

**Entrevista**

# Entrevista com Tiago Ventura – Ciências Sociais Computacionais: entre fronteiras e definições

DOI: <https://doi.org/10.14244/agenda.2024.3.10>

**iD Murilo de Oliveira Junqueira**

Doutor em Ciência Política pela Universidade de São Paulo (USP). Professor do Programa de Pós-Graduação em Ciência Política e da Faculdade de Ciências Sociais da Universidade Federal do Pará (UFPA). Pesquisador visitante do Instituto de Pesquisa Econômica Aplicada do Distrito Federal (IPEA-DF).

E-mail: mjunqueira@ufpa.br

Orcid: <https://orcid.org/0000-0003-3033-379X>

218

**iD Lucy Oliveira**

Doutora em Ciência Política pela Universidade Federal de São Carlos (UFSCar). Professora de Ciência Política no Departamento de Ciências Sociais (DCSo/UFSCar) e do Programa de Pós-Graduação em Ciência Política da Universidade Federal de São Carlos (PPGPol/UFSCar).

E-mail: gosilva@ufscar.br

Orcid: <https://orcid.org/0000-0002-2556-8730>

Tiago Ventura é paraense, cientista político, formado em Direito e apaixonado por Matemática. Esses interesses diversos o levaram ao campo das Ciências Sociais Computacionais (CSC's) durante seu doutorado na Universidade de Maryland, nos Estados Unidos. Com passagens por empresas de tecnologia e laboratórios de pesquisas interdisciplinares, além de um prêmio de melhor artigo no 14º Encontro da Associação Brasileira de Ciência Política (ABCP) de 2024. Tiago Ventura conversou com os coordenadores do dossiê da revista Agenda Política intitulado “**Ciências Sociais Computacionais: rupturas e contribuições para a Ciência Política**”, onde propôs debates e reflexões sobre ética na pesquisa com tecnologia e dados digitais, Inteligência Artificial e as potencialidades e desafios das Ciências Sociais Computacionais) e também indicou caminhos possíveis para quem quer ingressar no campo das CSC's que tem oportunidades na academia e no mercado. Atualmente, Ventura é professor na Universidade de Georgetown (Washington, Estados Unidos).



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

**Murilo de Oliveira Junqueira e Lucy Oliveira:** Professor Tiago Ventura, de forma resumida, o que são as Ciências Sociais Computacionais? É usar um *software*, aprender a codificar no R, usar inteligência artificial (IA) nas pesquisas?

**Tiago Ventura:** Eu considero que as Ciências Sociais Computacionais (CSC's) são um pouco de tudo isso na prática. Tem uma definição que eu gosto muito e é do professor Mathew Salganik da Universidade de Stanford, Estados Unidos. Salganik é cientista social computacional e organizador/fundador do *Summer Institute of Computational Social Science* (SICSS). Ele diz que CSC's são qualquer coisa que seja legal (*anything that is fun*). Essa definição é mais metafórica e abstrata, mas possui um espírito grandioso por trás: evitar impor limites a um campo novo, que vem se construindo nos últimos anos, é interdisciplinar e que está sendo impulsionado pelo barateamento do uso de computadores, pelo aumento da potência das máquinas e pela maior digitalização e acesso mais a dados.

Numa definição mais concreta, podemos dizer que as CSC's são a combinação entre cientistas sociais, interessados em entender o comportamento humano. E, as Ciências da Computação e de Dados, a partir do uso de modelos computacionais sofisticados em larga escala e do desenvolvimento de capacidade de trabalhar com dados não estruturados. Importante dizer que esses são dados diferentes dos que os cientistas sociais estão acostumados a trabalhar. Assim, na minha visão, as CSC's na prática são a combinação de perguntas substantivas das Ciências Sociais com técnicas da computação e ciência de dados. E essas áreas se nutrem mutuamente. Não é só cientistas sociais usando técnicas de programação e análise de dados, mas também cientistas da computação aprendendo com cientistas sociais sobre como fazer pesquisa, sobre desenho de pesquisa e sobre comportamento humano. Esse encontro permite aperfeiçoar áreas de conhecimento mais tradicionais da Computação e vice-versa.

**Murilo de Oliveira Junqueira e Lucy Oliveira:** Você vem da área das humanidades. Como você chega nas Ciências Sociais Computacionais (CSC's)?

**Tiago Ventura:** Minha trajetória é interessante. Sempre gostei muito de Matemática durante no colégio, sendo meu tópico favorito. Mas eu também sempre gostei muito de política, desde muito novo. Então, essa divisão acabou me levando a realizar duas graduações: Direito e Engenharia. Eu comecei, na verdade, as duas graduações porque gostava das áreas de saber, mas rapidamente deixei Engenharia, pois meu interesse pela Política prevaleceu, decidi me dedicar integralmente ao Direito. Contudo, logo percebi, na graduação em Direito, que não desejava advogar e nem praticar a profissão. Após a graduação, fiz mestrado no Instituto de Estudos Sociais e Políticos da Universidade do Estado do Rio de Janeiro (IESP/UERJ) em Ciência Política, quando comecei a entender melhor a área e ampliar meus horizontes, pois o mestrado me apresentou à pesquisa aplicada quantitativa. Então, eu percebi que era possível fazer ciência sobre questões políticas com métodos matemáticos robustos e

rigor científico, isso virou minha cabeça e eu decidi seguir esse caminho. Fiz diversos cursos de Matemática e Ciências da Computação durante o mestrado no IESP, o que fundamentalmente me incentivou a vir fazer um doutorado nos Estados Unidos em Ciência Política, já trabalhando a maior parte do meu doutorado no Laboratório de Ciências Sociais Computacional da Universidade de Maryland, com o orientador Ernesto Calvo. Lá, investi intensamente na formação mais forte em Matemática e Ciência Computacional, entrando de vez no universo das CSC's.

**Murilo de Oliveira Junqueira e Lucy Oliveira:** As Ciências Sociais Computacionais (CSC's) seriam uma área para Cientistas Sociais que trabalham com pesquisa quantitativa?

**Tiago Ventura:** Não. Acho que as CSC's também são para Cientistas da Computação que querem entender problemas sociais. Eu acho que os dois lados existem. Você tem doutorados das CSC's que recebem pessoas com graduação e mestrado em Ciências da Computação que desejam compreender melhor o comportamento humano, como, por exemplo, interação entre humanos e tecnologia, *Human-Computer Interaction*, vieses (*bias*) em sistemas algorítmicos e efeitos da tecnologia no comportamento humano. Essas são áreas das CSC's que Cientistas da Computação focam suas pesquisas e precisam entender questões relacionadas ao comportamento humano. E com essa revolução que a gente está vivendo em relação às Inteligências Artificiais (IAs), essas áreas têm crescido bastante e eu também entendo que fazem parte do diverso grupo de pesquisas e disciplinas que formam a CSC.

**Murilo de Oliveira Junqueira e Lucy Oliveira:** Teria espaço para pesquisadores com experiência qualitativa?

**Tiago Ventura:** Eu acredito que sim. Creio que faz parte do que descrevi como “tenda larga” das CSC's, onde cabe todo mundo que queira entender mais dessa relação entre tecnologia, comportamento humano e o uso de técnicas de análises baseadas em modelagem computacional. Importante dizer que eu não faço pesquisa qualitativa. Então, eu tenho dificuldade de te descrever exatamente como atuar. Mas, eu sei que em áreas como HCI (*Human-Computer Interaction*) tem diversas pesquisas que são qualitativas, pois as pessoas tentam entender os efeitos da tecnologia em comportamento humano usando grupos focais e entrevistas. E até no próprio mercado privado, sobretudo na parte de UX (*user experience*). Entretanto, é importante dizer que o pesquisador qualitativo ou quantitativo que vai trabalhar com CSC's precisa entender como funcionam os modelos de computação. Imagine um *UX researcher* fazendo grupo focal sobre experiências humanas com *chatbots*. Mesmo realizando pesquisa qualitativamente, é necessário conhecer como esses modelos funcionam, o que está por trás deles, portanto, tudo isso integra as áreas que se relacionam entre si. Novamente, não faço pesquisa qualitativa, mas minha impressão é que existe um diálogo para isso, embora represente uma parte pequena do campo.

**Murilo de Oliveira Junqueira e Lucy Oliveira:** Como é atuar como brasileiro nos Estados Unidos nessa área? E como você vê o campo no Brasil?

**Tiago Ventura:** Eu considero difícil comparar o Brasil com os Estados Unidos que, ao lado da China, está entre os países mais ricos do mundo. É uma comparação injusta. Porém, uma questão que nos interessa e que é possível destacar nessa diferença - à qual damos pouca atenção - é quanto a formação na graduação. No Brasil, ela é menos aberta à interdisciplinaridade do que nos Estados Unidos, e isso impacta a área das CSC's. O foco altamente disciplinar da academia brasileira é algo que precisa ser repensado. Nos Estados Unidos, a estrutura também é disciplinar, porém menos burocratizada, e os custos para colaborar com outros departamentos e obter uma formação interdisciplinar são menores. A academia brasileira é mais burocratizada do que a norte-americana, me refiro ao processo de contratação, no trânsito de alunos entre departamentos diferentes e permitir que professores transitem entre distintos departamentos.

Nos Estados Unidos, os estudantes têm uma possibilidade maior de fazer diferentes *majors*, ou seja, diferentes ênfases. Na verdade, é bem comum ter estudantes de doutorado na Ciência Política que possuem formação mais forte nas áreas de ciências exatas - Matemática e Estatística, diferente do que na formação brasileira. E isso ocorre porque tem, de fato, a possibilidade de em sua graduação você dialogar com diversas áreas, o que é mais difícil no modelo brasileiro.

221

Obviamente, há muitos discentes em Ciências Sociais que optam pelas humanidades porque não querem fazer matemática. Também existe dificuldade curricular e institucional para quem deseja aprender matemática de fato. Ou seja, uma pessoa que iniciou a graduação em Ciências Sociais e achou interessante entender problemas humanos utilizando computador, programação e dados da internet: onde ela aprende isso na universidade? É difícil, pois o currículo já está muito estruturado. Não há essa possibilidade de você fazer um *major* em Ciências Sociais com o *minor* (sub-graduação) em Ciências da Computação ou Matemática. Isso é bem mais comum na academia norte-americana. Isso impacta a formação dos alunos já na graduação. Alunos internacionais que não tiveram esse *background* em matemática na graduação precisam trabalhar em dobro no doutorado. Eu fiz isso, e outros colegas internacionais que vieram para os EUA também fizeram. Isso poderia ser evitado com mais interdisciplinaridade na graduação brasileira.

Por outro lado, os programas de doutorado aqui diferem dos do Brasil. Ambos são integrados ao mestrado. Nos dois ou três primeiros anos, há muita formação, seguida da defesa da tese. Em geral, esses primeiros anos de formação do doutorado nos Estados Unidos, na Ciência Política, pelo menos, são mais intensos em formação quantitativa — mais computacionais por isso. No Brasil, o doutorado é mais separado do mestrado, então você faz menos curso no início, pois já cursou um mestrado mais intenso. Essas são duas diferenças importantes entre as experiências brasileira e americana.

**Murilo de Oliveira Junqueira e Lucy Oliveira:** Pensando em quem quer fazer pesquisa em CSC, quais são hoje os tópicos quentes do campo?

**Tiago Ventura:** Então, destaco dois tópicos. O primeiro foi muito quente e agora estamos numa fase de “final de onda”, que envolve desinformação e mídias sociais. Portanto, são análises que buscam entender a circulação de conteúdos em geral: como eles viajam nas mídias sociais; o quanto de desinformação existe nas plataformas; o ambiente informacional das pessoas online e *offline*; as estruturas de redes de usuários e conteúdos; além de compreender como os humanos se comportam e o que eles consomem nas experiências *online*. Esse foi o *boom* das CSCs, um tema quente por muitos anos, que ainda é importante, mas tende a perder relevância nos próximos anos.

O segundo é o tema da Inteligência Artificial. Esse será o assunto no qual as CSC's irão dedicar bastante tempo nos próximos anos, com abordagens distintas. Tais como: i) vieses (*bias*) em modelos de IA- vieses étnico-raciais, políticos, considerando as características de cada país. Em relação aos vieses (*bias*), tenho uma colega estudando *bias* através da perspectiva do nacionalismo em *chatbots* e modelos de criação de texto; ii) treinamento de modelos: muitas pessoas tentam aperfeiçoar e tornar modelos mais acessíveis e de alta performance e iii) uso de modelos computacionais para experimentos, considero que esta será uma abordagem bastante promissora, tendo em vista de como utilizar esses modelos para construir o que a literatura chama de *tailored experiments*. Você pode usar esses modelos para fazer experimentos adaptativos, que se ajustam aos *inputs* fornecidos pelos participantes. É uma área que pode contribuir para melhorar os experimentos em Psicologia Política (*cognitive science*).

A popularização desses modelos de IA's ou os *Large Languages Models* (LLM's), para usar o termo técnico, será um tema bem importante das CSC, na minha opinião, com muitas possibilidades ainda a explorar, como a construção de agentes, melhores modelos de classificação, experimentos etc.

**Murilo de Oliveira Junqueira e Lucy Oliveira:** Temos visto o avanço do uso de *Large Languages Models* (LLM's) na pesquisa. Quais seriam os principais perigos e oportunidades do uso das Inteligentes Artificiais (IA's) generativas na Ciência?

**Tiago Ventura:** Essa área é toda muito nova, e a gente precisa de algum tempo para fazer um diagnóstico mais robusto e entender esse processo. Não há uma resposta definitiva sobre isso. Contudo, há alguns pontos fundamentais. O primeiro é a questão da transparência, representando um dilema central nessa nova revolução de modelos de linguagem, pois a maior parte deles são modelos de proprietários, sob domínio dos setores privados. As empresas com capacidade tecnológica de construir esses modelos em larga escala, com bilhões de parâmetros, alto desempenho e que superam as dificuldades inerentes à geração de linguagem nesses tipos de modelos, possuem grande capacidade de investimento alto e, por isso, desenvolvem esses modelos de forma privada. Esse dilema é também ético, incluindo milhares de pessoas e pesquisadores, que se utilizam desses modelos para suas pesquisas, mas esses modelos são todos “*black boxes*”. E não é porque a gente não saiba como os modelos funcionam, mas que não temos acesso a eles, a como foram treinados, nem sabemos o que é injetado para ajustar seus parâmetros. Por outro lado, penso que outras questões éticas que surgem agora são semelhantes às que enfrentávamos sobre o acesso a dados digitais e as revoluções tecnológicas anteriores. Esse é um

desafio nosso, das CSCs e da Ciência da Computação: compreender de forma mais crítica a revolução tecnológica que vivemos, focando nos efeitos sobre os seres humanos e agentes sociais, e conduzir pesquisas para responder essas questões. Não existe uma solução mágica neste momento, precisamos amadurecer o que o campo já vem realizando, temos que discutir os dilemas sobre acesso a dados digitais, proteção da privacidade e escala dos vieses em recomendações algorítmicas. Atualmente, esses temas têm sido escalados para outros níveis, mas que sempre permearam as preocupações das Ciências Sociais.

**Murilo de Oliveira Junqueira e Lucy Oliveira:** Pensando em transparência, vivemos no ano passado uma situação mundial *sui generis*. Mais de 60 países realizaram eleições em algum nível e, ao mesmo tempo, as plataformas digitais fizeram movimentos de restrição ao acesso aos dados que geram e coletam, inclusive para pesquisadores. Essa é uma conta que não fecha. Como você vê esse dilema?

**Tiago Ventura:** Acho que esse é um problema grave, as plataformas vem reduzindo acesso de pesquisadores aos dados resultantes das coletas dos usuários. Considero importante os esforços dos pesquisadores devem ser em pressionar os governos para aumentar a regulação em torno das plataformas. Temos a experiência da regulação na União Europeia, com muito esforço para fazê-la funcionar. Ainda não está muito claro quanto que essa medida vai funcionar de fato. Tem diversos pesquisadores muito sérios envolvidos nisso. Mas, é um avanço importante porque, por outro lado, não existe nenhuma regulamentação parecida nos Estados Unidos. E, até onde eu saiba, no Brasil também não existe, apesar do governo ter discussões parecidas nesse tema.

As plataformas têm enfrentado dificuldades de acessar os dados produzidos e coletados por elas e os pesquisadores não conseguem entender e ter domínio sobre os acontecimentos e todo o ambiente digital, o que impede a formulação de recomendações políticas assertivas sobre o funcionamento dessas plataformas. Isso se aplica também aos LLMs. Embora o *Facebook* disponibilize parte de seus modelos de acesso aberto (*open access*) - que permite às pessoas rodarem localmente - a principal produtora desses modelos de LLM's, a *Open IA*, lançou seus modelos mais recentes fechados, não *open access*. Hoje, as pessoas usam o *ChatGPT* para diversas pesquisas acadêmicas, mas não sabemos exatamente o que está nesses modelos, qual o banco de dados de treinamento e como podemos interpelar melhor seu conteúdo. Assim, a questão da falta de transparência acontece dos dois lados: não apenas nos dados de mídias sociais, mas também na transparência dos modelos de linguagem, tema que será o próximo ponto quente das CSC's.

**Murilo de Oliveira Junqueira e Lucy Oliveira:** Como poderia ser feita uma regulação? Atualmente, analisamos os seus impactos nas redes sociais, e nem sempre eles são positivos, mas não está claro o que poderia ser feito em termos de regulação.

**Tiago Ventura:** Não sabemos o impacto exatamente. As pesquisas mais recentes apontam acerca dos efeitos das mídias sociais em polarização não apresentam resultados robustos, pois não apontam para uma mesma direção. Há pesquisas sobre desativação (retirar usuários *do Facebook*), e o efeito disso em polarização varia em cada país. As pesquisas que eu faço sobre o efeito do *WhatsApp* e desinformação, ao retirar pessoas de sua rede, mostram que essa ação tem impactos muito pequenos ou praticamente nulos na polarização. Como ainda temos pouco acesso aos dados e pouca capacidade de colaboração com essa indústria, desconhecemos o efeito exato que eles causam na sociedade e com efeitos na política. Existe certo descolamento entre o que a sociedade acredita ser o impacto das mídias sociais e o que a pesquisa científica tem investigado sobre esses efeitos. As evidências indicam que essa percepção pode ser superestimada. Contudo, precisamos de mais pesquisas sobre esses temas para sermos mais assertivos. Não conseguimos avançar porque os dados não estão disponíveis e há pouca possibilidade de colaboração por parte das empresas. Um exemplo recente da importância dessa colaboração foi o conjunto de estudos que a *Meta* realizou com alguns acadêmicos sobre as eleições de 2020 nos Estados Unidos, publicados no ano passado.

Há diversos problemas nessa interação com a indústria, mas o fato concreto é que, graças a ela, pesquisadores produziram experimentos e pesquisas de alta qualidade. Essa foi uma situação em que a *Meta*, por interesses próprios, decidiu colaborar com um grupo de acadêmicos, sem exigência legal ou regulamentação. Assim, vivemos um cenário que depende das empresas para acesso a dados e possibilidades de pesquisa que podem existir entre academia e indústria. Hoje, o que temos de melhor é a experiência europeia e o que a União Europeia está tentando fazer. Os países ao redor do mundo deveriam olhar para a experiência do *Digital as a Service* e avaliar como adaptá-la a seus contextos nacionais.

224

**Murilo de Oliveira Junqueira e Lucy Oliveira:** A partir da sua experiência nos Estados Unidos, que recomendações você daria para implementar de forma mais efetiva a CSC na pesquisa e nas instituições brasileiras? Começar por onde?

**Tiago Ventura:** O Brasil é considerando um território de suma importância na região, levando em conta ainda o quanto de avanços já temos nas universidades brasileiras. Nesse sentido, o investimento na graduação é fundamental, precisamos construir departamentos mais interdisciplinares, sendo importante para o desenvolvimento do campo das CSC's e dar a possibilidade para alunos de graduação que quiserem se especializar e aprender em outros campos. Também é necessário deixar de ter medo da área da Ciências da Computação, da Matemática e da Estatística. Isso se faz ofertando aos alunos cursos interessantes e mostrando o quanto essa demanda é também do mercado, não exclusiva da academia ou da ciência. Discentes das Ciências Sociais que se formam com capacidade de programar e criar modelos matemáticos têm, de fato, maior empregabilidade em setores de tecnologia. Então, a longo prazo, isso é bom para eles também. Hoje, vivemos em um mundo que é mais fácil começar do

que era antes. Temos uma grande quantidade de cursos online, como *DataCamp* e *Coursera*, pelos quais as pessoas podem aprender mais de programação, matemática e estatística.

Além disso, outro ponto fundamental no nível individual, ou seja, se você é um aluno de Ciências Sociais, quer aprender CSC e ainda não está em um departamento interdisciplinar, tente verificar o que está acontecendo concretamente em sua universidade. Veja se há um curso de Ciências de Dados no Departamento de Ciências da Computação. Minha impressão é que isso hoje é muito mais comum no Brasil do que era quando eu estava na graduação. Isso é fundamental superar os desafios à sua formação. Você pode pensar: não sei se eu consigo fazer um curso de ciência de dados porque eu nunca trabalhei com isso! Acho que você pode tentar, nem que seja como aluno ouvinte. Faça o curso, tenta buscar também disciplinas de outros departamentos, porque esses cursos não só te ensinam coisas novas, mas também novas formas de pensar e de aprender. Os cursos de Ciências Sociais têm uma estrutura de aprendizado. Você lê o livro um artigo, você assiste às aulas, discute o artigo e faz um projeto ao final, representando um trabalho em equipe. Quando você vai para um curso de Ciências da Computação ou Estatística, é outro modelo, é diferente a forma de ensino-aprendizado, podendo ser bem distinta. Toda semana há uma lista de exercícios, ao resolver uma lista de exercícios, a maneira como se exercita a mente se difere de quando se realiza um trabalho em equipe para apresentar. Não há juízo de valor das áreas, nem destaque de uma forma seja melhor que a outra, portanto, é apenas mostra que as formas de aprendizado são diferentes.

225

A minha sugestão para quem tem interesse na temática da CSC's, diz respeito a arriscar em tentar aprender e entrecruzar os dois campos e outras áreas de saberes. No começo, vai ser difícil, você vai pensar: o que exatamente está acontecendo? Mas arrume um colega que queira aprender a mesma coisa que você e façam o curso juntos, aprendendo lado a lado. Arriscar em outras áreas, desafiar-se e aproveitar as inúmeras oportunidades online disponíveis hoje é um bom começo para quem deseja aprofundar seus conhecimentos sobre CSC's. Bolsas estudantis no exterior e, eventualmente, fazer uma formação completa fora para aprender mais sobre esse campo em outros países, retornando ao Brasil para construir essas pontes.

Em suma, pensar essa formação interdisciplinar de ponta a ponta, desde a graduação até a pós-graduação, portanto, os esforços como a construção de centros em comum e colaborações de pesquisa são essenciais.

**Murilo de Oliveira Junqueira e Lucy Oliveira:** Por fim, você indicaria qual livro ou atividade para quem vai ler essa entrevista e já vai querer começar nas CSC's?

**Tiago Ventura:** Indicaria duas literaturas. O livro seria o “*Bit by Bit*” e *Summer Institute of Computational Social Science* (SICSS), ambos do professor Matthew Salganik. Há alguns anos os pesquisadores da Fundação Getúlio Vargas (FGV) do Rio de Janeiro estão organizando a SICSS no Brasil e é uma oportunidade muito boa, porque os cursos são gratuitos e você se integra num projeto internacional. São dois pontos que para mim são fundamentais para começar bem.