem muita infelicidade, pois somos meras extensões delas e não podemos parasitá-las. Contudo, isso não significa que a chegada do

mundo pós-evolutivo vá, necessariamente nos tornar felizes. A parabióse não é a realização de uma utopia (Teixeira, 2010: 12).

No mundo da convergência digital, figuram supermáquinas ou superci-borgues com um enorme poder computacional que processam uma quantidade formidável de dados. Essas supermáquinas ou máquinas híbridas “terão um cérebro mais poderoso que o humano, sendo capazes de processar dados e programas estatísticos que nossos atuais supercomputadores ainda não conseguem” (Teixeira, 2010: 73-74).

Em outras palavras, as máquinas não apenas processarão imensas bases de dados e informações como ainda produzirão *conhecimento*. Para Teixeira, esse será um *conhecimento artificial*, construído a partir da correlação de base de dados – inclusive correlações aleatórias – o que gerará previsões e projeções em diversos campos. “Uma forma de transformar informação em conhecimento é cruzar dados”, afirma (Ibid.: 75).

Entramos, portanto, num mundo em que a máquina não apenas compreende dados – o que é o objetivo maior da *web* semântica – como, num passo seguinte, os correlaciona, o que, na visão do autor, irá gerar novas proposições para a humanidade.

Entre o excesso informativo e a emergência de uma sociedade informada por máquinas inteligentes e constituídas como uma extensão da mente humana, vemos que o campo jornalístico – ainda que conservador – não se coloca à parte do cenário. Pelo contrário, algumas das mais importantes marcas jornalísticas globais – The New York Times, The Guardian, BBC, Reuters, entre outras, buscam aproximações a esse mundo semântico, pulverizado por *apps* e plataformas.

Ao final de 2011, Jeff Sonderman, analista do Poynter Institute, conhecido no campo pela seriedade de seus estudos, publicou três tendências que pautariam o jornalismo a partir de 2012. Todas elas refletem, de alguma forma o uso de algoritmos e de sistemas semânticos. São elas:

- a emergência de narrativas para além do texto autoral, sustentadas por diversas camadas informativas a exemplo do Storify<sup>11</sup> e do ProPublica<sup>12</sup>. A tendência, portanto, é que o público seja um colaborador na captação noticiosa, que o jornalista assuma o papel de ouvinte da sociedade e de curador, e que as notícias sejam narradas pela sociedade por meio de um jornalista ao invés da tradicional narrativa de um jornalista para a sociedade;
- o uso do Facebook e outras plataformas de mídias sociais (tipicamente constituídas por algoritmos) como o melhor meio de publicação e disseminação de notícias; e

11. [www.storify.com](http://www.storify.com).

12. [www.propublica.org](http://www.propublica.org).

# D

## A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria

- o predomínio de *tablets* e *e-readers* (representando a mobilidade do leitor) como o principal meio de acesso informativo.

O pesquisador de Columbia C.W. Anderson afirma que a curadoria jornalística realizada pelo algoritmo pode ser inserida em um campo de estudo emergente – a sociologia do algoritmo (Anderson, 2011: 529). Ao falar sobre *jornalismo de algoritmos*, Anderson afirma que o algoritmo está desempenhando um papel sociotécnico cada vez mais importante na mediação entre jornalistas, audiência e produtos de mídia, e esta mediação tem implicações tanto sociológicas como normativas. Sua pesquisa foi extensiva nas redações dos principais jornais norte-americanos e conclui com uma visão positiva do uso de algoritmos pelas redações: é um caminho tecno-social eficaz para que os jornais conheçam e compreendam melhor o seu público, incorporando seu uso no DNA do trabalho jornalístico contemporâneo.

A consultoria MediaBistro, por meio de seu blog 10.000 Words – Where Journalism Meets Technology (Rabaino, 2011), também indicou mudanças paradigmáticas para o futuro, passando por:

- automatização de sistemas de coleta de informações e publicação (gerenciados por algoritmos), reduzindo o volume de papel nas redações e ampliando as possibilidades de conexão e sincronização dos conteúdos. Destaca-se o uso de GoogleDocs, iPads e Evernote;
- compartilhamento público de APIs<sup>13</sup> e códigos-fonte dos conteúdos informativos de modo a ampliar o alcance da informação produzida pela redação, não importando a ideia de *branded content*.
- consolidação do sistema de captação *open news*, a exemplo do recém-lançado programa do The Guardian, onde a pauta é postada aberta na rede para a recepção de colaborações dos leitores;
- consolidação de sistemas publicadores (CMS) inteligentes baseados mais em ontologias e folksonomias que em taxonomias.

Nicholas Carr (2011), veterano autor da área tecnológica, preconiza que, em breve, os negócios informativos deverão estar *appificados*, ou seja, direcionando seu conteúdo cada vez mais para *apps* acessíveis por dispositivos móveis e que agregam diferentes fontes noticiosas, reduzindo o peso das páginas *web* de cada marca informativa. No dizer de Carr:

Hoje, como resultado da computação em nuvem e de outros avanços, os *apps* se parecem muito com produtos midiáticos. Eles são sustentados por publicidade, aceitam assinaturas, são atualizados constantemente e o conteúdo que disponibilizam é tão importante quanto suas funcionalidades. Na medida em que as empresas de mídia tradicional passaram a distribuir seus sistemas em formatos

13. APIs - *Application Programming Interface* (ou Interface de Programação de Aplicativos) é um conjunto de rotinas e padrões estabelecidos por um *software* para a utilização das suas funcionalidades por aplicativos que não pretendem envolver-se em detalhes da implementação do *software*, mas apenas usar seus serviços.

digitais – a exemplo dos códigos (APIs) – elas começam a parecer empresas de software. Oferecem não apenas conteúdo original, mas também um conjunto de ferramentas e funcionalidades *online* que permitem aos consumidores ver, manipular e ampliar tais conteúdos numa miríade de possibilidades (Carr, 2011).

Podemos inferir de tais exemplos que o *caminho do meio* parece ser a opção em uso por parte das empresas informativas. Mesmo assim, fica evidenciado que estas ainda transitam na indefinição quando necessitam colocar num contexto estratégico e de modelo de negócio decisões importantes como investir num CMS proprietário – que implica numa mudança cultural e profissional; na *apfificação* do conteúdo – que implica na negação do *branded content*; ou até no desinvestimento em sua página *web*.

Uma vez discutido o cenário, nos dedicamos na terceira parte deste trabalho a refletir sobre as possibilidades de posicionamento do jornalista e das empresas informativas diante do panorama cibercultural vigente, sem, contudo, questionar a velha máxima de sobrevivência do jornalismo. Propomos, sim, um posicionamento resiliente e adaptativo aos tempos futuros.

### PARTE 3. JORNALISTAS E SEUS *MODUS OPERANDI* NUM CONTEXTO SÓCIO-MAQUÍNICO: A CURADORIA

Em termos objetivos podemos afirmar que hoje o jornalista transita entre o excesso de informação, a necessidade de relatar o mundo para a sociedade, o aprendizado tecnológico, o predomínio de dispositivos móveis como (não) localização da audiência e a proximidade às vezes intrusiva do velho leitor. Ou seja, um panorama tipicamente da cibercultura vigente.

O excesso de informação exige organização e contextualização, o aprendizado tecnológico demanda investimento corporativo e disponibilidade intelectual, a mobilidade e a proximidade solicitam novos formatos narrativos. Quantos desses requisitos se aproximam do processamento informativo vigente nas empresas informativas embasadas numa proposta clássica?

Ainda que possamos classificar como indícios alguns experimentos realizados por veículos de comunicação ou por plataformas curadas por algoritmos<sup>14</sup>, parece-nos que num contexto de abundância informativa em que a máquina processa mais rapidamente e melhor as informações, inclusive podendo fazer correlações e gerar, futuramente, um conhecimento artificial, a comunicação, o jornalismo e seus profissionais permanecem indispensáveis por pelo menos dois motivos.

O primeiro motivo: nenhum robô ainda passou no teste de Turing<sup>15</sup>, ou seja, ainda não conseguem *conversar* como um humano, interagindo de forma

14. Em 2012, no Brasil, a Editora Abril está desenvolvendo seu CMS proprietário e integrador de suas marcas – o Alexandria, concebido dentro da proposta de uso de marcadores semânticos e de folksonomia jornalística.

15. É um teste proposto por Alan Turing em uma publicação de 1950 chamada *Computing Machinery and Intelligence* cujo objetivo era determinar se máquinas podiam pensar. Disponível em: <[http://pt.wikipedia.org/wiki/Teste\\_de\\_Turing](http://pt.wikipedia.org/wiki/Teste_de_Turing)>.

# D

## A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria

natural e estabelecendo empatia (ou antipatia) como outro humano poderia gerar, nem fazendo correlações profundas e imprevisíveis.

Foi exatamente o que o jornal norte-americano The New York Times percebeu em maio de 2011, conforme publicado no site do Instituto Poynter.org (Sonderman, 2011). O jornal, à época, trocou seus robôs por jornalistas de verdade. Abandonou a publicação automática de conteúdos em sua conta no Twitter (@nytimes), substituindo os robôs por seres humanos, por uma semana, a título experimental, para checar se uma abordagem humana interativa seria mais eficaz em termos de comunicação, engajamento e audiência. Uma das editoras, Liz Heron, afirmou que se tratava de uma tentativa de observar como se daria a abordagem *cyborg* com a abordagem humana.

16. O jornal não divulgou os resultados do experimento.

Após uma semana, com o fim do experimento<sup>16</sup>, o jornal voltou ao formato antigo e os robôs retomaram a alimentação da conta no Twitter. As conversas com os leitores sobre o noticiário foram interrompidas (abandono do *Reply*) e as *hashtags* (#) eliminadas do cenário. Os retuítes também não puderam ser efetuados, já que a decisão de *recomendar* um conteúdo só pode ser decidida por um editor. As manchetes das notícias voltaram a ser idênticas às publicadas no *site*, sem edição específica para um espaço reduzido que apenas comporta 140 caracteres usuais (SMS). Mesmo sem a divulgação oficial, concluiu-se que sistemas, robôs e algoritmos não dialogam com a audiência e pouco influenciam a formação da opinião coletiva, apesar de otimizarem os custos e os recursos da redação.

Por outro lado, o próprio The New York Times registra um sucesso significativo com as postagens em sua página no Facebook, cuja tecnologia não privilegia a automação do conteúdo. O próprio jornal declara que a relação estabelecida com sua audiência por meio dos comentários às postagens é algo precioso para o seu conteúdo e sua consolidação.

O segundo motivo: no espaço semântico, robôs inteligentes percebem o comportamento do usuário e fornecem a ele somente as informações que ele próprio deseja e escolhe. Chega-se, portanto, ao nível máximo da personalização, o que, para além das vantagens evidentes, traz algumas implicações indesejáveis.

A principal é que o cidadão fica enalacrado naquilo que Eli Pariser (2011) chamou de *“the filter bubbles”*. Numa palestra recente ele explicou dessa forma:

Sou politicamente progressista, mas (...) gosto de ouvir o que eles (os conservadores) estão pensando (...) Assim, fiquei surpreso quando notei um dia que os conservadores tinham desaparecido da minha página (timeline) do Facebook. E o que estava acontecendo era o seguinte: o Facebook estava computando e compreendendo os links nos quais cliquei, e foi notando que, na verdade, eu clicava

17. TED. Technology, Entertainment, Design. Palestra online. Disponível em: <[http://www.ted.com/talks/eli\\_pariser\\_beware\\_online\\_filter\\_bubbles.html](http://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles.html)> Acesso em: jun. 2011.

mais informações sobre os meus amigos liberais em comparação aos meus amigos conservadores. E sem me consultar sobre isso, editou a informação que julgou não me interessar. Eles desapareceram. (...) (Pariser, 2011)<sup>17</sup>

Pariser diz que não apenas o Facebook, mas também o Yahoo News, Huffington Post, The Washington Post e o The New York Times estão flertando com a personalização de várias maneiras. Significa dizer que caminhamos em direção a um mundo em que a *web* está nos mostrando o que queremos ver, mas não necessariamente o que nós precisamos ver. O autor sustenta que estamos vendo uma passagem de tocha dos *gatekeepers* humanos para os algoritmos. Com o agravante que os algoritmos não têm embutido o tipo de ética que os editores tinham. E são os algoritmos que vão decidir o que iremos assistir, ler, ouvir, saber, conforme aquilo que julgarmos relevantes, mas não necessariamente consoante aquilo que for relevante do ponto de vista de toda uma sociedade, por exemplo. Pariser afirma que “precisamos ter certeza de que eles também nos mostram coisas que são desconfortáveis” (2011).

É nesse ponto que iniciamos uma reflexão para a configuração de um jornalista-curador<sup>18</sup>: o estabelecimento de uma relação entre sofisticação do algoritmo e correspondente intervenção humana especializada (o jornalista) no processo de sua construção. Quanto mais informações circunstanciais, sociais e comportamentais se fizerem necessárias para a modelagem do algoritmo, mais deveria ser exigida a participação do elemento humano como alimentador do modelo e, especialmente, como refinador ao longo da vida útil do algoritmo.

Observamos isso, sobretudo, quando a curadoria realizada pelos algoritmos menos complexos tende a olhar para trás: considera o comportamento passado do usuário, o que ele comentou, recomendou, apreciou, leu. Conhecendo padrões e preferências, esse passo a passo matemático traz mais informações similares e afins para seu usuário, a partir de uma varredura rápida e eficiente pelas bases de dados.

A curadoria humana pura e simples (sem os procedimentos matemáticos), por outro lado, é mais livre para olhar para o futuro. Um curador de conteúdos é capaz de agregar novas e inusitadas perspectivas à informação, oferecendo aos seus usuários a surpresa, o inesperado ou simplesmente aquilo que o usuário nem imaginaria existir no mundo e sobre o mundo, ampliando seu próprio entendimento de mundo.

A curadoria jornalística, que faz uso de algoritmos sofisticados associados às expertises de re-mediação do jornalista, parece ser o campo para o entendimento e caracterização da curadoria da informação na Comunicação. A repórter Jane Wakefield, em artigo publicado pela BBC News, chega a reiterar:

18. São múltiplas as definições de curadoria, originalmente calcadas nas Artes e no Direito. Neste texto assumimos aquela de Steven Rosenbaum: “Curadoria é quando o homem agrega um valor qualitativo àquele conjunto de informações que está sendo compilado e organizado (ainda que seja num primeiro momento reunido por máquinas)”(Rosenbaum, 2011, location 395).

# D

## A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria

“Algoritmos podem ser mais espertos que o homem, mas eles não possuem necessariamente a noção de perspectiva do homem” (Wakefield, 2011).

Anderson (op. cit., 2011) qualifica esse processo como “promiscuidade algorítmica”: a capacidade da máquina de agregar e tratar os dados computacionais e os dados humanos da mesma forma, reduzindo e esquematizando estes últimos.

Os leitores de informações assim são reduzidos às suas preferências, ao que consomem. A *agenda setting* passa a ser pessoal, única, personalizada e determinada por seus desejos. Processo indesejável do ponto de vista da comunicação social como um todo, em que diferentes pontos de vista, fontes, perspectivas e recortes são fundamentais para o alargamento da visão de mundo desses leitores e para a construção do conhecimento da humanidade (Pariser, 2011; Basulto, 2012).

Ao refletirmos sobre a curadoria jornalística partimos do pressuposto de que o profissional, por vivenciar a proximidade com aquele que quer e deve ser informado, possui um cabedal de variáveis importantes sobre processos de escolha, gostos, direcionamento da formação de opinião, variação de padrões sociais, entre outros dados, que possibilitam o uso de modelos algorítmicos menos alienantes e, portanto, mais adequados à função social do campo da Comunicação.

É possível inferir que o cenário, inicialmente de face adversa, é propício a mais um salto paradigmático no campo jornalístico e à evidência do papel humano especializado como ponte necessária entre informação e sociedade.

Pensando nas dimensões de ação da curadoria jornalística, é plausível concluir que – apesar da possibilidade potencial do exercício da curadoria por qualquer cidadão conectado – no campo da comunicação digital tal potencial é reduzido ao conjunto de profissionais que transitam em domínios simultâneos da arte do re-mediador (= interconexão de mídias), do estabelecimento de relações interpessoais pró-ativas (= seleção de públicos/audiências), da concepção ou uso de plataformas tecnológicas para tratamento e disponibilização de dados (= recorrência a bases de dados) e, principalmente, da capacidade intelectual e informativa para curar de modo único e diferenciado (= agregação de valor). Todo esse conjunto de competências pode ser associado aos algoritmos-curadores, de forma que não apenas entre em jogo o histórico passado de apropriação e preferências informativas por parte do usuário, mas ainda, e principalmente, o resultado dessas ações específicas do comunicador.

Nessa perspectiva, o algoritmo curador agirá a partir de dois vetores combinados: a marcação formal compreendida no publicador (CMS), e a ação do jornalista que vivencia a relação com o ambiente e a sociedade, criando marcadores randômicos e pessoais.

O que inferimos é que a curadoria no jornalismo é capaz de se favorecer da evolução tecnológica, mas ela não é fruto de um determinismo. Ao contrário, ela é uma ação deliberada de quem assume papéis de re-mediação na sociedade, seja um ativista, um blogueiro, um pesquisador, um jornalista ou um comunicador.

Falamos, portanto, de uma ação curatorial especializada. E que faz sentido, sobretudo, ao considerarmos o processo cada vez mais proeminente de perda da diversidade de conteúdos ofertada por meios de comunicação tradicionais em espaços digitais. Boczkowski (2010) aponta um quadro preocupante em relação a tal problema: existe um crescimento da proporção de conteúdos *genéricos* compartilhado pelos principais meios de comunicação (e mesmo entre provedores terceiros, como agências noticiosas); ao mesmo tempo em que há uma redução de empresas de mídia independentes de conglomerados jornalísticos a desempenhar a sua função de *cão de guarda* de uma forma vigorosa, resultando em *mais do mesmo* em termos informativos em sítios digitais que se propõem a manter seus leitores bem informados sobre os acontecimentos do mundo.

O salto da reprodução mecânica para a digital seja de notícias ou de outras áreas do trabalho simbólico introduziu, no início do século 21 uma era de abundância informativa, caracterizada pela preocupação sobre o crescente custo-benefício entre qualidade e quantidade de informação. Ter mais informações disponíveis do que recursos para processá-las e o temor que esse crescimento em quantidade cause a deterioração da qualidade do produto resultante e das experiências de consumo não são novidade, mas estão ficando cada vez mais acentuadas nos últimos anos (Boczkowski, 2010: loc.<sup>19</sup> 2445, tradução nossa).

Temos claro que nem todo curador (no sentido etimológico do termo) é um potencial comunicador no mundo digital, ou um jornalista das redações; por outro lado, não deixamos de observar que a ação comunicacional em rede digital está cada vez mais ancorada em processos curatoriais que vão desde simples sistemas de recomendação à construção de complexos algoritmos, passando todos eles pela mediação inerente à atividade comunicacional.

Em *Facts are Sacred: The power of data*, editado pelo jornal britânico The Guardian, Rogers (2011) revela que a curadoria realizada por um especialista (o jornalista, no caso) prevê habilidades bem distintas, entretanto: inclui horas compilando e analisando dados em tabelas Excel e documentos em PDFs disponíveis na Internet, procurando um padrão informativo e com valor-notícia para, então, a partir desse trabalho, criar um conjunto informativo perspectivado de uma maneira até então inédita (Roger, 2011: loc. 64) .

19. No Kindle, a localização é feita pela numeração de *locations*, cada qual relativa a um trecho curto do texto, de modo que o correspondente a uma página inclui uma sequência de cerca de 15 *locations*.

# D

## A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria

O autor afirma que a abundância de dados digitais transformou o jornalismo e, por extensão, a própria comunicação. O chamado jornalismo de dados (*data journalism*) torna-se, muitas vezes, curadoria, como afirma:

*O jornalismo em base de dados transformou-se em curadoria? Algumas vezes, sim. Hoje existe uma tal quantidade de dados disponível no mundo que procuramos oferecer em cada notícia os fatos principais – e encontrar a informação correta pode se transformar numa atividade jornalística tão intensa quanto buscar os melhores entrevistados para uma matéria (...) Qualquer um pode fazer isso... Especialmente com ferramentas gratuitas como o Google Fusion Tables, Many Eyes, Google Charts ou Timetric – e você pode acessar postagens dos leitores no seu grupo do Flickr (...) Mas, a tarefa mais importante é pensar sobre os dados obtidos mais como jornalista do que como um analista. O que é interessante sobre tais dados? O que é novo? O que aconteceria se eu mesclasse com novos dados? Responder tais perguntas é da maior importância. Funciona se pensarmos numa combinação disso tudo (Rogers, 2011: loc. 56-71, tradução livre e grifos nossos).*

Steven Rosenbaum (2011) preconiza uma curadoria ampla na sociedade, mesclando cidadãos, profissionais, ferramentas e plataformas, sem qualquer distinção de função ou necessidade. Para ele, a rede e seus usuários se autoformatam. Nessa linha, temos a discordar quanto a não consideração das possibilidades de exploração comunicacional desse processo curatorial hoje disponível em rede.

Beiguelman (2011) considera o curador de informação aquele que faz uso das ferramentas, associadas a filtros e plataformas que facilitam a atividade de agenciamento, gerando uma inteligência distribuída.

A partir de todo o cenário aqui descrito é possível afirmar que o jornalista em especial, e o próprio campo da Comunicação se constituem no cenário mais propício às atividades de curadoria da informação pelas próprias características. Também fica evidenciada a ação de curadoria alavancada pelas ferramentas baseadas em algoritmos, que funcionariam como parceiros do comunicador. A questão que se coloca, por fim, refere-se ao *status* de tal parceria e suas perspectivas futuras, abrindo-se mais um cenário para debates.

### CONSIDERAÇÕES FINAIS

Procuramos ao longo deste trabalho organizar as principais ideias que hoje estão em plena discussão sobre a influência da chamada terceira onda da *web* nos campos da informação jornalística e da Comunicação como um todo. Um cenário que consideramos típico da cibercultura vigente.

Para isso, foi necessário apresentarmos e explicarmos como funcionam a WS, os algoritmos e os *apps* – elementos técnicos fundantes para a existência dessa recente camada técnico-informativa, também conhecida como *Web 3.0*. Em seguida, buscamos correlacionar tais elementos com o processo de produção de conteúdos informativos e de ações comunicativas como um todo, indicando as potenciais mudanças. Por fim, discutimos a proposição de que o comunicador em geral e o jornalista em particular rumam para a absorção de mais um papel em sua atuação, delineado pelas características técnicas e interativas decorrentes do cenário em que a máquina pode assumir um papel mais ativo no processo de produção de conhecimento e opinião da sociedade na rede.

Evidentemente, a discussão aqui apresentada é complexa e, no momento, pouco conclusiva. A cena cibercultural que o mundo informativo adentra exigirá, novamente, um reposicionamento cultural e comportamental, tanto dos profissionais de comunicação, como das empresas informativas. A introdução da atividade/papel de curadoria deve persistir ainda pelos próximos anos. O contexto exigirá, ainda, investimentos significativos em sistemas CMS, ontologias e vocabulários controlados, ou seja, em áreas indiretas do *core business* cujo retorno não é imediato. E, conseqüentemente, exigirá um repensar dos modelos de negócio sobre os quais, na atualidade, ainda se discutem aspectos pré-*web 3.0*. Desse conjunto, emerge como ponto agregador a (re)valorização da palavra como matéria-prima essencial para a prática informativa no mundo semântico.

Novamente um cenário de incertezas? Algo a discutir, já que o mundo digital da Comunicação parece conviver com a incerteza desde os primórdios da rede, estando esta incorporada em sua rotina. Mas certamente um cenário de extrema valorização da palavra e, portanto, de seus artífices – jornalistas e comunicadores – que têm diante de si a perspectiva de consolidação de seu papel na sociedade digital maquínica que parece emergir. ■

## REFERÊNCIAS

- ANDERSON, C.W. Deliberative, Agonistic, and Algorithmic Audiences: Journalism's Vision of its Public in an Age of Audience Transparency. *International Journal of Communication* 5 (2011), 529–547. ISBN 1932–8036/20110529.
- AUER, Sören; LEHMAN, Jens. Creating knowledge out of interlinked data. *Semantic Web Journal* 1 (2010): 97-104.
- BANNIER, Stijn & VLEUGELS, Chris. How Web 3.0 combines user-generated and machine-generated content. In: A. Alas et al (Eds.), *Transforming Culture in the Digital Age: Conference proceedings* (pp. 69-77). Tartu, 2010.
- BEIGUELMAN, Giselle. O fim do virtual. *Revista Select*. São Paulo, 2011.

# D

## A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria

- BERTOCCHI, Daniela. Narrativas jornalísticas no contexto da web semântica.. In: *Anais do 2º Seminário de Ciberjornalismo de Mato Grosso do Sul*, 2010.
- BOCZKOWSKI, Pablo. *News at work – Imitation in an age of information abundance*. The University of Chicago Press, Ltd., London . Kindle Edition, 2010.
- BRESLIN, John; Passant, Alexandre; DECKER, Stefan. *The Social Semantic Web*. Heidelberg: Springer-Verlag, 2009.
- DI LUCCIO, Flavia. Do Iluminismo à Web Semântica : reflexões sobre a comunicação com base em uma única língua. Tese (Doutorado em Psicologia) – *Pontifícia Universidade Católica do Rio de Janeiro*, Rio de Janeiro, 2010.
- JOHNSON, Clay. *The Information Diet: A Case for Conscious Consumption*. O’Reilly Media, 2012. Kindle Edition.
- KING, Brandy E.; REINOLD, Kathy. *Finding the Concept, Not Just the Word*. A librarian’s guide to ontologies and semantics. Oxford, UK: Chandos, 2008.
- LATAR, Noam Lemelshtrich; NORDFORS, David. Digital Identities and Journalism Content – How Artificial Intelligence and Journalism May Co-Develop and Why Society Should Care. *Innovation Journalism* 6, no. 7 (2009): 47.
- LAURENTIZ, Silvia. Tags e metatags? De Ted Nelson a Tim Berners-Lee. *Revista Porto Arte*. Porto Alegre, May 2010.
- LÉVY, Pierre. *A ideografia dinâmica: rumo a uma imaginação artificial?* Editora Loyola, 1998.
- PARISER, Eli. *The Filter Bubble: What the Internet Is Hiding from You*. Publisher: Penguin Press HC, 2011. Kindle Edition.
- ROGERS, Simon. *Facts are sacred – The power of data*. The Guardian Books, London, Kindle Edition, 2011.
- ROSENBAUN, Steven. *Curation Nation – How To Win In a World Where Consumers are Creators*. McGrawHill, Kindle Edition, 2011.
- SIEGEL, David. *Pull – The power of Semantic Web to transform your business*. Portfolio Hardcover, 2010.
- SKIENA, Steven S. *The Algorithm Design Manual*, 2. Edição, Springer-Verlag London, 2008.
- TEIXEIRA, João de Fernandes. *A mente pós-evolutiva – A filosofia da mente no universo do silício*. Petrópolis, RJ: Ed. Vozes, 2010.
- . *Inteligência Artificial: uma odisséia da mente*. São Paulo, Editora P, 2009.
- TURKLE, Sherry. *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books, 2011.
- WEINBERG, David. *Too Big to Know*. Basic Book. Kindle Edition, 2012.

### Endereços eletrônicos

ADOLPHS, Peter, Xiwen Cheng, Tina Klüwer, Hans Uszkoreit, and Feiyu Xu. Question Answering Geographic Information and Social Network Powered by the Semantic

- Web. In: *Proceedings of the Seventh Conference on International Language Resources and Evaluation*, Valletta, Malta: European Language Resources Association (ELRA), 2010. Disponível em: <<http://www.lrec-conf.org/proceedings/lrec2010/summaries/611.html>>.
- ANGELETOU, Sofia et al. Bridging the Gap Between Folksonomies and the Semantic Web: An Experience Report. In: Workshop: Bridging the Gap between Semantic Web and Web 2.0, *European Semantic Web Conference*, 93, 2007. Disponível em: <<http://kmi.open.ac.uk/people/marta/papers/semnet2007.pdf>>.
- AROYO, Lora; HOUBEN, Geert-jan. User modeling and adaptive Semantic Web. *Semantic Web Journal* 1 (2010): 105-110. Disponível em: <<http://www.semantic-web-journal.net/content/user-modeling-and-adaptive-semantic-web>>.
- BASULTO, Dominic. Are Newspapers Civic Institutions or Algorithms? 16 jan. 2012. <http://bigthink.com/ideas/41976>
- BEIGUELMAN, Gisele. Curadoria de informação. Palestra, ECA-USP, 2011. Disponível em: <<http://www.slideshare.net/gbeiguelman/curadoria-informacao>>.
- BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The Semantic Web. *Scientific American*, 284, no. 5 (May 2001): 34-43. Disponível em: <<http://www.nature.com/doi/10.1038/scientificamericano501-34>>.
- BERNERS-LEE, Tim. Levels of Abstraction: Net, Web, Graph. *W3.org*, 2010. Disponível em: <<http://www.w3.org/DesignIssues/Abstractions.html>>.
- BOURENANE, Amel, SZONIECKY, Samuel and SALEH, Imad. Les perspectives du tagging sémantique, 2009. Disponível em: <<http://www.ieml.org/spip.php?article154>>.
- LATAR, Noam Lemelshtrich, and NORDFORS, David. The Future of Journalism: Artificial Intelligence And Digital Identities. In: Journalism Research & Education Section, *The International Association For Media and Communication Research (IAMCR)*, 1-31. Braga, Portugal, 2011.
- CARR, Nicholas. Nicholas Carr: 2012 will bring the appification of media. Nieman Journalism Lab. Disponível em: <<http://www.niemanlab.org/2011/12/nicholas-carr-2012-will-bring-the-appification-of-media/>>. Acesso em: 4 mar. 2012.
- GABLER, Neal. The Elusive Big Idea. *The New York Times*. 2011. Disponível em: <<http://www.nytimes.com/2011/08/14/opinion/sunday/the-elusive-big-idea.html>>. Acesso em: 14 fev. 2012.
- HITZLER, Pascal, and van HARMELEN, Frank. A reasonable Semantic Web. *Semantic Web Journal* 1 (2010): 39-44. Disponível em: <<http://knoesis.wright.edu/faculty/pascal/resources/publications/HH-ARSW.pdf>>.
- LÉVY, Pierre. From Social Computing to Reflexive Collective intelligence, 2009. Disponível em: <<http://www.ieml.org/IMG/pdf/2009-Levy-IEML.pdf>>.
- LOPES, Ilza Leite. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. *Ciência da Informação* 31, no. 1 (January 2002): 41-52. Disponível em: <<http://>

# D

## A cena cibercultural do jornalismo contemporâneo: web semântica, algoritmos, aplicativos e curadoria

[www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-19652002000100005&lng=pt&nrm=iso&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652002000100005&lng=pt&nrm=iso&tlng=pt).

RABAINO, Lauren. 5 new year's resolutions for newsrooms. *10.000 Words*, 30 dez. 2011. Disponível em: <[http://www.mediabistro.com/10000words/five-new-years-resolutions-for-newsrooms\\_b9610](http://www.mediabistro.com/10000words/five-new-years-resolutions-for-newsrooms_b9610)>. Acesso em: 4 mar. 2012.

SONDERMAN, Jeff. Three trends from 2011 that will reshape digital news in 2012. *Poynter*, 30 dez. 2011. Disponível em: <<http://www.poynter.org/latest-news/top-stories/157639/three-trends-from-2011-that-will-reshape-digital-news-in-2012/>>. Acesso em: 4 mar. 2012.

———. Why The New York Times replaced its Twitter 'cyborg' with people this week. *Poynter*, 24 maio 2011. Disponível em: <http://www.bbc.co.uk/news/technology-14306146>.

TEIXEIRA, João de Fernandes. *O que é inteligência artificial*. São Paulo, 1994. Disponível em: <[http://www.filosofiadamente.org/images/stories/pdf/o\\_que\\_e\\_inteligencia\\_artificial.pdf](http://www.filosofiadamente.org/images/stories/pdf/o_que_e_inteligencia_artificial.pdf)>.

WAKEFIELD, Jane. When algorithms control the world. *BBC News*. Disponível em: <<http://www.bbc.co.uk/news/technology-14306146>>. Acesso em: 14 fev. 2012.

WORLD WIDE WEBB CONSORTIUM, 2004. OWL Web Ontology Language: Use Cases and Requirements. W3C. Disponível em: <<http://www.w3.org/TR/2004/REC-webont-req-20040210/#onto-def>>.

---

Artigo recebido em 12 de março e aceito em 24 de abril de 2012.