



Análítica Quantitativa Criativa Aplicada ao Ensino e à Prática do Jornalismo de Dados para Alunos da Graduação em Comunicação Social¹

Ricardo FOTIOS²

Escola Superior de Propaganda e Marketing, São Paulo, SP

RESUMO

A quantidade de dados disponíveis nas redes virtuais é inusitadamente grande e o jornalismo investigativo ganhou importantes aliados com as novas ferramentas de busca e de tabulação de informações do *big data*. Nos últimos dez anos, pelo menos duas grandes reportagens tomaram conta do noticiário global baseadas, exclusivamente, na divulgação de dados: *Wikileaks* e *Swissleaks*. No entanto, ter acesso aos bancos de dados não estabelece vínculo direto com a precisão e a pertinência do material jornalístico. Cabe ao novo profissional de comunicação saber como manipulá-los e interpretá-los adequadamente. Para tanto, este trabalho propõe uma visão científica dessas informações no ensino e na prática do Jornalismo de Dados (JD), possibilitando análise quantitativa e abordagem criativa na execução da reportagem assistida por computadores (RAC).

PALAVRAS-CHAVE: jornalismo de dados; RAC; big data; análise; redes virtuais

1. INTRODUÇÃO

Desde meados dos anos 1990 que o jornalismo profissional se beneficia das reportagens assistidas por computadores. À época, muitas redações já contavam com seus próprios bancos de dados digitais e possibilitavam ao profissional de comunicação algum tipo de interação com as máquinas e seus arquivos, estimulando sinapses jornalísticas em que os repórteres pudessem coordenar e cruzar informações diferentes para chegar a outras inéditas (TOLEDO, 2011).

Mas foi com a ampliação da internet e de suas redes virtuais que a RAC ganhou escala e importância. Em 2003, a Associação Brasileira de Jornalismo Investigativo (Abraji) inaugurou um curso sobre o tema e, em 2004, o jornal *O Globo* venceu o Prêmio Esso com “Os Homens de Bens da Alerj”, reportagem sobre o enriquecimento de deputados estaduais coordenada pela jornalista Angelina Nunes. A produção foi totalmente baseada nos dados encontrados nos arquivos do Tribunal Regional Eleitoral (TRE) do Rio de Janeiro. Este foi o primeiro reconhecimento em premiação nacional da pertinência informativa dos bancos de dados.

¹ Trabalho apresentado no DT 06 – Interfaces Comunicacionais do XX Congresso de Ciências da Comunicação na Região Sudeste, realizado de 19 a 21 de junho de 2015.

² Professor Licenciado da Universidade Metodista de SP, Aluno Especial da Disciplina CRP5215 - Comunicação Digital e Educação a Distância da ECA-USP, e-mail: fotios@bol.com.br.



O acesso aos poucos e precários bancos de dados existentes até meados de 2004 era feito de forma fragmentada e com boa parte do trabalho realizado de maneira tradicional, com pouco auxílio de recursos tecnológicos avançados ou da internet. O principal método de investigação de dados estava debruçado na dedicação pessoal com boa dose de faro jornalístico. O repórter não dispunha de métodos ou processos científicos para orientar e conduzir a tarefa de garimpagem de notícias em meio aos dados de um banco.

A partir da segunda metade dos anos 2000, a expansão do conceito *Web 2.0* e seus princípios, responsáveis por tornaram a internet do acervo estático em plataforma multitarefas, passaram a permitir a manipulação de banco de dados diretamente dos terminais pessoais. A quantidade de informação disponível na rede mundial de computadores aumentou de forma exponencial.

“Hoje as notícias estão fluindo na medida em que acontecem, a partir de múltiplas fontes, testemunhas oculares, blogs, e o que aconteceu é filtrado por uma vasta rede de conexões sociais, sendo classificado, comentado e, muito frequentemente, ignorado” (LORENZ, 2011).

A grande concentração de dados nas redes virtuais e as ferramentas técnicas para sua exploração, no entanto, ainda esbarram no perfil pouco analítico do jornalista. Voltado para as ciências humanas, o contato deste profissional com a matemática esteve limitado às áreas da estatística. Mesmo assim, o repórter se dedicava muito mais a repercutir os dados com suas fontes do que a buscá-los e interpretá-los. Com a velocidade demandada pela informação em tempo real e a emergência das redes virtuais, o jornalista se vê diante de um ambiente rico de informações, mas que ele aproveita pouco por falta de familiaridade e até por certo desprezo pelos cálculos e análises. Ainda é muito comum ouvir de estudantes de comunicação que dizem ter escolhido a área exatamente porque não gostam e não sabem manipular números. Argumentos tais que não encontram respaldo na realidade das redações, onde o Jornalismo de Dados é frequentemente acionado.

Exemplos deste novo interesse são as reportagens conhecidas por Wikileaks e Swissleaks, iniciadas em 2006 e 2013, respectivamente, em que vazamentos de grandes bancos de dados de empresas, instituições financeiras e governos do mundo todo permitem inúmeras pautas nas redações ao redor do planeta. O momento informativo “é o maior da história do jornalismo. Existirá um antes e um depois desse instante”³.

³ Jornal El País, 8 de fevereiro de 2010.



“A filosofia do Wikileaks repousa sobre um princípio central: os segredos são feitos para serem desvendados. Toda informação escondida tem, assim, a vocação de ser atualizada e colocada ao alcance dos cidadãos” (RAMONET, 2012, p. 89).

Buscando quebrar esta barreira invisível que limita o comunicólogo contemporâneo no aprendizado do JD, este trabalho foi buscar nas ciências analíticas os métodos e os processos que possam orientar e contribuir para que as práticas do bom jornalismo sigam sendo prioridade na evolução das reportagens baseadas no *big data*.

2. OBJETIVO

Primeiramente, o propósito de trazer a analítica quantitativa para o mundo do JD é fornecer um processo acessível aos estudantes de comunicação, possibilitando a visão abrangente da pauta e dos dados que vão abastecê-la, uma vez que “o *big data* e a analítica nele baseada prometem mudar praticamente todos os setores de atividade e todas as funções de negócios nos próximos dez anos” (DAVENPORT, 2014, p.7).

Um segundo e não menos importante objetivo é o de envolver a análise criativa dos dados neste método quantitativo, retirando com o devido cuidado e proporção a obrigatoriedade da exatidão das fórmulas, que tanto afastam os repórteres das ciências matemáticas, e colocando em seu lugar a obrigatoriedade de correção das informações, tema mais próximo e mais facilmente assimilado pelo repórter.

Estas duas metas principais, associadas às práticas rotineiras de softwares clientes e online, como tabulação, infografia e técnicas de otimização para busca orgânica (*search engine optimization - SEO*), produzem um novo cenário para esta disciplina nas salas de aula. O reconhecimento de linguagens e formatos de dados também contribui para a readequação do jornalista em um contexto que valoriza a fonte primária de dados em detrimento daquela previamente manipulada pelas fontes, além de inserir o futuro profissional na utilização da internet e redes virtuais para além do entretenimento pessoal e da simples consulta. O repórter passa ele própria a formar bancos de dados.

A análise direta de dados comportamentais dos brasileiros em arquivos do IBGE (Instituto Brasileiro de Geografia e Estatística), por exemplo, pode revelar informações inusitadas e desconhecidas do público e que não aparecem nos resumos e *press kits* do órgão. Tal análise em primeiro nível rompe com a lógica do filtro oficial e estabelece novos critérios que ficam



sob o controle do jornalista. Traz, adicionalmente, revelações ampliadas e regionalizadas, comumente desprezadas pelo noticiário nacional.

3. O PROCESSO ANALÍTICO

Não é a primeira vez na história da comunicação que a matemática e o raciocínio lógico servirão de suportes para processos informacionais. Nos anos 1840, quando da instalação dos telégrafos em larga escala, nos Estados Unidos e na Inglaterra, a era da informação rápida e objetiva se iniciava. A mensagem, que antes levava dois dias para chegar ao destinatário, poderia alcançar seu destino em poucos segundos pelos fios telegráficos. Não por acaso, foi também a estreia dos bancos de dados, pois empresas telegráficas tentaram manter registros de todas as mensagens enviadas. “Tratava-se de um armazenamento de informações sem precedentes” (GLEIK, 2013, p.158)

O arquivamento de conteúdo e a agilidade da transmissão não foram as únicas demandas informativas do telégrafo. Cientistas do mundo todo passaram a elaborar complexos modelos de codificação de mensagens para reduzir a quantidade dos caracteres dos textos e poder pagar menos por sua transmissão. Não menos necessária, a criptografia também vinha resolver a questão da privacidade do conteúdo transmitido, no caso de comunicação pessoal. Têm origem naquela época os boletins de ações, loterias e meteorológicos utilizados até hoje em jornais, telejornais, websites e demais veículos de comunicação.

“Os fenômenos da atmosfera, os mistérios dos meteoros, as causas e os efeitos das combinações no céu não são mais questões de superstição e pânico para o marido, o marinheiro e o pastor” (ARCHER, 1848 apud GLEIK, 2013, p. 156).

Desde então, a matemática e a lógica alcançaram muitas conquistas científicas e a codificação dos tempos atuais se faz de forma quase automática, por meio de algoritmos. Mas, para trazer a lógica matemática para o JD, foi preciso investigar empiricamente alternativas de análise de dados, numéricos ou não, a fim de estabelecer um formato novo e dedicado à disciplina emergente. O método avaliado foi o de três estágios e seis passos da analítica quantitativa e a primeira aplicação se deu na turma do sétimo semestre dos cursos de jornalismo da Escola Superior de Propaganda e Marketing (ESPM-SP) e da Universidade Metodista de São Paulo (Umesp), em 2014.

O primeiro estágio deste processo analítico quantitativo aplicado ao JD foi o da Formulação e nele estavam contidos os dois primeiros passos propostos: o reconhecimento do problema e a



revisão das descobertas anteriores. O primeiro passo assemelha-se ao já conhecido modelo de pauta jornalística em que se lista perguntas pertinentes à investigação dos dados; o segundo passo consiste em avaliar quais das perguntas listadas no passo anterior já foram respondidas no passado e se as respostas encontradas têm valor para a análise atual.

Ao avaliar um histórico de desemprego em uma determinada região, por exemplo, o histórico dos meses e anos anteriores podem ser fundamentais para a reportagem encontrar pontos de divergências e de convergências com os dados atuais, o problema novo a ser investigado. É a informação nova gerada a partir da informação conhecida.

Do segundo estágio, nomeado de Solução do Problema, constaram três outros passos: a modelagem, a coleta de dados e a análise dos dados. Antes de buscar os dados propriamente ditos, este processo prevê uma listagem de variáveis possíveis com lacunas a serem preenchidas no passo seguinte, o de coleta de dados. Desta forma, o repórter não se sente buscando uma agulha no palheiro diante de planilhas com milhares de dados, números, nomes, pontos percentuais e afins. O passo da análise de dados neste estágio diz respeito à avaliação das informações coletadas, a comparação entre as variáveis sugeridas na modelagem e sua evolução entre uma variável e outra.

Seguindo na reportagem hipotética das taxas de desemprego citada anteriormente, o segundo estágio da analítica elaborada para o JD trata de estabelecer quais os setores, períodos, faixas etárias, gêneros e demais perfis de público a ser investigado devem ser confrontados e como as variáveis devem ser preenchidas com dados de bancos primários e, posteriormente, comparados para identificar informações relevantes.

O terceiro estágio, o da Comunicação, compreendeu na apresentação dos resultados. Este passo tende a ser o mais confortável para o futuro jornalista, pois trata da reportagem propriamente dita e quais recursos deve utilizar para comunicar melhor as descobertas do segundo estágio. O uso de gráficos, tabelas, ilustrações e demais argumentos informacionais devem aparecer nesta última etapa do processo.

Após a aplicação do processo analítico dos seis passos nas duas turmas de Jornalismo da ESPM-SP e da Umesp, o método foi atualizado para as aulas de 2015 com conexões mais efetivas com as abordagens criativas inerentes à elaboração da pauta, apuração das informações e produção do material jornalístico. Assim, o processo analítico foi associado a quatro níveis metodológicos: Preparação, Imersão, Incubação e Ideia.



A construção dos fundamentos para a solução do problema está contida na fase da Preparação e pode estar presente nos três primeiros passos analíticos: reconhecimento do problema, revisão das descobertas anteriores e modelagem.

Já o nível da Imersão sugere engajamento intenso com a solução do problema e com os dados disponíveis e pode ser invocada nos passos da modelagem, coleta de dados e análise de dados. Opcionalmente, pode existir o nível da Incubação, que é a internalização do problema na mente na expectativa de que se desenvolvam prováveis conexões inusitadas abaixo do nível da consciência, percorre o passo da análise de dados quando este encontra alguma barreira informativa ou de interpretação.

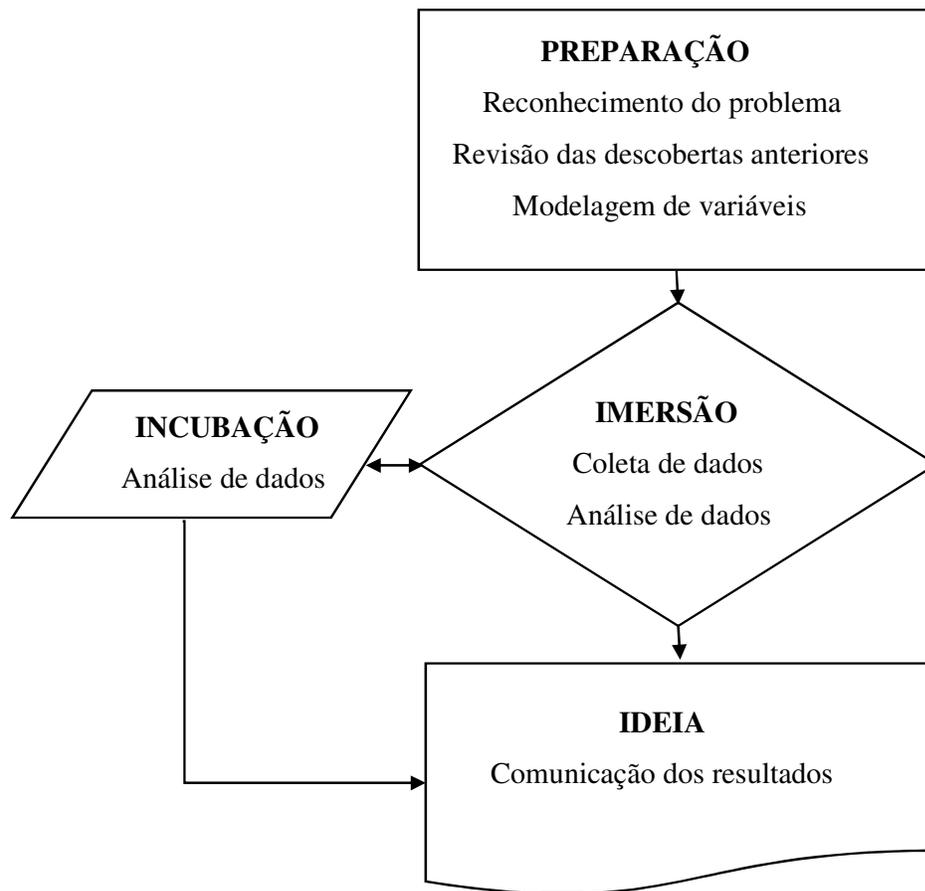
Finalmente, o nível da Ideia representa o grande avanço na compreensão de como o problema pode ser resolvido por meio da análise quantitativa. Eclode no passo da análise de dados e se manifesta com vigor na apresentação dos resultados.

4. ANALÍTICA CRIATIVA

O modelo de processo analítico quantitativo criativo descrito anteriormente foi aplicado com sucesso em diversos casos de uso em sala de aula, notadamente os de análise comportamental a partir da Pesquisa Nacional por Amostra de Domicílios (Pnad), na evolução do setor automotivo com banco de dados da Associação Nacional dos Fabricantes de Veículos Automotores (Anfavea) e no acompanhamento dos títulos cinematográficos e suas bilheterias no país descritos pela Agência Nacional do Cinema (Ancine). Em nenhuma das atividades práticas os estudantes precisaram recorrer a assessorias de imprensa ou a outras fontes secundárias de obtenção de dados e informações.

A compreensão das informações contidas nos dados analisados evoluiu de forma mais rápida e natural quando os seis passos analíticos foram seguidos em comparação a turmas de estudantes que não contavam com processos para a condução da RAC. Da mesma forma, a resistência a fórmulas e cálculos foi fortemente reduzida quando o valor daquele resultado era percebido como fundamental para fechar o ciclo da análise e, conseqüentemente, da finalização da reportagem investigativa, sem o prejuízo da naturalidade do material final.

Adicionalmente, foi possível perceber uma melhor aceitação de programas de computadores como *Excel*, *Analytics* e *Access*. Desta forma, com o vínculo entre teoria e prática devidamente estabelecido, foi possível produzir o diagrama estrutural do processo analítico quantitativo criativo como segue:



5. CONSIDERAÇÕES FINAIS

De posse de um processo específico e idealizado para a prática do Jornalismo de Dados, que une métricas matemáticas de avaliação aos roteiros subjetivos da produção intelectual, alunos de Jornalismo e futuros repórteres ultrapassam a barreira da carência de ferramentas diagnosticadas no início da era dos bancos de dados, nos anos 1990 e meados dos anos 2000 nos modelos informacionais. Ao mesmo tempo que não pretende ser um método definitivo nem exclusivo, a analítica aplicada ao JD abre uma nova competência da comunicação social voltadas para a investigação de fatos noticiosos, em que a fonte primária deixa de ser um entrevistado ou autoridade, mas o conjunto de dados que estes personagens geram rotineiramente simplesmente por acessar, gravar informações ou consultar as diversas redes virtuais disponíveis no *big data*.



6. REFERÊNCIAS

RAMONET, I. **A explosão do jornalismo**: das mídias de massa às massas de mídias. São Paulo: Ed.Publisher, 2012.

GLEICK, J. **A informação**: uma história, uma teoria, uma enxurrada. São Paulo: Ed. Companhia das Letras, 2013.

DAVENPORT, T. **Dados demais!** Como desenvolver habilidades analíticas para resolver problemas complexos, reduzir riscos e decidir melhor. São Paulo: Ed. Elsevier, 2014.

SCHÖNBERGER, V M; CUKIER, K. **Big Data**: como extrair volume, variedade, velocidade e valor da avalanche de informação cotidiana. São Paulo: Ed. Elsevier, 2013.

TOLEDO, J.R. O que é RAC e qual a sua importância? **Blog Amanhã no Globo**, Rio de Janeiro, jul. 2011. Disponível em: <<http://oglobo.globo.com/blogs/amanhanoglobo/posts/2011/07/06/o-que-rac-qual-sua-importancia-390735.asp>>. Acesso em: 10 mai. 2015.

LORENZ, M.. Por que jornalistas devem usar dados? **Manual de Jornalismo de Dados**, São Paulo, mar. 2012. Disponível em: <http://datajournalismhandbook.org/pt/introducao_1.html>. Acesso em: 10 mai. 2015.