
Plan Overview

A Data Management Plan created using DMPTool

Title: Impacto social da UFSCar: análise a partir dos egressos

Creator: Marisa Cubas Lozano - **ORCID:** [0000-0002-0888-9145](https://orcid.org/0000-0002-0888-9145)

Affiliation: Universidade Federal de São Carlos (ufscar.br)

DMP ID: <https://doi.org/10.48321/D1W31F>

Funder: Universidade Federal de São Carlos (ufscar.br)

Template: Digital Curation Centre

Project abstract:

A fim de entender como a UFSCar impacta na sociedade foram levantados dados públicos de egressos no Currículo Lattes. Para viabilizar a pesquisa no período de doutorado, foi realizado um recorte e recuperadas as informações de egressos do curso de Biblioteconomia e Ciência da Informação. Os dados do Currículo Lattes foram recuperados de forma automatizada a partir da ferramenta LattesMachina e tratados com o software proprietário VantagePoint e script da Linguagem de Programação Python.

Start date: 12-01-2020

End date: 06-02-2023

Last modified: 05-24-2022

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Impacto social da UFSCar: análise a partir dos egressos

Data Collection

What data will you collect or create?

Foram coletados dados dos Currículos Lattes dos egressos do curso de graduação de Biblioteconomia e Ciência da Informação da Universidade Federal de São Carlos. Para o tratamento dos dados foram usados o software pago Vantage Point e um script em linguagem de programação Python (<https://github.com/Leluke/python-xml-parser/tree/master>).

How will the data be collected or created?

Foi realizada a extração de forma automatizada do Currículos Lattes a partir da ferramenta LattesMachina* em 08 abr. 2021. São recuperados dados como formação acadêmica, atividades profissionais e acadêmicas (como participação em cursos e eventos), e produção intelectual, artística e técnica, em formato .xml.

*

MATIAS, Mesailde Souza de Oliveira. **Base referencial para o povoamento de repositórios institucionais: coleta automatizada de metadados da Plataforma Lattes**. 2015. Dissertação (Mestrado em Gestão de Organizações e Sistemas Públicos) - Universidade Federal de São Carlos, São Carlos, SP, 2015. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/6932>. Acesso em: 16 ago. 2021.

MATIAS, Mesailde Souza De Oliveira; AMARAL, Roniberto Morato Do; MATIAS, Paulo. Proxy customizado para acesso ao web service da Plataforma Lattes. *In*: WORKSHOP DE TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO DAS INSTITUIÇÕES FEDERAIS DE ENSINO SUPERIOR, 11., 2017, Salvador. *Anais* [...]. Salvador: UFBA, 2017. Disponível em: <http://www.xiwitcifes.ufba.br/modulos/submissao/Upload-353/86143.pdf>. Acesso em: 17 ago. 2021.

Documentation and Metadata

What documentation and metadata will accompany the data?

Os dados já são disponibilizados com os metadados.

Para leitura dos dados pelo software Vantage Point é necessário uso de um filtro no formato .config que será depositado junto com os dados (arquivo 20220511config-importacao-xml-synclattes.config).

Ethics and Legal Compliance

How will you manage any ethical issues?

Os dados utilizados estão disponibilizados de forma pública na Plataforma Lattes.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

Os dados estão disponíveis de forma pública na Plataforma Lattes (<https://lattes.cnpq.br/>) e aqueles que foram coletados para essa pesquisa serão disponibilizados em repositório digital sob uma licença Creative Commons.

Storage and Backup

How will the data be stored and backed up during the research?

Durante a realização da pesquisa, os dados ficarão armazenados no computador pessoal da pesquisadora e um servidor de uso exclusivo do Grupo de Pesquisa NIT/Materiais UFSCar.

How will you manage access and security?

O acesso aos dados está restrito à pesquisadora e ao orientador da pesquisa.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

Os dados coletados e aqueles gerados após o tratamento serão preservados a longo prazo:
- arquivos .xml dos Currículos Lattes dos egressos (xml-egressos-bci.rar),
- planilha .csv gerada a partir do script do Python (teste.csv),
- filtro para leitura dos arquivos .xml no *software* Vantage Point (20220511config-importacao-xml-synclattes.config)

What is the long-term preservation plan for the dataset?

Os dados coletados e os tratados serão disponibilizados em repositórios digitais de acesso público.

Data Sharing

How will you share the data?

Os dados serão disponibilizados em repositórios digitais de acesso público e divulgados em publicações científicas, como artigos e trabalhos apresentados em eventos.

Are any restrictions on data sharing required?

Os dados não têm restrição de acesso.

Responsibilities and Resources

Who will be responsible for data management?

Durante a pesquisa, a pesquisadora e o orientador da pesquisa serão os únicos responsáveis pelo gerenciamento dos dados.

What resources will you require to deliver your plan?

Foi necessário acesso ao LattesMachina, limitado a pesquisadores vinculados ao NIT/Materiais UFSCar devido o convênio realizado com o CNPq, para obtenção dos dados. O acesso ao servidor para armazenamento, onde o software Vantage Point está instalado, foi possibilitado pela participação da pesquisadora no Grupo de Pesquisa NIT/Materiais UFSCar.

Planned Research Outputs

Text - "Dados dos egressos - Lattes"

São 534 arquivos .xml referente aos Currículos Lattes dos egressos do curso de Biblioteconomia e Ciência da Informação da UFSCar

Text - "Filtro para leitura do dados "

Filtro criado para facilitar a interpretação dos .xml obtidos dos Currículos Lattes pelo software Vantage Point.

Text - "dados sobre atuação profissional"

Documento .csv gerado a partir do script em Linguagem Python com dados sobre a atuação profissional extraídos dos arquivos .xml extraídos dos Currículos Lattes dos egressos de Biblioteconomia e Ciência da Informação da UFSCar.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Dados dos egressos - Lattes	Text	2022-05-21	Open	Zenodo	25 MB	Creative Commons Zero v1.0 Universal	None specified	No	Yes
Filtro para leitura do dados	Text	2022-05-21	Open	Zenodo	1 MB	Creative Commons Zero v1.0 Universal	None specified	No	No
dados sobre atuação profissional	Text	2022-05-21	Open	Zenodo	1 MB	Creative Commons Zero v1.0 Universal	None specified	No	Yes