

Ciências Sociais Computacionais: Origens, Métodos e Desafios

Introdução

As Ciências Sociais Computacionais (CSC) emergem como uma área interdisciplinar que combina métodos computacionais e grandes bases de dados para investigar fenômenos sociais em escala inédita. A transição da era analógica para a digital nas últimas décadas criou novas possibilidades de coleta e análise de dados sociais (Salganik, 2019). Cada vez mais aspectos da interação humana estão registrados em formato digital (por exemplo, textos, comportamentos em redes sociais, registros administrativos), o que tem transformado a pesquisa social ao permitir inovações metodológicas e análises em larga escala (Salganik, 2019). Em contrapartida, esse avanço também levanta questões quanto à confiabilidade, relevância e ética dessas novas abordagens (Salganik, 2019).

Nesse contexto, a presente pesquisa busca delinear o campo das Ciências Sociais Computacionais, respondendo à seguinte pergunta de pesquisa: como os métodos computacionais e dados massivos têm alterado a prática das ciências sociais e quais oportunidades e desafios decorrem dessa transformação? Os objetivos incluem traçar as origens e definições da área, apresentar novas fontes de dados (“big data”) e metodologias (como a modelagem preditiva), discutir aplicações exemplares (análise de texto, experimentos digitais e estudos de caso recentes) e por fim examinar criticamente os limites e implicações éticas envolvidos. A justificativa para tal mapeamento reside na crescente influência dos dados digitais na pesquisa social contemporânea – compreender os princípios, potencialidades e dilemas das CSC é fundamental para pesquisadores em métodos quantitativos nas ciências sociais.

Nos últimos anos, a área ganhou proeminência rapidamente. Lazer et al. observam que a produção em CSC explodiu na última década, com milhares de estudos baseados em dados observacionais em larga escala, experimentos online e simulações computacionais antes impraticáveis (Lazer et al., 2020). Esses trabalhos têm avançado o entendimento sobre temas importantes – da

desigualdade social à difusão de doenças – evidenciando o potencial científico do novo paradigma (Lazer et al., 2020). Por outro lado, estruturas institucionais e éticas não acompanharam totalmente esse crescimento, havendo lacunas em termos de treinamentos integrados e diretrizes para uso de grandes dados (Lazer et al., 2020). Assim, este relatório também enfatiza os dilemas metodológicos (por exemplo, explicação vs. predição) e os desafios éticos (como viés algorítmico e privacidade) que emergem nesse campo. A estrutura do texto segue uma progressão lógica: primeiro são apresentadas as origens conceituais das CSC e sua definição; em seguida exploram-se as novas fontes de dados e métodos computacionais; posteriormente discutem-se estudos de caso exemplares; e finalmente refletem-se os limites e perspectivas futuras. Espera-se que a síntese fornecida sirva de base teórica para estudantes de pós-graduação compreenderem o escopo e as responsabilidades inerentes às Ciências Sociais Computacionais.

Origens e Definições das Ciências Sociais Computacionais

As Ciências Sociais Computacionais têm raízes na convergência de disciplinas e na evolução tecnológica. O termo ganhou destaque a partir de fins dos anos 2000, notadamente com a publicação de Lazer et al. (2009) na revista *Science*, que delineou uma nova agenda de pesquisa baseada em rastros digitais e colaboração entre cientistas sociais, computacionais e físicos. Conceitualmente, pode-se definir CSC como “o desenvolvimento e aplicação de métodos computacionais ao estudo de dados comportamentais humanos em grande escala” (Lazer et al., 2020). Em outras palavras, trata-se de aplicar técnicas de ciência de dados, estatística computacional e simulação a conjuntos de dados massivos sobre fenômenos sociais (muitas vezes capturados automaticamente por sistemas digitais) para compreender padrões coletivos complexos.

Essa abordagem difere da análise social quantitativa tradicional em vários aspectos-chave. Primeiro, a ênfase recai sobre dados granulares e não-estruturados – como texto, imagens, geolocalização e redes de relacionamento – enquanto a pesquisa quantitativa clássica privilegiava dados agregados ou provenientes de surveys estruturados. Como destacam Lazer et al., a CSC lida com linguagem, localização, redes e até vídeos, aplicando modelos estatísticos capazes de capturar dependências complexas nos dados (Lazer et al., 2020). Já as ciências sociais tradicionais tendiam a considerar observações independentes (caso a caso) e variáveis pré-definidas em matriz retangular. Segundo, as CSC envolvem frequentemente equipes interdisciplinares: um “comunidade” de sociólogos, cientistas da computação, estatísticos, físicos, entre outros, coalesceu sob o guarda-chuva das CSC (Lazer et al., 2020). Essa colaboração ampliou o repertório metodológico, incorporando *machine learning*, análise de redes, processamento de linguagem natural, simulação baseada em agentes, etc., às perguntas clássicas das ciências sociais.

Historicamente, pode-se apontar também origens anteriores: por exemplo, a

sociologia computacional dos anos 1990 já explorava simulações (como o modelo de segregação de Schelling) e a análise de redes sociais se beneficiou de algoritmos computacionais desde cedo. Contudo, a disponibilidade de “big data” sociais a partir dos anos 2000 (revolução da Internet, mídias sociais, sensores digitais) criou um ponto de virada. Matthew Salganik (2019) ilustra essa mudança com uma experiência pessoal: em 2005, ele conduziu um experimento online que coletou centenas de observações enquanto ele dormia, algo impraticável em experimentos de laboratório tradicionais (Salganik, 2019). Essa anedota simboliza a ruptura proporcionada pela era digital – volume e velocidade de dados incomparáveis ao passado – permitindo novas perguntas e métodos em pesquisa social.

O crescimento institucional da área confirma sua consolidação. Houve proliferação de eventos científicos dedicados, como conferências, workshops e *summer schools* ao redor do globo (Lazer et al., 2020). Por exemplo, o Summer Institute in Computational Social Science (SICSS), iniciado em 2017 por Matthew Salganik (Princeton) e Chris Bail (Duke), treina anualmente estudantes e jovens pesquisadores do mundo inteiro em técnicas de CSC, disponibilizando abertamente materiais didáticos online. Também surgiram periódicos especializados e linhas de financiamento voltadas a projetos interdisciplinares que utilizem dados massivos para questões sociais. Em síntese, as Ciências Sociais Computacionais se definem pela abordagem metodológica inovadora – alavancando poder computacional e grandes bases de dados – aplicada ao estudo rigoroso de problemas sociais clássicos (e novos). Essa integração entre as “ciências dos dados” e as “ciências sociais” representa não apenas uma extensão das ferramentas quantitativas, mas uma transformação epistemológica na forma de fazer ciência social, cujos contornos e implicações começam a se delinear e serão discutidos nas seções seguintes.

Big Data e Novas Fontes de Dados Sociais

Uma característica central das CSC é o uso de *Big Data* – conjuntos de dados extremamente volumosos, variados e gerados de forma contínua – como matéria-prima para inferências sociais. Diferentes das bases de dados tradicionais obtidas por pesquisas planejadas (surveys, censos, etc.), essas novas fontes de dados são muitas vezes subprodutos de atividades digitais cotidianas: registros de telefonia móvel, interações em redes sociais, históricos de buscas na internet, transações financeiras eletrônicas, dados de sensores GPS, entre outros. Tais fontes fornecem observações comportamentais em escala populacional e de alta resolução temporal e espacial, permitindo análises antes inviáveis. Entretanto, por não terem sido coletados com finalidade de pesquisa, esses dados “orgânicos” apresentam desafios de tratamento, seleção e interpretação.

Mapa preditivo de riqueza em Ruanda construído a partir de dados de telefones celulares (Blumenstock et al., 2015). Cada região é colorida de acordo com o nível estimado de riqueza, inferido por um modelo de aprendizado de máquina aplicado aos padrões de uso de celular da população.

Um exemplo emblemático do potencial do *big data* nas ciências sociais é

o estudo de Blumenstock, Cadamuro e On (2015) sobre pobreza em Ruanda. Esses autores exploraram dados anonimizados de metadados de telefonia móvel – registros de chamadas e recargas de crédito de milhões de usuários – para inferir o nível socioeconômico dos indivíduos e mapear a distribuição geográfica da pobreza. O método consistiu em primeiro realizar um *survey* tradicional com 1.000 usuários de celular (coletando informações demográficas e de riqueza auto-reportada) e então treinar algoritmos de *machine learning* para encontrar padrões nos dados de uso telefônico que estivessem correlacionados à riqueza (Blumenstock; Cadamuro; On, 2015; Blumenstock; Cadamuro; On, 2015). Entre os padrões identificados estavam, por exemplo: pessoas mais ricas faziam um número maior de chamadas, recarregavam valores maiores de crédito e tinham horários de chamada mais regulares (horário comercial) em comparação aos mais pobres (Blumenstock; Cadamuro; On, 2015). Utilizando milhares de variáveis extraídas dos metadados e técnicas de aprendizado supervisionado, o modelo conseguiu predizer com boa acurácia a riqueza individual a partir do comportamento telefônico (Blumenstock; Cadamuro; On, 2015).

De posse desse modelo preditivo, Blumenstock et al. aplicaram-no aos 1,5 milhão de usuários da operadora, extrapolando as previsões para todo o país. Isso permitiu gerar mapas de pobreza de alta resolução, identificando diferenças socioeconômicas em nível local (setores, vilarejos) (Blumenstock; Cadamuro; On, 2015) – algo impraticável apenas com as pesquisas domiciliares convencionais, que em Ruanda ocorriam a cada 3–5 anos e a custo elevado. Os resultados mostraram, de forma inovadora, que os dados de celular de 2009 previram melhor a distribuição de riqueza em 2010 do que a pesquisa domiciliar anterior de 2007 (Blumenstock; Cadamuro; On, 2015). Além disso, o custo estimado para produzir essas estimativas foi em torno de US\$ 10 mil, significativamente menor que o de um censo nacional completo (Blumenstock; Cadamuro; On, 2015). Em suma, o estudo demonstrou que *big data* privados podem complementar – ou em parte substituir – medidas oficiais, fornecendo indicadores socioeconômicos atualizados e granularizados a baixo custo. Seus autores vislumbram aplicações similares usando dados de redes sociais, imagens de satélite e outras fontes digitais para medir vulnerabilidade populacional e orientar políticas públicas (Blumenstock; Cadamuro; On, 2015).

Os novos tipos de dados sociais disponíveis vão muito além da telefonia móvel. Plataformas online como Facebook, Twitter, Instagram e YouTube geram quantidades massivas de dados comportamentais sobre interações sociais, opiniões, preferências culturais e formação de redes. Por exemplo, logs do Twitter possibilitaram rastrear a disseminação global de informações e hashtags em tempo real; dados do Facebook já foram usados para estimar traços de personalidade ou satisfação com políticas; históricos de buscas no Google anteciparam surtos de gripe (embora com ressalvas metodológicas). Registros administrativos digitalizados (escola, saúde, benefícios) também são um componente valioso do *big data* governamental, permitindo análises longitudinalmente extensas de trajetórias individuais. A riqueza dessas fontes, contudo, vem acompanhada de problemas de representatividade e viés: nem todos os grupos populacionais estão igualmente presentes no mundo digital (ex.: idosos sem acesso, comunidades

rurais desconectadas), e padrões de auto-seleção podem enviesar as inferências (ex.: usuários de rede social não representam a população geral).

Outra distinção fundamental das CSC é a possibilidade de combinar diferentes modalidades de dados para obter insights mais abrangentes. A integração de dados de satélite com dados de telecomunicação – como no exemplo de Blumens- tock – ou de dados textuais com dados de redes (sociométricos) permite análises multimodais. Estudos têm predito pobreza combinando imagens noturnas de satélite com registros de celular (Jean et al., 2016); outros uniram dados de mobilidade (GPS) com informação epidemiológica para modelar a propagação de doenças. Essas abordagens integrativas são parte da inovação metodológica das CSC, mas demandam alta capacidade computacional e cuidados adicionais na harmonização de fontes heterogêneas.

Em resumo, a era do *big data* trouxe novas fontes de evidência empírica às ciências sociais, revolucionando o tipo de pergunta que pode ser respondida. Podemos agora observar comportamentos sociais em escala massiva e em tempo quase real, algo inimaginável há poucas décadas. Como notam alguns autores, isso expande o foco tradicional de “linhas e colunas” (casos e variáveis numéricas independentes) para um universo de dados não estruturados e correlacionados (texto, redes, imagens) que exigem novos modelos estatísticos (Lazer et al., 2020). Nas seções seguintes, veremos como essa abundância de dados também promoveu novas metodologias de análise – em particular a ênfase na modelagem preditiva – e possibilitou experimentos e estudos antes impraticáveis. Entretanto, antecipa-se que mais dados não eliminam as dificuldades clássicas (de inferência causal, por exemplo) e trazem novos desafios que serão debatidos oportunamente.

Modelagem Preditiva vs. Explicativa

A disponibilidade de dados massivos e o uso de técnicas de *aprendizado de máquina* nas CSC ampliaram a ênfase em modelos preditivos, suscitando debates sobre suas relações e diferenças em relação aos modelos explicativos tradicionais nas ciências sociais. Historicamente, grande parte da modelagem estatística nas ciências sociais buscou explicar relações causais ou testar hipóteses teóricas – por exemplo, estimar o efeito de educação na renda controlando por outros fatores, ou verificar teorias de comportamento eleitoral via regressão. Nesse enfoque explicativo, importam a interpretação dos coeficientes, a significância estatística e o ajuste do modelo aos dados observados, pressupondo muitas vezes um modelo paramétrico “verdadeiro” subjacente aos dados.

Com as CSC, ganhou força um enfoque alternativo: o da modelagem preditiva, em que o objetivo principal é obter a melhor previsão possível de resultados futuros ou desconhecidos, frequentemente aplicando métodos computacionais (*machine learning*) que priorizam acurácia preditiva (por exemplo, árvores de decisão, florestas aleatórias, redes neurais) possivelmente à custa de interpretabilidade. Essa distinção foi formalizada por Galit Shmueli (2010) em seu influente artigo “To Explain or to Predict?”, que adverte para a confusão comum entre os dois objetivos. Em muitos campos, argumenta Shmueli, assume-se equivo-

cadamente que modelos com alto poder explicativo também terão alto poder preditivo, quando na realidade expiação causal e predição são tarefas distintas que requerem abordagens diferentes (Shmueli, 2010). A autora esclarece que enquanto a modelagem explicativa se preocupa em estimar corretamente relações e coeficientes (buscando evitar vieses para inferir causa e efeito), a modelagem preditiva foca em minimizar erros de previsão em novos dados, recorrendo a validação cruzada, conjuntos de teste e eventualmente aceitando modelos complexos ou de “caixa-preta” se eles predizem melhor (Shmueli, 2010). Shmueli salienta que conflitar os dois usos pode estagnar o progresso científico – é preciso decidir qual é o objetivo (explicar ou predizer) e então seguir práticas coerentes com cada finalidade (Shmueli, 2010).

Em alinhamento com essa ideia, o estatístico Leo Breiman, em seu célebre artigo “Statistical Modeling: The Two Cultures” (2001), criticou a ortodoxia estatística centrada exclusivamente em modelos de dados (ex.: suposições de regressão linear), propondo maior abertura a métodos orientados a algoritmos e previsão. Breiman identifica duas culturas no uso de modelos estatísticos: (1) a cultura do *data modeling*, que assume que os dados foram gerados por um modelo estocástico específico pré-definido (e.g., “ $y = f(X) + \varepsilon$ ” seguindo uma forma funcional conhecida) e então ajusta os parâmetros desse modelo aos dados; e (2) a cultura do *algorithmic modeling*, que trata o mecanismo gerador dos dados como desconhecido, aplicando algoritmos flexíveis para obter a melhor correspondência entre entradas e saídas sem se preocupar em explicitar uma forma paramétrica (Breiman, 2001; Breiman, 2001; Breiman, 2001; Breiman, 2001; Breiman, 2001; Breiman, 2001). Breiman observa que, no início dos anos 2000, 98% dos estatísticos alinhavam-se à primeira cultura, construindo modelos baseados em suposições (linearidade, normalidade etc.), validando-os por testes de aderência, enquanto apenas 2% (muitas vezes em outras áreas, como ciência da computação) seguiam a abordagem algorítmica com ênfase preditiva (Breiman, 2001). Ele argumenta que essa dependência excessiva de modelos de dados levou a “teorias irrelevantes e conclusões questionáveis” em muitos casos, e fez com que estatísticos ignorassem problemas empíricos interessantes que poderiam ser melhor abordados com métodos preditivos mais flexíveis (Breiman, 2001). Breiman advoga, portanto, por diversificar as ferramentas estatísticas, incorporando técnicas de *machine learning* não paramétricas que mostravam sucesso em previsão em campos como reconhecimento de padrões, ao invés de insistir apenas em modelos inferenciais tradicionais (Breiman, 2001).

Em termos práticos, a distinção preditivo vs. explicativo implica diferenças em todas as etapas da análise. Por exemplo, na seleção de variáveis, um pesquisador explicativo incluirá variáveis baseando-se em teorias e evitará colinearidade para interpretar efeitos, enquanto um preditivo selecionará automaticamente os preditores que melhor reduzem o erro, mesmo que sejam *proxies* sem significado claro. Na avaliação do modelo, o explicativo foca em R^2 , significância de coeficientes e testes de hipóteses; x Além disso, há uma diferença de atitude quanto à complexidade do modelo: explicações elegantes geralmente preferem parcimônia (modelo mais simples possível), ao passo que previsões eficazes muitas vezes se beneficiam de modelos complexos ou combinados (*ensembles*) se isso

reduzir erros preditivos. Shmueli (2010) e outros autores destacam que nenhum enfoque é “melhor” em absoluto – eles servem a propósitos distintos. O ideal, no campo das CSC, é reconhecer quando se quer gerar conhecimento causal e quando se quer simplesmente fazer previsões úteis, aplicando o ferramental adequado a cada caso.

Nas Ciências Sociais Computacionais, onde frequentemente lida-se com dados extensos e de alta dimensionalidade, houve uma proliferação de estudos de natureza preditiva: previsões de resultados eleitorais a partir de mídias sociais, detecção de crimes ou surtos de violência antes que ocorram, identificação de usuários propensos a abandonar um serviço, etc. Esses estudos trazem valor preditivo imediato (por exemplo, para orientar uma intervenção), mas às vezes são criticados por carecerem de interpretação sociológica ou teoria subjacente – o risco de “modelos-caixa-preta”. Por outro lado, há também uso de modelos preditivos como auxiliares para testar teorias, em abordagens conhecidas como modelagem híbrida (estimativa de parâmetros com técnica de ML seguida de interpretação causal, por exemplo).

O debate estimulou reflexões importantes: a complementaridade entre explicação e predição. Como notou Shmueli, ao clarificar as distinções, fica evidente que ambos os modos podem se informar mutuamente – modelos preditivos podem sugerir novas hipóteses explicativas, e teorias podem inspirar melhores preditores. Ainda assim, é crucial não confundir os critérios de sucesso: um modelo que prediz bem não está necessariamente explicando o por quê de um fenômeno, e um modelo com excelente ajuste causal interno pode falhar ao extrapolar para novos contextos. Em suma, as CSC trouxeram para o primeiro plano essa questão metodológica. Estudos emblemáticos como os de Breiman (2001) e Shmueli (2010) servem de norte para pesquisadores navegarem entre os objetivos de entender vs. prever nas suas análises – uma escolha que deve ser consciente.

Texto como Dados: processamento de linguagem natural

Grande parte dos dados gerados na sociedade digital contemporânea consiste em texto – posts em redes sociais, artigos de notícias online, transcrições de discursos, respostas abertas em enquetes, entre outros. Tradicionalmente, cientistas sociais qualitativos faziam análise de conteúdo manual desses textos (por exemplo, codificando temas em discursos políticos), enquanto quantitativos os tratavam apenas de forma limitada (contagem de palavras-chave). Nos últimos anos, porém, surgiu o paradigma de “texto como dados”, no qual técnicas computacionais de processamento de linguagem natural permitem extrair informações estruturadas de enormes coleções textuais para análise estatística. Esse avanço transformou a forma de conduzir pesquisas em ciência política, sociologia, comunicação e economia, ao viabilizar a análise sistemática de conteúdo em larga escala antes impossível ou excessivamente custosa (Grimmer; Stewart, 2013; Grimmer; Roberts; Stewart, 2022).

Conforme ressaltam Grimmer e Stewart (2013), “aqui reside a promessa da análise automatizada de textos: reduzir substancialmente os custos de analisar grandes coleções” (Grimmer; Stewart, 2013). Em seu artigo “Text as Data”, os autores revisam um amplo leque de novos métodos para tratar textos – de contagem de palavras a modelos de tópicos – mostrando que, em muitos casos, essas ferramentas já cumprem parte do potencial de tornar viável o que antes era inviável, como analisar todos os discursos no Congresso em um ano, ou milhares de mensagens em mídias sociais (Grimmer; Stewart, 2013). Contudo, eles enfatizam igualmente que há armadilhas: métodos automáticos não substituem a leitura cuidadosa humana, e requerem validações extensivas e específicas a cada problema (Grimmer; Stewart, 2013). Em outras palavras, embora possamos agora processar textos em escala massiva, o pesquisador não está liberado do trabalho de interpretação e verificação – ao contrário, deve aplicar pensamento substantivo para garantir que os algoritmos estejam realmente capturando o fenômeno de interesse e não se desviando por correlações espúrias ou ambiguidades de linguagem (Grimmer; Stewart, 2013).

A análise computacional de texto geralmente se divide em duas grandes tarefas: classificação e escala contínua (*scaling*) (Grimmer; Stewart, 2013; Grimmer; Roberts; Stewart, 2022). Na classificação, busca-se categorizar documentos em grupos predefinidos. Por exemplo, classificar tweets como “positivos”, “negativos” ou “neutros” em sentimento; identificar se artigos de jornal tratam de política doméstica ou internacional; ou atribuir posições ideológicas a manifestos partidários (esquerda/direita). Essa tarefa pode ser conduzida por métodos de aprendizado supervisionado, onde um subconjunto de textos é manualmente rotulado e usado para treinar um classificador (como regressão logística, SVM, redes neurais) que então rotula automaticamente novos textos. Também inclui métodos de dicionário, em que se define listas de palavras associadas a categorias e classifica-se o texto pelo número de ocorrências (por exemplo, medir “positividade” de um discurso contando palavras de um léxico positivo). Grimmer e Stewart alertam, entretanto, que aplicar dicionários fora do contexto para o qual foram desenvolvidos pode levar a erros graves (Grimmer; Stewart, 2013) – por exemplo, um dicionário de sentimento treinado em inglês cotidiano pode falhar em textos políticos formais ou em outras línguas. Técnicas mais sofisticadas para classificação incluem *deep learning* (redes neurais de linguagem) que aprendem representações semânticas complexas, hoje comuns na indústria (p.ex. BERT, GPT), embora sua interpretabilidade seja baixa.

Já a tarefa de escala visa posicionar textos ou autores em um espaço contínuo de dimensões latentes. Por exemplo, estimar a posição ideológica (esquerda–direita) de parlamentares a partir de seus discursos – aqui não há categorias fixas, mas um espectro latente a ser inferido. Métodos como *Wordscores* e *Wordfish* (amplamente usados em ciência política) quantificam o texto projetando-o num eixo ideológico baseado na frequência de palavras, sem necessidade de treinamento supervisionado. Modelos de tópicos latentes (como LDA, *Latent Dirichlet Allocation*) também se enquadram nessa ideia de *scaling*, no sentido de que revelam dimensões (tópicos) subjacentes a um corpus e atribuem a cada documento uma distribuição nesses tópicos – permitindo medir, por exem-

plo, sobre quais assuntos uma dada coleção de artigos trata e em que proporção. Esses métodos não supervisionados foram revolucionários ao descobrir estruturas ocultas em dados textuais, mas requerem interpretações cuidadosas dos tópicos extraídos e também validação (por exemplo, conferir se tópicos fazem sentido substantivo, ou correlacionar com metadados conhecidos). Conforme Gentzkow, Kelly e Taddy (2019) afirmam, a proliferação de textos digitais motivou o desenvolvimento de métodos estatísticos específicos para texto, e aplicações diversas têm florescido na economia e demais ciências sociais (Gentzkow; Kelly; Taddy, 2019). Eles destacam que texto tem características particulares – alta dimensionalidade, estrutura sequencial, presença de ambiguidades e ironias – que o diferenciam de outros dados, exigindo abordagens próprias (Gentzkow; Kelly; Taddy, 2019).

Uma aplicação ilustrativa de “texto como dado” foi o estudo de Gentzkow e Shapiro (2010) que mediou viés ideológico da mídia nos EUA comparando o vocabulário de jornais com o de congressistas de diferentes partidos, revelando quais jornais se alinhavam mais à retórica Democrata ou Republicana. Outro exemplo é o de Grimmer (2010), que usou modelagem de tópicos para descobrir os temas das correspondências parlamentares enviadas por senadores dos EUA aos seus eleitores, inferindo as estratégias de comunicação política. Também merece menção a análise de redes sociais online via texto: por exemplo, detectar comunidades ou polarização no Twitter a partir de hashtags e menções combinadas à análise de conteúdo das mensagens.

Apesar dos sucessos, as CSC reconhecem que a análise automatizada de texto tem limitações importantes. A linguagem humana é rica e contextual; sarcasmo, metáforas e variações culturais escapam facilmente dos algoritmos. Há sempre o risco de “viés do medidor”: isto é, o modelo de análise capturar algo diferente do pretendido. Por exemplo, um modelo de tópicos pode agrupar textos por autor ou estilo (vocabulário jurídico vs. coloquial) em vez de por assunto substantivo, se essas diferenças estilísticas predominarem nos dados. Por isso, inconsistentemente, recomenda-se a “validação extensiva e específica ao problema” (Grimmer; Stewart, 2013) – o pesquisador deve inspecionar manualmente amostras dos resultados, comparar com codificações humanas, e ajustar métodos conforme necessário. Grimmer e Stewart concluem que, para que métodos de texto automatizado se tornem ferramentas padrão confiáveis para cientistas sociais, novos métodos de validação e refinamento devem continuar a ser desenvolvidos, bem como uma integração pedagógica que permita aos pesquisadores não-computacionais utilizá-los com consciência de seus potenciais e limites (Grimmer; Stewart, 2013). Em suma, “texto como dado” abriu um horizonte promissor: democratizou o acesso ao conteúdo de milhões de documentos, antes restrito aos analistas qualitativos; mas trouxe consigo a responsabilidade de não abdicar do rigor teórico e da interpretação cuidadosa. Nas palavras de Grimmer e Stewart, os métodos automáticos devem ser vistos como formas de “ampliar e *augment*” a leitura humana, e não de a substituir (Grimmer; Stewart, 2013). Com essa filosofia, a análise computacional de conteúdo continuará avançando de mãos dadas com as ciências sociais tradicionais, enriquecendo-as sem, espera-se, descaracterizá-las.

Experimentos Sociais Digitais em Larga Escala

Outra inovação marcante possibilitada pelas CSC é a realização de experimentos sociais randomizados em larga escala através de plataformas digitais. Tradicionalmente, experimentos controlados na área social enfrentavam limitações logísticas (amostras pequenas, dificuldade de acompanhamento, alto custo). Contudo, com bilhões de usuários em plataformas online, tornou-se viável conduzir experimentos de campo com número de participantes na casa dos milhões, muitas vezes integrados às interações cotidianas das pessoas. Esses “mega-experimentos” fornecem um nível de poder estatístico sem precedentes para detectar efeitos sutis e estudar dinâmicas sociais complexas, embora também suscitem questões éticas e de interpretação.

Um exemplo pioneiro e emblemático é o estudo de Bond et al. (2012), conhecido como o “experimento com 61 milhões de pessoas” em parceria com o Facebook. Nesse experimento, realizado durante a eleição legislativa dos EUA de 2010, os autores testaram se mensagens de incentivo ao voto veiculadas em redes sociais poderiam aumentar efetivamente a participação eleitoral. A metodologia foi a seguinte: no dia da eleição, 60 milhões de usuários do Facebook foram aleatoriamente designados a receber no topo de seu feed uma mensagem social incentivando a votar – com um botão “Eu Votei”, um link para localizar a seção eleitoral e, crucialmente, mostrava ícones de alguns amigos do usuário que já haviam clicado “Eu Votei” (Bond et al., 2012). Paralelamente, 600 mil usuários receberam uma mensagem informativa quase idêntica, mas sem as informações sobre amigos (ou seja, sem o aspecto social), e outros ~600 mil formaram o grupo de controle, que não recebeu nenhuma mensagem relacionada ao voto (Bond et al., 2012). Assim, tratou-se de um experimento randomizado clássico, mas em escala massiva e inteiramente online.

Os resultados, cruzados com registros públicos de votação, demonstraram um efeito positivo embora modesto da intervenção. O grupo exposto à mensagem social teve uma taxa de comparecimento ligeiramente superior: os dados sugerem um aumento direto no comparecimento de cerca de 60 mil votantes devido à mensagem, além de um efeito indireto de contágio social envolvendo 280 mil votantes adicionais, totalizando aproximadamente 340 mil votos extras atribuídos à intervenção (Bond et al., 2012). Em termos relativos, esses 340 mil votos representaram cerca de 0,14% do eleitorado americano – um efeito pequeno em porcentagem, mas significativo em números absolutos e potencialmente relevante em disputas acirradas (Bond et al., 2012). Uma descoberta notável foi que a mensagem com componentes sociais (mostrar amigos) surtiu efeito significativamente maior do que a mensagem puramente informativa. Além disso, ao analisar as redes de amizade, constatou-se que amigos próximos (laços fortes) exerceram influência muito maior: usuários com fortes conexões que viram amigos votando foram mais propensos a votar, e esses laços fortes explicaram cerca de 4 vezes mais influências adicionais do que a mensagem em si (Bond et al., 2012). Ou seja, o contágio social ocorreu predominantemente entre amigos íntimos, indicando que as redes pessoais off-line refletidas on-line amplificaram o estímulo (amigos próximos motivando uns aos outros a votar), enquanto laços

fracos tiveram efeito marginal.

Esse experimento forneceu evidências claras de que mobilizações políticas online podem ter impacto no mundo real, validando empiricamente suposições antes debatidas. Bond et al. concluíram que a mobilização política via redes sociais funciona: não só aumentou a autoexpressão política (usuários clicando “Eu Votei”), mas também levou a ações concretas verificáveis (votos adicionais) (Bond et al., 2012). Destacaram ainda que estudos prévios não haviam detectado efeito de mensagens online possivelmente por falta de poder estatístico – aqui, a amostra gigantesca permitiu identificar um efeito relativamente pequeno, mas existente (Bond et al., 2012). O estudo também reforçou a superioridade de abordagens que exploram o elemento social (o engajamento pelas conexões pessoais) em comparação a comunicações unidirecionais: ver rostos familiares aumentou significativamente a eficácia da mensagem de mobilização (Bond et al., 2012).

Para além dos resultados substantivos, o experimento de 61 milhões demonstrou a viabilidade de integrar experimentos randomizados às plataformas digitais. Desde então, tornou-se comum empresas de tecnologia realizarem testes A/B contínuos com seus usuários para avaliar mudanças de produtos ou algoritmos. No meio acadêmico, outros experimentos de larga escala notáveis incluem: o estudo de “contágio emocional” no Facebook (Kramer et al., 2014), que envolveu ~700 mil usuários para avaliar se reduzir a exposição a posts positivos/negativos afetava o humor e publicações dos indivíduos; experimentos de disseminação de informação em redes (como testes de intervenções contra *fake news* entre milhares de usuários); ou experimentos de economia comportamental via apps (por exemplo, testando diferentes mensagens de incentivo à poupança em aplicativos financeiros com centenas de milhares de usuários).

Os benefícios científicos desses mega-experimentos são consideráveis: permitem estimar efeitos minúsculos com precisão, testar hipóteses de interação complexa (por exemplo, como efeitos variam conforme características de rede de cada pessoa), e observar comportamentos em contexto natural em vez de em laboratório artificial. Entretanto, surgem também preocupações éticas e práticas. Muitos participantes não estão conscientes de que fazem parte de um experimento (como no caso do Facebook, onde a “intervenção” apareceu como parte normal da interface), o que suscita questões sobre consentimento informado e manipulação psicológica. Houve debate público sobre a aceitabilidade dessas práticas, especialmente após divulgação do experimento de contágio emocional em 2014 – visto por alguns como invasão na autonomia emocional dos usuários. As empresas argumentam que seus Termos de Serviço cobrem tais testes e que eles se justificam pelos ganhos de conhecimento e melhorias de produto; já acadêmicos pedem maior transparência e aderência a normas éticas tradicionais de pesquisa (consentimento, possibilidade de *opt-out*, revisão por comitês de ética).

Além da ética, existe o desafio da reproduzibilidade: experimentos dentro de plataformas privadas dependem do acesso concedido pelas empresas e do comportamento do algoritmo proprietário, o que dificulta replicar estudos independentemente. No caso do experimento de Bond et al., a colaboração com a

plataforma foi crucial para chegar aos 61 milhões de participantes e aos dados de amizades, algo inviável sem a parceria. Isso aponta para a necessidade de parcerias academia-indústria e de repensar modelos de governança de dados para pesquisa científica.

Em síntese, os experimentos digitais em larga escala inauguraram um novo patamar para a ciência social causal, unindo o controle experimental clássico à abrangência populacional. O estudo de Bond e colegas (2012) mostrou que é possível randomizar intervenções sociais significativas e medir seus efeitos no comportamento de milhões de pessoas no mundo real (Bond et al., 2012). Essa capacidade traz tanto entusiasmo – por permitir testar teorias em contextos amplos e ecológicos – quanto cautela, pois exige que padrões éticos acompanhem o poder experimental. A medida que as CSC avançam, espera-se que esses experimentos massivos sejam usados com responsabilidade, para que possamos extraír insights sobre influência social, cooperação, difusão e outros fenômenos de forma ética e científica.

Estudos de Caso em CSC: Difusão de Notícias e Previsões Algorítmicas

Para ilustrar de forma mais concreta os alcances das Ciências Sociais Computacionais, examinamos a seguir dois estudos de caso influentes: um sobre difusão de informações online (Vosoughi et al., 2018) e outro sobre predição algorítmica em decisões judiciais (Kleinberg et al., 2018). Cada um exemplifica uma dimensão distinta do campo – respectivamente, o uso de *big data* para entender dinâmicas sociais complexas (propagação de *fake news*) e a aplicação de *machine learning* para melhorar políticas públicas (decisões de fiança no sistema criminal).

(i) **Difusão de notícias verdadeiras vs. falsas nas redes sociais.** No estudo “The spread of true and false news online” publicado na *Science* em 2018, Soroush Vosoughi, Deb Roy e Sinan Aral analisaram cascatas de disseminação de notícias no Twitter ao longo de 11 anos (2006–2017) para comparar o alcance e velocidade de rumores verdadeiros vs. falsos (Vosoughi; Roy; Aral, 2018; Vosoughi; Roy; Aral, 2018). Eles construíram uma base com ~126 mil cascatas de boatos que circularam na rede, classificadas como verdadeiras ou falsas com base em verificações de seis agências de checagem de fatos independentes. Com essa ampla amostra, aplicaram técnicas de análise de redes temporais para medir métricas como: quantas pessoas cada notícia atingiu, quantos retweets máximos em sequência (profundidade da cascata), quão rápido chegou a certo número de usuários, etc. Os resultados foram surpreendentes e alarmantes: notícias falsas difundiram-se significativamente mais longe, mais rápido, mais profundamente e de forma mais ampla do que as verdadeiras em todas as categorias de informação analisadas (Vosoughi; Roy; Aral, 2018). Especificamente, *fake news* tiveram 70% mais probabilidade de serem retuitadas do que notícias verdadeiras (Vosoughi; Roy; Aral, 2018). Narrativas falsas atingiram grupos de 1.500 pessoas seis vezes mais rapidamente do que as verdadeiras atingiam o mesmo número de pessoas

(Vosoughi; Roy; Aral, 2018). Além disso, as cadeias de retweet das falsidades tendiam a ser muito mais longas: falsos boatos alcançavam profundidade 10 (dez retweets em sequência) cerca de 20 vezes mais rápido do que os verdadeiros (Vosoughi; Roy; Aral, 2018). E em termos de abrangência, as falsas notícias também envolviam um número maior de usuários únicos em cada nível de difusão (Vosoughi; Roy; Aral, 2018).

Os autores examinaram diversas possíveis explicações para esse diferencial. Uma hipótese comum era a atuação de *bots* automáticos espalhando desinformação. Porém, ao filtrarem da análise todos os perfis suspeitos de serem robôs, as diferenças permaneceram praticamente iguais (Vosoughi; Roy; Aral, 2018). Isso indica que os próprios humanos foram os principais responsáveis por impulsionar a maior virilidade das *fake news*, e não apenas exércitos de *bots* programados (Vosoughi; Roy; Aral, 2018). Por que as pessoas compartilham mais o que é falso? Vosoughi et al. sugerem que a novidade das notícias falsas – frequentemente apresentando informações surpreendentes ou sensacionalistas – pode engajar mais as pessoas, levando-as a retuitar para “socialmente” alertar os outros ou simplesmente pela novidade em si (Vosoughi; Roy; Aral, 2018). De fato, análises de conteúdo mostraram que as falsidades tendiam a provocar reações de surpresa e indignação maiores do que as verdades (que frequentemente eram desmentidos ou confirmações do *status quo*). Esse apelo emocional e de novidade dá vantagem competitiva à mentira na ecologia da atenção online.

Esse estudo de caso ilustra a capacidade das CSC de revelar padrões emergentes em sistemas sociais complexos. A propagação de rumores em redes sociais envolve milhões de agentes com interconexões diversas – um problema difícil de analisar sem ferramentas computacionais e dados abrangentes. Ao mobilizar técnicas de *big data*, Vosoughi e colegas quantificaram empiricamente um fenômeno que havia sido teorizado mas pouco evidenciado: a de que a desinformação se espalha de modo “mais rápido e mais profundo” que a informação verificada. Os achados tiveram ampla repercussão, alertando para os perigos das plataformas digitais como amplificadoras de boatos infundados e motivando pesquisas subsequentes em combate à desinformação. Também evidenciaram a importância de fatores humanos – como emoção e surpresa – na difusão viral, fornecendo insights sociopsicológicos. Trata-se, portanto, de um caso em que métodos de mineração de dados e análise de rede serviram para iluminar um desafio social da era digital (*fake news*), quantificando-o e orientando possíveis intervenções (por exemplo, estratégias de *fact-checking* mais atraentes, ou ajustes nos algoritmos de curadoria de conteúdo para reduzir a viralidade de conteúdo suspeito).

(ii) Previsões algorítmicas e decisões sociais: o caso das fianças judiciais. O segundo estudo, “Human Decisions and Machine Predictions” (Kleinberg et al., 2018), insere-se na interseção entre ciência de dados e políticas públicas, explorando como modelos preditivos podem melhorar decisões tomadas por humanos. Especificamente, os autores analisaram a decisão de juízes em audiências de custódia (que definem se um réu aguardará julgamento preso ou em liberdade) como um problema de predição de risco: o juiz tenta prever se o acusado, se solto, cometerá crimes durante a liberdade provisória ou faltará ao julgamento. O estudo aproveitou um enorme conjunto de dados do sistema

judicial de Nova York, com centenas de milhares de casos e suas características (histórico criminal, idade, tipo de crime, etc.) e se o réu foi solto ou preso, bem como os resultados (se cometeu delito ou compareceu ao tribunal). Com esses dados, os pesquisadores treinaram um modelo de *machine learning* para predizer o risco de cada acusado cometer crime se solto, usando apenas as informações disponíveis ao juiz no momento da decisão (Kleinberg et al., 2018). Em seguida, compararam as recomendações do algoritmo com as decisões reais dos juízes e seus respectivos resultados, simulando cenários hipotéticos.

As conclusões foram notáveis: o algoritmo conseguiu distinguir com mais acurácia quais réus representavam alto risco e quais eram seguros para serem liberados, em comparação aos acertos e erros dos juízes. Nas simulações de política pública, os autores estimaram que, se as decisões de soltura/prisão fossem informadas pelas previsões do modelo, poder-se-ia reduzir os crimes durante liberdade provisória em até ~25% sem aumentar a taxa de prisão preventiva, ou alternativamente reduzir em até ~42% o número de presos preventivos sem elevar a criminalidade (Kleinberg et al., 2018). Em todos os cenários, haveria ganhos simultâneos: menos crime e/ou menos encarceramento, ou seja, as previsões algorítmicas permitiriam superar o *trade-off* atual enfrentado pelos juízes (que soltam alguns indivíduos perigosos e prendem desnecessariamente alguns que não cometem delitos) (Kleinberg et al., 2018). Além disso – ponto crucial – essas melhorias poderiam ser obtidas reduzindo disparidades raciais. O modelo preditivo, aplicado de forma uniforme, recomendaria soltar mais acusados negros de baixo risco que atualmente são presos (possivelmente devido a vieses ou menor acesso a fiança), equilibrando o tratamento (Kleinberg et al., 2018). Assim, não apenas eficiência seria aumentada, mas também equidade no tratamento entre brancos e negros melhoraria segundo os resultados simulados.

O estudo de Kleinberg et al. exemplifica como a análise computacional pode diagnosticar ineficiências nas decisões humanas e propor soluções baseadas em dados. Os autores mostram que os juízes, sob diversas limitações cognitivas e de informação, não tomam decisões ótimas e que um simples modelo estatístico usando os mesmos dados poderia salvar centenas de vidas (por crimes evitados) ou poupar milhares de pessoas da prisão desnecessária (Kleinberg et al., 2018). No entanto, eles também enfatizam que integrar algoritmos na tomada de decisão não é trivial: é preciso considerar que os dados de treinamento vêm de decisões passadas dos próprios juízes (logo, contêm vieses condicionais), e que juízes podem ter objetivos múltiplos (preocupação especial com crimes violentos, por exemplo, ou aversão a liberar reincidentes) (Kleinberg et al., 2018). Para lidar com isso, os pesquisadores usaram técnicas econômétricas (quase-experimentos com aleatoriedade na alocação de casos a juízes lenientes ou rigorosos) para isolar efeitos, e discutem que qualquer implementação de algoritmo precisaria ser acompanhada de um marco normativo claro – definindo a função objetivo (minimizar crime? minimizar dias de prisão?) e garantindo contrapesos para equidade (Kleinberg et al., 2018). Em suma, o caso demonstra tanto o potencial dos métodos de predição em aprimorar políticas sociais quanto a necessidade de enquadrá-los em um contexto econômico-jurídico adequado.

Outros estudos de caso importantes no repertório das CSC poderiam ser

citados, como o trabalho de Lazer et al. (2014) sobre o fracasso do Google Flu Trends, ressaltando os perigos de confiar cegamente em *big data* sem teoria; o estudo de King, Pan e Roberts (2013) sobre censura na Internet chinesa, que usou análise de texto automatizada para inferir os critérios de censura governamental; ou pesquisas mais recentes sobre comportamento de *bots* e polarização, predição de resultados eleitorais por mídias sociais, entre muitos. Entretanto, os dois exemplos aqui detalhados – difusão de *fake news* e predição em justiça – já evidenciam a amplitude temática da área. As Ciências Sociais Computacionais aplicam-se desde questões macro (opinião pública, comunicação de massa) até microdecisões institucionais (tribunais, hospitais, escolas), utilizando um arsenal que vai da mineração de dados observacionais à experimentação *in vivo* e à simulação de cenários. Com isso, esses estudos de caso nos encaminham a refletir sobre os desafios transversais que permeiam o campo – em especial no tocante a vieses, limitações e ética, tema da próxima seção.

Desafios, Limites e Implicações Éticas

Apesar do entusiasmo gerado pelas inovações das Ciências Sociais Computacionais, há um consenso de que o campo enfrenta desafios significativos e diversas limitações que precisam ser reconhecidos e tratados com seriedade. Muitos desses desafios são metodológicos (como lidar com vieses em *big data*, problemas de causalidade, replicabilidade), enquanto outros são de ordem ética e social (privacidade, consentimento, equidade e efeitos indesejados das aplicações). Nesta seção, discutimos alguns dos principais pontos críticos, com destaque para a questão dos viéses algorítmicos e implicações éticas, ilustrados pelo estudo de Obermeyer et al. (2019).

Viéses e imparcialidade algorítmica: Uma preocupação central é que os algoritmos e modelos preditivos possam perpetuar ou amplificar injustiças sociais preexistentes se treinados em dados enviesados. Ziad Obermeyer e colegas (2019) forneceram um exemplo impactante ao revelar um viés racial importante num algoritmo amplamente utilizado na saúde dos EUA (Obermeyer et al., 2019). Esse algoritmo, adotado por planos de saúde para identificar pacientes de alto risco que deveriam receber programas de gestão de cuidados, teoricamente deveria prever quem são os pacientes mais doentes. Contudo, Obermeyer et al. descobriram que o algoritmo usava como variável *proxy* o custo financeiro em saúde do paciente (gastos médicos históricos) para inferir necessidade de cuidados futuros. Devido a disparidades de acesso e qualidade do sistema de saúde, pacientes negros tipicamente têm menos gastos médicos do que brancos igualmente enfermos (pois enfrentam barreiras e tratamentos inadequados). O resultado: para um mesmo escore de risco dado pelo algoritmo, pacientes negros estavam na verdade muito mais doentes que pacientes brancos (Obermeyer et al., 2019). Em outras palavras, o sistema calibrado em custo subestimava sistematicamente a gravidade dos pacientes negros. Quantitativamente, a pesquisa mostrou que, em um dado nível de risco, os pacientes afro-americanos apresentavam muito mais doenças não controladas; e que corrigir esse viés quase triplicaria a proporção

de pacientes negros elegíveis para programas de cuidado intensivo – de 17,7% para 46,5% (Obermeyer et al., 2019). O viés portanto afetava milhões de indivíduos, excluindo negros do acesso a cuidados adicionais justamente por serem subtratados historicamente (Obermeyer et al., 2019).

Esse caso expõe uma lição crucial: algoritmos são tão bons quanto os objetivos e dados com que são construídos. A escolha de uma variável “conveniente” (gasto) como *proxy* para a verdadeira variável de interesse (saúde) introduziu um viés oculto, que não era evidente nos critérios usuais de acurácia preditiva (afinal, o algoritmo predizia bem os gastos futuros – só que gastos não equivalem a necessidade médica) (Obermeyer et al., 2019). Situações análogas podem ocorrer em diversas aplicações de CSC: sistemas de seleção de candidatos a emprego que aprendem com históricos onde certos grupos estavam sub-representados ou foram discriminados; modelos preditivos de crime que, alimentados por dados policiais já enviesados por vigilância desigual de bairros, reforçam estereótipos geográficos e raciais (*feedback loop*); recomendações em plataformas que marginalizam conteúdo de minorias, etc. Assim, detectar e mitigar vieses algorítmicos tornou-se um subcampo vital. Estratégias incluem: ajustar o alvo do modelo (como sugerido por Obermeyer et al., privilegiar variáveis diretamente clínicas em vez de custo); introduzir correções ou reponderações para grupos sub-representados; auditar modelos quanto a métricas de justiça (paridade de falsas negativas/positivas entre grupos); e envolver especialistas *domain* e *stakeholders* no desenho de algoritmos éticos.

Privacidade e consentimento: Outro desafio evidente é que o uso de *big data* muitas vezes entra em tensão com a privacidade individual. Dados de mídias sociais, telefone, geolocalização e outros são tipicamente dados pessoais sensíveis. Pesquisadores de CSC precisam navegar questões como anonimização, consentimento e potencial de mau uso. Mesmo dados anonimizados podem, se combinados, reidentificar pessoas ou expor informações comportamentais que os indivíduos não esperavam ver escrutinadas por terceiros. Por exemplo, análises de redes sociais podem revelar orientação sexual ou opiniões políticas de usuários que nunca declararam isso explicitamente, por meio de inferências baseadas em amigos e curtidas. Há ainda o problema do consentimento informado em massa: raramente é factível obter consentimento específico de milhões de usuários cujos dados (públicos ou de API) são analisados – e muitos forneceram esses dados para fins totalmente distintos (p. ex., uso pessoal da rede social, não participação em pesquisa acadêmica). Isso levanta dilemas éticos: alguns argumentam que dados publicamente disponíveis podem ser usados desde que anonimizados; outros defendem um modelo de consentimento aberto e contínuo (como painéis de cidadãos conscientes); e há esforços para envolvimento do público nas decisões sobre pesquisa com dados agregados.

No caso do estudo de Blumenstock sobre telefonia, os autores sublinharam que a pesquisa foi conduzida respeitando padrões éticos e a privacidade dos assinantes, com aprovação tanto do comitê de ética universitário quanto acordos estritos de uso com a operadora (Blumenstock; Cadamuro; On, 2015). Este cuidado é exemplar: implica avaliar riscos de identificação e impactos possíveis sobre as comunidades envolvidas (por exemplo, se o mapa de pobreza fosse usado

para estigmatizar regiões). Muitos projetos de CSC buscam seguir códigos de ética adaptados (Menlo Report, Belmont Report atualizado para ICT, etc.), mas a diversidade de contextos torna desafiador criar diretrizes universais. Um princípio emergente é o “*Privacy by Design*”, incorporando técnicas de proteção de privacidade (agregação, ruído aleatório, criptografia de dados sensíveis) desde a concepção do estudo.

Transparência e reproduzibilidade: A CSC também enfrenta problemas de opacidade algorítmica – modelos complexos de *machine learning* podem ser difíceis de interpretar ou explicar para públicos não técnicos. Em contextos de política pública, essa falta de transparência pode minar a confiança ou dificultar a contestação de decisões automatizadas. O campo de “explicabilidade de IA” vem buscando métodos para tornar modelos mais interpretáveis, mas há um *trade-off* nem sempre resolvível entre complexidade (e acurácia) e interpretabilidade. Para cientistas sociais preocupados com mecanismos causais e narrativas interpretativas, essa “caixa-preta” é problemática. Assim, muitos defendem uso de modelos mais simples ou *post-hoc* explicações (como importância de variáveis, casos similares, etc.) quando resultados de algoritmos são comunicados.

A replicabilidade dos estudos de CSC é outro ponto delicado. Muitas análises dependem de dados proprietários (de empresas) ou de grande infraestrutura computacional. Isso dificulta que outros pesquisadores reproduzam exatamente os resultados, ao contrário do ideal científico de replicabilidade. Iniciativas de ciência aberta têm incentivado o compartilhamento de códigos, amostras de dados e resultados agregados, mas restrições contratuais e de privacidade às vezes impedem a disponibilização integral dos dados. Sem acesso aos mesmos dados, replicar é virtualmente impossível. Esse cenário pede novas abordagens, como parcerias replicadoras (pesquisadores secundários reanalizando dados sob supervisão das empresas, por ex.), desafios de previsão abertos com dados similares, ou formas de *sandbox* seguro onde múltiplos grupos possam analisar dados sensíveis de maneira controlada.

Dilemas epistemológicos: Do ponto de vista científico, discute-se se o enfoque das CSC não arrisca priorizar correlação em detrimento de compreensão. Com tantos dados e poder computacional, é tentador achar padrões e relações estatísticas sem uma teoria orientadora – e alguns críticos apontam para estudos de *big data* que encontraram “previsores surpreendentes” mas falharam em explicar o porquê, contribuindo pouco para o conhecimento teórico. A resposta dos entusiastas é que padrões empíricos podem inspirar novas teorias e que existe espaço tanto para abordagens guiadas por dados (“*data-driven*”) quanto guiadas por teoria nas CSC. Todavia, fica o alerta: quantidade não substitui qualidade dos dados. Muitos grandes conjuntos sofrem de viés de seleção (quem está online?), erro de mensuração (os *likes* medem realmente satisfação?) ou confundimento (variáveis latentes não observadas). Assim, desafios clássicos de inferência causal permanecem – e de fato podem se agravar, já que com *big data* praticamente tudo se correlaciona com algo. Identificar causas reais exige desenho de pesquisa rigoroso (experimentos, quase-experimentos, controles adequados) tanto quanto nas ciências sociais tradicionais.

Implicações sociais e éticas das aplicações: Finalmente, há o amplo

debate sobre as consequências sociais da aplicação de modelos computacionais. Se por um lado podem tornar políticas mais eficientes e justas (como sugerido no caso das fianças), por outro, um uso descuidado ou mal intencionado de *big data* e algoritmos pode prejudicar comunidades vulneráveis. Exemplos incluem: sistemas de pontuação de crédito ou “*score social*” opacos que excluem indivíduos; vigilância em massa governamental usando mídia social e reconhecimento facial (comprometendo liberdades civis); manipulação comportamental via algoritmos de recomendação (como no escândalo Cambridge Analytica, em que dados de perfis do Facebook foram usados para direcionamento político). Esses riscos impulsionam a demanda por regulação e governança. A União Europeia, por exemplo, incluiu no GDPR (Regulamento Geral de Proteção de Dados) direitos relativos à explicação de decisões automatizadas e consentimento para uso de dados. Há movimentos para responsabilização algorítmica, auditorias independentes de algoritmos em setores críticos (bancos, justiça, saúde). Acadêmicos argumentam por códigos de ética profissionais e capacitação de cientistas de dados em ética (por isso muitas edições do SICSS e outras escolas de verão incorporam módulos de ética em CSC).

Em termos de limites, é importante reconhecer que as CSC não conseguem responder a todas as perguntas. Fenômenos que não deixam rastro digital (ou cujos rastros são inacessíveis) permanecem difíceis de estudar – e muitas das variáveis de interesse das ciências sociais (por ex.: crenças íntimas, emoções profundas, dinâmica familiar) podem não ser inferíveis apenas de *big data* comportamental. Além disso, o campo enfrenta a “falácia do n=infinito”: ter milhões de observações não elimina problemas de viés sistemático nos dados. Como bem coloca Salganik (2019), às vezes melhor qualidade de dados (coletados de forma adequada à pergunta) é preferível a mais quantidade de dados (Salganik, 2019). Nesse sentido, métodos tradicionais como *surveys*, etnografias e experimentos controlados mantêm seu valor e frequentemente complementam análises computacionais.

Concluindo esta seção, vemos que as Ciências Sociais Computacionais estão em fase de construir seus pilares éticos e críticos. Reconhecer e enfrentar esses desafios é parte do amadurecimento do campo. Há otimismo de que muitos problemas podem ser mitigados – por exemplo, Obermeyer et al. não apenas denunciaram o viés, como colaboraram para recalibrar o algoritmo de saúde com melhor equidade; e diversos pesquisadores têm proposto formas de “algoritmos justos” e “IA ética”. Lazer et al. (2020) ressaltam que é preciso melhorar compartilhamento de dados, incentivos acadêmicos e treinamento, de modo a alinhar as instituições científicas do século XX com as demandas intelectuais das CSC (Lazer et al., 2020). Ou seja, a comunidade científica, empresas e formuladores de políticas devem trabalhar juntos para ajustar estruturas de ética, pedagogia e infraestrutura à nova realidade de pesquisa baseada em *big data*. Somente assim poderemos colher os frutos das CSC – em conhecimento e benefício social – minimizando seus riscos e evitando potenciais malefícios.

Considerações Finais

As Ciências Sociais Computacionais representam uma fronteira emergente que está reconfigurando a maneira de fazer pesquisa em ciências sociais. Ao longo deste relatório, vimos que a área se caracteriza pelo uso criativo de dados massivos e métodos computacionais avançados para explorar questões clássicas e contemporâneas sobre a sociedade. Resumindo os principais pontos discutidos:

- **Síntese do campo:** As CSC surgem da convergência entre a disponibilidade de grandes dados digitais e o rigor analítico das ciências sociais. Elas expandem as capacidades empíricas, permitindo observar fenômenos sociais em alta resolução (tanto granular quanto abrangente) e experimentar intervenções em escala populacional, algo antes inviável. Vimos exemplos de como isso lançou novas luzes sobre padrões de comportamento (como a difusão de *fake news*) e permitiu repensar políticas públicas com auxílio de algoritmos (como nas decisões judiciais). Ao mesmo tempo, as CSC não rompem completamente com a tradição – elas constroem sobre fundamentos de teoria social, estatística e metodologia estabelecidos, adicionando novas ferramentas ao ferramental existente. Em essência, trata-se de uma ampliação do horizonte de pesquisa, mais do que uma substituição das abordagens anteriores.
- **Dilemas metodológicos:** Um tema transversal foi a tensão entre explicação e predição, ou de forma mais ampla, entre a busca de entendimento causal versus a utilidade preditiva. Discutimos que as CSC devem equilibrar essas duas metas: aproveitar a capacidade preditiva dos modelos de *machine learning* sem perder de vista as perguntas substantivas de por que os fenômenos ocorrem. Esse dilema se manifesta, por exemplo, na preferência às vezes por modelos de “caixa-preta” altamente acurados, que contudo desafiam interpretação – cabendo ao pesquisador decidir conforme o objetivo do estudo. Houve progresso conceitual nesse debate, com reconhecimento de que explicação e predição são ambos válidos e complementares, mas requerem clareza de propósito para evitar confusões. Outro dilema metodológico é lidar com inferência causal em *big data* observacional: técnicas computacionais sofisticadas não eliminam a necessidade de desenhos robustos (experimentos, controles, análise longitudinal) para reivindicar causalidade. Assim, a integração de métodos – *data mining* para descobrir padrões, seguido de experimentos ou modelos causais para testá-los – pode ser um caminho produtivo.
- **Desafios éticos:** Talvez o aspecto mais crucial seja garantir que a revolução das CSC ocorra de maneira ética e benéfica. A coleta e uso extensivo de dados pessoais exige salvaguardas de privacidade e consentimento; o emprego de algoritmos em decisões que afetam vidas humanas (saúde, justiça, crédito) demanda transparência, equidade e possibilidade de contestação. Exemplificamos com o caso Obermeyer et al. como até algoritmos usados

com boas intenções podem ocultar vieses graves se não forem cuidadosamente auditados. Isso alerta para a responsabilidade dos cientistas de dados sociais em antecipar consequências e monitorar os efeitos de suas ferramentas no mundo real. Felizmente, a conscientização sobre essas questões cresce rapidamente. Existe hoje um esforço em toda a comunidade científica – incluindo em iniciativas educacionais como o SICSS – de enfatizar ética, responsabilidade e engajamento público na formação de novos pesquisadores. Em paralelo, órgãos reguladores começam a atuar, e uma demanda por “algoritmos auditáveis” e “IA confiável” ganha corpo.

- **Perspectivas futuras:** O futuro das Ciências Sociais Computacionais parece promissor e dinâmico. Em termos de dados, novas fontes continuarão a emergir – pensemos em dados de sensores de Internet das Coisas, registros genômicos combinados com comportamentais, ou dados de plataformas descentralizadas. Cada nova fonte trará oportunidades de insight, mas também a necessidade de novas técnicas e atenção a novos tipos de viés. No lado metodológico, espera-se uma maior hibridização: modelos que combinam *machine learning* e inferência causal (por exemplo, usando árvores de decisão para descobrir heterogeneidade de efeitos de políticas), ou parcerias entre cientistas sociais qualitativos e cientistas de dados para interpretar resultados complexos. Academicamente, o campo tenderá a se institucionalizar mais – com revistas específicas, linhas de pesquisa consolidadas e possivelmente teorias e modelos próprios para fenômenos descobertos via dados computacionais (assim como a econometria e estatística social criaram paradigmas próprios no século XX).

Um ponto de cautela é que ainda estamos nos “primeiros dias” dessa revolução. Como observou Salganik, a pesquisa social digital ainda não produziu mudanças de paradigma teóricas de grande porte, mas o ritmo acelerado de melhora das ferramentas e dados é o que mais entusiasma (Salganik, 2019). Em outras palavras, há um progresso incremental rápido – ano após ano, novos limites são rompidos (ex.: algoritmos de linguagem natural que entendem texto quase como humanos; experimentos online cada vez mais sofisticados) – e isso sugere que as contribuições intelectuais mais profundas podem estar por vir.

No entanto, se mal conduzidas, as CSC também poderiam enfrentar uma reação (*backlash*). Escândalos de uso indevido de dados ou erros graves de algoritmos podem erodir a confiança do público e de formuladores de política na pesquisa computacional. Portanto, a continuidade do campo depende de autorregulação responsável e de evidenciar claramente os benefícios sociais advindos dessas abordagens. Caso isso ocorra, podemos imaginar, por exemplo, sistemas urbanos inteligentes que usem dados em tempo real para melhorar transporte e reduzir desigualdades; ou políticas de saúde preditivas que alocam recursos com máxima eficiência sem vieses – tudo isso fundamentado por pesquisas de CSC sólidas.

Em conclusão, as Ciências Sociais Computacionais oferecem um novo repertório para entendermos e agirmos sobre fenômenos sociais, mas trazem junto novas

responsabilidades. Este relatório delineou tanto as oportunidades – em termos de novos dados, métodos e conhecimentos – quanto os desafios que precisam ser enfrentados. Para os estudantes e pesquisadores que ingressam nesse campo, a mensagem é de equilíbrio: empregar-se com o poder analítico inédito, mas manter o rigor crítico e o compromisso ético que sempre foram marcas das boas ciências sociais. Assim, poderá se concretizar a visão otimista de uma ciência social verdadeiramente ampliada pela computação, gerando insights inovadores sobre a vida em sociedade no século XXI.

Referências Bibliográficas

- BLUMENSTOCK, Joshua E.; CADAMURO, Gabriel; ON, Robert. Predicting poverty and wealth from mobile phone metadata. *Science*, v. 350, n. 6264, p. 1073–1076, 2015.
- BOND, Robert M. *et al.* A 61-million-person experiment in social influence and political mobilization. *Nature*, v. 489, n. 7415, p. 295–298, 2012.
- BREIMAN, Leo. Statistical modeling: The two cultures. *Statistical Science*, v. 16, n. 3, p. 199–215, 2001.
- GENTZKOW, Matthew; KELLY, Bryan; TADDY, Matt. Text as Data. *Journal of Economic Literature*, v. 57, n. 3, p. 535–574, 2019.
- GRIMMER, Justin; ROBERTS, Margaret E.; STEWART, Brandon M. *Text as Data: A New Framework for Machine Learning and the Social Sciences*. Princeton: Princeton University Press, 2022.
- GRIMMER, Justin; STEWART, Brandon M. Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, v. 21, n. 3, p. 267–297, 2013.
- JEAN, Neal *et al.* Combining satellite imagery and machine learning to predict poverty. *Science*, v. 353, n. 6301, p. 790–794, 2016.
- KLEINBERG, Jon *et al.* Human decisions and machine predictions. *Quarterly Journal of Economics*, v. 133, n. 1, p. 237–293, 2018.
- LAZER, David *et al.* Computational social science: Obstacles and opportunities. *Science*, v. 369, n. 6507, p. 1060–1062, 2020.
- OBERMEYER, Ziad; POWERS, Brian; VOGELI, Christine; MULLAINATHAN, Sendhil. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, v. 366, n. 6464, p. 447–453, 2019.
- SALGANIK, Matthew J. *Bit by Bit: Social Research in the Digital Age*. Princeton: Princeton University Press, 2019.
- SHMUELI, Galit. To Explain or to Predict? *Statistical Science*, v. 25, n. 3, p. 289–310, 2010.
- SICSS – Summer Institute in Computational Social Science. Resources and materials. Disponível em: <https://sicss.io>. Acesso em: 25 out. 2025.
- VOSOUGHI, Soroush; ROY, Deb; ARAL, Sinan. The spread of true and false news online. *Science*, v. 359, n. 6380, p. 1146–1151, 2018.