

## Plan Overview

---

*A Data Management Plan created using DMPTool-Stage*

**DMP ID:** <https://doi.org/10.48321/D1RC90>

**Title:** Presença das universidades públicas estaduais paulistas na produção científica e tecnológica da UFABC, UFSCar e UNIFESP

**Creator:** Denilson de Oliveira Sarvo - **ORCID:** [0000-0002-4374-6430](https://orcid.org/0000-0002-4374-6430)

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

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

**Template:** Digital Curation Centre

### **Project abstract:**

Esta pesquisa investiga a presença das universidades públicas estaduais paulistas na produção científica e tecnológica das universidades federais localizadas no Estado de São Paulo: Universidade Federal do ABC (UFABC), Universidade Federal de São Carlos (UFSCar) e Universidade Federal de São Paulo (UNIFESP). A pesquisa é apoiada em um estudo bibliométrico, investigando registros da produção científica, publicados no período de 2017-2020, informados nos Currículo Lattes de docentes que atuam na pós-graduação das universidades públicas federais paulistas. Foram obtidos como resultados um conjunto de registros da produção científica e tecnológica, obtidos por meio da ferramenta SyncLattes; um filtro de importação de arquivos no formato .XML para o software VantagePoint e indicadores bibliométricos e de rede sobre a presença das universidades públicas estaduais paulistas na produção das universidades federais investigadas.

**Start date:** 03-06-2019

**End date:** 03-06-2023

**Last modified:** 08-07-2023

### **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

---

## **Presença das universidades públicas estaduais paulistas na produção científica e tecnológica da UFABC, UFSCar e UNIFESP**

Serão coletados dados provenientes do Currículo Lattes dos docentes que atuam na pós-graduação da UFABC, UFSCar e UNIFESP.

Os dados serão extraídos a partir da ferramenta SyncLattes, retornando como resultados da pesquisa registros da produção científica (formato .CSV) e as páginas com os metadados dos Currículos Lattes (formato .XML).

Os dados serão importados e tratados no software VantagePoint, sendo gerados bases de informação com os registros importados (formato .VPT). Software de análise de rede serão empregados, sendo obtidos arquivos nas extensões correspondentes (.GDF, .Gephi etc).

Os dados serão coletados a partir da relação de ID Lattes dos docentes que atuam na pós-graduação da UFABC, UFSCar e UNIFESP, requisito para a extração de dados a partir da ferramenta SyncLattes.

Os dados coletados já são disponibilizados com os metadados. O uso da ferramenta SyncLattes é tratado em publicação já disponível. A ferramenta VantagePoint apresenta manual de uso. Para a importação dos dados foi necessário o uso de um filtro de importação, que serão disponibilizados junto com os dados em formato .config e .csv. O processo de tratamento e análise de dados estão previstos em publicações sobre a pesquisa.

Serão divulgados dados e informações que estão disponibilizadas em fontes públicas, não sendo publicados dados críticos e sensíveis.

Os dados serão obtidos por meio de fontes públicas, não sendo apresentados questões legais quanto ao uso e disponibilização. Para o depósito dos dados é prevista a aplicação de licenças de uso, como as Creative Commons.

Durante a realização da pesquisa, os dados ficarão armazenados em servidores de uso exclusivo do grupo NIT/Materiais da UFSCar e armazenado em serviços de armazenameto em nuvem.

O acesso aos dados de pesquisas estão restritos a pesquisadores vinculados no grupo de pesquisa NIT/Materiais da UFSCar.

Serão preservados a longo prazo os dados brutos e os resultantes das etapas de tratamento e enriquecimento de dados, utilizados para a obtenção dos indicadores propostos, com valor agregado e que apoiam as publicações relacionadas à pesquisa.

Os dados serão preservados em armazenamento local, no âmbito do grupo de estudo NIT/Materiais, e os dados relacionados à publicações serão disponibilizados em Repositórios de Dados públicos.

Os dados serão disponibilizados em repositórios digitais de acesso público

Não há restrições de acesso.

O gerenciamento dos dados será feito pelo pesquisador responsável pela pesquisa e elaboração do PGD.

Foram utilizados softwares para extração de dados: a SyncLattes, ferramenta com acesso conrolado aos pesquisadores vinculados ao grupo do NIT/Materiais da UFSCar; uso do software proprietário VantagePoint, possível a partir do vínculo do pesquisador no grupo NIT/Materiais da UFSCar; e uso de outros softwares de acesso

aberto, voltados à tratamento e análise de dados, como os softwares de análise de redes. Foi utilizado recursos de armazenamento locais e em nuvem para a gestão do ciclo de vida dos dados.

---

## Planned Research Outputs

### Text - "Filtro para a importação de arquivos XML do Currículo Lattes para o VantagePoint"

O *VantagePoint* utiliza para a importação de registros um arquivo de configuração denominado *Importer Filter*, que consiste no filtro de importação no qual é delimitado a estrutura do registo a ser interpretado pelo *software*. Para o desenvolvimento desta pesquisa, um filtro de importação para os arquivos XML provenientes do Currículo Lattes foi elaborado, considerando o documento XSD disponibilizado pelo CNPq, que contém os campos e definições do *XML Schema* utilizados pelo Currículo Lattes. A partir da análise e seleção de metadados, foram previstas para a importação 15 categorias de informações que contém 72 campos de metadados do Currículo Lattes, os campos que compõem o filtro de importação são apresentadas na documentação que acompanha o dataset.

---

### 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?
Filtro para a importação de arquivos XML do Curríc ...	Text	2022-05-18	Open	Zenodo	1 MB	Creative Commons Zero v1.0 Universal	None specified	No	No