



**GEMINIS**

[DOSSIÊ - TV PÓS DIGITAL]

# **CONVERSANDO COM UMA API: UM ESTUDO EXPLORATÓRIO SOBRE TV SOCIAL A PARTIR DA RELAÇÃO ENTRE O TWITTER E A PROGRAMAÇÃO DA TELEVISÃO**

**MÁRCIO CARNEIRO DOS SANTOS**

*Doutorando do programa de Tecnologias da Inteligência e Design Digital da PUC-SP. Mestre em Comunicação pela UAM- SP. MBA em Marketing pelo ISAN-FGV-Rio. Graduado em Comunicação Social pela UFMA. Professor do Curso de Comunicação Social da UFMA na área de Jornalismo em Redes Digitais. Coordenador do Laboratório de Convergência de Mídias – LABCOM ([www.labcomufma.com](http://www.labcomufma.com)).*

*E-mail: [mcszen@gmail.com](mailto:mcszen@gmail.com)*

## RESUMO

Discute-se o fenômeno do uso de redes sociais em paralelo ao consumo de TV, constituindo o que alguns autores chamam de “*backchannel*”, um espaço secundário da produção de conteúdo e discussão sobre os programas que estão sendo assistidos de forma síncrona à sua emissão, formando o que propomos denominar aqui de nuvem de sentido. São exploradas as técnicas de mineração de dados através da ferramenta personalizada Social Tracker, desenvolvida na linguagem de programação Python, para coletar dados sobre o problema, considerando-se que, devido à grande quantidade de informações geradas neste ambiente, é cada vez mais difícil fazer isso manualmente.

**Palavras-Chave:** TV Social, Mineração de Dados, Twitter.

---

## ABSTRACT

In this text we intend to discuss the phenomenon of using social networks in parallel to TV consumption, constituting what some authors called “*backchannel*”, a secondary space of content production and discussion about the programs being watched synchronously to its issuance, forming what we propose to call here a cloud of sense. We explore the techniques of data mining through the tool customized for us by using the Python programming language, called Social Tracker (ST), to study the problem, considering that due to the large amount of information generated in this environment, is increasingly difficult to collect them manually.

**Keywords:** TV Social, Mineração de Dados, Twitter.

## 1.0 – TV E INTERNET

No dia 3 de fevereiro de 2013, quando o Baltimore Ravens venceu o San Francisco 49ers, por 34 a 31, na final da temporada 2012 da liga de Futebol Americano, a NFL, diversos números povoaram o noticiário sobre o evento, conhecido como Super Bowl. Dados da Nielsen (2013), empresa de monitoramento de audiência, indicam que a CBS, rede de TV americana detentora dos direitos de transmissão do jogo, teve uma média de 108,69 milhões de espectadores com um pico de 164 milhões. Segundo a Folha (2013) cada comercial com 30 segundos de exibição custou aos anunciantes 4 milhões de dólares (oito milhões de reais). Houve uma discussão se esse teria sido ou não o recorde de público da história do evento, entretanto, posteriormente foi constatado que a audiência da TV foi menor que nas finais dos dois últimos anos anteriores.

Já na internet os números foram históricos. A transmissão online segundo o site especializado em marketing de vídeo digital REELSEO (2013) atraiu mais de 2 milhões de pessoas e outras métricas, talvez menos divulgadas, trouxeram indícios de que o que chamamos de TV Social ou *Social TV* também tinha conseguido índices expressivos, basicamente através da participação dos espectadores gerando comentários a partir do Twitter.

No seu blog oficial, com o título “The Super Tweets of #SB47<sup>1</sup>”, o Twitter (2013) registrou no mesmo dia os números do Super Bowl. Foram gerados 24.1 milhões de *tweets* durante o evento e o intervalo entre os tempos da partida. O pico do fluxo de mensagens chegou a 185 mil *tweets* por minuto (TPMs). O show da cantora Beyonce, que aconteceu no intervalo da partida, gerou 5.5 milhões de *tweets* com fluxo, no final do show, de 268 mil *tweets* por minuto.

A constatação mais básica a fazer é que um número enorme de pessoas estava assistindo a transmissão pela TV e usando o Twitter, simultaneamente, já que os núme-

---

1 O Twitter utiliza o recurso das hashtags (#) através do qual os usuários podem identificar um tema específico e escrever sobre ele criando um conjunto de conteúdos que se refere a essa palavra que funciona como um rótulo(tag) e pode ser reproduzido a partir de ferramentas de busca.

ros da transmissão online, apesar de terem batido o recorde histórico, chegaram apenas a cerca de 2% da audiência da transmissão televisiva normal ou “off-line”.

Nesse texto pretendemos discutir o fenômeno da utilização das redes sociais em paralelo ao consumo de TV, constituindo o que alguns autores como Proulx e Shepatin (2012) chamam de “backchannel” (canal de fundo), ou seja, um canal secundário de produção de conteúdo e discussão sobre os programas que estão sendo assistidos de forma síncrona à sua emissão, formando o que propomos chamar aqui de nuvem de sentido.

A conversação online sobre um determinado programa acontece antes, durante e depois que ele vai ao ar. O canal de fundo, entretanto, é definido como a conversação em tempo real que está acontecendo através das mídias sociais durante a sua transmissão (PROULX, SHEPATIN, 2012, p.11).

A ideia de nuvem vem do termo *cloud computing* utilizado no jargão da internet para descrever o conjunto de servidores e infraestrutura de rede que possibilita a execução de programas e aplicações de internet, bem como o arquivamento de informação, de forma descentralizada e não no computador do usuário. Hoje mais do que nunca, guardamos nossos arquivos nas nuvens computacionais de grandes empresas como Google, Apple e Microsoft e utilizamos soluções e ferramentas que não estão instaladas em nossas próprias máquinas, mas sim nos servidores deles e de outras companhias que nem conhecemos.

Todo o resultado dos processos de significação, a partir da enorme produção de conteúdo que os usuários da internet agora conseguem realizar, é enviado para esses servidores, entre eles os que rodam as aplicações de mídias sociais. Essa massa de sentido que paira nas nuvens computacionais contém muito do que pensamos e imaginamos sobre tudo e todos, inclusive sobre nós mesmos. Tal conjunto pode ser replicado através de vários outros canais e ferramentas digitais, sendo inclusive disponibilizado via infraestrutura de internet e captado através de APIs (*Application Programming Interfaces* – Interfaces de Programação de Aplicações) como a do Twitter, foco desse trabalho.

A análise e processamento desse material passaram a ser alvo também da atenção acadêmica que ainda explora novos métodos e ferramentas para realizar suas pesquisas sobre o tema. Por isso, aqui apresentamos também um relato de utilização da ferramenta *Social Tracker* (ST), ainda em desenvolvimento, para a coleta de dados a partir da API do Twitter e posterior manipulação e análise dos mesmos, oferecendo possibilidades mais adequadas ao uso acadêmico.

Nossa intenção é argumentar que uma das novas possibilidades interativas da

TV tradicional existe a partir de um processo de remediação<sup>2</sup>, no sentido de Bolter e Grusin (2000) que discutiremos adiante, com a Internet, que tem sido benéfico para ambas as partes, mesmo em contradição às ideias prévias de que os meios digitais impõem perdas ou diminuição de atenção dos espectadores em relação aos meios mais antigos.

Discutimos também como uma prática que implica em atividades síncronas da audiência ao redor de uma transmissão de TV, acontece a partir de um meio que tem normalmente apontada como uma de suas principais características justamente a possibilidade de acessos assíncronos aos conteúdos que disponibiliza.

Por fim exploramos as técnicas de mineração de dados (*data mining*) através da ferramenta por nós customizada utilizando a linguagem de programação Python, chamada de *Social Tracker* (ST), para estudar o problema, considerando que, devido ao grande número de informações geradas nesse ambiente, é cada vez mais difícil coletá-las de forma manual.

## 2.0 – TV SOCIAL E O BACKCHANNEL DE SIGNIFICAÇÃO

O conceito de TV Social ainda é impreciso. Harboe (2011, p.1) fala sobre uma “combinação de tecnologias que permitem experiências sociais ao redor do conteúdo da TV”<sup>3</sup>. Johns (2012, p.333), apesar de não usar diretamente o termo, faz referência à “visualização de duas telas<sup>4</sup> com o uso de redes sociais para criar um canal comunicacional de fundo em tempo real entre os espectadores dos programas de televisão<sup>5</sup>”, identificando essa prática como uma atividade distinta entre as possíveis ações dos fãs no ambiente online. Evangelia (2011, p.1) lembra que “o conceito de TV Social veio à tona nos últimos anos como a fase seguinte da TV Digital Interativa (IDTV), como sua evolução, a partir de um meio que permite a participação ativa dos espectadores para seu engajamento e expressão pessoal.”<sup>6</sup> Já Proulx e Shepatin (2012, p.13) definem *Social TV* como “a convergência entre televisão e mídias sociais”<sup>7</sup>.

2 Para Bolter e Grusin, a remediação entre os meios acontece quando o meio mais novo representa ou assimila o meio mais antigo, incorporando o conteúdo deste último. Os autores, entretanto, admitem a possibilidade de que a remediação também ocorre no sentido inverso, ou seja, o meio anterior incorporar parte ou características do meio mais novo. É o caso da TV aberta que cada vez mais reforça seus links com a internet através da replicação de seu conteúdo em sites, incentivando o diálogo online na web e através das redes sociais, inclusive apropriando-se do conteúdo desses ambientes para divulgação em seus programas de veiculação aberta.

3 “Social television, a combination of technologies that enable social experiences around TV content...” – Tradução do autor

4 “Two-screen veiwing”

5 “Use of social networks to create a real-time backchannel of communication among viewers of television programs ...” – Tradução do autor

6 “The concept of Social TV has come forward over the last years as the next phase of Interactive Digital TV (IDTV) in its evolution from a medium that promotes the viewers’ active participation to one that aims to their engagement and self-expression.” – Tradução do autor.

7 “... the convergence of television and social media.” – Tradução do autor

O fato é que, diferente dos que imaginaram caminhos separados para a TV e a Internet, a tendência da TV Social, reforçada por números como da FORRESTER RESEARCH (2013) indicam o contrário. A partir de uma pesquisa com cerca de três mil americanos adultos que tem atividades online, 48% disseram que usam um computador pessoal enquanto assistem TV para conversar, navegar e pesquisar sobre o que estão assistindo. “Como o número de pessoas engajadas em mídias sociais continua a crescer, o conjunto de conversações online sobre televisão, nessas plataformas, durante a exibição dos programas, também cresce” (Proulx ; Shepatin, 2012, p.27)<sup>8</sup>. O crescimento das vendas de *tablets* e *smartphones* nos grandes mercados do mundo, entre eles o Brasil, também serve como indício de que uma “segunda tela”, móvel e mais leve, de onde se pode interagir nas redes sociais, pode estar bem à mão de um número cada vez maior de espectadores de TV.

O fenômeno de duas mídias que inicialmente pareciam apenas competir pela atenção das pessoas não é inédito. Até a década de 50, principalmente nos Estados Unidos, havia a dúvida de como o cinema, na época com bilheterias decrescentes, sobreviveria ao avanço da televisão nos lares americanos. O que a princípio aparentava ser o pior inimigo de Hollywood, logo se tornou seu principal parceiro comercial. A estratégia do *high concept* e dos *blockbusters*, discutidos abaixo, fez da TV uma alavanca para os lançamentos da indústria do cinema e também um importante cliente das suas produções. A indústria do cinema adaptou-se ao contexto e passou por um processo de hibridização e extensão para áreas antes não exploradas. Essas mudanças começaram com os filmes da indústria americana do cinema pós 75, um período que alguns autores chamam de Nova Hollywood e que Mascarello caracteriza,

[...] pelo abandono progressivo da pujança narrativa típica do filme hollywoodiano até meados de 1960, e também por assumir a posição de carro chefe absoluto de uma indústria fortemente integrada, daí em diante, à cadeia maior de produção e do consumo midiáticos (cinema, TV, vídeo, jogos eletrônicos, parques temáticos, brinquedos, etc.) (MASCARELLO, 2006, p.57).

Esse objetivo de integração é conhecido pelo termo “*high concept*” que também é utilizado para se referir ao modelo de negócios iniciado em filmes desse período, marcado pelo lançamento dos *blockbusters* “Tubarão” de 75 e “Guerra nas Estrelas” e “Embalos de Sábado à Noite”, ambos de 77.

De novo usamos o texto de Mascarello para definir o filme *blockbuster* ou tradu-

8 “As the number of people engaging within social media continues to increase, the amount of online conversations about television while shows are airing within those platforms also increases.” – Tradução do autor

9 Apesar dos autores citados nesse trabalho considerarem como “segunda tela”, *smartphones* e *tablets*, para outros esses equipamentos seriam a terceira e a quarta tela respectivamente, sendo a primeira a TV tradicional e a segunda os *laptops* e *notebooks*. Esse entendimento pode ser encontrado, por exemplo, em relatórios como o da COMSCORE.

zindo para o português, “arrasa quarteirão”:

Filmes que em sua maioria tem custo de produção alto (normalmente por conta de cachês e efeitos especiais), custos de lançamento também elevados e às vezes próximos ou superiores aos custos de produção (em razão do número elevado de cópias e da publicidade massiva) e rápida “queima” do filme no circuito primário de exibição, não importando o quão positivo seja o boca-a-boca, já que eventuais prejuízos de bilheteria, através da lógica do *high concept*, poderão ser compensados nos mercados secundários de exibição, bem como através dos produtos conexos (MASCARELLO, 2006, p.349).

O que Mascarello chama de mercados secundários são em síntese o mercado da exibição nas redes de televisão, nas locadoras e o mercado de licenciamento de produtos baseados em personagens ou filmes para a indústria dos games, dos parques temáticos, de brinquedos e de qualquer outro segmento que tenha interesse em associar sua marca à de um filme de sucesso.

Voltando ao século XXI, a inicial competição entre TV e Internet anunciada como definitiva por alguns, mais uma vez nos mostra que o paradigma de um ecossistema midiático (Scolari, 2008) traduz melhor a situação que as transformações tecnológicas e os ambientes culturais por elas possibilitados (Jenkins, 2006) geraram a partir de uma rede complexa de interações entre diversos atores tais como as redes de TV, os espectadores, cada vez mais interessados e aptos a produzir e fazer circular conteúdo próprio, os aparatos tecnológicos e as soluções de infraestrutura e software que os suportam.

Cada definição de Comunicação está fundada numa metáfora. A Comunicação já foi vista sucessivamente como canal, instrumento, flecha, projétil, conflito, contrato, orquestra, espiral e rede. [...] Neste texto faremos uma aposta muito clara pela metáfora do ecossistema, o seja a Comunicação entendida como um conjunto de intercâmbios, hibridações e mediações dentro de um entorno onde confluem tecnologias, discursos e culturas (SCOLARI, 2008, p.26).

Nos termos de Bolter e Grusin (2000), a Internet, com seu enorme poder de mediação, ao mesmo tempo em que incorporou a linguagem e os produtos dos meios audiovisuais, também foi por eles incorporada, gerando diversos fenômenos e tendências híbridas. A Televisão Social entre elas. Os meios como entes de um sistema complexo acabam interagindo de diversas formas, às vezes não tão fáceis de prever ou identificar. O relativo equilíbrio desse sistema não traduz apatia ou monotonia, e sim pelo contrário, uma dinâmica bastante intensa que dificulta a aplicação de modelos mais

simples em termos teóricos convencionais, traduzida, por exemplo, pela reconfiguração das dinâmicas entre o público e o privado, controle e liberdade e ainda a própria fronteira entre produtores (formalmente, emissores, codificadores) e consumidores (formalmente, receptores, decodificadores) de conteúdo.

O advento das tecnologias recentes ressalta a natureza plural e mutável da ecologia midiática, já que as interfaces da mídia ordenam o acesso aos conteúdos entregues pelos canais digitais, a cabo, via satélite, através da Web, em reprodutores de mídia pessoais, através de DVDs e, finalmente, por meio de transmissão digital. Ao mesmo tempo, a ampla distribuição de interfaces de mídia traz consigo a criação de novas categorias televisuais e midiáticas com base no acesso e competência, que se tornam particularmente atraentes quando mapeadas contra as flexíveis estratégias de *microcasting*<sup>10</sup> que já estão resultando em articulações distintas e diferenciadas de plataforma e conteúdo (CHAMBERLAIN, 2012, p.20)<sup>11</sup>.

### 3.0 - TV SOCIAL E A SINCRONICIDADE NO CONSUMO DE CONTEÚDO

O surgimento específico da TV Social também incorpora uma aparente contradição entre processos síncronos e assíncronos de acesso ao conteúdo. Faz parte da retórica positiva e do imaginário ligado à internet, sua capacidade, por exemplo, de possibilitar o consumo de conteúdo, desvinculado de rígidos controles como as grades de programação das televisões tradicionais. A ideia do consumo *on demand*, em qualquer hora ou lugar, incorpora bem essa característica que também traduz a crença de um aumento de poder do espectador que, com a Internet, poderia agora realizar coisas antes quase impossíveis, como o simples desejo de ver um programa que foi exibido num horário em que não estava perto de um aparelho de TV.

Apesar das evidentes possibilidades que as redes e a infraestrutura da internet trazem, transfigurando-se num imenso acervo digital a ser consultado, a qualquer hora ou de qualquer lugar (desde que exista a conexão), algumas questões devem ser lembradas.

Uma das mais importantes é a questão do controle. A partir da inerente possibilidade de identificação e registro da navegação online, a partir de diversas ferramentas de análise de tráfego, métricas e monitoramento, é possível afirmar que houve uma redução na homogeneidade ou regularidade das formas de entrega de conteúdo, mas um importante aumento no esforço de caracterização e análise sobre esses consumidores, facilitado principalmente pelas formas específicas de manipulação e tráfego dos arquivos digitais. O controle foi talvez perdido em um aspecto do processo, o da distribuição, mas evidentemente intensificado em outro, o do con-

<sup>10</sup> Difusão de conteúdo para públicos cada vez mais segmentados.

<sup>11</sup> The advent of recent technologies underscores the plural and mutable nature of the media ecology, yet media interfaces order engagements with content delivered through digital cable, satellite, over the Web, on personal media players, through DVDs, and ultimately through digital broadcast. At the same time, the broad distribution of media interfaces brings with it the establishment of new televisual and media divides based on access and competency, which become particularly compelling when mapped against strategies of flexible *microcasting* that are already resulting in distinct and differentiated articulations of platform and content. Tradução do autor.

sumo, provavelmente ainda em nível inferior ao que as novas possibilidades de tratamento de grandes massas de dados parece delinear. A recente utilização de tecnologias de “*Big Data*” na reeleição do presidente Obama nos Estados Unidos, é um exemplo.<sup>12</sup>

#### 4.0 - TWITTER E TRANSMISSÕES AO VIVO

A partir da classificação estabelecida por Hansen, Sheiderman e Smith (2011, pos. 771) para as mídias sociais, definidas por eles como “o conjunto de ferramentas online que dão suporte a interações sociais entre seus usuários”, o Twitter está na categoria Blogs<sup>13</sup> e Podcasts<sup>14</sup> e na subcategoria “Microblogs e Fluxos de Atividades”.

Parecido com os blogs tradicionais devido ao seu foco em publicações recentes, difere deles porque seus posts, chamados de tweets, são restritos a 140 caracteres de texto. O Twitter recupera a ideia de acompanhar o fluxo das postagens através da subscrição como nos blogs a partir da possibilidade de “seguir” um outro usuário. Um canal personalizado de determinado usuário mostra os tweets mais recentes de todos os indivíduos que ele está seguindo, criando um fluxo de pequenos pedaços de informação (HANSEN;SHEIDERMANN;SMITH,2011, e-book pos.1210).

Além da ideia de seguir alguém (*follow*), o usuário do Twitter também pode ser seguido e trabalha com um conjunto de possibilidades utilizadas para organizar o ambiente da comunicação tais como *hashtags*<sup>15</sup>, *retweets*<sup>16</sup>, *réplicas*<sup>17</sup>, *menções*<sup>18</sup>, *trends*<sup>19</sup> e mensagens diretas<sup>20</sup> entre outras.

12 A tecnologia de Big Data foge ao escopo desse texto. Sobre o exemplo citado, ver a matéria da INFO de Dezembro de 2012, “Big Brother OBAMA” sobre a utilização da ferramenta pela equipe do presidente Barack Obama nas eleições americanas em <http://info.abril.com.br/arquivo/2012/dez.shtml>. Acessado em 13/02/2013.

13 Forma especial de página na internet caracterizando-se por uma forma de baixo custo para publicação de conteúdo digital apresentado como uma série de mensagens de texto, chamadas de “posts” em ordem cronológica reversa.

14 Publicação na internet parecida com o blog com a diferença fundamental de, ao invés do texto, utilizar arquivos de som para o registro das mensagens.

15 Palavra chave que serve como uma espécie de rótulo (*tag*) para marcar conteúdos sobre um mesmo assunto ou tópico.

16 Forma utilizada com o termo “RT” iniciando a mensagem e indicando sua origem a partir de outro usuário que é replicada através do retweet por ser considerada interessante por aquele que a “retuitou”.

17 Forma de se referir a um outro usuário, indicando uma conversa específica com ele, representada a partir da inserção do caractere “@” antes do seu nome.( Ex: @mcszen ). É uma forma de indicar, dentro do fluxo público que contém várias mensagens, de que aquela mensagem, especificamente , se direciona a alguém. Também chamada de marca de endereçamento.

18 Forma de citar determinado usuário na mensagem sem necessariamente estar dialogando com ele naquele momento. Também representada com o caractere “@” antes do nome do usuário mas não colocada no início da mensagem , o que caracterizaria uma réplica (*reply*).

19 Tópicos sobre os quais há grande interesse ou conversa no fluxo de mensagens do Twitter. Os trends podem ser identificados por região geográfica o que ajuda a conhecer os temas de maior atenção no Twitter em determinado momento e em determinada área.

20 Mensagem direcionada especificamente para outro usuário fora do fluxo geral de mensagens.

Para as pesquisas sobre TV Social utilizando o Twitter como referência, os números têm indicado que os maiores picos de interação acontecem a partir do comentário sobre transmissões ao vivo ou, quando pré-gravadas, inéditas (sem veiculação anterior).

Números do balanço do Twitter no Brasil em 2012 indicam que entre os 60 picos repentinos de interesse no ano (*top trends*) a categoria entretenimento recebeu aproximadamente 28% do total, esportes, 13%, e política, em torno de 9%. No entretenimento, as novelas e os *reality shows* tiveram uma enorme presença no Twitter: de 1º de março ao dia 31 de outubro de 2012, o termo Avenida Brasil foi recordista com milhões de *tweets* e, em seu último capítulo, apresentou um fluxo de 3.031 *tweets* por minuto. Da mesma forma, a *hashtag* #BBB12 foi recordista entre os realities, seguido do The Voice Brasil, A Fazenda e Mulheres Ricas (Twitter Blog, 2012).



**Figura 1** – Gráfico com números de Tweets por Minuto em eventos nacionais.  
Fonte: LABCOM

Nos eventos internacionais os números são ainda maiores como o caso do *SuperBowl* já citado, bem como a transmissão de debates da eleição americana na TV, os jogos da Eurocopa (futebol) e dos jogos Olímpicos de Londres.



**Figura 2** – Gráfico com números de Tweets por Minuto em eventos mundiais.  
Fonte: LABCOM

## 5.0 – FERRAMENTA SOCIAL TRACKER E COLETA DE DADOS DE EVENTOS NACIONAIS

Uma API – *Application Programming Interface* (Interface de Programação de Aplicações) é o conjunto de rotinas, padrões e instruções de programação que permite que os desenvolvedores criem aplicações que possam acessar determinado serviço na internet. O exame dessa questão está além dos limites desse texto mas o principal é saber que no caso do Twitter, através da sua API<sup>21</sup>, é possível criar aplicações que possam acessar os computadores do mantêm o serviço e assim coletar informações diversas sobre o fluxo de conteúdo que é gerado pelas pessoas que estão conectadas a ele.

Em especial, a API do Twitter e a política de privacidade desse serviço são bastante abertas. Hoje existem centenas de aplicações que se “alimentam” das informações geradas nele, incluindo aí as mais simples como buscas (*search*) de *trends*, informações sobre os usuários e muitas outras.

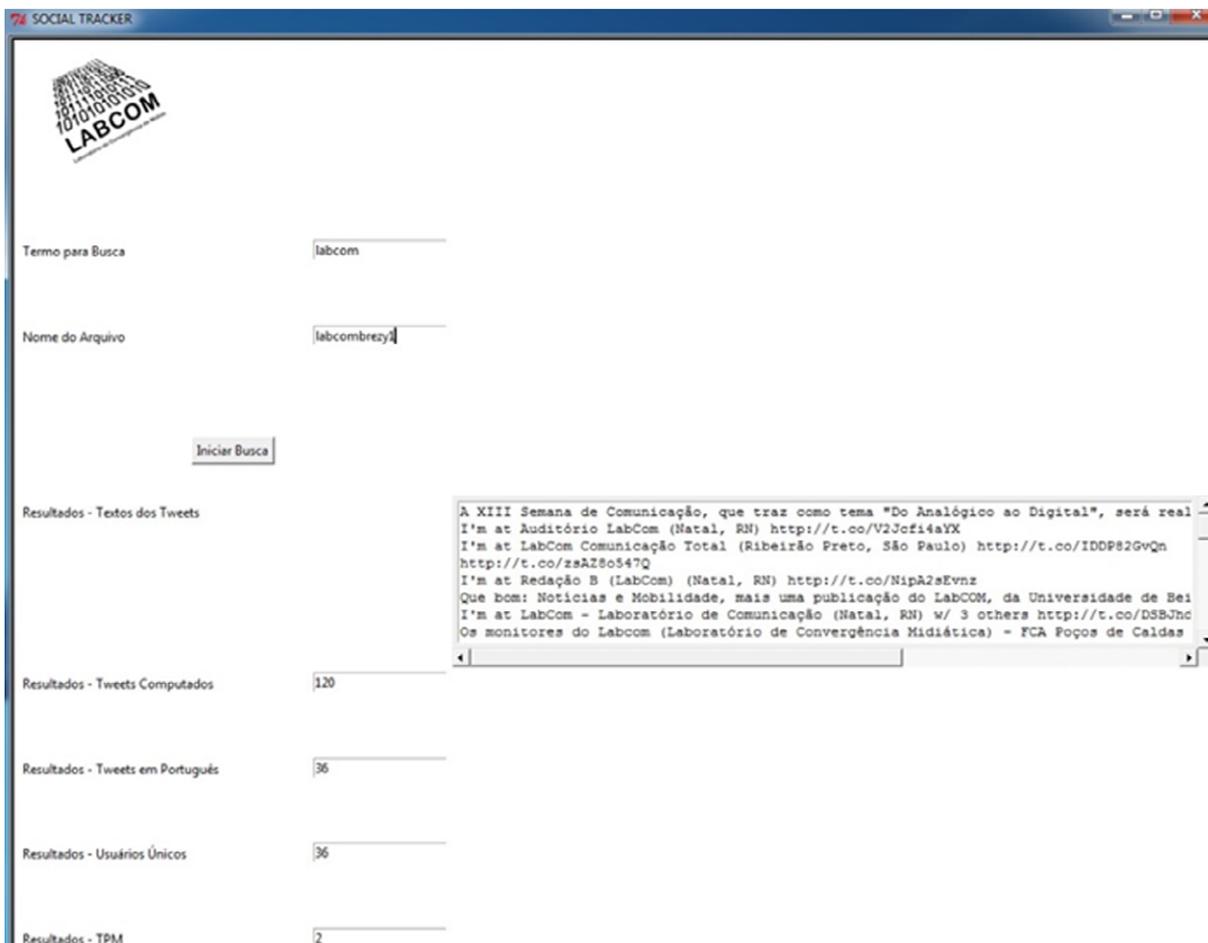
Uma lista básica dos principais recursos disponibilizados pela API do Twitter inclui também, além das citadas acima, a recuperação de mensagens diretas (*direct messages* – DM), amigos e seguidores (*friends*, *followers*), sugestões de usuários (*suggested users*), favoritos (*favorites*), listas (*lists*), buscas salvas (*saved searches*), lugares e dados geográficos (*places* e *geo*), relatório de mensagens automáticas ou indesejáveis (*spam reporting*), autenticação (*Oauth* – *Open Authentication*) e ajuda (*help*).

Na atual versão da API (1.1) há cerca de 100 tipos de requisições que podem

<sup>21</sup> <https://dev.twitter.com/docs/api/1.1>

ser acessadas nos grupos citados que permitem um amplo espectro de coleta de informações.

Nas pesquisas relacionadas à TV Social com abordagem quantitativa é fundamental a capacidade de coletar quantidades de mensagens relativamente grandes e depois conseguir processá-las também de alguma forma automatizada. A partir dessa constatação foi desenvolvida a ferramenta ST capaz de fazer consulta à API do Twitter através das diversas modalidades por ela permitidas. A primeira etapa do desenvolvimento foi focada no módulo de buscas (*search*) que além de retornar mensagens sobre determinado tema ou usuário, permite customizações por data e localização, além de não necessitar de autenticação prévia para o acesso.



**Figura 2** – Print de uma das telas da ferramenta ST

Fonte: LABCOM

No exemplo abaixo, colhido através da ferramenta ST, é possível ver a estrutura da resposta do servidor do Twitter a partir de uma solicitação de busca (*search*) sobre o próprio nome da ferramenta: “Social Tracker”. Para essa coleta foi feita uma postagem de teste por volta das 18 horas do dia 17 de fevereiro de 2013 e em seguida feita uma

coleta sobre o mesmo termo.

A primeira imagem mostra a postagem na forma tradicional dentro do fluxo de *tweets* da conta do LABCOM – Laboratório de Convergência de Mídias, responsável pelo desenvolvimento da ferramenta e, abaixo dela, segue a resposta original da API do Twitter com informações sobre a mesma postagem só que na forma estruturada com todos os possíveis detalhes sobre ela. Nem todos os campos estão completos porque alguns dados realmente não foram inseridos, como a menção sobre algum outro usuário, por exemplo.



**Figura 4** – Imagem do tweet de teste sobre Social Tracker feita na conta do LABCOM-UFMA  
Fonte: LABCOM

Na segunda imagem foram marcados alguns campos para destacar como as informações são apresentadas, como o tema da busca (*query*), a data da postagem com o horário no fuso padrão GMT que é 3 horas a frente do horário brasileiro (*created at*), o usuário que postou (*from\_user*), seu código de identificação (*from\_user\_id*), a língua da postagem (*iso language\_code*), o endereço da imagem do perfil (*profile\_image url*) e o texto da mensagem (*text*).

```

{
  "completed_in": 0.069,
  "max_id": 303248101462908928,
  "max_id_str": "303248101462908928",
  "next_page": "?page=2&max_id=303248101462908928&q=Social%20Tracker&rpp=100",
  "page": 1,
  "query": "Social+Tracker",
  "refresh_url": "?since_id=303248101462908928&q=Social%20Tracker",
  "results": [
    {
      "created_at": "Sun, 17 Feb 2013 21:02:36 +0000",
      "from_user": "mcazen",
      "from_user_id": 149853725,
      "from_user_id_str": "149853725",
      "from_user_name": "MCS",
      "geo": null,
      "id": 303248101462908928,
      "id_str": "303248101462908928",
      "iso_language_code": "pt",
      "metadata": {
        "result_type": "recent"
      },
      "profile_image_url": "http://a0.twimg.com/profile_images/1460824435/LabComReleituraFinal2_normal.JPG",
      "profile_image_url_https": "https://s10.twimg.com/profile_images/1460824435/LabComReleituraFinal2_normal.JPG",
      "source": "<a href='\"http://twitter.com/\"'>web</a>",
      "text": "Teste para a ferramenta Social Tracker desenvolvida pelo LABCOM para an\u00e9lise de conte\u00fado do Twitter.",
      "to_user": null,
      "to_user_id": 0,
      "to_user_id_str": "0",
      "to_user_name": null
    }
  ],
}

```

**Figura 5** – Imagem da resposta da API do Twitter recuperada por ST

Fonte: LABCOM

A vantagem de uma ferramenta desenvolvida de forma customizada é que ela normalmente permite um maior conjunto de possibilidades de coleta e manipulação de dados que a maioria das aplicações disponíveis. É preciso usar várias delas para conseguir o mesmo resultado. As ferramentas prontas obviamente tem uma interface gráfica mais trabalhada, entretanto, para a pesquisa acadêmica, o material coletado é o que importa.

A ferramenta ST ainda está em desenvolvimento mas já consegue coletar grande quantidade de *tweets* a partir de requisições de busca (*search*) para a API do Twitter. Algumas das funcionalidades já disponíveis são exatamente o levantamento em separado da massa de textos (mensagens) acoplados aos resultados da busca. Além disso, ST ainda permite verificar, a partir da amostra coletada, o fluxo em *Tweets* por Minuto (TPM) dos tópicos postados com maior frequência, além de verificar as principais fontes de *retweets* e menções.

## 6.0 - ST E A MEDIÇÃO DA ATIVIDADE SOCIAL NAS TRANSMISSÕES DE NOTICIÁRIO E DO CARNAVAL

### 6.1 - METODOLOGIA

Ainda que em caráter exploratório foram realizadas medições em dois tipos de transmissão televisiva. Foram escolhidas duas modalidades de transmissão. A primeira categoria, a partir do noticiário sobre um tema de grande repercussão e a segunda, a partir da transmissão ao vivo de um evento também de grande interesse, no caso, o desfile das escolas de Samba do Rio de Janeiro.

A medição a partir do noticiário aconteceu no dia 27 de janeiro de 2013, dia do incên-

dio de uma casa de shows em Santa Maria no Rio Grande do Sul, que teve intensa cobertura da mídia. Acompanhamos a cobertura da Rede Globo, rede de maior audiência nacional, que iniciou-se de maneira mais articulada a partir do programa Esporte Espetacular onde a partir do estúdio os apresentadores esportivos tiveram que assumir o contato e as chamadas dos correspondentes na região da tragédia. Apesar de não ter nenhum relacionamento com o tipo de programa que estava no ar, a importância do evento na avaliação da emissora alterou totalmente o andamento normal do programa, para que pudesse ser feita a inserção do noticiário sobre o incêndio na transmissão da TV.

No período da manhã, iniciando às 11 horas, até o final do Esporte Espetacular, por volta das 12:20, foram feitas 10 medições para acompanhar a variação do fluxo de *tweets* por minuto (TPMs) a partir da repercussão das informações que iam sendo dadas durante a transmissão na TV. Posteriormente, a noite, durante o Fantástico fizemos mais 5 medições para, num outro programa com transmissão ao vivo a partir do estúdio da emissora, termos uma ideia da expansão ou retração do interesse no Twitter sobre o tema.

A segunda medição aconteceu no início das transmissões do desfile das Escolas de Samba do Rio de Janeiro, nos dias 10 e 11 de fevereiro, acompanhando os desfiles do Salgueiro e da Unidos da Tijuca no primeiro dia e da Mangueira no segundo dia. A opção pelas primeiras escolas da transmissão deve-se ao fato de que, a partir do momento que avançamos pela transmissão durante a madrugada, o número de telespectadores vai diminuindo e conseqüentemente o número de *tweets* sobre o tema.

## 6.2 – RESULTADOS E CONSIDERAÇÕES FINAIS

Os dados coletados através de fontes secundárias bem como os resultados dessas medições estão publicados numa tabela no site do LABCOM. (<http://www.labcom-ufma.com/socialtv.htm>) .

No caso do acidente de Santa Maria foi observado um crescente interesse sobre o tema no Twitter à medida que a cobertura dentro do programa Esporte Espetacular foi acontecendo. É óbvio que outros programas em outras emissoras também ajudaram nesse fato. As medições iniciais começaram indicando um fluxo de 41 *tweets* por minuto, número que foi crescendo para alcançar um pico de 174 ao final do programa.

É importante lembrar que a API do Twitter responde a consultas informando apenas uma parte do fluxo total que recebe o que implica em dizer que os números apontados são ainda maiores. Os números indicados são uma amostra do universo de mensagens processadas pelo Twitter mas servem para apontar uma tendência de crescimento nesse intervalo de tempo.

Nas medições da noite, durante o programa Fantástico, apesar de já se terem

passado mais de 12 horas do início da cobertura da manhã, a intensidade do interesse foi diluída, com os fluxos variando em torno de 40 TPMs um pouco antes e depois da exibição das matérias sobre o acidente.

É impossível afirmar com certeza o que levou a essa redução. É fato que durante o programa da manhã havia a novidade da matéria, as pessoas estavam acordando no domingo e sendo pegadas de surpresa com as informações sobre o incêndio. Havia o choque inicial sobre a dimensão da tragédia que foi sendo delineada ao longo da cobertura com números e detalhes cada vez mais impactantes, bem como a intensidade das referências na programação ao fato, constituindo-se em algo bastante parecido com uma transmissão ao vivo sobre o evento, já que praticamente toda a duração do programa foi dedicada à cobertura de Santa Maria. Novidade, percepção crescente da gravidade da situação e intensidade do número de referências ao evento talvez possam explicar o reflexo no Twitter.

As transmissões da noite tinham uma situação distinta. O fato tinha sido amplamente divulgado por todos os canais e as pessoas de forma geral já conheciam a extensão dos danos causados. A saturação sobre o tema amplamente explorado durante todo o dia por todos os veículos de informação e a diluição da cobertura, agora já intercalada com outros assuntos, já não reforçavam a percepção das três características apontadas no parágrafo anterior. São hipóteses apenas que precisam ser mais aprofundadas.

Já na transmissão do Carnaval feita pela Rede Globo, os números chegaram a 200 TPMs com cerca de 1 hora após o início do desfile do Salgueiro (a primeira escola, não do desfile, mas a aparecer na cobertura da emissora), no primeiro dia das transmissões.

Esses números mais que dobraram (451 TPMs) após o início do desfile da Unidos da Tijuca, conhecida pela utilização de muita criatividade em suas comissões de frente e carros alegóricos. As dificuldades enfrentadas pela escola, que teve carro quebrado, princípio de incêndio e integrante passando mal, tudo isso durante o desfile, aparentemente refletiram na intensidade dos comentários do Twitter que chegaram a um pico de 684 TPMs ao final da apresentação.

O mesmo fenômeno (a relação das dificuldades enfrentadas pela escola durante o seu desfile e o crescimento do fluxo de mensagens do Twitter) foi constatado também no segundo dia durante o desfile da Mangueira. Com uma hora de transmissão as medições obtidas usando ST chegavam a 683 TPMs. Entretanto nos minutos finais do desfile ficou claro que a escola estava com dificuldades para terminar sua apresentação no tempo estipulado, o que ocasionaria perda de pontos na avaliação dos jurados e con-

seqüentemente diminuição das chances de título. A transmissão começou a enfatizar fortemente esse problema e o fluxo do Twitter quase que instantaneamente começou a subir com as pessoas comentando justamente essa situação. O pico medido ficou em 1356 TPMs, bem acima da média do dia anterior.

Como foi dito acima, os estudos usando a ferramenta ST estão ainda sendo feitos apenas em caráter exploratório já que o próprio software está incompleto. Apesar disso na situação atual já é possível coletar não só os textos das mensagens, como também aplicar algumas métricas básicas como o número de *tweets* por minuto, a listagem dos *retweets* e menções, as palavras mais frequentes e a divisão da massa coletada a partir da língua utilizada, bem como o número de usuários únicos nas postagens e a definição de um período para a coleta das publicações, resguardados os limites da própria API que retorna em média apenas o conteúdo dos últimos sete dias anteriores a data da consulta (*search*).

Novas funcionalidades vão permitir que outras modalidades de consulta sejam implementadas incluindo aí um módulo baseado em NLTK (*Natural Language Toolkit*<sup>22</sup>) ferramenta que trabalha com textos para que análises voltadas à parte de conteúdo das mensagens também possa ser feita.

A vantagem de uso da ST é justamente a de gerar arquivos em texto mais acessíveis para a análise dos pesquisadores bem como possibilitar a inserção dos mesmos em outras ferramentas complementares como planilhas e bancos de dados.

Utilizando as definições de Manovich (2006) sobre as características dos produtos das mídias digitais, a que ele chama de transcodificação, ou seja, a estruturação dos objetos de mídias digitais em duas camadas, uma cultural, de significação e outra lógica, a partir da forma com que as máquinas tratam as coisas; permite imaginar que as pesquisas em TV Social também precisam abordar e coletar informações sobre essas duas camadas, de forma a permitir uma compreensão sistêmica desses fenômenos. A ferramenta ST foi pensada com essa intenção para que o diálogo com as APIs possa ser feito de forma mais fácil e efetiva.

A recente associação entre Twitter e Nielsen para a criação de um novo conjunto de métricas de audiência na TV também baseada nos comentários do Twitter indica que a TV Social já conseguiu também bastante atenção do mercado e que novos formatos e modalidades híbridas de utilização e consumo do conteúdo da TV estão por vir.

---

22 [www.nltk.org](http://www.nltk.org)

## REFERÊNCIAS

- BOLTER, Jay; GRUSIN, Richard. **Remediation: Understanding New Media**. Cambridge: The MIT Press, 2000.
- CHAMBERLAIN, Daniel. Media interfaces, networked media spaces, and the mass customization of everyday space. In: KACKMAN, Michael; BINFIELD, Marnie; PAYNE, Matthew Tomas et al. (Eds.). **FlowTV**. New York: Routledge, 2011. Epub[Livro Eletrônico].
- EVANGELIA, Mantzari. **The Implementation of Social Interactive Television** . In: uxTV2008. Disponível em : <http://uxtv2008.org/program/social-tv-workshop-papers/evangelia.pdf> . Acessado em 17/04/2013.
- FORRESTER. **Use social media to boost your tv audience – A Social Computing Report**. Disponível em: <http://www.forrester.com/search?tmtxt=%20social%20TV#/Updated+Q1+2012+Use+Social+Media+To+Boost+Your+TV+Audience/quickscan/-/E-RES61227> . Acessado em 17/04/2013.
- Haboe, Gunnar . **The TV is watching you** . In: uxTV2008. Disponível em : <http://uxtv2008.org/program/social-tv-workshop-papers.html> . Acessado em 17/04/2013.
- HANSEN, Derek; SHNEIDERMAN, Ben; SMITH, Marc. **Analyzing social media networks with NodeXL: insights from a connected world**. Burlington-MA: Elsevier, 2011.
- JENKINS, Henry. **Convergence culture: Where old and new media collide**. New York: NYU Press, 2006.
- JOHNS, Mark . *Two Screen Viewing and Social Relationships. Exploring the invisible backchannel of TV viewing*. In: **Proceedings Cultural Attitudes Towards Communication and Technology 2012**. Murdoch University, Murdoch, 2012. pp. 333-343. Disponível em : <http://sammelpunkt.philo.at:8080/2159/> . Acessado em 17/04/2013.
- MANOVICH, Lev. **The Language of New Media**. Cambridge: MitPress, 2002.
- MASCARELLO, F. Cinema hollywoodiano contemporâneo. In: MASCARELLO, F. (Org.). **História do Cinema Mundial**. Campinas: Papirus, 2006.
- NIELSEN. **Super bowl xlvi: how we watch and connect across screens** . Disponível em : <http://www.nielsen.com/us/en/newswire/2013/super-bowl-xlvi-draws-108-7-million-viewers-26-1-tweets.html> . Acessado em 17/04/2013.

PROULX, Mike; SHEPATIN, Stacey. **Social TV: how marketers can reach and engage audiences by connecting television to the web, social media, and mobile**. New Jersey: John Wiley&Sons, 2012. Epub[LivroEletrônico].

REELSEO. **The Super Bowl Online Stream Amassed A Record Amount Of Viewers**. Disponível em : <http://www.reelseo.com/super-bowl-online-stream/> . Acessado em 17/04/2013.

SCOLARI, Carlos. **Hipermediaciones** - Elementos para una Teoría da Comunicación Digital Interactiva. Barcelona: Gedisa Editorial, 2008.

FOLHA. **Super Bowl de 2013 bate recorde de audiência** . Disponível em : <http://www1.folha.uol.com.br/esporte/1225569-super-bowl-de-2013-bate-recorde-de-audiencia.shtml> . Acessado em 17/04/2013.

TWITTER. **The super tweets of SB47**. Disponível em : <http://blog.twitter.com/2013/02/the-super-tweets-of-sb47.html> . Acessado em 17/04/2013.

TWITTER BLOG. **Este ano no Twitter**. Disponível em: <http://blog.pt.twitter.com/> . Acessado em 17/04/2013.