



## Dependência e Soberania nas Tecnologias de Inteligência Artificial: uma análise a partir dos conceitos de acumulação primitiva de dados e *Data Processing Inequality*

*Dependence and Sovereignty in Artificial Intelligence Technologies: An Analysis Based on the Concepts of Primitive Data Accumulation and Data Processing Inequality*

Pedro Odebrecht Khauaja <sup>a,\*</sup> 

**RESUMO:** Essa pesquisa propõe uma elaboração teórica da relação entre soberania digital e a Teoria da Dependência, com o uso de conceitos como acumulação primitiva de dados e Desigualdade de Processamento de Dados. O estudo investiga como o modelo neoliberal atual de desenvolvimento das tecnologias de inteligência artificial (IA) se baseia em relações exploratórias entre o Norte e o Sul Global, emulando relações coloniais. Essa hipótese é discutida por autores do campo digital, com a mobilização de conceitos como colonialismo digital. **O objetivo é resgatar uma leitura geopolítica da Teoria da Dependência, especialmente nas releituras contemporâneas de Theotonio Dos Santos (2020) e Claudia Wasserman (2022).** A pesquisa propõe que a Teoria da Dependência dialoga com o conceito de acumulação primitiva de dados, conforme elaborado por Lippold e Faustino (2022), que sugere uma analogia entre a economia de dados atual e a acumulação inicial de bens produtivos na forma de propriedade individualizada, como identificado por Marx. Para construir esse diálogo, o estudo costura a Teoria da Dependência com o princípio da Desigualdade de Processamento de Dados, que afirma que qualquer processo de transformação de dados não pode aumentar a informação sobre a variável medida (BEAUDRY, RENNER, 2012). Esse princípio explica como a produção do “novo” nas tecnologias de IA depende necessariamente do aporte contínuo de novos dados, e propõe uma explicação para o avanço colonial das tecnologias de IA no Sul Global como desdobramento da necessidade de novos dados para a acumulação primitiva de capital.

**Palavras-chave:** Soberania Digital; Teoria da Dependência; Acumulação Primitiva de Dados; Desigualdade de Processamento de Dados; Colonialismo Digital.

**ABSTRACT:** This research proposes a theoretical framework for understanding the relationship between digital sovereignty and Dependency Theory, using concepts such as primitive data accumulation and Data Processing Inequality. The study investigates how the current neoliberal model of artificial intelligence (AI) technology development is based on exploitative relations between the Global North and South, emulating colonial dynamics. This hypothesis is discussed by scholars in the digital field, through concepts such as digital colonialism. The aim is to revive a geopolitical reading of Dependency Theory, particularly in the contemporary interpretations by Theotonio Dos Santos (2020) and Claudia Wasserman (2022). The research suggests that Dependency Theory aligns with the concept of primitive data accumulation, as outlined by Lippold and Faustino (2022), which draws an analogy between the current data economy and

<sup>a</sup> Faculdade de Direito, Universidade Federal Fluminense, Niteroi, RJ, Brasil.

\* Correspondência para/Correspondence to Pedro Odebrecht Khauaja . E-mail: [Pedro\\_odebrecht@id.uff.br](mailto:Pedro_odebrecht@id.uff.br)

Recebido em/Received: 15/10/2024; Aprovado em/Approved: 27/11/2024.

Artigo publicado em acesso aberto sob licença [CC BY 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) 

the initial accumulation of productive assets in individualized private property, as identified by Marx. To construct this dialogue, the study integrates Dependency Theory with the principle of Data Processing Inequality, which states that any data transformation process cannot increase the information about the measured variable (Beaudry, Renner, 2012). This principle explains how the production of the "new" in AI technologies necessarily depends on the continuous input of new data, and proposes an explanation for the colonial-like expansion of AI technologies in the Global South as a consequence of the need for new data to support the primitive accumulation of capital.

**Keywords:** Digital Sovereignty; Dependency Theory; Primitive Data Accumulation; Data Processing Inequality; Digital Colonialism.

## INTRODUÇÃO

Essa pesquisa trata de uma elaboração teórica da relação entre a ideia de soberania digital e o arcabouço conceitual da Teoria da Dependência, através da mobilização dos conceitos de acumulação primitiva de dados e Data Processing Inequality, que pode ser traduzido inicialmente como “Desigualdade de Processamento de Dados”. Essa elaboração se dá em torno do problema de como o atual modelo de desenvolvimento neoliberal de tecnologias de inteligência artificial se baseia em relações de exploração estruturadas entre o Norte e o Sul Global, atualizando relações anteriores de colonialismo.

No geral, essa hipótese tem sido aventada por muitos autores no campo digital, através da mobilização de conceitos como “colonialismo digital”, “capitalismo de vigilância”, de “plataforma” e “colonialismo de dados”—quando usado nesse contexto específico. A distinção entre esses conceitos feita por Silveira (2021) me leva a optar pelo termo “colonialismo de dados” como o que melhor convém para a presente elaboração conceitual, por reforçar o elemento da coleta de dados como central.

O que os conceitos têm em comum é sua relação com um viés marxista, que liga o movimento colonial com a estruturação do sistema capitalista. Os conceitos citados, em certa medida, atualizam o debate para dentro da economia política do digital. Assim, busco resgatar em parte a leitura geopolítica da Teoria da Dependência, especialmente suas releituras mais contemporâneas, como as feitas por Theotonio Dos Santos (2020) ou Cláudia Wasserman (2022).

Acredito que a essência dessa teoria dialogue muito bem com o conceito proposto por autores do marxismo digital de acumulação primitiva de dados, conforme elaborado por Lippold e Faustino (2022), assim como por Couldry e Mejias (2019), qual seja, a ideia de que a economia de dados hoje se estrutura analogamente ao que Marx identifica como início do capitalismo: a acumulação inicial dos bens produtivos na forma de propriedade individualizada.

Para construir esse diálogo, buscarei costurar a ideia central da Teoria da Dependência com um princípio da teoria da informação, a Desigualdade de Processamento de Dados, que descreve uma propriedade fundamental das funções de processamento de dados (Beaudry e Renner, 2012). Em termos simples, esse princípio afirma que qualquer processo de transformação de dados não pode aumentar a quantidade de informação

sobre uma variável que está sendo medida ou transmitida. Em outras palavras, a informação sobre uma variável X não pode aumentar ao passar por uma função de processamento de dados.

Esse princípio pode ser usado para pensar uma propriedade fundamental das tecnologias de IA: a produção do “novo” depende, necessariamente, do input de novos dados. Aliando isso à dependência geopolítica, pretendo propor uma interpretação para o atual avanço colonial dessas tecnologias no Sul Global como sendo um desdobramento da necessidade de novos dados para a acumulação primitiva do capital.

## **ESTABELECENDO O PRINCÍPIO DA DESIGUALDADE NO PROCESSAMENTO DE DADOS**

O primeiro passo dessa pesquisa é estabelecer o que é a Desigualdade no Processamento de Dados. O termo será utilizado a partir do livro sobre teoria da informação de Cover e Thomas (2006), com a adição do trabalho mais recente de Beaudry e Renner (2012). Esses últimos oferecem uma prova intuitiva para o fenômeno que será mobilizada conceitualmente para analisar, a título de caso exemplificativo, as tecnologias de IA.

Para Cover e Thomas (2006), a Desigualdade no Processamento de Dados é um conceito fundamental dentro da teoria da informação, mais especificamente no contexto de cadeias de Markov e sistemas de comunicação. Esse princípio afirma que, em qualquer sistema de processamento de dados, a quantidade de informação disponível no *output* (resultado) de um sistema não pode exceder a quantidade de informação que estava disponível no *input* (entrada).

Em outras palavras, quando os dados passam por um processo de transformação, seja por filtragem, compressão ou qualquer outro tipo de manipulação, a informação contida no dado processado nunca pode aumentar em relação à informação original, e na maioria das vezes, ocorre uma redução. Cover e Thomas (2006) vão justamente chamar atenção para o elemento entrópico do problema da perda de informação, levantando que a entropia processual do fluxo de informações não pode ser completamente contida.

Nesse sentido, Dantas (2016) aponta que em Anthony Wilden e suas conceituações sobre a categoria “informação” há uma explicação entrópica, na forma da aceitação de toda informação como “variedade”. Nesse sentido, como a informação é, em última análise, um subconjunto do conjunto variedades, há sempre uma relação de materialidade: informação é matéria, energia, em que se condensa uma possível configuração do conjunto variedade. Assim, aplica-se o mesmo princípio entrópico que para todo fluxo de energia.

É a partir dessa ideia de uma entropia informacional que Beaudry e Renner (2012) oferecem uma prova conceitual para essa desigualdade. Os autores vão apontar que a

Desigualdade no Processamento de Dados deriva diretamente da definição matemática de entropia quando aplicada ao processamento da informação por qualquer meio que seja.

Quando aplicamos o conceito de entropia ao princípio da Desigualdade no Tratamento de Dados, isso significa que, ao transformar ou processar dados, a entropia—isto é, a incerteza ou perda de informação—não pode ser revertida. O processamento de dados pode introduzir redundâncias ou eliminar partes da informação original, o que leva a uma redução da entropia relacionada à variável de saída. Isso reflete o princípio fundamental de que a transformação dos dados nunca pode aumentar a quantidade de certeza ou informação sobre a fonte original.

Na prática, o Princípio da Desigualdade no Tratamento de Dados assegura que qualquer perda de informação durante o processamento é irreversível, o que está intimamente ligado ao conceito de entropia como uma medida de incerteza. Em sistemas de comunicação, por exemplo, à medida que dados são codificados, transmitidos e decodificados, há uma tendência de redução da entropia da informação original devido a ruídos e limitações nos canais de transmissão, ilustrando como a desigualdade no tratamento de dados opera dentro dos limites impostos pela entropia.

### **Consequências para as tecnologias de Inteligência Artificial**

Pensando nas tecnologias de processamento de dados em larga escala—*big data*—temos um princípio conceitual com consequências bastante concretas. Como aponta Morozov (2018), uma vez que se estabelece a eficiência desse processamento de dados como forma de analisar e manipular indivíduos em escala populacional, vimos um crescimento exponencial do uso dessas tecnologias e do seu desenvolvimento técnico.

Em teoria, ao analisar grandes volumes de dados, espera-se extrair insights valiosos, mas a Desigualdade de Processamento de Dados aponta necessariamente que o processamento desses dados tem limites inerentes em termos da quantidade de informação que pode ser extraída sem perda de conteúdo informacional. O princípio coloca em xeque a capacidade prometida com os algoritmos de *machine learning*—aprendizado de máquina—e seu papel no avanço da IA, por exemplo.

Ludemir (2021) aponta que existe uma grande aposta do campo do “progresso tecnológico” nessas tecnologias como forma de ultrapassar certas barreiras técnicas que seguram o avanço computacional. Sua revisão do campo de IA e aprendizado de máquina fornece um excelente panorama do aumento expressivo do uso cotidiano e profissional dessas tecnologias por atores que não são especialistas técnicos, uma marca geral dos processos de digitalização e sua democratização de acesso e uso.

O que autores como Sichman (2021) levantam, porém, são questões sobre a eticidade desse uso e as possíveis dependências que estão sendo estruturadas, além de aspectos

como o problema da coletivização dos riscos em troca da individualização dos lucros. Morozov (2018) vai apontar uma tendência clara de uso dessas tecnologias para manipulação social coletiva, em casos como eleições, e individual, como nos casos de marketing direcionado.

Um dos riscos, porém, é o de uma espécie de esperança excessiva, ou talvez um excesso de confiança. A IA é vendida como uma tecnologia revolucionária cujo uso solucionaria uma série de problemas de forma definitiva (LUDEMIR, 2021), o que, pensando a partir do Princípio da Desigualdade no Processamento de Dados, não parece ser o caso. Levando a sério o problema colocado por Clover e Thomas (2006), o uso das tecnologias de IA deve ser direcionado de forma a abarcar sua limitação criativa, tratando-as mais como formas de fazer uma reprodução qualificada e alterada de produções anteriores.

O avanço das tecnologias de inteligência artificial (IA) e análise preditiva é igualmente afetado pela DPI. Em sistemas que utilizam modelos preditivos, como os usados em recomendação de produtos, diagnóstico médico ou sistemas de reconhecimento facial, o desempenho desses sistemas é limitado pela quantidade de informação contida nos dados originais. Qualquer erro ou imperfeição nos dados de entrada inevitavelmente se propaga ao longo do processo de processamento, o que pode comprometer a precisão e a confiabilidade dos resultados.

O Princípio da Desigualdade no Processamento de Dados também implica que a qualidade das saídas geradas por sistemas de IA, incluindo previsões e classificações, é limitada pela qualidade e quantidade de informações disponíveis nos dados de entrada. **Isso significa que, independentemente da sofisticação de um algoritmo de IA, ele está inerentemente limitado pelas características dos dados que processa.** Um sistema de IA não pode, por exemplo, inferir informações com mais precisão do que a permitida pela quantidade de dados relevantes disponíveis na sua base de treinamento. Isso se torna especialmente problemático quando os dados de entrada são incompletos ou enviesados.

Para os chamados “Modelos de Linguagem em Larga Escala” (*Large Language Models* – LLMs), especificamente, a situação pode ser ainda mais preocupante. A Desigualdade no Processamento de Dados implica que a capacidade dos LLMs de gerar respostas precisas ou coerentes está limitada pela qualidade e pela completude dos dados nos quais esses modelos foram treinados. Os LLMs são treinados em vastos volumes de texto, mas a disparidade na qualidade e no viés dos dados pode resultar em respostas enviesadas, incompletas ou imprecisas. Uma vez que o processo de treinamento envolve o uso de algoritmos que buscam padrões e relações nos dados, qualquer falha ou deficiência nos dados de origem se propaga para as saídas do modelo, limitando sua utilidade e confiabilidade.

Durante o treinamento de LLMs, também é comum o uso de técnicas de redução de dimensionalidade, que comprimem a vastidão dos dados textuais em vetores de representação mais compactos. Embora essas técnicas sejam necessárias para tornar o processamento viável em termos computacionais, elas resultam em perda de

informações. Isso ocorre porque os dados brutos contêm nuances e detalhes que não são totalmente preservados nas representações compactas, limitando a profundidade e precisão das respostas geradas pelos modelos.

Devido à natureza iterativa e hierárquica do processamento de informações nos LLMs, o princípio da desigualdade no processamento de dados indica que tarefas mais complexas—que requerem compreensão profunda ou raciocínio sofisticado—podem ser particularmente sensíveis à perda de informações ao longo das camadas de processamento. Por exemplo, um LLM pode ter dificuldades em gerar inferências altamente precisas ou raciocinar sobre contextos específicos se os dados originais não fornecerem informações suficientes ou se essas informações forem perdidas durante o treinamento.

## ACUMULAÇÃO PRIMITIVA DE DADOS

Para que passemos à segunda parte da pesquisa, farei uma conceituação sobre a acumulação primitiva de dados. Esse conceito, baseado na ideia marxiana de acumulação primitiva de capital, foi desenvolvido por Couldry e Mejias (2019) e Zuboff (2019), no exterior, e Lippold e Faustino (2022) no Brasil, em trabalhos dialogando com ideias de colonialismo digital, colonialismo de dados ou imperialismo, e que será a principal base teórica para a análise conceitual mais à frente.

A acumulação primitiva de capital é um conceito desenvolvido por Marx (2013), no capítulo XXIV d'O Capital, e faz referência ao processo histórico que precedeu e possibilitou a consolidação do capitalismo como sistema econômico dominante, especialmente na Europa. Para Marx, a acumulação primitiva não é a acumulação de capital como ocorre dentro do próprio capitalismo (através do trabalho assalariado, extração de mais-valia e reinvestimento do lucro), mas sim uma fase anterior e violenta que cria as condições necessárias para a emergência do modo de produção capitalista.

Esse processo foi e é marcado por expropriação forçada de terras, concentração de riquezas em poucas mãos e criação de uma classe de trabalhadores livres, mas despossuídos de meios de produção e, portanto, forçados a vender sua força de trabalho. A perspectiva que entende essa acumulação como um processo contínuo é muito defendida dentro das teorias marxistas, como vemos em Brandão (2010).

Em Marx, existe uma relação muito clara desse processo com a tomada de terras. A privatização das terras comunais, como os cercamentos na Inglaterra, foi parte do que forçou os trabalhadores rurais a migrarem para as cidades e vender sua força de trabalho como mercadoria, mas também representou uma mudança radical na lógica de costumes e cultura que resulta na propriedade privada, ponto muito defendido por Thompson (2005).

Outro ponto importante é que a acumulação primitiva também envolveu a intervenção direta do Estado, através de um conjunto de leis e violências, tanto locais quanto globais, para consolidar a apropriação privada dos meios de produção (GONÇALVES,

2017). A estrutura do Estado Moderno desempenhou um papel central em garantir o direito à propriedade privada e em reprimir as resistências populares que surgiram contra o processo de expropriação.

Vale ressaltar que a acumulação primitiva não se limitou ao contexto interno das sociedades europeias. Ela está intimamente relacionada com o imperialismo e a expansão colonial, que, para Marx (2013), foi um dos principais mecanismos pelos quais o capitalismo emergente conseguiu se consolidar e expandir. No processo imperialista, as potências europeias expropriaram terras, recursos e a força de trabalho de sociedades fora da Europa, especialmente nas Américas, África e Ásia. O colonialismo, o saque de recursos e o tráfico de escravos forneceram uma quantidade imensa de riqueza que foi fundamental para alimentar o capital inicial da burguesia europeia (GONÇALVES, 2017).

A colonização de vastos territórios fora da Europa foi um meio crucial de expropriação e transferência de riquezas. Terras indígenas foram apropriadas, recursos naturais foram extraídos em grande escala, e sistemas econômicos locais foram destruídos para dar lugar à produção voltada para o mercado europeu. A exploração das colônias forneceu não apenas matérias-primas essenciais, mas também novos mercados para os produtos manufaturados nas metrópoles. Essa dinâmica, segundo Brandão (2010), não se encerra, e marca todo o desenvolvimento econômico do Brasil, atravessando império, república, ditadura.

O conceito de acumulação primitiva de dados, elaborado aqui por Lippold e Faustino (2022), é construído em diálogo com essa tradição marxista, e propõe uma abordagem importantíssima para os estudos sobre colonialismo digital, costurando a releitura que essa última elaboração teórica faz do movimento colonial com uma releitura para o conceito marxista. No campo do digital, a acumulação primitiva de dados refere-se ao processo inicial e frequentemente violento, ou pelo menos coercitivo, pelo qual grandes corporações e Estados extraem, apropriam-se e concentram dados provenientes de indivíduos, comunidades e nações, muitas vezes sem o consentimento ou sem uma justa compensação para os sujeitos de onde esses dados são originados (COULDRY, MEJIAS, 2019).

O conceito dialoga também com a ideia de uma sociedade cuja economia é baseada cada vez mais nas informações como mercadoria de troca, cujo valor agregado se deve ao seu uso justamente em processos de manipulação como os narrados por Morozov (2018). Bolaño e Vieira (2014) vão tratar isso como parte da economia política dos ambientes digitais, identificando que a informação migra de um fator de produção na direção de algo mais próximo de uma commodity. Esse movimento reconfigura o fluxo de informações, agora como um fluxo de mercadorias negociáveis por si só, de forma relativamente independente da sua “função”.

Esse processo está modelado no trabalho de Farboodi e Veldkamp (2021). Nesse modelo de mercado de informações, os autores identificam que há de fato um processo de conversão de informação em mercadoria, inicialmente de altíssimo valor, o que estrutura uma nova prática econômica, e que posteriormente passa por um ciclo

de desvalorização e normalização, se integrando aos outros mercados produtivos. Esse último ponto se conecta com o argumento central desta pesquisa, que é a caracterização da configuração das relações Norte-Sul no ambiente digital como uma de apropriação de matéria prima.

Nesse contexto, a acumulação primitiva de dados é o mecanismo através do qual a burguesia, proprietária das grandes plataformas digitais (SNIRCEK, 2016) e corporações multinacionais de tecnologia expropriam e concentram informações pessoais e coletivas, transformando esses dados em uma mercadoria fundamental para a economia digital contemporânea.

Esse mecanismo se efetiva de várias formas, mas tanto Lippold e Faustino (2021) quanto Couldry e Mejias (2019) e Zuboff (2019) chamam atenção para a interseção entre esse fenômeno e o colonialismo como interpretado pela tradição marxista. O ponto central da discussão é revelar, através dessa ferramenta conceitual, a estrutura por trás dos movimentos do campo capitalista digital em relação ao Sul Global. Desse movimento, destacam-se algumas questões que surgem em outros trabalhos sobre colonialismo em sua configuração digital, especialmente Silveira (2021) e Avelino (2021), que dialogam entre os conceitos de colonialismo de dados e colonialismo digital, mas também Kwet (2019) e Seto (2023), que trazem importantes considerações sobre o conceito de imperialismo.

Como optei por trabalhar com a ideia de colonialismo de dados, é importante analisar a diferença desse conceito em relação ao colonialismo digital. Ferreira (2021) aponta que quando falamos de um colonialismo de dados, apontamos para a dinâmica específica de coleta, extração e uso dos dados e sua transformação em informações commodificadas. Silveira (2021) trabalha com a distinção entre isso e o colonialismo digital, que trabalha com o controle sobre a infraestrutura digital que permite a existência de um ambiente digital concreto. Avelino (2021) usa o colonialismo digital exatamente nesse sentido, trabalhando com plataformas digitais enquanto agentes coloniais a partir do seu uso geopolítico das infraestruturas digitais.

Em ambos os casos, temos conceitos que dialogam com a posição marxista que relaciona o movimento colonial e o sistema econômico capitalista. Em comum, ambos tratam a atual estrutura de economia política digital de que falam Bolaño e Vieira (2014) como um desdobramento da configuração própria do capitalismo, em que países do Sul Global fornecem mão de obra e matéria prima em condições desiguais de força, e são mantidos nessa posição por um sistema de estruturas e superestruturas dedicado à manutenção dessa desigualdade de forças (AVELINO, 2021).

Silveira (2021), assim como Zuboff (2019) com seu capitalismo de vigilância, associam esse movimento com uma nova configuração sistêmica na figura do neoliberalismo. Assim, há uma reorganização do colonialismo em uma nova configuração, agora baseado na dinâmica de plataformas (SNIRCEK, 2016), manipulação em larga escala através de dados (MOROZOV, 2018) etc. O colonialismo na sua forma digitalizada, assim, envolveria de um lado o colonialismo digital, que garante desigualdade na

infraestrutura, e o colonialismo de dados, que configura a extração dessa nova commodity.

Kwet (2019) e Seto (2023) interpretam esse fenômeno dentro da lente do imperialismo. O primeiro é um exemplo de trabalhos que mobilizam a crítica feita à atuação norte-americana ao longo da segunda metade do séc. XX e seu caráter imperial, e trazem para o ambiente digital a ideia de que através da economia política das empresas multinacionais, países do Norte Global atuam na forma de um imperialismo velado.

Seto (2023), já no contexto brasileiro, traz a ideia de um subimperialismo de dados, colocando a configuração identificada pelo colonialismo de dados como parte desse imperialismo associado ao neoliberalismo, ou seja, a relação de controle do Sul pelo Norte em função da desigualdade econômica e produtiva, que permite um controle geopolítico.

São posições interessantes para realçar o caráter de violência identificado por diversos autores como sendo parte integral da economia política digital. Uma vez que, como vemos em Farboodi e Veldkamp (2021), a informação enquanto uma configuração materializada da variedade (DANTAS, 2016) integra a economia política do capitalismo, o ambiente digital passa a se sujeitar às mesmas configurações estruturais do próprio capitalismo (FAUSTINO, LIPPOLD, 2023).

Neto (2020) compara, por exemplo, a extração de dados pelas plataformas com o processo de mineração e extração de matéria prima. O autor trabalha com a ideia de que o processo de coleta de dados tem fundamentalmente a mesma montagem que a mineração, em que empresas extraem o que querem enquanto matéria prima do solo, destroem um ecossistema determinado, e abandonam rejeitos e dejetos. Para Neto (2020), não somos a matéria prima, e sim o rejeito, o que sobra após a extração dos dados. O processo que Morozov (2018) identifica como morte da política, para Neto, é a destruição do ecossistema.

Dentro do conceito de colonialismo temos uma relação de dependência informacional, em que países e comunidades no Sul Global se tornam fornecedores de dados brutos para as grandes corporações sediadas nos países desenvolvidos. Essas corporações, por sua vez, têm o poder de processar, armazenar e monetizar esses dados, muitas vezes sem compartilhar os benefícios econômicos ou tecnológicos com os países e indivíduos que forneceram as informações.

Schiav (2020) já nos aponta uma direção similar à desta pesquisa. Essa relação de dependência que os autores da Teoria da Dependência identificam na dinâmica colonial e imperialista do capitalismo se reproduz dentro da dinâmica do colonialismo de dados. Há uma gama de países periféricos cuja função na economia política é a de produção primária, o que reflete a mudança econômica que o Brasil inicia na década de 90 de desindustrialização e reprimarização da economia produtiva. Lippold e Faustino (2021) vão configurar essa ideia enquanto acumulação primitiva de dados pelo Norte Global.

A acumulação primitiva de dados ocorre através da coleta em larga escala de informações pessoais, muitas vezes sem o conhecimento ou consentimento informado dos indivíduos (ZUBOFF, 2019). Tello (2023) aponta que plataformas digitais, redes sociais, aplicativos de celulares e dispositivos conectados capturam uma vasta gama de dados, desde comportamentos online e preferências de consumo até localização geográfica, contatos e interações pessoais. As empresas de tecnologia, então, concentram esses dados para análise e monetização, sem contrapartida.

Assim como a acumulação primitiva de capital levou à concentração de riqueza em poucas mãos, a acumulação primitiva de dados leva à concentração de poder informacional em um número restrito de grandes corporações. Essas empresas *big techs* controlam uma infraestrutura global de dados, que abrange desde as plataformas em que os dados são coletados até as redes que possibilitam a circulação dessas informações. Elas detêm o poder de decidir como esses dados serão utilizados, explorados e vendidos, criando assim uma forma de monopólio informacional (MOROZOV, 2018).

Outro aspecto central da acumulação primitiva de dados é a apropriação e privatização de formas de conhecimento coletivo e cultural. Em seu livro sobre o tema, Lippold e Faustino (2023) mostram que dados gerados por comunidades inteiras, como informações sobre práticas culturais, linguísticas, ecológicas ou até genéticas de populações indígenas, são frequentemente capturados e transformados em mercadoria sem qualquer retorno ou benefício para essas comunidades. Esse processo reflete as dinâmicas de expropriação e colonização imperialistas, em que os recursos culturais e informacionais das periferias globais são apropriados pelas potências tecnológicas do Norte Global.

Essa dinâmica tem grandes implicações para o debate sobre soberania digital. Como mostra Ávila Pinto (2018), existe uma interseção tensionada entre as ideias de colonialismo e soberania digitais, especialmente em razão da própria natureza do ambiente digital de hiperconexão e globalidade. O problema maior é que se a soberania digital, sob a perspectiva marxista, pode ser entendida como a capacidade de um Estado ou sociedade de controlar e regular seu próprio espaço digital, em oposição à submissão às grandes corporações e potências hegemônicas que dominam a infraestrutura, os dados e as tecnologias digitais, então esse conceito está intrinsecamente ligado à luta pela autodeterminação e pela resistência contra a dominação econômica, política e cultural imposta pelo Norte Global.

Sob o prisma marxista de autores como Lippold e Faustino (2022), a falta de soberania digital reflete uma continuidade das relações de dependência e exploração que marcaram as dinâmicas pós-coloniais e imperialistas, agora reproduzidas no contexto das tecnologias digitais. A questão central reside no controle dos meios de produção, e no mundo digital, esses meios são representados pelas infraestruturas tecnológicas, como servidores, plataformas digitais, redes de comunicação e, mais recentemente, as redes de inteligência artificial.

Dentro da perspectiva do colonialismo digital como apresentada por Avelino (2021), esse controle se dá de tal forma que coloca o Sul Global numa posição muito clara dentro da divisão internacional do trabalho digital. Grohmann (2021) aponta que a sociologia empírica das relações trabalhistas não deixa dúvida que o Sul fornece mão de obra barata para as empresas do Norte. O que vemos dentro do colonialismo de dados, por outro lado, é algo que Neto (2020) também reconhece: além de mão de obra, fornecemos a matéria prima de dados.

Assim, a falta de soberania digital reflete, também, uma forma contemporânea de expropriação, onde o controle dos recursos digitais é monopolizado por atores externos que extraem valor econômico, em especial de países do Sul Global (JURI, YAMILA, 2023). A falta de controle sobre plataformas digitais e sua infraestrutura representa uma nova forma de alienação, em que países periféricos não possuem controle sobre os dados de suas populações, tornando-se dependentes de tecnologias estrangeiras e sujeitando-se a regimes de extração de valor que beneficiam majoritariamente as economias do Norte.

Além desse elemento, dialogando com a hegemonia cultural gramsciana, temos o problema da tensão entre soberania e hegemonia digitais (GUIMARÃES, FERREIRA JUNIOR, 2023). A hegemonia cultural é mantida através da padronização de plataformas digitais que disseminam conteúdos globais baseados nos valores e práticas do Norte Global. Isso inclui a imposição de normas de consumo, entretenimento e comunicação que desconsideram as especificidades culturais e as formas tradicionais de sociabilidade dos países do Sul Global. Dessa forma, a soberania digital também está relacionada à capacidade de resistência cultural e à preservação da diversidade frente à uniformização promovida pelas grandes corporações tecnológicas.

Isso tudo aponta na direção de uma tensão clara entre as ideias de soberania digital e a prática do colonialismo digital. Há uma clara relação de violência estruturante nessas práticas digitais, que atravessa não só a coleta de dados e seu uso, mas a mobilização da desigualdade social como ferramenta para estruturar relações de trabalho digital violentas (GROHMANN, 2020). O que proponho na pesquisa agora é o uso do ferramental teórico da Teoria da Dependência para analisar os desdobramentos desse tipo de estrutura do capital.

## **A TEORIA DA DEPENDÊNCIA COMO FERRAMENTA ANALÍTICA PARA A SOBERANIA DIGITAL**

Para pensar a Teoria da Dependência, campo bastante difuso e plural, escolhi o trabalho de Dos Santos (2020) para fornecer um panorama geral e contextual, e o trabalho de Wasserman (2022) como fundamentação principal para elaborar conceitualmente a relação entre a Teoria da Dependência e a soberania digital no contexto do colonialismo de dados.

Pensando a Teoria da Dependência como uma abordagem das ciências sociais que busca explicar as dinâmicas de subdesenvolvimento e dependência econômica dos países periféricos em relação às nações centrais no sistema capitalista global, seu papel analítico em relação à acumulação primitiva de dados de Lippold e Faustino (2022) é, em linhas introdutórias, explicar a relação estruturada entre o Norte e o Sul Global nas relações de produção da economia digital como o resultado das relações desiguais de troca e exploração entre países desenvolvidos (centro) e subdesenvolvidos (periferia), em que os últimos são mantidos em uma condição de subordinação econômica, política e tecnológica.

A Teoria da Dependência como um corpo teórico consolidado foi desenvolvida e ampliada nas décadas de 1960 e 1970. Os pensadores dessa “segunda onda” argumentavam que o subdesenvolvimento era mantido por uma estrutura de dependência, onde as elites dos países periféricos eram aliadas das economias centrais, funcionando como intermediárias para os interesses imperialistas (DOS SANTOS, 2020). Eles se distanciaram de teorias do desenvolvimento lineares e progressivas, entendendo que o sistema capitalista global é estruturado de forma a preservar essas desigualdades.

Intelectuais como o próprio Dos Santos (2020), Ruy Mauro Marini e Vânia Bambirra adotaram uma interpretação mais radical e marxista da dependência, argumentando que o desenvolvimento na periferia é sempre limitado pelas estruturas capitalistas globais, que exigem a manutenção de uma divisão internacional do trabalho. Esses autores viam a periferia do capitalismo como forçada a fornecer matérias-primas e trabalho barato, o que perpetuava a exploração e a extração de valor pelas nações centrais. Defendiam que a superação da dependência só poderia ocorrer através de uma ruptura revolucionária com o capitalismo.

A Teoria da Dependência tem um impacto grande no pensamento e na prática de economia política no Brasil, especialmente na virada dos anos 90 para 2000 (WASSERMAN, 2022), especialmente considerando que Fernando Henrique Cardoso estava ligado aos debates do campo e que o período é marcado por alterações bruscas no papel econômico brasileiro. Esse paradigma teórico, porém, passa a ser questionado em função dos casos de integração econômica global de países periféricos que saíram da posição subalterna— a Coreia do Sul sendo um exemplo clássico. Obras como *Parasita*, filme de 2019, colocam em xeque essa alegada superação da Dependência.

O ponto mais central de um framework de análise baseado na Teoria da Dependência, porém, é sua relação com os debates de soberania digital. Essa relação surge de uma análise crítica das dinâmicas de poder, controle e subordinação presentes na economia digital global. As dependências no âmbito digital que o conceito de soberania digital revela refletem a lógica estrutural de exploração e dominação que a teoria original identificava nas trocas comerciais e no controle sobre os meios de produção industrial.

Assim como na teoria original, a dependência digital que marca o desenvolvimento do século XXI é uma forma de dependência estrutural, em que as nações da periferia

digital global não têm acesso aos meios de infraestrutura ou conhecimento técnico que lhes permitiriam uma maior autonomia tecnológica. A expropriação de dados dos países periféricos também reflete as dinâmicas de subordinação que a Teoria da Dependência descreveu no campo econômico (WASSERMAN, 2022). As nações periféricas fornecem o insumo (dados), mas não controlam o processo produtivo (o uso desses dados para a criação de valor).

O que acredito ser mais interessante nessa interseção entre soberania digital, dependência e colonialismo de dados é o resultado da inserção do Princípio da Desigualdade no Processamento de Dados na equação. Isso porque, como apontam Clover e Thomas (2006), a essência dessa iniquidade entrópica é justamente que para a produção de qualquer informação além da contida nos dados originais necessariamente tem que haver a inserção de novos dados no processamento.

Ou seja, há um claro limite conceitual ao poder de produção das tecnologias de IA em função da estagnação de dados. A partir de um certo momento, como conclui-se a partir de Beaudry e Renner (2012), qualquer outro processo de transformação adicionado na cadeia de processamento faz com que a entropia informacional gere perdas sensíveis demais. Isso é similar ao resultado de Shumailov, et al. (2024), cujas pesquisas apontam um risco elevado de colapso dos sistemas de IA, especialmente LLMs e geradores de imagem, se alimentados com inputs gerados por si mesmo ou outras IAs.

Isso significa que os sistemas de IA entram em um feedback loop em que a entropia informacional atinge níveis sensíveis demais, e o conteúdo resultante se torna, aos nossos olhos, quebrado. Com esse risco aparecendo na dinâmica das IAs, parece se confirmar o que a Desigualdade no Processamento de Dados implica conceitualmente: que para a continuação do desenvolvimento dessas tecnologias, é obrigatória a entrada de dados realmente novos, e não só reprocessamentos de dados existentes.

O ponto, portanto, é que o Sul Global representa hoje uma espécie de atualização digital daquilo que representava durante o processo colonial e imperialista, qual seja, a fronteira dos insumos necessários. Como vamos com Assis (2014), a passagem da dinâmica colonial para uma dinâmica de “colonialidade”, nos termos do autor, envolve uma mudança na qualidade dessas relações, mas uma continuidade de elementos característicos, especialmente na desigualdade produtiva.

Ainda que Assis (2014) entenda que mesmo a Teoria da Dependência precise de algum grau de atualização ao explicar as novas relações de terras, Schiav (2020) e Avelino (2021), por exemplo, entendem que ao falarmos de um colonialismo de dados estamos justamente trazendo esses elementos de continuidade. São esses elementos que levam Kwet (2019) a buscar o conceito de imperialismo, assim como Seto (2023) o de subimperialismo, para qualificar a economia política do digital em função da dinâmica desigual de produção e extração de matéria prima.

Assim, há uma montagem estrutural e econômica em que lugares como a América Latina oferecem ao capital digital nova matéria prima necessária, o que explica a

reestruturação das relações de dependência explicadas pela Teoria da Dependência (DOS SANTOS, 2020) em uma atualização para uma dependência digital, em diálogo com a dinâmica de um colonialismo de dados.

### **Sul Global: latifúndio de dados e trabalho precarizado**

Se por um lado, temos como consequência da necessidade de novos dados gerada pela Desigualdade no Processamento de Dados esse movimento neocolonial digital que estrutura as relações de dependência digital, por outro temos o reaparecimento de outra dinâmica muito própria do modelo de mundo imperialista e colonial, e que a Teoria da Dependência também ajuda a explicar: a criação das relações de trabalho baseadas em trabalho precarizado e violência.

Como mostra Dos Santos (2020), um dos temas frequentes desse ferramental de análise era a forma com que a estrutura de mundo capitalista forçava o Sul Global a funcionar com base em um trabalho precarizado e semi-servil, fornecendo mão de obra barata no contexto globalizado. Os trabalhos de Grohman e Araújo (2021) confirmam exatamente essa hipótese, inclusive analisando especificamente as tecnologias de IA. Na sua pesquisa, encontraram uma dinâmica muito característica da economia colonial sendo reproduzida no treinamento de máquinas.

Como o autor aponta em trabalho anterior (GROHNMANN, 2020), a racionalidade neoliberal que domina o campo da economia digital reproduz, através da plataformação do trabalho, a relação de trabalho do mundo colonial. Quando somamos isso à essa ideia de que o Sul Global passa a ser uma espécie de latifúndio extrativista de dados, o que temos é a atualização digital da dinâmica econômica colonial descrita pela Teoria da Dependência.

## **CONCLUSÕES**

Assim, temos uma relação clara entre as consequências práticas da Desigualdade no Processamento de Dados, o colonialismo de dados e a acumulação primitiva, e a questão da soberania digital, tudo isso mediado pelas relações de dependência já analisadas conceitualmente pela Teoria da Dependência.

O Princípio da Desigualdade no Processamento De Dados limita a capacidade dos sistemas de IA de extrapolar ou melhorar significativamente a qualidade da informação processada em relação aos dados de entrada. Para tecnologias de IA em geral, isso implica uma necessidade contínua de atenção à qualidade dos dados, à propagação de erros e à capacidade de processamento. Se temos uma economia baseada nessas tecnologias, e uma economia marcada também pela necessidade de um grande fluxo informacional, criamos a mistura de onde surge a acumulação primitiva de dados.

Esse conceito de Lippold e Faustino (2022) é excelente em revelar a consequência da necessidade material de dados, que é o atual movimento de expropriação e

concentração de informações que define a economia digital contemporânea, replicando as dinâmicas de exploração da acumulação primitiva de capital. No contexto do colonialismo de dados, esse processo perpetua a exploração desigual entre o Norte e o Sul globais, em que grandes corporações e potências tecnológicas se apropriam de dados provenientes das periferias globais sem redistribuir os benefícios econômicos ou tecnológicos.

Essa dinâmica ajuda a explicar o atual dilema dos Estados-Nação em relação à soberania digital, cujo equilíbrio com as necessidades de inserção em mercados digitais globalizados é delicado. Na perspectiva marxista, a soberania digital é uma questão de controle dos meios de produção e dos recursos digitais, que se encontra profundamente imbricada nas relações de poder entre o Norte e o Sul Global. O enfraquecimento da soberania digital é um requisito básico do movimento do capital digital de acumulação primitiva de dados, mas também da construção das relações de dependência digital que marcam o avanço da digitalização.

A Teoria da Dependência, originalmente formulada para analisar a relação entre países centrais e periféricos no contexto do sistema capitalista internacional, oferece uma lente poderosa para compreender as relações de poder no campo digital. O conceito de soberania digital surge também como uma resposta à crescente dependência dos países periféricos em relação às potências tecnológicas globais, refletindo essas antigas dinâmicas de expropriação e exploração econômica.

Assim, é possível pensarmos em uma espécie de atualização digital para a Teoria da Dependência, que explique as relações de dependência digital a partir, dentre outras coisas, da acumulação primitiva de dados fruto da Desigualdade no Processamento de Dados. Esse movimento de acumulação primitiva marca a dinâmica de colonização digital que transforma os países periféricos em latifúndios informacionais e obriga relações de trabalho digital precarizado, largamente baseadas nas diferenças de capacidade técnica e na mobilização da desigualdade social e econômica da periferia.

A luta pela soberania digital, assim como a luta pela superação da dependência econômica, é uma batalha pela autonomia, pelo controle sobre os recursos (neste caso, dados) e pela capacidade de moldar um futuro tecnológico que não reproduza as desigualdades e assimetrias do passado. É parte integrante da luta por uma nova forma de organização econômica e social que resista à hegemonia capitalista e construa alternativas de poder popular e controle democrático sobre as tecnologias digitais.

## REFERÊNCIAS

ASSIS, Wendell Ficher Teixeira. 2014. Do colonialismo à colonialidade: Expropriação territorial na periferia do capitalismo. Cadernos CRH, 27(72). <https://doi.org/10.1590/S0103-49792014000300011>

AVELINO, Rodolfo. 2021. Colonialismo digital: Dimensões da colonialidade nas grandes plataformas. In J. F. Cassino, J. Souza, & S. A. da Silveira (Orgs.), *Colonialismo de dados: Como opera a trincheira algorítmica na guerra neoliberal*. Autonomia Literária.

ÁVILA PINTO, Renata. 2018. Soberanía Digital o Colonialismo Digital? Nuevas tensiones alrededor de la privacidad, la seguridad y las políticas nacionales. *Sur - Revista Internacional de Derechos Humanos*, v. 15, n. 27, p. 15-28.

BEAUDRY, Normand J.; RENNER, Renato. 2012. An intuitive proof of the data processing inequality. *Quantum Inf. Comput.* Vol. 12, No. 5&6, 0432-0441.

BOLAÑO, Cesar Ricardo Siqueira, VIEIRA, Eloy. 2014. Economia política da internet e os sites de redes sociais. *Revista Eletrônica Internacional de Economia Política da Informação, da Comunicação e da Cultura*, 16(2), 71–84. Disponível em <https://periodicos.ufs.br/eptic/article/view/2168>.

BRANDÃO, Carlos. 2010. Acumulação primitiva permanente e desenvolvimento capitalista no Brasil contemporâneo. In: ALMEIDA, Alfredo Wagner B.; ACSERARD, Henri et al. (orgs.). 2010. *Capitalismo globalizado e recursos territoriais: fronteiras da acumulação no Brasil contemporâneo*. Rio de Janeiro: Lamparina.

COULDRY, Nick., & MEJIAS, Ulisses. A. 2019. *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.

COVER, Thomas M.; THOMAS, Joy A. 2006. *Elements of Information Theory*. 2. ed. Hoboken, N.J.: Wiley-Interscience.

DANTAS, Marcos. 2016. Dialética da informação: Uma leitura epistemológica no pensamento de Vieira Pinto e Anthony Wilden (II) | *Dialectics of Information: An epistemological approach to Vieira Pinto and Anthony Wilden (II)*. *Liinc Em Revista*, 12(1). <https://doi.org/10.18617/liinc.v12i1.892>

DOS SANTOS, Theotonio. 2020. *Teoria da Dependência: Balanço e Perspectivas*. Florianópolis: Insular Livros.

FARBOODI, Maryam, VELDKAMP, Laura. 2021. A model of the data economy (NBER Working Paper No. 28427). National Bureau of Economic Research. <https://doi.org/10.3386/w28427>

FAUSTINO, Deivison; LIPPOLD, Walter. 2023. *Colonialismo digital: por uma crítica hacker-fanoniana*. 1. ed. São Paulo: Boitempo.

FERREIRA, Sandro Rodrigo da Silva. 2021 O que é (ou o que estamos chamando de) ‘Colonialismo de Dados’?. *PAULUS: Revista De Comunicação Da FAPCOM*, 5(10). <https://doi.org/10.31657/rcp.v5i10.458>

GONÇALVES, Guilherme Leite. 2017. Acumulação primitiva, expropriação e violência jurídica: expandindo as fronteiras da sociologia crítica do direito. *Revista Direito e Práxis*, Rio de Janeiro, v. 8, n. 2, p. 686-724, jun. Disponível em: <https://doi.org/10.12957/dep.2017.28770>.

GROHMANN, Rafael. 2020. Plataformização do trabalho: entre dataficação, financeirização e racionalidade neoliberal. *Revista Eletrônica Internacional de Economia Política da Informação da Comunicação e da Cultura*, São Cristovão, v. 22, n. 1, p. 106–122.

GROHMANN, Rafael; ARAÚJO, Willian Fernandes. 2021. O chão de fábrica (brasileiro) da inteligência artificial: a produção de dados e o papel da comunicação entre trabalhadores de Appen e Lionbridge. *Palavra Chave*, v. 24, n. 3, e2438, jul./set.

GUIMARÃES, Larissa Maria de Almeida; FERREIRA JÚNIOR, Amarildo. 2023. Cultura digital, políticas culturais e soberania de dados no Brasil. In: MANZANO, Marcelo (org.). 2023. *Novas e velhas desigualdades na era digital: dimensões da fronteira*. São Paulo: Hucitec Editora, Fundação Perseu Abramo, Instituto Lula.

JURI, Eliana; YAMILA. 2023. La gobernanza de los datos de la soberanía territorial a la soberanía digital. *Quaestio Iuris (QI)*, v. 16, n. 2, p. 802. DOI: <https://doi.org/10.12957/rqi.2023.72434>.

KWET, Michael. 2019. Digital colonialism: US empire and the new imperialism in the Global South. *Race & Class*, 60(4), 3-26. <https://doi.org/10.1177/0306396818823172>

LIPPOLD, Walter.; FAUSTINO, Deivison. 2022. Colonialismo digital, racismo e acumulação primitiva de dados. *Germinal: marxismo e educação em debate*, [S. l.], v. 14, n. 2, p. 56–78. Disponível em: <https://periodicos.ufba.br/index.php/revistagerminal/article/view/49760>. Acesso em: 31 jul. 2024.

LUDERMIR, Teresa Bernarda. 2021. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. *Estudos Avançados*, v. 35, n. 101, p. 133-150, jan.-abr.

MARX, Karl. 2013. *O capital: livro I*. Tradução de Rubens Enderle. São Paulo: Boitempo.

MOROZOV, Evgueny. 2018. *Big Tech: a ascensão dos dados e a morte da política*. Trad. Claudio Marcondes. São Paulo: Ubu.

NETO, Moyses Pinto. 2020. Nuvem: Plataforma: Extração. *Revista Percursos*, v. 21, n. 45, p. 05-23.

SCHIAV, Iara Franco. 2020. O colonialismo de dados sob a ótica da teoria da dependência: As especificidades e desafios dos países periféricos no novo contexto sociotécnico. Trabalho apresentado no 20º Congresso Brasileiro de Sociologia, UFPA, Belém, PA. GT 15 - Sociologia Digital.

SETO, Kenzo Soares. 2023. Subimperialismo de dados: uma crítica ao colonialismo de dados diante das Big Techs sul-americanas. *Revista Eletrônica Internacional De Economia Política Da Informação Da Comunicação E Da Cultura*, 25(2), 165–184. <https://doi.org/10.54786/revistaepitic.v25i2.19199>

SHUMAILOV, Iliia; SHUMAYLOV, Zakhar; ZHAO, Yiren; et al. 2024. AI models collapse when trained on recursively generated data. *Nature*, v. 631, p. 755–759. Disponível em: <https://doi.org/10.1038/s41586-024-07566-y>.

SICHMAN, Jaime Simão. 2021. Inteligência Artificial e sociedade: avanços e riscos. *Estudos Avançados*, v. 35, n. 101, jan.-abr, p. 47-62.

SILVEIRA, Sergio Amadeu da. 2021. A hipótese do colonialismo de dados e o neoliberalismo. In J. F. Cassino, J. Souza, & S. A. da Silveira (Orgs.), *Colonialismo de dados: Como opera a trincheira algorítmica na guerra neoliberal*. Autonomia Literária.

SRNICEK, Nick. 2016. *Platform Capitalism*. Polity Press.

TELLO, Andrés. 2023. Sobre o colonialismo digital: Dados, algoritmos e colonialidade tecnológica do poder no sul global. *Inmediac. Comum*, vol.18, n.2, pp.89-110.

THOMPSON, Edward Palmer. 2005. *Costumes em comum: estudos sobre a cultura popular tradicional*. São Paulo: Companhia das Letras.

WASSERMAN, Claudia. 2022. *A Teoria da Dependência*. 1. ed. Rio de Janeiro: FGV Editora.

ZUBOFF, Shoshana. 2019. *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.