

# Extração Automatizada de Microdados da Plataforma Nilo Peçanha como Suporte à Gestão da Rede Federal de Educação Profissional e Tecnológica

**Abdiel A. Gouveia Silva<sup>1</sup>, Khawan Silva<sup>1</sup>, Matheus Louro<sup>1</sup>, Nicolas Falcão<sup>1</sup>, Orlando P. S. Júnior<sup>1</sup>, Pedro Clarindo da S. Neto<sup>1</sup>, Tiago Lacerda<sup>1</sup>, Alceu A. Cardoso<sup>1</sup>**

<sup>1</sup> Instituto Federal de Mato Grosso – Campus Cuiabá Cel. Octayde Jorge da Silva R. Zulmira Canavarros, 95 – Centro – 78005-390 – Cuiabá – MT – Brasil

{abdiel.s, khawan.fellipe, matheus.louro, nicolas.f}@estudante.ifmt.edu.br, {orlando.junior, pedro.neto, tiago.lacerda, alceu.aparecido}@ifmt.edu.br

**Abstract.** *The Brazilian Law number 14,129/2021 has spurred digital transformation within public administration, promoting efficiency and integration. Within the Federal Network of Professional, Scientific, and Technological Education (EPCT), the Nilo Peçanha Platform (PNP) serves as the official repository for strategic educational microdata. However, technical and methodological challenges hinder the extraction and full use of this information. To address this scenario, a web extractor module is proposed to automate the filtering and organization of data. This initiative aims to improve accessibility and provide educational managers with better support for informed decision-making.*

**Resumo.** *A Lei nº 14.129/2021 fortaleceu a transformação digital na administração pública, incentivando eficiência e integração. Na Rede Federal de Educação Profissional, Científica e Tecnológica, a Plataforma Nilo Peçanha centraliza microdados educacionais estratégicos. Contudo, dificuldades técnicas e metodológicas limitam a extração e o uso dessas informações. Para superar esse cenário, propõe-se um módulo extrator web que automatiza a filtragem e organização dos dados. A iniciativa visa facilitar o acesso às informações e apoiar decisões mais qualificadas dos gestores educacionais.*

## 1. Introdução

A promulgação da Lei nº 14.129/2021 [BRASIL 2020], marco regulatório do Governo Digital no Brasil, estabeleceu diretrizes para a eficiência, transparência e integração dos processos administrativos por meio da inovação tecnológica e da governança digital [Ulrich et al. 2021].

No contexto da educação profissional e tecnológica, a governança digital tem sido utilizada para aprimorar a gestão acadêmica e administrativa [Guimarães 2005], enquanto a implementação de tecnologias digitais visa otimizar processos e garantir a qualidade dos serviços educacionais [Campos et al. 2024].

Nesse cenário, a Plataforma Nilo Peçanha (PNP) consolidou-se como o sistema oficial para coleta, validação e disseminação dos microdados da Rede Federal, englobando informações detalhadas sobre matrículas, corpo docente, servidores e execução orçamentária das instituições [SETEC/MEC 2025]. Contudo, a extração e o tratamento desses dados apresentam desafios técnicos significativos, devido ao grande volume e à complexidade das informações disponíveis.

Para superar essas dificuldades, propõe-se o desenvolvimento de uma ferramenta *web* intuitiva que automatiza a extração, filtragem e organização dos microdados da Plataforma Nilo Peçanha, facilitando o acesso e o uso desses dados pelas unidades da Rede Federal de Educação Profissional e Tecnológica, com o objetivo de aprimorar a capacidade gerencial e o suporte à tomada de decisão dos gestores educacionais [Pires et al. 2024].

## 2. Metodologia

A análise de dados é uma estratégia fundamental para obter informações precisas e apoiar a tomada de decisões assertivas [Lages et al. 2025]. Esse processo consiste em transformar dados brutos em informações estratégicas, tornando-os essenciais para o processo decisório. Para alcançar esse objetivo, é necessário aplicar métodos de extração e transformação de dados, de modo a garantir sua utilização adequada [Casagrande and da Silva 2024].

No contexto deste trabalho, a PNP representa a fonte primária de microdados que foram extraídos, filtrados e tratados para a obtenção de informações específicas de uma unidade, possibilitando análises mais direcionadas. A PNP é um ambiente virtual voltado à coleta, validação e disseminação das estatísticas oficiais da Rede Federal EPCT. Seu principal objetivo é disponibilizar dados consolidados para o cálculo de indicadores de gestão [Ministério da Educação 2025].

Os dados coletados correspondem ao período de 2017 a 2024 e são disponibilizados em planilhas no formato CSV, contendo registros acadêmicos de diferentes unidades de ensino. Para a realização deste estudo, foi necessário aplicar processos de extração e transformação de dados referentes à unidade de ensino e incluir informações relacionadas aos departamentos de um determinado campus. Esse procedimento possibilitou o mapeamento dos cursos e sua vinculação aos respectivos departamentos, permitindo análises mais detalhadas da estrutura acadêmica local [de Arruda and Mesquita 2024].

Inicialmente, foram utilizados *softwares* de planilhas eletrônicas, como Microsoft Excel e LibreOffice Calc, para realizar a transformação dos dados, por meio dos recursos disponibilizados pelas próprias ferramentas. Entretanto, devido ao grande volume de informações, o processo tornou-se lento e pouco eficiente, evidenciando a necessidade de adotar métodos mais robustos para lidar com a alta quantidade de dados.

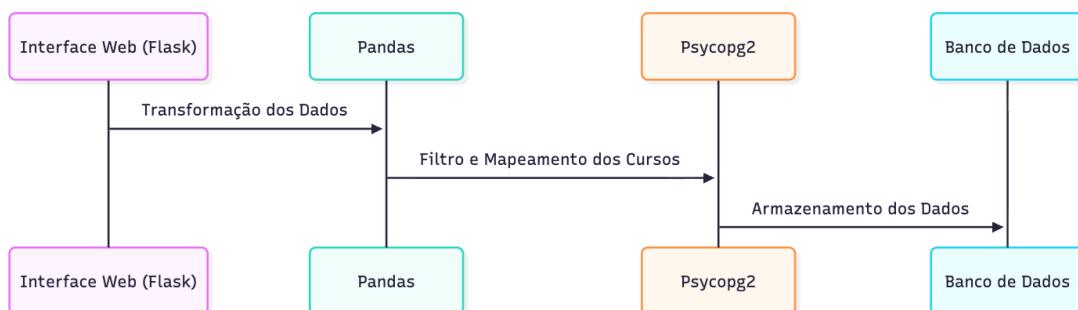
### 2.1. Módulo Extrator de Dados

O Módulo Extrator de Dados foi desenvolvido com o objetivo de automatizar o processo de extração e transformação dos dados provenientes da Plataforma Nilo Peçanha. Para sua implementação, foi utilizada a linguagem de programação *Python*, em conjunto com bibliotecas específicas como *Flask*, para a construção do servidor *web*; *Pandas*, para as operações de transformação e manipulação dos dados; e *Psycopg2*, para a integração e inserção das informações em banco de dados.

*Pandas* foi a biblioteca central para as etapas de manipulação, filtragem e limpeza dos microdados. A partir de seus recursos, foi possível aplicar filtros para registros de uma determinada unidade (campus), leitura dos arquivos CSV disponibilizados pela PNP, tratar inconsistências nos valores, padronizar categorias e criar colunas que possibilitaram análises detalhadas dos dados. Dessa forma, a biblioteca desempenhou papel essencial na transformação dos dados extraídos em um conjunto consolidado e pronto para análise.

O *Flask* foi utilizado como *framework* para estruturar o extrator em forma de aplicação *web* leve e modular, permitindo a criação de uma interface de fácil acesso e utilização. Na prática, o *Flask* funcionou como um servidor simples, fazendo a ponte entre o *upload* dos arquivos CSV e a execução das rotinas de limpeza e filtragem com *Pandas*, além da atualização do banco de dados por meio da *psycopg2*, biblioteca que possibilita a comunicação direta com o Sistema de Gerenciamento de Banco de Dados(SGBD). O processo é demonstrado na Figura 1.

Além da eficiência técnica, o módulo adotará práticas de segurança da informação alinhadas à Lei Geral de Proteção de Dados (Lei nº 13.709/2018), assegurando o tratamento ético e responsável das informações processadas. Considerando que os microdados da Plataforma Nilo Peçanha são públicos e anonimizados, o sistema buscará garantir conformidade com os princípios da LGPD mediante o uso de autenticação de usuários, gestão de permissões e acompanhamento das operações realizadas, de modo a preservar a integridade e a confiabilidade das informações.



**Figura 1. Fluxo de Dados para extração**

### 3. Resultados

Iniciativas semelhantes buscam aprimorar o acesso e a integração de dados educacionais. Entre essas iniciativas, destaca-se o *Extrator PNP* [Ministério da Educação 2025], ferramenta que oferece mecanismos de consulta e disseminação de microdados para fins de transparência e análise. Além disso, propostas como a de [RIBEIRO 2023], apresentam processos ETL aplicados a dados educacionais, reforçando a relevância de ferramentas de automação e análise de dados no contexto da Rede Federal.

O Módulo Extrator de Dados reduziu significativamente o tempo de processamento dos microdados da Plataforma Nilo Peçanha, passando de uma média de 30 a 40 minutos em procedimentos manuais para cerca de 5 minutos na versão automatizada. Essa plataforma representa maior eficiência, confiabilidade e precisão no tratamento das informações.

Foi desenvolvida uma interface gráfica simples, construída com *HTML*, *CSS* e *JavaScript*, para tornar a utilização do extrator mais acessível. Essa interface conta com uma seção que direciona o usuário ao site da PNP, facilitando o acesso e o download dos microdados, além de uma área dedicada à execução das rotinas de filtragem e tratamento implementadas. A tela inicial do extrator de dados é ilustrada na Figura 2. Módulo Extrator de Dados da PNP.



**Figura 2. Interface web do Extrator de Dados da PNP**

A plataforma *web* demonstrou-se de uso simples e acessível, permitindo o envio simultâneo de até três tabelas por meio de uma interface intuitiva. Com apenas um comando, os dados são filtrados, transformados e carregados no banco de dados, facilitando a utilização da ferramenta por diferentes tipos de usuários .

Além da agilidade e da usabilidade, o módulo garantiu padronização e consistência, incluindo a organização dos cursos por departamento da unidade. Essa estrutura possibilitou análises mais específicas e integração a ferramentas de *BI*(*Business Intelligence*), fortalecendo a produção de indicadores estratégicos para apoio à gestão.

#### 4. Conclusão

Este artigo detalhou o desenvolvimento e a implementação de uma interface *web* automatizada para extração de dados para a Plataforma Nilo Peçanha. Alinhada à Lei nº 14.129/2021 e aos princípios da governança digital, esta iniciativa visa modernizar a gestão educacional através da análise de dados, transformando microdados brutos em informações estratégicas para a tomada de decisões.

O Módulo Extrator de Dados automatiza a coleta, tratamento e filtragem de microdados da PNP, superando as limitações de ferramentas tradicionais. A ferramenta proporcionou trazer dados que podem gerar indicadores acadêmicos cruciais como Eficiência Acadêmica, Taxa de Evasão e Relação Aluno-Professor (RAP), e sua interface intuitiva

facilita a integração com sistemas de *BI*, fortalecendo a governança digital e promovendo uma gestão mais eficiente e transparente.

Os resultados demonstram o impacto positivo da solução na gestão acadêmica e administrativa. Como trabalhos futuros, propõe-se expandir o extrator para outras fontes de dados, aprimorar a interface de gestão para autonomia do usuário, incluir análise quantitativa de impacto nas rotinas administrativas e integrar novos indicadores, assegurando a continuidade e a possibilidade de adaptação do sistema em diferentes contextos.

## Referências

- BRASIL (2020). Decreto nº 10.332, de 28 de abril de 2020. institui a estratégia de governo digital para o período de 2020 a 2022.
- Campos, V. d. O., Oliveira, A. G. d., Casagrande, A. L., Faria, E. B. d., and Figueiredo, J. M. d. (2024). Formação para uso de ambiente de gestão escolar por meio da computação em nuvem. In *Anais do XXXII Seminário de Educação (SEMIEDU)*.
- Casagrande, V. M. and da Silva, T. P. (2024). Uma introdução à análise de dados usando pandas, matplotlib e seaborn. *SBC*, 1:8.
- de Arruda, L. I. and Mesquita, M. P. (2024). Ferramentas de baixo custo para integração e geração de dashboards de visualização de dados. *SBC*, 1:3.
- Guimarães, T. A. (2005). Governança eletrônica e administração pública: uma análise da literatura. *Revista de Administração Pública*, 39(5):1049–1068.
- Lages, S., da Silva Neto, P. C., de Faria, E. B., de Figueiredo, J. M., and Oliveira Confessor, A. O. (2025). Integrating an online management system into technical schools: A step toward digital transformation. *Conference on Digital Government Research*, 26.
- Ministério da Educação (2025). Pnp — plataforma nilo peçanha: observatório de dados e informações. <https://www.gov.br/mec/pt-br/pnp>. Acesso em: 11 set. 2025.
- Pires, S. P. S., Santos, G. F., and Oliveira, R. F. S. (2024). Impactos da governanÇa digital na educaÇÃo brasileira: Uma revisÃo de literatura no scielo. *Periódicos New Science Publ.*
- RIBEIRO, L. A. (2023). Proposta de processo etl para tratamento e integração de dados educacionais do censo escolar. *Dissertação de Mestrado em Computação Aplicada*, 92. 92 f. Dissertação (Mestrado em Computação Aplicada).
- SETEC/MEC (2025). Plataforma nilo peçanha (pnp). <https://www.gov.br/mec/pt-br/pnp>. Acesso em: 11 set. 2025.
- Ulrich, J., Maciel, C., and Cappelli, C. (2021). Políticas e estratégias para a transformação dos serviços públicos digitais nos governos federal e do estado de mato grosso. In *Anais do IX Workshop de Computação Aplicada em Governo Eletrônico (WCGE)*.