

Nome do Projeto Padronização e migração do inventário patrimonial do IFPE para o sistema SIADS

Objetivos do Projeto

Planejar e idealizar uma solução tecnológica e processual para padronizar, otimizar e integrar o inventário patrimonial do IFPE com o Sistema Integrado de Administração de Serviços (SIADS) do governo federal.

Descrição Detalhada do Produto do Projeto

Desenvolver um protótipo de solução tecnológica e metodológica que permita a padronização, digitalização e migração dos bens patrimoniais do IFPE do sistema SUAP e/ou outros sistemas de inventário patrimonial, para o SIADS, garantindo conformidade legal, transparência e eficiência nos processos de inventário.

Escopo da Migração

A migração será realizada em quatro fases, cada uma com objetivos, ferramentas e entregas definidas.

Fase 1 - Extração dos Dados

Objetivo

Converter arquivos PDF oriundos do levantamento de inventário em formato tabular (CSV), preservando o máximo possível da estrutura original.

Processo

- Utilização da ferramenta **tabula** para extração de tabelas.
- Geração de um arquivo CSV por campus ou por lote de páginas.
- Registro de eventuais perdas de formatação ou colunas inconsistentes.

Entregas

- Arquivos CSV extraídos dos PDFs.
- Relatório de consistência da extração (campos faltantes, colunas quebradas, páginas ignoradas).

Critérios de Aceitação

- 100% das páginas dos PDFs foram processadas.
- Todos os itens do PDF aparecem pelo menos uma vez no CSV.
- Não há colunas ilegíveis ou misturadas sem mapeamento claro.

Fase 2 - Limpeza e Agrupamento (Sem IA)

Objetivo

Eliminar ruídos nos dados, corrigir variações textuais e consolidar descrições semelhantes.

Ferramenta

- **OpenRefine** (clustering + facetas).

Processo

- Aplicação de técnicas de clustering (key-collision e nearest-neighbor).
- Agrupamento de descrições divergentes: Ex.: "Cad. Giratória", "Cadeira Girat.", "Cadeira Giratoria" - **Cadeira Giratória**
- Filtragem por campus, setor e tipo para localizar padrões divergentes.

Entregas

- Planilha única consolidada e limpa.
- Lista de descrições únicas (catálogo de itens).
- Relatório de métodos de clustering aplicados.

Critérios de Aceitação

- Todas as variações textuais foram avaliadas (100% revisadas).
- Redução significativa de descrições: ex.: de 10.000 linhas - ~500 itens únicos.
- Nenhum item fica sem descrição corrigida.

Fase 3 - Padronização Semântica

Objetivo

Criar uma lista oficial e padronizada de categorias e descrições conforme exigências do SIADS.

Processo

- Extração da lista de descrições únicas da Fase 2.
- Uso de IA local (**LM Studio**) ou online (ChatGPT/Gemini/Claude) para padronizar a nomenclatura para conformidade semântica, nos casos em que o OpenRefine não resolveu.

Entradas

- Lista com descrições originais únicas

Saídas

- Tabela “De - Para”:
 - **Nome original**
 - **Nome padronizado**

Critérios de Aceitação

- 100% das descrições únicas categorizadas.
- Todas as categorias são válidas para o SIADS.
- Não existe item sem padronização.
- Revisão humana aprova a lista final.

Fase 4 - Cruzamento Final e Geração da Base Migrável

Objetivo

Aplicar o “de-para” a toda a base, gerando o arquivo final no padrão do SIADS.

Processo

- Aplicação via:
 - PROCV/VLOOKUP no Excel ou Join no Python/OpenRefine
- Inserção das descrições padronizadas em todas as linhas.

Entregas

- Arquivo final padronizado (CSV ou XLSX).
- Documentação do de-para.
- Relatório de migração (percentual de sucesso e inconsistências tratadas).

Critérios de Aceitação

- 100% dos itens possuem descrição padronizada.

<ul style="list-style-type: none"> Os dados presentes inicialmente, com exceção dos tratados por nós, ainda estão presentes na planilha. O cliente valida que os dados representam o inventário real.
Origem dos Dados
<ul style="list-style-type: none"> PDFs fornecidos pelos campi do IFPE. Planilhas adicionais de setores (se existirem). Tabelas oficiais das categorias SIADS (Tesouro Nacional).
Arquitetura Funcional
<p>A solução adotada é não-sistêmica (não envolve desenvolvimento de software), mas segue uma arquitetura de processos, com ferramentas interoperáveis:</p> <p>Entrada</p> <ul style="list-style-type: none"> PDFs - Tabula <p>Processamento</p> <ul style="list-style-type: none"> CSV - OpenRefine <p>Padronização semântica</p> <ul style="list-style-type: none"> Lista única - IA Local ou Online (LM Studio, ChatGPT etc.) <p>Saída</p> <ul style="list-style-type: none"> Planilha padronizada - CSV final <p>Armazenamento</p> <ul style="list-style-type: none"> Pastas versionadas no Google Drive (ou outra nuvem definida pelo cliente). <p>Segurança</p> <ul style="list-style-type: none"> Dados mantidos localmente no OpenRefine. IA local (LM Studio) para evitar envio de dados sensíveis.
Escopo Contemplado
<ul style="list-style-type: none"> Levantamento e análise dos dados originais. Definição do padrão de nomenclatura e categorias. Padronização e limpeza dos dados. Construção do fluxo de migração. Testes de carga e consistência. Validação com o cliente. Documentação completa das regras e procedimentos adotados.
Escopo não contemplado
<ul style="list-style-type: none"> Migração parcial de dados; a integração será obrigatoriamente total para todos os campi. Implantação direta no SIADS (responsabilidade da instituição). Desenvolvimento de sistema próprio completo (o projeto cria fluxo/processo, não um software robusto). Correção manual de dados que dependerem de informações não fornecidas pelo cliente. Responsabilidade sobre falhas externas decorrentes de acesso, permissões ou funcionamento do SIADS.
Premissas
<ul style="list-style-type: none"> O SUAP disponibilizará dados via relatórios (PDF). Todos os campi participarão ativamente do processo de revisão de cadastros. O SIADS manterá o formato e requisitos de integração atualmente definidos
Restrições
<ul style="list-style-type: none"> Tecnológica: O SUAP não possui API pública, exigindo extração de dados via relatórios.

- Orçamentária: Custo zero para aquisição de softwares. Soluções devem utilizar tecnologias de código aberto ou já disponíveis no IFPE.
- Escopo: Migração total e indivisível ("tudo ou nada").

Data Inicial	Data final	Custo Estimado
28/09/2025	12/12/2025	R\$ 18.590,00
Elaborado por: Ana Beatriz Gomes da Silva		
Recife, 02 de dezembro de 2025		