

UNIVERSIDADE DE CAXIAS DO SUL
ÁREA DO CONHECIMENTO DE CIÊNCIAS EXATAS E ENGENHARIAS

OTÁVIO VAILATTI SIMIONATO

**CONSTRUÇÃO DE UM BANCO DE DADOS DE INDICADORES
BASEADOS EM DADOS ABERTOS SOBRE ECOSISTEMAS DE
INOVAÇÃO**

CAXIAS DO SUL

2025

OTÁVIO VAILATTI SIMIONATO

**CONSTRUÇÃO DE UM BANCO DE DADOS DE INDICADORES
BASEADOS EM DADOS ABERTOS SOBRE ECOSISTEMAS DE
INOVAÇÃO**

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Bacharel na Área do Conhecimento de Ciências Exatas e Engenharias da Universidade de Caxias do Sul.

Orientador: Prof. Daniel Luís Notari

CAXIAS DO SUL

2025

OTÁVIO VAILATTI SIMIONATO

**CONSTRUÇÃO DE UM BANCO DE DADOS DE INDICADORES
BASEADOS EM DADOS ABERTOS SOBRE ECOSSISTEMAS DE
INOVAÇÃO**

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Bacharel na Área do Conhecimento de Ciências Exatas e Engenharias da Universidade de Caxias do Sul.

Aprovado(a) em __/__/2025

BANCA EXAMINADORA

Prof. Daniel Luís Notari
Universidade de Caxias do Sul - UCS

Prof. Marinês Refosco Severo
Universidade de Caxias do Sul - UCS

Prof. Rafael de Lucena Perini
Universidade de Caxias do Sul - UCS

AGRADECIMENTOS

Agradeço à Universidade de Caxias do Sul e aos professores que contribuíram para minha formação ao longo desta jornada. Em especial, ao professor Daniel Luís Notari, pela orientação, apoio e disponibilidade durante o desenvolvimento deste trabalho. Agradeço também à minha mãe Cleusa, meu pai Alberto, e toda minha família e amigos, pelo incentivo, compreensão e apoio em todos os momentos.

“O sonho é que leva a gente para frente. Se a gente for seguir a razão, fica aquietado, acomodado.”

Ariano Suassuna

RESUMO

A inovação tem se tornado um elemento essencial para o desenvolvimento econômico e social do mundo, tendo os ecossistemas de inovação como ambientes chave para a promoção desse processo. Esses ecossistemas reúnem diferentes atores, como governo, universidades, empresas e sociedade civil, criando um ambiente propício para a geração e difusão de conhecimento. Considerando a importância de mensurar o desempenho desses ambientes, este trabalho tem como objetivo a construção de um banco de dados, com dados extraídos de fontes abertas, sobre os ecossistemas de inovação do Rio Grande do Sul (RS), para análise de indicadores quantitativos. A pesquisa teve como base teórica os conceitos de ecossistemas de inovação, dados abertos e os indicadores de inovação utilizados para realizar esse tipo de medida. A metodologia aplicada foi o processo de Extração, Transformação e Carga (ETL), envolvendo a identificação, coleta, tratamento e armazenamento de dados oriundos de diferentes fontes públicas. O banco de dados foi implementado utilizando o Sistema Gerenciador de Banco de Dados PostgreSQL, e estruturado a partir de um modelo relacional capaz de integrar as informações dos indicadores selecionados. O resultado foi um banco de dados composto de 10 tabelas e 54 colunas totais, contendo mais de 5000 registros ao todo, além de um painel feito utilizando a ferramenta PowerBI, para facilitar a visualização dos 23 indicadores construídos neste trabalho.

Palavras-chave: Ecossistemas de inovação; Indicadores; Dados abertos; Banco de dados.

LISTA DE FIGURAS

Figura 1 – Hélice quádrupla	16
Figura 2 – Hélice quántupla	17
Figura 3 – Posição do Brasil no Índice de Inovação Global (GII) nos anos 2020-2024	20
Figura 4 – Diagrama Entidade Relacionamento (ER) do banco de dados	28
Figura 5 – Consultas teste	29
Figura 6 – Diagrama ER final do banco de dados	39
Figura 7 – Painel com indicadores dos ambientes de inovação	44
Figura 8 – Painel com indicadores das startups	44
Figura 9 – Painel com indicadores das instituições	45

LISTA DE TABELAS

Tabela 1 – Comparativo dos indicadores	25
Tabela 2 – Fontes dos Indicadores de Ecossistemas de Inovação no RS	26
Tabela 3 – Conjunto final de indicadores	31
Tabela 4 – Quantia total de registros em cada tabela	42
Tabela 5 – Total de valores por indicador	42
Tabela 6 – Indicadores do GII	51

LISTA DE SCRIPTS

<i>Script 1</i>	Função para coleta de eventos do portal Inova RS	32
<i>Script 2</i>	Função para extração das datas inicial e final dos eventos	33
<i>Script 3</i>	Função para realizar request ao PowerBI	33
<i>Script 4</i>	Função para tratar dados do painel Observatório de Inovação	34
<i>Script 5</i>	Função para tratar dados do painel Rede RS Startup	35
<i>Script 6</i>	Função para união dos arquivos de instituições	36
<i>Script 7</i>	Funções de normalização e cálculo de similaridade de palavras	37
<i>Script 8</i>	Função para realizar requisições a API do Sucupira	37
<i>Script 9</i>	Função para filtragem dos arquivos por instituições do RS	38
<i>Script 10</i>	Procedure para criar as tabelas do banco de dados	55
<i>Script 11</i>	Código base para leitura dos arquivos CSV em SQL	57
<i>Script 12</i>	Procedure para carregar os ecossistemas de inovação para o banco	58
<i>Script 13</i>	Procedure para carregar dados dos eventos	58
<i>Script 14</i>	Procedure para carregar dados das cidades	59
<i>Script 15</i>	Procedure para carregar dados dos investimentos em PeD	59
<i>Script 16</i>	Procedure para carregar dados das instituições	60
<i>Script 17</i>	Procedure para carregar dados das matriculas	60
<i>Script 18</i>	Procedure para carregar dados das bolsas concedidas	62
<i>Script 19</i>	Procedure para carregar dados dos ambientes	62
<i>Script 20</i>	Procedure para carregar dados das startups	63
<i>Script 21</i>	Procedure para carregar dados das pessoas	64
<i>Script 22</i>	Procedure para carregar dados das produções	66
<i>Script 23</i>	Procedure para carregar dados das empresas de alto crescimento	68
<i>Script 24</i>	Procedure para carregar dados das entidades governamentais	68
<i>Script 25</i>	Script para executar todas as procedures	69

LISTA DE ABREVIATURAS E SIGLAS

CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
ER	Entidade Relacionamento
ETL	<i>Extract, Transform and Load</i>
GII	Índice de Inovação Global
IBGE	Instituto Brasileiro de Geografia e Estatística
ICT	Institutos de Ciência e Tecnologia
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
INPI	Instituto Nacional da Propriedade Industrial
ISO	Organização Internacional de Normalização
MCTIC	Ministério da Ciência, Tecnologia, Inovações e Comunicações
MIT	Instituto de Tecnologia de Massachusetts
OECD	Organização para a Cooperação e Desenvolvimento Econômico
OKFN	<i>Open Knowledge Foundation</i>
OMC	Organização Mundial do Comércio
ONU	Organização das Nações Unidas
PCT	Tratado de Cooperação em Matéria de Patentes
PeD	Pesquisa e Desenvolvimento
PIB	Produto Interno Bruto
PISA	Programa Internacional de Avaliação de Alunos
PME	Pequenas e Médias Empresas
PPC	Paridade do Poder de Compra
PT	Parque Tecnológico
QS	Quacquarelli Symonds
RS	Rio Grande do Sul
SEBRAE	Serviço Brasileiro de Apoio às Micro e Pequenas Empresas
SENAI	Serviço Nacional de Aprendizagem Industrial
SGBD	Sistema de Gerenciamento de Banco de Dados
STEM	Ciência, Tecnologia, Engenharia e Matemática
TCC	Trabalho de Conclusão de Curso
TIC	Tecnologias da Informação e Comunicação
TLD	<i>Top-Level Domain</i>
UCS	Universidade de Caxias do Sul
UIT	União Internacional de Telecomunicações

UNESCO Organização das Nações Unidas para a Educação, a Ciência e a
Cultura

WIPO Organização Mundial de Propriedade Intelectual

SUMÁRIO

1	INTRODUÇÃO	12
1.1	OBJETIVOS	14
1.2	ESTRUTURA DO TRABALHO	14
2	BANCO DE DADOS SOBRE ECOSISTEMAS DE INOVAÇÃO	15
2.1	ECOSISTEMAS DE INOVAÇÃO	15
2.2	DADOS ABERTOS	18
2.3	INDICADORES DE ECOSISTEMAS DE INOVAÇÃO	19
2.4	TRABALHOS RELACIONADOS	20
3	PROPOSTA DE SOLUÇÃO	24
3.1	INDICADORES	24
3.2	EXTRAÇÃO DE DADOS E BANCO DE DADOS	26
4	DESENVOLVIMENTO DA PROPOSTA DE SOLUÇÃO	30
4.1	DIFICULDADES ENCONTRADAS	30
4.2	EXTRAÇÃO E TRANSFORMAÇÃO	31
4.3	CARREGAMENTO DOS DADOS	39
4.4	RESULTADOS DA CARGA DOS DADOS	42
4.5	VISUALIZAÇÃO DOS DADOS	43
5	CONCLUSÃO	46
	REFERÊNCIAS	47
A	INDICADORES DE INOVAÇÃO DA ORGANIZAÇÃO MUNDIAL DE PROPRIEDADE INTELECTUAL	51
B	PROCEDURES PARA CARGA E TRATAMENTO DOS DADOS NO BANCO	55

1 INTRODUÇÃO

A palavra ecossistema é utilizada para designar um sistema ambiental formado por populações de diferentes espécies vivendo em determinada área por um período definido, em conjunto com fatores abióticos como solo, água, atmosfera e luz (SCHWAMBACH; SOBRINHO, 2017) . Um ambiente natural pode ser definido como um ecossistema se esse é autossuficiente e ecologicamente estável, sendo resistente e resiliente, ou seja, capaz de manter sua estrutura e funções caso ocorra algum desarranjo, e também preparado para se reestruturar na possibilidade de ocorrer um forte abalo em sua organização (Secretaria do Meio Ambiente de São Paulo, 2025).

Por outro lado, o conceito de inovação refere-se à criação de produtos (bens ou serviços), processos, ideias ou métodos que sejam capazes de oferecer soluções para problemas já existentes (Sebrae, 2025). Conforme escrito no Manual Oslo, da Organização para a Cooperação e Desenvolvimento Econômico (OECD), criações desse gênero podem ser ditas inovações se e somente se, são totalmente novas, ou, possuem uma significativa melhoria em relação às existentes atualmente, e tenha sido colocado no mercado ou posto em uso por parte do criador, visando o aumento da produtividade ou a geração de riquezas (OECD, 2018).

Inovações podem ser classificadas em três subcategorias em relação a novidade que traz consigo. Pode ser incremental, quando essa caracteriza uma melhoria em algo existente, radical, se for algo capaz de redefinir o mercado, e disruptiva, quando altera totalmente o paradigma vigente, tornando o produto anterior obsoleto (OECD, 2018). Inovar é algo extremamente importante, e precisa ser feito por quem busca encontrar novas soluções, criar novas tecnologias, gerar riquezas e melhorar o estilo de vida da sociedade (AEVO, 2024).

Tendo os conceitos de ecossistema e inovação definidos, pode-se concluir então, que os ecossistemas de inovação são comunidades formadas por diferentes atores, que vão de governos, empresas e universidades, a empresários ou até mesmo a sociedade civil como um todo, os quais trabalham juntos em prol do crescimento de ambientes de Pesquisa e Desenvolvimento (PeD), buscando solucionar demandas não exploradas ou insatisfeitas do mundo moderno (IDIA, 2025). A capacidade de se adequar às constantes mudanças que o mundo e o mercado sofrem, iniciando novas pesquisas e buscando diferentes soluções, mostram também a resiliência e resistência desses ambientes (Cubo Itaú, 2024).

Conforme é colocado no Marco Legal do Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), ecossistemas de inovação são “espaços que agregam infraestrutura e arranjos institucionais e culturais, que atraem empreendedores e recursos financeiros, constituem lugares que potencializam o desenvolvimento da sociedade do conhecimento e compreendem, entre outros, parques científicos e tecnológicos, cidades inteligentes, distritos de inovação e polos tecnológicos”(MCTIC, 2018).

Junto dos ecossistemas de inovação, existem também os ambientes de inovação, lugares que incentivam a troca de conhecimento e o aprendizado coletivo entre empresas, centros de pesquisa,

universidades e governo, com o intuito de fomentar a pesquisa e inovação, para promover o desenvolvimento da economia do local (MINEIRO; SOUZA; CASTRO, 2021). O MCTIC os coloca como espaços que promovem a inovação e o empreendedorismo, formando ambientes típicos da nova economia baseada no conhecimento (MCTIC, 2018). No Brasil, esses locais estão dispostos como Parque Tecnológicos (PTs), *hubs*, aceleradoras, incubadoras, *smart farms*, entre outros, bem como espaços virtuais advindos da transformação digital (Ministério da Agricultura e Pecuária, 2025).

O termo PeD é conceituado como os trabalhos criativos e sistemáticos a fim de aumentar o saber humano cultural e social, mas também para se encontrar novas aplicações para o conhecimento atual disponível, onde tais atividades podem ser direcionadas a objetivos específicos ou gerais, porém, sempre buscando novas descobertas (OECD, 2015). O investimento na PeD contribui de grande forma para o desenvolvimento de novos artifícios e melhores tecnologias, que abrem caminho para a implementação de inovações em produtos e serviços (ROCHA *et al.*, 2016).

A OECD também classifica, no Manual Frascati, a PeD de três formas distintas, sendo elas, pesquisa básica, pesquisa aplicada, e desenvolvimento experimental. Na básica, são colocadas as investigações que buscam gerar conhecimento sem foco em uma aplicação prática. De outro modo, a pesquisa aplicada já coloca a prática em pauta como forma de buscar o conhecimento. E por fim, o desenvolvimento experimental usa da pesquisa e experiência para se chegar nas novidades ou melhorias dos produtos e processos (OECD, 2015).

A análise desses ecossistemas é crucial para seu fortalecimento e para manter a inovação em constante desenvolvimento no mundo. Os dados abertos são uma grande fonte de informações para serem utilizados para mensurar ecossistemas, e a análise deles pode ser feita para os entender melhor. O termo refere-se a uma metodologia voltada para a divulgação de dados em formatos reutilizáveis, buscando promover a transparência, o engajamento político dos cidadãos e a criação de soluções pela sociedade (Governo Digital - Brasil, 2025).

A utilização desses dados pode beneficiar diferentes partes da sociedade. O compartilhamento deles com o coletivo social objetiva aumentar a transparência do funcionamento público para o cidadão, além de visar um maior engajamento político e em projetos sociais do mesmo (Governo Digital - Brasil, 2025). Junto disso, serviços públicos podem utilizá-los para regular seus gastos, por exemplo, por meio do compartilhamento de conhecimento e conteúdos (Data Europa Academy, 2025). Essa troca de informações pode ser utilizada também pelos ecossistemas de inovação, fazendo com que seja possível conceber ideias inovadoras e a criação de novos modelos de negócio (RUNESON; OLSSON; LINÅKER, 2021).

Atualmente, existem centenas de ecossistemas espalhados pelo mundo, com mais de 250 desses localizados na Europa (EIE Datahub, 2025). Também, no Brasil, existem 64 parques tecnológicos em operação, com programação de aumentar esse número para 112, entre novos locais, que estão em estado de implementação ou planejamento (InovaData-BR, 2025).

Apesar disso, estudar os dados desses é algo bastante complexo, tendo em vista que eles muitas vezes não são de fácil acesso, e estão espalhados por diversos lugares, como portais governamentais, relatórios de instituições, bases de dados acadêmicas e plataformas de organizações pri-

vadas, podendo serem apresentados de várias formas diferentes (PDFs, planilhas, APIs, etc). Esses problemas fazem com que o estudo dessas informações seja dificultado, gerando dificuldades para pesquisadores e gestores desses ecossistemas, uma vez que a análise desses dados é essencial para entender como o ambiente está desempenhando e se desenvolvendo, assim como identificar novas oportunidades de crescimento e inovação.

1.1 OBJETIVOS

O presente trabalho tem como objetivo a criação de um banco de dados por meio do levantamento e organização de dados abertos sobre os ecossistemas de inovação.

Os objetivos específicos do trabalho são:

1. Identificar e mapear fontes de dados abertos.
2. Coletar e tratar de forma correta estes dados.
3. Projetar e implementar um banco de dados para armazenar e relacionar as informações.
4. Construir consultas para conseguir extrair o conhecimento necessário.

1.2 ESTRUTURA DO TRABALHO

Este projeto está organizado em 4 capítulos. O segundo capítulo apresenta a revisão bibliográfica, além dos conceitos fundamentais para o trabalho. O Capítulo 3 aborda a proposta de solução do trabalho. O capítulo 4 mostra o desenvolvimento do banco de dados, junto dos resultados obtidos. Por fim, o capítulo 5 apresenta as conclusões do trabalho.

2 BANCO DE DADOS SOBRE ECOSISTEMAS DE INOVAÇÃO

Inovação é um termo bastante complexo, sem uma definição única entre quem debate o tema (KOGABAYEV; MAZILIAUSKAS, 2017). A inovação é o processo de repensar ideias e tecnologias, a fim de transformá-los em bens ou serviços (PAIVA *et al.*, 2018). O ato de inovar é potencializado pelo investimento em PeD e capital humano, além do apoio de empresas, e instituições estudantis ou do governo (KOCHETKOV, 2023).

Para Schumpeter (1883-1950), inovação e economia estão fortemente ligadas, afirmando que inovação é o mecanismo primordial para a evolução do capitalismo. O autor também expressa a inovação como o centro das mudanças que ocorrem no sistema econômico, colocando ela como um eterno processo de mutação industrial, destruindo o antigo e criando o novo (ŚLEDZIK, 2013; PAIVA *et al.*, 2018). A concepção de novos produtos, bem como a formação de novos mercados que ocorrem por conta da inovação, corroboram, também, para aumentar a diversidade e resistência da estrutura econômica (KOGABAYEV; MAZILIAUSKAS, 2017).

Além da economia, a inovação tem grande influência sobre a sociedade como um todo, através da criação de novos meios para atender necessidades sociais, promover inclusão, e melhorar a qualidade de vida das pessoas (DOMANSKI; HOWALDT; KALETKA, 2020). Um exemplo recente é a pandemia da COVID-19, onde ocorreram investimentos pesados em inovação, o que foi peça chave para a evolução de tecnologias medicinais, e conseqüentemente, surgimento de novos métodos para tratamento da doença (LEPORE; FRONTONI; MICOZZI, 2023).

A união de diferentes atores e mecanismos foi peça chave para a evolução da sociedade, através do incentivo a inovação para alcançar novas ideias, e, com base nisso, foi cunhado o termo ecossistema de inovação, locais que possuem grande importância para a evolução econômica e tecnológica mundial (GRANSTRAND; HOLGERSSON, 2019). A existência deles gera uma saudável competitividade do sistema industrial e de empresas, além de melhorar o modelo produtivo como um todo (SPINOSA; KRAMA; HARDT, 2018).

O capítulo atual tem como objetivo aprofundar os principais conceitos deste trabalho. O capítulo é dividido em 4 seções distintas, onde a primeira aprofunda o conceito de ecossistemas de inovação, a segunda o conceito de dados abertos, a terceira com indicadores de inovação e de ecossistemas de inovação, e a quarta seção é feita uma análise de outros trabalhos relacionados a esse.

2.1 ECOSISTEMAS DE INOVAÇÃO

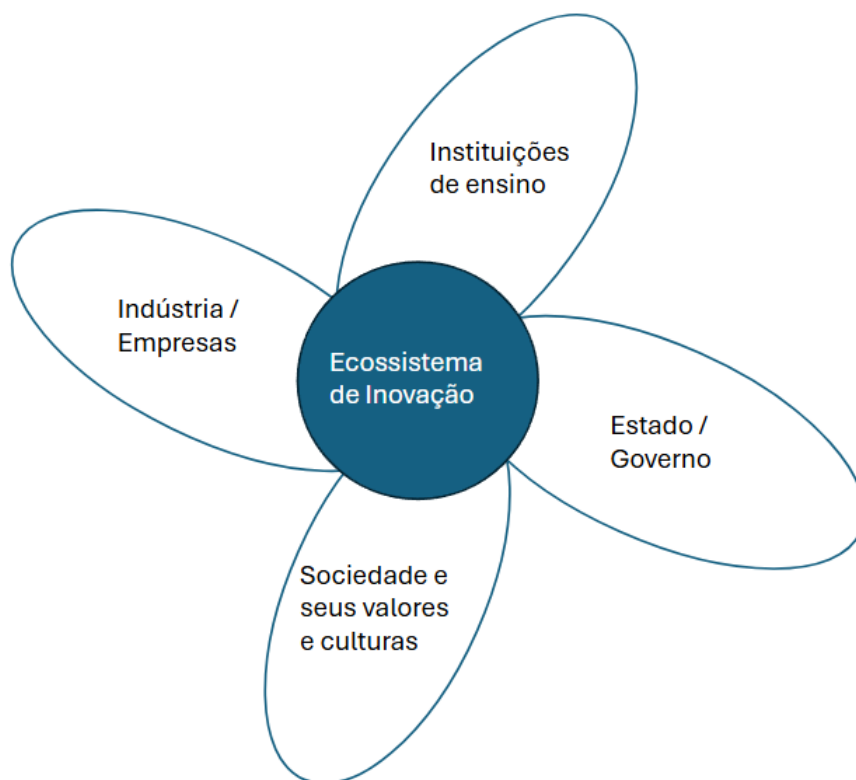
Em seu estudo no Japão, Freeman (1987) observou que bancos, empresas privadas, e o Ministério do Comércio Internacional e Indústria, além de outros órgãos do governo, estudavam possíveis mudanças futuras da sociedade, e trabalhavam em conjunto para atingir objetivos de longo prazo (FREEMAN, 1987). Com isso, Freeman definiu o conceito de sistemas nacionais de inovação, referindo-se a essa interação entre empresas, governo e universidades para impulsionar a inovação (FREEMAN,

1987).

Granstrand e Holgersson (2020) propõem um conceito baseado na tríade formada por atores, artefatos e instituições, referenciando um ecossistema de inovação como um “conjunto em constante evolução de atores, artefatos e instituições, as atividades e relações, incluindo complementares e substitutas, as quais são importantes para a performance inovativa de um ator”.

Em sua tese, Perini (2023) trabalha os conceitos de hélices propostas por Carayannis e Campbell. A ideia da hélice quádrupla coloca o ecossistema de inovação organizado em uma hélice de quatro pás (ver Figura 1), uma expansão da tríplice hélice, unindo universidade, indústria, governo, mais a sociedade civil, e o contexto onde ela está inserida, como novo ator, retratando a importância do corpo social para a inovação, afim de promover o conhecimento para a sociedade do conhecimento. Os ecossistemas são ditos então como conjuntos de sistemas complexos, multiagente e em constante evolução, onde os diferentes atores atuam em coopetição, evoluindo e especializando-se constantemente em prol da inovação (PERINI, 2023).

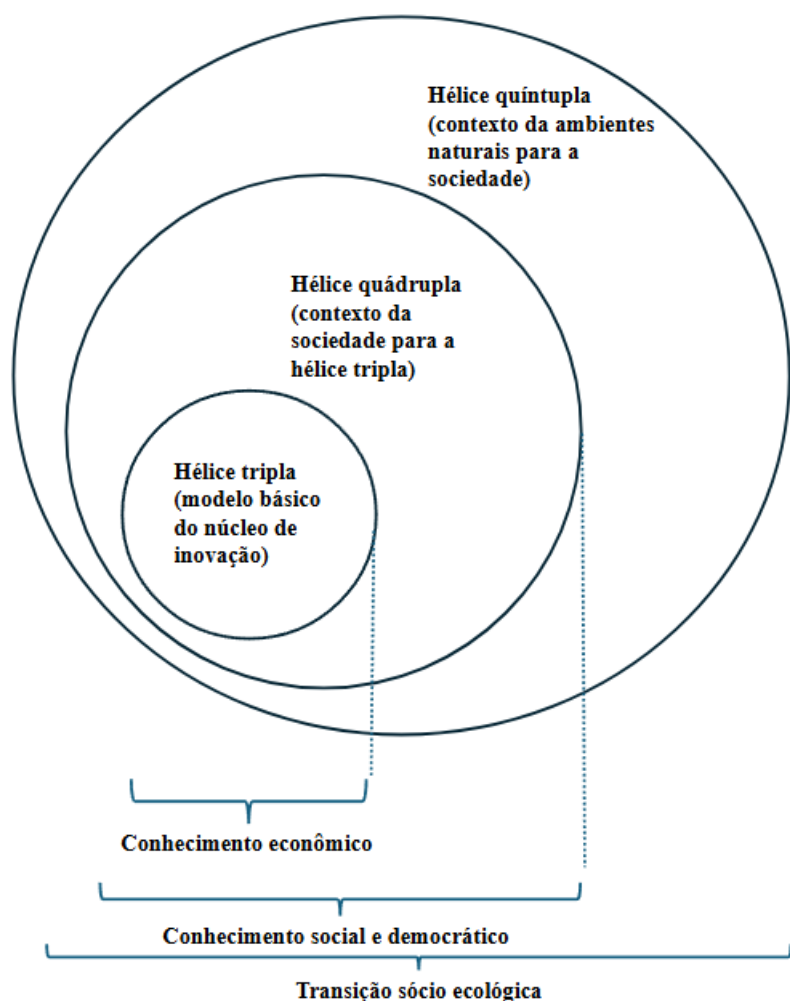
Figura 1 – Hélice quádrupla



Fonte: O autor (2025)

Perini (2023) também aborda o modelo da hélice quintupla, proposto por Carayannis e Campbell (2012), com o novo componente fundamental para a inovação sendo o meio ambiente, tendo sua inclusão motivada pelas mudanças climáticas que o planeta vem sofrendo. O objetivo dessa alteração é colocar a natureza em foco, para que a inovação seja voltada, também, para a preservação da mesma, com o desenvolvimento de novas tecnologias verdes e sustentáveis (PERINI, 2023; CARAYANNIS; BARTH; CAMPBELL, 2012). A Figura 2 demonstra os modelos de hélice tripla, quádrupla e quintupla, bem como as áreas de conhecimento que cada uma traz a um ecossistema de inovação.

Figura 2 – Hélice quántupla



Fonte: O autor (2025), com base em Carayannis e Campbell (2012)

Por ser um conceito bastante complexo e abrangente, diversos ambientes de inovação podem fazer parte de um único ecossistema de inovação. PTs, *smart cities*, e diferentes ambientes de inovação são locais que promovem a inovação, tecnologia e empreendedorismo entre diferentes atores que participam deles (Ministério da Agricultura e Pecuária, 2025).

PTs são organizações que reúnem empresas, instituições de ensino e Institutos de Ciência e Tecnologias (ICTs), com o objetivo de , através de projetos e ações estruturadas, promover a inovação, ciência, empreendedorismo e tecnologia (Anprotec, 2025; DUARTE; OUTROS, 2025). Os PTs possuem dois elementos fundamentais, a presença de uma entidade gestora, que coordena o ambiente, e a existência de um espaço físico para que os atores membros possam trabalhar de forma integrada e cooperada em projetos e atividades. Na Serra Gaúcha, destaca-se a presença do TecnoUCS, que objetiva levar o conhecimento feito na Universidade de Caxias do Sul para o poder público, mundo empresarial e a sociedade, com projetos, serviços e negócios inovadores (TecnoUCS , 2025).

Outro tipo de ecossistema de inovação, as cidades inteligentes são cidades em que o uso de soluções digitais busca tornar redes e serviços tradicionais mais eficientes, para favorecer o crescimento econômico e melhorar a qualidade de vida de seus habitantes (Comissão Europeia, 2025).

Os ecossistemas de inovação englobam também outros tipos de ambientes de inovação, tais como incubadoras de empresas, *hubs* de inovação, e aceleradoras de *startups*. As incubadoras oferecem assistência a empreendedores, a fim de que eles consigam desenvolver novas ideias e transformá-las em negócios de sucesso (BUSATO, 2020). Por outro lado, as aceleradoras focam em empresas emergentes, as *startups*, já que essas surgem com ideias e produtos inovadores ao mercado, e, com o apoio das aceleradoras, conseguem ter uma esfera propícia para o seu desenvolvimento (BUSATO, 2020). Já as *hubs* são ambientes que focam na interação entre os atores do ecossistema de inovação, com o objetivo de gerar negócios e ideias novas, ousadas e transformadoras (Sebrae, 2025).

O Brasil conta com vários ecossistemas de inovação bastante desenvolvidos. Além do exemplo do TecnoUCS, o RS tem ainda a presença do Tecnopuc em Porto Alegre, organizado pela PUCRS, e que é formado por 290 organizações, e apoia mais de 1000 *startups*, com 6500 pessoas envolvidas (Tecnopuc, 2025). Com sede em Recife, o Porto Digital, maior distrito de inovação da América Latina, conta com mais de 470 empresas e 21 mil colaboradores, e possui uma unidade localizada em Portugal para conectar a inovação brasileira com o mercado europeu (Porto Digital, 2025).

O acesso à informação é um fator essencial para que os ecossistemas de inovação sigam se desenvolvendo, de forma sustentável e eficiente. Nesse ponto, os dados abertos servem como base para isso, oferecendo informações que auxiliam na tomada de decisões, promovem a colaboração, e estimulam a geração de novas ideias dentro desses ambientes.

2.2 DADOS ABERTOS

Dados abertos são definidos como as informações ou conteúdos de livre acesso público, para as pessoas usarem, reutilizarem e redistribuírem, sem restrições legais, sociais ou tecnológicas (OKFN, 2025). A *Open Knowledge Foundation* (OKFN) cita três fatores fundamentais para dados serem caracterizados como abertos, são eles:

1. Disponibilidade e acesso: A informação deve ser disponibilizada na íntegra, e sem barreiras para o seu acesso.
2. Reutilização e redistribuição: Os dados devem ser passíveis de reutilização e redistribuição, além de serem interoperáveis.
3. Participação universal: Todos podem ter acesso aos dados, sem discriminar o usuário ou os fins para os quais serão usados.

No Brasil, a Lei de Acesso à Informação (2011) estabeleceu o direito de qualquer pessoa de solicitar e receber informações sobre qualquer órgão público do governo. Adicionalmente, a criação da Infraestrutura Nacional de Dados Abertos (2012) como unidade central no controle de iniciativas sobre dados abertos no governo federal, estabeleceu um ecossistema funcional para os dados abertos (Governo Federal, 2025).

O Decreto Nº 8.777/2016, consolidou a Política de Dados Abertos, estabelecendo como alguns de seus objetivos, o fomento à inovação, ao desenvolvimento de novas tecnologias, a pesquisa científica e o surgimento de novos negócios. Os dados abertos são definidos no decreto como os “dados acessíveis ao público, representados em meio digital, estruturados em formato aberto, processáveis por máquina, referenciados na internet e disponibilizados sob licença aberta que permita sua livre utilização, consumo ou cruzamento, limitando-se a creditar a autoria ou a fonte” (Governo Federal, 2025).

Além das leis e decretos, a criação do Portal Brasileiro de Dados Abertos, em 2012, facilitou o acesso a informações sobre órgãos governamentais diversos, através da construção de um catálogo com mais de 14 mil conjuntos de dados abertos detalhados e organizados (Governo Digital - Brasil, 2025). Existem ainda outros exemplos de plataformas que disponibilizam dados para uso recreativo, de diferentes áreas, como o Lattes e o Sucupira, focados em projetos de estudo e pesquisa, e produção científica, ou o site do Instituto Brasileiro de Geografia e Estatística (IBGE), que dispõe de dados populacionais e sociais, como atividade econômica, demografia e educação (Governo Digital - Brasil, 2025; CAPES, 2025).

A utilização de dados abertos beneficia não somente à sociedade, mas também a empresas e órgãos de administração pública (MACEDO; LEMOS, 2021). O estudo dos dados abertos pode resultar em melhorias de serviços públicos, tanto em qualidade, como em eficiência e transparência, ou na redução de gastos pelo governo, além disso, o crescimento da economia e de empregos criados também é perceptível (Data Europa EU, 2025). A participação dos cidadãos em políticas públicas também se torna mais fácil, uma vez que eles possuem acesso aos dados e podem contribuir para as ações que melhor atendem às suas necessidades (Data Europa EU, 2025).

Os dados abertos podem desempenhar um papel fundamental na realização de pesquisas. No caso dos ecossistemas de inovação, é possível utilizá-los para a construção de indicadores, afim de mensurar o nível de inovação e desenvolvimento desses sistemas.

2.3 INDICADORES DE ECOSISTEMAS DE INOVAÇÃO

Indicador é um termo usado para designar uma variável utilizada com fins de mensuração e promoção de atividades avaliativas (PRADO; CASTANHA, 2020). De forma geral, indicadores são uma forma de representar medidas quantitativas e qualitativas de processos e atividades, através de uma simplificação de dados brutos, dimensionando propriedades específicas, e designando parâmetros para comparar unidades de análise semelhantes ao longo do tempo e localização (PRADO; CASTANHA, 2020; OECD, 2018).

O Manual Oslo da OECD (2018), coloca um indicador de inovação como uma medida de um fenômeno inovativo observado em uma amostra de determinado tempo ou local, normalmente padronizados para permitir comparações. A utilidade dos indicadores de inovação é vista tanto em questões de discussão políticas e sociais, quanto no campo empresarial, influenciado o comportamento de empresários. Ademais, os indicadores de inovação são bastante importantes para a análise

exploratória das atividades de inovação, acompanhar tendências, e avaliar o desempenho da inovação em diferentes níveis (OECD, 2018).

A Organização Mundial de Propriedade Intelectual (WIPO) faz uso de 78 indicadores, que incluem critérios sobre política, educação, infraestrutura e criação de conhecimento, para medir a inovação de economias da forma mais completa possível (ver Apêndice A). A Figura 3 mostra como o Brasil ficou ranqueado no GII, organizado pela WIPO, em 2024, e nos quatro anos anteriores. É possível perceber um significativo aumento na capacidade inovativa do país desde 2022, ficando na 50ª posição dentre 133, além de ter se tornado mais eficiente os recursos inovativos em resultados de fato. Porém, pode-se constatar uma estagnação com relação aos inputs de inovação, mostrando que há espaço para melhora nesse quesito (World Intellectual Property Organization (WIPO), 2025).

Figura 3 – Posição do Brasil no GII nos anos 2020-2024

Ano	Posição GII	Recursos Inovativos	Produções Inovativas
2020	62°	59°	64°
2021	57°	56°	59°
2022	54°	58°	53°
2023	49°	59°	49°
2024	50°	58°	49°

Fonte: O autor (2025), com base em dados fornecidos pela WIPO

Um importante fator para a evolução dos ecossistemas de inovação a presença de incubadoras e ambientes de inovação no ecossistema. Lideranças de diferentes ecossistemas de inovação do país, comentam sobre como a importância das regiões do Brasil e a influência delas sobre os ecossistemas, a presença de *startups* e *startups* decorrentes de projetos acadêmicos, e as quantidades totais de pessoas, mestres e doutores, empresários e estudantes, são também, fatores consideráveis ao analisar um ecossistema (Ministério da Ciência, Tecnologia e Inovação, 2024).

Para ampliar a percepção sobre indicadores de inovação, será feita uma análise de outros artigos e pesquisas, que relacionam-se com o objetivo deste trabalho. Dessa maneira, busca-se compreender quais indicadores são utilizados de forma mais frequente em estudos sobre ecossistemas de inovação.

2.4 TRABALHOS RELACIONADOS

Esta seção do trabalho tem como foco analisar outros estudos realizados sobre indicadores de ecossistemas de inovação, quais os mais utilizados, e quais as dificuldades encontradas. Os trabalhos observados aqui foram retirados de diversas fontes, através de pesquisas na plataforma *Google Scholar*, com as palavras chave “ecossistemas de inovação”, “inovação”, “indicadores”, e seus referentes em inglês.

No geral, os trabalhos concordam que existe uma grande dificuldade em realizar um estudo sobre os ecossistemas de inovação, tendo em vista que os conceitos e atividades de inovação são bastantes complexas, além de estarem em constante mudança.

La Rovere, Santos e Vasconcellos (2021), trabalham em cima de indicadores de ecossistemas de inovação brasileiros, e relatam a complexidade de concretizar o estudo no país. A falta de consenso referente aos conceitos e dimensões dos ecossistemas, bem como obstáculos para se conseguir os dados necessários são fatores que atrapalham (ROVERE; SANTOS; VASCONCELLOS, 2021). Valotto e May (2020), e Beaudry e Solar-Pelletier (2020) concordam com o quesito das definições inconsistentes e a falta de padronização de métricas, o que acaba por complicar a identificação de como esses ecossistemas auxiliam no impulsionamento da inovação (VALOTO; MAY, 2020; BEAUDRY; SOLAR-PELLETIER, 2020).

Cobben, Ooms e Roijackers (2023) focam mais em indicadores que possam medir a saúde e o sucesso dos ecossistemas de inovação, relatando o pouco entendimento que se tem para conceituar e mensurar esses campos (COBBEN; OOMS; ROIJACKERS, 2023). Por outro lado, Kosorukova et al. (2023) concentraram seu estudo em uma avaliação comparativa dos indicadores de ecossistemas de inovação da região asiática, buscando encontrar quais fatores aparentam ser mais favoráveis para desenvolvimento desses ecossistemas. O trabalho de Kosorukova et al. (2023) não exemplifica dificuldades encontradas na mensuração dos indicadores, mas coloca a importância de entender as conexões existentes entre os ecossistemas ao realizar comparações (KOSORUKOVA *et al.*, 2023).

No artigo de Cukier, Kon e Krueger (2015), os autores debatem a dificuldade de comparar *hubs* tecnológicos, dadas as diferenças existentes na maturidade e na evolução entre eles. Com isso em mente, estudam os fatores que têm grande influência nos ecossistemas, e criam um modelo para medir a maturidade e a evolução deles (CUKIER; KON; KRUEGER, 2015). Silva (2023) constrói um modelo para medição de desempenho dos ecossistemas, porém, para isso, baseia-se na capacidade de inovação dos atores de cada ecossistema, avaliando quais aspectos possuem maior influência, e quais agentes têm uma maior atuação dentro do ecossistema (SILVA, 2023).

No quesito de indicadores utilizados, La Rovere, Santos e Vasconcellos (2021) fazem uso de métricas na esfera nacional e regional. Os autores colocam parâmetros consolidados na literatura em conjunto com medidas oficiais de pesquisas por amostras, como remuneração da mão de obra com educação superior, número de mestres e de doutores por cem mil habitantes, gastos em pesquisa e desenvolvimento (PeD) como porcentagem do Produto Interno Bruto (PIB), número de licenciamentos, de patentes e de empresas de alto crescimento, parcela de *venture capital* no financiamento a empresas de base tecnológica, a capacidade inovadora de diversos setores industriais de determinado país, os investimentos realizados em inovação, o percentual de empresas com redes de cooperação estabelecidas (ROVERE; SANTOS; VASCONCELLOS, 2021).

La Rovere, Santos e Vasconcellos (2021) apontam também o mapeamento de estruturas de apoio (incubadoras, aceleradoras, PTs, etc.), e os aspectos como aversão ao risco, percepção de fracasso e tolerância à diversidade como possíveis métricas. Para a questão dos *outputs* de inovação, os autores sugerem o mapeamento de empresas de alto crescimento, e de *startups* que foram criadas e se

mantiveram ativas no ecossistema de inovação (ROVERE; SANTOS; VASCONCELLOS, 2021).

O artigo de Valotto e May (2020) tem enfoque maior em indicadores de *startups* que visam entrar em ecossistemas de inovação, trazendo um estudo das métricas de inovação, ecossistemas de inovação, e *startups*. O trabalho coloca, como principal medida de mensuração da inovação, a existência de PeD e os assuntos ligados a ela, como gastos, participação em programas públicos de suporte à inovação, formação educacional e participação em projetos em conjunto com instituições acadêmicas e de pesquisa (VALOTO; MAY, 2020).

Beaudry e Solar-Pelletier (2020) realizaram um estudo com enfoque no projeto do governo canadense *The Innovation Superclusters Initiative*, que promove a criação de *superclusters* para promover a inovação no país. As autoras indicam que, por conta das características desses *superclusters*, eles podem ser classificados como ecossistemas de inovação. No que diz respeito aos indicadores, o artigo critica indicadores que define como genéricos, número de empresas, novos produtos, processos e empregos, fáceis de medir, mas que não possuem um impacto real nas análises. Por isso, as autoras indicam a necessidade de se construir novas métricas que se encaixem melhor na realidade dos ecossistemas de inovação (BEAUDRY; SOLAR-PELLETIER, 2020).

O estudo realizado por Cobben, Ooms e Roijackers (2023) propõe a medição da capacidade absorptiva, robustez, produtividade e a inovação nos diferentes níveis dos ecossistemas de inovação como maneira de mensurar a saúde e o sucesso dos mesmos. Ainda, os autores sugerem a utilização e medição de oportunidades de co-criação, e de antecedentes como governança, estrutura organizacional, resiliência e fluidez, focando em dimensões como comunicação, alinhamento de valor, contribuição, reputação e satisfação dos membros, e diferenças culturais (COBBEN; OOMS; ROIJACKERS, 2023).

Kosorukova et al. (2023) compararam, em seu estudo, os indicadores de ecossistemas de inovação em cinco países asiáticos (Coréia do Sul, Japão, Índia, China, e Cazaquistão). Os indicadores utilizados relacionam-se em sua maioria com o nível de atividade de inovação e o desenvolvimento qualitativo dos ecossistemas de inovação. Os autores comentam que, para realizar a comparação, foi criada uma pontuação com base nos diferentes indicadores, que incluía indicadores quantitativos (*startups*, espaços colaborativos, aceleradoras, *startups* influentes e eventos relacionados a *startups*) e qualitativos (base de clientes, políticas governamentais, empresas unicórnio, marcas globais de *coworking* e uma quantidade crítica de *startups*) (KOSORUKOVA *et al.*, 2023).

Cukier, Kon e Krueger (2015) colocam a educação em foco, dizendo ser um fator muito importante para os ecossistemas. Os autores falam que o incentivo à educação e ao empreendedorismo, além do livre acesso a esses recursos por todos, são coeficientes significativos na mensuração da maturidade do ecossistema. Além disso, a cultura do próprio ecossistema e a qual ele está inserido, e a interconectividade entre os atores dele, configuram fatores que indicam evolução do ecossistema. Investimentos e financiamentos, eventos, presença e número de PTs, incubadoras e aceleradoras, oportunidades de mercado para novas empresas, e número de empregos gerados, são outros aspectos que Cukier, Kon e Krueger (2015) colocam como indicadores de evolução (CUKIER; KON; KRUEGER, 2015).

Em sua tese, Silva (2023), abordou um estudo diferente, buscando avaliar um ecossistema de

inovação “de baixo para cima”, mensurando primeiro os agentes participantes desses ecossistemas. Com isso, o autor chegou a conclusão da existência de 7 atores principais: *startups*, Pequenas e Médias Empresas (PMEs), Incubadoras, PTs, Universidades, Indústria e Governo, e, construiu um conjunto de 227 indicadores, elencados em 12 dimensões, para realizar as medições, tanto quantitativas, quanto qualitativas (SILVA, 2023).

Ao avaliar o desempenho do ecossistema, Silva (2023) utilizou do método SWARA, onde é necessário definir pesos diferentes para cada agente. Dessa forma, o autor definiu as universidades como o ator mais influente, seguido pelos PTs, e, após, as incubadoras. A ordem de importância segue com indústria, *startups*, governo e PMEs (SILVA, 2023).

A análise dos trabalhos relacionados revela uma grande influência de dados relacionados à educação na mensuração da qualidade e maturidade de um ecossistema de inovação, como evidenciado por Cukier, Kon e Krueger (2015), e Silva (2023) em seus trabalhos. A presença e número de incubadoras, aceleradoras e PTs, é algo comentado por todos os autores, ressaltando a importância dessas estruturas para um ecossistema de inovação, e mostrando como podem ser bastante úteis para avaliá-los. Junto dessas estruturas, outro grupo mencionado por todos, são as *startups*, invariavelmente ligadas às estruturas de apoio, e que podem indicar o nível de evolução e inovação dos ecossistemas.

O tema da PeD também é posto em destaque por Valotto e May (2020) e La Rovere, Santos e Vasconcellos (2021), como maneira de medir a produção e retenção do conhecimento. Juntamente, na definição dos ecossistemas de inovação, desde o modelo da tríplice hélice até a quádrupla hélice, o governo é um ator principal, portanto, a participação de agentes governamentais, ou políticas de investimento em inovação, também se caracterizam como bons indicadores.

Por seu trabalho se estender por cinco anos e consistir na construção de novos indicadores para os ecossistemas de inovação, algo que até o momento, não foi concluído, Beaudry e Solar-Pelletier (2020) apenas apontam fatores importantes para um ecossistema, como a troca de informações e conhecimento, e as conexões entre os agentes do ecossistema, assim como Cobben, Ooms e Roijackers (2023), que seguem o mesmo pensamento.

Com base nessa análise, nos comentários dos líderes das plataformas de inovação brasileiras, e em conjunto as definições de indicadores de inovação fornecidas pela OECD e WIPO, chegou-se a um conjunto de 30 indicadores de medidas quantitativas para este trabalho, e que serão aprofundados no capítulo seguinte.

3 PROPOSTA DE SOLUÇÃO

O capítulo atual trata da proposta de solução para este trabalho. A solução irá focar em ecossistemas de inovação da região sul, mais especificamente do RS. Então, para ser possível fazer a medição desses ecossistemas de inovação, será criado um banco de dados para armazenar as informações necessárias, que serão retiradas de fontes de dados abertos. Ao longo desta seção, serão apresentados os indicadores que vão ser utilizados no trabalho, as possíveis fontes de onde os dados serão retirados e como isso será feito, e uma exemplificação do modelo de banco que será construído. Ao final, será apresentado o cronograma para o Trabalho de Conclusão de Curso (TCC) II.

3.1 INDICADORES

Para realizar a mensuração dos atributos dos ecossistemas de inovação, foi selecionado um grupo de trinta indicadores, baseados nos aspectos discutidos pelos trabalhos relacionados, e pelos indicadores já utilizados pela WIPO e OECD. A Tabela 1 mostra um comparativo dos indicadores do atual trabalho, com os utilizados, ou discutidos, pelos trabalhos correlatos.

Ao analisar a tabela, é possível perceber a importância dada à produção de conhecimento, bem como o compartilhamento dela, algo evidenciado pelos indicadores de "número de eventos de inovação realizados", "número de pesquisadores" e "número de programas de internacionalização de *startups*", com seis citações cada. Além disso, os indicadores "parcerias universidade-empresa" e "número de mestres e doutores por 100 mil habitantes" e "número de pesquisas e publicações científicas por instituição", foram mencionadas cinco vezes, e também corroboram para demonstrar a importância dada à troca e geração de informações.

Os indicadores "número de *startups* ativas / empresas inovadoras", "número de aceleradoras" e "número de *startups* exportadoras ou com atuação internacional", referenciadas 5 vezes entre os trabalhos, evidencia a relevância que as *startups*, e o desenvolvimento delas, têm dentro do ecossistema de inovação.

Por outro lado, apesar da presença das estruturas serem bastante referenciadas, os indicadores de número de *startups* incubadas, aceleradas e em PTs são pouco citados, aparecendo uma, duas e três vezes, respectivamente, entre todos os trabalhos. As outras métricas pouco citadas pelos artigos são "número de agentes do governo", e o "setor predominante das *startups*", os quais são nomeados somente duas vezes.

Tabela 1 – Comparativo dos indicadores

Indicadores	Kosorukova et al. (2023)	La Rovere, Santos e Vasconcelos (2021)	Valotto e May (2020)	Cukier, Kon e Krueger (2015)	Silva (2023)	Cobben, Ooms e Roijakkers (2023)	Beaudry e Solar-Pelletier (2020)
Localização	✓	✓			✓		
Gastos com PeD	✓	✓	✓		✓		
Número de universidades				✓	✓	✓	✓
Número de mestres e doutores por 100 mil habitantes			✓	✓	✓	✓	✓
Número de pesquisadores	✓		✓	✓	✓	✓	✓
Empresas de alto crescimento		✓		✓	✓		
Parcerias universidade - empresas		✓		✓	✓	✓	✓
Número de projetos em andamento					✓	✓	✓
Número de alunos em cursos de graduação e pós na área de exatas					✓	✓	✓
Número de bolsas de iniciação científica e tecnológica					✓	✓	✓
Presença de agentes governamentais		✓	✓		✓		
Número de agentes governamentais		✓	✓		✓		
Volume de investimento público em inovação		✓	✓	✓	✓		
Recursos captados por leis de incentivo			✓	✓	✓		
Volume de investimento anjo e capital de risco		✓	✓	✓	✓		
Número de <i>startups</i> ativas / Empresas inovadoras	✓	✓	✓	✓	✓		
Valor de mercados das <i>startups</i>		✓	✓		✓		
Número de incubadoras		✓	✓	✓	✓		
Número de <i>startups</i> incubadas		✓	✓		✓		
Número de aceleradoras	✓	✓	✓	✓	✓		
Número de <i>startups</i> aceleradas		✓		✓	✓		
Número de PTs		✓		✓	✓	✓	✓
Número de empresas instaladas em PTs					✓	✓	✓
Setor predominante das <i>startups</i>		✓			✓		
Número de <i>startups</i> decorrentes de projetos acadêmicos		✓	✓		✓		
Número de patentes registradas			✓		✓		
Número de pesquisas e publicações científicas por ano			✓	✓	✓	✓	✓
Número de eventos de inovação realizados	✓	✓	✓		✓	✓	✓
Número de programas de internacionalização de <i>startups</i>	✓	✓		✓	✓	✓	✓
Número de <i>startups</i> exportadoras ou com atuação internacional		✓		✓	✓	✓	✓

3.2 EXTRAÇÃO DE DADOS E BANCO DE DADOS

Para realizar a obtenção dos dados dos ecossistemas de inovação, será realizado o processo de *Extract, Transform and Load (ETL)*, ou em português Extrair, Transformar e Carregar. O processo refere-se à obtenção de informações através da combinação de diversas fontes de dados para carregá-las em um *data lake*, *data warehouse* ou bancos de dados. A realização dessa técnica tem como objetivo dar suporte à descoberta, geração de relatórios e análise de dados (Oracle, 2025; IBM, 2025).

A divisão do processo se dá, então, por três partes: extração, transformação e carregamento. Na extração, os dados brutos são identificados, e copiados de suas fontes originais, para uma área de preparação. A parte da transformação envolve o tratamento dos dados brutos, validando, autenticando, removendo valores duplicados, formatando, traduzindo, e outros procedimentos, para deixá-los prontos para serem gravados. O carregamento, terceira e última parte, é onde os dados, já tratados, são movidos do ambiente de transformação, para o armazenamento de destino, que para este trabalho, é um banco de dados (Oracle, 2025; IBM, 2025).

A Tabela 2 exemplifica possíveis fontes de dados, de onde serão extraídas as informações para serem transformadas e inseridas no banco de dados. Por enquanto, ainda não foram escolhidas fontes para os indicadores "presença de agentes governamentais", "número de agentes governamentais", "recursos captados por leis de incentivo", "volume de investimento anjo e capital de risco", "número de eventos de inovação realizados" e "número de programas de internacionalização de *startups*".

Tabela 2 – Fontes dos Indicadores de Ecossistemas de Inovação no RS

Indicadores	Fontes	Links
Localização	Site do Inova.RS	< https://programainova.rs.gov.br/inicial >
Gastos com PeD	Observatório de Inovação RS FormP&D governo Gov RS Outros sites de prefeituras com dados de despesas	< https://sict.rs.gov.br/observatorio > < https://formpd.mcti.gov.br > < https://l1nq.com/GovernoRS >
Número de universidades	Site OpenSense Rede RSStartup	< https://opensense.com.br/radar_ecosystem/# > < https://startup.rs.gov.br >
Número de mestres e doutores por 100 mil habitantes	Lattes Observatório de Inovação RS Painel Mestres e Doutores	< https://lattes.cnpq.br > < https://sict.rs.gov.br/observatorio > < https://mestresdoutores2024.cgee.org.br/painel-busca >
Número de pesquisadores	Lattes Observatório de Inovação RS	< https://lattes.cnpq.br > < https://sict.rs.gov.br/observatorio >
Empresas de alto crescimento	Site OpenSense Rede RSStartup Observatório de Inovação RS	< https://opensense.com.br/radar_ecosystem/# > < https://startup.rs.gov.br > < https://sict.rs.gov.br/observatorio >
Parcerias universidade - empresas	Fontes específicas das universidades	
Número de projetos em andamento	Fontes específicas das universidades	

Continua na próxima página

Tabela 2 – Fontes dos Indicadores de Ecossistemas de Inovação no RS (continuação)

Indicadores	Fontes	Links
Número de alunos em cursos de graduação e pós na área de exatas	Observatório de Inovação RS Censo da Educação Superior (INEP)	< https://sict.rs.gov.br/observatorio > < https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-da-educacao-superior >
Número de bolsas de iniciação científica e tecnológica concedidas	Censo da Educação Superior (INEP)	< https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-da-educacao-superior >
Presença de agentes governamentais	-	
Número de agentes governamentais	-	
Volume de investimento público em inovação	Gov RS Outros sites de prefeituras com dados de despesas	< https://11nq.com/GovernoRS >
Recursos captados por leis de incentivo	-	
Volume de investimento anjo e capital de risco	-	
Número de <i>startups</i> ativas / Empresas inovadoras	Site OpenSense Rede RSStartup Observatório de Inovação RS	< https://opensense.com.br/radar_ecosystem/# > < https://startup.rs.gov.br > < https://sict.rs.gov.br/observatorio >
Valor de mercados das <i>startups</i>	Sites diversos das empresas ou sobre mercado financeiro	
Presença e Número de incubadoras	Site OpenSense Rede RSStartup Observatório de Inovação RS	< https://opensense.com.br/radar_ecosystem/# > < https://startup.rs.gov.br > < https://sict.rs.gov.br/observatorio >
Número de <i>startups</i> incubadas	Site OpenSense Rede RSStartup Observatório de Inovação RS	< https://opensense.com.br/radar_ecosystem/# > < https://startup.rs.gov.br > < https://sict.rs.gov.br/observatorio >
Presença e Número de aceleradoras	Site OpenSense Rede RSStartup Observatório de Inovação RS	< https://opensense.com.br/radar_ecosystem/# > < https://startup.rs.gov.br > < https://sict.rs.gov.br/observatorio >
Número de <i>startups</i> aceleradas	Site OpenSense Rede RSStartup Observatório de Inovação RS	< https://opensense.com.br/radar_ecosystem/# > < https://startup.rs.gov.br > < https://sict.rs.gov.br/observatorio >
Presença e Número de PTs	Site OpenSense Rede RSStartup Observatório de Inovação RS	< https://opensense.com.br/radar_ecosystem/# > < https://startup.rs.gov.br > < https://sict.rs.gov.br/observatorio >
Número de empresas instaladas em PTs	Site OpenSense Rede RSStartup Observatório de Inovação RS	< https://opensense.com.br/radar_ecosystem/# > < https://startup.rs.gov.br > < https://sict.rs.gov.br/observatorio >
Setor predominante das <i>startups</i>	Site OpenSense Rede RSStartup Observatório de Inovação RS	< https://opensense.com.br/radar_ecosystem/# > < https://startup.rs.gov.br > < https://sict.rs.gov.br/observatorio >
Número de <i>startups</i> decorrentes de projetos acadêmicos	-	
Número de patentes registradas	Instituto Nacional da Propriedade Industrial (INPI)	< https://www.gov.br/inpi/pt-br >

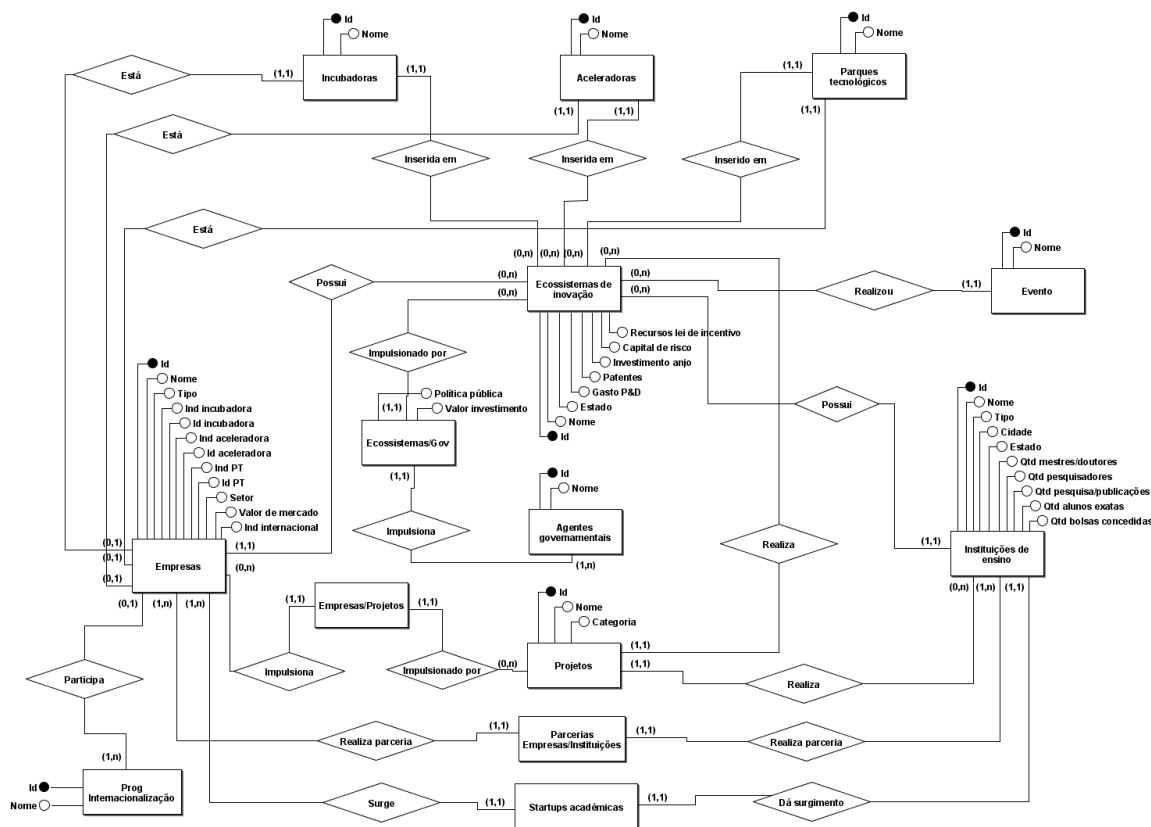
Continua na próxima página

Tabela 2 – Fontes dos Indicadores de Ecossistemas de Inovação no RS (continuação)

Indicadores	Fontes	Links
Número de pesquisas e publicações científicas por instituição	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) / Fundação Sucupira	< https://www.gov.br/capes/pt-br >
Número de eventos de inovação realizados	-	
Número de programas de internacionalização de startups	-	
Número de startups exportadoras ou com atuação internacional	Fontes diversas específicas de cada startup	

O banco de dados que vai ser criado será relacional, e utilizará do sistema PostgreSQL para seu gerenciamento. A implementação física será executada no PostgreSQL, utilizando SQL para a criação de esquemas, tabelas, índices, etc. A Figura 4 mostra um diagrama ER, ilustrando as tabelas que serão criadas, bem como seus atributos, e as relações entre as mesmas. Serão criadas 8 tabelas principais, as quais terão como chave primária um campo "Id", de valor numérico sequencial, e 4 tabelas oriundas de relações, tendo como chaves primárias os campos "Id" de suas tabelas pai.

Figura 4 – Diagrama ER do banco de dados



Fonte: O autor (2025)

A Figura 5 apresenta quatro exemplos de consultas que serão realizadas no banco de dados desenvolvido, com o objetivo de demonstrar como os dados necessários para a análise das métricas dos ecossistemas de inovação poderão ser extraídos. As consultas foram elaboradas com base nos indicadores definidos ao longo da pesquisa, considerando a estrutura relacional adotada. Assim, espera-se facilitar a obtenção das informações necessárias para as análises desenvolvidas ao longo do trabalho.

Inicialmente, é possível ver que em todas as consultas, exceto a quarta, um ecossistema de inovação pode ser informado pelo parâmetro "&id_eco". A consulta de teste um faz uma busca das *startups* que um ecossistema de inovação possui. A conexão é feita através da chave de ligação "id_ecossistema" que a tabela de "empresas" possui, juntamente com o indicador de "tipo"1, o qual mostra que uma empresa é *startup*.

A segunda consulta tem como objetivo encontrar o número total de alunos matriculados em Ciência, Tecnologia, Engenharia e Matemática (STEM). A busca é feita somando os alunos de todas as instituições presentes dentro do ecossistema de inovação. A finalidade da consulta três é obter todos os projetos que são apoiados por empresas, por ecossistema, através de uma busca pela tabela de ligação "empresas_projetos". A última consulta visa encontrar o número total de incubadoras, aceleradoras e PTs, por ecossistema de inovação, o que é feito agrupando essas estruturas por cada ecossistema.

Figura 5 – Consultas teste

```

/* 1 - Startups de um determinado ecossistema */
SELECT stp.id
      , stp.nome as startup
      , ei.nome as ecossistema
FROM empresas as stp
      , ecossistemas_de_inovacao as ei
WHERE stp.tipo = 1
      AND stp.id_ecossistema = ei.id
      AND ei.id = &id_eco;

/* 2 - Quantidade total de alunos matriculados em exatas */
/* em todo o ecossistema */
SELECT SUM(NVL(inst.qtd_alunos_exatas, 0)) as alunos_em_exatas
FROM instituicoes_de_ensino as inst
      , ecossistemas_de_inovacao as ei
WHERE inst.id_ecossistema = ei.id
      AND ei.id = &id_eco;

/* 3 - Projetos apoiados por */
/* empresas, por ecossistema */
SELECT emp.id,
      emp.nome AS empresa,
      proj.nome
FROM empresas emp,
      projetos proj,
      empresas_projetos epr,
      ecossistemas_de_inovacao ei
WHERE emp.id = epr.id_empresa
      AND proj.id = epr.id_projeto
      AND emp.id_ecossistema = ei.id
      AND ei.id = &id_eco;

/* 4 - Número de incubadoras, aceleradoras e PTs*/
/* por cada ecossistema */
SELECT ei.nome
      , COUNT(inc.id) as Incubadoras
      , COUNT(accel.id) as Aceleradoras
      , COUNT(pt.id) as Parques_tecno
FROM ecossistemas_de_inovacao as ei
      , aceleradoras as accel
      , aceleradoras as accel
      , parques_tecnologicos as pt
WHERE ei.id = inc.id_ecossistema
      AND inc.id_ecossistema = accel.id_ecossistema
      AND accel.id_ecossistema = pt.id_ecossistema
GROUP BY ei.nome;

```

Fonte: O autor (2025)

4 DESENVOLVIMENTO DA PROPOSTA DE SOLUÇÃO

O capítulo aborda a descrição dos processos de extração, transformação e carregamento propostos no capítulo 3, além da visualização dos dados obtidos. Nele, serão descritas as dificuldades que foram identificadas durante a realização do projeto, junto do conjunto final de indicadores obtido com suas fontes, e as ferramentas e métodos utilizados nesta etapa.

4.1 DIFICULDADES ENCONTRADAS

Para a extração dos dados utilizados na montagem do banco, foram utilizadas diferentes fontes, com sua maioria sendo ferramentas e sites governamentais, junto de sites específicos para determinados indicadores. Durante a execução desse processo, porém, foram encontradas algumas dificuldades que ocasionaram em modificações no conjunto inicial de 30 indicadores, proposto no capítulo 3.

O principal obstáculo na obtenção dos dados foi a dispersão deles. Dados de um mesmo indicador, por exemplo, não possuíam uma fonte centralizada, e estavam espalhados por diferentes plataformas. Isso acabou na retirada dos indicadores “parcerias universidade - empresas”, “recursos captados por leis de incentivo”, “volume de investimento anjo e capital de risco”, “valor de mercados das startups”, “número de startups decorrentes de projetos acadêmicos”, “número de programas de internacionalização de startups” e “número de startups exportadoras ou com atuação internacional”, tendo em vista que, para se obter uma quantidade minimamente expressiva de dados que satisfizesse esses indicadores, seria necessário uma busca mais aprofundada e extensa, algo que demandaria um tempo hábil maior que o disponível.

Outros indicadores sofreram modificações, visando adaptá-los às limitações, e buscando garantir sua aplicabilidade na análise dos ecossistemas de inovação. Por conta dos dados dispersos, o indicador “gastos com PeD”, teve seu foco reduzido somente a dados de prefeituras de cidades onde existe uma unidade da Universidade de Caxias do Sul (UCS). Posteriormente, o indicador “volume de investimento público em inovação” foi unificado a este. Também em razão desse entrave, os indicadores “presença de agentes governamentais” e “número de agentes governamentais”, tiveram sua abrangência restringida, focando as duas medidas na presença e número de unidades do Serviço Brasileiro de Apoio às Micro e Pequenas Empresas (SEBRAE) e Serviço Nacional de Aprendizagem Industrial (SENAI) pelo estado do RS.

O indicador “número de mestres e doutores por 100 mil habitantes” e “número de pesquisadores” tinham como fonte a plataforma Lattes, e apesar de existir uma ferramenta para a extração de dados própria da plataforma, o seu acesso requer uma liberação especial, à qual não foi possível conseguir para a realização deste estudo. Dessa forma, optou-se por fazer uma busca manual na plataforma, tendo como base os professores do curso de Ciência da Computação da UCS, em conjunto com os participantes do projeto *City Living Lab*.

Ao buscar dados para o indicador “número de projetos em andamento”, através da API da plataforma Sucupira, passando os valores referentes a “inovação” e “pesquisa”, no filtro de “natureza do projeto”, a API parava de trazer dados após um determinado registro, e, somente após a retirada do valor referente a “pesquisa” a extração passou a funcionar corretamente. Esse indicador recebeu também valores da pesquisa manual realizada na plataforma Lattes, incluindo todos os projetos de pesquisa em andamento dos profissionais analisados.

4.2 EXTRAÇÃO E TRANSFORMAÇÃO

Após mitigar os problemas identificados, o conjunto de 30 indicadores proposto no capítulo 3, teve seu tamanho reduzido para 23 indicadores. A Tabela 3 apresenta o conjunto final, junto das fontes utilizadas para a obtenção dos dados.

Tabela 3 – Conjunto final de indicadores

Indicadores	Fontes	Links
Localização	Site Inova RS	< https://programainova.rs.gov.br/inicial >
Número de eventos de inovação realizados	Site Inova RS	< https://programainova.rs.gov.br/inicial >
Número de <i>startups</i> ativas / Empresas inovadoras	Rede RS Startup	< https://startup.rs.gov.br >
Presença e Número de incubadoras / aceleradoras / PTs	Observatório de Inovação RS	< https://sict.rs.gov.br/observatorio >
Número de <i>startups</i> incubadas / aceleradas / instaladas em PTs	Rede RS Startup	< https://startup.rs.gov.br >
Setor predominante das <i>startups</i>	Rede RS Startup	< https://startup.rs.gov.br >
Número de universidades	Observatório de Inovação RS; Painel Estatístico Censo da Educação Superior (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP)); Plataforma Sucupira / CAPES	< https://sict.rs.gov.br/observatorio > < https://www.gov.br/inep/pt-br > < https://sucupira.capes.gov.br >
Número de alunos em cursos de graduação e pós na área de exatas	Painel Estatístico Censo da Educação Superior (INEP);	< https://www.gov.br/inep/pt-br >
Número de projetos em andamento	Plataforma Sucupira / CAPES; Plataforma Lattes	< https://sucupira.capes.gov.br > < https://lattes.cnpq.br >
Número de patentes registradas	Plataforma Sucupira / CAPES;	< https://sucupira.capes.gov.br >
Número de pesquisas e publicações científicas por instituição	Plataforma Sucupira / CAPES;	< https://sucupira.capes.gov.br >
Número de mestres e doutores	Plataforma Lattes	< https://lattes.cnpq.br >
Número de pesquisadores	Plataforma Lattes Plataforma Sucupira / CAPES;	< https://lattes.cnpq.br > < https://sucupira.capes.gov.br >
Presença de agentes governamentais	SEBRAE; SENAI	< https://sebrae.com.br > < https://www.senairs.org.br >
Número de agentes governamentais	SEBRAE; SENAI	< https://sebrae.com.br > < https://www.senairs.org.br >
Gastos com PeD	Sites de prefeituras (relatórios de despesas públicas)	

Continua na próxima página

Tabela 3 – Conjunto final de indicadores (continuação)

Indicadores	Fontes	Links
Número de bolsas de iniciação científica e tecnológica concedidas	Diário Oficial do Estado do RS	< https://diariooficial.rs.gov.br >
Empresas de alto crescimento	Revista Exame	< https://exame.com/negocios-em-expansao/ranking/1 >

A extração das informações dos nomes dos ecossistemas e das cidades que compõem cada um, pelo site do Inova RS, foi feita de forma manual, levando em conta o pequeno volume de dados. Porém para os dados de eventos, também retirados do site, foi criado um *script* capaz de acessá-lo, e retirar as informações a partir do código-fonte em HTML da página. O *script* utiliza das bibliotecas "Selenium" e "BeautifulSoup" do Python, para conseguir realizar a retirada das informações.

O *Script 1* mostra o *script* feito para extrair os dados dos nomes e datas, onde inicialmente o filtro de data era encontrado, utilizando a função "find_element" e passando o nome do elemento dentro do código-fonte da página como parâmetro, e após, preenchido com a data "01/01/2025". Em seguida, o botão para aplicar os filtros era clicado, e o documento HTML atualizado era obtido. A partir do novo código-fonte, os elementos contendo os nomes dos eventos e suas datas eram buscados, através das funções "find" e "find_all". A informação exata das datas era extraída pela função "extrair_data", apresentada no *Script 2*. Seu objetivo é remover o texto extra que o elemento contém, retornando somente a data na formatação "dia/mês/ano".

Script 1 – Função para coleta de eventos do portal Inova RS

```

1 def get_eventos(file, ecossistema, url):
2     ...
3     # Campo de Data Inicial
4     campo_dta_inicial = driver.find_element(by=By.NAME, value="datahoraini")
5     campo_dta_inicial.clear()
6     campo_dta_inicial.send_keys("01/01/2025")
7     campo_dta_inicial.send_keys(Keys.TAB) # confirma a data
8     time.sleep(3)
9     # Clicar no botao Buscar
10    filtrar = driver.find_element(By.XPATH, "//button[text()='Filtrar']")
11    filtrar.click()
12    time.sleep(2)
13
14    html = BeautifulSoup(driver.page_source, "html.parser")
15    eventos = html.find("div", class_="matriz-ui-pagedlist-body conteudo-
        lista__body")
16    for evento in eventos.find_all("article", class_="conteudo-lista__item"):
17        nome_ev = None
18        dta_ini = None
19        dta_fim = None

```

```

20     # Nome do evento
21     h2 = evento.find("h2", class_="conteudo-lista__item__titulo")
22     if h2 and h2.find("a"):
23         nome_ev = h2.find("a").string.strip()
24
25     # Pegar as datas
26     li_data = evento.find("li", class_="conteudo-lista__item__info-
27         evento__data")
28     if li_data:
29         times = li_data.find_all("time")
30         if len(times) == 2:
31             dta_ini = extrair_data(times[0].get_text(strip=True))
32             dta_fim = extrair_data(times[1].get_text(strip=True))
33             if dta_fim is None:
34                 dta_fim = dta_ini
35         elif len(times) == 1:
36             dta_ini = extrair_data(times[0].get_text(strip=True))
37             dta_fim = dta_ini

```

Fonte: O autor (2025)

Script 2 – Função para extração das datas inicial e final dos eventos

```

1 def extrair_data(texto):
2     if not texto:
3         return None
4
5     # Procura por padrao DD/MM/YYYY
6     match = re.search(r'\b(\d{2}/\d{2}/\d{4})\b', texto)
7     if match:
8         return match.group(1)
9     return None

```

Fonte: O autor (2025)

As fontes do “Observatório de Inovação RS”, “RedeRS Startup” e “Painel Estatístico Censo da Educação Superior” estavam no formato de *PowerBIs*. Por isso, criou-se o *Script 3*, capaz de realizar uma requisição à API *PowerBI*, por meio do uso da biblioteca "*requests*", e capturando o arquivo JSON de retorno do mesmo. Na requisição eram enviados a URL da API do *PowerBI*, os *headers* necessários, e um modelo do arquivo JSON contendo os filtros e a estrutura dos dados desejados.

Script 3 – Função para realizar request ao PowerBI

```

1 def request_bi(url, headers, json_model):
2     req = requests.post(url, headers=headers, json=json_model)
3     req.encoding = req.apparent_encoding
4
5     return req.json()

```

Fonte: O autor (2025)

Porém, a abordagem utilizada após realizar a requisição difere entre as fontes. O arquivo JSON retornado *PowerBI* retorna possui um vetor de dicionários, com os valores dos dados, e um segundo vetor, contendo os registros retornados, onde, em cada registro, três variáveis podem aparecer:

1. C: Contém os números referentes dos dados nos dicionários do registro pesquisado.
2. R: Indicador de repetição de dados. Aparece caso um dado do registro anterior no vetor possui uma mesma informação. Quando isso ocorre, o dado repetido fica omitido em C. Cada combinação de dados repetido recebe um valor diferente, sem padrão definido.
3. Ø: Indicador de dado faltante. Quando isso ocorre, o dado fica omitido em C. Caso o mesmo dado esteja faltando em dois registros subsequentes, "Ø" deixa de aparecer no segundo registro, criando um indicador de repetição de dado novo. Cada combinação de dados faltantes recebe um valor diferente, sem padrão definido.

Dessa forma, para os dados advindos do "Observatório de Inovação RS", foram identificados todos os valores de "R" e "Ø" possíveis, levando em conta que os conjuntos de dados retornados dentro de cada arquivo JSON eram relativamente curtos.

O *Script 4* então descompactava os arquivos JSON, até se obter o vetor "DM0" contendo os valores de "C", "R" e "Ø", e o vetor "ValueDicts" com os dicionários. Após, era analisado qual caso os valores se encaixavam, para depois disso, sim, obter os dados de "Nome", "Sigla", "Cidade" e "Site" do vetor "C". Em caso de repetição de algum destes valores, aplicava-se o número 99, indicando que o valor do registro anterior deveria ser reutilizado. Quando algum dado não estava disponível, era atribuído o valor "None", representando a ausência de informação e mantendo o campo nulo no registro final.

Script 4 – Função para tratar dados do painel Observatório de Inovação

```
1 def extrai_observatorio(folder, casos):
2     CASOS = {
3         # (R, O): [Sigla, Nome, Link, Cidade]
4         (None, None): [0, 1, 3, 6],
5         (44, None): [0, 1, 99, 3],
6         (None, 4): [0, 1, 2, 5],
7         (36, 152): [0, 1, None, 2],
8         ....
9     }
10    all_observatorio = []
11    for idx, eco in enumerate(ecossistemas):
12        path = os.path.join(folder, f"{files[idx]}.json")
13        with open(path, "r", encoding="utf-8") as f:
14            data = json.load(f)
15            ds = data["results"][0]["result"]["data"]["dsr"]["DS"][0]
16            dm = ds["PH"][0]["DM0"]
17            dicts = ds["ValueDicts"]
18            for reg in dm:
```

```

19     c, r, o = reg.get("C", []), reg.get("R"), reg.get("O")
20     indices = casos.get((r, o))
21     if indices:
22         # Pega o indice dos campos no registro (99 = anterior, se
           nulo = nulo)
23         value = None if indices[N] is None else (sigla if indices[N]
           == 99 else c[indices[N]])
24         # Busca o valor no dict
25         value_dict = None if sigla is None else dicts["DN"][sigla]
26         all_observatorio.append([eco, value_dict])
27     return all_observatorio

```

Fonte: O autor (2025)

Para o caso das fontes "Rede RS Startup" e "Painel Estatístico Censo da Educação Superior", por essas possuírem um número maior de dados, foi realizado um procedimento diferente. Primeiro, uma requisição era feita buscando conseguir apenas os nomes, tanto das instituições de ensino, para o caso do painel do INEP, quanto para as *startups* do Rede RS. Após, um segundo *script* era executado, percorrendo todos os nomes, e realizando uma requisição única para cada. Assim, apenas um registro retornava no arquivo JSON, evitando a aparição de variáveis de repetição e de dados faltantes, visto que, caso o último acontecesse, o dado era apresentado na variável "C", porém sem valor nenhum.

O *Script 5* apresenta o *script* utilizado para o caso das *startups*, que segue mesma lógica do caso do painel do INEP. Inicialmente a requisição a API era feita, e logo após, ocorria a descompactação do arquivo JSON, e os dados eram extraídos e salvos em um único vetor. Ao final, as informações do vetor foram armazenados em arquivos, sendo utilizados formatos distintos: CSV para as *startups* e JSON para as instituições.

Script 5 – Função para tratar dados do painel Rede RS Startup

```

1 def busca_startups():
2     ...
3     for reg in registros:
4         ecossistema, name_val = reg
5         json_model = monta_json(name_val, ecossistema)
6         req = requests.post(url, headers=headers, json=json_model)
7
8         data = req.json()
9         ds = data["results"][0]["result"]["data"]["dsr"]["DS"][0]
10        vd = ds.get("ValueDicts", {})
11        # sempre 1 registro, pega direto os valores
12        value = vd.get("DN", [""])[n] if vd.get("DN") else ""
13        all_rows.append([
14            ecossistema, values
15        ])
16
17    return all_rows

```

Fonte: O autor (2025)

Após a retirada dessas informações, um último tratamento foi feito, unindo os dados de instituições de ensino e pesquisa, retirados do "Observatório de Inovação RS", com os dados de matrículas vindo do painel do INEP. O resultado dessa fusão foi a inclusão de uma nova coluna "Matrículas" no CSV que continha os dados do observatório.

No *Script 6* é possível ver como o *script* extraía as linhas dos dois arquivos, utilizando da biblioteca "pandas", e de suas funções "*DataFrame*" para ler o conteúdo do arquivo JSON, e "*read_csv*" para ler o arquivo CSV. Ao percorrer os dados encontrados no CSV, verificava se, para cada um, existia uma instituição correspondente nos dados retirados do INEP, através de um *loop* interno, e da chamada da função "similar". Caso existisse uma correspondência, o valor do campo "QtIngresso" do arquivo JSON era salvo na nova coluna "matrículas" no arquivo final, montado também utilizando a função "*DataFrame*" do "pandas".

Script 6 – Função para união dos arquivos de instituições

```
1 def merge_instituicoes(arq_json, arq_csv, arq_saida):
2     df_json = pd.DataFrame(data_json)
3     df_csv = pd.read_csv(arq_csv, sep=';', encoding='utf-8')
4     df_csv.columns = df_csv.columns.str.strip()
5     ...
6     for idx, r_csv in df_csv.iterrows():
7         # Ve se tem um registro parecido no json
8         match = next(
9             (i for i, r_json in df_json.iterrows()
10              if similar(r_json['Instiuiacao'], r_csv['Nome'])),
11             None
12         )
13
14     # Se tem, salva a quantidade de matriculas no novo arquivo
15     if match is not None:
16         r_json = df_json.loc[match]
17         r_csv['Matriculas'] = str(r_json.get('QtIngresso', None))
18
19
20     resultado.append(r_csv.to_dict())
21
22     pd.DataFrame(resultado).to_csv(arq_saida, sep=';', index=False, encoding='
utf-8')
```

Fonte: O autor (2025)

As funções responsáveis por calcular a similaridade dos nomes são observadas no *Script 7*. Um passo importante, a normalização dos nomes, era feito através da biblioteca "re" do *Python*, removendo caracteres especiais e palavras desnecessárias, enquanto a similaridade era calculada através da função "*SequenceMatcher*" da biblioteca "diffib". Duas instituições só são consideradas semelhantes caso seus nomes atinjam uma porcentagem de similaridade maior que 95%.

Script 7 – Funções de normalização e cálculo de similaridade de palavras

```
1 def norm(s):
2     if pd.isna(s):
3         return ""
4     s = str(s).lower() # Minusculo
5     s = re.sub(r'\s*-\s*.*$', '', s) # Remove tudo depois do primeiro hifen
6     s = re.sub(r'^\w\s]', '', s) # Remove acentos e caracteres especiais
7     return s
8
9 def similar(a, b, nome_min=0.95):
10    return SequenceMatcher(None, norm(a), norm(b)).ratio() >= nome_min
```

Fonte: O autor (2025)

Para efetuar a coleta dos dados da base Sucupira, foram feitas *requests* à sua API. Inicialmente, foram realizadas três consultas diferentes, buscando extrair as informações de produções do tipo "Patente", projetos de natureza "Inovação", e publicações da área de "Ciências Exatas e da Terra", sempre filtrando o "ano base" como 2024. Tais requisições retornaram um arquivo JSON com os dados buscados. O *Script 8* demonstra a realização das requisições através da biblioteca "*requests*", passando também como parâmetros o número da página e o número de registros que aparecem por página. Dessa forma, foi necessário executar uma requisição para cada página, obtendo-se assim os dados retornados, que eram transformados em um arquivo JSON, e a partir daí, o vetor "*content*", que continha realmente os dados, tinha suas informações extraídas e armazenadas em único vetor, para serem salvas juntas em um novo arquivo.

Script 8 – Função para realizar requisições a API do Sucupira

```
1 def req_sucupira():
2     ...
3     while True:
4         try:
5             url = f"{api_url}?query={query}&page={page}&size={size}"
6             r = requests.get(url)
7             if r.status_code == 200:
8                 data = r.json()
9                 items = data.get("content", r) # usa 'content' se existir, senao
10                    assume lista
11                 if not items:
12                     break
13                 all_data.extend(items)
14                 page += 1
15             except requests.exceptions.RequestException as e:
16                 print(f"An error occurred: {e}")
```

Fonte: O autor (2025)

Depois de conseguir os dados necessários da base Sucupira, uma última requisição foi feita, com a intenção de extrair os nomes e siglas de todas as insituições de ensino do RS, utilizando o

mesmo método do *Script 8*. Obtidos os dados das instituições, foi executado um *script*, apresentado no *Script 9*, sobre os outros dados advindos da plataforma Sucupira, filtrando somente pelas instituições que foram obtidas nessa última extração, através do método "*merge*", da biblioteca "pandas". A função primeiramente carrega os dois arquivos que serão integrados, e após, realiza uma operação de junção pela coluna "sigla", garantindo que apenas registros presentes em ambos sejam mantidos no resultado. Ao final, um novo arquivo CSV é gerado, contendo o resultado da fusão dos repositórios, ou seja, somente produções feitas no RS.

Script 9 – Função para filtragem dos arquivos por instituições do RS

```
1 def merge_instituicoes(arq_json , arq_csv , arq_saida):
2     instituicoes = pd.read_csv(arq_instituicoes , sep=';' , encoding='utf-8')
3     outros_dados = pd.read_csv(arq_dados , sep=';' , encoding='utf-8')
4
5     combinado = pd.merge(
6         outros_dados ,
7         instituicoes ,
8         left_on="Sigla" ,
9         right_on="Sigla" ,
10        how='inner' ,
11        suffixes=("_dados" , "_instituicao")
12    )
13
14    combinado = combinado[
15        [
16            "Sigla" ,
17            "Nome_instituicao" ,
18            "Nome_dados" ,
19            "Responsavel"
20        ]
21    ]
22
23    combinado = combinado.rename(columns={
24        "Sigla": "Sigla da IES" ,
25        "Nome_instituicao": "Instituicao" ,
26        "Nome_dado": "Dado"
27    })
28
29    combinado.to_csv(arq_saida , sep=';' , index=False , encoding='utf-8')
```

Fonte: O autor (2025)

Para realizar a extração das unidades do SEBRAE e SENAI no RS, um *script* parecido com o *Script 1* foi montado, gravando as informações de nome da unidade e a cidade onde está localizada. A extração dos dados do indicador "empresas de alto crescimento" foi feita do mesmo modo, também gravando seus dados em um CSV.

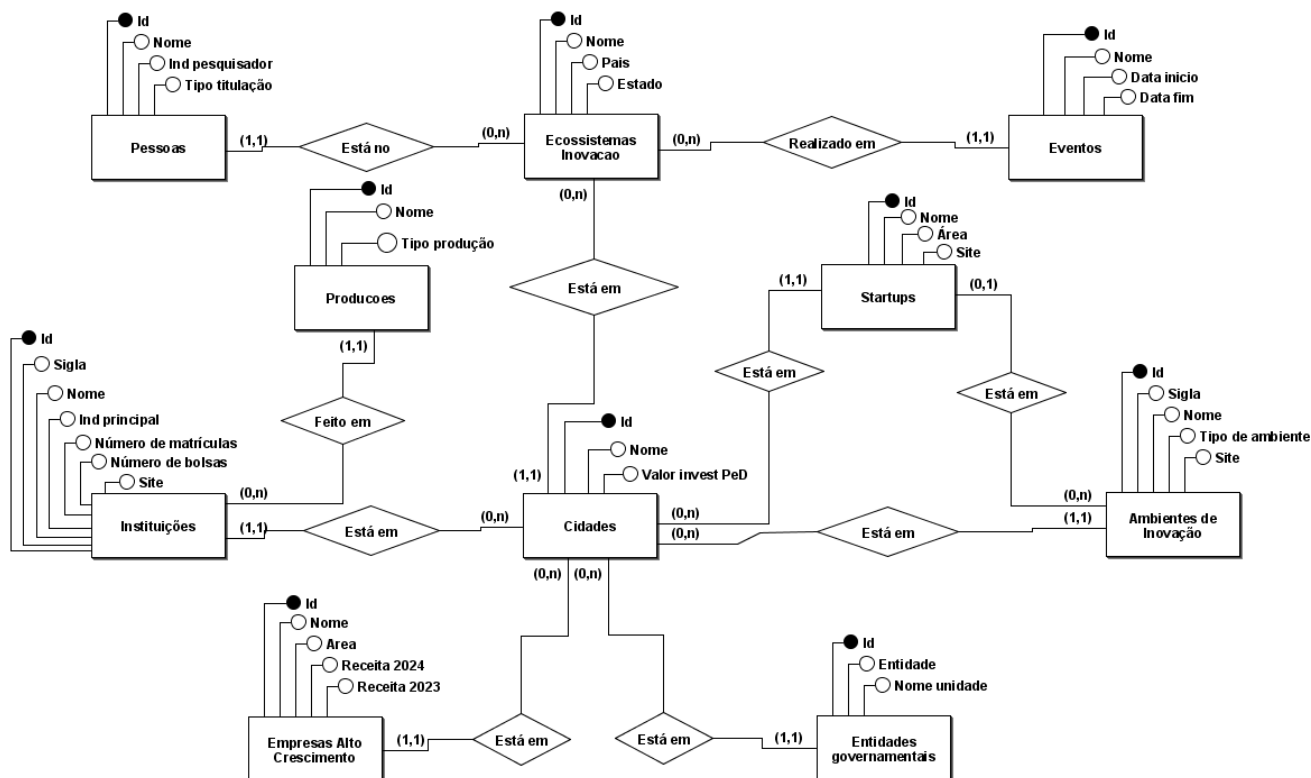
Para os dados provenientes dos sites das prefeituras, a extração foi realizada manualmente, com foco nos valores de despesas de secretarias ou órgãos voltados à inovação e PeD, até o dia

primeiro de novembro. Da mesma forma, as informações sobre bolsas concedidas também foram coletadas manualmente, devido ao volume de dados disponíveis ser pequeno. Os dados extraídos da plataforma Lattes seguiram o mesmo procedimento, porém isso foi necessário devido a um *Captcha* que impedia o acesso à página por meio do uso das funções da biblioteca "Selenium".

4.3 CARREGAMENTO DOS DADOS

Com as alterações realizadas no conjunto de indicadores, e as informações obtidas no processo de extração e transformação, o modelo do banco de dados proposto no capítulo 3 também sofreu mudanças. A Figura 6 mostra o novo diagrama ER, com as tabelas criadas, suas colunas, chaves primárias e as relações entre elas. O modelo final conta com 10 tabelas, todas com chave primária sendo um campo "id", de valor numérico sequencial. O Apêndice B – *Script 10* mostra o *script* SQL utilizado para a criação dessas tabelas no banco de dados.

Figura 6 – Diagrama ER final do banco de dados



Fonte: O autor (2025)

Após realizadas as transformações nos dados, os arquivos de dados gerados eram todos do tipo CSV, portanto, o carregamento das informações contidas neles seguiu uma mesma lógica (ver Apêndice B – *Script 11*). Todos os *scripts* criados para fazer esse *upload* tem pelo menos um parâmetro de entrada, sendo ele o diretório do arquivo CSV que será lido. As exceções são as *procedures* para carregar dados dos ambientes de inovação e das produções, que também possuem um indicador de tipo como parâmetro de entrada. A ordem básica que cada um segue é:

1. Abre o arquivo utilizando a função *"pg_read_file"* e armazena seu conteúdo.
2. Converte o conteúdo em um vetor de linhas por meio da função *"string_to_array"*, utilizando a quebra de linha como separador.
3. Itera sobre o vetor de linhas, separando as colunas através do delimitador *','*, também com *"string_to_array"*, e armazenando-as em um vetor.
4. Realiza as transformações e ações necessárias com os dados do vetor de colunas.

O carregamento dos dados dos nomes de cada ecossistema de inovação é o único que difere levemente dessa lógica, conforme mostra o Apêndice B – *Script 12*. O motivo disso é por conta de seu arquivo, que continha apenas uma linha com todos os nomes, separados por *','*.

Os *scripts* criados para popular as tabelas de "eventos"(ver Apêndice B – *Script 13*) e "cidades"(ver Apêndice B – *Script 14*) possuem escopos semelhantes. Para os dois, antes de inserir os dados em cada tabela, o "id" do ecossistema correspondente é buscado através da coluna "nome", utilizando o operador *"LIKE"*, e a função *"UPPER"*. A carga dos dados de investimento foi realizada pelo Apêndice B – *Script 15*, *procedure* semelhante às anteriores, com a diferença na busca do "id", que nesse caso é feita na tabela das cidades, e na realização de um update na coluna *"vlr_invest"* da tabela com o valor vindo do arquivo.

Os dados de instituições foram carregados da mesma maneira. A busca pelo "id" da cidade correspondente ao nome do arquivo era feita, e após, os valores, inicialmente apenas de instituições registradas pelo INEP, eram inseridos na tabela, conforme mostrado no Apêndice B – *Script 16*.

Na sequência, dois *scripts* eram executados, para complementar as colunas da tabela "institui-coes". Primeiro, o Apêndice B – *Script 17* atualizava o valor da coluna *"num_matriculas"* da tabela, e, após, o Apêndice B – *Script 18* carregava a quantidade de bolsas por instituição, atualizando o valor dos registros pela "sigla". Por conta dos dados de instituições serem oriundos de fontes diferentes, um tratamento especial no Apêndice B – *Script 17* foi feito, de modo que, caso uma instituição não fosse encontrada através da busca pelo nome, um novo registro era inserido na tabela. Ainda, uma busca pela sigla era feita, para determinar se a instituição existia nos dados retirados do INEP, e caso não, era salvo com a coluna *"tip_inst"* com valor 2, indicando ser uma instituição de pesquisa.

A inserção dos dados na tabela *"ambientes_inov"*, referente às incubadoras, aceleradoras e PTs, ocorreu de mesma maneira das instituições. A diferença da *procedure* criada para isso é a presença do parâmetro de entrada *"tip_amb"*, referindo-se aos tipos de ambientes de inovação, onde o valor 1 indica ser uma aceleradora, 2 uma incubadora, e 3 um PT. Dessa maneira, foram realizadas três chamadas do Apêndice B – *Script 19*, uma para cada tipo, recebendo também o caminho dos arquivos correspondentes.

O arquivo utilizado para a carga das *startups* continha uma coluna indicando se a empresa participava de algum ambiente de inovação, onde tal coluna possuía a sigla do ambiente. Porém, nesse caso, alguns nomes não possuíam correspondência exata com os dados de ambientes carregados, algo que poderia acontecer por dois motivos distintos. Uma das razões ocorre por conta do ambiente não

ser do tipo de nenhum dos três extraídos para esse trabalho, como centros ou núcleos de inovação. A outra razão está relacionada a pequenas variações na escrita das siglas, como a inclusão ou remoção de preposições (“de”, “do”, “da”, etc.) ou mesmo alterações na forma de grafia. Para tratar esses casos, então, o Apêndice B – *Script 20* verifica a existência de um ambiente pela sigla e pela cidade, e caso não encontre, verifica ainda somente pela sigla, e se mesmo assim não encontrar um valor correspondente, insere um novo registro na tabela de ambientes, gravando-o com o tipo 4 (outros). Porém, por conta disso, a ocorrência de registros “duplicados” é passível, com dois registros referindo-se à um mesmo ambiente, porém com siglas diferentes.

A tabela de pessoas refere-se aos dados dos professores do curso de ciência da computação da UCS e dos participantes do *City Living Lab*, extraídos da plataforma Lattes. O *script* para carga dos dados, apresentado no Apêndice B – *Script 21*, realiza duas operações “CASE”, uma para determinar o tipo de titulação, atribuindo o valor 1 para “especialização”, 2 para “mestrado”, 3 para “doutorado” e 4 para casos não informados ou outros tipos de titulação, além de obter o valor que indica se é pesquisador, sendo 1 para sim e 0 para não. Também, um *loop* era realizado sobre a coluna que continha os nomes dos projetos de pesquisas de cada um, inserindo um novo registro na tabela “producoes” para cada, considerando nesse caso a instituição UCS para todas inserções.

O Apêndice B – *Script 22* mostra o *script* responsável por carregar os dados de patentes, projetos e publicações, extraídos através da API do Sucupira. A *procedure* realiza um *loop* interno para cada produção, com o objetivo de inserir um registro na tabela para cada autor separadamente. Ainda, se a produção for do tipo “projeto”, é inserido um registro na tabela “pessoas”, considerando cada autor um pesquisador. Para o caso desses dados, fez-se necessária a realização de uma validação para a inclusão de um registro repetido, pois a API retornava alguns registros duplicados, contendo mesmo dado de nomes, autores, ano, etc., mas com um número identificador diferente.

No Apêndice B – *Script 23*, uma variável do tipo “ROWTYPE” da tabela de empresas de alto crescimento é utilizada. Cada campo recebe o valor contido nas colunas vindas do arquivo, removendo espaços em branco no início e fim, através da função “TRIM”, e, no caso dos valores de receita, ocorre a remoção do “.” como separador de milhar, e a transformação para *numeric*. Além disso, para popular o campo “id”, a função “NEXTVAL” é utilizada, buscando o próximo valor na sequência do campo na tabela.

A última *procedure* criada para carregar os dados extraídos para o banco de dados foi o Apêndice B – *Script 24*. Responsável por inserir as informações sobre as entidades governamentais, o *script* é bastante simples, realizando a busca do “id” da cidade correspondente e inserindo os dados na tabela “entidades_gov”. Ponto importante do código é a transformação do nome da cidade vindo do arquivo, retirando os acentos e caracteres especiais através da extensão “UNACCENT” do PostgreSQL, tornando todas as letras maiúsculas e removendo espaços em branco com as funções “UPPER” e “TRIM”.

Para a execução de todas estas *procedures* criadas, um *script* final foi feito, para evitar executar as chamadas de maneira individual. Dessa forma, o Apêndice B – *Script 25* realiza todas as chamadas em um só local, criando as tabelas e populando-as de uma só vez.

4.4 RESULTADOS DA CARGA DOS DADOS

Após finalizar todas as etapas do processo de ETL, o banco de dados conta com 10 tabelas, um total de 54 colunas, e mais de 5000 registros incluídos, contemplando todos os indicadores mostrados na Tabela 3.

Ao todo, 15 procedures foram criadas para tratar as tabelas e carregar os dados extraídos e transformados. A Tabela 4 mostra o número total de linhas por cada tabela do banco. É possível ver como os indicadores que possuíam fontes mais estruturadas, com dados agrupados e disponibilizados por sites governamentais, resultaram em uma maior quantidade de registros ao fim de todo o processo.

A Tabela 5 apresenta os números totais de registros e valores obtidos para cada indicador. A maioria dos indicadores refere-se ao número de registros gravados no banco de dados, com exceção dos "alunos matriculados", "bolsas concedidas" e "gastos com PeD", que possuem valores correspondentes às somas acumuladas.

Tabela 4 – Quantia total de registros em cada tabela

Tabela	Quantidade de registros
Ecosistemas de Inovação	8
Cidades	496
Eventos	29
Instituições	184
Ambientes de inovação	117
Startups	1472
Pessoas	366
Produções	2358
Empresas de alto crescimento	19
Entidades governamentais	220

Tabela 5 – Total de valores por indicador

Indicador	Total
Localização	8 Ecosistemas de Inovação e 496 Cidades
Número de eventos de inovação realizados	29
Número de <i>startups</i> ativas / Empresas inovadoras	1472
Presença e Número de incubadoras	66
Presença e Número de aceleradoras	4
Presença e Número de PTs	24
Número de <i>startups</i> incubadas	97
Número de <i>startups</i> aceleradas	29
Número de <i>startups</i> instaladas em PTs	200
Setor predominante das <i>startups</i>	53 áreas de atuação diferentes

Continua na próxima página

Tabela 5 – Total de valores por indicador (continuação)

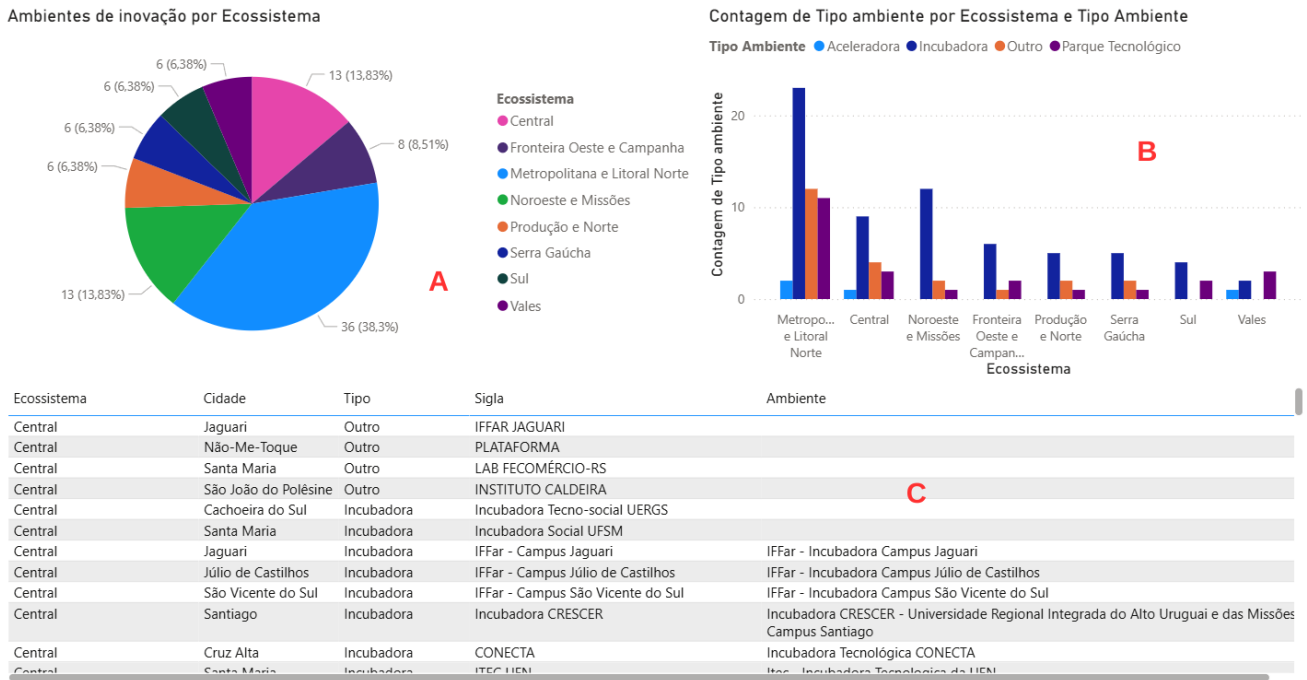
Indicador	Total
Número de universidades	41 Instituições diferentes e 139 Campus totais
Número de alunos em cursos de graduação e pós na área de exatas	10760
Número de projetos em andamento	588
Número de patentes registradas	241
Número de pesquisas e publicações científicas por instituição	776
Número de mestres e doutores	85
Número de pesquisadores	332
Presença e número de agentes governamentais	220
Gastos com PeD	R\$174.565.033,14
Número de bolsas de iniciação científica e tecnológica concedidas	400
Empresas de alto crescimento	19

4.5 VISUALIZAÇÃO DOS DADOS

Para auxiliar a visualização dos indicadores propostos, foi desenvolvido o painel Visualizador de Ecossistemas de Inovação no *PowerBI* (O Autor, 2025). Dessa forma, tornou-se possível gerar gráficos que apresentam os dados de maneira mais clara, atrativa e intuitiva. A Figura 7 contém o gráfico de pizza "A", que mostra o número de ambientes de inovação por ecossistema de inovação, onde é possível ver uma maior concentração no ecossistema "Metropolitana e Litoral Norte". Além disso, o gráfico de barras "B" contabiliza o número total dos ambientes separados por tipos, por cada ecossistema, permitindo ver que as incubadoras estão em maior número em quase todos os ecossistemas, à exceção do "Vales", que possui mais PTs. O painel apresenta ainda a tabela "C" com os dados dos ambientes.

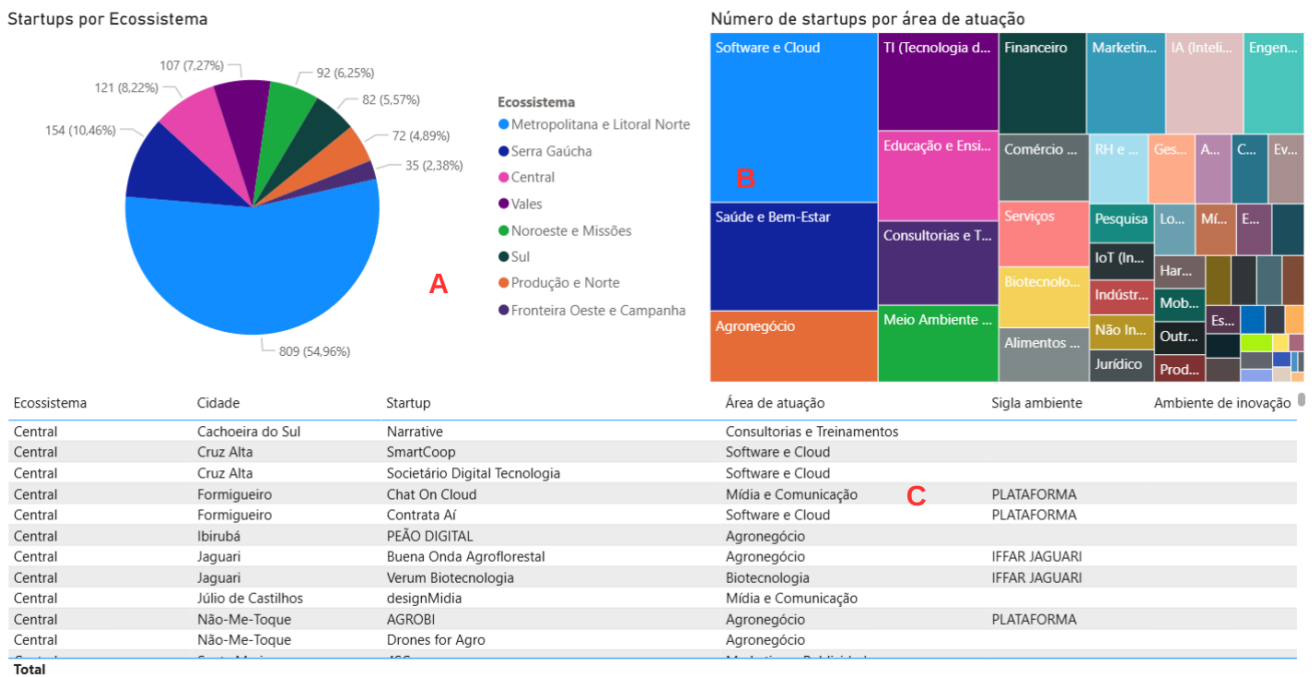
A Figura 8 mostra outro exemplo de painel, dessa vez contendo os dados extraídos sobre startups. Semelhante à Figura 7, também conta com o gráfico de pizza "A" para mostrar o número de startups por ambiente, e, de mesma forma, pode-se perceber uma maior concentração delas no ecossistema "Metropolitana e Litoral Norte". O painel possui também a *treemap* "B", indicando a "popularidade" de cada área de atuação entre as startups, e, por fim, também possui a tabela "C" com os dados nomeados de todas as startups.

Figura 7 – Painel com indicadores dos ambientes de inovação



Fonte: O autor (2025)

Figura 8 – Painel com indicadores das startups



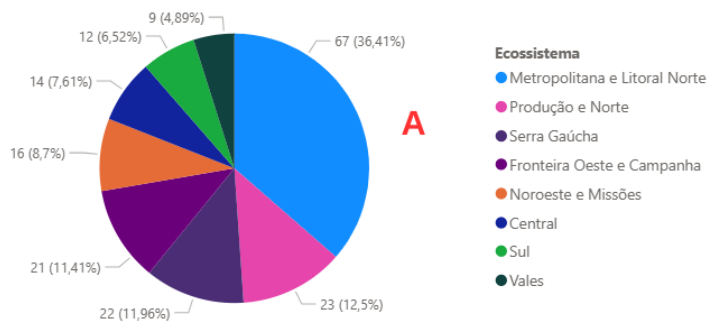
Fonte: O autor (2025)

Como último exemplo, o painel com informações das instituições de ensino e pesquisa, que pode ser visto na Figura 9, apresenta também o gráfico de pizza "A" com os totais por ecossistema, mais uma vez com predominância da região metropolitana e do litoral norte. Também conta com dois gráficos de barras, onde o "B" exibe o número de bolsas concedidas, e o "C" o número de matrículas por instituições, agrupadas por suas siglas. A tabela "D" retrata as produções extraídas da API da

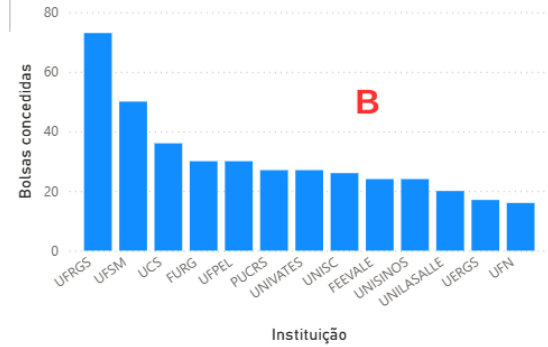
plataforma Sucupira, organizadas também pelas siglas de suas respectivas instituições.

Figura 9 – Painel com indicadores das instituições

Instituições por Ecosistema

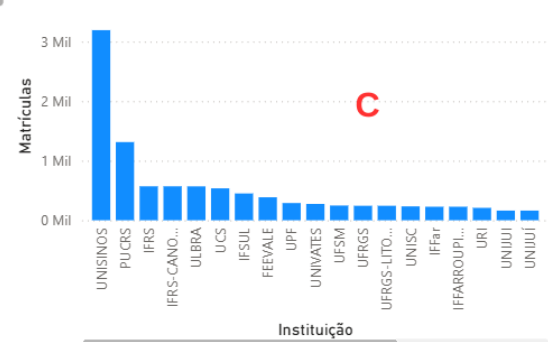


Bolsas concedidas por Instituição



Sigla IEP	Tipo	Produção
FURG	Dissertação	δ13CDIC COMO INDICADOR DE PROCESSOS BIOGEOQUÍMICOS I AUSTRAL AO LONGO DO ESTREITO DE BRANSFIELD, ANTÁRTICA
UCS	Projeto de Pesquisa	ZYMOMONAS – Desenvolvimento de filmes ativos e nanocompós aldônicos produzidos por células imobilizadas de Zymomonas mc
UFRGS	Projeto de Pesquisa	WORKSHOP DE VISUALIZAÇÃO E MANIPULAÇÃO DE DADOS DE I PROGRAMADORES
FURG	Projeto de Pesquisa	WATER MAASURE
UCS	Projeto de Pesquisa	VULNERA - VULNERABILIDADE E CLÍNICA-ESCOLA: UMA RESPOS
FURG	Tese	Voltametria de micropartículas imobilizadas aplicada a identificaç
PUCRS	Dissertação	VISUALIZAÇÃO DE DADOS FALTANTES PARA MELHORIA DA QUAL
FEEVALE	Projeto de Pesquisa	VIGILÂNCIA EPIDEMIOLÓGICA MOLECULAR DE VARIANTES DO S. BRASIL-ARGENTINA
FURG	Patente	VIDA DE BICO ENERGIZADO DE CONTATO EM SOLDAGEM

Matrículas por Instituição



Fonte: O autor (2025)

5 CONCLUSÃO

Os ecossistemas de inovação desempenham um papel fundamental no desenvolvimento econômico, tecnológico e social, reunindo diversos atores na busca pela inovação. A análise desses ambientes é essencial para compreendê-los e para fortalecer o processo de evolução inovativa que promovem. No entanto, essa análise é algo bastante complexo de ser realizado, tendo em vista a dispersão e falta de padronização dos dados disponíveis nas mais diversas fontes, tornando ainda mais necessária a criação de ferramentas capazes de organizar essas informações de maneira estruturada.

Este trabalho teve como objetivo a criação de um banco de dados que fosse capaz de apresentar indicadores dos ecossistemas de inovação do RS, utilizando-se apenas de dados abertos para tal. Ao final do projeto, pôde-se perceber que mesmo com todas as inconsistências existentes das fontes de dados abertos, foi possível identificar, extrair, transformar e integrar as informações de diferentes plataformas públicas, institucionais ou independentes, resultando em uma base unificada e funcional.

A criação de fato do banco de dados permitiu concluir o objetivo principal do trabalho. A utilização da linguagem "*Python*" e suas bibliotecas, juntamente do Sistema de Gerenciamento de Banco de Dados (SGBD) PostgreSQL, permitiu a realização do processo de ETL, resultando na criação de um banco de dados relacional, com 10 tabelas, mais de 5000 registros, e informações de 23 indicadores diferentes.

Da mesma forma, todos os objetivos específicos foram alcançados. As fontes de dados abertas foram mapeadas, analisadas e utilizadas conforme sua disponibilidade, tendo seus dados coletados, limpos e transformados. O banco de dados foi modelado e implementado com sucesso, e consultas SQL foram criadas para permitir a extração de indicadores relevantes. Além disso, foram construídos painéis, utilizando a ferramenta *PowerBI*, para facilitar a visualização dos resultados, reforçando a utilidade prática do sistema desenvolvido.

Os resultados obtidos demonstram que a utilização de dados abertos, mesmo com suas limitações, pode servir como base sólida para analisar os ecossistemas de inovação do estado. A consolidação das informações permitiu construir uma visão integrada desses ambientes, possibilitando comparações e análises que antes eram inviáveis devido à fragmentação das fontes. Assim, o banco de dados desenvolvido tem potencial para apoiar os diversos atores envolvidos em seus processos de análise e definição de ações.

Por fim, este projeto pode ainda ser expandido em diversas frentes. A busca por novas fontes permitira atender novos indicadores, possibilitando também expandir a análise, de modo que englobe todo território nacional. Além disso, aprimoramentos no processo de ETL, bem como a atualização contínua do banco de dados e dos painéis de visualização, também podem ser melhorias que tornariam a solução muito mais robusta e completa.

REFERÊNCIAS

- AEVO. **Por que inovar nas empresas é tão importante?** 2024. Acesso em: 07 jun. 2025. Disponível em: <<https://blog.aevo.com.br/por-que-inovar-nas-empresas/>>.
- Anprotec. **Parques Tecnológicos Consolidados.** 2025. Acesso em: 6 abr. 2025. Disponível em: <<https://anprotec.org.br/site/lideres-tematicos/parques-consolidados/>>.
- BEAUDRY, C.; SOLAR-PELLETIER, L. **The Superclusters Initiative: An Opportunity to Reinforce Innovation Ecosystems.** Institute for Research on Public Policy (IRPP), 2020. Acesso em: 27 abr. 2025. Disponível em: <<https://irpp.org/wp-content/uploads/2020/10/The-Superclusters-Initiative-An-Opportunity-to-Reinforce-Innovation-Ecosystems.pdf>>.
- BUSATO, L. H. **Atores, políticas e fomento à inovação: SNI, lei da inovação, incentivos, venture, capital, anjos, incubadoras, aceleradoras, startups.** 1. ed. São Paulo: Contentus, 2020. E-book. Disponível em: <https://plataforma.bvirtual.com.br>. Acesso em: 05 abr. 2025.
- CAPES. **Plataforma Sucupira.** 2025. Acesso em: 07 jun. 2025. Disponível em: <<https://sucupira.capes.gov.br>>.
- CARAYANNIS, E. G.; BARTH, T. D.; CAMPBELL, D. F. J. The quintuple helix innovation model: global warming as a challenge and driver for innovation. **Journal of Innovation and Entrepreneurship**, v. 1, n. 2, p. 1–12, 2012. Acesso em: 18 abr. 2025. Disponível em: <https://www.researchgate.net/publication/257884675_The_Quintuple_Helix_innovation_model_global_warming_as_a_challenge_and_driver_for_innovation>.
- COBBEN, D.; OOMS, W.; ROIJAKKERS, N. Indicators for innovation ecosystem health: A delphi study. **Journal of Business Research**, v. 162, p. 113860, 2023. Acesso em: 27 abr. 2025. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0148296323002187>>.
- Comissão Europeia. **Cidades Inteligentes.** 2025. Acesso em: 6 abr. 2025. Disponível em: <https://commission.europa.eu/eu-regional-and-urban-development/topics/cities-and-urban-development/city-initiatives/smart-cities_en>.
- Cubo Itaú. **Ecosistema de inovação: o que é e como funciona.** 2024. Acesso em: 08 jun. 2025. Disponível em: <<https://blog.cubo.network/ecossistema-de-inovacao>>.
- CUKIER, D.; KON, F.; KRUEGER, N. Towards a software startup ecosystems maturity model. In: **16th International Conference on Product-Focused Software Process Improvement (PROFES).** Springer International Publishing, 2015. p. 600–606. Acesso em: 26 mai. 2025. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-319-26844-6_45>.
- Data Europa Academy. **What is Open Data?** 2025. Acesso em: 22 mar. 2025. Disponível em: <<https://data.europa.eu/en/dataeuropa-academy/what-open-data>>.
- Data Europa EU. **Benefits and Value of Open Data.** 2025. Acesso em: 6 abr. 2025. Disponível em: <<https://data.europa.eu/en/publications/datastories/benefits-and-value-open-data>>.
- DOMANSKI, D.; HOWALDT, J.; KALETKA, C. A comprehensive concept of social innovation and its implications for the local context – on the growing importance of social innovation ecosystems and infrastructures. **European Planning Studies**, v. 28, n. 3, p. 454–474, 2020. Acesso em: 6 abr. 2025. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/09654313.2019.1639397>>.

DUARTE, A. H.; OUTROS. **Gerenciamento de projetos inovadores**. [S.l.]: Brasport, 2025. E-book. Disponível em: <https://plataforma.bvirtual.com.br>. Acesso em: 05 abr. 2025.

EIE Datahub. **EIE Data Hub Platform**. 2025. Acesso em: 22 mar. 2025. Disponível em: <https://eie-datahub.eisma.eu/#>>.

FREEMAN, C. **Política tecnológica e desempenho econômico: Lições do Japão**. Londres: Pinter, 1987.

Governo Digital - Brasil. **Dados Abertos**. 2025. Acesso em: 22 mar. 2025. Disponível em: <https://www.gov.br/governodigital/pt-br/dados-abertos>.

Governo Federal. **Legislação — Dados Abertos**. 2025. Acesso em: 08 jun. 2025. Disponível em: <https://www.gov.br/agricultura/pt-br/acao-a-informacao/dadosabertos/legislacao-dados-abertos>.

GRANSTRAND, O.; HOLGERSSON, M. Innovation ecosystems: A conceptual review and a new definition. **Technovation**, v. 90–91, p. 1–12, 2019. Acesso em: 6 abr. 2025. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0166497218303870>.

IBM. **ETL - Extração, transformação e carregamento de dados**. 2025. Acesso em: 13 jun. 2025. Disponível em: <https://www.ibm.com/br-pt/topics/etl>.

IDIA. **What is an Innovation Ecosystem**. 2025. Acesso em: 22 mar. 2025. Disponível em: <https://www.idiainnovation.org/what-is-an-innovation-ecosystem>.

InovaData-BR. **Parques Tecnológicos no Brasil**. 2025. Acesso em: 22 mar. 2025. Disponível em: <https://www.inovadata-br.ufv.br/parks-page>.

KOCHETKOV, D. M. Innovation: A state-of-the-art review and typology. **International Journal of Innovation Studies**, v. 7, n. 4, p. 263–272, 2023. Acesso em: 08 jun. 2025. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2096248723000267>.

KOGABAYEV, T.; MAZILIAUSKAS, A. The definition and classification of innovation. **HOLISTICA – Journal of Business and Public Administration**, v. 8, n. 1, p. 59–72, 2017. Acesso em: 08 jun. 2025. Disponível em: <https://doi.org/10.1515/hjbpa-2017-0005>.

KOSORUKOVA, I. *et al.* Development of an entrepreneurial start-up ecosystem: Socio-economic and legal factors influencing the achievement of esg principles (case of asian countries). **Journal of Law and Sustainable Development**, v. 11, n. 4, p. 1–25, 2023. Acesso em: 27 abr. 2025. Disponível em: https://www.researchgate.net/publication/373402555_Development_of_an_Entrepreneurial_Start-Up_Ecosystem_Socio-Economic_and_Legal_Factors_Influencing_the_Achievement_of_ESG_Principles_Case_of_Asian_Countries.

LEPORE, D.; FRONTONI, E.; MICOZZI, A. Uncovering the potential of innovation ecosystems in the healthcare sector after the covid-19 crisis. **Health Policy**, v. 127, n. 1, p. 1–8, 2023. ISSN 0168-8510. Acesso em: 6 abr. 2025. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0168851022003098>.

MACEDO, D. F.; LEMOS, D. L. d. S. Dados abertos governamentais: iniciativas e desafios na abertura de dados no brasil e outras esferas internacionais. **AtoZ: Novas Práticas em Informação e Conhecimento**, v. 10, n. 2, p. 14–26, 2021. Acesso em: 6 abr. 2025. Disponível em: <https://revistas.ufpr.br/atoz/article/view/77737/43552>.

MCTIC. **Marco Legal de Ciência, Tecnologia e Inovação**. 2018. Acesso em: 22 mar. 2025. Disponível em: <https://antigo.mctic.gov.br/mctic/export/sites/institucional/arquivos/ASCOM_PUBLICACOES/marco_legal_de_cti.pdf>.

MINEIRO, A. A. d. C.; SOUZA, T. Assis de; CASTRO, C. Carvalho de. The quadruple and quintuple helix in innovation environments (incubators and science and technology parks). **Innovation & Management Review**, Emerald Publishing Limited, v. 18, n. 3, p. 292–307, 2021.

Ministério da Agricultura e Pecuária. **Ambientes de Inovação - Agrohubs Brasil**. 2025. Acesso em: 22 mar. 2025. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/inovacao/agrohub-brasil/ambientes-inovacao>>.

Ministério da Ciência, Tecnologia e Inovação. **Painel indica que ecossistema de inovação precisa de constância de investimentos e maior integração entre os seus atores**. 2024. Acesso em: 26 mai. 2025. Disponível em: <<https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2024/08/painel-indica-que-ecossistema-de-inovacao-precisa-de-constancia-de-investimentos-e-maior-integracao-entre-o>>

O Autor. **Visualizador de Ecossistemas de Inovação**. 2025. Disponível em: <<https://bit.ly/VisualizadorEcoInov>>.

OECD. **Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development**. Paris: OECD Publishing, 2015. (The Measurement of Scientific, Technological and Innovation Activities). Disponível em: <<https://doi.org/10.1787/9789264239012-en>>.

_____. **Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation**. 4. ed. Paris: OECD Publishing, 2018. (The Measurement of Scientific, Technological and Innovation Activities). Disponível em: <<https://doi.org/10.1787/9789264304604-en>>.

OKFN. **Open Knowledge Foundation – For a fair, sustainable and open future**. 2025. Acesso em: 07 jun. 2025. Disponível em: <<https://okfn.org/>>.

Oracle. **O que é ETL?** 2025. Acesso em: 13 jun. 2025. Disponível em: <<https://www.oracle.com/br/integration/what-is-etl/>>.

PAIVA, M. S. d. *et al.* Inovação e os efeitos sobre a dinâmica de mercado: uma síntese teórica de smith e schumpeter. **Interações (Campo Grande)**, v. 19, n. 1, p. 155–170, 2018. ISSN 1518-7012. Acesso em: 6 abr. 2025. Disponível em: <<https://www.scielo.br/j/inter/a/DVkwShDFG99PSxN3tjrndcq/>>.

PERINI, R. d. L. **Análise integrada do sistema de capitais de Detroit: inovação e transformação de uma cidade industrial em cidade do conhecimento**. Tese (Doutorado) — Universidade de Caxias do Sul, 2023. Acesso em: 08 jun. 2025. Disponível em: <<https://repositorio.ucs.br/xmlui/bitstream/handle/11338/12607/Tese%20Rafael%20de%20Lucena%20Perini.pdf?sequence=1&isAllowed=y>>.

Porto Digital. **Porto Digital – Parque Tecnológico e Distrito de Inovação**. Recife, Pernambuco, Brasil: Porto Digital, 2025. Acesso em: 6 abr. 2025. Disponível em: <<https://www.portodigital.org>>.

PRADO, M. A. R. d.; CASTANHA, R. C. G. Indicadores: Conceitos fundamentais e importância em ct&i. In: GRÁCIO, M. C. C. *et al.* (Ed.). **Tópicos da Bibliometria para Bibliotecas Universitárias**. Marília, SP: Oficina Universitária; Cultura Acadêmica, 2020. p. 50–71. Acesso em: 18 abr. 2025. Disponível em: <https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/download/187/997/1761?inline=1>.

ROCHA, L. A. *et al.* O impacto dos investimentos em pesquisa e desenvolvimento no desempenho das empresas. **Revista de Economia Contemporânea**, v. 20, n. 1, p. 58–91, 2016. Acesso em: 08 jun. 2025. Disponível em: <<https://doi.org/10.1590/198055272013>>.

ROVERE, R. L. L.; SANTOS, G. de O.; VASCONCELLOS, B. L. X. Desafios para a mensuração de ecossistemas de inovação e de ecossistemas de empreendedorismo no Brasil. **Revista de Empreendedorismo e Gestão de Pequenas Empresas**, v. 10, n. 1, p. 1–24, 2021. ISSN 2316-2058. Acesso em: 27 abr. 2025. Disponível em: <<https://dialnet.unirioja.es/servlet/articulo?codigo=7929441>>.

RUNESON, P.; OLSSON, T.; LINÅKER, J. Open data ecosystems – an empirical investigation into an emerging industry collaboration concept. **Journal of Systems and Software**, v. 182, p. 111088, 2021. Acesso em: 08 jun. 2025. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0164121221001850>>.

SCHWAMBACH, C.; SOBRINHO, G. C. **Biologia**. 1. ed. Curitiba: Intersaberes, 2017. E-book, Disponível em: <<https://plataforma.bvirtual.com.br>>, Acesso em: 13 mar. 2025.

Sebrae. **O que é inovação? Descubra seu conceito, tipos e muito mais**. 2025. Acesso em: 19 mar. 2025. Disponível em: <https://sebrae.com.br/sites/PortalSebrae/ufs/sp/licitacao_uf/o-que-e-inovacao-descubra-seu-conceito-tipos-e-muito-mais,1184cfe7f0c1c810VgnVCM1000001b00320aRCRD>.

Secretaria do Meio Ambiente de São Paulo. **Ecossistema**. 2025. Acesso em: 19 mar. 2025. Disponível em: <<https://semil.sp.gov.br/educacaoambiental/prateleira-ambiental/ecossistema/>>.

SILVA, D. J. C. d. **Título da Tese**. Tese (Doutorado) — Universidade Federal de Santa Maria, 2023. Acesso em: 26 mai. 2025. Disponível em: <https://repositorio.ufsm.br/bitstream/handle/1/31419/TES_PPGADMINISTRACAO_2023_SILVA_DEOCLECIO.pdf?sequence=1&isAllowed=y>.

SPINOSA, L.-M.; KRAMA, M.-R.; HARDT, C. Desenvolvimento urbano baseado em conhecimento e ecossistemas de inovação urbanos: uma análise em quatro cidades brasileiras. **EURE (Santiago)**, v. 44, n. 131, p. 193–214, 2018. ISSN 0250-7161. Acesso em: 6 abr. 2025. Disponível em: <https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0250-71612018000100193&lng=en&nrm=iso&tlng=en>.

Tecnopuc. **Tecnopuc – Parque Científico e Tecnológico da PUCRS**. 2025. Acesso em: 6 abr. 2025. Disponível em: <<https://tecnopuc.pucrs.br>>.

TecnoUCS . **Quem Somos**. 2025. Acesso em: 6 abr. 2025. Disponível em: <<https://www.ucs.br/site/tecnoucs/quem-somos/>>.

VALOTO, D. de S.; MAY, M. R. Indicadores de inovação tecnológica para seleção de startups em ecossistemas de inovação: uma revisão bibliográfica. **Revista Livre de Sustentabilidade e Empreendedorismo**, v. 5, p. 165–193, 2020. Acesso em: 27 abr. 2025. Disponível em: <<https://dialnet.unirioja.es/servlet/articulo?codigo=7929441>>.

World Intellectual Property Organization (WIPO). **WIPO – World Intellectual Property Organization**. 2025. Acesso em: 6 abr. 2025. Disponível em: <<https://www.wipo.int/portal/en/>>.

ŚLEDZIK, K. Schumpeter's view on innovation and entrepreneurship. In: HITTMÁR, (Ed.). **Management Trends in Theory and Practice: Scientific Papers**. Žilina: Faculty of Management Science and Informatics, University of Žilina, 2013. p. 144–147. ISBN 978-80-554-0736-4. Acesso em: 6 abr. 2025. Disponível em: <<https://repozytorium.bg.ug.edu.pl/info/article/UOG41f8eec4650c4ee08e43b47fec31d43/Schumpeter%25E2%2580%2599s%2Bview%2Bon%2Binnovation%2Band%2Bentrepreneurship?lang=en&r=publication&tab=>>>.

A INDICADORES DE INOVAÇÃO DA ORGANIZAÇÃO MUNDIAL DE PROPRIEDADE INTELECTUAL

Tabela 6 – Indicadores do GII

Categoria	Subcategoria	Nº	Título do Indicador	Fontes
Instituições	Ambiente institucional	1	Estabilidade operacional para negócios	WIPO / diversas
Instituições	Ambiente institucional	2	Eficiência do governo	WIPO / diversas
Instituições	Ambiente regulatório	3	Qualidade regulatória	WIPO / diversas
Instituições	Ambiente regulatório	4	Estado de direito	WIPO / diversas
Instituições	Ambiente de negócios	5	Estabilidade política para negócios	WIPO / diversas
Instituições	Ambiente de negócios	6	Políticas e cultura empreendedoras	WIPO / diversas
Capital humano e pesquisa	Educação	7	Gasto do governo em educação (% do PIB)	WIPO / diversas
Capital humano e pesquisa	Educação	8	Financiamento por aluno do ensino médio (% do PIB per capita)	WIPO / diversas
Capital humano e pesquisa	Educação	9	Expectativa de vida escolar, do primário ao superior (anos)	WIPO / diversas
Capital humano e pesquisa	Educação	10	Notas médias do Programa Internacional de Avaliação de Alunos (PISA) em leitura, matemática e ciências	WIPO / diversas
Capital humano e pesquisa	Educação	11	Relação aluno-professor no ensino médio	WIPO / diversas
Capital humano e pesquisa	Ensino superior	12	Matrícula no ensino superior (% bruto)	WIPO / diversas
Capital humano e pesquisa	Ensino superior	13	Formandos em STEM (% do total de formandos do ensino superior)	WIPO / diversas
Capital humano e pesquisa	Ensino superior	14	Taxa de mobilidade internacional de estudantes (%)	WIPO / diversas
Capital humano e pesquisa	PeD	15	Pesquisadores em equivalência de tempo integral (por milhão de hab.)	WIPO / diversas
Capital humano e pesquisa	PeD	16	Despesa bruta em PeD (% do PIB)	WIPO / diversas
Capital humano e pesquisa	PeD	17	Média de despesas em PeD das três maiores empresas (% da receita)	WIPO / diversas
Capital humano e pesquisa	PeD	18	Pontuação média das três melhores universidades no ranking Quacquarelli Symonds (QS)	QS World University Rankings
Infraestrutura	Tecnologias da Informação e Comunicações (TICs)	19	Índice de acesso a TIC	União Internacional de Telecomunicações (UIT) / diversas
Infraestrutura	TICs	20	Índice de uso de TIC	UIT / diversas
Infraestrutura	TICs	21	Índice de serviços online do governo	Organização das Nações Unidas (ONU) e-Government Survey
Infraestrutura	TICs	22	Índice de participação eletrônica	ONU e-Government Survey
Infraestrutura	Infraestrutura geral	23	Produção de eletricidade (GWh por milhão de habitantes)	WIPO / diversas
Infraestrutura	Infraestrutura geral	24	Índice de desempenho logístico	Banco Mundial

Continua na próxima página

Tabela 6 – Indicadores do GII (continuação)

Categoria	Subcategoria	Nº	Título do Indicador	Fontes
Infraestrutura	Infraestrutura geral	25	Formação bruta de capital fixo (% do PIB, média de 3 anos)	Banco Mundial
Infraestrutura	Sustentabilidade ecológica	26	PIB por fornecimento total de energia (milhares de dólares Paridade do Poder de Compra (PPC) por kg)	Agência Internacional de Energia
Infraestrutura	Sustentabilidade ecológica	27	Participação de fontes de baixa intensidade de carbono na energia (%)	Agência Internacional de Energia
Infraestrutura	Sustentabilidade ecológica	28	Certificados Organização Internacional de Normalização (ISO) 14001 por bilhão de dólares PPC	ISO
Sofisticação mercado	do Crédito	29	Financiamento para <i>startups</i> e <i>scaleups</i>	WIPO / diversas
Sofisticação mercado	do Crédito	30	Crédito doméstico ao setor privado (% do PIB)	Banco Mundial
Sofisticação mercado	do Crédito	31	Empréstimos de instituições de microfinanças (% do PIB)	Microfinance Information Exchange
Sofisticação mercado	do Investimento	32	Capitalização de mercado de empresas listadas (% do PIB)	Banco Mundial
Sofisticação mercado	do Investimento	33	Número de negócios de capital de risco investidos por milhão de hab.	WIPO / diversas
Sofisticação mercado	do Investimento	34	Número de negócios de capital de risco recebidos por milhão de hab.	WIPO / diversas
Sofisticação mercado	do Investimento	35	Valor total do capital de risco recebido (% do PIB)	WIPO / diversas
Sofisticação mercado	do Comércio e escala de mercado	36	Tarifa média aplicada ponderada, todos os produtos (%)	Banco Mundial
Sofisticação mercado	do Comércio e escala de mercado	37	Diversificação industrial doméstica	WIPO / diversas
Sofisticação mercado	do Comércio e escala de mercado	38	Escala do mercado interno (PIB em bilhões de dólares PPC)	Banco Mundial
Sofisticação dos negócios	Trabalhadores do conhecimento	39	Emprego em serviços intensivos em conhecimento (% do total)	WIPO / diversas
Sofisticação dos negócios	Trabalhadores do conhecimento	40	Empresas que oferecem treinamento formal (% das empresas)	Banco Mundial
Sofisticação dos negócios	Trabalhadores do conhecimento	41	Despesa em PeD realizada por empresas (% do PIB)	WIPO / diversas
Sofisticação dos negócios	Trabalhadores do conhecimento	42	PeD financiada por empresas (% da PeD total)	WIPO / diversas
Sofisticação dos negócios	Trabalhadores do conhecimento	43	Mulheres empregadas com pós-graduação (% dos empregados com pós-graduação)	WIPO / diversas
Sofisticação dos negócios	Vínculos de inovação	44	Publicações de pesquisa coautoria público-privada (% do total)	Scopus / Elsevier
Sofisticação dos negócios	Vínculos de inovação	45	Colaboração entre empresas e universidades em PeD	WIPO / diversas
Sofisticação dos negócios	Vínculos de inovação	46	Amplitude dos clusters	WIPO / diversas
Sofisticação dos negócios	Vínculos de inovação	47	Acordos de joint ventures e alianças estratégicas por bilhão de dólares PPC	WIPO / diversas
Sofisticação dos negócios	Vínculos de inovação	48	Famílias de patentes registradas em ao menos dois escritórios (por bilhão de dólares PPC)	WIPO

Continua na próxima página

Tabela 6 – Indicadores do GII (continuação)

Categoria	Subcategoria	Nº	Título do Indicador	Fontes
Sofisticação dos negócios	Absorção de conhecimento	49	Pagamentos por uso de propriedade intelectual (% do comércio total)	Organização Mundial do Comércio (OMC)
Sofisticação dos negócios	Absorção de conhecimento	50	Importações de alta tecnologia (% do comércio total)	Comtrade / ONU
Sofisticação dos negócios	Absorção de conhecimento	51	Importações de serviços TIC (% do comércio)	Comtrade / ONU
Sofisticação dos negócios	Absorção de conhecimento	52	Investimento direto estrangeiro líquido (% do PIB)	Banco Mundial
Sofisticação dos negócios	Absorção de conhecimento	53	Pesquisadores em empresas (%)	WIPO / diversas
Resultados em conhecimento e tecnologia	Criação de conhecimento	54	Pedidos de patentes residentes (por bilhão de dólares PPC)	WIPO
Resultados em conhecimento e tecnologia	Criação de conhecimento	55	Pedidos de patentes Tratado de Cooperação em Matéria de Patentes (PCT) (por bilhão de dólares PPC)	WIPO
Resultados em conhecimento e tecnologia	Criação de conhecimento	56	Pedidos de modelos de utilidade residentes (por bilhão de dólares PPC)	WIPO
Resultados em conhecimento e tecnologia	Criação de conhecimento	57	Artigos científicos e técnicos (por bilhão de dólares PPC)	Scopus / Elsevier
Resultados em conhecimento e tecnologia	Criação de conhecimento	58	Índice H de citações científicas	Scopus / Elsevier
Resultados em conhecimento e tecnologia	Impacto do conhecimento	59	Crescimento do PIB por pessoa empregada (% , média de 5 anos)	Banco Mundial
Resultados em conhecimento e tecnologia	Impacto do conhecimento	60	Avaliação total dos unicórnios do país (% do PIB)	CB Insights / Crunchbase
Resultados em conhecimento e tecnologia	Impacto do conhecimento	61	Gastos com software de computador (% do PIB)	WIPO
Resultados em conhecimento e tecnologia	Impacto do conhecimento	62	Manufatura de média e alta tecnologia (% do total)	UNIDO
Resultados em conhecimento e tecnologia	Difusão de conhecimento	63	Receitas por uso de propriedade intelectual (% do comércio total)	OMC
Resultados em conhecimento e tecnologia	Difusão de conhecimento	64	Índice de Complexidade Econômica	Instituto de Tecnologia de Massachusetts (MIT)
Resultados em conhecimento e tecnologia	Difusão de conhecimento	65	Exportações de alta tecnologia (% do comércio total)	Comtrade / ONU
Resultados em conhecimento e tecnologia	Difusão de conhecimento	66	Exportações de serviços TIC (% do comércio total)	Comtrade / ONU

Continua na próxima página

Tabela 6 – Indicadores do GII (continuação)

Categoria	Subcategoria	Nº	Título do Indicador	Fontes
Resultados em conhecimento e tecnologia	Difusão de conhecimento	67	Certificados ISO 9001 por bilhão de dólares PPC	ISO
Resultados criativos	Ativos intangíveis	68	Valor dos ativos intangíveis como % do valor total das empresas	WIPO / diversas
Resultados criativos	Ativos intangíveis	69	Número de classes em pedidos de marcas residentes (por bilhão de dólares PPC)	WIPO
Resultados criativos	Ativos intangíveis	70	Valor global de marcas das 5.000 maiores (% do PIB)	Brand Finance
Resultados criativos	Ativos intangíveis	71	Número de desenhos industriais residentes (por bilhão de dólares PPC)	WIPO
Resultados criativos	Bens e serviços criativos	72	Exportações de serviços culturais e criativos (% do comércio total)	Comtrade / ONU
Resultados criativos	Bens e serviços criativos	73	Número de longas-metragens nacionais produzidos (por milhão de hab.)	Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO)
Resultados criativos	Bens e serviços criativos	74	Investimento em mídia e telecomunicações (por milhão de hab.)	PwC Outlook
Resultados criativos	Bens e serviços criativos	75	Exportações de bens criativos (% do comércio)	Comtrade / ONU
Resultados criativos	Criatividade online	76	Domínios <i>Top-Level Domain</i> (TLD) genéricos e com código de país (por milhão de hab.)	WIPO / diversas
Resultados criativos	Criatividade online	77	Commits no GitHub (por milhão de habitantes)	GitHub
Resultados criativos	Criatividade online	78	Downloads globais de apps móveis (por bilhão de dólares PPC)	App Annie / WIPO

B PROCEDURES PARA CARGA E TRATAMENTO DOS DADOS NO BANCO

Script 10 – Procedure para criar as tabelas do banco de dados

```

1 CREATE OR REPLACE PROCEDURE public.cria_tabelas ()
2 BEGIN
3     Create Table ecossistemas_inovacao (
4         id        serial primary key,
5         nome      varchar unique not null,
6         pais      varchar not null,
7         estado    varchar
8     );
9
10    Create Table eventos (
11        id        serial primary key,
12        nome      varchar not null,
13        dta_ini   date,
14        dta_fim   date,
15        eco_id    int not null,
16        unique (nome, eco_id),
17        foreign key (eco_id) references ecossistemas_inovacao(id)
18    );
19
20    Create Table cidades (
21        id        serial primary key,
22        nome      varchar unique not null,
23        eco_id    int not null,
24        vlr_invest numeric(14,2),
25        foreign key (eco_id) references ecossistemas_inovacao(id)
26    );
27
28    Create Table instituicoes (
29        id        serial primary key,
30        sigla     varchar not null,
31        nome      varchar not null,
32        tip_inst  int not null,
33        ind_principal int not null,
34        num_matriculas int,
35        num_bolsas int,
36        site      varchar,
37        cid_id    int not null,
38        unique (nome, cid_id),
39        foreign key (cid_id) references cidades(id)
40    );
41    Comment on column instituicoes.tip_inst is
42    'Indica tipo da instituicao: 1 = Ensino, 2 = Pesquisa';

```

```

43      Comment on column instituicoes.ind_principal is
44      'Indica se e o campus principal: 0 = Nao, 1 = Sim';
45
46      Create Table ambientes_inov (
47          id          serial primary key,
48          sigla       varchar not null,
49          nome        varchar,
50          tip_amb     int not null,
51          site        varchar,
52          cid_id      int not null,
53          unique (sigla, cid_id),
54          foreign key (cid_id) references cidades(id)
55      );
56      Comment on column ambientes_inov.tip_amb is
57      '1 = Aceleradora, 2 = Incubadora, 3 = Parque tecnologico, 4 = Outros';
58
59      Create Table startups (
60          id          serial primary key,
61          nome        varchar not null,
62          area        varchar,
63          site        varchar,
64          cid_id      int not null,
65          amb_id      int,
66          unique (nome),
67          foreign key (cid_id) references cidades(id),
68          foreign key (amb_id) references ambientes_inov(id)
69      );
70
71      Create Table pessoas (
72          id          serial primary key,
73          nome        varchar not null,
74          ind_pesq    int,
75          tip_titulo  int,
76          eco_id      int not null,
77          foreign key (eco_id) references ecossistemas_inovacao(id)
78      );
79
80      Comment on column pessoas.ind_pesq is
81      'Indica se e pesquisador: 0 = Nao, 1 = Sim';
82      Comment on column pessoas.tip_titulo is
83      '1 = Especializacao, 2 = Mestrado, 3 = Doutorado, 4 = Outro ou nao
84         informado';
85
86      Create Table producoes (
87          id          serial primary key,
88          nome        varchar not null,
89          tip_prod    int not null,
90          autor       varchar not null,
          inst_id      int not null,

```

```

91         unique (nome, autor),
92         foreign key (inst_id) references instituicoes(id)
93     );
94     Comment on column producoes.tip_prod is
95     '1 = Patente, 2 = Projeto de Pesquisa, 3 = Tese, 4 = Dissertacao';
96
97     Create Table empresas_alto_cresc (
98         id            serial primary key,
99         nome          varchar not null,
100        area          varchar,
101        receita_24    numeric(14,2),
102        receita_23    numeric(14,2),
103        cid_id        int not null,
104        unique (nome, cid_id),
105        foreign key (cid_id) references cidades(id)
106    );
107
108     Create Table entidades_gov (
109         id            serial primary key,
110        entidade      varchar not null,
111        nome_uni      varchar,
112        cid_id        int not null,
113        unique (nome_uni, cid_id),
114        foreign key (cid_id) references cidades(id)
115    );
116
117 END;
```

Fonte: O autor (2025)

Script 11 – Código base para leitura dos arquivos CSV em SQL

```

1  ...
2  DECLARE
3      i            int;
4      wconteudo   text;
5      wlin        text;
6      wlinhas     text [];
7      wcols       text [];
8      ...
9  BEGIN
10     wconteudo := pg_read_file(PI_ARQ);
11     -- Tira a quebra de linha
12     wconteudo := replace(wconteudo, E'\r', '');
13     -- Quebra por linhas
14     wlinhas := string_to_array(wconteudo, E'\n');
15
16     -- Itera pelas linhas
17     FOR i IN 2..array_upper(wlinhas, 1) LOOP
18         -- Pega linha atual (sem espacos antes ou depois)
```

```

19     wlin := trim(wlinhas[i]);
20     -- Reseta so para evitar duplicados
21     wcols := NULL;
22     IF wlin IS NOT NULL AND wlin <> '' THEN
23         -- Separa as colunas em Ecosistema e Nome
24         wcols := string_to_array(wlin, ';');
25         IF wcols IS NOT NULL THEN
26             ...
27         END IF;
28     END IF;
29 END LOOP;
30 END

```

Fonte: O autor (2025)

Script 12 – Procedure para carregar os ecossistemas de inovação para o banco

```

1 CREATE OR REPLACE PROCEDURE public.carrega_ecossistemas(IN pi_arq text)
2 DECLARE
3     wconteudo text;
4     wcols      text[];
5     i          int;
6 BEGIN
7     wconteudo := pg_read_file(PI_ARQ);
8     -- Tira a quebra de linha
9     wconteudo := replace(wconteudo, E'\r', '');
10    -- Separa pelos ";"
11    wcols := string_to_array(wconteudo, ';');
12
13    -- Loop pra pegar os nomes
14    FOR i IN array_lower(wcols, 1)..array_upper(wcols, 1) LOOP
15        INSERT INTO ecossistemas_inovacao(NOME, PAIS, ESTADO)
16        VALUES(wcols[i], 'Brasil', 'Rio Grande do Sul');
17    END LOOP;
18 END

```

Fonte: O autor (2025)

Script 13 – Procedure para carregar dados dos eventos

```

1 CREATE OR REPLACE PROCEDURE public.carrega_eventos(IN pi_arq text)
2 DECLARE
3     weco_id int;
4 BEGIN
5     ....
6     BEGIN
7         SELECT ecoi.id INTO STRICT weco_id
8         FROM ecossistemas_inovacao ecoi
9         WHERE UPPER(ecoi.nome) LIKE UPPER(wcols[1]);
10    EXCEPTION WHEN NO_DATA_FOUND THEN
11        RAISE EXCEPTION 'Ecosistema nao encontrado: %', wcols[1];

```

```

12     END;
13
14     BEGIN
15         INSERT INTO eventos (NOME , DTA_INI , DTA_FIM , ECO_ID)
16         VALUES(wcols[2] , TO_DATE(wcols[3], 'DD/MM/YYYY') , TO_DATE(wcols
           [4], 'DD/MM/YYYY') , weco_id);
17     EXCEPTION WHEN OTHERS THEN
18         RAISE EXCEPTION 'Nao foi possivel incluir evento: %' , wcols[2];
19     END;
20     ....
21 END

```

Fonte: O autor (2025)

Script 14 – Procedure para carregar dados das cidades

```

1 CREATE OR REPLACE PROCEDURE public.carrega_cidades(IN pi_arq text)
2 DECLARE
3     ....
4     weco_id    int;
5 BEGIN
6     ....
7     BEGIN
8         SELECT ecoi.id
9         INTO STRICT weco_id
10        FROM ecossistemas_inovacao ecoi
11        WHERE UPPER(ecoi.nome) LIKE UPPER(wcols[1]);
12    EXCEPTION WHEN NO_DATA_FOUND THEN
13        RAISE EXCEPTION 'Ecosistema nao encontrado: %' , wcols[1];
14    END;
15
16    INSERT INTO cidades (NOME, ECO_ID)
17    VALUES(wcols[2], weco_id);
18    ....
19 END

```

Fonte: O autor (2025)

Script 15 – Procedure para carregar dados dos investimentos em PeD

```

1 CREATE OR REPLACE PROCEDURE public.carrega_invest(IN pi_arq text)
2 DECLARE
3     ....
4     wcid_id    int;
5 BEGIN
6     ....
7     BEGIN
8         SELECT cid.id
9         INTO STRICT wcid_id
10        FROM cidades cid
11        WHERE UPPER(TRIM(cid.nome)) LIKE UPPER(TRIM(wcols[1]));

```

```

12      EXCEPTION WHEN NO_DATA_FOUND THEN
13          RAISE EXCEPTION 'Cidade nao encontrada: %', wcols[1];
14      END;
15
16      BEGIN
17          UPDATE cidades
18              SET vlr_invest = wvlr_invest
19              WHERE id = wcid_id;
20      END;
21      ....
22  END

```

Fonte: O autor (2025)

Script 16 – Procedure para carregar dados das instituições

```

1  CREATE OR REPLACE PROCEDURE public.carrega_instituicoes (IN pi_arq text)
2  DECLARE
3      ....
4      wcid_id    int;
5  BEGIN
6      ....
7      BEGIN
8          SELECT cid.id
9              INTO STRICT wcid_id
10             FROM cidades cid
11             WHERE UPPER(TRIM(cid.nome)) LIKE UPPER(TRIM(wcols[3]));
12      EXCEPTION WHEN NO_DATA_FOUND THEN
13          RAISE EXCEPTION 'Cidade nao encontrada: %', wcols[3];
14      END;
15
16      BEGIN
17          INSERT INTO instituicoes (sigla, nome, tip_inst, cid_id,
18                                  ind_principal)
19              VALUES (wcols[1], wcols[2], 1, wcid_id, 1);
20      EXCEPTION
21          WHEN OTHERS THEN
22              RAISE EXCEPTION 'Erro ao inserir instituicao: %, Detalhe: %',
23                  wcols[2], SQLERRM;
24      END;
25      ....
26  END

```

Fonte: O autor (2025)

Script 17 – Procedure para carregar dados das matriculas

```

1  CREATE OR REPLACE PROCEDURE public.carrega_matriculas (IN pi_arq text)
2  DECLARE
3      ....
4      wcid_id    int;

```

```

5      wtip_inst int;
6      wnum_mats int;
7 BEGIN
8      ....
9      wnum_mats := NULLIF(wcols[3], '')::int;
10
11     BEGIN
12         SELECT cid.id
13             INTO STRICT wcid_id
14             FROM cidades cid
15             WHERE UPPER(TRIM(cid.nome)) LIKE UPPER(TRIM(wcols[5]));
16     EXCEPTION WHEN NO_DATA_FOUND THEN
17         RAISE EXCEPTION 'Ecosystema nao encontrado: %', wcols[1];
18     END;
19
20     BEGIN
21         UPDATE instituicoes
22             SET num_matriculas = wnum_mats
23                 , site = wcols[6]
24             WHERE UPPER(nome) LIKE UPPER(TRIM(wcols[4]))
25                 AND cid_id = wcid_id;
26
27     IF NOT FOUND THEN
28         -- Faz update nas matriculas pela sigla para evitar
29         -- registros com valor nulo onde nao deveria ser
30         UPDATE instituicoes
31             SET num_matriculas = wnum_mats
32             WHERE UPPER(sigla) LIKE UPPER(TRIM(wcols[2]));
33         -- Mesmo assim insere um novo registro
34         -- com o campus novo
35     BEGIN
36         SELECT i.tip_inst
37             INTO STRICT wtip_inst
38             FROM instituicoes i
39             WHERE UPPER(TRIM(i.sigla)) LIKE UPPER(TRIM(wcols[2]))
40             LIMIT 1;
41     EXCEPTION
42         WHEN NO_DATA_FOUND THEN
43             wtip_inst := 2;
44     END;
45
46     BEGIN
47         INSERT INTO instituicoes (SIGLA, NOME, TIP_INST,
48             IND_PRINCIPAL, NUM_MATRICULAS, SITE, CID_ID)
49             VALUES(wcols[2], wcols[4], wtip_inst, 0, wnum_mats,
50                 wcols[6], wcid_id);
51     EXCEPTION WHEN OTHERS THEN
52         RAISE EXCEPTION 'Erro ao inserir instituicao: %, Detalhe
53             : %', wcols[4], SQLERRM;

```

```

51         END;
52     END IF;
53     ....
54 END

```

Fonte: O autor (2025)

Script 18 – Procedure para carregar dados das bolsas concedidas

```

1 CREATE OR REPLACE PROCEDURE public.carrega_bolsas(IN pi_arq text)
2 DECLARE
3     ...
4     wcid_id      int;
5     winst_id     int;
6     wnum_bolsas int;
7     wsigla_inst text;
8 BEGIN
9     ...
10    IF wcols[3] <> '-' THEN
11        wnum_bolsas := NULLIF(wcols[3], '')::int;
12        wsigla_inst := TRIM(split_part(wcols[1], '-', 1));
13
14        BEGIN
15            UPDATE instituicoes
16                SET num_bolsas = COALESCE(num_bolsas, 0) + wnum_bolsas
17                WHERE UPPER(sigla) LIKE UPPER(wsigla_inst);
18
19            IF NOT FOUND THEN
20                RAISE EXCEPTION 'Instituicao nao encontrada: %', wcols[1];
21            END IF;
22        END;
23    END IF;
24    ...
25 END

```

Fonte: O autor (2025)

Script 19 – Procedure para carregar dados dos ambientes

```

1 CREATE OR REPLACE PROCEDURE public.carrega_ambientes(IN pi_arq text, IN
2     pi_tip_amb integer)
3 DECLARE
4     ...
5     wcid_id      int;
6 BEGIN
7     ...
8     BEGIN
9         SELECT cid.id
10            INTO STRICT wcid_id
11           FROM cidades cid
12          WHERE UPPER(TRIM(cid.nome)) LIKE UPPER(TRIM(wcols[4]));

```

```

12      EXCEPTION WHEN NO_DATA_FOUND THEN
13          RAISE EXCEPTION 'Cidade nao encontrada: %', wcols[4];
14      END;
15
16      BEGIN
17          INSERT INTO ambientes_inov (sigla, nome, tip_amb, site, cid_id)
18          VALUES (wcols[2], wcols[3], pi_tip_amb, wcols[5], wcid_id);
19      EXCEPTION
20          WHEN OTHERS THEN
21              RAISE EXCEPTION 'Erro ao inserir ambiente: %, Detalhe: %', wcols
22                  [3], SQLERRM;
23      END;
24      ....
25  END

```

Fonte: O autor (2025)

Script 20 – Procedure para carregar dados das startups

```

1  CREATE OR REPLACE PROCEDURE public.carrega_startups (IN pi_arq text)
2  DECLARE
3      ....
4      wcid_id    int;
5      wamb_id    int;
6  BEGIN
7      ....
8          BEGIN
9              SELECT cid.id
10             INTO STRICT wcid_id
11             FROM cidades cid
12             WHERE UPPER(TRIM(cid.nome)) LIKE UPPER(TRIM(wcols[2]));
13          EXCEPTION
14              WHEN NO_DATA_FOUND THEN
15                  RAISE EXCEPTION 'Cidade nao encontrada: %', wcols[2];
16              WHEN OTHERS THEN
17                  RAISE EXCEPTION 'ERRO Cidade: %; %', wcols[2], SQLERRM;
18          END;
19
20          IF UPPER(wcols[10]) LIKE 'NAO INFORMADO' THEN
21              wamb_id := NULL;
22          ELSE
23              BEGIN
24                  SELECT amb.id
25                 INTO STRICT wamb_id
26                 FROM ambientes_inov amb
27                 WHERE UPPER(TRIM(amb.sigla)) LIKE '%' || REPLACE(UPPER(TRIM
28                     (wcols[10])), ' ', '%') || '%'
29                 LIMIT 1;
30          EXCEPTION WHEN NO_DATA_FOUND THEN
31              -- Ve se ja tem com mesma sigla, mas cidade diferente

```

```

31         BEGIN
32             SELECT amb.id
33                 INTO STRICT wamb_id
34                 FROM ambientes_inov amb
35                 WHERE UPPER(TRIM(amb.sigla)) LIKE '%' || REPLACE(
36                     UPPER(TRIM(wcols[10])), ',', '%') || '%'
37                 LIMIT 1;
38         EXCEPTION WHEN NO_DATA_FOUND THEN
39             -- Insere ambiente como "Outro"
40             INSERT INTO ambientes_inov (sigla, tip_amb, cid_id)
41             VALUES (wcols[10], 4, wcid_id);
42             -- Pega o ID novo do ambiente adicionado
43             SELECT amb.id
44                 INTO STRICT wamb_id
45                 FROM ambientes_inov amb
46                 WHERE UPPER(TRIM(amb.sigla)) LIKE '%' || REPLACE(
47                     UPPER(TRIM(wcols[10])), ',', '%') || '%'
48                 AND amb.cid_id = wcid_id
49                 LIMIT 1;
50         END;
51     WHEN OTHERS THEN
52         RAISE EXCEPTION 'ERRO Ambiente: %; %', wcols[10], SQLERRM;
53     END;
54 END IF;
55
56 BEGIN
57     INSERT INTO startups (nome, area, site, cid_id, amb_id)
58     VALUES (wcols[9], wcols[6], wcols[3], wcid_id, wamb_id);
59 EXCEPTION
60     WHEN OTHERS THEN
61         RAISE EXCEPTION 'Erro ao inserir startup: %, Detalhe: %',
62             wcols[9], SQLERRM;
63     END;
64 END IF;
65
66 BEGIN
67     INSERT INTO startups (nome, area, site, cid_id, amb_id)
68     VALUES (wcols[9], wcols[6], wcols[3], wcid_id, wamb_id);
69 EXCEPTION
70     WHEN OTHERS THEN
71         RAISE EXCEPTION 'Erro ao inserir startup: %, Detalhe: %', wcols
72             [9], SQLERRM;
73     END;
74
75     ....
76 END

```

Fonte: O autor (2025)

Script 21 – Procedure para carregar dados das pessoas

```

1 CREATE OR REPLACE PROCEDURE public.carrega_pessoas(IN pi_arq text)
2 DECLARE
3     ....
4     weco_id    int;
5     winst_id   int;
6     wtitulo    int;
7     wpesq     int;
8     wprojs     text[];
9     wproj_salvo int;
10    wproj_grava text;
11 BEGIN
12     BEGIN
13         SELECT inst.id
14             INTO STRICT winst_id
15             FROM instituicoes inst
16             WHERE UPPER(TRIM(inst.sigla)) LIKE 'UCS'
17                 AND inst.ind_principal = 1;
18     EXCEPTION WHEN NO_DATA_FOUND THEN
19         RAISE EXCEPTION 'Instituicao UCS nao encontrada';
20 END;
21     ....
22     BEGIN
23         SELECT eco.id
24             INTO weco_id
25             FROM ecossistemas_inovacao eco
26             WHERE UPPER(eco.nome) LIKE UPPER(TRIM(wcols[1]));
27     EXCEPTION WHEN NO_DATA_FOUND THEN
28         RAISE EXCEPTION 'Ecosistema nao encontrado: %', wcols[1];
29 END;
30
31     wtitulo := CASE
32                 WHEN TRIM(wcols[3]) = 'Especializacao' THEN 1
33                 WHEN TRIM(wcols[3]) = 'Mestrado' THEN 2
34                 WHEN TRIM(wcols[3]) = 'Doutorado' THEN 3
35                 ELSE 4
36             END;
37
38     wpesq := CASE
39                 WHEN UPPER(TRIM(wcols[4])) LIKE 'SIM' THEN 1
40                 ELSE 0
41             END;
42
43     INSERT INTO pessoas(nome, ind_pesq, tip_titulo, eco_id)
44     VALUES (wcols[2], wpesq, wtitulo, weco_id);
45
46     IF wcols[5] <> '' THEN
47         wprojs := string_to_array(trim(replace(wcols[5], '"', '')), '\');
48         -- Loop para inserir projetos
49         FOR j IN 1 .. array_upper(wprojs, 1) LOOP

```

```

50         wproj_grava := TRIM(wprojs[j]);
51     BEGIN
52         INSERT INTO producoes (nome, tip_prod, autor, inst_id)
53         VALUES (wproj_grava, 2, wcols[2], winst_id);
54     EXCEPTION
55         WHEN UNIQUE_VIOLATION THEN
56             wproj_salvo := 1;
57         WHEN OTHERS THEN
58             RAISE EXCEPTION 'Erro ao inserir producao: % - %,
59                 Detalhe: %', wproj_grava, wcols[2], SQLERRM;
60     END;
61 END LOOP;
62 END IF;
63     ....
64 END

```

Fonte: O autor (2025)

Script 22 – Procedure para carregar dados das produções

```

1 CREATE OR REPLACE PROCEDURE public.carrega_producoes(IN pi_arq text, IN
2     pi_tip_prod integer)
3 DECLARE
4     ....
5     weco_id    int;
6     winst_id   int;
7     wtip_prod  int;
8     wautores   text[];
9     waut_grava text;
10    wexiste    boolean;
11 BEGIN
12     ....
13     BEGIN
14         SELECT inst.id
15         INTO STRICT winst_id
16         FROM instituicoes inst
17         WHERE UPPER(TRIM(inst.sigla)) LIKE UPPER(TRIM(wcols[1]))
18         AND inst.ind_principal = 1;
19     EXCEPTION WHEN NO_DATA_FOUND THEN
20         RAISE EXCEPTION 'Instituicao nao encontrada: %', wcols[1];
21     END;
22     -- Se for publicacao, tem campo extra indicando se e tese ou dissertacao
23     IF pi_tip_prod = 3 THEN
24         wtip_prod := CASE WHEN wcols[4] = 'TESE' THEN 3 ELSE 4 END;
25         -- Tira os colchetes e separa pela virgula
26         wautores := string_to_array(trim(replace(replace(wcols[5], '[' , ''),
27             ']', '')), ',');
28     ELSE
29         wtip_prod = pi_tip_prod;

```

```

29         wautores := string_to_array(trim(replace(replace(wcols[4], '[' , ''),
30             ']', '')), ',');
31     END IF;
32     -- Loop para inserir autores separadamente
33     FOR j IN 1 .. array_upper(wautores, 1) LOOP
34         waut_grava := TRIM(replace(wautores[j], '''' , ''));
35         BEGIN
36             INSERT INTO producoes (nome, tip_prod, autor, inst_id)
37             VALUES (wcols[3], wtip_prod, waut_grava, winst_id);
38         EXCEPTION
39             WHEN UNIQUE_VIOLATION THEN
40                 registro_duplicado_na_api := 1;
41             WHEN OTHERS THEN
42                 RAISE EXCEPTION 'Erro ao inserir producao: % - %, Detalhe: %
43                 ', wtip_prod, wcols[3], SQLERRM;
44         END;
45     IF pi_tip_prod = 2 THEN
46         SELECT EXISTS (
47             SELECT 1
48             FROM pessoas
49             WHERE unaccent(UPPER(nome)) LIKE unaccent(UPPER(
50                 waut_grava))
51             ) INTO wexiste;
52     IF NOT wexiste THEN
53         BEGIN
54             SELECT eco.id
55             INTO weco_id
56             FROM instituicoes inst
57             JOIN cidades c
58             ON c.id = inst.cid_id
59             JOIN ecossistemas_inovacao eco
60             ON eco.id = c.eco_id
61             WHERE inst.id = winst_id;
62         EXCEPTION WHEN NO_DATA_FOUND THEN
63             RAISE EXCEPTION 'Ecosistema nao encontrado para
64             pesquisador: %', waut_grava;
65         END;
66     INSERT INTO PESSOAS(nome, ind_pesq, tip_titulo, eco_id)
67     VALUES (waut_grava, 1, 4, weco_id);
68     END IF;
69     END IF;
70     END LOOP;
71     ....
72 END

```

Script 23 – Procedure para carregar dados das empresas de alto crescimento

```

1 CREATE OR REPLACE PROCEDURE public.carrega_alto_cresc (IN pi_arq text)
2 DECLARE
3     ....
4     wcid_id    int;
5     r_alto_cresc empresas_alto_cresc%ROWTYPE;
6 BEGIN
7     ....
8     BEGIN
9         SELECT cid.id
10            INTO wcid_id
11           FROM cidades cid
12          WHERE UPPER(cid.nome) LIKE UPPER(TRIM(wcols[3]));
13     EXCEPTION WHEN NO_DATA_FOUND THEN
14         RAISE EXCEPTION 'Cidade nao encontrada: %', wcols[3];
15     END;
16
17     r_alto_cresc := NULL;
18
19     SELECT nextval('empresas_alto_cresc_id_seq') INTO r_alto_cresc.id;
20
21     r_alto_cresc.nome := TRIM(wcols[2]);
22     r_alto_cresc.area := TRIM(wcols[5]);
23     r_alto_cresc.receita_24 := replace(TRIM(wcols[6]), '.', '')::numeric;
24     r_alto_cresc.receita_23 := replace(TRIM(wcols[7]), '.', '')::numeric;
25     r_alto_cresc.cid_id := wcid_id;
26
27     INSERT INTO empresas_alto_cresc
28     VALUES (r_alto_cresc.*);
29     ....
30 END

```

Script 24 – Procedure para carregar dados das entidades governamentais

```

1 CREATE OR REPLACE PROCEDURE public.carrega_gov (IN pi_arq text)
2 DECLARE
3     ....
4     wcid_id    int;
5     wnome_cid  text;
6 BEGIN
7     ....
8     BEGIN
9         wnome_cid := unaccent(replace(UPPER(TRIM(wcols[3])), ''', ''));
10
11         SELECT cid.id
12            INTO STRICT wcid_id

```

```

13         FROM cidades cid
14         WHERE unaccent(UPPER(TRIM(cid.nome))) LIKE wnome_cid;
15     EXCEPTION WHEN NO_DATA_FOUND THEN
16         RAISE EXCEPTION 'Cidade nao encontrada: %', wcols[3];
17     END;
18
19     BEGIN
20         INSERT INTO entidades_gov (entidade, nome_uni, cid_id)
21         VALUES (wcols[1], wcols[2], wcid_id);
22     EXCEPTION
23         WHEN OTHERS THEN
24         RAISE EXCEPTION 'Erro ao inserir unidade: %, Detalhe: %', wcols
25             [2], SQLERRM;
26     END;
27     ....
28 END

```

Fonte: O autor (2025)

Script 25 – Script para executar todas as procedures

```

1 CREATE OR REPLACE PROCEDURE public.exec_all()
2 BEGIN
3     -- Cria tabelas
4     CALL cria_tabelas();
5     -- Ecossistemas de inovacao
6     CALL carrega_ecossistemas('dados_ecos/ecossistemas_rs.csv');
7     -- Eventos
8     CALL carrega_eventos('dados_ecos/eventos.csv');
9     -- Cidades
10    CALL carrega_cidades('dados_ecos/cidades.csv');
11    -- Investimento em P&D
12    CALL carrega_invest('dados_ecos/invest.csv');
13    -- Instituicoes
14    CALL carrega_instituicoes('dados_ecos/univs.csv');
15    -- Matriculas
16    CALL carrega_matriculas('dados_ecos/matriculas.csv');
17    -- Bolsas
18    CALL carrega_bolsas('dados_ecos/bolsas.csv');
19    -- Aceleradoras
20    CALL carrega_ambientes('dados_ecos/aceleradoras.csv', 1);
21    -- Incubadoras
22    CALL carrega_ambientes('dados_ecos/incubadoras.csv', 2);
23    -- Parques
24    CALL carrega_ambientes('dados_ecos/parques.csv', 3);
25    -- Startups
26    CALL carrega_startups('dados_ecos/startups.csv');
27    -- Pessoas
28    CALL carrega_pessoas('dados_ecos/pessoas.csv');
29    -- Patentes

```

```
30 CALL carrega_producoes('dados_ecos/patentes_instituicoes.csv', 1);
31 -- Projetos
32 CALL carrega_producoes('dados_ecos/projetos_instituicoes.csv', 2);
33 -- Publicacoes
34 CALL carrega_producoes('dados_ecos/publicacoes_instituicoes.csv', 3);
35 -- Alto crescimento
36 CALL carrega_alto_cresc('dados_ecos/alto_cresc.csv');
37 -- SEBRAE
38 CALL carrega_gov('dados_ecos/ent_gov.csv');
39 END;
```

Fonte: O autor (2025)