

## DADIFICAÇÃO E TRANSDISCIPLINARIDADE NOS ESTUDOS DO LETRAMENTO: O JORNALISMO DE DADOS COMO CONTEXTO DE PESQUISA

## DATAFICATION AND TRANSDISCIPLINARITY IN LITERACY STUDIES: DATA JOURNALISM AS A RESEARCH CONTEXT

Marcelo El Khouri Buzato  
Universidade Estadual de Campinas, Campinas, São Paulo, Brasil

Cleyton Carlos Torres  
Universidade Estadual de Campinas, Campinas, São Paulo, Brasil

**Resumo:** Aborda-se o conceito e a prática do letramento de dados, visando caracterizá-lo como instanciação da transdisciplinaridade do(s) letramento(s) em sentido mais geral, apresentando-se, a partir dessa caracterização, perguntas e hipóteses de uma proposta de pesquisa focada no jornalismo de dados. Situamos a discussão no contexto da dadificação digitalizada da atividade social, argumentando que ela afeta a formulação de opiniões e argumentos baseados em textos jornalísticos que caracteriza o exercício da cidadania na esfera pública. Nossa pesquisa bibliográfica evidenciou que, embora predominem concepções de letramento de dados de natureza funcional e/ou baseada em habilidades (disciplinares), há também propostas baseadas no modelo ideológico/sociocultural do letramento. Concluímos apresentando a hipótese de que, nas práticas de jornalismo de dados, haveria uma tendência à subordinação das interpretações qualitativas aos saberes quantitativos, mas que se podem criar situações em que as lacunas disciplinares evidenciam o caráter transdisciplinar dos textos e dos eventos para os próprios produtores e leitores.

**Palavras-chave:** Letramento; Transdisciplinaridade; Dadificação; Letramento de Dados; Jornalismo

**Abstract:** The concept and practice of data literacy is approached in order to characterize it as an instantiation of the transdisciplinary of literacy (s) in a more general sense, and, based on this characterization, questions and hypotheses are presented of a research project focused on data journalism. We place the discussion in the context of the digitized datafication of social activity, arguing that it affects the formulation of opinions and arguments based on journalistic texts that characterize the exercise of citizenship in the public sphere. Our bibliographical research has shown that, although the conceptualizations of data literacy are predominantly functional and/or skills-based, there are also proposals based on the ideological/sociocultural model of literacy. We conclude by hypothesizing that, in data journalism practices, there would be a tendency for qualitative interpretations to be subordinate to quantitative knowledge, but it may be possible to create situations in which disciplinary shortcomings show the transdisciplinary nature of the texts and the events for the producers and readers themselves.

**Keywords:** Literacy; Transdisciplinary; Datafication; Data Literacy; Journalism

## 1. INTRODUÇÃO

Neste trabalho, discutiremos o conceito e a prática do letramento de dados, caracterizando-os como transdisciplinares. Para tanto, é importante estabelecermos, inicialmente, o que entendemos por transdisciplinaridade, para, em seguida, qualificar o(s) letramento(s) como fenômeno(s) transdisciplinar(es). Estabelecidas essas premissas, trataremos o problema da dadificação das práticas socioculturais contemporâneas e os conceitos associados de *big data* e ciência de dados, caracterizando o(s) letramento(s) de dados como um problema perverso (*wicked problem*), “que desafia uma definição completa e não pode ser resolvido usando-se os modos existentes de investigação e tomada de decisão” (BERNSTEIN, 2015, s/p., tradução nossa)<sup>1</sup>; logo, torna-se um objeto ainda mais evidentemente transdisciplinar. Ao revisar algumas concepções de letramento de dados vigentes na literatura pertinente, argumentamos que seu caráter transdisciplinar é tão ou mais visível do que no caso de outros letramentos, porque assenta-se justamente sobre a cisão entre quantitativismo e qualitativismo, duas importantes estratégias de fronteirização entre disciplinas.

Levamos o problema do(s) letramento(s) de dados, em seguida, para a esfera do jornalismo de dados, uma prática discursiva que consideramos ter o duplo potencial de permitir a apropriação dos saberes e benefícios da dadificação pelo cidadão letrado e/ou de deslocar o cidadão iletrado em quantitativismo à posição de um leigo crédulo e incapaz de estabelecer um diálogo crítico com “os fatos” que lhe são “mostrados” em lugar de “as histórias” que lhe eram contadas no jornal.

Concluimos o trabalho com algumas perguntas de pesquisa voltadas para o jornalismo de dados enquanto nova esfera de letramentos sociais, discutindo, brevemente, as hipóteses levantadas para a pesquisa de Torres (2019), em andamento, e suas implicações.

## 2. LETRAMENTO E TRANSDISCIPLINARIDADE

Aparentemente, foi Jean Piaget (1972, p. 138) quem primeiro utilizou o termo “transdisciplinaridade”, referindo-se a um estágio superior de produção do conhecimento em que as interações e reciprocidades entre disciplinas especializadas seriam sobrepujadas por um sistema mais geral, no qual não haveria fronteiras disciplinares. Nicolescu (2002) dá fim à ilusão piagetiana, mostrando que o transdisciplinar não poderia, de fato, dispensar fronteiras internas senão reconhecendo que a transdisciplinaridade não almeja a totalidade, mas aceita o exercício permanente da abertura. Kochelmans (1979), por sua vez, já conectava transdisciplinaridade e democracia ao postular que a transdisciplinaridade, como prática e como ideal, implica pesquisa e educação socialmente relevantes.

Desde meados do século XX, os estudos do letramento demonstram a transdisciplinaridade do seu objeto complexo e mutante. Por ser complexo, atrai diferentes perspectivas disciplinares, tais como a dos estudos clássicos e literários (PARRY, 1971); a da comunicação e midialogia (McLUHAN, 1972); a da antropologia (HEATH, 1983; STREET, 1984); a da pedagogia (FREIRE, 1970; SOARES, 1995); a da linguística (KATO, 1986) e a da psicologia (SCRIBNER; COLE, 1981) – além das abordagens claramente transdisciplinares como as da teoria dos sistemas complexos (LEMKE, 1995) e dos estudos sociais da ciência e tecnologia (LATOURETTE, 1986), entre outras.

O(s) letramento(s) é(são) mutante(s) porque entrelaça(m) variações de tecnologias, competências, valores, significados, esquemas de ação, propósitos e tipos de textos (LEMKE, 2010; BUZATO, 2009) em diferentes épocas e lugares, sendo impossível torná-lo fixo a partir

---

<sup>1</sup> “(...) defy complete definition and cannot be solved using existing modes of inquiry and decision making (...)”. (BERNSTEIN, 2015, s/p.).

de uma somatória de olhares especializados. Acrescente-se a esses entrelaçamentos, hoje, o das agências e inteligências humanas e não-humanas mutuamente agenciadas (BUZATO, 2013).

Assim, o(s) letramento(s) ilustra(m) a ideia de Morin (2004, p. 2) de que “a transdisciplinaridade não procura o domínio sobre as várias outras disciplinas, mas a abertura de todas elas àquilo que as atravessa e as ultrapassa”. Isso vale mesmo que sua gênese tenha sido grafocêntrica (GOODY; WATT, 1963) e apesar da necessidade eventual de reduzir o “trans” a “multi” para fins políticos (THE NEW LONDON GROUP, 1996) e pedagógicos (COPE; KALANTZIS, 2000) ou, ainda, compartimentalizar variedades do objeto (“letramento matemático”, “letramento de informação”, “letramento escolar”, “letramento de dados” etc.) para tentar geri-lo por facetas.

Defendemos que, mais do que outros letramentos específicos, o letramento de dados convoca urgentemente uma discussão sobre transdisciplinaridade. Primeiramente, porque, pensado como prática sociocultural, ele atinge, em cheio, uma delimitação central na “geopolítica” das disciplinas ditas exatas e humanas, colocando em questão o direito de cada lado da divisa de definir quais letramentos de dados contam socialmente e como esse processo se dá. Isso é especialmente urgente, porque a cultura da dadificação está em franca difusão em domínios e comunidades discursivas anteriormente refratárias ao quantitativismo, alheias às suas formas de representação comportando-se de forma pouco crítica em relação à pragmática dos números e estatísticas.

### 3. DADIFICAÇÃO

Dadificar (*to datafy*, SCHÄFER; Van ES, 2017) um fenômeno é transformá-lo em quantidades que podem ser tabuladas e analisadas estatisticamente. Digitalizar (*to digitize*), é converter informação analógica em dígitos binários legíveis por computadores. A dadificação é muito anterior à digitalização, porém o encontro dessas duas operações epistêmicas disparou uma profusão sem precedentes de dados e metadados gerados por sensores ambientais, sistemas de gestão informatizados e, especialmente, pela atividade sociodiscursiva registrada nos circuitos das mídias sociais digitais, fazendo surgir um novo campo de conhecimento chamado Ciência de Dados (*Data Science*) (EMC EDUCATION SERVICES, 2015).

O objeto central da Ciência de Dados é o *big data*, isto é, a operação sobre grandes volumes de dados, gerados e transmitidos instantaneamente desde uma grande variedade de fontes, que são processados em laboratórios equipados com mecanismos de aprendizagem de máquina e análise estatística avançada (KITCHIN, 2014), por pessoas com formação em estatística, informática e analítica de dados. Os dados e metadados utilizados são obtidos, em boa parte das vezes, no segundo plano, invisível ao olhar leigo, das atividades cotidianas dos cidadãos no ciberespaço, mas também por sensores diversos instalados nos espaços públicos e privados de convívio (MAYER-SCHÖNBERGER; CUKIER, 2013).

Assim como a digitalização, a dadificação é consequência da implantação do quantitativismo, já nos primórdios da própria ciência moderna, por volta do século XIII, como via preferencial de acesso aos fatos (científicos) invisíveis a olho nu. Ambas não são mais do que formas de exponencializar a eficiência do quantitativismo e, por tabela, os discursos ortodoxos da – e sobre a – ciência projetados sobre o cotidiano. A despeito de seu cientificismo, a quantificação digitalizada do mundo não é um empreendimento neutro em seus efeitos sociais, tampouco isento de mediações semióticas complexas (LATOURET, 1999, 2000). Ao contrário: a quantificação é “ação social que, como o discurso, pode ter múltiplos propósitos e sentidos (que) como para a língua/linguagem (...) são estabelecidos pelo uso”

(ESPELAND E STEVENS, 2008, p. 407, tradução nossa)<sup>2</sup>.

A quantificação “torna visíveis pessoas, objetos ou características que antes eram invisíveis”, possibilitando “o escrutínio de fenômenos complexos ou díspares de maneira que permitam o julgamento” (ESPELAND; STEVENS, 2008, p. 415, tradução nossa)<sup>3</sup>. Na medida em que “ver algo é o primeiro passo para controlá-lo”<sup>4</sup> (*ibidem*)<sup>5</sup>, a quantificação estabelece formas de autoridade sobre o que é real, e, portanto, sobre o que pode ou não ser mudado/questionado.

Assim como palavras e imagens, quantificações são também performativas, no sentido da teoria dos atos de fala (AUSTIN, 1962), segundo Espeland e Stevens (2008). O seu componente locucionário envolve convenções matemáticas, padrões metodológicos e formas de expressão próprias. O seu componente ilocucionário aparece como capacidade de “direcionar atenção, persuadir e, especialmente, criar novas categorias de apreensão do mundo” (ESPELAND; STEVENS, 2008, p. 404, tradução nossa)<sup>6</sup>. Já entre seus efeitos perlocucionários, estão, por exemplo, os seguintes: países entram em recessões (ou saem delas), dependendo do valor do seu PIB<sup>7</sup>; traços censitários das populações produzem “comunidades estatísticas” (IGO, 2007) que adquirem força política; doenças são diagnosticadas e tratadas, ou rebaixadas a psicossomatismos, a partir de métricas médicas; e, é claro, quantidades de *likes*, *views* e seguidores criam micro celebridades e influenciadores (KHAMIS *et al.*, 2017) que ajudam a eleger, ou derrubar, governantes.

O enorme avanço do quantitativismo e do estatisticismo<sup>8</sup> representado pelo *big data* traz não apenas preocupações para com a privacidade dos sujeitos e o agravamento de estigmas e preconceitos sociais embutidos em algoritmos (O’NEIL, 2017), mas também certa forma de ver/interpretar o mundo – até então restrita a laboratórios, empresas de análise de riscos e *bureaus* de estatística governamentais – , que se pretende apta a explicar e melhor gerir os mundos subjetivo e intersubjetivo dos cidadãos, suplantando, por conseguinte, formas discursivas verbais e não-verbais tidas, tradicionalmente, como suficientes para o exercício da cidadania.

Não haveria motivo para alarde não fossem afirmações como a de Anderson (2008, parágrafo 7, tradução nossa), editor da badalada revista *Wired*, acerca do *big data*:

Fora com todas as teorias de comportamento humano, da linguística à sociologia. Esqueçam taxonomia, ontologia e psicologia. Quem sabe por que as pessoas fazem o que elas fazem? O ponto é que elas o fazem, e nós podemos rastrear e medir isso com uma fidelidade sem precedentes. Com dados suficientes, os números falam por si mesmos.<sup>9</sup>

<sup>2</sup> “(...) Social action that, akin to speech, can have multiple purposes and meanings (that) as with language (...) are established through use”. (ESPELAND E STEVENS, 2008, p. 407).

<sup>3</sup> “Quantification makes visible people, objects or characteristics that may formerly have been invisible. It permits scrutiny of complex or disparate phenomena in ways that enable judgment”. (ESPELAND E STEVENS, 2008, p. 415).

<sup>4</sup> Ver também Latour (1986).

<sup>5</sup> “Seeing something is the first step to controlling it”. (*ibidem*).

<sup>6</sup> “(...) directing attention, persuading, and creating new categories for apprehending the world”. (ESPELAND E STEVENS, 2008, p. 404).

<sup>7</sup> O Produto Interno Bruto (PIB) é obtido pela soma dos valores monetários de todos os bens e serviços finais realizados em um país em um determinado período de tempo.

<sup>8</sup> Duncan (1984) define o estatisticismo como a noção de que quantificar é o mesmo que fazer pesquisa, baseada, por sua vez, na fé ingênua na estatística em si mesma como metodologia científica que sustentaria modelos causais da realidade. Essa formulação é importante para contradizer alguns dos discursos recentes sobre o poder científico do *big data*.

<sup>9</sup> “Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it

Ainda que fortemente rechaçada pelos estudos críticos de dados (BOYD; CRAWFORD, 2012), essa atitude resume a ambição latente do cientista de dados: estabelecer a supremacia da correlação sobre a causalidade. Como amostras muito maiores melhoram muito os intervalos de confiança estatística, é possível, para esses analistas, tratar padrões e tendências numéricas como fatos científicos e sociais, mesmo sem identificarem uma lógica subjacente de causa e consequência. Não há nada de anticientífico nisso, exceto a crença, ingênua, de que os dados são de fato dados, e não construídos pela seleção do que, como e por que se mede e se considera relevante. Essa ideologia quantitativista que consiste em acreditar que a quantificação é imune à ideologia interessa a uma discussão sobre letramento(s) de dados enquanto objeto(s) transdisciplinar(es) por diversos motivos.

Ademais, aquilo que foi construído à base de correlações numéricas por uma inteligência artificial, em um *corpus* de bilhões de tuítes, certamente tem sentidos, nessa escala de significação (LEMKE, 2000), que não podem ser apreendidos por leitores humanos a partir de sua experiência corporificada. Dessa forma, a propagação desses sentidos maquinais no ecossistema cognitivo das comunidades humanas, que inclui computadores, gerará, sempre, interpretações muito parciais que, ainda assim, podem ter repercussões políticas, estéticas e éticas imprevisíveis e danosas (O'NEIL, 2017).

Isso quer dizer que precisamos aprender e ensinar pessoas a construir significados via dados, atravessando não apenas sistemas semióticos, gêneros e instituições, como de antes, mas também cruzando escalas de produção de sentido (LEMKE, 2000) e sistemas de interpretação constituídos por agentes híbridos, humanos e não-humanos (BUZATO, 2017), de modo a não confundir causalidade com correlação – por exemplo, quando se publica ou interpreta um texto escrito de cem palavras, uma equação de reta a partir de cem pontos de dados ou um grafo das cem palavras mais frequentes em um milhão de tuítes. É nisso que reside, em parte, a real possibilidade de abertura transdisciplinar do(s) letramento(s) de dados.

Essa abertura não é simples, porque o conhecimento extraído dos dados por mecanismos de aprendizagem de máquina é opaco. A máquina “aprende” a fazer coisas que o cientista de dados não pode descrever em palavras, e, logo, não pode ensinar a humanos. Assim, se aplicarmos tecnologia de *big data* a, por exemplo, todas as redações já escritas para o ENEM, poderemos ensinar a máquina a tirar a nota máxima, mas não um aluno a escrever. Por outro lado, um aluno poderá, facilmente, convencer a máquina de que fez uma ótima redação se entregar-lhe uma sequência de palavras encaixadas nos padrões desejáveis de coerência, ainda que essas palavras não tenham qualquer relação causal estabelecida entre si que torne o texto plausível (COPE; KALANTZIZ, 2016).

Assim como ambicionamos que um cidadão saiba ler coisas nas entrelinhas de uma notícia de jornal, ou a leia desconfiando, no bom sentido, de cada escolha feita pelo redator, devemos achar maneiras de tornar o letrado em dados disposto e apto a interpelar os dados e as visualizações e narrativas de dados como produções discursivas, não signos transparentes (BUZATO, 2017). Ainda que não se desmintam, por definição, axiomas matemáticos/estatísticos, é preciso desconfiar que “dados crus” (*raw data*) é um oxímoro (BOYD; CRAWFORD, 2012).

É dessas e outras dimensões da dadificação que, como propomos, deve tratar uma concepção transdisciplinar de letramento de dados, assim como as agendas de pesquisa e educação aí implicadas.

---

with unprecedented fidelity. With enough data, the numbers speak for themselves”. (ANDERSON, 2008, paragraph 7).

#### 4. DEFININDO LETRAMENTO DE DADOS

As definições de letramento de dados disponíveis na literatura atual são, em geral, fruto de pesquisa ou prática pedagógica nos campos da biblioteconomia/ciência da informação e estatística (SCHIELD, 2004), mas também começam a aparecer nos campos da comunicação (GRAY *et al.*, 2016; D'IGNAZIO; BHARGAVA, 2015) e da educação (VAHEY *et al.*, 2010; WOLFF *et al.*, 2015; ATHANASES *et al.*, 2013). Em grande parte dos casos, porém, as definições tendem a assentar-se sobre uma ótica funcional em que o fenômeno é visto como um conjunto de habilidades discretas, por vezes supostamente “trazidas” de outros letramentos quantitativos e técnicos.

Schild (2004, p. 8, tradução nossa)<sup>10</sup>, por exemplo, define letramento de dados como a habilidade de “acessar, avaliar, manipular, sumarizar e apresentar dados” que congrega os letramentos de informação (capacidade de reconhecer a informação necessária e ter a habilidade de localizá-la, avaliá-la e usá-la eficazmente) e estatístico (utilizar análises estatísticas como evidência argumentativa e saber como comunicá-las, considerando o contexto da sua interpretação).

Nessa definição, a série de infinitivos “acessar, avaliar, manipular, sumarizar e apresentar dados” claramente remete a habilidades genéricas mas, ainda assim, Schild (2004) as encaixa em uma prática discursiva em que o letrado deve buscar dados numa fonte de registro técnico e entregar conhecimento de forma culturalmente aceitável a alguém. Schild (2004) é, contudo, um bibliotecário, preocupado em definir letramento de dados no contexto do ensino de graduação, para futuros profissionais de análise simbólica, sobretudo nos campos das ciências exatas e biológicas, e não para o cidadão comum, que, cada vez mais, é chamado a posicionar-se sobre sua realidade social a partir de representações quantitativas produzidas por outrem.

Barghava e D'Ignazio (2014, p. 2), especialistas em “computação cívica” e professores da disciplina de *data storytelling*, dão um passo adiante, rumo a uma concepção retórico-discursiva de letramento de dados, *i.e.* ler, manipular, analisar e “argumentar com dados” (*argue with data*), o que envolveria usar dados e análises para dar sustentação a um argumento (ou uma discussão) mais amplo(a), isto é, criar uma narrativa de dados (ver também BUZATO, 2018) que seja persuasiva, e, ao mesmo tempo, “discutir com os dados”, questionando escolhas tendenciosas na sua geração e manipulação.

Avançando em direção a uma concepção afeita ao modelo sociocultural/ideológico, e agregando a ele uma dimensão crítica, inspirada explicitamente na visão freireana de letramento, D'Ignazio e Barghava (2015) advogam o ensino de letramento de dados como a instrução sobre habilidades técnicas, como as elencadas por Schild (2004), mas voltadas para a emancipação sociopolítica dos letrados. Nesse percurso, terminam por descrever uma habilidade específica no letramento de dados como prática social, que denominam “letramento de algoritmos”, isto é, a capacidade de compreender o funcionamento de algoritmos executando experimentos de engenharia reversa que permitam desvendar as decisões subjetivas de seus autores. Essa concepção de letramento de dados vai ao encontro do que, em outro lugar, Buzato (2017) se refere como sendo letramentos de dados críticos.

Todas essas concepções aparecem somadas, embora não exatamente integradas, num importante relatório produzido pela *Data Pop Alliance*, consórcio liderado pela *Rockefeller Foundation*, *Harvard Humanitarian Initiative* e *MIT Media Lab*, que tem por missão “juntar pesquisadores, especialistas, praticantes e ativistas para promover uma revolução *Big Data* centrada em pessoas”<sup>11</sup>. Aparenta, assim, buscar um trabalho efetivamente transdisciplinar e democrático em torno do letramento de dados, que, no relatório, define-se como “o desejo e a

<sup>10</sup> “(...) access, assess, manipulate, summarize, and present data (...)”. (SCHIELD, 2004, p. 8).

<sup>11</sup> Disponível em: <<http://datapopalliance.org/about/vision-and-members/>>. Acesso em: 20 Jul. 2016.

habilidade de engajar-se construtivamente na sociedade através de, ou a respeito de, dados”. (DATA POP ALLIANCE, 2015, p. iv, tradução nossa)<sup>12</sup>.

A inserção do “desejo” nessa definição visa, segundo os autores, destacar o papel da tecnologia como amplificadora da intenção humana, que se serve de “habilidades” adquiridas em diferentes níveis e relacionadas a diferentes posições e objetivos, as quais, no entanto, não estabeleceriam uma dicotomia entre letrados e não letrados em dados e, por extensão, entre os que vêm de disciplinas/práticas de base quantitativa ou qualitativa. “Engajar-se construtivamente na sociedade” pressupõe, segundo explicam os autores, um senso de propósito e agência não só no sentido de comunicar ou argumentar com, ou via, dados, mas de mudar algo no mundo. A expressão “ou a respeito de”, anteriormente destacada, ressalta, novamente, uma abertura do conceito às interpretações daqueles que pretendem simplesmente criticar ou apontar lacunas nos dados, na forma como são tratados, na sua capacidade de abarcar o fenômeno em tela etc., ainda que não disponham de conhecimentos disciplinares que permitam analisá-los ou representá-los efetivamente.

A relação entre dadificação enquanto processo social vinculado a interesses sociais e visões disciplinares específicos e o letramento de dados como prática de cidadania, do ponto de vista do consórcio, fica expressa no seguinte comentário: “se o letramento de dados fosse apenas uma questão de transformar dados em informação, uma sociedade de analistas juniores da NSA (e seus pares da Amazon ou da Google) seriam uma elite letrada”<sup>13</sup>, mas, “em uma era em que as preocupações sobre governos empoderados por analítica de dados e manipulações (como nos casos da experimentos sociais do Facebook [...]), estão aumentando, pode-se sentir intuitivamente que tal elite talvez não seja a mais progressiva e inclusiva” (DATA POP ALLIANCE, 2015, p. 6, tradução nossa)<sup>14</sup>.

Essa concepção não só abre o letramento de dados para uma abordagem transdisciplinar em que as facetas técnicas, psicológicas, sociais e discursivas não se recolhem a fronteiras rígidas, como implica que se convém mapear os espaços de apropriação da ciência de dados e das tecnologias do *big data* nas mais variadas esferas de atividade social. Em nosso caso, em particular, interessa pensar o letramento de dados no contexto do jornalismo, por sua ligação histórica e intrínseca com o exercício da cidadania e a construção do interesse público.

## 5. DO JORNALISMO DE DADOS

### 5.1. JORNALISMO E ESFERA PÚBLICA

O conceito de esfera pública como dimensão da prática discursiva que atua como mediadora entre o Estado e a sociedade é o lócus da instituição e gestão do que chamamos de opinião pública. Surgida no final do século XVII com a consolidação do ideal da cidadania baseado nas lutas das burguesias urbanas contra o absolutismo, a esfera pública é, na formulação de Habermas (1974, p. 29, tradução nossa), “um domínio de nossa vida social em que algo que se aproxima da condição de opinião pública pode ser formado”<sup>15</sup>. Materializada

<sup>12</sup> “(...) the desire and ability to constructively engage in society through and about data (...)”. (DATA POP ALLIANCE, 2015, p. iv)

<sup>13</sup> “(...) if ‘data literacy’ were just the ability to turn data into information, a society of junior NSA analysts (and their Amazon or Google counterparts) would be a highly data literate society (...)”. (DATA POP ALLIANCE, 2015, p. 6).

<sup>14</sup> “(...) in an era where concerns over data analytics-enabled government surveillance and corporate manipulations (as in the cases of the Facebook social experimentation ...) are rising; one can feel intuitively that such a society may not be the most progressive and inclusive (...)”. (ibidem).

<sup>15</sup> “(...) a realm of our social life in which something approaching public opinion can be formed”. (HABERMAS, 1974, p. 29).

na prática de cidadãos dotados de racionalidade e liberdade de expressão, a esfera pública é, muitas vezes, figurada como grupos reunidos a discutir notícias de jornal, TV e rádio, sopesando o valor do interesse público frente a interesses privados de indivíduos e corporações.

O jornalismo adquire, nessa prática, uma configuração com traços próprios de autoridade, pois caberia ao jornalismo oferecer ao leitor aquilo que ele, com critérios e valores próprios, julga se tratar de fatos relevantes em sintonia com os dizeres contemporâneas da sociedade (LIMA JUNIOR; ROCHA, 2010). A produção da esfera pública hoje envolve, contudo, produtiva ou perturbadoramente, discursos tirados de *blogs* e *vlogs* de “influenciadores” de vários tipos, cuja autoridade é, sobretudo, quantitativa, medida pelo número acessos, seguidores e reenvios obtidos.

Habermas (1974) já apontava que a esfera pública entrou em decadência nas sociedades contemporâneas à medida em que foi sendo “colonizada” por interesses privados, sobretudo comerciais, os quais influenciam os veículos jornalísticos, retirando do leitor, em última instância, a liberdade de expressar-se sobre o interesse comum. Hoje, podemos dizer que a esfera pública agrega um ecossistema em que a mídia tradicional e as mídias sociais se relacionam de maneira ambígua, ora criticando, ora apoiando, ora manipulando uma à outra (MARWICK; LEWIS, 2017), sendo essa, também, uma das facetas da relação entre quantitativismo e qualitativismo no jornalismo de dados.

Junto com as notícias, o jornalismo atualmente fomenta a participação dos sujeitos, fornecendo visualizações e narrativas de dados produzidas a partir de dados abertos, *i.e.*, de livre acesso, estocados em repositórios do governo, e de dados e metadados “raspados” de mídias sociais (ANDRIOLO, 2012). Esses dados, tanto estruturados (tabelas, listas, cadastros etc.), quanto não estruturados (*posts*, vídeos, *podcasts* etc.) podem ser submetidos, pelo próprio cidadão, ou por jornalistas profissionais, ao processo de transformação do dado em informação, da informação em conhecimento e do conhecimento em sabedoria (ACKOFF, 1989), revertendo-se, supostamente, a favor de algum interesse público.

Jornalistas de dados podem, dessa forma, não só obter autoridade para falar pelo cidadão por meios quantitativos, como “garimpar”, pela exploração relativamente livre de dados heterogêneos vindos de fontes desconexas, correlações inesperadas que viram pauta, trazendo para a esfera pública *insights* novos sobre problemas de interesse comum, ainda que daí não surjam vínculos causais que facultem deliberações racionais.

Os textos que possibilitam a circulação desses discursos de base quantitativista são representações verbo-visuais complexas e especializadas, chamadas visualizações e narrativas de dados (SEGEL; HEER, 2010; WANG *et al.*, 2015; BUZATO, 2018). Trata-se de gráficos, quadros e diagramas, muitas vezes interativos, de formatos inusitados para o leitor comum, que tornam correlações numéricas acessíveis à interpretação humana pela via da percepção visual (CLEVELAND; MCGILL, 1984; PINKER, 1990). Entendidas como visualizações de fenômenos, e não propriamente leituras de representações, as interpretações aí realizadas pretendem-se neutras e objetivas, como, de resto, pretende-se o próprio jornalismo. No entanto, trata-se de gêneros híbridos de narrativas e bancos de dados imbuídos, como quaisquer outros, de dimensões retóricas e estéticas, que estão abertos ao jogo discursivo, guardada sua especificidade, *i.e.*, o entrecruzamento do quantitativismo e da correlação com o qualitativismo e a causalidade.

É em torno da produção e recepção (interpretação, apropriação) desses gêneros que Torres (2019) vem desenvolvendo um projeto de pesquisa em jornalismo de dados.

## 5.2. O JORNALISMO DE DADOS COMO INDUTOR SOCIAL DO(S) LETRAMENTO(S) DE DADOS

O consórcio *Data-Pop Alliance* (2015; p. ii) define o jornalismo de dados como uma nova forma de jornalismo, estimulada sobretudo pelo movimento aberto de dados, em que as narrativas jornalísticas são complementadas por visualizações de dados. Já Barbosa e Torres (2013) são mais incisivos no sentido de explicar que não se trata apenas de inserir visualizações em notícias, mas de negociar a estrutura, composição e apresentação dos conteúdos jornalísticos a partir do trabalho quantitativo com os dados.

Exemplos de *breakthroughs* do jornalismo de dados não faltam. Um dos mais recentes e ruidosos foi o caso dos *Panama Papers*<sup>16</sup>, em que 370 jornalistas de mais de 100 veículos reunidos analisaram, com a ajuda de cientistas de dados, 12 milhões de documentos que revelaram a existência de 214 mil empresas de *offshore*<sup>17</sup> ligadas a um único escritório de advocacia utilizado para lavagem de dinheiro e evasão de divisas.

Há também os grandes veículos que apostam em editorias ou laboratórios de dados, entre eles<sup>18</sup> o *New York Times*, o *The Guardian*<sup>19</sup> e *O Estado de São Paulo*<sup>20</sup>. Além disso, há diversas iniciativas de setores da sociedade civil, algumas das quais identificadas com a cultura *hacker* e ativismos políticos, como o *Wikileaks*<sup>21</sup>. E, finalmente, iniciativas individuais de cidadãos, como a estudante de jornalismo Débora Sögur Hous, que usou dados do Portal da Transparência<sup>22</sup> para detectar inconsistências nas atribuições de bolsas de estudos na Universidade Federal do Paraná (UFPR). Seus achados geraram repercussão em grandes veículos da imprensa, trazendo, assim, uma contribuição sua, investida de dupla “autoridade”, dos números e da mídia tradicional, à esfera pública.

“Contar histórias” com dados (*data storytelling*) é, portanto, uma das formas pelas quais a cultura digital borra a divisa entre produtor e consumidor de textos midiáticos, extrapolando saberes disciplinares, ainda que por caminhos improvisados ou amadores. Tal transdisciplinaridade aparece não só nos textos daí resultantes, mas também nos papéis, métodos e saberes envolvidos na sua produção e recepção (interpretação, apropriação).

## 5.3. INVESTIGANDO O LETRAMENTO DE DADOS NO CONTEXTO DO JORNALISMO DE DADOS

Uma agenda de pesquisa em letramentos de dados que privilegie o fenômeno da linguagem em uso em situações de relevância social, ao gosto da Linguística Aplicada, precisa contemplar as práticas situadas nas quais se tornam evidentes os pressupostos culturais (ideológicos) e os papéis dos atores humanos e das tecnologias envolvidas em cada esfera ou comunidade discursiva, neste caso, a do jornalismo de dados.

Cabem nessa agenda, entre outras perguntas sobre a produção e a recepção (interpretação, apropriação) de visualizações e narrativas de dados jornalísticas, as seguintes:

<sup>16</sup> *The Panama Papers: Exposing the Rogue Offshore Finance Industry*. Disponível em: <http://bit.ly/2InE7IC>. Acesso em: 25 fev. 2019.

<sup>17</sup> Nomenclatura dada a empresas ou contas bancárias instauradas nos chamados paraísos fiscais, com o intuito de obterem vantagens financeiras na redução ou sonegação de impostos.

<sup>18</sup> Disponível em: <https://nytlabs.com/>. Acesso em: 25 fev. 2019.

<sup>19</sup> Disponível em: <https://www.theguardian.com/data/>. Acesso em: 19 fev. 2019.

<sup>20</sup> Disponível em: <http://blog.estadoadados.com/>. Acesso em: 25 fev. 2019.

<sup>21</sup> Disponível em: <https://wikileaks.org/>. Acesso em: 19 fev. 2019.

<sup>22</sup> Portal mantido pelo Governo Federal, desde 2004, com informações (dados), acerca de diversos aspectos da gestão pública, sobretudo quanto aos gastos realizados. Disponível em: <http://www.portaltransparencia.gov.br/>. Acesso em: 20 fev. 2019.

(i) de que forma os saberes, as habilidades, as prioridades e a autoridade do jornalista e do cientista de dados são negociados nessa nova forma de fazer notícia? (ii) Como e para que esses atores utilizam as tecnologias e os saberes disciplinares envolvidos? (iii) Como as visões quantitativas e qualitativas da realidade são negociadas nessa prática e nesses textos? (iv) Como as lacunas de formação disciplinar do leitor, do jornalista e do cientista de dados interferem nos eventos de produção e interpretação desses textos e na constituição dos objetos por eles referidos?

Tanto no lócus da produção quanto no da recepção, essas perguntas convocam métodos etnográficos e participativos, complementados por entrevistas, grupos focais e outros instrumentos de pesquisa que permitam partilhar o ponto de vista desses sujeitos sobre a natureza transdisciplinar da prática.

Partindo do pressuposto de que a tensão entre o quantitativismo e o qualitativismo vai transparecer, discursivamente, nas relações estabelecidas entre as pessoas, as tarefas/atividades e os textos, em ambos os contextos, nossa hipótese é a de que os saberes disciplinares quantitativos vão, preferencialmente, se impor aos qualitativos, o que afetará as relações de poder no seio da prática. Nesse caso, interessa saber sob que condições, mesmo que induzidas pelo pesquisador, os praticantes reverterão as lacunas percebidas em suas visões disciplinares não em critério de subordinação de uma visão à outra, mas na abertura do fenômeno visualizado/narrado. Essa informação, entendemos, pode ser útil para o desenho de estratégias educacionais voltadas para a formação de leitores de dados engajados, críticos e aptos à transdisciplinaridade.

## 6. CONSIDERAÇÕES FINAIS

Neste trabalho buscamos caracterizar o(s) letramento(s) de dados como conceito e como prática transdisciplinar, destacando o jornalismo de dados como lócus privilegiado para a sua pesquisa. É importante registrar que não vemos o jornalismo de dados como melhor ou pior, nem mais emancipador ou mais alienador do que os jornalisismos canônicos, mas que essa é uma esfera em que os significados sociais, ideológicos, acadêmicos e políticos da dadificação ficam mais evidentes.

A modesta agenda de pesquisa aqui sugerida está, direta ou indiretamente, vinculada a uma agenda educacional para a formação de cidadãos aptos a superar a cisão entre quantitativismo e qualitativismo no diálogo racional na esfera pública, a despeito dos descaminhos e colonizações já mencionadas. Essa, porém, é uma agenda a ser construída coletivamente, em outro lugar/oportunidade, por praticantes disciplinares (educadores e outros) que tenham o espírito e a mente abertos à transdisciplinaridade. Nenhuma agenda, contudo, será efetiva, sem o acesso democrático aos dados, hoje, em grande parte, reservados ao uso das empresas mantenedoras de mídias sociais, ou divulgados publicamente de forma desorganizada e refratária às capacidades de trabalho com dados do cidadão comum. Dito de outra forma, não basta que haja “a disposição” das comunidades para o engajamento com dados: os dados também precisam estar “à disposição” das comunidades, especialmente os que elas mesmas produzem, por vezes, sem sabê-lo.

## Referências

AUSTIN, John. *How to do things with words*. Cambridge: Harvard University Press, 1962.

ACKOFF, Russel. From data to wisdom. *Journal of Applied Systems Analysis*, [S.L]. v. 15, p. 3-9, 1989.

ANDRIOLO, Eric. *Desvendando “Data Scraping”*: Entenda como raspar dados pode facilitar o trabalho jornalístico. Blog. Disponível em: <https://knightcenter.utexas.edu/pt-br/blog/00-9586-desvendando-o-data-scraping-entenda-como-raspar-dados-pode-facilitar-o-trabalho-jornali>. Acesso em: 15 maio. 2019.

ATHANASES, Steven; BENNETT, Lisa; WAHLEITHNER, Juliet. Fostering Data Literacy Through Preservice Teacher Inquiry in English Language Arts. *The Teacher Educator*, [S.L]. v. 48, n. 1, p. 8–28, jan. 2013.

BARBOSA, Suzana Oliveira; TORRES, Vitor. O paradigma ‘Jornalismo Digital em Base de Dados’: modos de narrar, formatos e visualização para conteúdos. *Galaxia*, [S.L]. n. 25, p. 152-164, jun. 2013.

BHARGAVA, Rahul.; D’IGNAZIO, Catherine. “Designing Tools and Activities for Data Literacy Learners”. In: BHARGAVA, Rahul.; D’IGNAZIO, Catherine. *Data Literacy Workshop at Web Science*. Oxford, UK: University of Oxford, 30 jun., 2015.

BOYD, Danah.; CRAWFORD, Kate. Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, [S.L]. v. 15, n. 5, p. 662–679, jun. 2012.

BUZATO, Marcelo E. K. “Data Storytelling e a dadificação de tudo: um gênero bastardo de mãe narrativa e pai banco de dados”. In: LIMA-LOPES, R. E. DE; BUZATO, M. E. K. (Org.). *Gênero Reloading*. Campinas, SP: Pontes Editores, p. 95–123, 2018.

BUZATO, Marcelo E. K. “Critical Data Literacies: going beyond words to challenge the illusion of a literal world”. In: TAKAKI, N. H.; MONTE MOR, W. (Org.). *Construções de sentido e letramento digital crítico na área de línguas e linguagem*. Campinas, SP: Pontes Editores, p. 119–142, 2017.

BUZATO, Marcelo E. K. “Mapping Flows of Agency in New Literacies: Self and Social Structure in a Post-social World”. In: JUNQUEIRA, E. S.; BUZATO, M. E. K. (Org.). *New Literacies, New Agencies: a Brazilian perspective*. New Literacies and Digital Epistemologies. New York: Peter Lang, p. 22–49, 2013.

BUZATO, Marcelo E. K. Letramento e inclusão: do estado-nação à era das TIC. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, São Paulo. v. 25, n. 1, p. 01-38, 2009.

CLEVELAND, William.; MCGILL, Robert. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, [S.L]. v. 79, n. 387, p. 531–554, set. 1984.

COPE, Bill; KALANTZIS, Mary. Big Data Comes to School: Implications for Learning, Assessment, and Research. *AERA Open*, [S.L]. v. 2, n. 2, 1 abr. 2016.

COPE, Bill; KALANTZIS, Mary. *Multiliteracies: literacy learning and the design of social futures*. London; New York: Routledge, 2000.

DATA POP ALLIANCE. *Beyond Data Literacy: Reinventing Community Engagement and Empowerment in the Age of Data*. (White Paper) 2015. Disponível em: <http://bit.ly/2tM5uUM> . Acesso em: 15 ago. 2017.

D'IGNAZIO, Catherine; BHARGAVA. "Rahul. Approaches to Building Big Data Literacy". In: D'IGNAZIO, Catherine; BHARGAVA. *BLOOMBERG DATA FOR GOOD EXCHANGE*, 2014, New York. *Anais...* New York: Semantic Scholar, 2015, [s/p].

EMC EDUCATION SERVICES. *Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Indianapolis, IN, USA: John Wiley & Sons, Inc, 2015, [s/p].

ESPELAND, Wendy; STEVENS, Mitchell. A Sociology of Quantification. *European Journal of Sociology*, [S.L]. v. 49, n. 03, p. 401, dez. 2008.

FREIRE, Paulo. *Pedagogia do oprimido*. Rio de Janeiro: Paz e Terra, 1970.

GRAY, Jonathan *et al.* "Ways of Seeing Data: Towards a Critical Literacy for Data Visualizations as Research Objects and Research Devices". In: KUBITSCHKO, S.; KAUN, A. (Org.). *Innovative Methods in Media and Communication Research*. London, UK: Palgrave, p. 227–251, 2016.

HABERMAS, Jurgen. The Public Sphere: An Encyclopedia Article. *New German Critique*, [S.L]. n. 3, p. 49-55, 1974.

HEATH, Shirley Brice. *Ways with words: language, life, and work in communities and classrooms*. Cambridge: Cambridge University Press, 1983.

IGO, Sarah Elizabeth. *The averaged American: surveys, citizens, and the making of a mass public*. Cambridge, Mass.; London: Harvard University Press, 2008.

KATO, Mary. *No mundo da escrita: uma perspectiva psicolinguística*. São Paulo: Ática, 1986.

KITCHIN, Rob. Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, [S.L]. v. 1, n. 1, 1 abr. 2014.

KHAMIS, Susie; ANG, Lawrence; WELLING, Raymond. Self-branding, 'micro-celebrity' and the rise of Social Media Influencers. *Celebrity Studies*, [S.L]. v. 8, n. 2, p. 191–208, 3 abr. 2017.

KOCKELMANS, Joseph. "Why interdisciplinarity?" In: KOCKELMANS, Joseph. (Ed.). *Interdisciplinarity and higher education*. University Park, PA: Pennsylvania State University Press, p. 123-160, 1979.

LATOUR, Bruno. *Pandora's hope: essays on the reality of science studies*. Cambridge, Mass.: Harvard University Press, 1999.

LATOUR, Bruno. Visualization and cognition: drawing things together. *Knowledge and society*, [S.L]. v. 6, p. 1–40, 1986.

- LEMKE, Jay L. Letramento metamidiático: transformando significados e mídias. *Trabalhos em Linguística Aplicada*, [S.L]. v. 49, n. 2, p. 455–479, dez. 2010.
- LEMKE, Jay L. Across the Scales of Time: Artifacts, Activities, and Meanings in Ecosocial Systems. *Mind, Culture, and Activity*, [S.L]. v. 7, n. 4, p. 273 – 290, 2000.
- LEMKE, Jay L. *Textual politics: discourse and social dynamics*. London; Bristol, PA: Taylor & Francis, 1995.
- LIMA JUNIOR, Walter Teixeira; ROCHA, Ana Paula. Visualização de informações estruturada por bancos de dados digitais: o Jornalismo em sintonia com a complexidade informativa contemporânea. *Líbero*, [S.L]. v. 13, n. 26, p. 51-62, dez. de 2010.
- MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. *Big data: a revolution that will transform how we live, work, and think*. Boston: Houghton Mifflin Harcourt, 2013.
- MARWICK, Alice; LEWIS, Rebecca. *Media Manipulation and Disinformation Online*. New York, N.Y: Data & Society Research Institute, 2017.
- MIRANDA, Clarissa Mazon.; SCHAEFER, Ricardo.; MEDEIROS, Vicente Reis. “O jornalismo como mediador de consciência e operador de realidade”. In: FUNDAÇÃO ANTONIO MENEGHETTI (Org.). *Ontopsicologia: ciência interdisciplinar*. Recanto Maestro: Fundação Antonio Meneghetti, p. 197-215, 2015.
- MCLUHAN, Marshall. *A galáxia de Gutenberg: a formação do homem tipográfico*. São Paulo, Editora Nacional, Editora da USP, 1972.
- MORIN, Edgar. *Os sete saberes necessários à educação do futuro*. Tradução Catarina Eleonora F. Da Silva. São Paulo: Cortez, 2004.
- NICOLESCU, Basarab. *Manifesto of transdisciplinarity*. Albany, NY: State University of New York Press, 2002.
- O’NEIL, Cathy. *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: B/D/W/Y Broadway Books, 2017.
- PARRY, Milman. *The Making of Homeric Verse: The Collected Papers of Milman Parry*. Oxford: Clarendon Press, 1971.
- PIAGET, Jean. “The epistemology of interdisciplinary relationships”. In: CENTRE FOR EDUCATIONAL RESEARCH AND INNOVATION (CERI), *Interdisciplinarity: Problems of teaching and research in universities*. Paris, France: Organisation for Economic Co-operation and Development, p. 127-139, 1972.
- PINKER, Stephen. “A theory of graph comprehension”. In: FREEDLE, R. (Org.). *Artificial intelligence and the future of testing*. Hillsdale, NJ: Erlbaum, p. 73–126, 1990.
- SCHÄFER, Mirko Tobias; VAN ES, Karin (Org.). *The Datafied Society. Studying Culture through Data*. Amsterdam: Amsterdam University Press, 2017.

SCHIELD, Milo. Information Literacy, Statistical Literacy, Data Literacy. *IASSIST Quarterly*, [S.L]. v. 28, p. 6–11, 2004.

SCRIBNER, Sylvia; COLE, Michael. *The psychology of literacy*. Cambridge, Mass.: Harvard University Press, 1981.

SEGEL, Edward; HEER, Jeffrey. Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, [S.L]. v. 16, n. 6, p. 1139–1148, nov. 2010.

SOARES, Magda. Língua escrita, sociedade e cultura: Relações, dimensões e perspectivas. *Revista Brasileira de Educação*, [S.L]. n. 0, p. 05-16, 1995.

STREET, Brian. *Literacy in theory and practice*. Cambridge [Cambridgeshire]; New York: Cambridge University Press, 1984.

THE NEW LONDON GROUP. A Pedagogy of Multiliteracies: Designing Social Futures. *Harvard Educational Review*, [S.L] v. 66, n. 1, p. 60–93, abr 1996.

TORRES, Cleyton Carlos. *A linguagem digital na era do Big Data: a visualização jornalística de dados como vetor ao data literacy*. (Projeto de pesquisa de Doutorado em Linguística Aplicada) Universidade Estadual de Campinas, Campinas, 2019.

VAHEY, Philip *et al.* “Thinking With Data: A Cross-Disciplinary Approach To Teaching Data Literacy And Proportionality”. In: THE ANNUAL CONFERENCE OF THE AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, [s/n.], 2010, Denver, CO. *Anais...* Denver, CO: AERA, 2010, [s/p.].

WANG, Lidong; WANG, Guanghui; ALEXANDER, Cheryl. Big Data and Visualization: Methods, Challenges and Technology Progress. *Digital Technologies*, [S.L]. v. 1, n. 1, p. 33–38, 2015.

WOLFF, Annika; KORTUEM, Gerd; CAVERO, Jose. “Urban Data in the primary classroom: bringing data literacy to the UK curriculum”. In: DATA LITERACY WORKSHOP, [s/n.], jun. 2015, Oxford, UK. *Anais...* Oxford, UK: Web Sci, 2015 [s/p.].

Marcelo El Khouri Buzato  
[marcelo.buzato@gmail.com](mailto:marcelo.buzato@gmail.com)

Cleyton Carlos Torres  
[ton@blogmidia8.com](mailto:ton@blogmidia8.com)

Recebido em: 09 de Julho de 2019

Aceito em: 15 de Agosto de 2019

Publicado em: Setembro de 2019