

Ciência de Dados Aplicada aos Dados Abertos da Universidade Federal de Ouro Preto a partir dos Conjuntos de Dados de Assistência Estudantil

Lucineia Souza Maia¹, Rian Lucas Dias Guedes²

¹Departamento de Sistemas e Computação
Universidade Federal de Ouro Preto– João Monlevade – MG – Brazil

²Colegiado de Sistemas de Informação
Universidade Federal de Ouro Preto– João Monlevade – MG – Brazil

lucineia@ufop.edu.br, rian.guedes@aluno.ufop.edu.br

Abstract. *Since it was established in 2011, Law 12,527 guarantees Brazilian citizens access to public information from government agencies and institutions in all spheres. According to the aforementioned legislation, some information can be requested through Citizens' Information Systems while others are made available a priori in information transparency services. In this case, one of the mechanisms of the Federal University of Ouro Preto to provide transparency is the UFOP Dados Abertos Portal. In it, data sets were organized, grouped as: Administration, institutional, planning, people management, student assistance, graduation, extension, research and post-graduation. However, going against the very nature of data, the data available on this Portal is in its raw state, that is, in the form of tables in which the correlation, analysis and visualization of information is not intuitive. In this sense, the objective of this project is to use Data Science to identify insights about the intersections between UFOP's open student assistance, extension, research and graduate data sets. Therefore, the methodology will employ as a methodological procedure the case study combined with Data Science techniques. As a result, it is expected to reveal the influence of student assistance actions for the development of extension and research at UFOP by the benefited students.*

Keywords: *Open Data, Access to Information Law, Data Science.*

Resumo. *Desde que estabelecida em 2011, a Lei 12.527 garante aos cidadãos brasileiros o acesso às informações públicas dos órgãos e instituições governamentais em todas as esferas. De acordo com a referida legislação, algumas informações podem ser requeridas por meio de Sistemas de Informações aos Cidadãos enquanto outras são disponibilizadas a priori em serviços de transparência de informações. Nesse caso, um dos mecanismos da Universidade Federal de Ouro Preto para prover transparência é o Portal UFOP Dados Abertos. Nele foram organizados conjuntos de dados agrupados como: Administração, Institucional, Planejamento, Gestão de Pessoas, Assistência Estudantil, Graduação, Extensão, Pesquisa e Pós - Graduação. Contudo, indo de encontro com a própria natureza de dados, os dados disponibilizados nesse Portal estão em estado bruto, ou seja, em forma de tabelas as quais a*

correlação, a análise e visualização das informações não é intuitiva. Nesse sentido, o objetivo desse projeto é utilizar Ciência de Dados para identificar insights a respeito das intersecções entre os conjuntos de dados abertos de assistência estudantil, extensão, pesquisa e pós-graduação da UFOP. Logo, a metodologia empregará como procedimento metodológico o estudo de caso combinado com as técnicas de Ciência de Dados. Como resultados espera-se revelar a influência das ações de assistência estudantil para o desenvolvimento da extensão e da pesquisa na UFOP pelos discentes beneficiados.

Palavras-chave: Dados Abertos, Lei de Acesso à Informação, Ciência de Dados.

1. Introdução

A Lei nº 12.527 de 18 de novembro de 2011, Lei de Acesso à Informação (LAI), determina que todos os cidadãos têm direito ao acesso às informações públicas. Algumas dessas informações podem estar prontamente publicizadas, outras podem ser requeridas pelos cidadãos. Esse projeto trata do primeiro caso, de informações divulgadas independentemente de solicitações dos cidadãos. Dentre essas informações de acompanhamento de programas, ações e projetos [BRASIL 2018].

Isto posto, a LAI determina que as informações previamente divulgadas devem ser transparentes e estar em local de fácil acesso nos sites oficiais dos órgãos aos quais a Lei se aplica. Desse modo, a UFOP implementou o Portal de Dados Abertos que, de acordo com o Plano de Dados Abertos 2020 - 2022, tem como objetivo “disponibilizar para toda a comunidade dados e informações públicas da Instituição, possibilitando a concretização do acesso a informação (regulamentado pela Lei de Acesso à Informação Pública - Lei nº 12.527/2011) e o acompanhamento do funcionamento da Universidade” [UFOP 2021]. O Portal visa “promover a manutenção e atualização de dados da UFOP, garantindo os princípios da publicidade e da transparência na administração pública, com o compromisso de divulgação permanente dos dados de interesse público produzidos nas diversas esferas de sua competência” [UFOP 2021].

O Portal de Dados Abertos da UFOP [UFOP 2021] possui atualmente 56 conjuntos de dados. Esse conjuntos de dados foram estabelecidos considerando “o grau de relevância das informações para o cidadão, o alinhamento com os instrumentos de planejamento estratégico da Instituição, consulta pública e a disponibilidade de dados já gerenciados pelos Sistemas Integrados de Gestão da UFOP” [UFOP 2021]. Os conjuntos de dados foram organizados em 8 grupos: Administração; Assistência Estudantil; Extensão; Graduação; Institucional; Pesquisa e Pós - Graduação; Planejamento e; Gestão de Pessoas. Cada um desses grupos possui quantidades diferentes de conjuntos de dados. O grupo Assistência Estudantil, por exemplo, possui 4 conjuntos de dados: Bolsa Remunerada; Moradia Estudantil; Bolsa Alimentação e; Bolsa Transporte.

Os dados disponibilizados no Portal de Dados Abertos da UFOP [UFOP 2021] são registros “brutos” ou “evidências relevantes” referentes aos conjuntos de dados. Isso vai de acordo com a definição de dados “coleção de evidências relevantes sobre um fato” [SORDI 2008]. Os dados tem natureza estruturada, não obstante, os dados disponibilizados no Portal estão organizados estruturadamente em tabelas, o que torna sua

interpretação difícil de ser realizada. Desse modo, é necessário transformar esses dados em informação.

A informação, segundo Sordi [SORDI 2008] é a interpretação de um conjunto de dados segundo um propósito relevante e de consenso para o público-alvo (leitor). Ainda de acordo com o autor, “um conjunto de dados não produz necessariamente uma informação, logo é necessário processar esses dados para gerar informação”. No processamento dos dados são utilizadas operações aritméticas e lógicas (algoritmos).

A Ciência de Dados tem se debruçado em desenvolver estratégias para processar dados (em grandes quantidades) e extrair *insights* a partir deles, desse modo, as informações geradas a partir de combinações de conjuntos de dados podem revelar comportamentos impossíveis de serem previstos sem seu poder analítico, aumentando seu valor. Segundo [A.BEAL 2004], o valor da informação aumenta quando há combinação de informações.

Portanto, esse projeto visa aplicar Ciência de Dados no Portal de Dados Abertos da UFOP tendo como partida os conjuntos de dados de Assistência Estudantil. Em alguns desses conjuntos de dados é possível identificar os beneficiários, desse modo, pode-se localizar as outras ocorrências desses indivíduos nos conjuntos de dados de Pesquisa e Pós-Graduação e de Extensão e analisar as relações entre elas.

1.1. Objetivos

O objetivo geral deste projeto de Iniciação Científica é utilizar Ciência de Dados para identificar *insights* a respeito das intersecções entre os conjuntos de dados abertos de assistência estudantil, extensão, pesquisa e pós-graduação da UFOP. Os objetivos específicos são:

- Interpretar o conhecimento extraído por meio da análise dos dados.
- Usar as estratégias de Ciência de Dados para compreender problemas reais.
- Identificar convergências entre os dados.

1.2. Justificativa / Relevância

Esse projeto justifica-se por (1) aplicar computação na solução de problemas e (2) tratar de temáticas de Sistemas de Informação, mais necessariamente do componente curricular Gestão da Informação.

Os avanços na tecnologia da informação tornaram possível obter, gerir e usar grandes quantidades de dados [GORDON and GORDON 2006]. Desse modo, a primeira justificativa, diz respeito as técnicas e modelos de Ciência de Dados que serão utilizados. Nesse caso, serão utilizados algoritmos para combinar dados a partir dos seguintes conjuntos de dados de Assistência Estudantil: Bolsa Remunerada, Bolsa Alimentação e Bolsa Transporte. Esses dados serão confrontados com os conjuntos de dados de Extensão: Bolsista de Extensão e Voluntários de Extensão e também com os conjuntos de dados de Pesquisa e Pós-Graduação: Alunos de Pós-Graduação e Bolsistas de Iniciação Científica. Com base nisso, ficará nítido e intuitivo para o usuário as implicações da Assistência Estudantil (grupo de dados em questão) na participação do aluno na Extensão e na Pesquisa. Desse modo, o aspecto da LAI que trata da transparência das informações será expandido.

A segunda justificativa refere-se ao conhecimento de Gestão da Informação aplicado neste projeto. Nele, será possível distinguir com clareza dado de informação, destacando as características de cada um desses conceitos. No contexto de Sistemas de Informação, a Informação é o elemento que torna necessário o desenvolvimento dos sistemas. Ou seja, os sistemas foram feitos para entregar informação, logo é tão importante projetos que entreguem informações com qualidade para os usuários. No Portal de Dados Abertos da UFOP não é objetivo fazer o tratamento dos dados para gerar interfaces que facilitam a interpretação, por isso, os dados lá disponibilizados precisam passar por tratamentos para que possam ser intuitivos e claros para os usuários.

2. Metodologia

Quanto à sua natureza, esta pesquisa está classificada como aplicada. Segundo Silveira e Cordova [SILVEIRA and CÓRDOVA 2009], a pesquisa aplicada "[objetiva gerar conhecimentos para a aplicação prática, dirigidos à solução de problemas específicos]". Quanto aos objetivos da pesquisa, este projeto tem a característica de ser uma pesquisa exploratória. Segundo Gil, "Estas pesquisas tem como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses. Pode-se dizer que estas pesquisas [exploratórias] têm como objetivo principal o aprimoramento de ideias ou a descoberta de intuições" [Gil 2002].

Por fim, quanto aos procedimentos, esta pesquisa empregará o estudo de caso. Gil [Gil 2002] afirma que o estudo de caso "[consiste no estudo profundo e exaustivo de um ou poucos objetos, de maneira que permita seu amplo e detalhado conhecimento]". Para o autor, o estudo de caso permite uma visão global do problema ou identifica possíveis fatores que influenciam ou que são por ele influenciados na amostra.

O planejamento do estudo de caso norteará as atividades desse projeto. Para isso, algumas atividades do Processo de Experimentação de Wohlin [WHOLIN 2000] combinadas com as etapas de Ciência de Dados foram feitas.

3. Fundamentação Teórica

3.1. Lei de Acesso à Informação

Pela ótica do direito e da história, o direito de acesso à informação no Brasil ainda pode ser considerado recente. Isso porque no período da ditadura militar a regra era o sigilo absoluto, cujos fundamentos eram baseados em razões de segurança nacional. Vale lembrar que até mesmo antes de tal período, tanto em âmbito nacional, quanto internacional, não havia preocupação com a transparência dos atos do Estado. Com o fim da ditadura militar e a promulgação da carta magna de 1988, essa discussão passou a ser mais relevante [GRUMAN 2012]. Contudo, somente em 18 de novembro de 2011 foi sancionada a Lei de Acesso a Informação (LAI), que corresponde a Lei n. 12.527.

O principal objetivo da LAI de acordo com o artigo terceiro [BRASIL 2018] é assegurar o direito fundamental do acesso à informação e que devem ser executados em conformidade com os princípios da administração pública, que são:

I - observância da publicidade como preceito geral e do sigilo como exceção;

II - divulgação de informações de interesse público, independentemente de solicitações;

III - utilização de meios de comunicação viabilizados pela tecnologia da informação;

IV - fomento ao desenvolvimento da cultura de transparência na administração pública;

V - desenvolvimento do controle social da administração pública.

Segundo Esmarioto Carpes e Clicéres Mack Dal Bianco, uma das afirmações da LAI de maneira subentendida é que a lei diz que todo cidadão brasileiro poderá solicitar qualquer informação pública, exceto aquelas classificadas como sigilosas, tanto no Governo Federal, quanto no Estadual ou no Municipal, e em qualquer órgão público dos três poderes: Executivo, Legislativo e Judiciário, sendo aplicada esta regra, também, para as entidades sem fins lucrativos que recebam recursos públicos [E.CARPES and BIANCO 2013].

Dessa forma, a LAI tornou-se um importante instrumento de fomento a transparência pública, bem como, de consolidação do regime democrático de direito. Nesse sentido, Marcelo Gruman afirma que a "cultura do acesso permite cada vez mais a participação da sociedade na avaliação das ações do Estado e mesmo na proposição de mudanças de rumo" [GRUMAN 2012]. Logo, é imprescindível a citação do inciso XX-XIII do artigo quinto da Constituição Federal, que diz: "todos tem direito a receber dos órgãos públicos informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo da lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível a segurança da sociedade e do Estado" [BRASIL 1988].

Consequentemente, com a aprovação e implementação da LAI, surgiu em 2011 os dados abertos governamentais (*Open Government Data*) garantindo a transparência de informações, melhoria do controle social e desenvolvimento do relacionamento entre o governo e a sociedade, ressaltando que a promoção da transparência se deve a Lei da Transparência 131/2009. De acordo com o artigo primeiro da Lei Complementar 131/2009 referente a Lei da Transparência [BRASIL 2009] a transparência será assegurada pelos fatores:

I – incentivo à participação popular e realização de audiências públicas, durante os processos de elaboração e discussão dos planos, lei de diretrizes orçamentárias e orçamentos;

II – liberação ao pleno conhecimento e acompanhamento da sociedade, em tempo real, de informações pormenorizadas sobre a execução orçamentária e financeira, em meios eletrônicos de acesso público;

III – adoção de sistema integrado de administração financeira e controle, que atenda a padrão mínimo de qualidade estabelecido pelo Poder Executivo da União e ao disposto no art. 48-A." (NR).

Mediante a este parecer acerca da definição de transparência informacional, foi uma premissa ao surgimento do Sistema de Informação ao Cidadão (SIC) e o Sistema Eletrônico de Serviços de Informação ao Cidadão (e-SIC). De acordo com o Ministério da Economia [da Economia 2022], O SIC é a unidade responsável por atender os pedidos de acesso à informação feitos ao Ministério da Economia, com base na Lei nº 12.527 de 18 de novembro de 2011 (Lei de Acesso à Informação), não sendo necessário jus-

tificar, enquanto o e-SIC permite aos cidadãos fazer pedidos de informações públicas e manifestações de ouvidoria num único local [GOV.BR 2022].

3.2. Ciência de Dados

De acordo com Cao, a Ciência de Dados é o estudo de dados que no ponto de vista disciplinar se sintetiza com Estatística, Informática, Gestão e Sociologia para fazer com que os dados gerem ideias de tomada de decisão em organizações. Logo, é uma área nova e com inovações em contextos de caráter ético, social, econômico, político e organizacional que segue em debate na sociedade contemporânea [Cao 2017].

Segundo [Provost and Fawcett. 2013] essa área é um conjunto de princípios fundamentais que apoiam e orientam a extração baseada em princípios de informações e conhecimento dos dados. Logo, os autores idealizaram uma lista com conceitos fundamentais acerca do *Data Science* que são:

- Extrair conhecimento útil de dados para resolver problemas de negócios pode ser tratado sistematicamente seguindo um processo com etapas razoavelmente bem definidas.
- A avaliação dos resultados da ciência de dados requer uma consideração cuidadosa do contexto em que eles serão usados.
- A relação entre o problema de negócios e a solução analítica muitas vezes pode ser decomposta em problemas menores tratáveis por meio da estrutura de análise do valor esperado.
- A tecnologia da informação pode ser usada para encontrar itens de dados informativos dentro de um grande corpo de dados.
- Entidades que são semelhantes em relação a características ou atributos conhecidos geralmente são semelhantes em relação a características ou atributos desconhecidos.
- Se você olhar muito para um conjunto de dados, encontrará algo, mas pode não generalizar além dos dados que você está observando.
- Para tirar conclusões causais, deve-se prestar muita atenção à presença de fatores de confusão, possivelmente invisíveis.

Certamente, a Ciência de Dados, *Big Data* e análises avançadas têm sido cada vez mais reconhecidas como as principais forças motrizes para a inovação, economia e educação da próxima geração. Nessa perspectiva, os dados permitem processos a serem realizados que solidificam os conceitos de Ciência de Dados, entre eles (1) Visualização de Dados, (2) Análise Exploratória, (3) Modelagem de Dados e (4) Estrutura de Dados [Cao 2017].

Em suma, é uma área interdisciplinar que está em evolução que se baseia em âmbitos e necessita de adquirir estratégias de diferentes contextos para resultados mais precisos, como mostra a figura:

4. Ferramentas

4.1. Python

De acordo com a própria organização, é uma programação interpretada, orientada a objetos e de alto nível de linguagem com semântica dinâmica além dos seus dados serem

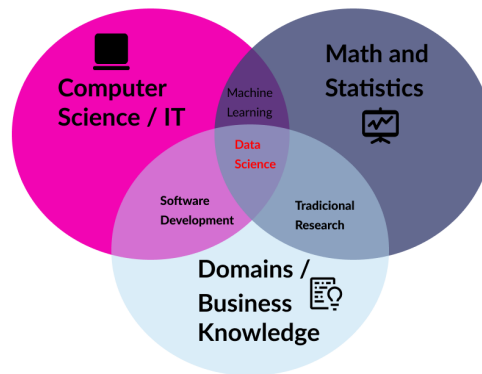


Figura 1. Adaptado de [CONWAY 2022].

integrados de alto nível com estruturas combinadas e digitação e vinculação dinâmica [Org. 2022]. A escolha se deve por ser *Open Source* e também por ser uma das principais linguagens utilizadas na Ciência de Dados para análises e visualizações de dados devido a um ecossistema de pacotes e bibliotecas disponíveis.

5. Desenvolvimento

5.1. Definição do Estudo de Caso

O estudo de caso será norteado por meio da aplicação de *Data Science* nos dados abertos disponibilizados pela UFOP, sendo analisados os dados de Assistência Estudantil, Pós-Graduação, Pesquisa e Extensão.

A situação – problema do estudo de caso é sobre como a apresentação tabular de dados públicos podem ocultar a extração e obtenção de conhecimentos necessários para entender a assistência estudantil e como sua aplicação é relevante para manter investimentos do governo brasileiro, logo, a falta de intuição e facilitação ao usuário por meio dos dados e uma argumentação para planejar um desenvolvimento metodológico á problemática. Os dados estão agrupados por uma base de dados, conforme o quadro:

Desse modo, alinhado aos objetivos gerais e específicos da pesquisa, este estudo de caso irá aplicar *Data Science* com o intuito de extrair os melhores conhecimentos possíveis por meio da organização desses dados e também facilitar a visualização dos dados para os usuários. Logo, é definido as questões norteadoras, que são perguntas que durante a investigação serão respondidas ao longo do processo da investigação [GONCALVES 2001]. Para tal, algumas questões norteadoras para a análise dos dados foram colocadas, são elas:

- Os alunos que tem assistência estudantil atuam mais em projetos de extensão ou pesquisa?
- Os alunos que recebem assistência estudantil também possuem acesso a bolsas de extensão?

Nome da Categoria de Dados	Quantidade de Conjuntos de Dados	Nome dos Conjuntos
Assistência Estudantil	4	Moradia Estudantil Bolsa Renumerada Bolsa Alimentação Bolsa Transporte
Pesquisa e Pós-Graduação	8	Equipamentos Grupos de Pesquisa PROPP Alunos Orientadores Pesquisadores Bolsistas Projetos
Extensão	6	Bolsistas Voluntários Cursos Projetos Empresas Juniores Programas

- Os alunos contemplados com assistência estudantil buscam projetos de extensão de maneira voluntária?
- Os alunos que recebem assistência estudantil são propensos a serem bolsistas ou voluntários tanto em projetos de extensão quanto em projetos de pesquisa?
- Os alunos que recebem assistência estudantil também conseguem ingressar na área de pesquisa e pós-graduação?
- Em qual tipo de projeto os alunos contemplados com assistência estudantil atuam mais, em projetos de extensão ou pesquisa?

Todas essas questões serão analisadas nas perspectivas de anos e departamentos/cursos, esperando - se que a partir dos conjuntos de dados seja possível chegar as respostas que a pesquisa propõe. As questões norteadoras foram formuladas com base na questão de assistência estudantil, logo, os conjuntos de dados que não apresentam dados dos bolsistas em questão foram descartados, sendo:

- Cursos de Extensão
- Programas de Extensão
- Projetos de Extensão
- Equipamentos Multiusuário
- Grupos de Pesquisa
- Pesquisadores contemplados com Auxílio Financeiro da PROPP
- Orientadores de Pós - Graduação
- Pesquisadores

5.1.1. Definição do Objeto, Universo e Amostra do Estudo

- **Objeto de Estudo:** São os dados abertos.
- **Universo do Estudo:** São os dados abertos da UFOP que se dividem em três categorias: Assistência Estudantil, Extensão e Pesquisa e Pós - Graduação.
- **Amostra do Estudo:** São os dados existentes dentro do objeto de estudo, sendo a quantidade de conjuntos de dados em cada categoria dos dados abertos da UFOP, totalizando dezoito conjuntos de dados.

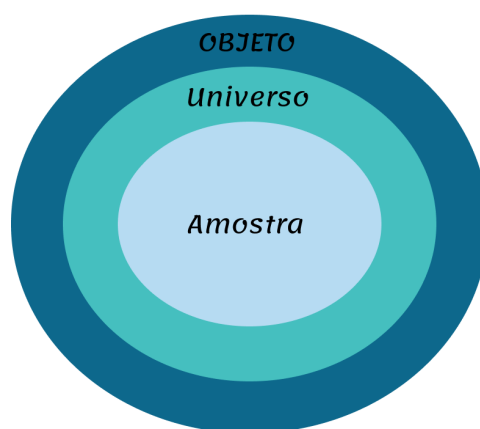


Figura 2. Elaborado pelo autor

5.1.2. Caracterização dos Conjuntos de Dados

Assistência Estudantil: apresenta quatro conjuntos de dados.

Moradia Estudantil: Alunos que foram contemplados com moradia, sendo nos primeiros anos da disponibilização dos dados de maneira pública e transparente no formato PDF e consistiam em apenas o nome dos estudantes. Com o passar do tempo, esses arquivos tiveram seu formato alterado para CSV com os seguintes campos:

Campo	Tipo	Descrição
ID	Inteiro	identificador do aluno contemplado
Nome	Texto	nome completo do estudante contemplado
Curso	Texto	nome do curso do aluno contemplado
Tipo_Moradia	Texto	categoria da moradia estudantil

Observação: Alguns arquivos CSV contém um campo do tipo data chamado "Data de Ingresso" que consiste em descrever a data em que o estudante começou a usufruir da moradia estudantil.

Bolsa Renumerada: Alunos que foram contemplados pela PRACE para manter a permanência na Universidade Federal de Ouro Preto, existindo esse recurso desde 2010. Apresenta arquivos no formato CSV e PDF, com os seguintes campos:

Campo	Tipo	Descrição
Nome	Texto	nome do bolsista
Código_curso	Texto	código do curso do bolsista
Descrição/bolsa	Texto	tipo de bolsa renumerada
Data_inicio	Data	data de início da bolsa
Data_fim	Data	data de término da bolsa

Observação: A partir de 2020, teve a inclusão do campo "Nome Social" e os cursos foram modificados para códigos.

Bolsa Alimentação e Bolsa Transporte: Alunos que foram contemplados com Bolsa Alimentação para a utilização no Restaurante Universitário e / ou com Bolsa Transporte para custear o deslocamento. Apresenta arquivos no formato CSV e PDF, com os seguintes campos:

Campo	Tipo	Descrição
Nome	Texto	nome do bolsista
Código_curso	Texto	código do curso do bolsista
Data_inicio	Data	data de início da bolsa
Data_fim	Data	data de término da bolsa

Observação: A bolsa alimentação é registrada e documentada desde 2003 e teve o campo "Data Fim" removido em 2019 enquanto a bolsa transporte tem seus registros documentados desde 2011.

Extensão: apresenta seis conjuntos de dados.

Bolsistas de Extensão e Voluntários de Extensão: Registros documentados desde 2014 no formato CSV que mostram alunos que recebem bolsas de projetos de extensão e voluntários de projetos de extensão, com os seguintes campos:

Campo	Tipo	Descrição
Nome	Texto	nome do bolsista ou voluntário
Curso	Texto	curso do bolsista ou voluntário
Setor	Texto	setor de lotação do coordenador do curso
Data_inicio	Data	data de início das atividades
Data_Fim	Data	data de término das atividades
Projeto	Texto	título da ação extensionista a qual o aluno está vinculado

Observação: Voluntários e Bolsistas são termos que se diferenciam devido ao bolsista receber um tipo de benefício e o voluntário prestar ações de extensão por interesse próprio sem expectativas de troca de benefícios.

Empresas Juniores: Arquivo CSV único com os seguintes campos:

Campo	Tipo	Descrição
Projeto_EJ	Texto	nome da empresa junior com projeto EJ registrado
Data_inicio	Data	data de reconhecimento do projeto EJ pela PROEX
Professor-orientador	Texto	nome do professor orientador do projeto EJ
Cursos de vinculação da EJ	Texto	nome dos cursos de graduação vinculados ao projeto EJ

Observação: O campo "Data de Início" representa a data de criação da Empresa Junior.

Pesquisa e Pós - Graduação: apresenta oito conjuntos de dados.

Alunos de Pós-Graduação: Arquivo CSV único com os seguintes campos:

Campo	Tipo	Descrição
ID	Inteiro	identificador do aluno
Nome	Texto	nome completo do aluno
Programa	Texto	nome do programa de pós - graduação
Nível	Texto	nível do curso de pós - graduação
Data_Início	Data	data de ingresso na pós - graduação
Data_Defesa	Data	data de defesa na pós - graduação
Orientador	Texto	nome completo do orientador
Situação	Texto	situação atual do aluno

Observação: Os registros são desde 1992 e os dados estão documentados de maneira incompleta, com muitos dados desconhecidos ou não preenchidos.

Bolsistas de Iniciação Científica: Arquivo CSV único com os seguintes campos:

Observação: Os registros documentados são somente do ano de 2017 e do ano de 2018.

Projetos de Pesquisa: Arquivo CSV único com os seguintes campos:

Observação: Os registros documentados são somente do ano de 2018 e do ano de 2019.

6. Conclusões

Campo	Tipo	Descrição
ID	Inteiro	identificador
Bolsista	Texto	nome completo do bolsista
Curso_bolsista	Texto	curso do bolsista
Programa	Texto	programa do bolsista
Ano	Inteiro	ano do projeto de ic
Orientador	Texto	orientador do bolsista
Setor	Texto	setor do bolsista
Linha_Pesquisa	Texto	linha de pesquisa do projeto
ID_projeto	Inteiro	identificador do projeto
Título	Texto	título do projeto de ic
Fomento	Texto	agência de fomento do projeto de ic
Tipo_bolsa	Texto	tipo da bolsa

Campo	Tipo	Descrição
ID	Inteiro	identificador
Título	Texto	título do projeto
Situação	Texto	situação do projeto
Ano	Inteiro	ano de início do projeto
Data_inicio	Data	data de início do projeto
Data_fim	Data	data de término do projeto
Natureza	Texto	natureza do projeto
Equipe_servidor	Texto	servidores envolvidos no projeto
Equipe_aluno	Inteiro	alunos envolvidos no projeto
Equipe_aluno_pos	Texto	alunos de pós-graduação envolvidos no projeto
Equipe_colaborador	Texto	colaboradores envolvidos no projeto

Referências

- A.BEAL (2004). Gestão estratégica da informação: como transformar a informação e a tecnologia da informação em fatores de crescimento e de alto desempenho nas organizações. *Atlas*.
- BRASIL, P. D. R. (1988). Constituição da república federativa do brasil de 1988. Disponível em: <http://www.planalto.gov.br/ccivil03/Constituicao/Constituicao.htm>art5xxxiii.
- BRASIL, P. D. R. (2009). Lei complementar nº 131, de 27 de maio de 2009. Disponível em: <http://www.planalto.gov.br/ccivil03/leis/lcp/lcp131.htm>.
- BRASIL, P. D. R. (2018). Lei número 12.527, de 18 de novembro de 2011 (lei de acesso a informacao – lai). regula o acesso a informacoes previsto no inciso xxxiii do art. 5o, no inciso ii do § 3o do art. 37 e no § 2o do art. 216 da constituicao federal. Disponível em: <http://www.planalto.gov.br/ccivil03/ato2011-2014/2011/lei/12527.htm>.
- Cao, L. (2017). Data science: A comprehensive overview. page 1–42.
- CONWAY, D. (2022). Drew conway official website. Disponível em: <http://drewconway.com/>.

- da Economia, M. (2022). Serviços de informação ao cidadão. Disponível em: <https://www.gov.br/economia/pt-br/canaisatendimento/ouvidoria/sic>.
- E.CARPES and BIANCO, C. (2013). Acesso à informação e transparência na administração pública.
- Gil, A. (2002). Como elaborar projetos de pesquisa. 5:61.
- GONCALVES, E. P. (2001). Iniciação á pesquisa científica. Alínea, Campinas, 2 edition.
- GORDON, S. and GORDON, J. (2006). *Sistemas de Informação: Uma abordagem gerencial*. LTC, Rio de Janeiro, 3 edition.
- GOV.BR (2022). Acesso á informação e ouvidoria. Disponível em: <https://landpage.cgu.gov.br/redirectfalabr/>.
- GRUMAN, M. (2012). Lei de acesso á informação: notas e um breve exemplo.
- Org., P. (2022). What is python? executive summary. Disponível em: <https://www.python.org/doc/essays/blurb/>.
- Provost, F. and Fawcett., T. (2013). Data science and its relationship to big data and data-driven decision making.
- SILVEIRA, D. and CORDOVA, F. (2009). *Unidade 2 - A pesquisa científica*. Editora da UFRGS, Porto Alegre.
- SORDI, J. (2008). *Administração da Informação: fundamentos e práticas para uma nova gestão do conhecimento*. Saraiva, São Paulo.
- UFOP (2021). Ufop dados abertos. Disponível em: <http://www.dados.ufop.br>.
- WHOLIN, C. (2000). *Experimentation in Software Engineering[S.I.]*. Kluwer Academic, London.