

PLANO DE GESTÃO DE DADOS (PGD)

1. Coleção de dados

Tipo, formato e volume quanto dos dados coletados:

Mapa Bibliográfico de Gramsci no Brasil - pesquisa realizada em 2016 (1ª ed.), 2018 (2ª ed.) e 2019 (3ª ed.) pela equipe do Núcleo de Estudos e Pesquisas em Filosofia, Política e Educação (NuFiPE/UFF). Trata-se do levantamento bibliográfico das obras que, no Brasil, discorrem acerca de Antonio Gramsci como referencial teórico e/ou a respeito de sua biografia.

O material é divulgado em formato textual, em um arquivo de tipo PDF, disponibilizado online na página da International Gramsci Society-Brasil¹ e compondo um volume de aproximadamente 2 mb.

Quanto aos dados a serem criados:

A base coletada em formato PDF será transformada em um dado estruturado (em formatos CSV, JSON-LD e RDF), seguindo os princípios FAIR, e disponibilizada em repositório com identificador persistente (DOI) e controle de versionamento (utilizando o repositório do projeto no Zenodo associado ao GitHub), contendo marcação com metadados de origem e proveniência de modo a assegurar informações sobre o ciclo de vida dos dados de pesquisa e a relação com a organização e a comunidade responsável pelo conteúdo (COX et al., 2021). Além disso, serão adicionados metadados descritivos sobre o conteúdo das publicações para utilização no grafo de conhecimento (*knowledge graph*) a ser produzido.

Formatos e software para compartilhamento e acesso de longo prazo aos dados:

¹ Disponível em: <https://storage.googleapis.com/production-hostgator-brasil-v1-0-9/739/971739/23GwhSwA/93b61e567faa4c7d906c5568c73acb77?fileName=IGS-Mapa-Bibliogr%C3%A1fico-de-Gramsci-no-Brasil-Site-Setembro-de-2019-1.pdf>

Os dados coletados dispõem de um endereço (URL), mas ainda não contam com identificadores persistentes, controle de versionamento e metadados de proveniência e autoridade acionáveis por máquinas, como preconizado pelas boas práticas para dados na web, editadas como recomendação pela W3C (2017) e pelos princípios de dados FAIR, conforme indicado por Wilkinson et. al. (2016).

Em razão dessas lacunas, a pesquisa pretende transformar esses dados adotando um modelo de publicação compatível com os princípios FAIR, tanto para os dados e quanto para os metadados a serem trabalhados.

Os dados e as transformações realizadas serão organizados por meio de um script em Python e divulgados em um Caderno Jupyter nos repositórios do projeto (Zenodo e GitHub), facilitando a reprodutibilidade da pesquisa e o acesso de longo prazo.

Sobre a reutilização dos dados:

Os dados coletados podem ser reutilizados uma vez que foram publicados pela IGS-Brasil a qual tem, dentre suas finalidades estabelecidas em Estatuto (2015, p. 02), “[...] intercambiar informações relacionadas às suas pesquisas, aos seus projetos editoriais, às suas particulares perspectivas analíticas e suas atividades em geral dentro do universo do pensamento gramsciano” e “envidar esforços para garantir o acesso do público brasileiro à bibliografia gramsciana publicada em âmbito nacional e internacional”.

Os dados criados poderão ser reutilizados conforme a licença estabelecida para todos os conteúdos produzidos no âmbito desse projeto, a saber: [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/deed.pt_BR) (https://creativecommons.org/licenses/by/4.0/deed.pt_BR).

Obtenção e organização dos dados coletados:

Os dados serão coletados e criados utilizando um script Python a ser disponibilizado por meio de um Caderno Jupyter, facilitando a reprodutibilidade da pesquisa e a reutilização dos dados. Será realizado download do arquivo PDF da URL <https://storage.googleapis.com/production-hostgator-brasil-v1-0-9/739/971739/23GwhSwA/93b61e567faa4c7d906c5568c73acb77?fileName=IGS-Mapa-Bibliogr%C3%A1fico-de-Gramsci-no-Brasil-Site-Setembro-de-2019-1.pdf>, e se seguirão as transformações necessárias para tornar os dados textuais brutos em dados estruturados e marcados com os metadados adequados aos princípios FAIR.

As pastas e arquivos serão nomeadas em formato de *slug* (texto em letras minúsculas separados por hífen), visando a fácil previsibilidade de URL/URI nos repositórios do projeto no GitHub e no Zenodo. Os arquivos dentro dos repositórios serão separados em pastas inicialmente da seguinte forma: arquivos do projeto (projeto), dados coletados (dados-coletados), Caderno Jupyter contendo os scripts em Python (caderno-jupyter), dados criados (dados-criados) e publicações (publicacoes). A estrutura prévia do repositório está disponível em: <https://github.com/emiliefaedo/gramsci-kg>.

O controle de versão será realizado por meio dos repositórios do GitHub e do Zenodo que já estão associados e dispõem de sistemas robustos de versionamento. Como conceito associado ao controle de versão será utilizado o modelo de versionamento semântico (Semantic Versioning 2.0.0 - <https://semver.org/>) para definir as versões dos arquivos.

O processo de garantia de qualidade, além da revisão por pares, especialistas na área, contará com a verificação de tipos de dados em Python, realizada por meio da biblioteca em Pydantic, a qual permite a validação dos dados de entrada e fornece um conjunto de validadores predefinidos para vários tipos de dados, incluindo strings, números e datas. Além de possibilitar a definição de validadores personalizados usando instrução interna. O processo também será documentado no Caderno Jupyter do projeto visando a possibilidade de reutilização em trabalhos derivados.

2. Documentação e metadados

Será utilizado como padrão para os metadados um subconjunto dos vocabulários da Dublin Core (<http://purl.org/dc/terms/>), adequado à descrição de cada recurso bibliográfico disponibilizado no arquivo coletado. No desenvolvimento do trabalho, caso seja identificada a necessidade de metadados adicionais, será dada preferência aos conjuntos de metadados disponibilizados como recomendações pela W3C e aos identificadores publicados na Wikidata ([Wikidata identifiers](#)), no intuito de facilitar a compreensão e o reuso dos dados em novas pesquisas.

Os metadados serão divulgados dentro dos objetos digitais gerados durante o projeto com a base de dados tratada para a elaboração do grafo de conhecimento da produção bibliográfica sobre o pensamento gramsciano no Brasil, em formato CSV, JSON-LD e RDF, com identificadores persistentes e metadados de proveniência e autoridade, alinhados com os princípios FAIR.

Destaca-se que os vocabulários de metadados do Dublin Core, as recomendações da

W3C (PROV-O, SKOS, RDF, JSON-LD e outros) e os identificadores da Wikidata contam com vasta utilização e documentação por uma comunidade internacional e especializada na criação e curadoria de descritores, o que tende a facilitar com que os dados criados no âmbito deste projeto sejam lidos e interpretados no futuro.

3. Ética e conformidade legal

Os dados coletados estão publicados em acesso público, portanto, não haverá a necessidade de proteger a identidade de participantes, nem serão utilizados dados confidenciais. Os dados são de autoria da equipe do Núcleo de Estudos e Pesquisas em Filosofia, Política e Educação (NuFiPE/UFF) e estão publicados pela IGS-Brasil.

Quanto aos direitos autorais, os dados criados serão disponibilizados com licença que permite seu reuso para qualquer fim, desde que citada a fonte, uma vez que será utilizado para licenciamento de todos os conteúdos produzidos ao longo da pesquisa a licença [Creative Commons Atribuição 4.0 Internacional \(CC BY 4.0\)](#).

Esta licença permite que o trabalho seja copiado e redistribuído em qualquer suporte ou formato. Também permite que o trabalho seja adaptado, re combinado e transformado a partir do material produzido, para qualquer fim, mesmo que comercial. Tais características de licenciamento são compatíveis com os princípios de ciência aberta.

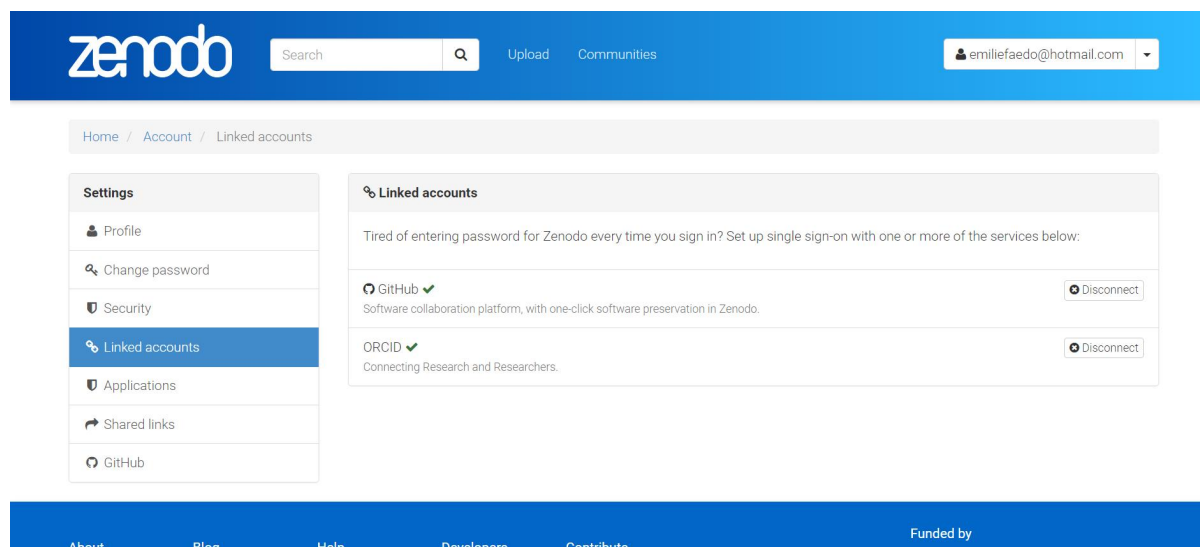
Não haverá período de embargo e restrição para os dados criados, uma vez que o conteúdo será disponibilizado, de forma versionada, à medida em que for sendo produzido.

Quanto aos direitos de propriedade intelectual, cumprindo os itens do edital, foi realizada pesquisa de busca com os termos da pesquisa nas seguintes bases de propriedade intelectual: Instituto Nacional da Propriedade Industrial (INPI); LATIPAT-Espacenet; WIPO - Search International and Nacional Patent Collections; USPTO - United States Patent and Trademark Office. O resultado das buscas consta como anexo no projeto de pesquisa e não infringem direitos de propriedade intelectual.

4. Armazenamento e backup

Uma vez que os dados da pesquisa serão coletados, tratados e criados a partir de um script Python usando um Caderno Jupyter, seu armazenamento e backup serão produzidos a cada ciclo de utilização, tendo os dados tratados e o script registrados diretamente no repositório do GitHub. As entregas parciais serão publicadas como um *release* no GitHub e

espelhadas automaticamente no repositório do Zenodo, como já configurado (imagem a seguir):



Além disso, será também gerada uma cópia local no laptop de trabalho, atualizada a cada nova modificação na árvore de arquivos do repositório do GitHub, usando a ferramenta de sincronização GitHub Desktop.

A responsável pelo backup e recuperação dos dados será a pesquisadora. Em caso de incidentes, os dados podem ser recuperados na máquina local ou no repositório do GitHub, bem como novamente produzidos por meio do script Python adequado à reprodutibilidade da pesquisa.

Quanto ao gerenciamento do acesso e segurança, reitera-se que não serão utilizados dados confidenciais. Os riscos para a segurança dos dados serão gerenciados por meio de login e senha na máquina de trabalho local e nos repositórios do GitHub e do Zenodo. Não haverá coleta de dados em campo e a pesquisa não conta com colaboradores, sendo de responsabilidade exclusiva da pesquisadora.

5. Seleção e Preservação

Além de serem depositados no repositório LattesData (CNPq), todos os dados criados serão preservados nos repositórios GitHub e Zenodo, conforme citado anteriormente, e serão mantidos enquanto estes repositórios estiverem disponíveis.

A previsibilidade de reuso dos dados criados se dá em diversas áreas do conhecimento, especialmente em pesquisas no campo das Ciências Humanas e Sociais que se dediquem a estudar o pensamento gramsciano no Brasil. Os códigos gerados no script em Python também

poderão ser aproveitados de modo integral ou em partes, tanto para a realização de trabalhos similares com outras bases bibliográficas, quanto para o tratamento e validação dos dados para a publicação de outros grafos de conhecimento.

Os repositórios utilizados são mantidos por organizações comerciais (GitHub - Microsoft) e destinadas à pesquisa (Zenodo - CERN; LattesData - CNPq), sendo repositórios de livre acesso, sem custos para os usuários pesquisadores.

Além dos repositórios gerenciados diretamente pela pesquisadora, os dados e metadados postados no LattesData poderão contar com plano de preservação de longo prazo, a critério do CNPq, visando garantir a curadoria dos dados para além do período de execução e divulgação do projeto de pesquisa.

6. Compartilhamento de dados

Os dados serão disponibilizados ao longo da execução do projeto, à medida em que forem sendo trabalhados. O registro do DOI será atribuído sempre que necessário e também em cada *release* publicado no Zenodo, garantindo um identificador único e persistente para cada arquivo ou conjunto de arquivos do projeto. Destaca-se que o Zenodo dispõe de um recurso de atribuição de DOI, com controle de versão.

Por exemplo, o presente Plano de Gestão de Dados (PGD), já está publicado no Zenodo com o DOI da primeira versão do arquivo de trabalho (10.5281/zenodo.8256065, versão 0.1.0a) disponibilizado no endereço: <https://doi.org/10.5281/zenodo.8256065>, bem como as versões subsequentes.

No caso de arquivos que recebam DOI ou outros identificadores persistentes de origens diversas, o Zenodo permite indicar a referência externa no momento do upload, mantendo a relação de identificação persistente do recurso, alinhado às “Boas Práticas para Dados na Web” (W3C, 2017).

Tendo em visto que os dados coletados são públicos e que a pesquisa utilizará a licença CC BY 4.0 para todos os conteúdos produzidos, não haverá restrição de acesso no compartilhamento dos dados.

7. Responsabilidades e Recursos

A pesquisadora é a responsável exclusiva pela gestão dos dados da pesquisa e sua divulgação.

Detalhes do projeto de pesquisa

Título: O microcosmo dos subalternos: elementos para pensar a subalternidade brasileira.

Resumo: As elaborações contemporâneas do pensamento de Antonio Gramsci, considerado teórico clássico do marxismo, fornecem indubitável contributo à compreensão crítica de questões sociopolíticas hodiernas relacionadas, por exemplo, à produção e reprodução da dominação hegemônica de classes, bem como da subalternidade em suas diversas e complexas dimensões. No que tange à presente proposta de pesquisa, o que se propõe é avançar na compreensão filológica do par categorial subalternidade/hegemonia. Mais especificamente, estudar o desenvolvimento dos *critérios metodológicos* acerca das experiências e práticas da subalternidade produzida pela dominação hegemônica. E, a partir disso, tematizar os desenvolvimentos gramscianos em torno da subalternidade, em uma abordagem interseccional da dominação de classes, raça e gênero no Brasil. Parte-se da hipótese de que a subalternidade é produzida pela dominação hegemônica e, por isso, a história dos subalternos é episódica, por vezes desagregada, mas, apesar disso, há movimentações e “rupturas moleculares” que não podem ser desconsideradas, pelo contrário, devem ser estudadas, divulgadas, depuradas e ampliadas. Apresenta-se, desse modo, um conjunto de questões preliminares que norteiam o problema de pesquisa: a noção de subalternidade, desenvolvida por Gramsci, configura-se como um novo conceito de classe? Quem são os subalternos na obra gramsciana? E no Brasil? Quem são e como estudar os subalternos em tematização com a realidade brasileira? E para quê? Metodologicamente, o que se propõe é uma reconstrução filológica do termo, e, em um exercício de tradutibilidade, avançar na investigação das particularidades da subalternidade brasileira, em seus contornos de país de capitalismo dependente e na singularidade de seus grupos sociais subalternos, inseridos na luta hegemônica. De modo a organizar a representação da problemática apontada, será produzido um grafo de conhecimento (*knowledge graph*), como um recurso educacional aberto, das publicações sobre o pensamento gramsciano no Brasil, com destaque para as categorias subalternidade e hegemonia.

Palavras-chave: Antonio Gramsci, subalternidade, hegemonia, *knowledge graph*.

Instituição executora: Universidade Federal Fluminense.

Vigência: 01 de janeiro de 2024 a 31 de dezembro de 2024.

Recursos financeiros: recursos financeiros disponibilizados integralmente pelo CNPq.

Chamada e número do processo: Chamada Pública MCTI/CNPq nº 14/2023 - Apoio a Projetos Internacionais de Pesquisa Científica, Tecnológica e de Inovação.

Área do conhecimento: Teoria Política

Hipótese de pesquisa/objetivo dos dados: A hipótese da pesquisa, bem como seu objetivo, considera que seja possível tematizar a noção gramsciana de subalternidade em uma abordagem interseccional da dominação de classe, raça e gênero no Brasil, especialmente no que concerne às possibilidades metodológicas de estudo dos subalternos.

Autoria: Emilie Faedo Della Giustina (pesquisadora responsável)

E-mail: emiliefaedo@hotmail.com

ID Lattes: 8900673299422814 / <http://lattes.cnpq.br/8900673299422814>

ORC ID: <https://orcid.org/0000-0003-1966-9989>

Instituição responsável pelos dados: Universidade Federal Fluminense - UFF.

Telefone de contato: +55 (42) 99959-8366

Importante: O proponente declara que deu ciência, a todos os envolvidos na geração ou coleta de dados, do compromisso de depósito dos dados de pesquisa no repositório LattesData e que todos estão de acordo.

REFERÊNCIAS

COX, SJD; GONZALEZ-BELTRAN, AN; MAGAGNA B; MARINESCU M-C. **Ten simple rules for making a vocabulary FAIR**. PLoS Comput Biol 17(6): e1009041, 2021.

<https://doi.org/10.1371/journal.pcbi.1009041>

INTERNATIONAL GRAMSCI SOCIETY. **Estatuto da International Gramsci Society - Brasil**. 2015. Disponível em <https://storage.googleapis.com/production-hostgator-brasil-v1-0-9/739/971739/23GwhSwA/8c29ddaa51774557afaeab3af18f4732?fileName=Estatuto%20IGS.pdf>. Acesso em 17 de agosto de 2023.

WILKINSON, M; DUMONTIER, M; AALBERSBERG, I. et al. **The FAIR Guiding Principles for scientific data management and stewardship**. Sci Data 3, 160018, 2016.

<https://www.nature.com/articles/sdata201618>

W3C. **Data on the Web Best Practices**. 2017. Disponível em:

<https://www.w3.org/Translations/DWBP-pt-BR/>. Acesso em 17 de agosto de 2023.