



MINISTÉRIO DA EDUCAÇÃO
Secretaria Executiva

Plano de Dados Abertos – PDA MEC – 2016/2018

Versão: 01.8

Código:
PDAMEC-01-
2016-2018

Data da versão:
06/12/2016

Dicionário - PRONATEC

Conjunto de Dados

PRONATEC

Nome ou título do conjunto de dados

Dados da Rede Federal de EPCT (Educação Profissional, Científica e Tecnológica)

Soluções escolhidas para publicar dados abertos (formatos de saída dos dados):

Solução escolhida	Pré-requisito identificado	Motivo da escolha
Publicar dados em arquivos CSV	Mecanismo de ETL	Possui baixa complexidade e curto prazo de implementação frente ao prazo estipulado pelo Acórdão TCU e são de conhecimento e uso mais abrangente por parte dos consumidores de dados abertos de acordo com o exposto no Cartilha Técnica para Publicação de Dados Abertos no Brasil v1.0.
	Servidor web para arquivos	

CSV

A Cartilha Técnica para Publicação de Dados Abertos no Brasil v1, publicada pela Secretaria de Logística e Tecnologia da Informação – SLTI do Ministério do Planejamento Orçamento e Gestão – MP, define CSV como um formato para armazenamento de dados tabulares em texto. A codificação é muito simples: cada linha do arquivo representa uma linha na tabela, e as colunas são separadas por vírgula. As planilhas CSV têm a vantagem de poderem ser facilmente abertas e manipuladas em qualquer editor de planilhas, inclusive por uma pessoa que não tenha conhecimentos de programação. Além disso, são igualmente fáceis de serem consumidas por aplicações. Algumas de suas vantagens que determinam a escolha desse formato para a disponibilização de Dados Abertos do MEC são:

- Simplicidade (registros em estrutura tabular);
- Facilidade de geração (qualquer banco de dados exporta);
- Facilidade de consumo (qualquer editor de planilhas manipula);
- Sucinto (os arquivos gerados são menores).

Descrição dos campos

Documentação de quais são e o que significam cada coluna e seus tipos de dados de acordo com o conjunto de dados selecionados para abertura na ficha dos Gestores publicadas no PDA do MEC.

Instituições da Rede Federal de EPCT			
Nome	Campo	Tipo	Descrição
Código da Unidade	CODIGO_UNIDADE_ENSINO	Alfanumérico	Código da Unidade de Ensino da Rede Federal de EPCT.
Nome da Rede	NOME_UNIDADE_ENSINO	Alfanumérico	Nome da Unidade de Ensino da Rede Federal de EPCT.
Data de autorização de funcionamento	DT_PUBLICACAO_ATO_AUTORIZATIVO	Data	Data de autorização de funcionamento da Unidade da Rede Federal de EPCT.



MINISTÉRIO DA EDUCAÇÃO
Secretaria Executiva

Plano de Dados Abertos – PDA MEC – 2016/2018

Versão: 01.8

Código:
PDAMEC-01-
2016-2018

Data da versão:
06/12/2016

Região	NOME_REGIAO_UNIDADE	Alfanumérico	Região das Unidades da Rede Federal de EPCT.
UF	SIGLA_UF_UNIDADE	Alfanumérico	UF das Unidades da Rede Federal de EPCT.
Município	NOME_MUNICIPIO_UNIDADE	Alfanumérico	Município das Unidades da Rede Federal de EPCT.
Mesorregião	NOME_MESORREGIAO_UNIDADE	Alfanumérico	Mesorregião das Unidades da Rede Federal de EPCT.
Microrregião	NOME_MICRORREGIAO_UNIDADE	Alfanumérico	Microrregião das Unidades da Rede Federal de EPCT.

Matrículas Atendidas da Rede Federal de EPCT			
Nome	Campo	Tipo	Descrição
Código da Unidade	CODIGO_UNIDADE	Alfanumérico	Código da Unidade de Ensino da Rede Federal de EPCT.
Nome da Unidade	NOME_UNIDADE	Alfanumérico	Nome da Unidade de Ensino da Rede Federal de EPCT.
Tipo de Curso	TIPO_CURSO	Alfanumérico	Tipo de curso que recebeu matrículas (Concomitante ou Subsequente).
Curso	NOME_CURSO	Alfanumérico	Nome do curso.
Tipo de Oferta	TIPO_OFERTA	Alfanumérico	Tipo de oferta do curso (FIC ou Técnico).
Modalidade de Ensino	NOME_MODALIDADE	Alfanumérico	Nome da modalidade de ensino (Presencial ou EAD – Ensino a Distância).
Quantidade de matrículas	QUANT_MATRICULA	Numérico	Quantitativo de matrículas atendidas da Rede Federal de EPCT.
Região	NOME_REGIAO_MATRICULA	Alfanumérico	Região das matrículas atendidas na Rede Federal de EPCT.
UF	SIGLA_UF_MATRICULA	Alfanumérico	UF das matrículas atendidas na Rede Federal de EPCT.
Município	NOME_MUNICIPIO_MATRICULA	Alfanumérico	Município das matrículas atendidas na Rede Federal de EPCT.
Mesorregião	NOME_MESORREGIAO_MATRICULA	Alfanumérico	Mesorregião das matrículas atendidas na Rede Federal de EPCT.
Microrregião	NOME_MICRORREGIAO_MATRICULA	Alfanumérico	Microrregião das matrículas atendidas na Rede Federal de EPCT.

Novas Matrículas da Rede Federal de EPCT			
Nome	Campo	Tipo	Descrição
Código da Unidade	CODIGO_UNIDADE	Alfanumérico	Código da Unidade de Ensino que recebeu novas matrículas.
Nome da Unidade	NOME_UNIDADE	Alfanumérico	Nome da Unidade de Ensino que recebeu novas matrículas.
Tipo de Curso	TIPO_CURSO	Alfanumérico	Tipo de curso que recebeu novas matrículas (Concomitante ou Subsequente).
Curso	NOME_CURSO	Alfanumérico	Nome do curso.
Tipo de Oferta	TIPO_OFERTA	Alfanumérico	Tipo de oferta do curso (FIC ou Técnico).
Modalidade de Ensino	NOME_MODALIDADE	Alfanumérico	Nome da modalidade de ensino (Presencial ou EAD – Ensino a Distância).
Quantidade de novas matrículas	QUANT_NOVAS_MATRICULAS	Numérico	Quantitativo de novas matrículas da Rede Federal de EPCT.
Região	NOME_REGIAO_NOVAS_MATRICULAS	Alfanumérico	Região das novas matrículas da Rede Federal de EPCT.
UF	SIGLA_UF_NOVAS_MATRICULAS	Alfanumérico	UF das novas matrículas da Rede Federal de EPCT.
Município	NOME_MUNICIPIO_NOVAS_MATRICULAS	Alfanumérico	Município das novas matrículas da Rede Federal de EPCT.
Mesorregião	NOME_MESORREGIAO_NOVAS_MATRICULAS	Alfanumérico	Mesorregião das novas matrículas da Rede Federal de EPCT.
Microrregião	NOME_MICRORREGIAO_NOVAS_MATRICULAS	Alfanumérico	Microrregião das novas matrículas da Rede Federal de EPCT.



MINISTÉRIO DA EDUCAÇÃO
Secretaria Executiva

Plano de Dados Abertos – PDA MEC – 2016/2018

Versão: 01.8

Código:
PDAMEC-01-
2016-2018

Data da versão:
06/12/2016

Concluintes da Rede Federal de EPCT			
Nome	Campo	Tipo	Descrição
Código da Unidade	CODIGO_UNIDADE	Alfanumérico	Código da Unidade de Ensino dos concluintes da Rede Federal de EPCT
Nome da Unidade	NOME_UNIDADE	Alfanumérico	Nome da Unidade de Ensino dos concluintes da Rede Federal de EPCT
Tipo de Curso	TIPO_CURSO	Alfanumérico	Tipo do curso dos concluintes (Concomitante ou Subsequente).
Curso	NOME_CURSO	Alfanumérico	Nome do curso.
Tipo de Oferta	TIPO_OFERTA	Alfanumérico	Tipo de oferta do curso (FIC ou Técnico).
Modalidade de Ensino	NOME_MODALIDADE	Alfanumérico	Nome da modalidade de ensino (Presencial ou EAD – Ensino a Distância).
Quantidade de concluintes	QUANT_CONCLUINTES	Numérico	Quantitativo de concluintes da Rede Federal de EPCT.
Região	NOME_REGIAO_CONCLUINTES	Alfanumérico	Região dos concluintes da Rede Federal de EPCT.
UF	SIGLA_UF_CONCLUINTES	Alfanumérico	Sigla da UF dos concluintes da Rede Federal de EPCT.
Instituição	CODIGO_INSTITUICAO_CONCLUINTES	Alfanumérico	Código da Instituição dos concluintes da Rede Federal de EPCT.
Município	NOME_MUNICIPIO_CONCLUINTES	Alfanumérico	Município dos concluintes da Rede Federal de EPCT.
Mesorregião	NOME_MESORREGIAO_CONCLUINTES	Alfanumérico	Mesorregião dos concluintes da Rede Federal de EPCT.
Microrregião	NOME_MICRORREGIAO_CONCLUINTES	Alfanumérico	Microrregião dos concluintes da Rede Federal de EPCT.

Metadados

Metadados são dados sobre os dados, ou seja, são informações que possibilitam organizar, classificar, relacionar e inferir novos dados sobre o conjunto de dados. A quantidade e a qualidade dos metadados de um conjunto de dados podem determinar a utilidade de um conjunto de dados. Quanto melhores os metadados, mais valor agregado se oferta ao conjunto de dados, além de melhorar a classificação e uso de filtros e buscas. A seguir são definidos os conjuntos de metadados deste conjunto de dados.

Descrição dos Metadados para catalogação no dados.gov

PRONATEC	
Título	Dados da Rede Federal de EPCT
Descrição	Lista de todas as instituições da Rede Federal de EPCT, com detalhamento de nome, município, data de autorização de funcionamento, quantidade de matrículas atendidas, quantidade de novas matrículas e quantidade de concluintes por iniciativa do Pronatec, unidade, tipo de curso, curso, tipo de oferta e modalidade de ensino da rede federal de EPCT. Quantidade de Professores EBTT e Técnicos Administrativos em Educação (TAE) por campus, instituição, qualificação e nível funcional. Valor total disponibilizado por instituição via matriz orçamentária ou descentralização e situação das obras de expansão da Rede e de reestruturação.
Catálogo origem	dados.gov.br
Órgão responsável	Secretaria Executiva (SE) do Ministério da Educação (MEC).
VCGE	http://vocab.e.gov.br/2011/03/vcge#educacao-profissional-tecnologica
Etiquetas	PRONATEC, Ensino Técnico, Profissionalização
Autoria	Coordenação-Geral de Planejamento e Gestão CGPG/DDRSETEC/MEC



MINISTÉRIO DA EDUCAÇÃO
Secretaria Executiva

Plano de Dados Abertos – PDA MEC – 2016/2018

Versão: 01.8

Código:
PDAMEC-01-
2016-2018

Data da versão:
06/12/2016

Documentação	Arquivo PDAMEC-01-2016-2018 Versão 01.8
Cobertura geográfica	Todos os municípios brasileiros que tenham beneficiários do PRONATEC.
Cobertura temporal	2011 a 2016
Granularidade geográfica	Municipal
Granularidade temporal	Semestral
Frequência de atualização	Semestral
Referências	Este conjunto de dados pode ser relacionado com dados do CENSO da Educação Superior (INEP)
Metodologia	O processo de criação desse conjunto de dados seguiu a metodologia de "Barramento de Dados" do Escritório de Gestão de Dados e Informação Estratégica (EGDIE/SE/MEC) que consiste em modelar, de forma multidimensional ODS (Operational Data Store), os dados do programa sob o aspecto negocial e realizar cargas por meio de ETL para atualização. A partir deste repositório podem ser construídos DWs (Data Warehouse) com versionamento via SCD (Slowly Changing Dimensions) que retrata as dimensões que sofrem atualizações em seus campos e os classifica pelo tipo de mudança existente em cada uma delas.
Recursos:	
- Identificador	dados.gov.br – Ministério da Educação
- Título	PRONATEC
- Formato	CSV
- Descrição	Arquivos nos formatos CSV contendo dados negociais do Programa PRONATEC referente às Redes Federais de EPCT dos anos 2011 a 2016.

Observação sobre os conjuntos de dados

Para cada conjunto de dados especificados neste documento nos itens 2, 3 e 4, foram gerados arquivos (CSV) anuais, ou seja, um arquivo por ano para cada conjunto de dados de matrículas, excetuando apenas o conjunto de dados do Item 1, com a lista das Instituições da Rede Federal de EPCT.

Conceitos aplicados aos status de matrículas

Considerando que para cada conjunto de dados foi gerado um arquivo CSV, alguns conceitos foram aplicados quanto ao status das matrículas para que os registros pudessem ser agrupados por ano.

Matrículas Novas

É o conjunto de dados referente aos registros de matrículas de ingressantes. Para o agrupamento por ano foi utilizado o seguinte conceito:

- *Matrículas nova/ingressante são matrículas cuja data de início do CURSO pertence ao mesmo ano do agrupamento ou do arquivo gerado;*
- *Exemplo: no arquivo de Matrículas Novas 2011, a matrícula 0001 tem como início do curso a data 01/01/2011 portanto, esta foi considerada, neste conjunto de dados como matrícula nova/ingressante em 2011.*



MINISTÉRIO DA EDUCAÇÃO
Secretaria Executiva

Plano de Dados Abertos – PDA MEC – 2016/2018

Versão: 01.8

Código:
PDAMEC-01-
2016-2018

Data da versão:
06/12/2016

Matrículas Concluintes

É o conjunto de dados referente aos registros de matrículas de concluintes de cursos da Rede Federal de EPCT. Para o agrupamento por ano foi utilizado o seguinte conceito:

- Matrículas concluintes são matrículas cuja data de término do CURSO pertence ao mesmo ano do agrupamento ou do arquivo gerado;
- Exemplo: no arquivo de Matrículas Concluintes 2012, a matrícula 0002 tem como término do curso a data 02/02/2012 portanto, esta foi considerada, neste conjunto de dados, como matrícula concluinte em 2011.

Matrículas Atendidas

É o conjunto de dados referente a todos os registros de matrículas de cursos da Rede Federal de EPCT, ou seja, todas as matrículas atendidas, sejam estas novas/ingressantes ou concluintes. Para o agrupamento por ano foi utilizado o seguinte conceito:

- Matrículas atendidas são aquelas cuja data de início e término do CURSO pertence ao mesmo ano do agrupamento ou do arquivo gerado;
- Exemplo: no arquivo de Matrículas Atendidas 2013, a matrícula 0003 tem como início e término do curso a data 02/02/2013 portanto, esta foi considerada, neste conjunto de dados, como matrícula atendida em 2011.