

COMUNICAÇÃO E HARD SCIENCES: Possibilidades epistêmicas para o quantitativismo em ciências da comunicação¹
COMMUNICATION AND HARD SCIENCES: Epistemological possibilities of mathematical methods in communication sciences

Luiz Guilherme Antunes²
Daniel Torres Guinezi³

Resumo: No presente artigo buscaremos dimensionar as decorrências epistemológicas, metodológicas e práticas do uso instrumental de metadados gerados por tecnologias digitais, em especial por Smartphones, na pesquisa científica em Ciências da Comunicação. A proposta toma como referência histórica a transformação ocorrida nas Ciências Econômicas a partir do século XIX, época em que o Marginalismo surge, utilizando-se da moeda como instrumento de medição que transforma a economia em uma ciência aplicada. De forma análoga, os metadados coletados por serviços de informação digital registram intenções, sentimentos e demandas sociais que podem ser utilizadas nas Ciências Sociais para a construção de novos modelos teóricos, com o potencial de deslocar a posição e o status atuais do campo. Mais do que uma simples aplicação de técnica, tal proposta implica em profundas transformações na maneira como a pesquisa é realizada hoje em dia, e pode levar a perigosas distorções se não for usada adequadamente.

Palavras-Chave: Epistemologia; Big Data; Comunicação

Abstract: This article seeks the epistemological, methodological and practical consequences on the usage of Metadata generated by Mobile Technologies in the scientific research of Communication Sciences. The proposal uses the transformation occurred in Economical Sciences from the 19th century as historical reference, a time when the usage of the coin as a measurement tool shifts the economy into an applied Science. Similarly, the metadata collected by digital information services register intentions, sentiments and social demands that might be used in the applied social sciences in order to build new theoretical models with the potential of shifting the position and the current status of the field. More than a technical application, this proposal implies deep transformations in the way that the research is currently make, leading to possible biases if not adequately used.

Keywords: Epistemology; Big Data; Communication

¹ Trabalho apresentado ao Grupo de Trabalho Epistemologia da Comunicação do XXVI Encontro Anual da Compós, Faculdade Cásper Líbero, São Paulo - SP, 06 a 09 de junho de 2017.

² Professor Livre Docente do Programa de Pós-graduação em Ciências da Comunicação da Escola de Comunicações e Artes da Universidade de São Paulo (USP) e coordenador do Grupo de Pesquisa Datacracia. Email: luli@luli.com.br

³ Mestrando do Programa de Pós-Graduação em Ciências da Comunicação da Escola de Comunicações e Artes da Universidade de São Paulo (USP) e integrante do Grupo de Pesquisas Datacracia. Email: daniel.guinezi@usp.br

1. Introdução

Os meios de comunicação digital ampliam as possibilidades de investigação no campo das ciências da comunicação. Podemos avaliar este processo em termos tanto de objeto quanto de métodos.

Enquanto objeto, a internet constitui um novo meio de comunicação e interação. Possui um dinamismo próprio distinto daquele imposto através dos meios de comunicação analógicos como televisão, jornal e rádio. Isso faz da comunicação em sua modalidade digital um objeto à parte, aberto para infindáveis possibilidades analíticas. O que caracteriza o processo comunicativo no meio digital? Qual é a nova relação estabelecida entre receptores e produtores de conteúdo? De que forma o caráter interativo altera a natureza da relação que temos com a informação? Abre-se todo um horizonte de questionamentos que se enquadram facilmente na pesquisa do campo.

A tecnologia digital traz também possibilidades de inovação metodológica para o campo. O acesso facilitado aos componentes do *córpus*, que ficam registrados em vídeos do Youtube, postagens do Twitter, notícias em portais ou imagens do Instagram, traz, por si só, novas possibilidades reflexivas e analíticas. Mas mais do que isso, o comportamento do usuário no ambiente digital gera os chamados metadados. Metadados são, em essência, dados sobre dados. Isto é, são “a informação que nós criamos, armazenamos e compartilhamos ao interagir com objetos digitais” (RILEY, 2017, pg.1) podendo usufruir deles no teste de hipóteses, na elaboração de modelos teóricos ou até em reflexões de caráter crítico:

Se escuta música no Spotify, fotos são postadas no Instagram, vídeos localizados no Youtube; usam o Quicken para gerir as finanças pessoais; o email, as mensagens de texto e as redes sociais para se comunicar com os outros; nos celulares, armazenam longas listas de contatos. Todo esse conteúdo gera metadados - informação sobre a criação, o nome, tópico e os features daquele conteúdo. Metadados são fundamentais na funcionalidade de sistemas que entregam conteúdo, possibilitando que os usuários encontrem itens de interesse, armazenando informações essenciais sobre eles, e compartilhando estas informações com outros. (RILEY, 2017, pg. 2)⁴

⁴ No original: “People listen music in Spotify, post photos in Instagram; locate videos in Youtube; manage finance through Quicken; connect with others via email, text, and social media; and store lengthy contacts list in their mobile devices. All of this content comes with metadata - information about item’s creation, name, topic, features and the like. Metadata is key to the functionality of the systems holding the content, enabling

É com a possibilidade de matematização das pesquisas no campo através de metadados que este artigo se ocupa. Para isso procuraremos exibir as potencialidades que se abrem com a aproximação do campo à ciência de dados, bem como ressaltar as diversas limitações impostas pelo objeto comunicacional contemporâneo quando se realizam tentativas de descrição quantitativa dos modelos.

A hipótese inicial, assumindo provisoriamente enquanto objeto da comunicação o indivíduo em relação à mídia, é de que a mudança da mídia analógica para a digital muda também a natureza dos modelos teóricos com os quais o campo pode operar. Tal hipótese é proveniente do modelo quadripolar da geração de conhecimento, segundo o qual “a geração de conhecimento se processa em quatro níveis ou polos: epistemológico, teórico, metodológico e técnico” (MARTINS, 2009, pg. 4). Entendendo que o uso de metadados constitui um deslocamento no polo técnico, devemos dimensionar as consequências disso para os outros três, já que constituem instâncias interligadas. Uma mudança técnica gera invariavelmente transformações metodológicas que, por sua vez, exigirão reinterpretações de caráter teórico e epistemológico para a investigação no campo.

De maneira geral, avaliamos que a partir desse deslocamento podemos migrar de um modelo de *soft science*, pautado em uma matriz analítica de caráter metafísico e puramente dedutiva, para outro mais indutivo, que sem abandonar a complexidade reflexiva incorpora hipóteses passíveis de falseabilidade empírica. Trata-se de um deslocamento que só é possível a partir do momento em que nosso objeto migra para o mundo digital, uma esfera onde comportamentos e dinamismos sociais deixam rastros, sendo estes passíveis de mensuração e modelagem matemática.

Essa conclusão parte de um paralelo do campo da comunicação com outro ramo das ciências sociais aplicadas que matematizou-se no final do século XIX: a economia. A avaliação comparada dos dois campos não só é rica por mostrar o que possibilitou o desenvolvimento da economia como uma *Hard Science*, mas também por ressaltar diferenças fundamentais entre a natureza dos objetos de cada campo, diferenças estas que constituem problemáticas profundas para quaisquer tentativas de se lidar com o objeto comunicacional a partir de abordagens quantitativas.

users to find items of interest, record essential information about them, and share that information with others” (RILEY, 2017, pg. 2) - tradução livre dos autores.

2. Ciências “duras” e “moles”

Thomas Kuhn, em *A Estrutura das Revoluções Científicas* (1994), propõe uma divisão na estrutura lógica do avanço científico. De acordo com o autor, o desenvolvimento da ciência tem início em uma fase caracterizada por divergências acerca do objeto, dos métodos e princípios a serem adotados. Assim que surge consenso, os princípios acordados seriam então transformados em *paradigmas*. À medida que a comunidade científica apresenta divergências aos princípios paradigmáticos, algumas proposições são falseadas e substituídas por novas.

Segundo Kuhn, a revolução científica ocorreria apenas a partir do momento em que há mudança no paradigma vigente, quando há ampla aceitação dos novos princípios, que passam a substituir os anteriores. A troca dos princípios até então vigentes no campo por novos constitui um exemplo de avanço da fronteira do conhecimento. Uma vez que se tem contato com a fronteira, não é necessário se estudar o processo histórico que levou a ela, já que a nova teoria teria como pressuposto a incorporação ou refutação das propostas anteriores.

A proposta de Kuhn é desenvolvida por Arida (1996) ao propor uma divisão da pesquisa científica entre o que o autor denomina *Hard Science*, ligada à tradição acadêmica anglicana, e *Soft Science*, relacionada à academia francesa.

Modelos teóricos de *Hard Science*, usualmente ligados às ciências exatas e naturais, costumam não estar temporalmente localizados. Isso, porque eles ignoram a história do pensamento, lidando com o que se chama de “fronteira do conhecimento”. A “fronteira do conhecimento” seria o conjunto de teorias em seu estado mais desenvolvido. Sob essa lógica, cada novo avanço teórico implica abandono do estágio anterior de conhecimento, que seria de alguma forma invalidado.

De outro lado, a comunicação enquanto ciência social aplicada opera muito mais próxima ao polo da *Soft Science*. Isso, porque “o modelo *Soft Science* desconhece a noção de fronteira” (ARIDA, 1996, pg.17). No geral, os modelos teóricos da área de comunicação constituem interpretações, pensamentos muito ligados à chave de leitura que os autores possuem. Trata-se de um entendimento do conhecimento que insere cada teoria em seu contexto de produção, já que este é determinante para as proposições. Além disso, são por

vezes proposições de caráter metafísico, que não configurariam saber científico na noção popperiana por não serem passíveis de falseabilidade.

De acordo com a proposição de Arida, a física constituiria um exemplo de *Hard Science* por excelência, uma vez que o treinamento para se tornar físico não requer o conhecimento da história da formação das ideias. Pelo contrário, os manuais de física possuirão tudo o que se precisa saber, já que ao aprender o que existe de mais atual na teoria o aluno teria internalizado automaticamente o conjunto de avanços científicos que o precederam, incorporados no status presente da teoria. Estudar a história das ideias seria, no caso de teorias consolidadas, mera questão de curiosidade, não constituindo-se como exigência para dominar o campo. Tal estudo faria sentido apenas em ocasiões que houvessem dúvidas com relação à validade da teoria. Nesse caso, o conhecimento do contexto poderia contribuir para verificar (ou falsear) propostas.

Áreas que operam mais próximas ao polo da *Hard Science* costumam proceder segundo o racionalismo crítico apresentado por Karl Popper em *A Lógica da pesquisa científica* (1975). Segundo a definição Popperiana, o conhecimento somente pode ser caracterizado como científico estritamente quando feito através de proposições passíveis de falseabilidade. O progresso das ciências, de acordo com essa visão, somente poderia ocorrer a partir do sucessivo falseamento do(s) modelo(s) vigente(s), cuja objetividade precisa ser garantida tanto pela replicabilidade dos procedimentos metodológicos como pela previsibilidade dos resultados.

De acordo com Popper:

Um sistema seja suscetível de ser dado como válido, de uma vez por todas em sentido positivo; exigirei, porém, que sua forma lógica seja tal que se torne possível validá-lo através de recurso a provas empíricas, em sentido negativo: deve ser possível refutar, pela experiência, um sistema científico empírico (POPPER, 2001, pg.42)

Tal visão colide com o desenvolvimento das ciências humanas e sociais, uma vez que as interações humanas, como já defendia David Hume no século 18, tendem a ser complexas e mutuamente dependentes, a ponto de não serem passíveis de determinação ou falsificação:

A dependência mútua dos homens é tão grande, em todas as sociedades, que raramente alguma ação humana é inteiramente completa por si só, ou mesmo

realizada sem alguma referência às ações de outros, que são os requisitos [necessários] para a resposta completa á intenção do agente.(HUME, 2007, pg.64)⁵

O componente humano ficaria, então, a cargo das *Soft Sciences*. Segundo Arida (1996), dois princípios estruturantes do conhecimento construído por *Soft Science* o diferenciam daquele constituído pelas *Hard Sciences*: a reconciliação problemática e a intradução. Por reconciliação problemática, entende-se que as matrizes básicas de cada teoria são irreconciliáveis, de modo que “entre a matriz A e B, é problemático tanto contrapor A com B quanto sintetizá-las” (ARIDA, 1996, pg.17). Já por intradução expressa-se a ideia de que “toda transcrição de matrizes básicas envolve perdas irremediáveis” (ARIDA, 1996, pg.18). Isto é, uma vez que o contexto de produção constitui fator determinante nas proposições, é natural que apenas no próprio autor tenha-se a plena construção conceitual. Quaisquer tentativas de tradução implicariam perdas irremediáveis.

A partir destes dois princípios, entendemos que a pesquisa que pretende engendrar-se no conhecimento caracterizado por *Soft Science* precisaria familiarizar-se com boa parte das matrizes básicas (e suas eventuais contradições), como também recorrer aos clássicos daquele campo.

A título de exemplo, levemos em consideração três linhas teóricas da comunicação: a Teoria Administrativa norte-americana, a Escola de Frankfurt e os Estudos Canadenses. Pelo princípio da reconciliação problemática pode-se perceber que trata-se de campo muito mais próximo à *Soft Science*. Enquanto a teoria administrativa recorre ao *behaviorismo* para entender melhor os efeitos da comunicação de massa (e a partir disso buscar torná-la mais eficiente), a Escola de Frankfurt parte de uma matriz marxista para descrever o processo de reificação desempenhado pelos meios de comunicação no que Adorno denominaria “Indústria Cultural”. Por fim, os estudos canadenses de Marshall McLuhan remontam a Harold Innis e à noção de centralidade dos meios técnicos para a organização social. Tentativas de atualizar uma teoria a partir de proposições originárias de outra escola incorrerão em problemas. Não há cenário em que uma proposição da escola administrativa, que toma como unidade analítica o indivíduo, possa desconstruir ou mesmo atualizar

⁵ No original: “The mutual dependence of men is so great, in all societies, that scarce any human action is entirely complete in itself, or is performed without some reference to the actions of others, which are requisite to make it answer” (HUME, 2007, p. 64) - tradução livre dos autores.

concepções frankfurtianas, já que estas, por sua vez, partem da noção marxista de classe na construção da teoria crítica.

A irreconciliação de métodos (*behaviorismo* vs. materialismo histórico dialético), de unidade analítica (indivíduo vs. classes sociais), ou mesmo de finalidade do conhecimento (prático vs. crítico) impossibilita o progresso desse campo da ciência nos moldes de Popper ou Kuhn. Isso não diminui sua importância nem impede que essas três linhas de pensamento constituam o núcleo do que denomina-se Teoria da Comunicação, exibindo como de fato trata-se de um campo próximo à *Soft Science*.

Feita essa diferenciação inicial entre as duas vertentes do conhecimento científico, a questão que se coloca é: o que pode ocorrer com as ciências da comunicação diante da emergência de métodos computacionais e quantitativos? Como isso altera o próprio objeto da produção teórica no campo? Seria o quantitativismo responsável por um deslocamento em direção à *Hard Science* ou o surgimento de mais uma matriz analítica a conviver com as demais?

3. Uma trajetória de matematização

Durante muito tempo os estudos em economia estavam próximos à filosofia, versando sobre a natureza humana, a origem da riqueza, a estrutura do sistema econômico, a divisão do trabalho, o valor dos bens. Enquanto ramo da filosofia, era conhecida ainda por “Economia Política”, estando por muito tempo ligada a uma matriz de reflexão dedutiva.

Marcam esse período os Fisiocratas franceses, tendo em Quesnay seu principal nome, e o início da Economia Clássica, onde despontam Adam Smith, David Ricardo, John Stuart Mill, Thomas Malthus e tantos outros. Apesar de profundas distinções, fisiocratas e clássicos comungavam um mesmo entendimento da Economia Política: seria o ramo da ciências morais que busca as origens da riqueza e da sua distribuição. Esse caráter fica evidente nas palavras de Smith:

Economia política, considerada um ramo da ciência do estadista ou do legislador, propõe dois objetos distintos: primeiro, suprir renda ou produtos em abundância para o povo, ou, mais apropriadamente, possibilitar que provenham tal renda ou provento por si sós; e segundo, suprir o Estado ou *Commonwealth* com uma renda suficiente para os serviços públicos. Ela se propõe a enriquecer tanto o povo quanto o soberano. (SMITH, 1983, pg.37)

Apesar da convergência de propósitos, as propostas econômicas ainda estavam bastante próximas da definição de *Soft Science*. A corrente Fisiocrata, por exemplo, tinha como princípio fundamental o raciocínio de que toda riqueza era proveniente da terra. Isso significa que a terra, e apenas ela, é capaz de criar riqueza em uma nação visto que “o excedente é um dom da natureza” (HUNT, 2005, pg.35). Sendo a agricultura a única fonte de riqueza, o Estado deveria deixar funcionarem livremente as trocas na sociedade para distribuir essa riqueza, tributando estritamente as atividades ligadas à terra.

Por outro lado, a teoria do valor de Adam Smith defendia que “o pré requisito para qualquer mercadoria ter valor era que ela fosse o produto do trabalho humano” (HUNT, 2005, pg.47). O valor de uma dada mercadoria é dado segundo a quantidade de trabalho por ela comandada, sendo a manufatura, e não a terra, a atividade que cria valor em uma sociedade.

O trabalho era o primeiro preço, o dinheiro da compra inicial que era pago por todas as coisas. Não foi com o outro nem com a prata, mas com o trabalho que toda a riqueza do mundo foi inicialmente comprada. (SMITH, 1983, pg. 30)

Este é um dos vários exemplos a evidenciar as contradições inerentes ao uso do método científico (na concepção popperiana) no estudo de humanidades. Não há proposição da corrente clássica a respeito da origem do valor que possa ser comparada com uma proposição fisiocrata, já que cada uma parte de pressupostos básicos diferentes, nenhuma das proposições é falseável, nem tampouco parte de operações indutivas de natureza empírica. Constituem, portanto, modelos criados para descrever a criação de valor na sociedade a partir de operações dedutivas de caráter lógico e metafísico.

A trajetória da Economia começa a mudar a partir da linha econômica chamada *marginalista*. Com base na teoria utilitária de Jeremy Bentham (1969), essa linha de pesquisa desloca a unidade analítica da economia de um modelo pautado na noção de classe, para outro centrado no indivíduo, chamado de individualismo metodológico. Ao descrever a sociedade como um conjunto de indivíduos maximizadores de utilidade, é dado um passo fundamental em direção a uma ciência de base mais empírica e mais quantitativa.

De acordo com essa proposta de enfoque, a sociedade deixaria de ser um objeto abstrato e sujeito às mais variadas especulações para se transformar em um coletivo de unidades individualizadas. Ainda assim, o individualismo metodológico e o utilitarismo não

são suficientes na transformação da Economia Política em uma ciência de caráter quantitativo nos moldes popperianos. A Utilidade, por constituir uma medida abstrata, ainda não poderia ser auferida em termos absolutos. Para contornar o problema, economistas passaram a buscar vias alternativas. Uma das mais relevantes para a nossa proposta é a ideia de *marginalismo* – a substituição das medidas absolutas de utilidade, que não poderiam ser facilmente medidas, por variações marginais da utilidade nas trocas de mercadorias. Ao focar a variação em detrimento do valor absoluto, torna-se mais fácil comparar diferentes decisões de consumo, já que estas passam a ter base não na utilidade total, mas na utilidade marginal gerada por cada novo bem adquirido.

A utilidade deve ser considerada como medida idêntica à adição feita à felicidade de uma pessoa. (...) Devemos, assim, discriminar cuidadosamente a utilidade total gerada por qualquer bem, da utilidade associada a alguma fração particular deste bem. A utilidade total do alimento que comemos consiste em sustentar a vida, e pode ser considerada elevada. Contudo, se subtraímos um décimo de tudo o que comemos diariamente, nossa perda seria mínima. (JEVONS, 1965, pg.45)⁶

Para tornar a medida marginal mensurável, era necessário um instrumento de medida universalmente aceito. A moeda, por sua disponibilidade e facilidade de mensuração, tornou-se o instrumento preferencial. Outras unidades de valor, como reputação, confiabilidade, risco e retorno a longo prazo surgiram posteriormente, em sua maioria baseados na troca de mercadorias com base no valor mutuamente aceito da moeda.

O objeto da Economia, vale ressaltar, é a avaliação de tomadas de decisão em contextos específicos, normalmente no que se convencionou chamar de *mercados*. Estes seriam caracterizados pela existência de dois tipos de agentes: os que compram e os que ofertam. O principal elemento de mediação da relação entre esses agentes é a moeda. A priori, qualquer ação fora dos mercados não diria respeito aos estudos econômicos, tampouco decisões que não podem ser medidas em unidades monetárias. Em outras palavras, o objeto de estudo não é o indivíduo, mas o seu comportamento em situações de interação de mercado.

⁶ No original: “Utility must be considered as measured by, or even identical with, the addition made to a person’s happiness. (...) We must now carefully discriminate between the total utility arising from any commodity and the utility attaching to any particular portion of it. The total utility of the food we eat consists of maintaining life, and may be considered as great. But if we were to subtract a tenth part of what we eat daily, our loss would be slight.” (JEVONS, 1965, pg.45) - tradução livre dos autores.

O principal motivo pelo qual o objeto econômico adquire caráter quantitativo é a restrição de sua análise à decisões intermediadas por moeda. Nas palavras de Alfred Marshall:

Economia é um estudo dos homens tal como vivem, agem e pensam nos assuntos ordinários da vida. Mas diz respeito, principalmente, aos motivos que afetam de um modo mais intenso e constante a conduta dos homens na parte comercial da sua vida. (...) o motivo mais constante para a atividade dos negócios é o desejo da remuneração, a recompensa material do trabalho. Essa remuneração poderá ser gasta egoística ou altruisticamente, para fins nobres ou mesquinhos, e nisto influi a variedade da natureza humana. Mas o motivo é dado por uma determinada soma de dinheiro: e é esta exata e determinada medida em dinheiro que permitiu à Economia avançar sobre os demais ramos do estudo do homem. (MARSHALL, 1920, pg.33)

Há que se ter clareza da moeda enquanto instrumento apto a aferir indiretamente grandezas econômicas. Recorrendo novamente a Marshall, “é essencial notar que o economista não se arroga a possibilidade de medir diretamente as inclinações do espírito, mas só indiretamente através de seus efeitos” (MARSHALL, 1920, pg.34). Com base nesse pressuposto os economistas criaram uma forma de ciência baseada em modelos matemáticos, utilizando-se sempre de valores monetários como via para análise dos fenômenos de natureza econômica.

A partir da virada marginalista, o processo de matematização das ciências econômicas avança de maneira acelerada, culminando no que se denomina Economia Moderna. Esta, por sua vez, caracteriza-se fundamentalmente por operar de maneira muito mais próxima à uma *Hard Science* do que a Economia Política⁷. São abandonados do escopo econômico os debates de caráter moral e político, convertendo-se em uma ciência que busca fundamentalmente, nos termos de Lionel Robbins (1945), encontrar as alocações mais eficientes de recursos para fins que sejam dados. A economia torna-se uma ciência aplicada.

Cálculo, estatística e econometria convertem-se nas ferramentas fundamentais desta nova versão do economista, que a partir delas passa a aplicar modelos e testar hipóteses. O individualismo metodológico consolida-se enquanto núcleo da teoria econômica, o pressuposto que norteia o avanço da ciência. Com ele, o problema da irreconciliação de

⁷ A diferenciação da Economia Política para Economia aqui trabalhada parte da divisão feita por Marshall (1920), que contrapõe o caráter filosófico e moral da primeira ao aplicado da segunda.

métodos é superado, possibilitando a criação de uma fronteira do conhecimento que aproxima a economia das ciências naturais.

4. Em busca de um instrumento

Novas maneiras de se pensar e de conviver estão sendo elaboradas no mundo das telecomunicações e da informática. As relações entre os homens, o trabalho, a própria inteligência dependem, na verdade, da metamorfose incessante de dispositivos informacionais de todos os tipos. (...) Não se pode mais conceber a pesquisa científica sem uma aparelhagem complexa que redistribui as divisões antigas entre experiência e teoria. (LÉVY, 1993, pg.4)

Tecnologias digitais podem propiciar ao campo da comunicação novos instrumentos de medição e previsão, capazes de promover uma transformação de magnitude similar à ocorrida pelo uso do valor marginal da moeda nos estudos econômicos. O princípio que guia este raciocínio é de que a lógica pela qual a moeda instrumentaliza a apreensão do mundo econômico é similar à lógica pela qual as tecnologias digitais passam a operar nas interações sociais contemporâneas. Mas, afinal, o que de fato caracteriza esta lógica?

A moeda, em essência, traduz de forma simplificada e reduzida diversas decisões de oferta e demanda dos agentes econômicos em quantidades. Essa simplificação, mesmo com seus defeitos e limitações, habilita estudos de razoável precisão a respeito de como as interações baseadas na moeda refletem de maneira agregada a média das motivações nos mercados, permitindo que a teoria econômica construa modelos para descrever dinâmicas sociais.

Tais modelos só são possíveis, contudo, devido ao caráter distribuído e fiduciário da moeda. Por ser distribuída, a moeda opera como o tecido mediador da totalidade das transações comerciais. Consequentemente, acaba por registrar quantitativamente estas transações. Por seu caráter fiduciário, ela é capaz de captar motivações de maneira orgânica. Diferentemente de um questionário que induz uma resposta, a transação da moeda revela motivações dos agentes pelo simples fato de que eles acreditam no seu valor. Trata-se de um fator fundamental na capacidade dos modelos econômicos refletirem com precisão dinâmismos sociais que permaneceriam ocultos sem a moeda.

No ambiente contemporâneo, o smartphone representa cada vez mais o papel de um “companheiro” digital. Presente na maior parte das atividades cotidianas, ele funciona como

despertador, relógio, comunicador, GPS, câmara, rádio, TV, acesso às redes sociais e, naturalmente, telefone. Há cerca de 4 bilhões de usuários desses dispositivos no mundo⁸, cujo tráfego de dados cresceu 50% apenas no período entre 2015 e 2016. No Brasil o número de usuários chega a cerca de 170 milhões⁹, aproximadamente 85% da população.

O smartphone contemporâneo está mais para um computador portátil dotado de sensores do que para um telefone, e o número de sensores tende a crescer. Modelos bastante populares já possuem sensores de movimento, posição, velocidade, proximidade e geolocalização. Modelos mais modernos tem esses sensores e diversos outros.

Aplicativos para smartphones são muito populares no Brasil¹⁰: seus principais usos são comunicação (*WhatsApp*, 93% dos pesquisados, *Skype*, 11% e *Snapchat*, 6,5%), redes sociais (*Facebook*, 79%, *Instagram*, 37%, *Twitter*, 14% e *LinkedIn*, 5%), Vídeo e música (*YouTube*, 60%, *Netflix*, 9,5% e *Spotify*, 5,5%), Jogos (35%), Mapas (19%) e Bancos (10%).

A maior parte dos aplicativos, serviços de informação e mídias sociais é oferecida de forma gratuita para seus usuários. Com isso, tecnologias digitais como o smartphone passam a estar presentes na mediação das relações sociais, operando enquanto sensores de hábitos e comportamentos de usuários. Trata-se do novo tecido que permeia a instância comunicacional da vida social e, com isso, registra suas características.

Inicialmente, devemos ressaltar algumas distinções entre as peculiaridades de cada instrumento. Se a moeda é em si o meio e a medida, as tecnologias obedecem outra lógica. Diferentemente, celulares e navegadores são instâncias de mediação das interações, não constituindo em si instrumentos de medidas. A mensuração ocorre através dos rastros deixados pelas interações tecnologicamente mediadas, rastros estes chamados de *Metadados*. Os *Metadados*, ou "dados sobre dados", são a informação coletada pelos serviços e aplicativos digitais que pode ser armazenada e interpretada de forma nunca antes possível. Da mesma forma que a moeda simplifica e traduz uma série de intenções na interação econômica, os metadados agregados por serviços de interação social podem ser armazenados e gerenciados em grandes bases de dados.

⁸ Dados do relatório Ericsson Mobility Report, de nov/2016 - <http://bit.ly/2lgK8MJ>

⁹ Dados da 27ª Pesquisa Anual de Administração e Uso de TI nas Empresas realizada pela FGV-SP - <http://bit.ly/2ko73UD>

¹⁰ Dados da pesquisa IBOPE com 2000 entrevistados. 49% dos respondentes usam mais de 5 aplicativos diferentes por dia. - <http://bit.ly/2lgOcwn>

Uma consulta a um serviço via Internet nada mais é que uma operação de transferência de dados entre máquinas. Essa transação deixa vários tipos de rastros, que vão da análise da atividade do servidor à instalação de códigos no website que são disparados a cada interação do usuário. O registro desses rastros, normalmente invisível para o usuário, pode identificar a origem e localização geográfica do visitante, o tempo dispensado no serviço, o preenchimento de formulários, as interações de conversa, comentário ou apreciação de conteúdos disponibilizados por outro, visualizações de vídeos, download de documentos e compartilhamento de conteúdos, entre outros. Os dados pessoais de identificação costumam ficar em sigilo, interpretados apenas em grupo, como base estatística.

Através da coleta desses dados em grandes quantidades, as interações sociais, normalmente repletas de sutilezas e dependentes de contextos e interlocutores, podem ser compiladas e interpretadas. Um exemplo são os modelos capazes de inferir humores e opiniões, uma área de pesquisa estatística chamada de *Sentiment Analysis and Opinion Mining* a partir de redes sociais, como define Bing Liu, da Universidade de Illinois:

Sentiment analysis and opinion mining é o campo de estudo que analisa a opinião das pessoas, sentimentos, avaliações, atitudes e emoções a partir da linguagem escrita. É uma das áreas de pesquisa mais ativas no processamento de linguagem natural e também vastamente utilizada na pesquisa de dados. Essa pesquisa expandiu-se para além da Ciência da Computação para as Ciências de Administração e Ciências Sociais, devido à sua importância para os negócios e sociedade como um todo. A crescente importância da *Sentiment Analysis* coincide com o crescimento de mídias sociais, como avaliações, discussões em fóruns, blogs, micro-blogs, Twitter e redes sociais. Pela primeira vez na história humana há um grande volume de dados de opinião, registrado digitalmente para análise. (LIU, 2012, pg.12)¹¹

Os metadados gerados pelas tecnologias de comunicação digital despontam, portanto, como uma oportunidade de mapear e visualizar facetas de dinâmicas sociais. Assim como a

¹¹ No original: “*Sentiment analysis and opinion mining is the field of study that analyzes people's opinions, sentiments, evaluations, attitudes, and emotions from written language. It is one of the most active research areas in natural language processing and is also widely studied in data mining, Web mining, and text mining. In fact, this research has spread outside of computer science to the management sciences and social sciences due to its importance to business and society as a whole. The growing importance of sentiment analysis coincides with the growth of social media such as reviews, forum discussions, blogs, micro-blogs, Twitter, and social networks. For the first time in human history, we now have a huge volume of opinionated data recorded in digital form for analysis.*” (<http://bit.ly/2ktuihQ>) - Tradução livre dos autores.

moeda, smartphones estão distribuídos pela a sociedade, mediando a relação dos usuários com empresas, governos, produtos, veículos de informação e, naturalmente, outros usuários.

Os comportamentos sociais mediados por tecnologias digitais costumam ser bastante espontâneos, o que garante a qualidade dos dados coletados por seus provedores de serviços. É assim que o ambiente digital é transformado em um tecido que permeia a esfera pública, viabilizando o mapeamento de opiniões, relações entre usuários, preferências, engajamento com conteúdos, localização geográfica, deslocamentos e toda uma infindável lista que só tem limites na criatividade de se pensar vias para mensuração e cruzamento dos dados. Uma pesquisa feita pela Comissão Federal de Comércio dos EUA, analisando nove das maiores empresas de coleta e comercialização de dados (*Data Brokers*, ou corretores de dados) nos Estados Unidos¹² chegou às seguintes conclusões:

O segmento de corretores de dados é complexo, com múltiplas camadas intercambiando informações. Os corretores de dados fornecem dados não só para os usuários finais, mas também a outros corretores de dados;

Corretores coletam dados de várias fontes, em grande parte sem conhecimento de quem os fornece. Os dados recolhidos podem incluir informações de falência, registro de votação, dados de compra do consumidor, atividades de navegação na internet, dados de garantia e outros detalhes de interações cotidianas, de fontes comerciais e públicas. Enquanto cada fonte pode fornecer alguns dados sobre as atividades do consumidor, corretores podem aglutiná-los para formar um composto detalhado de cada pessoa; e

Corretores combinam e analisam dados para fazer inferências, incluindo relações potencialmente sensíveis. Categorias potencialmente sensíveis incluem etnias, níveis de renda, idade, histórico de moradias, nível de escolaridade e suas combinações.

Em outras palavras, os dados já são coletados. Eles são, em sua maioria, restritos a serviços comerciais que os comercializam para potenciais interessados no envio de mensagens direcionadas de marketing.

¹² As empresas pesquisadas foram: Acxiom, Corelogic, Datalogix, eBureau, ID Analytics, Intelius, PeekYou, Rapleaf e Recorded Future. A pesquisa está disponível em (<https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf>)

5. Desafios epistemológicos e práticos

Apesar promissora, a simples existência de um instrumento capaz de mensurar comportamentos midiáticos é insuficiente para um deslocamento quantitativo dos modelos em ciências da comunicação. É importante ressaltar o surgimento de problemas de caráter epistemológico devido a natureza do objeto comunicacional. Eles podem ser classificados em três categorias:

- a) **Objeto:** características do consumo de mídia e interação social mediada por redes sociais;
- b) **Finalidades:** intenções da produção no campo; e
- c) **Instrumento:** Métricas múltiplas, de validade questionável e frequentemente de fontes privadas

Começamos pela natureza do objeto, talvez o principal desafio epistemológico. Mais uma vez retomando o paralelo com as ciências econômicas, a capacidade de tradução matemática dos comportamentos só é possível dadas as características do homem nos mercados (*Homo economicus*): trata-se de um indivíduo, em linhas gerais, racional e maximizador de utilidade, portanto razoavelmente previsível e determinável em seus padrões sociais.

A lógica econômica não atribui essas características ao comportamento humano genérico, mas estritamente às motivações dos agentes de mercado em sua oferta e/ou demanda de bens. Tendo como pressuposto um indivíduo racional e maximizador de utilidade, a previsibilidade dos comportamentos permite a enunciação de leis de tendência facilmente traduzidas em modelos matemáticos.

No mundo da comunicação a natureza das interações sociais é bastante diferente, o que dificulta muito a descrição matemática dos modelos teóricos. Diferentemente dos mercados, uma exposição pessoal ou interação social não se trata de um comportamento com caráter estritamente racional. Em verdade, os hábitos de mídia são muito mais marcados por fatores ambientais, cognitivos e subjetivos do que propriamente racionais. O caráter indeterminado da interação social online dificulta a determinação de padrões ou tendências de comportamento. Diferente da economia, a natureza do comportamento com a mídia não é marcada por previsibilidade e nem tampouco por interações facilmente previsíveis.

Não obstante, é preciso ressaltar que as proposições teóricas através de um determinismo matemático constitui apenas uma forma do pensar quantitativo. O uso de “leis” é algo que as ciências econômicas importaram da física, e sua aplicação ortodoxa frente à complexidade da esfera social sempre foi questionada. Por outro lado, ferramentas estatísticas e computacionais são cada vez mais utilizadas na procura de relações repletas de indeterminação. Por mais que um modelo teórico de comunicação possa ser enunciado através de uma fórmula, as possibilidades de análise estatística são cada vez maiores. Em vez de buscar uma lei que totalize a compreensão de um dado fenômeno, a estatística deve permitir a avaliação de padrões e tendências em circunstâncias pontuais, mensurando até mesmo o nível de incerteza da situação analisada.

Esse tipo de matriz interpretativa está bastante alinhada com as proposições teóricas de Bruno Latour (2007), que vislumbram um modelo de análise sociológica empírico e, ao mesmo tempo, repleto de indeterminação. Deixa-se de buscar uma teoria totalizante em detrimento do entendimento contextual, singular e único dos fenômenos.

Passemos agora para o ponto dois, acerca das pretensões. A partir da consolidação marginalista na economia, o núcleo do campo se restringiu a buscar alocações eficientes de meios para dados fins. Com isso, ficam excluídos da ciência econômica ortodoxa debates de caráter moral sobre os fins, que passam principalmente para outros ramos das ciências sociais. Debater se a busca por crescimento do PIB é a melhor via para o desenvolvimento seria, nesse sentido, um debate alheio ao escopo da economia, já que esta versaria estritamente sobre a melhor alocação de recursos possível para um dado crescimento almejado.

A pesquisa no campo da comunicação, no entanto, conta com uma rica multiplicidade de pretensões. Há linhas em que predomina o debate crítico sobre a cultura e a mídia, reflexões de caráter puramente teórico sobre as características do processo comunicacional, e até mesmo linhas de caráter experimental que utilizam neurociência e psicologia cognitiva no entendimento dos aspectos subjetivos do comportamento.

Se nas ciências econômicas o consenso ao redor das pretensões facilita a matematização do campo, em comunicação a questão não é tão simples. Quando a ciência se porta enquanto ferramenta na busca da melhoria de um processo, seja ele a alocação de recursos ou a busca por um modelo de comunicação mais eficiente, não há dúvidas de que a estatística e o cálculo são extremamente valiosos. Por outro lado, quando o fim com que se

pesquisa é propor um debate de caráter moral e crítico, há uma grande propensão em restringir-se ao processo argumentativo e retórico tradicional das ciências humanas. Ainda que a matemática e a busca por métodos mais objetivos de construção do conhecimento de fato encontrem seu lugar de excelência na busca por melhoria de processos, não devemos descartá-las de imediato na produção reflexiva. Com possibilidades cada vez maiores de se apreender aspectos da esfera social, a teoria crítica passa a contar não só com o arsenal retórico, mas também com outro de caráter empírico.

Esse vislumbre de uma teoria crítica mais próxima ao modelo de *Hard Science* não é novo, e carrega muita controvérsia. A grande dificuldade constitui a procura por um modelo mais objetivo, com procedimentos replicáveis e previsibilidade de resultados. Hoje, na produção científica em ciências da comunicação predominam esforços hermenêuticos de diversas ordens, que mesmo quando carregados de brilhantismo permanecem ligados ao raciocínio interpretativo pessoal de cada autor.

A busca por uma construção reflexiva de base empírica, por outro lado, implica um esforço maior de afastamento na busca por análises que, se jamais serão plenamente objetivas, poderão ser facilmente checadas pelo mais questionador dos leitores, que Bruno Latour chamará de “Opositor Inflexível” (LATOUR, 1998, pg.91). Em última instância, conciliar as construções hermenêuticas da *Soft Science* com recursos quantitativos e computacionais que fortaleçam a produção científica no campo da comunicação é objetivo final deste artigo.

Explicitando, enfim, o último dos problemas de caráter epistemológico, temos a questão das métricas. Trata-se de um ponto que afasta definitivamente o processo de matematização das ciências econômicas das possibilidades que se abrem para o campo da comunicação. Enquanto que a moeda foi um medidor das motivações utilitárias dos indivíduos, é impossível pensarmos uma única medida oriunda dos metadados provenientes de interações sociais mediadas por tecnologias digitais. Por outro lado, as associações entre diferentes perfis em redes sociais permitem mapear os diferentes núcleos de aglutinação.

A grande possibilidade de registros que podem ser extraídos das interações sociais mediadas por tecnologias digitais permite um potencial maior no uso de dados digitais para elaboração de análises estatísticas do que para a proposição de leis de tendência dos comportamento. Ficaria a cargo da inventividade de cada pesquisador pensar formas mais sofisticadas de usos para estes dados na pesquisa científica.

6. A inevitável matematização das ciências sociais aplicadas

A intenção deste ensaio é chamar a atenção dos pesquisadores de humanidades para as novas ferramentas e possibilidades oriundas das interações sociais mediadas por redes sociais digitais e as possibilidades que se abrem para o campo das ciências da comunicação a partir do uso de ferramentas computacionais e quantitativas na análise desses dados.

Partimos do entendimento que os metadados provenientes das interações mediadas pelo ambiente digital constituem recursos capazes de apreender dimensões até então ocultas do processo comunicacional. Por mediar a relação entre os usuários e o ambiente digital, os serviços de mídias sociais passam a operar como sensores de comportamentos, sendo estes, por sua vez, produtores de rastros passíveis de mensuração, armazenamento e análise.

Esta ambição não deve ser confundida com um pretense deslocamento do campo em direção ao modelo de produção científica positivista, estritamente ligado a estrutura de investigação das *Hard Sciences*. De fato, com a possibilidade de mensuração podemos adotar procedimentos próximos ao modelo de pesquisa determinista nos moldes positivistas. Contudo, mostramos também que a natureza do objeto comunicacional possui marcas distintas que, por mais que agora possam ser tratados com mais objetividade, através de metodologias rigorosas e replicáveis de análise dos dados, isso não implica de maneira alguma no abandono o caráter hermenêutico da produção no campo. Em verdade, o caráter interdisciplinar é justamente a marca distintiva do campo, já que "o campo da comunicação se constitui a partir da multiplicidade de discursos que originam e configuram a unidade do discurso da comunicação" (BACCEGA, 1998, pg.104)

É importante, acima de tudo, levar em conta a existência de uma dimensão política nas transformações científicas. Se a matematização das ciências econômicas só foi possível por que um conjunto de pressupostos tornou-se hegemônico, hegemonia esta que fez com que debates sobre a natureza da ação humana se tornassem "superados" e exteriores ao escopo da ciência, o mesmo vale para a comunicação. Foi só a partir da consolidação de um núcleo duro, citando Lakatos (1989), que a economia passou a avançar a partir de testes econométricos, criando o que convencionou-se chamar hoje de Ortodoxia. Da mesma forma, a adoção ou não dos procedimentos aqui mencionados depende do nível de consagração que seus pressupostos analíticos receberão, bem como da inclinação dos membros do campo a

adotarem instrumentos computacionais e quantitativos como vias para a reflexão sobre fenômenos comunicacionais.

Este ensaio busca contribuir para um debate que se inicia sobre o uso de metadados na pesquisa em Ciências Humanas. O dimensionamento epistemológico é fundamental para se compreender que, assim como qualquer instrumento, as tecnologias digitais não apreendem a realidade em si, mas apenas uma versão dela. O comportamento do internauta não é um análogo perfeito do comportamento exterior a esfera digital. As questões são muitas e abrem espaço para um necessário debate sobre o status das ciências sociais aplicadas diante da revolução digital.

Referências

- ADORNO, Theodor W. ; HORKHEIMER, Max. **Dialética do Esclarecimento**. Rio de Janeiro: Jorge Zahar Editor, 1985.
- ARIDA, P. A **História do pensamento econômico como teoria e retórica**. In Rego, José Marcio (org.) Retórica na Economia. São Paulo: Editora 34, 1996.
- BACCEGA, Maria Aparecida. **Comunicação e Linguagem: discursos e ciência**. São Paulo: Moderna, 1998.
- BENTHAM, Jeremy (1780). **An introduction to the principles of morals and legislation**. Nova York: Pegasus, 1989,
- LIU, Bing. **Sentiment analysis and opinion mining**. Synthesis Lectures on Human Language Technologies, Vol.5, No. 1, pgs.1-164: Graeme Hirst, 2012.
- HUME, David. **An Enquiry concerning Human Understanding**. Oxford University Press, Nova York, Estados Unidos, 2007.
- HUNT, Emery Kay. **História do Pensamento Econômico: Uma perspectiva Crítica; 2a edição**. Rio de Janeiro: Elsevier, 2005.
- JEVONS, S. **The Theory of political economy**. Nova York: Reprints of Economic Classics, 1965.
- KUHN, T. S. **A estrutura das revoluções científicas**. São Paulo: Editora Perspectiva, 1994.
- LAKATOS, I. **Falsificação e metodologia dos programas de investigação científica**. Trad. E.P.T.M. Mendes. Lisboa: Edições 70, 1999.
- _____. **La metodologia de los programas de investigación científica**. Madrid: Alianza, 1989.
- LATOUR, B. **Ciência em Ação**. São Paulo: Editora Unesp, 1998.
- _____. **Resembling the social: an introduction to actor-network theory**. Oxford: Oxford University Press, 2007.
- LÉVY, Pierre. **As tecnologias da inteligência: o futuro do pensamento na era da informática**. São Paulo: Editora 34, 1993.
- MARTINS, Gilberto de Andrade; THEÓPHILO, Carlos Renato. **Metodologia da investigação científica para Ciências Sociais Aplicadas**. 2a edição. São Paulo: Editora Atlas, 2009.
- MARSHALL, Alfred (1920). **Princípios de economia**. Vol I, livro 4. São Paulo: Nova Cultural, 1982
- McLUHAN, Marshall. **Galáxia de Gutemberg**. São Paulo: Editora Nacional, 1977
- POPPER, K. R. **A Lógica da pesquisa científica**. Tradução de Leonidas Hegenberg e Ocatny Silveira da Mota. São Paulo: Editora Cultrix, 1975.
- RILEY, Jenn. **Understanding Metadata: what is metadata and what is it for?** Baltimore: Niso Primer, 2017.
- ROBBINS, Lionel. **An essay on the nature and significance of the economic science**. Londres: Macmilan and co, 1945.
- SMITH, Adam (1776). **A riqueza das nações**. São Paulo: Abril Cultural, 1983