

Um balanço da utilização das ferramentas de auxílio à pesquisa qualitativa (CAQDAS): Trinta anos de MST na Mídia Impressa (1984-2014).

Lidia Maria Soares Pires Cardel¹

Paulo Henrique Dantas Pita²

Marina Ferreira de Araújo Fernandes³

Lara Rosa Meirelles Barros⁴

Caio Araújo dos Santos⁵

RESUMO

A finalidade deste artigo é relatar a relevância da utilização dos CAQDAS (computer assisted qualitative data analysis software) na coleta e na análise dos discursos da Mídia Impressa Nacional publicados desde o aparecimento do MST (Movimento dos Trabalhadores Rurais Sem Terra), 17 de abril de 1984, até o ano de 2014, quando o movimento completou 30 anos de existência. Este levantamento justifica-se pela importância do MST na formação de um novo cenário sobre a questão fundiária no Brasil, como também, na capacidade deste movimento social em influenciar o Estado Brasileiro na elaboração e produção de “políticas públicas transversais” (Brito Ivo, 2005), que impulsionaram a mitigação da pobreza no meio rural, a concepção de novas tecnologias para a agricultura familiar e a criação de novos espaços educacionais voltados para a população ligada à agricultura familiar e agroecológica. Por outro lado, visando complexificar o objetivo desta investigação, esta pesquisa também teve como meta compreender o processo de burocratização do MST (Movimento dos Trabalhadores Rurais Sem Terra) a partir do momento em que este Movimento Social escolheu aproximar-se da estrutura burocrática do Estado, elegendo suas lideranças para cargos no legislativo, como também, cedendo seus quadros para governos municipais, estaduais e federais.

ABSTRACT

¹ Professora Associada III do Departamento de Sociologia da Universidade Federal da Bahia (UFBA), membro permanente da Pós-Graduação em Ciências Sociais (UFBA), pesquisadora convidada da Universidade de Brasília e Coordenadora do Núcleo de Estudos Ambientais e Rurais.

² Mestrando do Programa de Pós-graduação em Sociologia – UFBA, Bacharel em Ciências Sociais com habilitação em Sociologia pela Universidade Federal da Bahia, pesquisadora do Núcleo de Estudos Ambientais e Rurais.

³ Bacharel em Ciências Sociais com habilitação em Sociologia pela Universidade Federal da Bahia, pesquisadora do Núcleo de Estudos Ambientais e Rurais.

⁴ Bacharel em Humanidades pela Universidade Federal da Bahia, graduanda em Ciências Sociais na UFBA, bolsista PIBIC do Núcleo de Estudos Ambientais e Rurais da Universidade Federal da Bahia.

⁵ Graduando em Psicologia pelo IPS-Universidade Federal da Bahia, bolsista PIBIC do Núcleo de Estudos Ambientais e Rurais da Universidade Federal da Bahia.

The purpose of this article is to report the relevance of the use of CAQDAS (computer-aided qualitative data analysis software) in the collection and analysis of national print media discs since the appearance of the MST (Movimento dos Trabalhadores Rurais Sem Terra), April 17, 1984, through the year 2014, When the movement has completed 30 years of existence. This research is justified by the importance of the MST in the formation of a new scenario on land problems in Brazil, as well as in the capacity of this social movement to influence the Brazilian State in the elaboration and production of "transversal public policies" (Brito Ivo, 2005).), Which promoted rural poverty alleviation, a conception of new tools for family farming, and the creation of new educational spaces for a population linked to family and agroecological agriculture. In order to complicate the objective of this research, this research also had as its purpose the process of bureaucratization of the MST from the moment this Social Movement chose to approach the bureaucratic structure of the State, electing its leadership to positions in the legislature, as well as , Giving his paintings to municipal, state and federal governments

INTRODUÇÃO

Objetivamos neste artigo analisar metodologicamente quais os ganhos e quais as limitações da utilização dos softwares na pesquisa sociológica com recortes temporais de médio e longo alcance. No caso específico da nossa pesquisa, procuraremos relatar quais foram os ganhos e as dificuldades encontradas ao utilizarmos tais recursos na coleta e análise sobre os discursos da Mídia Impressa Nacional publicados desde o aparecimento do Movimento dos Trabalhadores Sem Terra (MST), em 17 de abril de 1984, até o ano de 2014, quando o movimento completou 30 anos de existência. Este levantamento justifica-se pela importância do MST na formação de um novo cenário sobre a questão fundiária no Brasil, como também, na capacidade deste movimento social em influenciar o Estado Brasileiro na elaboração e produção de “políticas públicas transversais” (Brito Ivo, 2005)⁶, que impulsionaram a mitigação da pobreza no meio rural, a criação de novas tecnologias para a agricultura familiar e a criação de

⁶ Em IVO, Anete B. L. . La destitución de lo social: Estado, gobierno y políticas sociales. *Estudios Sociológicos*, Ciudad de México, v. XXIII, n.68, p. 347-374, 2005.

novos espaços educacionais voltados para a população ligada à agricultura familiar e agroecológica. Por outro lado, visando complexificar o objetivo da investigação proposta, a pesquisa central também tem como meta compreender o processo de burocratização⁷ do MST a partir do momento em que este Movimento Social escolheu aproximar-se da estrutura burocrática do Estado, elegendo suas lideranças para cargos no legislativo, como também, cedendo seus quadros para governos municipais, estaduais e federais.

Assim sendo, com a pesquisa ainda em curso, esperamos encontrar no discurso da Mídia Impressa dos últimos 30 anos, elementos esclarecedores para elucidar o papel do MST no cenário político nacional contemporâneo. Entretanto, em função do recorte temporal dilatado para uma investigação qualitativa de cunho sociológico, foi necessário delimitar a abrangência das Mídias escritas a serem pesquisadas. Deste modo, por meio da democratização da produção textual, e pela importância alcançada por alguns grupos editoriais em nível nacional, optamos por trabalhar com o acervo de três grandes Jornais brasileiros, quais sejam, a Folha de São Paulo, o jornal O Globo, e o jornal O Estado de São Paulo, que já possuem todo o seu material digitalizado e disponibilizado para pesquisas desta natureza.

A BASE EMPÍRICA DA PESQUISA

Diante do volume dos dados coletados até o momento apenas na Mídia o Estado de São Paulo⁸ e na busca por operacionalizar metodologicamente o objeto pesquisado, optamos pela utilização dos softwares de apoio a análise de dados qualitativos (Computer Assisted Qualitative Data Analysis Software – CAQDAS) e quantitativos. Assim, após a análise amostral da primeira coleta de dados⁹ reunidos por meio do software R¹⁰,

⁷ No sentido weberiano, onde o processo de burocratização está intimamente ligado ao processo da dominação legítima.

⁸ A base de dados do Estado de São Paulo possui 8.264 arquivos em PDF, e cada arquivo possui de duas a três matérias sobre a temática estudada. Estes arquivos foram inicialmente coletados em formato JPEG e convertidos para PDF com caracteres reconhecidos por meio do software ABBYY.

⁹ Para analisar todo o material coletado, utilizamos a técnica de análise de discurso que nos facultará a condição da análise objetiva e subjetiva dos dados coletados. Estas técnicas transversalizadas de utilização de softwares e análise de discurso nos darão a condição de termos acesso ao contexto

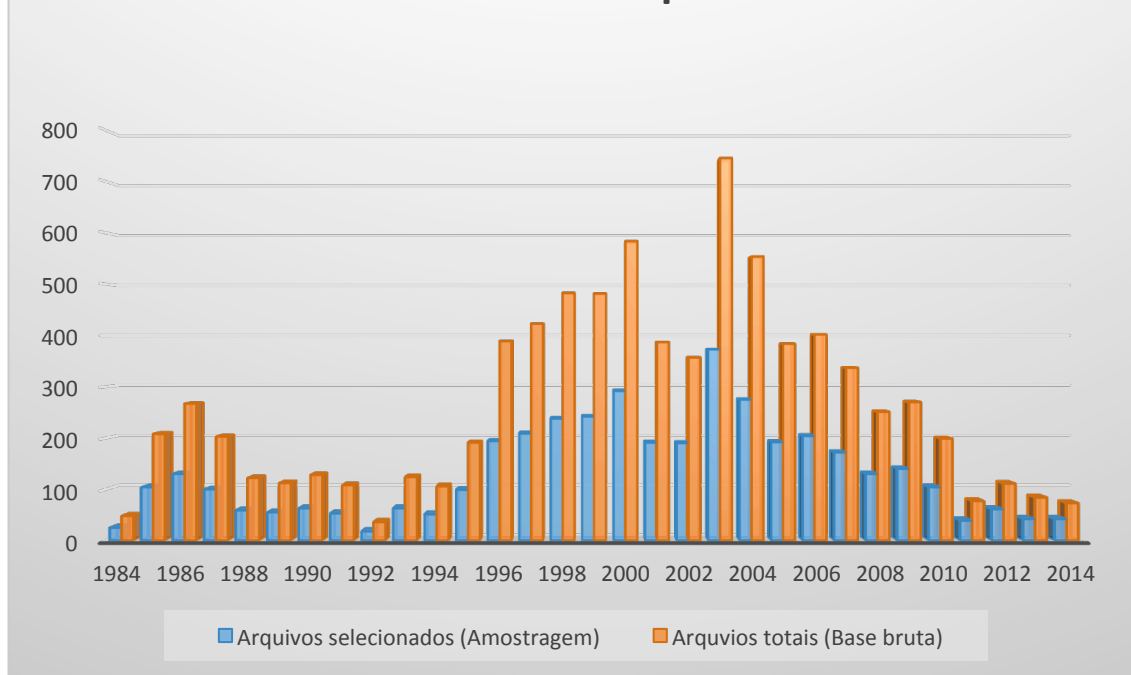
passamos a utilizar o software Atlas.ti para a análise qualitativa dos dados textuais, imagéticos e audiovisuais, e do software Sphinx que nos possibilita a elaboração de um questionário digital que serve para aglutinar quantitativamente os dados da mesma temática, possibilitando a criação de tabelas, gráficos, cruzamentos de variáveis e análises estatísticas¹¹. Para tanto, construímos um manual de códigos por meio da leitura prévia de 10 % do total de 8.264 arquivos com uma ou mais matérias recolhidas no acervo do Jornal O Estado de São Paulo entre os anos de 1980 (momento que antecede a criação do MST, em 1984, mas que já nos apontava a luta pela terra no campo) à 2014, totalizando 834 matérias lidas, debatidas e codificadas à luz da teoria sociológica e do arcabouço epistemológico da Sociologia Rural, da Sociologia das Mídias e dos Movimentos Sociais para a elaboração de palavras-chave.

imediate da enunciação, e no sentido mais amplo, compreender a dimensão e o contexto sócio-histórico e ideológico do objeto estudado.

¹⁰ . Por meio deste software elaboramos um script para a realização de download em massa de matérias relacionadas às palavras-chave (MST, Sem Terra, Movimento dos Trabalhadores Sem terra) da pesquisa para executar a coleta temática de matérias nos acervos coletado dos jornais que estão sendo estudados.

¹¹ Para clarificar as etapas de utilização dos softwares é importante demonstrar o desenho metodológico que elaboramos para nossa pesquisa: 1. Coleta de dados em acervos digitais por meio de script e software R; 2. Formatação da base de dados; 3. Elaboração de amostragens aleatórias (divisão em decênios); 4. Elaboração de novas amostragens aleatórias dentro dos decênios; 5. Levantamento de palavras chaves de matérias em cada decênio; 6. Equalizamento de palavras chave em categorias para estruturação de questionário digital no Sphinx e códigos no Atlas.TI.; 7. Criação de um manual de preenchimento e operacionalização das categorias elaboradas.; 8. Aplicação de questionários e codificação das matérias visando relação com eixos de análises elaborados.

Gráfico comparativo de arquivos selecionados X arquivos totais.



Fonte: Gráfico gerado com dados coletados pela pesquisa.

Assim, por meio deste gráfico, demonstramos que o recorte amostral possível para a nossa pesquisa - que possui uma base de dados de grande extensão com curto período para a sua execução - mesmo utilizando softwares com grande capacidade de abrangência analítica, não supera mais de 50% do total dos arquivos coletados pelo Script R. Isso se dá pelas limitações impostas pela temática e pelo recorte temporal da pesquisa, ou seja, ao nos propormos a investigar toda a produção textual de uma Mídia escrita num período de 30 anos, nos deparamos com uma extensão de dados (Big data) que se fossem trabalhados *in totum*, consumiriam um tempo inexecuível de organização do material. E, outra questão central em termos metodológicos é que por ser uma análise sociológica, priorizamos a abordagem analítica que vai além de uma abordagem estatística e descritiva.

Neste sentido, a literatura sobre a qualidade e a abrangência destes softwares nos indicaram a possibilidade da criação de modelos conceituais e redes de informações sobre a trajetória do MST entre 1984 a 2014, evidenciando a conexão entre os dados, além de oferecerem recursos para a geração dos relatórios e a visualização dos resultados. Porém, no decorrer da pesquisa, tomamos consciência epistemológica de que

nenhum destes softwares estabeleceriam os caminhos a serem seguidos, como também, não seriam ferramentas epistemológicas sobre como analisar os dados ou como elaborar a codificação. Estas são responsabilidades do pesquisador, que deve estar munido de um aparato teórico sólido sobre o tema pesquisado, no nosso caso, sobre o papel social dos movimentos sociais e as estruturas sócio-políticas do mundo rural brasileiro.

Ainda que se questione o quanto legítimo em termos científicos seja a utilização dos Caqdas na pesquisa, negar o quão facilitadores são, no que tange ao tratamento de dados, denota uma falta de aproximação e contato com a guinada digital que as ciências humanas vêm tomando. Por intermédio do software, passa-se a ter recursos sumamente importantes para o pesquisador, pois, além da possibilidade de se armazenar e agrupar uma quantidade grandiosa de informações acerca do objeto investigado, ainda é possível a extração de seus resultados e o cruzamento das informações obtidas. O problema central que nos deparamos no processo foi como equalizar os percalços e as oportunidades nessa utilização a fim de criar um método de fato eficiente e válido para trabalharmos com uma realidade de dados bastante extensa. O desafio não está posto necessariamente em relação ao quesito “tamanho dos dados”, e sim na complexidade que se expande quando pensamos em minimizar os possíveis erros metodológicos durante o processo. No caso da nossa pesquisa, esbarramos na extensão temporal da temática estudada. Nos propomos, no início, a trabalhar os 30 anos da história do MST de forma monolítica. Entretanto, o tempo histórico e político brasileiro entre o período de 1984 à 2014 apresentou mudanças e rupturas sociais, econômicas e culturais profundas, e os nossos dados refletiam estas complexidades. Após nos vermos impedidos empiricamente de criar categorias totalizantes destes trinta anos da história do MST, decidimos dividir a base de dados em três decênios (1984-1994; 1995-2005; 2006-2014). Esta proposta seguiu o critério das mudanças sócio-políticas presente na história recente do país, como a abertura política iniciada em 1984 (o Movimento das Diretas Já), as eleições presidenciais de Fernando Henrique Cardoso (1994-2002), de Luís Inácio Lula da Silva (2003-2011) e de Dilma Rousseff (2011-2016), como também, o apogeu da visibilidade dos movimentos sociais por meio de políticas públicas restitutivas e redistributivas (Honneth e Frazer, 2006)

Dessa forma, mediante o exposto acima, constatamos que uma pesquisa que vise como resultado um trabalho metodologicamente confiável e com relevância científica precisa

problematizar o método previamente com equidade, a fim de dar conta antecipadamente da maioria dos problemas que poderão se tornar presentes em seu estudo. Essa etapa é o que chamaremos de fase exploratória, onde o tema a ser pesquisado possui uma amplitude que deverá ser remodelada epistemologicamente pelos conceitos teóricos e metodológicos do campo científico escolhido. É importante que o pesquisador visualize esse momento com sensatez que o distancie do senso comum erudito¹², pelo fato desse momento singular de construção do modelo e do estado da arte se tonar posteriormente a base de todas as diretrizes que serão tomadas dali para frente. Nele, é imprescindível que se tenha noção da realidade de dados que irá processar, assim como o tipo de análise qualitativa que será feita, uma vez que o tipo de investigação interfere na ferramenta a ser utilizada¹³. Para Lage (2011), a seleção do Caqdas deve estar diretamente relacionada à abordagem metodológica eleita pelo pesquisador, considerando que todo software de análise qualitativa é elaborado sobre um tipo de metodologia específico, proposto pelo seu desenvolvedor. Todavia, na prática, outros fatores precisam ser levados em consideração e ponderados pelo grupo de pesquisa que estiver trabalhando. No caso da pesquisa que estamos elaborando, os dados brutos coletados se mostraram inoperantes numa primeira abordagem. Quando começamos a lançar mão do conhecimento sociológico adequado, as ferramentas epistemológicas da sociologia das Mídias, da sociologia rural e dos movimentos sociais foram moldando o método a ser utilizado para depurar o material coletado. E assim, os *links* e os *insights* começaram a dar forma e esteio ao tratamento dos dados. Foi só após superarmos estas dificuldades que a pesquisa começou a apresentar os primeiros resultados.

Em termos práticos, foi necessário adequar (para além da metodologia, pensando também em quesitos operacionais ligados a informática) os dados à plataforma (software) de análise que usamos em relação ao tipo de documento, tamanho, forma de armazenamento e recursos analíticos disponíveis. Entre um software e outro, por mais que se busque cumprir com o mesmo tipo de análise, há uma diferença considerável no que diz respeito a seus recursos e possibilidades. Aqui entra a perspectiva do recorte empírico da pesquisa e a escolha do arcabouço teórico e epistemológico.

¹² .Como aponta Bourdieu, 2002.

¹³ O medo de apagões sistemáticos nos aplicativos e o receio de perder repentinamente meses e até anos de pesquisa pelo programa “travar” por meio de *bugs* e *crash* geram dúvidas pertinentes sobre a confiabilidade dos Caqdas na área de pesquisas qualitativas.

Lewins & Silver (2009), recomendam que o pesquisador responda a um conjunto de questões antes de partir para a escolha de seu Caqdas. De forma resumida, algumas delas são: qual o tipo e volume de dados de sua pesquisa? Qual é a sua abordagem teórica para o tratamento dos dados? Qual é o tipo de questionamento ou consultas que precisará fazer aos dados? Será necessário usar algum processo de codificação (em caso afirmativo, quais as suas características)? Será preciso trabalhar com temas e/ou categorias de informações? A pesquisa é executada em grupo? Em síntese, prepara-se um terreno fértil para a adoção das técnicas mais propícias a cada contexto; e essa é tarefa intransferível do pesquisador.

Superada a definição da ferramenta, todos os olhares direcionam-se à amostra, tendo ela total importância entre a fase exploratória e o início do estado da arte sobre a temática explorada. Ao considerar a amostra o principal objeto deste trabalho, idealmente, a ferramenta precisa se adequar a ela e não o contrário. No nosso caso, enfrentamos dificuldades quanto a isso por conta de alguns fatores. O primeiro, devido ao tamanho elevado do nosso universo de matérias. Ao propor trabalhar com trinta anos de veiculações nos jornais O Estado de São Paulo, Folha de São Paulo e O Globo sobre todos os assuntos veiculados relativos ao MST, não tínhamos o completo conhecimento de que o processo de coleta nos renderia alguns milhares de matérias pela frente. Com isso, enfrentamos alguns percalços inesperados, devido a tamanha despreensão nossa da complexidade de alguns quesitos. Por exemplo, o armazenamento. Armazenar um, dois gigabytes é mais simples, uma vez que diferentes provedores oferecem estes recursos gratuitamente. Por outro lado, armazenar dez, vinte gigabytes é um processo completamente diferente, não só por conta da operação em si, mas também pelo tamanho incompatível com a maioria dos pens drives disponíveis no mercado. Isso se agrava quando pensamos que além de armazenar, precisaríamos manipulá-los conjuntamente. A saída foi armazená-los em nuvem, utilizando um servidor pago para isso. Caso esta resolução tivesse sido pensada antecipadamente, poderíamos chegar a uma outra solução, inclusive, gratuita. Outro entrave diz respeito ao quesito codificação. Possuímos um sujeito a analisar (o Sem Terra) que “muda de figura” nos discursos da Mídia ao longo dos 30 anos de forma latente. Isso traz um desafio na codificação que é exatamente o de criar categorias que se proponham a dar conta dessas diferentes leituras a partir da reportagem, e que viabilizem uma análise concisa posteriormente. O desafio estabelece-se, sobretudo, no momento em que essas categorias de análise devem abarcar

toda a realidade posta, para que seja viável uma análise de conteúdo/discurso¹⁴, inclusive atentando-se para que não se caia no fetichismo da codificação exacerbada, um problema comum quando se lida com uma gama de dados extensos e heterogêneos. É fundamental, nesse sentido,

considerar que os principais softwares atuais para apoio à análise de dados qualitativos foram desenhados com base na premissa de que o pesquisador precisa tanto de proximidade quanto de distância de seus dados (RICHARDS, 1998; BAZELEY, 2007 apud (LAGE & GODOY, 2008)).

Vencida a fase exploratória constatamos que esse é o momento mais trabalhoso do estudo, sendo essa etapa muito mais laboral do que organizacional/analítica. É necessário frisar que todo o trabalho presente nessa fase inicial consiste na diminuição de impactos negativos que venham a emergir durante a pesquisa em si, e que viabilizem uma análise posterior eficiente e rica em substrato para a análise.

Uma das técnicas mais utilizadas para superar as dificuldades da análise de uma grande base de material é a construção de uma base amostral. No nosso caso, com o auxílio dos softwares de análise qualitativo e quantitativo, foi possível trabalhar com 50 % dos dados coletados, um percentual bem acima da média utilizada e que nos trouxe ganhos analíticos sólidos e confiáveis. Com efeito, uma amostra tão alta lança luz e converge na direção de uma percepção mais apurada e ao mesmo tempo precisa acerca da realidade dos dados, sobretudo no momento inicial.

Assim, visando compreender a história e os processos sociológicos que envolvem a relação do MST com o Estado, essa pesquisa tem um viés inovador ao propor trazer à luz, por meio do discurso sociológico, todas as matérias já publicadas (nas Mídias escritas O Estado de São Paulo, A Folha de São Paulo e O Globo) sobre esse movimento social desde a sua fundação em 1984 até o ano de 2014, quando a sua relação com o governo torna-se tensa em função da escolha explícita da então presidente Dilma Rousseff pela política de incentivo ao agronegócio. Para tanto, optamos por iniciar a coleta de dados pelo jornal O Estado de São Paulo, pela sua complexidade em relação a temática pesquisada. Na fase exploratória, por meio de

¹⁴ No caso dessa pesquisa, após levantarmos mais de uma centena de categorias empíricas referente ao objeto estudado criamos um questionário único que nos possibilitou instrumentalizar, por meio de uma análise sociológica apurada, conceitos-chave que emergiram da análise sobre a enorme complexidade social encontrada no material analisado.

rodadas de leitura¹⁵ das amostragens das matérias, foram encontrados elementos fundamentais que justificaram a sua escolha como ponto de partida na pesquisa, a exemplo do caderno “Questão Agrária”, criado desde a década de 80, com a finalidade de reunir, independente da região do país ou dos agentes envolvidos, todas as notícias sobre as ações relacionadas ao MST e às questões fundiárias em nível nacional.

Ao optarmos por pesquisar sociologicamente trazendo para o escopo do trabalho todas as matérias já publicadas desde a fundação do movimento em 1984 até 2014, após 30 anos de atuação no cenário nacional sobre a luta pela terra, já sabíamos de antemão, posteriormente à revisão bibliográfica, que não havia ainda nenhuma investigação desta magnitude. Obviamente, o desafio só foi lançado por termos como ferramentas de pesquisa os programas Atlas.ti, Sphinx e R como ferramentas fundamentais para buscar, organizar e analisar o volume de matérias coletada através da busca por palavras-chave e categorias específicas, uma vez que pelo recorte temporal e pela quantidade de dados seria extremamente laborioso, ou mesmo inviável, trabalhar estimadamente com mais 10.000¹⁶ matérias por Mídia sem os recursos desses softwares. Obviamente, estas ferramentas não trabalham por si e aqui entra o papel decisivo do pesquisador. Assim, por meio das leituras prévias e da criação de categorias sociológicas e de mediação, conseguimos até o momento atual de pesquisa, trabalhar com 50% do universo das matérias coletadas.

Em suma, o que buscamos com esta pesquisa? Em um sentido amplo, ela visa dar continuidade aos estudos realizados na linha de pesquisa sobre Mídia e Movimentos Sociais, direcionando para a análise da Mídia impressa pela sua importância enquanto construtora de um discurso multifocal do cotidiano brasileiro. No sentido específico dos estudos sobre o MST, esperamos contribuir para a compreensão da dinâmica política interna e externa desse movimento identificando os atores sociais e os interlocutores que são chamados para falar a favor e/ou contra o MST. Será possível, inclusive, elencar mudanças de comportamento de ações e narrativas do MST ao longo dos seus

¹⁵ A técnica de rodadas de leitura se refere à fase exploratória da referida pesquisa. A base de dados foi dividida em 3 decênios para facilitar essa etapa da pesquisa, e por cada decênio foram realizadas 5 amostragens, o que resultou em 15 rodadas de leitura. Cada rodada envolveu a leitura coletiva dos pesquisadores que levantaram as categorias, os posicionamentos discursivos, os atores e a representação social cunhada sobre o MST em cada decênio. Desse modo, através dessa fase da pesquisa, foi possível encontrar padrões da mídia impressa sobre o MST por cada 10 anos.

¹⁶ A estimativa foi feita levando em conta o total de 8.264 arquivos encontrados na Mídia impressa O Estado de São Paulo. Cada arquivo possui uma ou mais matérias como também gráficos e imagens.

30 anos, assim como as suas transformações referentes às pautas e estratégias de ações relativas à reforma agrária e as temáticas que tangenciam as políticas públicas sobre a estrutura fundiária brasileira.

Responder essas questões e tabular todas as matérias por meio da análise de discurso, de gráficos e de nuvem de palavras¹⁷ são aparatos metodológicos fundamentais dentro das Ciências Humanas. Na Sociologia contemporânea, o estudo sobre movimentos sociais e Mídia têm permitido compreender os movimentos para além de sua atuação interna, pois a interação destes com os meios de comunicação é central para o entendimento e a consolidação das suas dinâmicas.

Segundo Penn, Gemma,

A semiologia e a análise de conteúdo [e discurso] são consideradas, muitas vezes, como sendo instrumentos de análise radicalmente diversos, mas, como afirmam tanto Leiss et al. (1977), como Curan (1976), há muitas razões para uma aproximação. Os semiólogos podem incorporar os procedimentos sistemáticos de amostragem da análise de conteúdo. Isto levará, de alguma maneira, a discutir as críticas de que o enfoque produz resultados autoconfirmadores, e de que não é legítimo generalizar as conclusões de uma análise semiológica para outro material. A sistematização mais aprimorada da análise, que a análise de conteúdo [e discurso] defende, pode levar também a ajudar o semiólogo a combater acusações de seletividade (por exemplo, na construção de inventários de denotação e matrizes de possíveis sintagmas). A análise resultante deverá ser mais fidedigna (replicável) e menos dependente de idiosincrasias e habilidades de determinados analistas. (2002:339).

Seguindo os passos analíticos dessa autora, ressaltamos que a utilização dos softwares (Atlas.ti, Sphinx e R) tem possibilitado, para essa pesquisa, a criação de modelos conceituais e redes de informações sobre a trajetória do MST nos jornais analisados desde 1984 até 2014, evidenciando a conexão entre os dados, e oferecendo recursos para a geração de relatórios e visualização dos resultados. Através da análise de conteúdo, de discurso e da definição de categorias que são apresentadas na narrativa da Mídia impressa, e, inclusive, com a utilização das técnicas de produção deste discurso nos deparamos com vários veiculadores sobre o objeto estudado: o sujeito formal (sem autoria), o identificado (jornalistas e articulistas), os atores sociais integrantes do Movimento, os interlocutores preferenciais da Mídia (aqueles atores que são recorrentes nas matérias), como também, é possível identificar quais os eventos noticiados e quais

¹⁷ As nuvens de palavras ou *word cloud* serão fundamentais para a análise através de recursos visuais.

aqueles que são sub-analisados (como, por exemplo, marchas, ocupações, assentamentos, congressos).

Como todo discurso, a Mídia Impressa possui um viés hermenêutico e valorativo. Por detrás da linguagem aparente utilizada e que se autopropaga como neutra e objetiva, existe um ou vários grupos, classes e frações de classe que se posicionam e defendem seus pontos de vista e interesses sociais, econômicos e culturais. Na linguagem simbólica e polissêmica da Mídia (e de todas as formas de linguagem), esconde-se um sentido que convém desvendar. Assim, o método da análise de conteúdo possibilita uma leitura profunda¹⁸ das comunicações, indo além da leitura aparente¹⁹.

Com relação à análise quantitativa, o que serve de referencial é a frequência com que surgem certas características do conteúdo. Através do questionário digital feito no Sphinx, é possível tabular questões multifatoriais (políticas, sociais, econômicas, históricas, entre outras) sobre o MST, quantificando informações desde a sua gênese até o ano de 2014. É justamente a presença ou ausência de uma informação ou característica de conteúdo ou de um conjunto de aspectos num determinado fragmento de mensagem que está sendo levado em consideração na nossa pesquisa.

ANÁLISE CRÍTICA E TEÓRICA

A utilização de tecnologia vem acompanhando o processo histórico de desenvolvimento de pesquisa em ciências sociais. Alguns desses instrumentos impactaram de forma profunda as metodologias e formas de fazer pesquisa, como máquinas de escrever, os gravadores, máquinas fotográficas, aparelhos celulares e, mais recentemente, os computadores e a internet. Uma grande parcela dessas ferramentas proporcionou mudanças operacionais nas pesquisas em sociologia, antropologia e ciências políticas, já outras foram incorporadas aos processos metodológicos e fizeram emergir novos campos de pesquisa como no caso das etnografias visuais e audiovisuais (Lage, 2011).

¹⁸ Os programas de computadores permitiram uma análise mais detalhada dos textos, no que diz respeito à sua mensuração, ou seja, à frequência de uma determinada unidade de análise; técnicas estatísticas avançadas puderam ser aplicadas. (BARDIN, 2009).

¹⁹ Ver Bardin, 2009 .

Uma importante metodologia de pesquisa que impactou os estudos sociais e se consolidou como um excelente instrumento de coleta de dados foram os questionários ou surveys. O uso de pesquisa survey em ciências sociais remonta a autores clássicos como Marx, Weber e Durkheim (BABBIE, 1999) que lançaram mão de dados qualitativos e quantitativos estabelecendo o campo de pesquisa das ciências sociais. A utilização desta técnica passou por percursos diversos, mas se consolidou como um instrumento importante de pesquisa nas várias correntes de estudos sociais norte-americanos, como por exemplo, na Escola Sociológica de Chicago do início do século XX. Os estudos sobre a intenção de voto, a pesquisa de opinião e as correntes migratórias foram as linhas de pesquisas que inicialmente construíram os maiores vetores de desenvolvimento dos Surveys (BABBIE, 1999; SIMÕES, 2007).

A tabulação dos dados criados por meio dos surveys podem ser organizadas de variadas formas. Até os anos 1990 utilizava-se cartões de papel onde eram feitas marcações de controle para posterior quantificação. Recentemente, alguns softwares considerados simples e de uso comum estão sendo utilizados com essa finalidade, como o excel. Outro software utilizado é o SPSS, que demanda alguns conhecimentos específicos, mas opera por uma lógica próxima a do excel. Uma variedade de outros softwares, mais específicos para pesquisa com questionários, tem alcançado visibilidade nos últimos 20 anos, um deles é o Sphinx. Esse programa foi o escolhido no desenho dessa investigação e permite ao pesquisador um entrosamento dos vários processos que envolvem a pesquisa *survey*, desde a construção do questionário, a coleta de respostas, passando pela tabulação e englobando alguns testes estatísticos (BARDIN, 2009). No entanto, a utilização do Sphinx contempla uma parte restrita da pesquisa centrado-se especificamente aos dados quantitativos. E como nossa pesquisa possui o duplo caráter quali/quantitativo (como já expusemos acima), utilizamos, em associação a esse software, uma outra ferramenta de auxílio à pesquisa, voltada para a análise qualitativa, qual seja, o Atlas.TI.

O software Atlas.TI é considerado um instrumental computacional por ser uma ferramenta que apoia análise de dados qualitativos em pesquisas acadêmicas. A utilização desse tipo de instrumento, de auxílio à análise em investigações científicas, vem crescendo desde os anos 1980. São variadas suas funcionalidades, mas em comum a maioria deles opera pela organização, armazenamento e indexação de dados, o que

pode facilitar a análise (LAGE & GODOY, 2008), principalmente em casos de pesquisas Big Data.²⁰

O uso de Caqdas nas práticas científicas das ciências sociais é recente no Brasil, mas tem alcançado importante projeção. Alguns programas de pós-graduação brasileiros, como o Centro de Pesquisas Quantitativas em Ciências Sociais da UFMG²¹ oferecem treinamento para utilização de metodologias quantitativas com auxílio de software. No curso de ciências sociais da UFBA, onde o nosso núcleo de pesquisa está alocado²², a utilização desses softwares tem uma emergência muito atual, sendo o aprendizado ainda restrito a alguns grupos de pesquisa de pós-graduação e de poucas disciplinas na graduação.

O Atlas.TI é uma das ferramentas mais completas de auxílio à pesquisa qualitativa. Com seu uso podemos armazenar grandes volumes de dados, codificar uma variedade de arquivos de textos, imagens e organizar essas codificações de forma a entender relações e contextos - essas são algumas das ferramentas que mais utilizamos nesta pesquisa. Dentre as ferramentas computacionais de uso nas ciências sociais muitas não possuem sistemas de desenvolvimento aberto ou compartilhado²³, sendo a restrição financeira para aquisição de licença, treinamento e desconhecimentos dos parâmetros de desenvolvimento do programa, pontos negativos na escolha dessas ferramentas. Um software que se destaca pela abertura do código-fonte para desenvolvimento e que tem sido utilizado em pesquisas na área de ciências humanas é o R. Esse programa que opera através de lógica de programação permite o download em massa de diferentes arquivos da internet, assim como trabalho com metadados e análises de rede (NASCIMENTO, 2016)

²⁰ Big Data é o termo utilizado para designar grande dados oriundos de fontes informacionais, como rastros digitais, meta dados ou dados de armazenamento automático. Para mais informações sobre big data e seus impactos nas ciências sociais podem ser obtidas nos textos de Nascimento, 2015 e Lupton, 2015.

²¹ Maiores informações em: www.fafich.ufmg.br/cpeqs e www.fafich.ufmg.br/mq.

²² Núcleo de Estudos Ambientais e Rurais (Nuclear), certificado pelo CNPq desde 1999.

²³ "Software livre, segundo a definição criada pela Free Software Foundation é qualquer programa de computador que pode ser usado, copiado, estudado, modificado e redistribuído com algumas restrições. A liberdade de tais diretrizes é central ao conceito, o qual se opõe ao conceito de software proprietário, mas não ao software que é vendido almejando lucro (software comercial). A maneira usual de distribuição de software livre é anexar a este uma licença de software livre, e tornar o código fonte do programa disponível." Informação disponível em <<http://www.softwarelivre.gov.br/>>.

O uso de ferramentas computacionais para auxílio em pesquisa qualitativa é de grande valia para determinados desenhos de investigação científica, como no caso da análise com Mídia digital que envolve grande volume de dados. No entanto, a utilização de software dá apoio a algumas metodologias mais do que a outras, como por exemplo, ferramentas como o Atlas.TI que permitem uma imersão inicial nos dados brutos e pode ser utilizado como importante suporte à pesquisas embasadas na teoria fundamentada em dados empíricos²⁴ (STRAUSS & CORBIN, 2008). Dessa forma, o software proporciona uma imersão no conjunto de dados que ajuda na criação das categorias e, posteriormente, juntamente com o arcabouço teórico específico da temática a ser estudada, colabora na estruturação das relações entre elas.

O uso de programas computacionais de auxílio à pesquisa é importante na medida em que com ele o cientista social amplia sua capacidade de organização de grandes volumes de dados em menor tempo, de forma sistemática, dentro da estrutura do programa, podendo assim auxiliar sua análise dos dados. Além do suporte no processo de análise, o software auxilia na coleta, armazenamento e visualização de resultados da pesquisa. Na coleta através de ferramentas de busca automatizada e softwares como o R, com o uso de programação, podemos fazer download em massa de diversos conteúdos presentes na internet, como postagens em rede sociais, notícias de jornal e banco de dados de artigos científicos, entre outros dados pertinentes. No caso desta pesquisa, o uso dos softwares supracitados tem auxiliado todas essas etapas, em particular a etapa da coleta, que dentro desse desenho de pesquisa, possibilitou-nos a ousar na construção do objetivo central, dando-nos a chance de optar por uma construção de uma ampla base de dados empíricos construída por três principais Mídias impressas brasileiras sobre o MST. Sem estas ferramentas a única forma de coleta de dados seria a manual, o que demandaria um longo tempo empreendido nessa fase. Nesta pesquisa realizamos a coleta de maneira automatizada o que permitiu trabalhar com um volume maior de dados, tendo maior confiabilidade na coleta e menor tempo dedicado à esta etapa.

²⁴ “O que Strauss e Corbin querem dizer quando usam o termo "teoria fundamentada"? Eles querem dizer teoria que foi derivada de dados, sistematicamente reunidos e analisados por meio de processo de pesquisa. Neste método, coleta de dados, análise e eventual teoria mantêm uma relação próxima entre si.” (CORBIN E STRAUS, 2008, P. 14)

Mediante ao exposto, constatamos que a pesquisa sobre os 30 anos do MST na Mídia impressa em O Estado de São Paulo desfrutou de uma trajetória própria, intimamente guiada pela própria complexidade do objeto em questão. A utilização de ferramentas Caqdas, com isso, representou não somente um auxílio relevante no que diz respeito à análise qualitativa de uma base de dados referentes a 30 anos do MST, como também, foi introduzida por nós de forma a adaptar-se às demandas do objeto em si. Assim, nossa imaginação sociológica permitiu encontrar a metodologia adequada para a apreensão de nossa questão prévia sobre o processo de burocratização do MST.

No primeiro momento, interpretamos o MST em sua totalidade e assim organizamos nossa extensa base de dados em 30 anos de história, a partir de sua representação no jornal O Estado de São Paulo por motivos já expostos aqui. Dessa forma, as amostragens construídas inseriram o MST em um contexto analítico amplo e toda a sua complexidade compôs, a princípio, uma só ferramenta de investigação. A cada momento inicial do trabalho, tínhamos a impressão de que o MST não encaixava em nossa metodologia, o que nos levou a refletir constantemente sobre o processo de construção da metodologia da pesquisa.

Após intensos debates e revisão bibliográfica, a estrutura do estudo inicial da base de dados foi pautada por elementos descritivos e analíticos. Os dados referentes à data, formato, título, e do caderno das matérias foram os primeiros a serem identificados. Logo em seguida, a parte analítica deu o tom da pesquisa. O mapeamento dos conflitos, a identificação do sujeito em questão, à dinâmica de interação do Estado com o movimento social, as narrativas, os vínculos identitários, os compromissos coletivos do MST e a tipologia de protesto/ repertórios de interação apresentaram-nos a necessidade de pesquisar o MST de acordo com o desenrolar de sua história, repleta de eventos com afinidades eletivas próprias.

A vigilância epistemológica permitiu que inseríssemos esta pesquisa em outra dinâmica de funcionamento. Ao analisarmos as matérias das amostragens do MST de 1984 a 2014, percebemos que a própria natureza do objeto nos exigia outra metodologia. O MST apresentava-se de forma distinta a cada decênio do seu surgimento o que nos levou a criar a organização da base de dados seguindo os eventos históricos da realidade brasileira como citado acima. Assim, num momento estafante da pesquisa, quando observamos que andávamos em círculos, tivemos o *insigth* de separar a base de

dados em três decênios (1984-1994, 1995-2004 e 2005- 2014), cada uma com 10 amostragens compostas por 15 reportagens, que foram lidas, categorizadas e serviram de partida para a construção das categorias analíticas que subsidiaram a elaboração das nuvens de conteúdo (*tag clouds*) apresentadas abaixo²⁵.

A divisão em três decênios efetivou-se a partir das informações exaradas em cada período: a abertura política brasileira no primeiro decênio, o início da redemocratização brasileira por meio de eleições diretas no segundo decênio, e a ascensão dos Partidos dos Trabalhadores no terceiro decênio. Observamos que a narrativa sobre o MST se transfigurava e o apresentava de forma distinta a cada decênio, e isso abriu um amplo leque de possibilidades de análise.

A primeira década revelou, antes de mais nada, o sujeito sem terra como elemento central da dinâmica da política fundiária brasileira. Categorias como “sem-terra”, “colonos”, “boia fria” e “trabalhador rural” surgiam nas matérias de 1984 a 1994, delineando o sujeito a ser trabalhado. Nesse decênio, a relação do MST com o Estado é dada quase exclusivamente pela via do conflito: prevalecem categorias como “invasão”, “violência jurídica”, “conflito”, assim como a enorme quantidade de confrontos entre ambos.

No primeiro decênio, o repertório de interação e as tipologias de protesto tornam-se evidentes, sobretudo pela via da ação direta: as categorias empíricas “tática de guerrilha”, “tática de ocupações”, “conflito armado” apareceram de forma recorrente no discurso da Mídia impressa para descrever as ações dos movimentos sociais. E nesse sentido, a narrativa do MST, seus vínculos de identidade e compromissos coletivos emergiam a partir da compreensão da função social da terra, princípio ideológico fundador do movimento, que tem como parâmetro as oposições “propriedade x posse”, “terras improdutivas” x “desapropriação”, categorias êmicas que deram norte aos sujeitos do MST na década de 1980.

²⁵ Foram utilizadas 14 matérias das amostragens de cada decênio, essas matérias foram rodadas em um programa de reconhecimento de caracteres onde somente as partes referentes a pesquisa foram selecionadas. Depois as matérias foram rodadas no Atlas.ti para que fosse gerada a frequência de palavras, dando subsídio as nuvens geradas no programa R.



Figura 1: Nuvem de palavras referente ao primeiro decênio (1984-1994).

O segundo decênio, por sua vez, apresentou-se a partir de algumas ambiguidades. Se por um lado, a política do conflito porta-se como regente da narrativa e ação do MST, por outro, as primeiras experiências de interação e integração com a arena estatal trariam a tona novos elementos de análise. Esse é o momento histórico em que o jornal O Estado de São Paulo apresenta o sujeito sem terra através de seu agente de mediação. O sujeito na luta pela terra não era mais o camponês, mas o movimento social.



Figura 2: Nuvem de palavras referente ao segundo decênio (1995-2005).

No terceiro decênio, as primeiras experiências de administrações petistas inseriram novas categorias de análise à pesquisa, e a década de 1990 assistiu à coexistência de termos como “aliança” e “conflito”. Ainda que em Estados como o Rio Grande do Sul, através da eleição do governador Olívio Dutra em 1998, tenha havido maior aproximação entre o MST e a arena do Estado, a década de 90 foi marcada por categorias como “radicalização” e “ausência de interlocução”, no que diz respeito à esfera federal.



Figura 3: Nuvem de palavras referente ao terceiro decênio (2005-2014).

Chamaram a nossa atenção as categorias relacionadas ao sujeito em questão. De 1995 a 2004, o MST apresentou-se como o mediador central das lutas camponesas. As matérias coletadas e analisadas, em sua maioria, se referiam ao sujeito “sem terra” como um coadjuvante do Movimento dos Trabalhadores Rurais Sem Terra. Confundia-se assim, no discurso midiático e no discurso dos mediadores, o sujeito e o movimento.

O repertório de interação e a tipologia de protesto são marcados pelo processo de burocratização do movimento por um lado, e por outro pelas rupturas internas e expurgos de lideranças que insistiram no processo de “radicalização”. Contraditoriamente, as estratégias da “invasão como principal forma de luta”, assim como os discursos sobre “negociação”, “apoio” e “participação no governo” fizeram parte tanto do discurso da Mídia impressa como das lideranças do MST. Nesse sentido,

os compromissos coletivos do movimento também comportaram uma conjunção de elementos que se tensionavam entre si: estiveram presentes a luta pela “função social da terra”, os discursos do movimento sobre o “balanço da reforma agrária”, como também, a participação política no interior do governo para auxiliar a “construção de políticas públicas”.

Destarte, o decênio de 2005 a 2014 apresentou-se sob a moldura da bifocalidade. A ascensão do Partido dos Trabalhadores ao governo federal estabeleceu um novo repertório de interação com o MST, trazendo para a cena um novo padrão de negociação, ainda que a rotinização dos protestos não necessariamente tenha conseguido reduzir os conflitos fundiários no cenário nacional. Com relação à nossa pesquisa, o estudo deste período impulsionou a reflexão sobre até que ponto a institucionalização do MST significou a contenção ou perda da sua autonomia.

Como um paradigma para a análise da terceira década, fomos guiados pela existência de uma matéria que denominamos de “matéria-tronco”, publicada em 24/05/09. Nela o sociólogo José de Souza Martins, especialista em estudos rurais, aponta que à revelia da presença de um desencontro profundo entre as motivações imediatas do sujeito sem terra e a ação dos movimentos sociais, a luta agrária tem resistido à ordem latifundista. Além de problematizar as multiplicidades de organizações voltadas para a luta pela terra, o referido sociólogo questiona a ausência no discurso dos mediadores interpretativos, da crítica à questão da renda fundiária e da reforma agrária de mercado.

Nesse sentido, a base de dados compilada pelos softwares fez-se eloquente. O repertório de interação entre a arena estatal e o MST tornou-se aparente por meio de categorias discursivas como “aproximação”, “negociação” e “aliança”. Termos como “gestão” e “políticas públicas” tornaram-se cada vez mais presentes no discurso do e sobre o MST. Com isso, a tipologia de protesto estabeleceu-se de outra forma. Ainda que as ocupações estivessem presentes, houve outra dinâmica emergindo, a exemplo de “audiências públicas” e “barulhaços”.

CONCLUSÃO

Por meio da codificação das matérias utilizando o Sphinx e o Atlas.ti foi possível cotejar as mudanças discursivas que permearam as três décadas de existência do MST. Os dados da fase exploratória apontam que houveram mudanças profundas nas estratégias e nos discursos sobre as práticas organizacionais do movimento. As ocupações voltaram-se contra o agronegócio e não mais para o latifúndio improdutivo, sendo que as categorias “Multinacionais”, “conglomerados industriais” “invasão de laboratórios” aparecem como ações mais frequentes. A luta contra a propriedade privada da terra, que a princípio, era a base de luta desse movimento, se transforma pela luta de uma reforma agrária “possível” e negociada no interior do Estado, ocupando cargos no executivo e no legislativo. Observamos, inclusive, que o discurso se ampliou para outras temáticas, como o discurso ambientalista e o produtivista, no sentido de inserir o camponês na agricultura de mercado, e não apenas defendendo as bases camponesas da economia moral do grupo doméstico familístico, voltado para a ética camponesa da economia solidária. (CHAYANOV, 1974)

Em função dessa inflexão observada nos trinta anos de ação do MST, relatados pela mídia nacional, tornou-se fundamental para o nosso trabalho analítico pensar quais os conceitos se tornaram necessários para operacionalizar a análise à luz da hipótese inicial, qual seja o processo de burocratização e transformação interna das estruturas organizacionais do Movimento.

Para além do uso dos softwares foi necessário salientar que o material coletado a partir da codificação deveria se relacionar com as hipóteses teóricas inicialmente escolhidas e posteriormente refinadas. Os dados, por serem extensos, demandaram o uso de tecnologias sofisticadas para serem tratados e nos ajudaram a traçar novos percursos analíticos trazendo a tona diferentes perspectivas da realidade social analisada, nos levando a concordar ou discordar dos objetivos originalmente lançados pela questão investigativa preliminar da pesquisa. O ponto positivo da utilização dos programas é a possibilidade de perceber variáveis que se associam aos conceitos teóricos da pesquisa, justamente por nos ajudar a visualizar de forma detalhada um lapso temporal sociologicamente dilatado. No caso dessa pesquisa, os trinta anos da história do MST, se mostraram heterogêneos e hermeneuticamente ricos na medida em que ao serem cotejados de forma mais refinada e apurada por meio da utilização dos

softwares já mencionados, alcançamos uma profundidade qualitativa e quantitativa, uma vez que a metodologia utilizada nos possibilitou analisar um pouco mais de 50% de um volume de 8264 arquivos com matérias jornalísticas, gráficos e imagens.

Em resumo, na nossa pesquisa sobre a mídia impressa e o MST, entre os anos de 1984 a 2014, foi possível constatar a inegável contribuição dos Caqdas para o ordenamento, organização, refinamento dos dados, na geração de relatórios e novas abordagens analíticas a respeito de um tema bastante explorado em pesquisa sociológica sobre movimentos sociais, mídias e questões rurais. Verificamos assim, que os Caqdas aumentaram a possibilidade de pesquisas com imensos volumes de dados estruturados e não estruturados, além de nos dar a possibilidade de análise de um número maior de variáveis para a investigação de uma determinada realidade social, que no caso desse trabalho, envolve compreender as mudanças e as afinidades eletivas da história recente do Movimento dos Trabalhadores Sem Terra inserida no complexo cenário da estrutura fundiária brasileira.

Referências bibliográficas

- BABBIE, E. (1999). *Métodos de pesquisa de survey*. . Belo Horizonte: UFMG.
- BARDIN, L. (2009). *Análise de Conteúdo* (70 ed.). Lisboa: LDA.
- BOURDIEU. (1965). *Um art moyen - essai sur les usages sociaux de la photographie*. Paris.
- BOURDIEU, P. (2002). Introdução a uma Sociologia Reflexiva. Em *O poder Simbólico* (5 ed., pp. 17-58). Rio de Janeiro: Bertrand Brasil.
- Caiuby Novaes, S., Cornélia, E., & M. J. (2005). *O imaginário e o poético nas Ciências Sociais*. Bauru: EDUSC.
- CHAYANOV, A. (1974). *La organizacion de la unidad economica campesina*. Buenos Aires: Nova Vision.

- FRASER, N., & HONNETH, A. (2006). *¿Redistribución o reconocimiento?: un debate político-filosófico*. Madrid.
- GEMMA, Penn. (2002). Análise Semiótica de Imagens Paradas . In: Pesquisa Qualitativa com texto imagem e som, org. Bauer M. & Gaskell G. Petrópolis, RJ: Vozes.
- IVO, A. B. (2005). La destitución de lo social: Estado, gobierno y políticas sociales. *Estudios Sociológicos, XXIII*, p. 347-374.
- Lage. (2011). Os softwares tipo CAQDAS e a sua contribuição para a pesquisa qualitativa em educação. *ETD – Educação Temática Digital*.
- LAGE, M., & GODOY, A. (2008). O uso do computador na análise de dados qualitativos: questões emergentes. *RAM- Revista de administração Mackenzie, 9(4)*, 75-98.
- LEWINS, A., & SILVER, C. (2009). *Choosing a CAQDAS package*.
- LUPTON, D. (2015). *Digital Sociology*. Routledge.
- NASCIMENTO, L. F. (2015). DIGITAL SOCIOLOGY de Deborah Lupton. *Cad. CRH, 28*, 671-673.
- NASCIMENTO, L. F. (2016). A Sociologia Digital: um desafio para o século XXI. *Sociologias, 216-241*.
- STRAUSS, A., & CORBIN, J. (2008). *Pesquisa qualitativa: técnicas e procedimentos para o desenvolvimento de teoria fundamentada*. Porto Alegre: Artmed.