

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

LUANA CRISTINA SILVA

PAINEL DE INDICADORES BIBLIOMÉTRICOS E DE SUSTENTABILIDADE PARA  
MONITORAMENTO CIENTÍFICO E TECNOLÓGICO SOBRE PLANTAS ENDÊMICAS  
DA AMAZÔNIA BRASILEIRA

São Carlos, SP

2026

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

LUANA CRISTINA SILVA

PAINEL DE INDICADORES BIBLIOMÉTRICOS E DE SUSTENTABILIDADE PARA  
MONITORAMENTO CIENTÍFICO E TECNOLÓGICO SOBRE PLANTAS ENDÊMICAS  
DA AMAZÔNIA BRASILEIRA

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de São Carlos, como requisito para a obtenção do título de Mestre em Ciência da Informação.

**Área de concentração:** Conhecimento, Tecnologia e Inovação.

**Linha de pesquisa:** Conhecimento e Informação para Inovação

**Orientador:** Prof. Dr. Leandro Innocentini Lopes de Faria

**Co-orientador:** Prof. Dr. Rogério Aparecido de Sá Ramalho

São Carlos, SP

2026

Cristina-Silva, Luana

Painel de indicadores bibliométricos e de sustentabilidade para monitoramento científico e tecnológico sobre plantas endêmicas da Amazônia brasileira / Luana Cristina-Silva -- 2026.  
150f.

Dissertação (Mestrado) - Universidade Federal de São Carlos, campus São Carlos, São Carlos  
Orientador (a): Leandro Innocentini Lopes de Faria  
Banca Examinadora: José Eduardo dos Reis, Roniberto Morato do Amaral  
Bibliografia

1. Indicadores bibliométricos; Painel de visualização; Plantas Endêmicas da Amazônia Brasileira. I. Cristina-Silva, Luana. II. Título.

Ficha catalográfica desenvolvida pela Secretaria Geral de Informática  
(SIn)

DADOS FORNECIDOS PELO AUTOR

Bibliotecário responsável: Arildo Martins - CRB/8 7180

## **FOLHA DE APROVAÇÃO**

Defesa de Dissertação de Mestrado da candidata Luana Cristina-Silva, realizada em 23/02/2026.

### **Comissão Julgadora:**

Prof. Dr. Leandro Innocentini Lopes de Faria (UFSCar)

Prof. Dr. Roniberto Morato do Amaral (UFSCar)

Prof. Dr. José Eduardo dos Reis (UNICEP)

## **AGRADECIMENTOS**

Agradeço às considerações e direcionamentos do Prof. Dr. Ramon Adrian Salinas Franco, e Prof. Dr. Roniberto Morato do Amaral, durante a banca de qualificação, essenciais para evolução e conclusão da dissertação.

Ao Prof. Dr. José Eduardo pela visão e considerações ao participar na banca de defesa, junto ao Prof. Dr. Roniberto.

Agradeço ao Prof. Dr. Leandro Faria, pela oportunidade de trilhar a carreira acadêmica sob sua preciosa orientação, e guiou em cada etapa com sabedoria, paciência, atenção e profissionalismo.

Igualmente, agradeço à coorientação do Prof. Dr. Rogério Ramalho, que proporcionou perspectiva de mercado e de conexões.

Aos colegas do NIT: Edenilza, Vanessa, Renan, Vinícius, Laura, Ingrid, Dênis, Nani, Allan, Emma, Fernando, Itunnu, Raquel e Shirley, que me acolheram e foram essenciais nesta jornada.

Aos colegas que ingressaram no mestrado em minha turma e tornaram-se amigos, principalmente o trio Vitória, Ana Carolina e César.

Aos colegas do i2i, Prof. Dr. Rogério, Dra. Aline e Gabriel.

Especialmente aos familiares que apoiaram meu amor pelos estudos, principalmente minha companheira Karla Barbosa que acreditou e me deu todo suporte necessário.

## **DEDICATÓRIA**

**À minha amada avó Teresinha, que me criou e deu a base para não somente viver, mas também sonhar. Ao meu pai Valdomiro, que apesar de não ter esta mesma oportunidade, me incentivou desde a graduação. E à minha parceira Karla, por todo apoio nesta trajetória e na vida.**

## RESUMO

A presente pesquisa visa monitorar a produção científica e tecnológica sobre plantas endêmicas da Amazônia devido à sua importância como bioinsumos naturais e aos riscos crescentes de extrativismo e degradação. Pesquisas provenientes da temática “plantas endêmicas do domínio fitogeográfico da Amazônia brasileira” abrangem diversas áreas do conhecimento, com diferentes objetivos e vieses. Neste cenário observa-se, em contrapartida, uma lacuna de indicadores bibliométricos para mapear e monitorar a produção científica e tecnológica desta temática para fins analíticos, diagnósticos ou preditivos. A hipótese é que o desenvolvimento de um sistema de monitoramento automatizado, baseado em um *pipeline* ETL capaz de integrar dados heterogêneos (*Scopus*, *Espacenet* e fontes de sustentabilidade), e composto por indicadores bibliométricos em intersecção com indicadores classificatórios de sustentabilidade, viabilizará o monitoramento periódico da dinâmica dos desenvolvimentos científicos e tecnológicos sobre plantas endêmicas da Amazônia. Deste modo, a pesquisa possui caráter quantitativo e utilizou métodos e técnicas provenientes da bibliometria. A busca por registros de artigos científicos ocorreu pela base de dados multidisciplinar internacional *Scopus* e por registros de patentes pela *Espacenet*. O conjunto de dados analisados contemplou 800 artigos científicos e 153 patentes, recuperados a partir de cinco expressões de busca principais (referentes aos 5 diferentes grupos taxonômicos, totalizando 1975 espécies endêmicas da Amazônia brasileira). Adicionalmente, um *pipeline* de *scripts* para o processo de modelagem dos dados de modo automatizado foi criado, com a finalidade de otimizar futuras etapas de atualização periódica dos indicadores. O Painel de Visualização com recursos dinâmicos foi criado utilizando o *Looker Studio* do *Google*. Os resultados demonstram que o painel proposto soluciona a lacuna de indicadores ao oferecer um sistema que integra múltiplas métricas de forma sistemática e periódica. Sua arquitetura, simplificada e intuitiva, facilita a redução do esforço cognitivo do usuário e estabelece um novo modelo analítico para o monitoramento de desenvolvimentos em Ciência, Tecnologia e Inovação sob a perspectiva da sustentabilidade na Amazônia. Em conclusão, a criação do Painel e do sistema que lhe dá suporte viabiliza o monitoramento sistemático e sustentável da produção científica e tecnológica sobre recursos naturais exclusivos da Amazônia, consolidando uma contribuição metodológica e prática para a Ciência da Informação e para a sociedade

**Palavras-chave:** Amazônia; Bibliometria; Informação Ambiental; Painel; Sustentabilidade.

## ABSTRACT

This research stems from the importance of natural bio-inputs from the Brazilian Amazon biome and the risks associated with extractivism and other degradation factors affecting these resources. Research concerning "endemic plants of the Brazilian Amazon phytogeographic domain" spans multiple fields of knowledge, featuring diverse objectives and perspectives. Within this context, however, a gap is observed regarding bibliometric indicators capable of mapping and monitoring the scientific and technological production on this subject for analytical, diagnostic, or predictive purposes. Accordingly, we hypothesize that a periodic monitoring system, comprising bibliometric and sustainability indicators, can facilitate the visualization of the dynamics of scientific and technological developments regarding endemic Amazonian plants. These indicators, intersecting with sustainability-aligned indicators, highlight both an opportunity and a research demand from the perspective of Information Science. The research is quantitative in nature and employed methods and techniques derived from bibliometrics. The search for scientific articles was conducted using the international multidisciplinary database Scopus, and the search for patents was conducted via Espacenet. Based on five main search queries (referring to 5 different taxonomic groups, totaling 1975 endemic species from the Brazilian Amazon), the analyzed dataset comprised 800 scientific articles and 153 patents. Additionally, a script pipeline for the automated data modeling process was created to optimize future stages of periodic indicator updates. A visualization dashboard with dynamic features was created using Google's Looker Studio. Partial results indicate that the proposed dashboard not only helps address the gap in indicator concentration but also offers several implications, given its simplified and intuitive design aimed at reducing the user's cognitive effort. Finally, this research corroborates the importance of monitoring contemporary scientific and technological activities related to the Amazon's exclusive natural resources, particularly from a sustainability perspective. Consequently, it highlights the relevance of this study and its contribution to Information Science and civil society as a whole.

**Keywords:** Amazon; Bibliometrics; Environmental Information; Dashboard; Sustainability.

## LISTA DE ILUSTRAÇÕES

<b>Figura 1</b> – Esquema de relação entre datas de pedido de patente por ano de prioridade .....	44
<b>Figura 2</b> – Exemplo de níveis de indentação dos subgrupos da IPC .....	46
<b>Figura 3</b> – Esquema do método científico hipotético-dedutivo por Koche .....	54
<b>Figura 4</b> – Quadro teórico para indicadores compostos .....	67
<b>Figura 5</b> – Definições prévias para seleção de dados .....	68
<b>Figura 6</b> – Arquitetura geral do do <i>pipeline</i> ETL e fluxo operacional .....	81
<b>Figura 7</b> – Fluxograma detalhado do <i>pipeline</i> de Extração, Transformação e Carga (ETL) .....	84
<b>Figura 8</b> – Aquisição e consolidação das fontes de dados brutos .....	85
<b>Figura 9</b> – Etapa 1 do <i>pipeline</i> (processamento e unificação base) .....	86
<b>Figura 10</b> – Etapa 2 do <i>pipeline</i> (enriquecimento de sustentabilidade) .....	88
<b>Figura 11</b> – Lógica de enriquecimento de sustentabilidade para artigos .....	89
<b>Figura 12</b> – Lógica de enriquecimento de sustentabilidade para patentes .....	90
<b>Figura 13</b> – Lógica de enriquecimento de ameaça, conservação e risco para artigos e patentes .....	91
<b>Figura 14</b> – Etapa 3 do <i>pipeline</i> (achatamento para BI) .....	93
<b>Figura 15</b> – Etapa 4 do <i>pipeline</i> (carga e apresentação) .....	95
<b>Figura 16</b> – Esquema de estrela aplicado aos dados de artigos .....	96
<b>Figura 17</b> – Esquema de estrela aplicado aos dados de patente .....	97
<b>Figura 18</b> – Visão geral dos indicadores compostos de sustentabilidade .....	100
<b>Figura 19</b> – Distribuição geográfica das publicações por instituições .....	102
<b>Figura 20</b> – Distribuição geográfica das publicações por instituições brasileiras .....	103
<b>Figura 21</b> – Número de publicações e de autores por instituição (Top 10) .....	104
<b>Figura 22</b> – Série temporal das publicações científicas sobre plantas endêmicas da Amazônia .....	105
<b>Figura 23</b> – Constância de publicação sobre plantas endêmicas da Amazônia por instituições ao longo dos anos .....	106
<b>Figura 24</b> – Número de publicações por autor .....	107
<b>Figura 25</b> – Gráficos do painel com filtro aplicado referente ao autor com maior número de publicações .....	109

<b>Figura 26</b> – Lei de Lotka aplicado à distribuição de contribuições dos autores .....	110
<b>Figura 27</b> – Protótipo do painel da produção científica sobre <i>Endopleura Uchi</i> .....	111
<b>Figura 28</b> – Frutos provenientes da árvore <i>Endopleura Uchi</i> .....	112
<b>Figura 29</b> – Relação de indicadores de publicações sobre a planta <i>Faramea occidentalis</i> .....	113
<b>Figura 30</b> – Plantas mencionadas em 5 ou mais pesquisas .....	114
<b>Figura 31</b> – Áreas tecnológicas das patentes com menções às plantas endêmicas da Amazônia brasileira .....	116
<b>Figura 32</b> – Visão geral da produção tecnológica .....	118
<b>Figura 33</b> – Top 5 das entidades aplicantes com maior número de patentes .....	119
<b>Figura 34</b> – Número de espécies endêmicas mencionadas em patentes .....	121
<b>Figura 35</b> – Espécie endêmica com maior número de menções nas patentes analisadas ( <i>Passiflora riparia</i> ) .....	122
<b>Figura 36</b> – Principais aplicantes e inventores com relação às patentes que mencionam a espécie <i>Passiflora riparia</i> .....	123
<b>Figura 37</b> – Série temporal das publicações segundo ODS .....	125
<b>Figura 38</b> – Relação de plantas endêmicas da Amazônia alinhadas aos ODS nas publicações científicas .....	126
<b>Figura 39</b> – Distribuição de publicações por Categoria de Risco CNCFlora .....	127
<b>Figura 40</b> – Distribuição geográfica das patentes verdes .....	129
<b>Figura 41</b> – Categorias de risco das plantas mencionadas nas patentes .....	133
<b>Figura 42</b> – Relação dos tipos de ameaças mais frequentes das plantas endêmicas mencionadas nas patentes .....	135
<b>Figura 43</b> – Relação dos tipos de conservação mais frequentes das plantas endêmicas mencionadas nas patentes .....	136
<b>Figura 44</b> – Mapa da publicação de “patente verde” mencionando a planta <i>Peltogyne altissima</i> categorizada como “em perigo” .....	137
<b>Figura 45</b> – Compêndio de indicadores da Lista Vermelha CNFlora da planta <i>Peltogyne altissima</i> classificada como “patente verde” .....	138
<b>Figura 46</b> – Página inicial do protótipo inicial do painel .....	142
<b>Figura 47</b> – Exemplo de página padrão com os indicadores bibliométricos de artigo proposto para o painel .....	145

## LISTA DE QUADROS

<b>Quadro 1</b> – Relação entre CD e CI .....	32
<b>Quadro 2</b> – Comparação entre bibliometria descritiva e avaliativa .....	37
<b>Quadro 3</b> – Indicadores derivados do indicador geral de produção (científica) .....	39
<b>Quadro 4</b> – Abordagens metodológicas da PT .....	41
<b>Quadro 5</b> – Tipos de abordagens para definir a localização da patente .....	42
<b>Quadro 6</b> – Fases e etapas do procedimento metodológico alinhado aos objetivos específicos .....	57
<b>Quadro 7</b> – Critérios de busca utilizados no site ReFlora .....	59
<b>Quadro 8</b> – Delimitação dos indicadores macro-bibliométricos descritivos de produção científica .....	69
<b>Quadro 9</b> – Delimitação dos indicadores macro-bibliométricos descritivos de produção tecnológica .....	70
<b>Quadro 10</b> – Delimitação dos indicadores indiretos sobre as plantas e sustentabilidades ...	72
<b>Quadro 11</b> – Visão geral dos indicadores simples e compostos finais .....	73
<b>Quadro 12</b> – Exemplo de estratégia de busca para mapear ODS (Grupo 3 - ODS 11) .....	75
<b>Quadro 13</b> – Matriz de Análise Integrada: <i>Google Looker Studio &amp; Workspace</i> .....	78
<b>Quadro 14</b> – Estrutura de scripts composição do Orquestrador 1 .....	86
<b>Quadro 15</b> – Estrutura de scripts de composição do Orquestrador 2 .....	92
<b>Quadro 16</b> – Estrutura de scripts de composição do Orquestrador 3 .....	93
<b>Quadro 17</b> – Descrição IPC geral e sustentável das patentes .....	132
<b>Quadro 18</b> – Relação de espécies por categoria de risco e número de ocorrências em patentes .....	134
<b>Quadro 19</b> – proposta de matriz para relacionamento de dimensões do CNCFlora .....	136
<b>Quadro 20</b> – Estrutura de navegação sugerida para o painel .....	143

## LISTA DE TABELAS

<b>Tabela 1</b> – Busca exploratória e resultados obtidos em diversas bases de dados .....	63
<b>Tabela 2</b> – Relação quantitativa de espécies pesquisadas e registros bibliométricos recuperados .....	99

## LISTA DE ABREVIATURA E SIGLAS

**API** – *Application Programming Interface* (Interface de Programação de Aplicações)

**BI** – *Business Intelligence* (Inteligência de Negócios)

**CI** – Ciência da Informação

**CIP** – Classificação Internacional de Patentes

**CNCFlora** – Centro Nacional de Conservação da Flora

**CT&I** – Ciência, Tecnologia e Inovação

**EMBRAPA** – Empresa Brasileira de Pesquisa Agropecuária

**ENV-TECH** – *Environmental Technology*

**ETL** – *Extraction, Transformation, Load* (Extração, Transformação e Carga)

**GI** – *Green Inventory* (Inventário Verde)

**GPSC-CDB** – Estratégia Global para Conservação de Plantas

**IBF** – Instituto Brasileiro de Florestas

**IBGE** – Instituto Brasileiro de Geografia e Estatística

**IC** – Inteligência Competitiva

**INPE** – Instituto Nacional de Pesquisas Espaciais

**INPI** – Instituto Nacional de Propriedade Industrial

**IPC** – International Patent Classification

**IT** – Inteligência Tecnológica

**IUCN** – União Internacional para a Conservação da Natureza

**LPI** – Lei da Propriedade Industrial

**LS** – *Looker Studio* (ou *Google Looker Studio*)

**MAIC** – Modelo Analítico de Indicadores Compostos

**MEC** – Ministério da Educação

**MI** – Monitoramento da Informação

**MMA** – Ministério do Meio Ambiente e Mudança do Clima

**MVP** – *Minimum Viable Product* (Produto Mínimo Viável)

**OCDE** – Organização para a Cooperação e Desenvolvimento Econômico

**ODS** – Objetivos de Desenvolvimento Sustentável

**OECD** – Organization for Economic Co-operation and Development

**OMPI** – Organização Mundial da Propriedade Intelectual

**ONU** – Organização das Nações Unidas

**PIB** – Produto Interno Bruto

**PT** – Prospecção Tecnológica

**SUDAM** – Superintendência de Desenvolvimento da Amazônia

**THE** – Times Higher Education

**WIPO** – *World Intellectual Property Organization*

# SUMÁRIO

<b>1 INTRODUÇÃO</b>	<b>12</b>
<b>1.1 PROBLEMÁTICAS: DESAFIOS E LACUNAS</b>	<b>13</b>
1.1.1 LACUNA DE INDICADORES BIBLIOMÉTRICOS	13
1.1.2 DESAFIO DE MODELAGEM DOS DADOS: HETEROGENEIDADE DE DADOS DE MÚLTIPLAS FONTES	14
1.1.3 DESAFIO DE AUTOMAÇÃO DO FLUXO DE DADOS PARA ANÁLISE VISANDO MONITORAMENTO PERIÓDICO	15
<b>1.2 QUESTÃO DE PESQUISA, OBJETIVOS E MÉTODOS</b>	<b>15</b>
1.2.1 OBJETIVO GERAL	16
1.2.2. OBJETIVOS ESPECÍFICOS	16
1.2.3 MÉTODOS	16
<b>1.3 ORIGINALIDADE E RELEVÂNCIA</b>	<b>17</b>
<b>2 REFERENCIAL TEÓRICO</b>	<b>20</b>
<b>2.1 A AMAZÔNIA</b>	<b>20</b>
2.1.1 CAPITAL NATURAL E COMERCIAL NA AMAZÔNIA	22
2.1.2 PARADIGMA DO DESENVOLVIMENTO SUSTENTÁVEL E A AMAZÔNIA	24
<b>2.2 CIÊNCIA, TECNOLOGIA E INOVAÇÃO</b>	<b>26</b>
2.2.1 DIMENSÃO CIENTÍFICA	26
2.2.2 DIMENSÃO TECNOLÓGICA	27
<b>2.3 CIÊNCIA DE DADOS</b>	<b>30</b>
2.3.1 RELAÇÃO ENTRE CIÊNCIA DE DADOS E CIÊNCIA DA INFORMAÇÃO	31
<b>2.4 CIÊNCIA DA INFORMAÇÃO E ESTUDOS MÉTRICOS</b>	<b>33</b>
<b>2.4.1 MONITORAMENTO DA CIÊNCIA ATRAVÉS DA BIBLIOMETRIA</b>	<b>35</b>
2.4.2 INDICADORES BIBLIOMÉTRICOS DE PRODUÇÃO CIENTÍFICA	36
2.4.3 MONITORAMENTO TECNOLÓGICO ATRAVÉS DA BIBLIOMETRIA DE PATENTES	39
2.4.4 INDICADORES DE SUSTENTABILIDADE	46
<b>2.5 ARQUITETURA E VISUALIZAÇÃO DA INFORMAÇÃO</b>	<b>49</b>
<b>3 MÉTODOS</b>	<b>50</b>
<b>3.1 DELINEAMENTO, TIPO E ABORDAGEM DA PESQUISA</b>	<b>50</b>

3.1.1 PLANO DE GESTÃO DE DADOS E FERRAMENTAIS.....	52
3.1.2 FASES E ETAPAS DA PESQUISA.....	53
<b>3.2 FASE 1: DELINEAMENTO, PROTOCOLO DE BUSCA E DEFINIÇÃO DE INDICADORES.....</b>	<b>55</b>
3.2.1 DEFINIÇÕES INICIAIS PARA EXPRESSÃO DE BUSCA .....	55
3.2.2 DEFINIÇÃO DOS TIPOS DE FONTES.....	57
3.2.3 DEFINIÇÃO DA BASE DE DADOS .....	58
3.2.4 DEFINIÇÃO DE INDICADORES.....	62
3.2.5 ESCOLHA DA FERRAMENTA DE VISUALIZAÇÃO: <i>LOOKER STUDIO</i> .....	72
<b>3.3 FASE 2: EXECUÇÃO DO PROTOCOLO E PIPELINE ETL DE MODELAGEM DE DADOS .....</b>	<b>76</b>
3.3.1 EXTRAÇÃO E FONTES DE DADOS BRUTOS.....	80
3.3.2 MODELAGEM E TRANSFORMAÇÃO DOS DADOS .....	81
3.3.3 MODELAGEM DIMENSIONAL.....	88
3.3.4 ESTRUTURAÇÃO E CARREGAMENTO .....	90
<b><u>4 RESULTADOS E DISCUSSÃO .....</u></b>	<b><u>94</u></b>
<b>4.1 INDICADORES BIBLIOMÉTRICOS DE PRODUÇÃO CIENTÍFICA .....</b>	<b>96</b>
4.1.1 NÚMERO DAS PUBLICAÇÕES POR AFILIAÇÃO E ANO .....	97
4.1.2 NÚMERO DE PUBLICAÇÕES POR AUTOR .....	102
4.1.3 NÚMERO DE PUBLICAÇÃO POR PLANTA .....	106
<b>4.2 INDICADORES BIBLIOMÉTRICOS DE PRODUÇÃO TECNOLÓGICA E INOVATIVA .....</b>	<b>110</b>
4.2.1 ÁREAS TECNOLÓGICAS DAS PATENTES .....	111
4.2.2 PATENTE POR AUTORIA.....	113
4.2.3 PATENTES POR ESPÉCIES.....	116
<b>4.3 INDICADORES DE SUSTENTABILIDADE.....</b>	<b>120</b>
4.3.1 INDICADORES DE SUSTENTABILIDADE BASEADOS NA PRODUÇÃO CIENTÍFICA.....	120
4.3.2 INDICADORES DE SUSTENTABILIDADE BASEADOS EM PATENTES.....	125
4.3.3 LISTA VERMELHA CNCFLORA NAS PATENTES.....	127
4.3.4 INTERSECÇÃO ENTRE PATENTES VERDES E LISTA VERMELHA CNCFLORA.....	132
<b>4.3 VALIDAÇÃO DA HIPÓTESE E CUMPRIMENTO DOS OBJETIVOS.....</b>	<b>134</b>
4.3.1 INTERSECÇÃO E FLUXO DE CONHECIMENTO: CIÊNCIA, TECNOLOGIA E SUSTENTABILIDADE.....	135
4.3.2 VALIDAÇÃO DA ARQUITETURA, USABILIDADE E SUSTENTABILIDADE DO PAINEL COMO MVP .....	135
4.3.3 COMPROVAÇÃO DA HIPÓTESE E ALCANCE DOS OBJETIVOS DA PESQUISA.....	139
<b><u>5 CONSIDERAÇÕES FINAIS .....</u></b>	<b><u>141</u></b>

**REFERÊNCIAS.....145**

**APÊNDICES .....158**

**ANEXOS .....180**

## 1 INTRODUÇÃO

O aumento da temperatura média da superfície terrestre é um dos impactos provenientes do aquecimento global, caracterizado pela sua larga extensão e provocado por fatores internos e externos. Nos fatores internos, consideram-se sistemas caóticos não lineares, como atividades solares inconstantes ou tectonismo; nos fatores externos, incluem-se as ações antropogênicas causadoras do efeito estufa, como as queimadas (Silva; Paula, 2015).

A ação humana é, portanto, uma das principais responsáveis pela mudança climática que acarreta o aquecimento global, cujo impacto mais notável se inicia a partir da Revolução Industrial, no final do século 18. Estima-se que, no Brasil, cerca de 58% dos gases emitidos que causam o efeito estufa provêm de queimadas e desmatamento (Brasil; INPE, 2024). Além disso, o extrativismo de recursos naturais com finalidade comercial também se destaca, integrando aspectos de economia e desenvolvimento ou crescimento econômico (Silva *et al.*, 2016).

Nesse contexto, o Brasil se destaca por sua vasta biodiversidade. A Amazônia, por sua vez, é o maior bioma brasileiro, abrigando 13.056 espécies de flora e diversos recursos naturais (Brasil; MMA, 2021). Além disso, o país detém a maior área de ecossistemas tropicais dentro de Unidades de Conservação (UCs), as quais são cruciais para a proteção da biodiversidade e para o sustento de populações humanas tradicionais (Lapola *et al.*, 2020, p. 427).

O bioma amazônico, conforme aponta o Instituto Brasileiro de Florestas (IBF), engloba uma vasta diversidade de ecossistemas, como florestas densas de terra firme, formações estacionais e de igapó, campos alagados, várzeas, savanas, refúgios montanhosos e formações pioneiras. Apesar de ser o bioma mais preservado do Brasil, estima-se que cerca de 16% de sua área já tenha sido devastada. Essa fragilidade intrínseca, somada à sua dependência de material orgânico próprio para subsistência, implica que intervenções e danos antrópicos podem causar impactos irreversíveis no ecossistema local (IBF, 2020).

A temática da Amazônia, embora presente na Ciência da Informação (CI), ainda é abordada de maneira incipiente. Conforme apontam (Assunção *et al.*, 2022), a CI se insere no campo do desenvolvimento sustentável ao utilizar a informação ambiental como ferramenta para a construção de uma sociedade sustentável. No entanto, o desenvolvimento sustentável na Amazônia, em particular, carece de maior aprofundamento na área da CI. Desse modo, verifica-se a pertinência do tema proposto para o campo da Ciência da Informação e a clara necessidade de discutir o desenvolvimento sustentável na Amazônia sob a perspectiva da CI.

Avançando na delimitação da pesquisa, o trabalho se norteia pela demanda por

indicadores capazes de relacionar Ciência, Tecnologia e Inovação (CT&I) com indicadores de sustentabilidade, com intuito de monitorar artigos e patentes sobre plantas endêmicas nativas do bioma amazônico. Algumas pesquisas e iniciativas se aproximam desta demanda, como na pesquisa de Carvalho (2023) que apontou lacunas e vieses em investigações científicas ecológicas, lacunas estas que se relacionam a determinadas áreas com biodiversidade negligenciadas da Amazônia, apontando para uma probabilidade de maior suscetibilidade a impactos de mudança climática.

Para este propósito, e buscando explorar o cenário em questão, a informação ambiental, quando alinhada à Ciência da Informação (CI), constitui-se como uma ferramenta essencial para viabilizar a compreensão da situação da Amazônia Legal, englobando seus recursos naturais e os diversos atores envolvidos. Dessa forma, torna-se relevante a mensuração de fenômenos, descrevendo e monitorando sistematicamente as condições do meio ambiente, sempre sob a perspectiva do desenvolvimento sustentável (Mueller, 2021; Assunção *et al.*, 2022).

## 1.1 PROBLEMÁTICAS: DESAFIOS E LACUNAS

A presente seção descreve a problemática central que guiou a pesquisa, destacando os desafios e as lacunas identificadas na área.

### 1.1.1 Lacuna de indicadores bibliométricos

A presente pesquisa parte da premissa de que, embora seja possível encontrar algumas métricas e indicadores sobre a Amazônia ou suas plantas, até o presente momento percebe-se uma lacuna de indicadores que mapeiam especificamente a produção científica e tecnológica sobre plantas endêmicas do bioma Amazônico, sobretudo de uma perspectiva sistemática e que possibilite a inter-relação com indicadores de sustentabilidade.

A pesquisa de Souza (2018) apresenta um mapeamento científico da Amazônia nas publicações científicas com suas respectivas temáticas e atores, em uma perspectiva mais geral, em termos de espécies de fauna e flora. Pode-se destacar ainda o Projeto *Synergize*<sup>1</sup>, que possui uma rede de diversos pesquisadores e instituições pelo Brasil e pelo mundo, com o objetivo de

---

<sup>1</sup> Projeto Synergize. Disponível em: [https://taoca.net/about\\_taoca/](https://taoca.net/about_taoca/). Acesso em: 04 dez. 2024.

colaborar com coleta e curadoria de dados sobre alguns tipos de espécies da fauna terrestre contidas na Amazônia.

O Projeto Dendrogene da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), por sua vez, objetiva possibilitar a integração de conhecimento científico referente à botânica da Amazônia para tomada de decisão em seu manejo, além de melhorar o processo de identificação desta botânica (Embrapa Amazônia Oriental; Brasil; MAPA, 1999). A dissertação de Yanai (2012) também pode ser destacada, uma vez que realizou análise do impacto da inovação tecnológica por meio de patentes de determinadas plantas naturais da Amazônia.

Contudo, percebe-se que os mapeamentos existentes são ou muito gerais (Souza, 2018), ou focado em fauna (Synergise) ou botânica geral (Dendrogene), ou tratam de plantas determinadas (Yanai, 2012), mas não do universo sistemático de plantas endêmicas com viés de sustentabilidade. Tais iniciativas, embora relevantes, confirmam que os esforços existentes se concentram em abordagens amplas. A lacuna, portanto, reside na ausência de indicadores bibliométricos que abranjam a totalidade das espécies de plantas endêmicas da Amazônia, disponíveis em algum artefato de informação, tal como um painel de visualização, que possam ser atualizados regularmente e permitam o monitoramento dos desenvolvimentos científicos e tecnológicos relacionados a essas espécies.

### **1.1.2 Desafio de modelagem dos dados: heterogeneidade de dados de múltiplas fontes**

O desafio da modelagem dos dados é central para a construção do painel e do sistema que o compõem, residindo na necessidade de integrar, harmonizar e relacionar, de forma sistemática, três conjuntos de dados intrinsecamente heterogêneos em termos de formato e estrutura:

1. Registros Bibliométricos: Dados de artigos científicos extraídos da base de dados multidisciplinar *Scopus* (textos completos, resumos, metadados de autoria, afiliação etc.).
2. Registros Patentométricos: Dados de documentos de patentes extraídos da *Espacenet* (classificações IPC, aplicantes, inventores, datas de concessão).
3. Dados Classificatórios de Sustentabilidade: Informações externas e normativas (ex: as listas de ameaça da CNCFlora, o alinhamento das publicações aos 17 ODS – Objetivos de Desenvolvimento Sustentável – e as classificações do IPC *Green Inventory*).

Essa multiplicidade de fontes exige o desenvolvimento de uma arquitetura de dados (como o esquema estrela, que será detalhado na metodologia) capaz de unificar as chaves de relacionamento, permitindo a intersecção analítica dos indicadores e a visualização no painel.

### **1.1.3 Desafio de automação do fluxo de dados para análise visando monitoramento periódico**

O terceiro desafio reside na necessidade de garantir a escalabilidade, a periodicidade e a confiabilidade do sistema de indicadores. O monitoramento contínuo dos desenvolvimentos científicos e tecnológicos sobre plantas endêmicas da Amazônia é de crucial importância, pois permite a identificação proativa de tendências, lacunas e possíveis impactos de novas tecnologias sobre os recursos naturais. No entanto, a coleta de dados de múltiplas fontes heterogêneas (1.1.2) exige uma intervenção manual intensa a cada ciclo de atualização, tornando o processo lento, custoso e suscetível a erros.

Portanto, a solução a ser empregada é o desenvolvimento de um fluxo de dados automatizado (*pipeline* ETL – *Extraction, Transformation, and Loading*). A automação por ETL soluciona este desafio ao padronizar e otimizar a aquisição, o tratamento e a modelagem dos dados. Isso garante que o painel de indicadores possa ser atualizado de forma sistemática e periódica, transformando uma complexa operação de *Data Science* em um sistema replicável e de baixa manutenção, essencial para o acompanhamento contínuo da dinâmica em Ciência, Tecnologia e Inovação (CT&I) na Amazônia. A automação por ETL, juntamente ao Plano de Gestão de Dados (PGD) contribuem para a sustentabilidade do sistema, dados e seus indicadores, por não depender de intervenção manual intensa a cada ciclo de monitoramento e reduzindo erros e esforço cognitivo.

## **1.2 QUESTÃO DE PESQUISA, OBJETIVOS E MÉTODOS**

De que maneira a integração de indicadores bibliométricos e de sustentabilidade, por meio de um painel de visualização sistematizado, pode viabilizar o monitoramento periódico dos desenvolvimentos científicos e tecnológicos sobre plantas endêmicas da Amazônia, contribuindo para a gestão da informação ambiental?

### 1.2.1 Objetivo geral

O objetivo geral da pesquisa é propor, desenvolver e avaliar um painel de monitoramento automatizado, baseado em *pipeline* ETL, composto por indicadores bibliométricos em intersecção com indicadores classificatórios de sustentabilidade, visando a gestão da informação sobre desenvolvimentos científicos e tecnológicos acerca de plantas endêmicas da Amazônia brasileira.

### 1.2.2. Objetivos Específicos

Para guiar a construção deste sistema e a elaboração de indicadores, foram definidos os seguintes objetivos específicos:

- A. Definir a estratégia de busca, as bases de dados (*Scopus*, *Espacenet*) e o protocolo de coleta para construir a amostra de artigos e patentes sobre plantas endêmicas;
- B. Elaborar e integrar os indicadores bibliométricos (de artigos e patentes) e os indicadores classificatórios de sustentabilidade (ODS, CNCFlora – Lista Vermelha de Ameaças, *IPC Green*), definindo suas relações de intersecção;
- C. Desenvolver o *pipeline* ETL (Extração, Transformação e Carga) automatizado para o tratamento e a harmonização dos dados heterogêneos provenientes de múltiplas fontes;
- D. Configurar e implementar o protótipo do painel de visualização (*dashboard*) com recursos dinâmicos, para apresentar os indicadores compostos de forma sistemática e periódica;
- E. Analisar e discutir os resultados dos indicadores compostos para mapear a dinâmica dos desenvolvimentos científicos e tecnológicos sob a perspectiva da sustentabilidade na Amazônia;
- F. Propor um protocolo de atualização e manutenção do sistema de monitoramento para garantir a replicabilidade e a gestão contínua dos dados.

### 1.2.3 Métodos

O delineamento da pesquisa adota o método científico hipotético-dedutivo e abordagem quantitativa (Rayward, 1996; Lakatos; Marconi, 2003) em que o problema central identificado é a lacuna de indicadores de pesquisas, tecnologias e inovações relacionadas às plantas endêmicas do bioma amazônico.; e por premissas: a) o potencial comercial e demais benefícios provenientes de espécies endêmicas da Amazônia podem resultar em produções científicas e criação de novas tecnologias e inovação; b) tais resultados tradicionalmente são publicados através de artigos científicos e patentes, dentre outros formatos; c) é passível de se mensurar tais resultados através de indicadores de CT&I; d) é possível integrar sistematicamente tais indicadores, somado a indicadores de sustentabilidade, a fim de realizar mapeamentos e monitoramentos para análises e diagnósticos.

Considerando o problema e as premissas estabelecidas, levanta-se a hipótese é que o desenvolvimento de um painel de monitoramento automatizado, baseado em um *pipeline* ETL capaz de integrar dados heterogêneos (*Scopus*, *Espacenet* e fontes de sustentabilidade), e composto por indicadores bibliométricos em intersecção com indicadores classificatórios de sustentabilidade, viabilizará (ou favorecerá de forma eficaz) o monitoramento periódico da dinâmica dos desenvolvimentos científicos e tecnológicos sobre plantas endêmicas da Amazônia

### 1.3 ORIGINALIDADE E RELEVÂNCIA

A contribuição desta pesquisa se desdobra em três dimensões cruciais, as quais, em conjunto, elevam o patamar do conhecimento e da prática em monitoramento de Ciência, Tecnologia e Inovação (CT&I) com foco em sustentabilidade e recursos naturais exclusivos:

- a) originalidade teórica e inovação no marco conceitual:
  - i) integração sistemática inédita: a principal inovação teórica reside na inédita e rigorosa integração sistemática de indicadores de sustentabilidade globalmente reconhecidos (como os Objetivos de Desenvolvimento Sustentável - ODS, a Lista Vermelha de Espécies Ameaçadas da IUCN e o IPC Green Patent Inventory) com as metodologias consolidadas de bibliometria e patentometria.
  - ii) foco em recursos naturais exclusivos: este arcabouço é aplicado especificamente ao tema de recursos naturais exclusivos, caracterizando um avanço na forma como a produção de conhecimento e tecnologia é avaliada, vinculando-a diretamente aos impactos ambientais e sociais.

Ao unir métricas de produção científica e tecnológica com métricas de impacto socioambiental, o trabalho estabelece um novo paradigma para a avaliação da pesquisa e inovação responsável.

- b) relevância metodológica e inovação em ferramentas de análise:
  - i) desenvolvimento de um *pipeline* automatizado: a pesquisa inova no campo metodológico por meio do desenvolvimento de um *pipeline* automatizado de processamento de dados. Este sistema é fundamental para a coleta, tratamento e integração dos diferentes tipos de indicadores (bibliométricos, patentométricos e de sustentabilidade), garantindo eficiência e reprodutibilidade.
  - ii) Modelo Analítico de Indicadores Compostos (MAIC): foi desenvolvido um modelo analítico de indicadores compostos. Este modelo permite a sintetização de dados complexos em métricas de fácil interpretação para a tomada de decisão. O MAIC representa um modelo ágil e altamente replicável, servindo como uma ferramenta robusta para o monitoramento contínuo de tendências em CT&I.
  - iii) protótipo de um painel dinâmico como ferramenta de comunicação: a materialização e comunicação desta inovação metodológica se dá por meio de um protótipo de painel dinâmico (*dashboard* interativo). Esta ferramenta de visualização de dados permite que stakeholders acessem e interajam com as informações de forma intuitiva, transformando dados brutos em inteligência estratégica acessível.
- c) relevância prática, social e potencial de replicabilidade:
  - i) instrumento de diagnóstico e previsão estratégica: o resultado da pesquisa é o fornecimento de um instrumento diagnóstico robusto e preditivo. Este instrumento é projetado para apoiar gestores, cientistas e formuladores de políticas públicas na avaliação do cenário atual da pesquisa e desenvolvimento (P&D) e na previsão de tendências futuras.
  - ii) foco na exploração responsável de bioinsumos na amazônia: a aplicação imediata e o impacto social da pesquisa centram-se na promoção da exploração responsável e sustentável de bioinsumos na Amazônia, contribuindo diretamente para a bioeconomia regional e a conservação da biodiversidade.

- iii) replicabilidade para múltiplos contextos: a arquitetura metodológica desenvolvida (o *pipeline* e o MAIC) é intrinsecamente modular e adaptável. Isso garante sua alta replicabilidade, permitindo que o modelo seja prontamente aplicado para monitorar P&D e inovação em outros biomas (como a Mata Atlântica ou o Cerrado) ou em contextos temáticos distintos, solidificando sua utilidade como um padrão de monitoramento de CT&I para o desenvolvimento sustentável.

## 2 REFERENCIAL TEÓRICO

A presente pesquisa dedica seu referencial teórico à exploração do contexto do bioma amazônico e seus recursos naturais. O objetivo é compreender a dinâmica entre seus benefícios potenciais (o chamado capital natural) e seus desafios (como a degradação e o extrativismo, representados pelo capital comercial), no paradigma do desenvolvimento sustentável. Tal análise é conduzida sob o viés da Ciência, Tecnologia e Inovação (CT&I) e à luz da Ciência da Informação (CI) (Pelsa; Balina, 2022). Nessa perspectiva, a informação é passível de quantificação, e utiliza-se desta quantificação para realizar análises tais como de manifestação futura ou comportamento em rede (Buckland, 1991; Lakatos; Marconi, 2003; Araújo, 2009).

Este capítulo apresenta o referencial teórico resultado do levantamento bibliográfico exploratório realizado para fins de contextualização para a pesquisa. No primeiro subcapítulo aborda-se a contextualização sobre a Amazônia brasileira com seu capital natural e comercial, em relação com o paradigma do desenvolvimento sustentável e o viés da informação ambiental. No segundo subcapítulo apresenta-se a CT&I dada sua relevância histórica e social para o Brasil e o mundo. No terceiro subcapítulo discorre-se a contextualização histórica da Ciência da Informação e os aspectos quantitativos voltados à bibliometria. O quarto subcapítulo explana alguns dos principais indicadores de sustentabilidade e sua importância tanto para o mapeamento quanto para o direcionamento dos avanços que se alinham ao desenvolvimento sustentável. Por fim, o quinto subcapítulo discorre sobre a arquitetura da informação e visualização da informação, aspectos imprescindíveis para a construção de um sistema complexo que relaciona a intersecção de diversos indicadores de diversas origens em diversos níveis.

### 2.1 A AMAZÔNIA

A Amazônia é o maior bioma do Brasil ocupando um território com mais de 4 milhões de km<sup>2</sup> e, segundo o Ministério do Meio Ambiente e Mudança do Clima (MMA), sua flora conta com a ocorrência de 13.056 espécies. As estimativas apontam que a Amazônia é a maior reserva de madeira tropical do planeta e recursos naturais diversos dos quais incluem enormes estoques de borracha, castanhas, peixes e minérios (Brasil, 2021). Além disso, o bioma amazônico possui diversos ecossistemas de florestas densas de terra firme, estacionais e de igapó, bem como campos alagados, várzeas, savanas, refúgios montanhosos e formações pioneiras, como aponta o IBF. Embora seja considerado o bioma mais preservado do Brasil,

infere-se que cerca 16% de sua área já foi devastada, proporção esta equivalente a duas vezes a área do estado de São Paulo (Instituto Brasileiro de Florestas, 2020).

A Amazônia também se destaca na participação de processos básicos para a estabilidade química na atmosfera terrestre e, conforme sua importância para o equilíbrio se reafirma, cresce também seu valor econômico. Com isso, se configura novas formas de dominação e colonialismo na região. Na Amazônia Ocidental, por exemplo, estima-se haver mais de 550 indústrias (nacionais e transnacionais). Em 2012, os números apontam que atividades de natureza não poluente empregava nesta região mais de 550 mil empregos diretos e indiretos, com notável expansão física e econômica, sobretudo nos setores elétrico-eletrônico, automotora, informática e biotecnologia (cosméticos, biofármacos e alimentação). Em termos de exportação, nesse período, atingiu-se mais de 50 países (Freitas; Freitas, 2016).

Neste sentido, os principais problemas ambientais enfrentados se dão pelo desmatamento, queimadas, garimpagem, agropastoreio e a biopirataria. Tal conjunto de ações é compreendido como responsável por graves mudanças climáticas no país e no planeta, como o aquecimento global, já que a Amazônia é considerada como um grande resfriador atmosférico e maior abrigo da biodiversidade do mundo (Instituto Brasileiro de Florestas, 2020). O número de focos de incêndio, como uma das causas antropogênicas, são progressivamente evidentes e apontam índices elevados, principalmente no ano de 2004, em que se atingiu mais de 218 mil focos só em seu bioma (Jesus *et al.*, 2020). Outra problemática advinda do fator humano no contexto amazônico se dá pelo extrativismo para atividades econômicas, compreendido dentro de um processo que abrange aspectos políticos, sociais e econômicos (Silva *et al.*, 2016). Nesse contexto, a Amazônia é caracterizada pela economia de fronteira, assim como ocorre na América Latina em geral, em que se perpassa pelo processo de mercantilização de seus bens naturais para transformar em mercadorias. Portanto, preconiza-se o crescimento econômico baseado na incorporação de sua terra e seus recursos naturais em uma perspectiva de infinitude (Becker, 2005).

Dentre as delimitações abrangendo a Amazônia, pode-se encontrar variações terminológicas referentes à distribuição geográfica, climática, política, biológica e ecossistêmica. Com isso, faz-se necessário apontar algumas das definições imprescindíveis para elucidação da pesquisa. Partindo da perspectiva de geopolítica tem-se a Amazônia Legal, que “corresponde à área de atuação da Superintendência de Desenvolvimento da Amazônia – SUDAM delimitada em consonância ao Art. 2º da Lei Complementar n. 124, de 03.01.2007”, instituída com o objetivo de “definir a delimitação geográfica da região política de atuação da SUDAM como finalidade promover o desenvolvimento incluyente e sustentável de sua área de

atuação e a integração competitiva da base produtiva regional na economia nacional e internacional” segundo o IBGE. Neste contexto, a região da Amazônia legal compreende 772 municípios de 9 estados brasileiros (Instituto Brasileiro de Geografia, 2024)

Pelo viés biológico, compreende-se o bioma amazônico, ou bioma da Amazônia. O termo “bioma” segundo o dicionário *Michaelis On-line* (2025) trata-se da

Unidade biótica imediatamente superior ao ecossistema, formada por todos os vegetais, animais e comunidades existentes em determinado espaço e caracterizada por tipos de vegetação semelhantes, porém em diferentes estágios de evolução, como, por exemplo, floresta decídua temperada, tundra, pastagem ou deserto.

Portanto, o bioma da Amazônia considera plantas animais e comunidades que existem nas florestas e espaços da Amazônia brasileira e dos demais países. Ainda na abordagem biológica, pode-se definir o domínio fitogeográfico de um determinado bioma. O termo “fitogeográfico” é relativo à “fitogeografia” e ao “geobotânico”, segundo o dicionário (“Fitogeográfico”, [S.d.]). Segundo o Jardim Botânico do Rio de Janeiro<sup>2</sup>, entende-se o domínio fitogeográfico da Amazônia como aquele que está presente nas regiões norte e centro-oeste brasileiros, abrangendo variedade de fitofisionomias. Este território ocupa 49,3% do Brasil, e se estende pela Bolívia, Peru, Colômbia, Guianas e Venezuela.

Contudo, o interesse na Amazônia se dá por diferentes razões, e considerando a riqueza da sua biodiversidade, compreende-se que uma das principais se vincula ao seu potencial de exploração de recursos naturais que resultarão em produtos e/ou soluções para a sociedade. Por exploração entende-se a conversão de capital natural à capital econômico, como apontam Espada e Sobrinho (2015).

### **2.1.1 Capital natural e comercial na Amazônia**

Entende-se por capital natural, recursos advindos da natureza que produz valor (economia e bem-estar) para as pessoas, podendo assim ser ecossistemas, espécies, ar, oceanos, processos naturais, dentre outros, termo este proveniente de princípios da economia (Brasil, 2024). O mapeamento, classificação e registro de tais recursos, pode ser um grande desafio para

---

<sup>2</sup> Disponível em: <https://floradobrasil.jbrj.gov.br/consulta/#CondicaoTaxonCP>. Acesso em: 06 fev. 2025.

a ciência, e iniciativas como o Programa Refflora<sup>3</sup> desempenham um papel crucial no registro do capital natural do bioma amazônico e demais biomas nacionais. O programa surge em 2010, inicialmente vinculado à Meta 1 da Estratégia Global para Conservação de Plantas (GPSC-CDB), com o lançamento da primeira versão online da lista de espécies de flora do Brasil, projeto que ficou popularmente conhecido como “Lista do Brasil” e que foi encerrado em novembro de 2015 com a publicação de artigos e suas respectivas bases de dados. No início de 2016 lançaram o projeto “Flora do Brasil de 2020”, voltado para a mesma meta da GPSC-CBD, porém agora com enfoque na divulgação de descrições morfológicas, chaves de identificação e ilustração das espécies de plantas, algas e fungos para o Brasil. Mais de 900 taxonomistas, brasileiros e estrangeiros, contribuíram colaborativamente, resultando em uma plataforma online para inclusão de dados de seus grupos de especialidades, disponibilizando assim, no final de 2020, ampla e irrestritamente, monografias de todas famílias de briófitas, samambaias, licófitas e gimnospermas do Brasil.

No entanto, ao longo do processo identificaram que as monografias de algas e fungos não foram capazes de abranger a real diversidade destes grupos no Brasil, estimulando uma nova fase no projeto denominada “Flora e Funga do Brasil” com o intuito de buscar uma terminologia biológica mais inclusiva e com maior engajamento da comunidade, tanto para melhoria da lista de espécies quanto para criação de novas monografias. Até o presente momento o projeto já registra 52242 espécies, dentre elas nativas, naturalizadas e cultivadas. Portanto, dado o histórico do projeto, os membros envolvidos e sua fundamentação que nos leva até a lista extensa e fidedigna, uma base claramente relevante para localizar espécies da flora amazônica, até mesmo as endêmicas.

Compreende-se por plantas endêmicas aquelas que são nativas e exclusivas à uma determinada região geográfica (Brasil, 1994).

O capital comercial de plantas endêmicas da Amazônia é apontado por Oliveira *et al.* (2019), dentre 3 principais áreas tecnológicas com patentes depositadas. Em primeiro lugar alimentos e em segundo medicamentos no Brasil e no mundo. Em terceiro lugar biotecnologia no Brasil, e agricultura no mundo. A busca por soluções na saúde através do uso de insumos naturais, bem como na integração da flora endêmica na cadeia de produção global, mostra-se como uma prática crescente e pode representar impactos adversos nos ecossistemas. Para fins medicinais Fernandes *et al.* (2021, p. 4729) inferem que “A busca por compostos bioativos

---

<sup>3</sup> **Flora e Funga do Brasil.** Jardim Botânico do Rio de Janeiro. Disponível em: <http://floradobrasil.jbrj.gov.br/>. Acesso em: 6 Fev 2025.

contra doenças é imprescindível e a riqueza da Amazônia constitui uma grande fonte a ser explorada”. Deste modo requerem, para um uso bem-sucedido ecologicamente, uma clara compreensão dos sistemas de demanda e produção, bem como de abordagens inovativas (Volenzo; Odiyo, 2020). Dado sua importância, ações e políticas para proteger e direcionar a exploração de seus recursos de modo responsável e sustentável são imprescindíveis.

### **2.1.2 Paradigma do desenvolvimento sustentável e a Amazônia**

No paradigma do desenvolvimento sustentável, na esfera econômica, duas teorias são abordadas: a do crescimento econômico e a do desenvolvimento econômico. Na teoria do crescimento econômico pela definição de Paul M. Romer, segundo Pelsa e Balina (2022, p. 91) “[...] ocorre sempre que as pessoas utilizam recursos e os reorganizam de formas mais valiosas. Assim, o crescimento econômico significa um aumento do rendimento nacional real/produção nacional”. E por outro lado, a teoria do desenvolvimento econômico “[...] é uma melhoria da qualidade de vida e dos padrões de vida, por ex. melhoria da alfabetização, expectativa de vida e abordagem às questões ambientais” que abrange uma gama de métricas mais amplas, para além do Produto Interno Bruto (PIB) per capita, tais como relacionados aos níveis de padrões ambientais, disponibilidade e qualidade de habitação, dentre outros aspectos.

Assim, enquanto a teoria do crescimento apoia valores voltados para obtenção de riquezas, na teoria do desenvolvimento os valores se constituem na ideia de que deve-se combater as desigualdades não só de hoje, mas para as futuras gerações também. E é nesse sentido que os autores apontam como os limites ecológicos de ambas teorias se distinguem, onde a teoria do crescimento ignora frequentemente a realidade de que os recursos naturais e serviços ecossistêmicos são limitados, enquanto a teoria do desenvolvimento aborda estas questões no sentido de um “ponto de partida”, a fim de construir um sistema novo e funcional de avaliação, desenvolvimento e políticas de emprego, com uma série de métricas e indicadores modernos que viabilizam respostas de questões realmente relevantes à população. Contudo, ambas teorias se complementam e o principal objetivo é o de “[...] garantir o desenvolvimento sustentável da sociedade, que inclui aspectos econômicos, sociais e ambientais para uma vida melhor da sociedade” (Pelsa; Balina, 2022, p. 100).

Na perspectiva histórica, o desenvolvimento sustentável ganhou maior destaque nas discussões internacionais em 1972, na Conferência Mundial sobre o Homem e o Meio Ambiente, realizada pela Organização das Nações Unidas (ONU) em Estocolmo. Em 1984 a

ONU direciona outra conferência que dá origem à Comissão Mundial sobre Meio Ambiente e Desenvolvimento, com o lema “Uma agenda global para a mudança”, encerrado em 1987. Os anos se passaram e as discussões subseqüentes não foram suficientes para obtenção de resultados. Novas conferências ocorreram nos anos seguintes, mas foi em 2015 que foi lançada a Agenda 2030, composta dos Objetivos de Desenvolvimento Sustentável (ODS) (Teixeira, 2020)

A Amazônia, por sua vez, assume importante papel na estabilidade climática do planeta junto à complexidade dos problemas de trópicos úmidos, propondo novos desafios à Política de Estado de Ciência e Tecnologia, e através da emergência da materialização das sociedades do saber, humanização e integração da Amazônia constituem-se pressupostos para a construção do seu desenvolvimento sustentável (Freitas; Freitas, 2016). Consoante ao bioma, dentre as frentes de impacto antropogênico destacam-se duas: a) o extrativismo para fins comerciais (podendo abranger ou não a biopirataria); e b) desmatamento, queimadas e mineração. Cada qual assume desafios em termos de sustentabilidade. Ao que se refere a primeira opção, entende-se que o

[...] extrativismo, desenvolvimento e sustentabilidade são temas cuja discussão está longe de ser esgotada. Por isso, é preciso investimento em pesquisa sobre a diversidade dos ecossistemas e categorias sociais que formam a Amazônia e sobre a maneira como tem se dado a interação entre homem e natureza nesse território (Silva *et al.*, 2016, p. 57).

Duas possibilidades são evidenciadas sobre o extrativismo de bioinsumos amazônicos, ora com foco na possibilidade do desenvolvimento econômico, ora com foco na preservação dos recursos naturais e modo de vida das populações tradicionais e originárias. De todo modo, estratégias de desenvolvimento têm sido configuradas para a Amazônia no intuito de integrá-la economicamente em níveis nacional e internacional. Dentre as complexas questões, conflitos e contradições acerca do extrativismo, teóricos apontam que o paradigma neste cenário tem orientado o direcionamento de uma exploração que promova um desenvolvimento menos degradante e capaz de proporcionar o crescimento econômico a fim de gerar renda à população local (Silva *et al.*, 2016).

Contudo, alinhar os princípios e valores do desenvolvimento sustentável à exploração dos recursos naturais amazônicos é, além de necessário, também um desafio. Nesse sentido, existem conceitos voltados à inovação sustentável, também conhecidos como invenção verde, tecnologia verde, eco-inovação e inovações ambientais, e que se alinham ao desafio de usufruto com responsabilidade ecológica (Menezes; Santos; Bortoli, 2016; D’Amato; Mazzanti; Nicolli,

2021; Semenova; Semenov; Storchevoy, 2023; Favot *et al.*, 2023).

## 2.2 CIÊNCIA, TECNOLOGIA E INOVAÇÃO

Na sociedade contemporânea, no paradigma da informação e com o advento da globalização, o grande desafio enfrentado é o de organizar e estruturar o caos do grande volume de dados e informações dos mais diversos segmentos, cada vez mais rápidos. Se antes houvera a percepção de controle, inicialmente neste cenário atual “ninguém parece estar no controle [...]” ou sequer está claro o que é “ter controle”, conforme o teórico Bauman (1999, p. 65–66).

Castells (2011), entretanto, aponta neste cenário o novo espaço industrial e a nova economia de serviços, que direcionam suas operações no sentido de suas unidades que geram informação, reintegrando por meio de sistemas de comunicação. O novo estado na globalização, portanto, afirma suas fontes de poder justamente no controle e na orientação estratégica do conhecimento. Desse modo, as “novas” tecnologias da informação são instrumentos que permitem a incorporação da lógica organizacional que transforma o significado social do espaço.

Um dos esforços globais a fim de sistematizar, representar e interpretar a CT&I é através da definição de indicadores. A Organização para a Cooperação e o Desenvolvimento Econômico (OECD na sigla em inglês) assume este importante papel, e em ação coletiva elabora manuais com atualização sob demanda, e sistemática. Esses manuais orientam na coleta, interpretação e desenvolvimento de indicadores, compartilhados com países membros e não membros da organização, possibilitando comparações internacionais. (OECD, 2007)

Seguindo a abordagem métrica para fins avaliativos, à luz da CI, discorre-se nos subtópicos seguintes as dimensões científicas e tecnológicas-inovativas.

### 2.2.1 Dimensão científica

Fazer ciência é transcender o senso comum, no intuito de compreender e representar o contexto e fatores existentes no mundo. Segundo Koche (2011, p. 29), “O homem quer ir além dessa forma de ver a realidade imediatamente percebida e descobrir os princípios explicativos que servem de base para a compreensão da organização, classificação e ordenação da natureza em que está inserido”.

Contudo, para o desenvolvimento constante da ciência, denota-se que só ocorre “em sociedades com certa ordem, sujeitas a complexos peculiares de pressuposições tácitas e construções institucionais” como indica (Merton, 2013, p. 159). Portanto, faz-se necessário participação ativa de pessoas capazes e com interesse nos propósitos científicos, para tal evolução. Segundo Merton, são as condições culturais apropriadas que assegurem o apoio à ciência, e mudanças nas estruturas institucionais podem interferir positiva ou negativamente no desenvolvimento e avanço da ciência. Assim, uma maneira de observarmos a dimensão científica é por meio das instituições de ensino e pesquisa, que fazem ou fomentam o fazer científico. No Brasil, atualmente, existem diversas configurações e instituições voltadas para tal propósito, conforme aponta o Ministério da Educação (MEC)<sup>4</sup>. Estas instituições podem se dividir em diferentes níveis, tais como federal, estadual ou municipal, podem ser privadas ou públicas, com diferentes focos de atuação e interesses.

Para além dessas entidades que compõem o fazer científico, existem também os resultados da produção científica, geralmente gerados por essas instituições de ensino e pesquisa. Um dos mais notórios é a publicação de artigos científicos. O artigo científico é, portanto, um dos principais meios de divulgação de resultados alcançados por pesquisadores no âmbito da comunicação científica, nas diversas áreas do conhecimento. Neste cenário, um dos principais objetivos do artigo é de que aquilo que foi produzido seja comunicado e chegue aos demais pesquisadores, como sugerem Dias, Dias e Moita (2021).

A observação da ciência através da comunicação científica é considerada legítima uma vez que, para que seja considerado como científico, a publicação (via de regra) passa por revisão de pares. Assim, além de ser essencial para a certificação é também importante para geração de novos conhecimentos, como defende Mueller (2021). Assim, Mueller aponta que publicações científicas e os elementos de seu processo é o interesse central ao estudarmos o fazer científico, podendo contribuir para a democratização do acesso e uso da informação e para a formação de políticas nacionais e internacionais, voltadas para tecnologia e ciência, e para a sociedade conseqüentemente.

### **2.2.2 Dimensão tecnológica**

---

<sup>4</sup> MEC. Disponível em: <https://www.gov.br/mec/pt-br>.

A mudança técnica advinda da inovação tecnológica respalda o progresso nas principais economias que se baseia no conhecimento. Crescimento, produtividade e competitividade são cada vez mais influenciados pelo aprimoramento de novas tecnologias, criação de novos produtos, atualização de processos ou serviços personalizados. E, no centro das inovações estão procedimentos, como o de estudo-da-arte, ou então rotinas que conseguem solucionar problemas de forma bem-sucedida (Tijssen, 2005).

A importância da inovação tecnológica para sociedade se dá, tanto no século XVII quanto na contemporaneidade, por ser a “[...] condição necessária para a passagem de um mundo de conhecimento misterioso e secreto sobre a natureza para um novo modelo público e coletivo de produção científica.” (Delfanti; Pitrelli, 2015, p. 66).

Entretanto, “inovações não decorrem exclusivamente de pesquisa científica”, de acordo com Schons, Prado Filho e Galdino (2020, p. 29). Resultados “palpáveis” e legítimos para representar inovação tecnológica podem se dar através da propriedade intelectual, por exemplo.

Neste contexto, concebe-se que a competição mundial é direcionada pela propriedade intelectual com base no domínio do conhecimento de força e inovação do processo produtivo. Isto denomina o paradigma “técnico-científico de produção” onde tem-se no “[...] progresso da inovação tecnológica a base de competitividade no mundo atual da globalização econômica”. Neste aspecto, a proteção da propriedade intelectual se torna valoroso para o mercado. (Chaves; Nogueira, 2008, p. 127).

A propriedade intelectual abrange três categorias. A primeira: propriedade industrial – onde registra-se patentes, desenhos industriais, marcas e indicações geográficas – a segunda o direito autoral – contemplando direitos do autor, direitos conexos e programas de computador – e a terceira a proteção *sui generis* – conhecimentos tradicionais, patrimônio genético, cultivar e topografia de circuitos integrados. Portanto, criações intelectuais voltadas à invenção e inovação são tuteladas pela Lei de Propriedade Industrial (Bochi; Gabriel Júnior; Moura, 2020).

Em uma perspectiva mais pragmática, a propriedade intelectual pode ser compreendida como um sistema que protege patentes, marcas, direitos autorais e demais ativos, elementos estes da criação humana e inovações tecnológicas predominantes no âmbito empresarial e industrial, como inferem Silva e Santana (2022). Neste sentido, a patente como um dos tipos de ativos da propriedade intelectual, visa garantir a proteção de tecnologias, processos ou produtos que apresentem requisitos de novidade, atividade inventiva e aplicação industrial. No Brasil, a propriedade industrial é regulada pela Lei nº 9.279 de 14 de maio de

1996<sup>5</sup>, que define direitos e obrigações sobre a propriedade registrada em âmbito nacional. Desse modo, regula-se legalmente o registro de marcas e patentes, concedido pelo Instituto Nacional de Propriedade Industrial (INPI) no país.

A patente, de acordo com o INPI (2021), é um título oficial e temporário concedido pelo Estado por força de lei ao titular e sucessores, sejam pessoas físicas ou jurídicas. O titular da patente, após ter o registro concedido, passa a ter os direitos exclusivos sobre o bem que pode ser um produto, um processo de fabricação ou mesmo um aperfeiçoamento de produtos e processos já existentes.

Para que terceiros explorem esta tecnologia patenteada, é necessário solicitar a permissão do titular por meio de uma licença. Após o prazo da vigência da patente finalizada (20 anos no caso do Brasil), esta passa a ter domínio público e o direito de exclusividade não mais vigora, permitindo a utilização por qualquer outra pessoa ou empresa.

O documento da patente, portanto, é composto por um grupo de folhetos com a base de informação tecnológica, e segundo o INPI (2021, p. 10)

“Existem dois tipos de folhetos, um é chamado de folheto de publicação ou primeira publicação e o outro é o folheto de concessão ou segunda publicação. Os folhetos são formados por uma "folha de rosto" e pela matéria relativa ao invento (Relatório descritivo, Reivindicações, etc). A folha de rosto é a primeira folha do documento e contém os dados identificadores do pedido, tais como: nome do depositante e inventor, data de depósito, etc.”

A patente conta com caracteres fundamentais, para além da propriedade limitada temporalmente, como também o território de validade, obrigações definidas ao titular, e o interesse público na divulgação da informação contida no pedido da patente, que permite

“[...] à sociedade o livre acesso ao conhecimento da matéria objeto da patente. Dessa forma, os concorrentes do inventor podem desenvolver suas pesquisas a partir de um estágio mais avançado do conhecimento, promovendo, assim, o desenvolvimento tecnológico do país.” (INPI, 2021, p. 11)

Mas por que patentear? Segundo o INPI (2021, p. 9) pode-se considerar que patentear é “[...] uma forma de incentivar a contínua renovação tecnológica, estimulando o investimento das empresas para o desenvolvimento de novas tecnologias e a disponibilização de novos produtos para a sociedade”.

Estudos como o de Silva e Santana (2022), apontam uma relação positiva e estatisticamente significativa entre o registro de patentes - assim como de marcas - para com o

---

<sup>5</sup> Disponível em: [https://www.planalto.gov.br/ccivil\\_03/leis/19279.htm](https://www.planalto.gov.br/ccivil_03/leis/19279.htm). Acesso em: 23 out. 2023

valor de mercado das empresas analisadas. Portanto, as patentes influem no desempenho de mercado, pois relaciona uma expectativa positiva de investidores sobre a aplicação da firma nos aspectos de tecnologia, bem como de inovação e de propriedade intelectual.

Oliveira e Avellar (2020) apontam que empresas tendem a realizar mais de um tipo de inovação simultaneamente – seja ela de processo ou produto. Tais inovações apresentaram impactos nas indústrias de baixo, médio e alto grau. Em que alto grau considera-se a redução de custos de produção, por exemplo. Deste modo, o registro de patente é uma formalização do processo ou produto inovativo de uma empresa.

### 2.3 CIÊNCIA DE DADOS

O atual paradigma da Sociedade da Informação, demarcado pelo uso acentuado da Tecnologia da Informação e Comunicação (TIC), propicia um volume de dados exponencial, também conhecido como *Big Data*. Este grande volume de dados são insumos para gerar indicadores, que respaldam a tomada de decisões em diversas organizações. Este fenômeno se apresenta, portanto, como um desafio às organizações, que precisam encontrar formas para extrair valor e desenvolver melhores serviços aos seus clientes, através deste grande volume de dados (Garoufallou; Gaitanou, 2021)

*Big Data*, no entanto, trata-se de um dos aspectos inseridos no campo da Ciência de Dados, que por sua vez busca estudar quanto a capacidades de apoiar na descoberta de informações potencialmente úteis, partindo de grandes ou complexas bases de dados no sentido de tomada de decisão embasada pelos dados (Saldanha; Barcellos; Pedroso, 2021).

Definir *Big Data*, portanto, é mais do que afirmar ser um simples amontoado de dados, já que envolve certa complexidade. Garoufallou e Gaitanou (2021) apontam *Big Data* como dados gerados constante, automática e rapidamente e que combinam-se de quatro importantes características: volume, velocidade, variedade e veracidade. Este fenômeno se apresenta então, como um desafio às organizações, que precisam encontrar formas para extrair valor e desenvolver melhores serviços aos seus clientes, através deste grande volume de dados. Contudo, *Big Data* incute na ciência a necessidade de análise de grande volume de dados no tocante “[...] à produção intelectual no campo científico construída em uma ciência computacional, chamada de *e-science*, baseada em vertentes de armazenamento e tecnologias emergentes na web” (Reis; Carvalho, 2022, p. 19).

*Big Data* trata-se, deste modo, de um dos aspectos inseridos no campo da Ciência de Dados, que segundo Saldanha, Barcellos e Pedroso (2021, p. 52), aborda aspectos como

“[...] estratégias para extração, transformação e carga dos dados, modelagem, construção e avaliação de algoritmos descritivos e preditivos, visualização de grandes quantidades de dados e deploy dos modelos em ambientes de produção para a tomada de decisão, entre outros”

A Ciência de Dados, por sua vez, é “um campo de estudo que se destaca pela capacidade de auxiliar a descoberta de informação útil a partir de grandes ou complexas bases de dados, bem como a tomada de decisão orientada por dados” (Saldanha; Barcellos; Pedroso, 2021, p. 52). Ou seja, CD é uma área do conhecimento cujo enfoque está no “[...] desenvolvimento de inovações tecnológicas e com a produção de conhecimento científico de natureza quantitativa em um contexto de massiva produção de dados”. Esta área se apropria então de ferramentas estatísticas e computacionais, possibilitando análises de grande volume de dados por sínteses de uma maneira dinâmica, em um curto espaço de tempo e geralmente de natureza preditiva por sínteses de uma maneira dinâmica, em um curto espaço de tempo e geralmente de natureza preditiva (Fernandes, 2020, p. 240).

### **2.3.1 Relação entre Ciência de Dados e Ciência da Informação**

Outra área do conhecimento que surge por intermédio do advento tecnológico e com objetivo de resolver problemas relacionados à informação, é a Ciência da Informação. Com a evolução da Ciência da Informação como uma área do conhecimento busca-se, dentre outros aspectos, resolver questões referentes ao problema do grande volume de informação, porém agora em um contexto universal e com uma variedade de formatos digitais, por meio das tecnologias de informação, conforme Saracevic (1995).

A relação entre as áreas de conhecimento Ciência de Dados (CD) e Ciência da Informação (CI) já são apontadas na literatura científica. Isto é possível dado a interdisciplinaridade da CI com outras áreas do conhecimento visando solucionar questões da sociedade contemporânea (Saracevic, 1995). Indo ao encontro, a CD considera a representação da informação, tecnologia e dados para práticas científicas dando importância característica de interdisciplinaridade e ressignificação do dado para a ciência (Reis; Carvalho, 2022, p. 19). Para melhor compreensão da relação entre ambas as áreas do conhecimento, contempla-se o Quadro 1.

Quadro 1 – Relação entre CD e CI

Aspecto	Ciência de Dados ( <i>Data Science</i> )	Ciência da Informação ( <i>Information Science</i> )	Intersecção e Relação
<b>Definição e Foco</b>	É uma ciência nova, intensiva em dados, focada na extração de conhecimento útil a partir de grandes volumes de dados ( <i>Big Data</i> ), utilizando métodos estatísticos e computacionais. Foca na inovação tecnológica e na produção de conhecimento quantitativo.	Campo multidisciplinar / interdisciplinar que lida com o ciclo da informação: coleta, análise, classificação, armazenamento, recuperação, disseminação e proteção da informação e documentos.	A CD é vista por alguns autores como uma "irmã gêmea" ou um subconjunto interdisciplinar da CI, complementando-se na preservação e expansão do conhecimento.
<b>Base Epistemológica</b>	Tende a uma perspectiva positivista e quantitativa. Adota uma postura ativa e de intervenção ( <i>design thinking</i> ), buscando gerar conhecimento acionável para tomada de decisão imediata.	Possui uma visão histórica, pragmática e social dos documentos e dados. Tradicionalmente, atua como mediadora neutra entre o conhecimento registrado e o usuário, focando na organização e acesso.	A CI pode contribuir para a CD trazendo a concepção histórica e social dos dados, contrapondo-se à visão puramente objetiva / neutra que a CD muitas vezes assume sobre os dados.
<b>Paradigmas Científicos</b>	Representa o Quarto Paradigma da Ciência ( <i>e-Science</i> ): a ciência baseada na exploração intensiva de dados, unificando experimento, teoria e simulação.	Evolui para integrar a perspectiva da <i>e-Science</i> , adaptando seus métodos de organização para lidar com o "dilúvio de dados" e a necessidade de curadoria digital.	Ambas são afetadas pelo "dilúvio de dados" e convergem na <i>e-Science</i> , onde a tecnologia e os dados ressignificam as práticas científicas.
<b>Metodologia e Técnicas</b>	Utiliza <i>Machine Learning</i> (aprendizado de máquina), mineração de dados, inteligência artificial, estatística e análises preditivas. É fortemente orientada para " <i>Analytics</i> " e descoberta de padrões.	Utiliza métodos de organização do conhecimento, ontologias, vocabulários controlados, indexação e estudos de usuários / fluxos informacionais.	Há um uso crescente de algoritmos de <i>Machine Learning</i> em problemas da CI (como recuperação da informação). A CI fornece padrões de metadados e qualidade de dados para a CD.
<b>Perfil Profissional</b>	Cientista de Dados: Combina habilidades de estatística, matemática, computação / programação e conhecimento substantivo do domínio.	Bibliotecário de Dados / Gestor da Informação: Foca na curadoria, gestão, preservação, metadados, ética, acesso aberto e instrução/alfabetização em dados.	Surge a necessidade de "alfabetização de dados" para bibliotecários e a colaboração em equipes multidisciplinares, onde o bibliotecário apoia o cientista na gestão e organização dos ativos de dados.
<b>Evolução Histórica</b>	Evoluiu a partir da computação e estatística, impulsionada pelo aumento da capacidade de processamento e armazenamento (nuvens) e pela geração massiva de dados na <i>Web</i> .	Evoluiu da Documentação e Biblioteconomia tradicional. As " <i>iSchools</i> " buscam agora integrar currículos de dados para se manter relevantes e "ambidestras"	Historicamente, as comunidades se desenvolveram paralelamente, mas o movimento de Acesso Aberto e a <i>e-Science</i> forçaram uma convergência e novas parcerias

Fonte: A autora (2025).

Nota: Adaptado de Fernandes (2020); e Reis e Carvalho (2022).

Contudo, embora as áreas sejam distintas elas caminham para uma simbiose necessária. Por parte da CD, que contribui com o poder computacional e analítico para lidar com o volume massivo (*Big Data*) e a predição; e por parte da CI, que oferece a estrutura teórica para garantir que esses dados sejam organizados, recuperáveis, confiáveis e eticamente geridos ao longo do tempo (curadoria e preservação).

A literatura evidencia ainda que a Ciência da Informação deve agir rapidamente para incorporar a competência em dados (*data literacy*) em seus currículos, sob o risco de perder relevância. Entretanto, a CI tem a vantagem de humanizar a Ciência de Dados, trazendo contextos sociais e éticos que muitas vezes faltam na abordagem puramente algorítmica desta última. (Fernandes, 2020)

## 2.4 CIÊNCIA DA INFORMAÇÃO E ESTUDOS MÉTRICOS

A Ciência da Informação é uma área do conhecimento que surge após o problema da explosão informacional, utilizando da tecnologia da informação como solução para gerenciar o grande volume informacional, conforme aponta Saracevic (2009).

Com a evolução da CI como uma área do conhecimento, busca-se, dentre outros aspectos, resolver questões referentes ao problema do grande volume de informação, porém, agora se tratando de um contexto universal e com uma variedade de formatos digitais, e ainda por meio das tecnologias de informação, conforme Saracevic (2009). Para além de tratar e organizar dados e informações, a CI nos suscita quanto à conceitos, modelos e ferramentas para gestão de informação em uma perspectiva que transcende o técnico, e compreende as minúcias desde a gênese dos dados, transformando-o em Informação que será mediada ao usuário que poderá gerar novos conhecimentos, considerando diferentes usuários, contextos sociais, culturais, econômicos e demais vertentes.

Seus paradigmas epistemológicos originam-se em duas raízes, conforme aponta Capurro e Hjørland (2007), sendo a primeira raiz na Biblioteconomia, com enfoque nos aspectos de caráter social e cultural humano – como, por exemplo, a capacidade de questionarmos o que não sabemos ou aquilo que acreditamos que sabemos – e a segunda raiz, que se dá pela computação, que de modo geral focava em estudar os problemas de transmissão de mensagens. Na segunda raiz considera-se, portanto, o caráter tecnológico e os impactos da computação em processos relacionados à informação, sobretudo da informação científica em documentos impressos. A CI perpassa, segundo o autor, pelos paradigmas físico, cognitivo e

social e permitiram a evidenciação de três processos hermenêuticos que condicionam quaisquer sistemas de informação, sendo estes, 1) usuário; 2) acervo e 3) sistema intermediário. Assim, a hermenêutica do usuário condiz quanto a sua capacidade de interpretar as necessidades em relação a si mesmos, ao intermediário e ao sistema.

No tocante à interdisciplinaridade da CI com demais áreas, fica evidente que além de factível, é ainda imprescindível, para solucionar questões da sociedade contemporânea, como aponta Saracevic (1996). Em consonância com a interdisciplinaridade da área, Rayward (1996) complementa ao afirmar que podemos considerá-la em estudos que abordem a informação para além de si, visando um sistema no qual está organizada.

Seguindo por esta abordagem, considera-se o delineamento teórico quantitativo da CI, que utiliza de conceitos matemáticos e de processamento de informação a fim de fundamentar, conceitual e metodologicamente, no qual este sistema é baseado. Neste contexto, o conceito de *informação-como-coisa* apresentado por Buckland (1991) traz à CI a perspectiva da informação como passível de ser mensurada e quantificada. Neste sentido, os estudos métricos da informação “[...] privilegiaram a produção científica e sempre desenvolveram pesquisas buscando medir índices, principalmente a partir de citações, com objetivo de avaliação [...]” (Araújo, 2017, p. 12–13).

Bibliometria, por sua vez, é um campo da Ciência da Informação que atua com desenvolvimento e aplicação de medidas e indicadores quantitativos para ciência e tecnologia advindos de informações de registros bibliográficos, define van Leeuwen (2005). Desse modo, compreende-se que as informações bibliográficas contemplam conhecimentos codificados.

O termo “bibliometria” foi utilizado pela primeira vez por Pritchard em 1969, conforme aponta a literatura levantada por Alvarado (2008). Com precedência do idioma francês, o termo foi proposto por Otlet em 1934 com a ideia de criar uma nova disciplina, *bibliometrie* ou então, *bibliologia*. Esta abordagem se denominou como teoria da bibliometria, que consiste “[...] na aplicação de técnicas estatísticas para a contagem e estabelecimento de padrões de regularidade em itens informacionais [...]”, baseando-se em leis empíricas desde 1920, e passou a ser mais difundida a partir de 1960. Posteriormente esta teoria ganha variações, como *cientometria* e *informetria*, por exemplo, mas de forma a preservar a lógica em que “[...] informação pode ser quantificada e que, por meio dessa quantificação, seria possível prever suas manifestações futuras [...]”. (Araújo, 2009, p. 194).

Partindo desta teoria, os indicadores bibliométricos podem ser subdivididos em indicadores bibliométricos de produção científica - podendo abranger os artigos científicos - e em indicadores bibliométricos de produção tecnológica/inovativa - podendo abranger a

propriedade intelectual e industrial como patentes.

O monitoramento da informação (MI), assim como a inteligência competitiva (IC) podem ser usados estrategicamente por empresas e organizações em vistas de obter resultados mais precisos e superar desafios. Através do monitoramento pode-se estabelecer um quadro inicial de fontes de informações e conhecimento essenciais para uma organização, alimentando ou atuando como matéria prima para aplicar IC. Fundamentalmente, MI compreende métodos ou técnicas de observação sistemática e constante de informações. Este trabalho é como um radar capaz de detectar e interpretar sinais. Alguns de seus princípios são voltados para planejamento e gerenciamento estratégico como um programa de pesquisa e desenvolvimento; implementar a atividade como um sistema formal, planejado, contínuo e coordenado; estabelecer parcerias com especialistas em TI (tecnologia da informação) para sistematizar monitoramentos; ou mesmo gerenciar a informação como uma função básica da monitoração (Hoffmann, 2011).

Contudo, os domínios de monitoramento descritos nas subseções seguintes podem oferecer uma integração de perspectivas entre produção científica e tecnológica, para com a sustentabilidade.

#### 2.4.1 MONITORAMENTO DA CIÊNCIA ATRAVÉS DA BIBLIOMETRIA

A aplicação da bibliometria, em geral, dentro de um procedimento de medição do desempenho da pesquisa, requer respostas a questões importantes, aponta van Leeuwen (2005). Assim, a primeira questão se baseia em como a visibilidade do tema se relaciona na literatura científica internacional, ou seja, até que ponto a aplicação da bibliometria se concentra em parte substancial da produção científica do tema de pesquisa.

O movimento para monitorar a atividade científica, em nível nacional e supranacional, se torna cada vez mais forte no mundo, com início notável à partir de 1990, sobretudo, na Europa e Estados Unidos da América. Agências criadas com função de criar e acompanhar tais indicadores, como Fundação Nacional da Ciência (EUA) e Observatório da Ciência e Tecnologia (França), reforça-se o caráter comparativo dos indicadores bibliométricos por países, regiões geográficas e instituições. Estes estudos partem de um nível macro de agregação, portanto, baseiam-se em nomes de países, cidades, universidades ou códigos postais, para obter informações de localização (endereço) das publicações. Contudo, esses dados carecem de precisão, fazendo com que os estudos permaneçam comparativos e descritivos. Além disso, através desses estudos pode-se obter *insights* tanto em nível geográfico – destilado de

informações de endereço disponíveis nas publicações científicas – quanto em nível cognitivo – campos ou disciplinas de pesquisas com base no periódico publicado (van Leeuwen, 2005).

Nesse sentido, vieses são encontrados nesta abordagem. Áreas como artes e humanidades são frequentemente publicadas em menor volume internacionalmente, bem como vinculado ao idioma inglês. Não significa necessariamente que essas áreas específicas não são veiculadas internacionalmente, mas sim que o inglês não é especificamente a língua franca dos desenvolvedores. A questão do idioma pode se estender para outras áreas do conhecimento acadêmico, tais como ciências médicas. Deste modo, difere-se bibliometria descritiva e avaliativa conforme Quadro 2 (van Leeuwen, 2005).

**Quadro 2** – Comparação entre bibliometria descritiva e avaliativa

<b>Característica</b>	<b>Bibliometria descritiva</b>	<b>Bibliometria avaliativa</b>
Abordagem	Análise de cima para baixo	Análise de baixo para cima
Aplicação	Instalações nacionais para monitoramento de P&D em perspectiva mais ampla	Avaliação de pesquisa (nacional) em disciplinas e organizações de pesquisa (universidades e afins)
Resultados	Geralmente de domínio público	Geralmente mantidos em segredo
Informações sobre pares	Pouco utilizada (maior distância com relação à própria pesquisa)	Componente necessário para avaliação

Fonte: A autora (2025).

Nota: Adaptado de van Leeuwen (2005).

de Sousa, Almeida e Bezerra (2024) também abordam sobre a subdivisão entre macro-bibliometria e micro-bibliometria. Na perspectiva macro, os estudos têm por objetivo obter uma investigação mais ampla de diversas áreas do conhecimento, e fazer comparações entre âmbito nacional e internacional. O interesse é significativo por instituições de ensino, pesquisadores, editores de periódicos científicos, ou gestores nacionais de políticas científicas. Aqui, percebe-se uma inquietação pelas dinâmicas da produção acadêmica-científica em algumas áreas e suas implicações futuras. Em suma, este ramo da bibliometria, a macro-bibliometria, corresponde a análise estatística à bibliografia geral.

#### **2.4.2 Indicadores bibliométricos de produção científica**

A produção científica é assim denominada com base no registro do conhecimento em um conjunto de publicações resultantes de pesquisas em diversas áreas e domínios com registro

em diversos suportes. Uma vez que o ato de publicar – ou seja, tornar público – o resultado de pesquisa é um compromisso incumbido ao pesquisador, entende-se, por conseguinte, que a pesquisa científica e publicação de resultados são atividades inseparáveis. Deste modo se constitui a espiral de produção e comunicação da ciência, pela socialização do conhecimento por meio de seu registro documental (Prado; Castanha, 2020).

A aplicação da mensuração bibliométrica surge primeiramente com enfoque em livros, considerando aspectos tais como quantidade de edições e exemplares, ou a quantidade de palavras por livro. Outros formatos de publicações bibliográficas foram, aos poucos, abordados nesta perspectiva métrica, incluindo artigos de periódicos (Araújo, 2006).

É notável a importância da comunicação científica no progresso do conhecimento. A troca dos resultados de pesquisas é um aspecto crucial que representa o esforço científico. Neste sentido, o fenômeno da literatura em série – através das publicações em periódicos tais como internacionais – são imprescindíveis para a ciência e, dentre as opções existentes, esses elementos são claramente importantes no processo de troca de conhecimento. Por elementos entende-se nomes, autores, endereços institucionais, periódicos, referências (no caso de citações), conceituações (palavras-chave e conjuntos). Em 2004, Van Raan estimou que cerca de um milhão de publicações por ano são adicionadas a bases científicas, volume este considerado suficientemente alto para realização de análises quantitativas que promovam resultados estatísticos significativos (Van Raan, 2005).

Ademais, artigos científicos tendem a ser mais proeminentes em áreas das ciências exatas e biológicas enquanto a área de humanas e sociais tendem aos livros. Portanto, o veículo de comunicação da publicação também deve ser considerado ao se comparar diferentes áreas do conhecimento (Faria *et al.*, 2011).

Indicadores, por sua vez, são recursos de medidas aplicadas às diversas finalidades com objetivo de mensuração. São ferramentas que dimensionam propriedades específicas atribuindo valores em uma determinada situação expressando alusão à uma certa realidade. Eles descrevem sinteticamente os padrões identificados em um estado dinâmico subordinado à potenciais variações em seu contexto. Subordinados aos indicadores de produção científica, há um conjunto de indicadores divididos em três categorias básicas: indicadores de produção, indicadores de citação e indicadores de ligação (Prado; Castanha, 2020).

#### 2.4.2.1 Indicadores de produção

Os indicadores de produção são estabelecidos por meio da contagem do número de

publicações por tipo de documento, por instituição, por área de conhecimento, país, dentre outros. Através da produtividade busca-se refletir o impacto junto à comunidade científica. Entretanto, o número de publicações reflete as características da produção científica, mas não determina sua qualidade. Indicadores de distribuição de produtividade de autores (Lei de Lotka) são derivados dos indicadores de produção, e comportam estudos frente a distribuição de uso de vocabulário (Lei de Zipf) ou classificação de periódicos e distribuição de revistas por assuntos (Lei de Bradford), conforme Quadro 3 (Prado; Castanha, 2020).

**Quadro 3** – Indicadores derivados do indicador geral de produção (científica)

Indicadores derivados	Definição
Lei de Lotka	distribuições de produtividade de autores
Lei de Zipf	indicadores de distribuição do uso de vocabulário
Lei de Bradford	indicadores de classificação de periódicos e distribuições de revistas por assunto
Número de publicações por tipos de documentos	contagem de publicações por tipos de documentos (livros, artigos, publicações científicas, relatórios etc)
Número de publicações por tipo de autoria	contagem de publicações por autoria de pesquisador, grupos de pesquisa, instituições ou países

Fonte: A autora (2025).

Nota: Adaptado de Prado e Castanha (2020); e de Sousa, Almeida e Bezerra (2024).

Esses indicadores, portanto, proporcionam visibilidade aos mais produtivos ou temáticas que se destacam em uma área do conhecimento. (De Sousa; Almeida; Bezerra, 2024)

#### 2.4.2.2 Indicadores de citação e ligação

Para Prado e Castanha (2020) os indicadores de citação baseiam-se no número de citações recebidas por uma publicação, determinando assim o impacto, influência ou visibilidade dos artigos científicos ou autores citados. A citação é considerada o meio mais conhecido de creditar aos autores. Dentro deste grupo, os indicadores mais utilizados são o número de citações – número de vezes que uma determinada publicação foi citada – e o total de citações por trabalho. Existem ainda índices como o de imediatez, de impacto, de atividade, de afinidade, de atração, dentre outros. Por fim, os indicadores de ligação, que tem por base a coocorrência de autores – tais como pesquisadores, instituições ou países (coautoria) –, citações de palavras de demais formas de relação observadas ao que tange à produção científica (cocitação ou coocorrência de palavras, por exemplo).

Os indicadores de ligação – também conhecido como colaboração científica – não se restringem à coautoria, portanto. Colaboração científica é uma prática mais ampla, contempla outras formas de trabalhar em conjunto e que não necessariamente se formaliza ou explicita-se. Exemplos de colaboração científica que transcende a coautoria são: indicação de leituras e de perspectivas analíticas distintas; discussão e esclarecimento de dúvidas sobre o tema estudado; compartilhamento de espaços ou recursos laboratoriais; assistência editorial, dentre outros. Assim sendo, a forma de mensurar as atividades colaborativas também se distingue ao que tange coautoria ou colaboração no geral. No caso dos artigos, a coautoria pode ser mensurada através de assinatura conjunta (seja este pesquisador, instituição ou país), pelo número e/ou ordem de autores, ou o tipo de relação entre os autores. Trazendo para colaboração, mensura-se por meio de agradecimentos, financiamentos, indicações em referências bibliográficas e em materiais e métodos (Hilário; Freitas, 2020).

#### 2.4.2.3 Indicadores alternativos

Outro indicador, proposto por volta de 2010, é o altimétrico. Altimetria se refere à métricas alternativas, cujo objetivo é mensurar o impacto das publicações de maneira complementar aos demais indicadores bibliométricos tradicionais. Este concentra-se na seleção de informações e mensuração de impacto das publicações científicas na *Web 2.0*. (Puerta-Díaz; Martí-Lahera; Martínez-Ávila, 2020)

#### **2.4.3 Monitoramento tecnológico através da bibliometria de patentes**

Dados de patentes são cada vez mais utilizados em indicadores de CT&I. Os documentos de patentes contem diversos elementos (descrições e referências altamente detalhadas) que podem ser utilizados como indicadores bibliométricos. Esses dados são e serão cada vez mais úteis como fonte de informação para fornecer uma medida aproximada da inovação (Okubo, 1997).

Outro interesse em monitorar tais indicadores se dá a partir de sua ligação com aspectos de crescimento econômico, com avanço significativo em meados da década de 80, com a disponibilização de bancos de dados eletrônicos com informações bibliográficas de patentes em grandes conjuntos de dados permitindo análises estatísticas. (Schmoch, 2005)

Dentro da Prospecção Tecnológica (PT) – processo prático de produção de inteligência direcionado a compreender aspectos da CT&I – três abordagens metodológicas são utilizadas (Quadro 4), segundo Kupfer e Tigre (2004).

**Quadro 4** – Abordagens metodológicas da PT

Abordagem metodológica	Visão de futuro	Plano lógico	Definição
Baseada em inferência (mais convencional)	“o futuro tende a reproduzir, em alguma medida, os fenômenos já ocorridos, não implicando em rupturas ou descontinuidades nas trajetórias evolutivas dos objetos analisados”	Monitoramento ( <i>Assessment</i> )	“consiste no acompanhamento da evolução dos fatos e na identificação dos fatores portadores de mudanças, realizados de forma sistemática e contínua.”
Geração sistemática de trajetórias alternativas	“o futuro é projetado por meio da construção de cenários em um processo de contraposição de determinadas variáveis.”	Previsão ( <i>Forecasting</i> )	“consiste na realização de projeções com base em informações históricas e modelagem de tendências”
Construção do futuro por consenso (baseada em intuição ou cognição coletiva)	“o futuro é construído a partir de visões subjetivas de especialistas ou outros grupos de indivíduos dotados de capacidade de reflexão sobre os objetos do exercício de prospecção.”	Visão ( <i>Foresight</i> )	“consiste na antecipação de possibilidades futuras com base em interação não estruturada entre especialistas, cada um deles apoiados exclusivamente em seus conhecimentos e subjetividades.”

Fonte: A autora (2025).

Nota: Adaptado de Kupfer e Tigre (2004).

Neste sentido, a matéria-prima central para a análise visando Inteligência Tecnológica (IT) são as informações tecnológicas que inclui relatórios técnicos, normas técnicas e, principalmente, documentos de patente (NIT, 2004).

O estudo bibliométrico de patentes, também chamado de patentometria, considera as informações bibliográficas codificadas em um registro de patentes (Van Leeuwen, 2005). A patente, como um dos elementos da propriedade intelectual, subordina-se à propriedade industrial, tal qual aos desenhos industriais, marcas ou indicações geográficas (Bochi; Gabriel Júnior; Moura, 2020)

Tal qual o artigo científico, as patentes possuem também uma estrutura universalizada. A padronização dos registros pode incluir informações sobre inventores, depositantes, áreas de conhecimento (partindo da Classificação Internacional de Patentes – CIP), vinculação por famílias de patentes ou descrição do objeto requerido para sua aplicação industrial. Subdivididas em patentes de invenção (PI) e modelos de utilidade (UI), no Brasil o sistema é

regido pela Lei da Propriedade Industrial (LPI) – Lei no 9.279/96 (Bochi; Gabriel Júnior; Moura, 2020).

No escopo da CI, os documentos de patentes anexam-se às técnicas bibliométricas em uma aplicação metodológica. Os estudos bibliométricos de patentes, ou patentometria, são indicadores relevantes para verificação de competitividade em CT&I, por exemplo, entre instituições públicas e privadas. Este viés métrico tem ganhado notoriedade tanto no ambiente acadêmico (como programas de pesquisas) como no governo, a fim de analisar o desenvolvimento e a distribuição tecnológica em níveis micro, meso e macro. (Bochi; Gabriel Junior; Moura, 2020)

#### 2.4.3.1 Geolocalização das patentes

Os documentos de patentes possuem a informação de países de referência, informações úteis para comparações e análises. Os países podem ser atribuídos com relação ao país de residência do requerente; o país de residência do inventor (geralmente o endereço do local de trabalho ou laboratório); ou o país de escritório de prioridade, ou seja, o país em que primeiro foi depositada a patente (OECD, 2009).

Dentre as abordagens para identificar a localização de uma patente, destaca-se a apontada pela OECD. Aqui, se considera primeiramente uma oposição entre patentes concedidas de fato *versus* patentes solicitadas. Posteriormente, quatro tipologias são definidas conforme Quadro 5. (Okubo, 1997)

**Quadro 5** – Tipos de abordagens para definir a localização da patente

Abordagem	Contribui para
Número de pedido de patentes por residentes (inventores que residem no país por determinado longo período)	Concepção sobre a produção tecnológica, além de viabilizar a extração de informações adicionais de coeficiente de inventividade (em que se calcula o número de pedidos de patentes por 10 mil habitantes)
Número de pedidos de patentes por não residentes (submissão de patentes realizada por inventores que não necessariamente residam no país em questão)	Concepção sobre a penetração tecnológica de outros países no país em questão
Número de pedidos de patentes nacionais (soma-se os pedidos de residentes e não residentes)	Concepção do tamanho do mercado de tecnologia que o país representa
Número de patentes requeridas no exterior (verifica-se inventores que residam no país em questão com aplicações em outros países)	Concepção da difusão tecnológica do país em questão para cenários internacionais

Fonte: A autora (2025).

Nota: Adaptado de Okubo (1997).

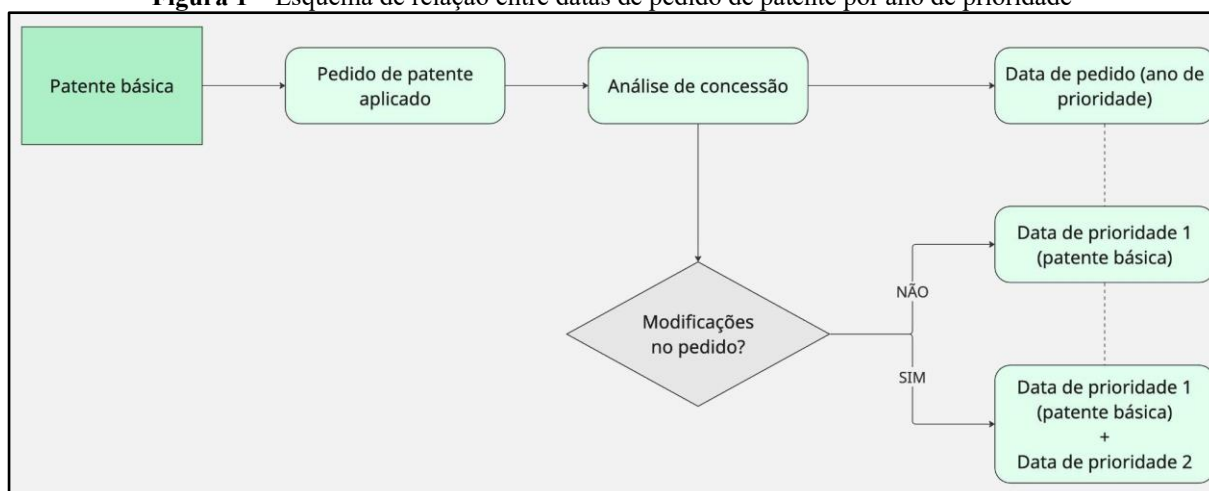
Neste sentido, a análise das diferentes origens das patentes, que podem ser nacionais ou internacionais, é um elemento crucial para fornecer *insights* específicos e multifacetados sobre a dinâmica da inovação tecnológica. Estas diversas fontes de titularidade e depósito refletem não apenas as capacidades intrínsecas de geração de conhecimento de cada região, mas também as estratégias de proteção de propriedade intelectual adotadas, qualidade de regulamentações, atratividade do processo de patenteamento, questões econômicas, de custos e regras de patenteamento, e afins. A origem da patente é um indicador-chave da penetração tecnológica em diferentes mercados ou regiões geográficas. Patentes depositadas por entidades estrangeiras, por exemplo, sugerem a aspiração de proteger a tecnologia em novos territórios e indicam uma potencial penetração no mercado local. Em contraste, um alto volume de patentes de origem nacional pode sinalizar a força e a autossuficiência de um país em gerar e proteger sua própria base tecnológica. Em suma, o mapeamento dessas origens é fundamental para delinear políticas públicas e estratégias empresariais voltadas para o fomento da inovação, a promoção da concorrência e a maximização dos benefícios econômicos e sociais da produção, difusão e penetração tecnológica (OECD, 2009; Okubo, 1997).

#### 2.4.3.2 Escala temporal

A data de pedido de uma patente não é a mesma que a data de concessão ou publicação (Figura 1). Isso ocorre porque há um atraso significativo entre o pedido e a concessão, podendo levar mais de 49 meses (caso de um escritório Europeu). Essa lacuna temporal pode influenciar significativamente na análise de patentes, atribuindo assim peso na escolha da escala (Schmoch, 2005).

Neste contexto, a data de uma patente pode ser a do seu pedido inicial (que se relaciona mais diretamente com a data da invenção e reflete as atividades de empresas ou instituições) ou uma data posterior, caso haja modificações após esse primeiro pedido. Nos casos de patentes com mais de uma data de prioridade, ocorre quando uma patente já aplicada para análise sofre alterações durante o processo e uma nova data é elaborada, entretanto conservando a data de primeira aplicação da patente básica (Schmoch, 2005).

**Figura 1** – Esquema de relação entre datas de pedido de patente por ano de prioridade



Fonte: A autora (2025).

Nota: Adaptado de *Schmoch* (2005) utilizando a ferramenta *Miro*.

Essas variações de ano de prioridade podem ser diferentes em cada região, com base no aparato legislativo. Entretanto, sugere-se utilizar apenas o ano mais antigo disponível no documento, para simplificar a contagem e reduzir duplicatas. *Schmoch* (2005)

Segundo a *OECD* (2009), a data de referência para uma patente pode ser: a) data de prioridade – primeira data de depósito de um pedido de patente, em qualquer lugar do mundo, para proteger uma invenção; b) data de aplicação em um determinado escritório – com prazos variando entre depósito em países estrangeiros e nacionais; c) a data da publicação – reflete o momento em que as informações sobre a invenção são divulgadas ao público e disponibilizadas para análises estatísticas; e por fim d) a data de concessão – quando os direitos das patentes são conferidos ao requerente pelo órgão autorizado.

Indicadores de desempenho inventivo baseados em datas de pedido/concessão de patentes apresentam vieses. Eles dependem de atrasos administrativos e estratégias do titular, dificultando a comparação internacional. O tempo entre a prioridade e o pedido/concessão varia por país, sendo que inventores nacionais depositam imediatamente, enquanto estrangeiros o fazem um ano depois. Contabilizar ambos pelo ano de pedido compara invenções de anos distintos, introduzindo distorções em épocas de rápida mudança ou crescimento tecnológico. Para escolher qual indicador de data de patente, deve-se considerar tais vieses e alinhar com o foco de pesquisa (*OECD*, 2009)

#### 2.4.3.3 Citação de patentes

Outro indicador no segmento de patentes é o número de citações, medindo assim o

impacto científico e, sobretudo, tecnológico. Embora não exista uma maneira absoluta ou relativa de se medir o valor de uma patente, as citações de umas patentes são amplamente utilizadas para mensurar C&T.

Neste sentido, considera-se que em um documento de patente pode possuir em sua primeira página referências de patentes já aprovadas dentro de um mesmo assunto, atribuído pelo avaliador da patente durante o processo de registro. Através da citação, descreve-se o estado da técnica, relacionando assim o que já foi feito em áreas semelhantes relacionando à invenção proposta e sua importância. Assim, o indicador de citação de patentes reflete a importância de uma patente para uma tecnologia que se busca proteger, e são comumente utilizadas em análises bibliométricas (Okubo, 1997)

#### 2.4.3.4 Classificação Internacional de Patentes (IPC)

A Classificação Internacional de Patentes (IPC) foi estabelecida em outubro de 1975 por meio do Acordo de Estrasburgo, com o objetivo de uniformizar a classificação de patentes em nível internacional. Disponibilizada pela WIPO (*World Intellectual Property Organization*) nos idiomas inglês e francês, em textos idênticos, a IPC serve como uma ferramenta essencial para escritórios de propriedade intelectual e demais usuários. (Brasil; MDIC; INPI, 2025)

Entre seus principais propósitos, a classificação visa organizar os documentos de patentes de forma ordenada, facilitando o acesso às informações tecnológicas e legais; servir de base para a disseminação seletiva de informações a todos os usuários; auxiliar na investigação do estado da técnica; e fornecer subsídio para a elaboração de estatísticas sobre a propriedade industrial, por meio da avaliação do desenvolvimento das diversas áreas. Além disso, a IPC é fundamental para estabelecer o que constitui novidade e para avaliar a atividade inventiva em um pedido de patente. (Brasil; MDIC; INPI, 2025)

A classificação é organizada de forma hierárquica, utilizando uma disposição e símbolos específicos. O nível superior é a Seção, que se divide em oito classes, representadas pelas letras maiúsculas de A a H. A Seção A trata de Necessidades humanas; a B, de Operações de processamento e transporte; a C, de Química e metalurgia; a D, de Têxteis e papéis; a E, de Construções fixas; a F, de Engenharia mecânica, iluminação, aquecimento, armas e explosão; a G, de Física; e a H, de Eletricidade. Dentro de cada Seção, há Subseções com títulos informativos; por exemplo, a Seção "Necessidades Humanas" inclui Subseções sobre "Agricultura" ou "Produtos Alimentícios". (Brasil; MDIC; INPI, 2025)

Na sequência, as Classes são o segundo nível hierárquico da classificação representado pelo símbolo por um número de dois dígitos (exemplo: H01 Elementos elétricos). As classes podem abranger uma ou mais Subclasses, sendo o terceiro nível hierárquico da classificação, sendo uma letra maiúscula (exemplo: H01S dispositivos utilizando o processo de amplificação da luz por emissão estimulada de radiação [*laser*] para gerar ou amplificar luz; dispositivos utilizando emissão estimulada de radiação eletromagnética em faixas de frequência outras que não a óptica). (Brasil; MDIC; INPI, 2025)

O Grupo é o desdobramento das Subclasses, cujo símbolo consiste no símbolo da Subclasse seguido de dois números separados por uma barra oblíqua. Nesta hierarquia entende-se o Grupo principal, quando o símbolo da subclasse possui um número, a barra e dois dígitos zeros (exemplo: H01S 3/00); e o Subgrupo, que é a subdivisão dos grupos principais em que o símbolo contempla um número após a barra inclinada de até três dígitos que não sejam os dois zeros (exemplo: H01S 3/02). Ainda possuem diferentes níveis de indentação, dentro da hierarquia dos subgrupos, conforme Figura 2 à seguir. (Brasil; MDIC; INPI, 2025)

**Figura 2** – Exemplo de níveis de indentação dos subgrupos da IPC

Exemplo: G01N33/483	• •	Análise física de material biológico
33/487	• • •	de material biológico líquido
33/49	• • • •	do sangue
33/50	• •	Análise química de material biológico, p. ex. sangue

Fonte: (Brasil; MDIC; INPI, 2025)

Cada documento de patente recebe pelo menos um código primário e frequentemente códigos secundários, atribuídos pelos examinadores dos escritórios de patente (Hinze; Schmoch, 2005).

Adotada por mais de 100 países, essa classificação é essencial para a recuperação de documentos de patentes, especialmente na busca pelo "estado da técnica", e auxilia na localização do conteúdo técnico de uma patente. Tal recuperação é indispensável para diversas partes interessadas, como autoridades emissoras de patentes, potenciais inventores, unidades de pesquisa e desenvolvimento, e outros interessados na aplicação ou desenvolvimento de tecnologia. (Brasil; ME; INPI, 2021)

Embora este sistema de classificação seja um método para analisar a área tecnológica de uma patente, em certos casos, pode ser recomendada a adoção de sistemas híbridos para análise, incorporando, por exemplo, palavras-chave (OECD, 2009).

#### 2.4.4 Indicadores de Sustentabilidade

No cenário multidimensional que compõe os desafios para o desenvolvimento sustentável, reafirmam-se ciência e tecnologia como eixo central do processo civilizatório. O Brasil se torna um signo ecológico mundial e ganha destaque na história e no futuro da humanidade (Freitas; Freitas, 2016). A composição entre a interdisciplinaridade e a complexidade referente à temática sustentabilidade, encontra-se a exigência de sistemas interligados e inter-relacionados para que se quantifique e agregue a quantificação e visualização das informações em seu processo comunicativo otimizado (Van Bellen, 2005).

Os indicadores de sustentabilidade devem ir além de indicadores ambientais ou sobre crescimento. Eles devem abordar aspectos sobre tempo, limites, suficiência, eficiência, equidade e qualidade de vida. O uso de diversos indicadores é justificado pelos diversos propósitos e diferentes visões de mundo e os indicadores de sustentabilidade mais abrangentes podem contribuir para sua compreensão de modo a transcender nações e culturas. (Meadows, 2021).

Foi com base nos sinais de mudanças climáticas evidentes, que discussões acerca de desenvolvimento sustentável, exploração consciente e preservação dos recursos naturais, ganham cada vez mais enfoque no mundo com um certo senso de urgência. Nesse sentido, a ONU, da qual o Brasil é integrante, propõe como desafio metas estruturadas em 17 vertentes, denominadas como Objetivos de Desenvolvimento Sustentável (ODS), como forma de apelo para ações humanitárias e para proteção do meio ambiente. Além disso, um prazo foi estabelecido para o cumprimento das metas que compõem tais objetivos, chamado “Agenda 2030”. Com vistas à Agenda 2030 surge, por exemplo, o “Projeto Amazônia 2030” por meio de iniciativa de pesquisadores com intuito de desenvolver um plano de desenvolvimento sustentável para a Amazônia Legal.

Os objetivos são: 1) Acabar com a pobreza em todas as suas formas, em todos os lugares; 2) Acabar com a fome, alcançar a segurança alimentar e melhoria da nutrição e promover a agricultura sustentável; 3) Assegurar uma vida saudável e promover o bem-estar para todas e todos, em todas as idades; 4) Assegurar a educação inclusiva e equitativa e de qualidade, e promover oportunidades de aprendizagem ao longo da vida para todas e todos; 5) Alcançar a igualdade de gênero e empoderar todas as mulheres e meninas; 6) Assegurar a disponibilidade e gestão sustentável da água e saneamento para todas e todos; 7) Assegurar o acesso confiável, sustentável, moderno e a preço acessível à energia para todas e todos; 8) Promover o

crescimento econômico sustentado, inclusivo e sustentável, emprego pleno e produtivo e trabalho decente para todas e todos; 9) Construir infraestruturas resilientes, promover a industrialização inclusiva e sustentável e fomentar a inovação; 10) Reduzir a desigualdade dentro dos países e entre eles; 11) Tornar as cidades e os assentamentos humanos inclusivos, seguros, resilientes e sustentáveis; 12) Assegurar padrões de produção e de consumo sustentáveis; 13) Tomar medidas urgentes para combater a mudança climática e seus impactos; 14) Conservação e uso sustentável dos oceanos, dos mares e dos recursos marinhos para o desenvolvimento sustentável; 15) Proteger, recuperar e promover o uso sustentável dos ecossistemas terrestres, gerir de forma sustentável as florestas, combater a desertificação, deter e reverter a degradação da terra e deter a perda de biodiversidade; 16) Promover sociedades pacíficas e inclusivas para o desenvolvimento sustentável, proporcionar o acesso à justiça para todos e construir instituições eficazes, responsáveis e inclusivas em todos os níveis; e 17) Fortalecer os meios de implementação e revitalizar a parceria global para o desenvolvimento sustentável. Por meio destes é possível monitorar, mapear e medir as ações realizadas em prol do desenvolvimento sustentável. (Nações Unidas Brasil, 2024)

Nesse sentido, bases de dados internacionais traçaram estratégias e metodologias específicas para viabilizar a recuperação de fontes relacionadas aos ODS, contribuindo assim para a investigação e monitoramento por pesquisadores, instituições e demais interessados no desenvolvimento da produção científica alinhado à cada meta. No caso da base *Scopus*, por exemplo, a equipe de cientista de dados da base elaborou expressões de buscas para consultas abrangendo os 16 primeiros ODS, parte da demanda advinda do *Times Higher Education* (THE). Com isto, é possível aplicá-las no modo de pesquisa avançada (*Scopus Advanced*) e recuperar resultados específicos de cada ODS (Apêndice B), combinando-os entre si ou combinando-os com outras expressões, de acordo com a necessidade do usuário ao realizar sua busca. As expressões estão disponíveis gratuitamente e com acesso livre via *Mendeley*.

O uso dos ODS para identificar sustentabilidade nas produções científicas já foi adotado por diversas fontes tais como a *United Nations Environment Program* (2015) e *Walz et al.* (2017) mas não há conhecimento de sua aplicação às publicações sobre plantas endêmicas da Amazônia.

Outra perspectiva que pode identificar aspectos de sustentabilidade nas publicações científicas é através do mapeamento de plantas que possuem alguma classificação de ameaça à extinção. Nesse sentido, o MMA publica listas oficiais de espécies da flora brasileira em diferentes categorias de ameaça de extinção, conforme a Portaria nº 443, de 17 de dezembro de 2014, ou a Portaria Nº 148, de 7 de junho de 2022. (Brasil; MMA, 2022)

Através dessas listas pode-se compreender aspectos como ameaças globais à biodiversidade de plantas, no processo de conservação da biodiversidade, ou superação dos desafios para implementação da “Estratégia Global para a Conservação de Plantas no Brasil” (Martinelli; Moraes, 2013). A Lista Vermelha de Espécies Ameaçadas já foi adotada por autores como Santos, Silva e Faria (2023). No contexto das plantas endêmicas da Amazônia, pode aplicar o conceito de Lista Vermelha para elaborar indicadores de publicações científicas sobre plantas ameaçadas de extinção.

No contexto das patentes, tem-se um viés que delimita a definição de inovação sustentável, também conhecido como “patente verde”. Neste contexto, Favot *et al.* (2023, p. 6) apontam diversas definições para a inovação ambiental, mas com duas características em comum que são a “[...] redução dos impactos negativos no meio ambiente e uso mais eficiente dos recursos”.

Para mensurar as eco-inovações, seria através da identificação das aplicações de patentes cuja abrangência seja voltada para tecnologias relacionadas com o ambiente, portanto as patentes verdes, que por sua vez

“[...] inclui pelo menos um código verde conforme definido pela **metodologia relacionada**. A definição de patente verde baseada apenas nos códigos de classificação é uma aproximação em comparação com a análise profunda do pedido único de patente [...]” (Favot *et al.*, 2023, p. 6; grifo nosso).

Por “metodologia relacionada” entende-se o tipo de metodologia adotada para levantar os dados das patentes verdes. Favot *et al.* (2023) relacionam três metodologias diferentes e sua taxa de recuperação em determinado recorte de dados de sua pesquisa. As metodologias, portanto, seriam 1] Inventário Verde com os códigos de classificação internacional de patentes (*Green Inventory* IPC), elaborado pela Organização Mundial da Propriedade Intelectual (OMPI), com uma recuperação de 80% das patentes; 2] ENV-TECH, elaborado pela OECD que retornou cerca de 40% dos resultados; e por fim 3] a combinação de ambas metodologias que resultou em 20% de recuperação.

O uso da classificação *Green Inventory* relacionada aos códigos da IPC já foi adotado por diversos autores, tais como Stefani Cativelli; Mañana Rodriguez; Lascurain Sanchez, 2022; Stefani Cativelli; Pinto; Lascurain Sanchez, 2020; Cui *et al.* (2022); D’Amato; Mazzanti; Nicolli (2021); Favot *et al.* (2023b); Liu e Wang (2023); Menezes, Santos e Bortoli (2016); Teixeira (2018); Toledo; De Campos (2018). Não há conhecimento da aplicação do conceito de patentes verdes para elaboração de indicadores de sustentabilidade sobre plantas endêmicas da Amazônia.

## 2.5 ARQUITETURA E VISUALIZAÇÃO DA INFORMAÇÃO

Conforme a OECD (2008), um indicador composto resulta da combinação de duas partes, e sua validade depende, principalmente, da qualidade das variáveis subjacentes.

Para a mensuração de múltiplos indicadores, especialmente em Ciência e Tecnologia (C&T), os métodos tradicionais de eficiência podem ser insuficientes. Dessa necessidade, surgem os indicadores compostos, pioneiros na orientação de políticas nacionais de C&T. No âmbito empresarial ou de inovação (nível micro), a utilização de scoreboards ou placares é bem estabelecida para decisões estratégicas, baseando-se em *benchmarking* e métricas agregadas (como a produtividade) para avaliar a posição frente à concorrência. Essa lógica se estende à C&T, resultando em “Painéis de P&D” ou “Painéis de patentes” (Grupp; Moguee, 2004).

O uso de *benchmarking* e painéis de indicadores enfrenta o problema da ausência de modelos teóricos claros que orientem a seleção e a ponderação dos indicadores, bem como a comparação entre diferentes países. A validação desses modelos é, portanto, essencial para prevenir manipulações no processo de seleção, ponderação e agregação. Embora Grupp e Moguee (2005) destaquem que o sucesso de uma análise baseada em painéis depende do domínio na seleção e design dos indicadores, os indicadores compostos têm a vantagem de simplificar situações complexas em um único valor, facilitando a comunicação, como no contexto político.

No monitoramento, o Esquema Estrela (Kimball; Ross, 2002) é frequentemente empregado na modelagem de dados, separando os dados em Tabelas de Fato (medições quantitativas e chaves de relacionamento) e Tabelas de Dimensão (contexto das medições). Este esquema é eficaz com ferramentas de *Business Intelligence* (BI) e oferece benefícios como suporte à performance – otimiza o desempenho de ferramentas de BI (ex: *Looker Studio*); integridade da chave – assegura a unicidade da chave de relacionamento entre os conjuntos de dados; análise contextualizada – permite a análise de intersecção, onde os Fatos são devidamente contextualizados pelas Dimensões.

A visualização da informação é um instrumento vital na Ciência da Informação, pois, além da estética, é crucial para reduzir a sobrecarga cognitiva do usuário e apoiar a tomada de decisão (Few, 2006; Tufte, 2007). Um Painel (*Dashboard*) dinâmico e interativo constitui uma solução eficaz de Inteligência Competitiva para o monitoramento periódico dos desenvolvimentos em Ciência, Tecnologia e Inovação (CT&I), transformando a simples apresentação de dados em uma ferramenta indispensável.

### 3 MÉTODOS

O presente capítulo visa detalhar o *corpus* metodológico que orientou o desenvolvimento do sistema de monitoramento automatizado de indicadores compostos. Dada a natureza do estudo, que envolve o desenvolvimento de um produto tecnológico (*dashboard*) e sua validação quantitativa, a metodologia foi estruturada em 3 fases progressivas e interconectadas. Esta abordagem em fases garante a rastreabilidade entre os Objetivos Específicos (OEs) da pesquisa e os procedimentos técnicos adotados, com foco particular na superação dos desafios de modelagem de dados heterogêneos e de automação do fluxo de informação.

#### 3.1 DELINEAMENTO, TIPO E ABORDAGEM DA PESQUISA

A presente investigação científica está alinhada a uma abordagem metodológica que combina as características de uma pesquisa quantitativa com um foco aplicado. O cerne do esforço de pesquisa reside na construção e validação de uma sistemática de *Extract, Transform, Load* (ETL), desenhada especificamente para estabelecer e visualizar a inter-relação dinâmica de um conjunto de indicadores.

Para materializar este objetivo central, o projeto culmina no desenvolvimento de um produto tecnológico tangível: um protótipo de painel de indicadores (*Minimum Viable Product* - MVP). Este painel concentra os indicadores compilados e processados pela sistemática ETL (Apêndice E).

O propósito fundamental deste empreendimento é oferecer uma solução robusta e inovadora para o problema de pesquisa identificado. Para isso, a metodologia emprega e integra de forma sinérgica métodos e técnicas avançadas, provenientes de duas áreas do conhecimento: a Bibliometria e a Ciência de Dados. A Bibliometria contribui com os fundamentos para a seleção e análise dos indicadores, enquanto a Ciência de Dados fornece as ferramentas e algoritmos para a manipulação, processamento (via ETL) e visualização eficaz dos dados, garantindo a representação dinâmica e interconectada das informações.

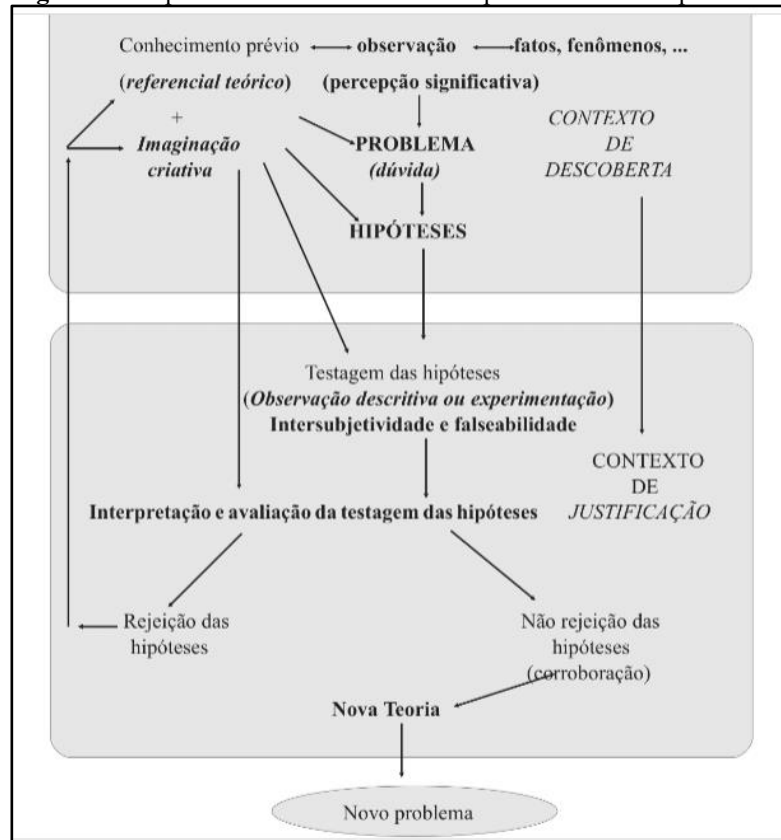
A abordagem quantitativa utiliza conceitos matemáticos e de processamento de informação (Rayward, 1996). Entende-se também, na perspectiva quantitativa, a característica estatística que apresenta como papel “[...] fornece uma descrição quantitativa da sociedade, considerada como um todo organizado” (Lakatos; Marconi, 2003, p. 108). O método estatístico constitui-se, neste trabalho, como o mecanismo de validação e prova que transcende a mera

descrição, permitindo a análise sistemática e o monitoramento dos dados bibliométricos e de sustentabilidade.

O tipo da pesquisa adota inicialmente o exploratório-descritivo, em que se realiza o levantamento das premissas e de sua caracterização, para descrever a natureza das premissas contidas no referencial teórico (Koche, 2011).

A pesquisa possui delineamento pelo método científico hipotético-dedutivo. Koche (2011) sugere um esquema para aplicação do método hipotético dedutivo, conforme Figura 3. No esquema proposto denota-se dois contextos, o primeiro da descoberta e o segundo da justificação. No contexto da descoberta, na ciência contemporânea, parte de questionamentos, problemáticas ou incompatibilidades teóricas observadas, que por sua vez fomentam a necessidade de buscar por respostas, críticas e/ou soluções. Nesse sentido, o problema da pesquisa é também a pergunta que não se pode responder com o conhecimento prévio. Outra consideração importante é o fato da ciência não se tratar de mera observação dos fenômenos, portanto um fato por si só não tem relevância a não ser que se relacione a um problema ou dúvida.

**Figura 3** - Esquema do método científico hipotético-dedutivo por Koche



Fonte: Koche (2011).

Partindo da elaboração do problema — lacuna de indicadores de pesquisas e tecnologias/ inovações sobre plantas endêmicas do bioma amazônico — é possível conjecturar as premissas a) o potencial comercial e demais benefícios provenientes de espécies endêmicas da Amazônia podem resultar em produções científicas e criação de novas tecnologias e inovação; b) tais resultados tradicionalmente são publicados através de artigos científicos e patentes, dentre outros formatos; c) é passível de se mensurar tais resultados através de indicadores de CT&I; d) é possível integrar sistematicamente tais indicadores, somado a indicadores de sustentabilidade, a fim de realizar mapeamentos e monitoramentos para análises e diagnósticos.

Dentre a problemática e premissas apontadas, é possível elaborar hipóteses, conforme orienta Koche (2011), a fim de obter potenciais soluções e respostas. Nesse sentido, tendo por perspectiva as dimensões de CT&I e estudos métricos da informação, esta pesquisa propõe a hipótese de que o desenvolvimento de um sistema de monitoramento automatizado, baseado em um *pipeline* ETL capaz de integrar dados heterogêneos (*Scopus*, *Espacenet* e fontes de sustentabilidade), e composto por indicadores bibliométricos em intersecção com indicadores classificatórios de sustentabilidade, viabilizará (ou favorecerá de forma eficaz) o monitoramento periódico da dinâmica dos desenvolvimentos científicos e tecnológicos sobre plantas endêmicas da Amazônia

Ao que concerne o segundo contexto do esquema de justificativa proposto por Koche (2011), a presente pesquisa orienta ao quadro teórico, conforme explorado no referencial teórico como fundamentação e plausibilidade das problemáticas, bem como a utilização dos estudos métricos da informação como instrumentos, técnicas de pesquisa e parâmetros (Araújo, 2009; Rayward, 1996; Schmoch, 2005; van Leeuwen, 2005)

### 3.1.1 Plano de gestão de dados e ferramentais

Ressalta-se a implementação do Plano de Gestão de Dados (PGD), considerando os princípios FAIR – Findable, Accessible, Interoperable, Reusable; ou seja, seguindo boas práticas na gestão e compartilhamento de dados de pesquisa com vistas à tornar dados científicos localizáveis, acessíveis, interoperáveis e reusáveis (Wilkinson *et al.*, 2016). Para a criação do PGD utilizou-se a ferramenta DMPTOOL<sup>6</sup> conforme sugestão da Fundação de

---

<sup>6</sup> DMPTOOL. Disponível em: <https://dmptool.org/plans>. Acesso em: 20 out. 2025.

Amparo à Pesquisa do Estado de São Paulo (FAPESP)<sup>7</sup>.

Os versionamentos e ferramentais utilizados na automatização no processo de Extração, Transformação e Carga<sup>8</sup> (ETL– *Extract, Transform, Load*) dos dados encontram-se no repositório do *GitHub* sob a publicação via Identificação de Objeto Digital (DOI–*Digital Object Identifier*)<sup>9</sup>. Justifica-se o uso do *GitHub* dado seu acesso gratuito e, de acordo com a *Microsoft* (2024), trata-se de um site para armazenamento de projetos *Git*, e funciona como um hub que possibilita compartilhamento e trabalho colaborativo em código. Um exemplo de sua relevância está no uso de instituições do próprio Governo Federal do Brasil, conforme o Ministério da Gestão e da Inovação em Serviços Públicos, sobre o Portal do Software Público Brasileiro (Brasil, 2023).

### 3.1.2 Fases e Etapas da Pesquisa

Partindo das definições prévias, a pesquisa segue considerando 3 fases, subdivididas em 11 etapas. As fases compõem os pontos macros da pesquisa - Planejamento e preparação inicial; Execução; Revisão e atualizações - e as etapas, que são os pontos subordinados a seguir para concluir cada fase da pesquisa, conforme Quadro 6.

Com base no exposto, inicia-se a fase 1 com o delineamento da pesquisa e protocolo de coleta de dados.

---

<sup>7</sup> FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO (FAPESP). FAQ - Plano de Gestão de Dados. Disponível em: <https://fapesp.br/16923/faq-plano-de-gestao-de-dados>. Acesso em: 20 out. 2025.

<sup>8</sup> O que é ETL. Disponível em: <https://cloud.google.com/learn/what-is-etl?hl=pt-BR>. Acesso em: 27 out. 2025.

<sup>9</sup> DOI para repositório do GitHub: <https://doi.org/10.5281/zenodo.16851246>. Acesso em: 20 out. 2025.

Quadro 6 - Fases e etapas do procedimento metodológico alinhado aos objetivos específicos

FASE	ETAPAS (Procedimentos e Ferramentas)	OBJETIVOS ESPECÍFICOS VINCULADOS
Fase 1 - Delineamento e Protocolo de Busca (3.2)	<b>1.1 Definição do Escopo e Bases de dados:</b> A. Escolha das espécies endêmicas (Reflora - CNCFlora) B. Escolha das fontes ( <i>Scopus e Espacenet</i> ).	<b>OE 1:</b> Definir a estratégia de busca, as bases de dados e o protocolo de coleta
	<b>1.2 Protocolo de Busca:</b> A. Elaboração das expressões de buscas principais (taxonômicas); B. Elaboração das expressões classificatórias de sustentabilidade ( <i>ODS, IPC Green Inventory</i> ).	<b>OE 1:</b> Definir a estratégia de busca, as bases de dados e o protocolo de coleta
	<b>1.3 Definição de indicadores:</b> Escolha e justificativa dos indicadores diretos (bibliométricos) e indiretos, e sua representação.	<b>OE 2:</b> Elaborar e integrar os indicadores bibliométricos e os indicadores de sustentabilidade
Fase 2 - Desenvolvimento, tratamento e construção do protótipo para painel (3.3)	<b>2.1 Busca e coleta de dados brutos:</b> A. Extração dos dados das fontes primárias ( <i>Scopus, Espacenet</i> ); Extração dos dados das fontes secundárias (CNCFlora, <i>ODS, IPC Green Inventory</i> ).	<b>OE 2:</b> Elaborar e integrar os indicadores bibliométricos e os indicadores de sustentabilidade
	<b>2.2 Desenvolvimento do Pipeline ETL:</b> Criação de scripts para a extração, transformação e padronização dos dados heterogêneos ( <i>Scopus e Espacenet</i> ).	<b>OE 3:</b> Desenvolver o pipeline ETL para a coleta, o tratamento e a harmonização dos dados heterogêneos
	<b>2.3 Modelagem de Dados:</b> Aplicação da lógica para cruzamento de dados e interseção entre os indicadores classificatórios de sustentabilidade com os respectivos artigos e patentes.	<b>OE 2:</b> Elaborar e integrar os indicadores bibliométricos e os indicadores de sustentabilidade
	<b>2.4 Carga (Loading) e Implantação da Plataforma Analítica de Indicadores:</b> Criação do protótipo do painel de visualização no <i>Looker Studio</i> e implementação dos recursos dinâmicos.	<b>OE 4:</b> Configurar e implementar o protótipo do painel de visualização ( <i>dashboard</i> )
Fase 3 - Revisão, análise e protocolo (3.4)	<b>3.1 Revisão e Validação do Sistema:</b> Testes de funcionalidade e acurácia da visualização, dos indicadores e da execução do pipeline ETL.	<b>OE 4:</b> Configurar e implementar o protótipo painel de visualização ( <i>dashboard</i> )
	<b>3.2 Análise dos Resultados:</b> Aplicação das técnicas de análise bibliométrica/patentométrica aos indicadores compostos.	<b>OE 5:</b> Analisar e discutir os resultados dos indicadores compostos para mapear a dinâmica
	<b>3.3. Proposição do Protocolo de Atualização:</b> Documentação do processo de replicação, manutenção e atualização periódica do sistema.	<b>OE 6:</b> Propor um protocolo de atualização e manutenção do sistema de monitoramento

Fonte: A autora (2025).

## 3.2 FASE 1: DELINEAMENTO, PROTOCOLO DE BUSCA E DEFINIÇÃO DE INDICADORES

A primeira fase da pesquisa compreende três etapas sequenciais: as definições iniciais para busca (3.2.1), a definição dos indicadores (3.2.2) e a definição de estrutura de visualização dos indicadores (3.2.3), detalhados a seguir.

### 3.2.1 Definições iniciais para expressão de busca

A primeira etapa consiste na elaboração da expressão de busca para levantamento dos dados bibliométricos nas bases. Considerou-se, portanto, duas abordagens possíveis: a primeira voltada para uma pesquisa abrangente, em que pode-se partir do geral - sem definir espécies ou quaisquer outros delineamentos - ou a abordagem de selecionar especificamente as espécies nativas que ocorrem exclusivamente no domínio fitogeográfico da Amazônia brasileira.

A primeira abordagem apresenta como vantagem a possibilidade de encontrar resultados mais abrangentes, independentemente de ser flora ou fauna, de abordar espécies endêmicas ou não endêmicas, ou diversos outros tipos de assuntos que podem abarcar a Amazônia. Por outro lado, a desvantagem seria uma demanda maior de dados que não necessariamente representaria insumos encontrados somente no bioma amazônico do Brasil. Na segunda abordagem aponta-se, a contraposto da opção anterior, a vantagem de delimitar os resultados da pesquisa de modo mais objetivo, cuja análise aborda plantas encontradas somente no bioma amazônico do Brasil, e obviamente restringe-se de outras plantas e insumos passíveis de análise reduzindo o escopo da busca.

Deste modo, optou-se pela segunda alternativa, utilizando a nomenclatura binominal em latim das espécies da flora endêmica do bioma amazônico no Brasil. Esta escolha baseia-se na premissa de que as fontes primárias (patentes e artigos) devem mencionar a espécie em questão, conforme Santos, Silva e Faria (2023). Quanto à patente, a menção à uma espécie pode se dar por uma série de motivos, tais como ser a fonte ou o alvo da invenção, e segundo Oldham, Barnes e Hall (2015) os requerentes de uma patente podem mencionar um grande número de famílias, gêneros ou espécies genéticas tanto na descrição quanto na reivindicação, a fim de incorporar qualquer um dos componentes advindos do bioinsumo.

Entretanto, essa estratégia de busca possui algumas limitações, como apontam Santos, Silva e Faria (2023, p. 2), já que ignoram as variantes de nomes científicos; espécies que podem

variar por regiões, idiomas étnicos e países; a possibilidade de que a patente mencione somente os mais elevados níveis taxonômicos ou somente os nomes isolados de compostos provenientes da espécie em questão. Por fim, as limitações podem se dar também a depender das capacidades de busca da base de dados de patentes, assim como espécies que podem “[...] aparecer nos campos de busca de patentes não porque sejam objeto das reivindicações da invenção, mas porque são alvo dela, como solução para uma espécie que causa doenças ou é considerada uma praga.”.

Para isto, definimos como fonte de pesquisa para elencar os nomes científicos binominal em latim das espécies, a lista disponível no *site* do Programa Re flora CNCFlora - Flora e Funga do Brasil<sup>10</sup> (apresentado previamente na Revisão Teórica). Todavia, o *site* possui uma interface de busca para o usuário, com campos a preencher de acordo com sua necessidade, cujos critérios foram definidos e explicitados no Quadro 7 abaixo. Cabe a ressalva de que os campos não mencionados abaixo, mas são passíveis de preenchimento, ficaram vazios para a busca em questão, visto que seu preenchimento não era pertinente ao escopo da busca.

**Quadro 7** - Critérios de busca utilizados no *site Re flora*

Campos de busca do site	Critérios definidos
Nome	- <b>Grupo:</b> Flora e Funga - <b>Família:</b> Todos
Forma de vida e substrato	- <b>Forma de vida:</b> Todos - <b>Substrato:</b> Todos
Abrangência geográfica	- <b>Ocorre no Brasil:</b> Sim - <b>Ocorrência:</b> Só ocorre em - <b>Endemismo:</b> Só endêmicas do Brasil - <b>Origem:</b> Todos
Distribuição	- <b>Região:</b> Qualquer - <b>Estado:</b> Qualquer - <b>Domínio fitogeográfico:</b> Amazônia
Vegetação	Todos
Buscar até	Espécies
Opções de busca	Buscar só nomes aceitos

Fonte: A autora (2023).

Contudo, o Projeto *Re flora* informa seu *status* de pesquisa em andamento até o presente momento da pesquisa, portanto os dados podem ser atualizados a qualquer momento. Ou seja, plantas que estão em processo de investigação aqui levantadas como endêmicas, podem sofrer

<sup>10</sup> Disponível em: [https://reflora.jbrj.gov.br/consulta/?grupo=6&familia=null&genero=&especie=&autor=&nomeVernaculo=&nomeCompleto=&formaVida=null&substrato=null&ocorreBrasil=SIM&ocorrencia=SOMENTE\\_OCORRE&endemismo=ENDEMICA\\_BRASIL&origem= TODOS&regiao=QUALQUER&ilhaOceanica=32767&estado=QUALQUER&domFitogeograficos=AMAZONIA&vegetacao=TODOS&mostrarAte=ESPECIES&opcoesBusca=NOME\\_ACEITO&loginUsuario=Visitante&senhaUsuario=&contexto=consulta\\_publica&pagina=1#CondicaoTaxonCP](https://reflora.jbrj.gov.br/consulta/?grupo=6&familia=null&genero=&especie=&autor=&nomeVernaculo=&nomeCompleto=&formaVida=null&substrato=null&ocorreBrasil=SIM&ocorrencia=SOMENTE_OCORRE&endemismo=ENDEMICA_BRASIL&origem= TODOS&regiao=QUALQUER&ilhaOceanica=32767&estado=QUALQUER&domFitogeograficos=AMAZONIA&vegetacao=TODOS&mostrarAte=ESPECIES&opcoesBusca=NOME_ACEITO&loginUsuario=Visitante&senhaUsuario=&contexto=consulta_publica&pagina=1#CondicaoTaxonCP). Acesso em: 14 dez. 2023.

futuras alterações em caso de ser localizado em outros espaços fora do bioma amazônico. Com isto, fixa-se nos resultados obtidos na data mencionada anteriormente (14 de dezembro de 2023).

A listagem com os 1975 nomes científicos foi exportada para a planilha do *Google* e as respectivas páginas do site salvas como arquivo em formato “*PDF*”. Ressalta-se ainda que, ao aplicar os filtros e critérios de busca mencionados no Quadro 2 acima, automaticamente a base descarta os táxons subordinados às algas. Portanto, dada a extensão de espécies levantadas (1975), optou-se por não incluí-las (algas) na presente pesquisa, deixando-as para uma oportunidade futura de pesquisa.

Na ferramenta *Google Planilhas* foi utilizado fórmula que permite a junção dos termos inserindo-os entre as aspas e combinando com operadores *Booleanos* “*OR*”, para finalmente obter as expressões de buscas (Apêndice A).

### 3.2.2 Definição dos tipos de fontes

Definida a estratégia de busca (3.2.1), a etapa subsequente consistiu na seleção das fontes de dados para a aplicação do protocolo. Entende-se por fontes os recursos informacionais que contém dados que viabilizam análises. No âmbito da CI, é amplamente utilizado e aceito a realização de análise de recursos informacionais advindos de publicação científica e tecnológica. Conforme contextualização abordada no referencial teórico da presente pesquisa, definimos, portanto, nossas fontes para análise a) artigos científicos; e b) patentes (Dias; Dias; Moita, 2021; Mueller, 2013; van Leeuwen, 2005)

Portanto, os tipos de fontes principais se concentram nos dados dos registros de artigos científicos e patentes publicados em bases de dados bibliométricas. Outras fontes secundárias foram obtidas em *sites* oficiais para complementar o cruzamento de dados, sendo estes: Organização Mundial da Propriedade Intelectual<sup>11</sup> (OMPI) – dados de Classificação Internacional de Patentes (CIP) e de patentes verdes (*Green Inventory*); Re flora Flora e Funga do Brasil<sup>12</sup>, do Jardim Botânico do Rio de Janeiro – base de validação taxonômica; e base de

---

<sup>11</sup> **OMPI** - Disponível em: <https://www.wipo.int/portal/en/index.html>. Acesso em: 11 fev. 2025.

<sup>12</sup> **Flora e Funga do Brasil**. Jardim Botânico do Rio de Janeiro. Disponível em: <http://floradobrasil.jbrj.gov.br/>. Acesso em: 11 fev. 2025

dados sobre níveis de ameaça de floras do CNCFlora<sup>13</sup> – status de conservação e níveis de ameaça (ex: "Em perigo", "Vulnerável").

Estas fontes secundárias, embora externas às bases bibliométricas, são essenciais para a elaboração dos indicadores compostos e a intersecção com o viés de sustentabilidade, dada a heterogeneidade inerente aos dados bibliométricos e classificatórios.

### 3.2.3 Definição da base de dados

Para definir a base para coletar os dados para análise, primeiramente considerou-se que esta deve possuir as seguintes características básicas: abrangência geográfica (compreender artigos e patentes de diversas origens); abrangência de conteúdo (ser multidisciplinar); recursos para bibliometria (disponibilizar opção de exportação de dados bibliométricos dos registros recuperados).

Dentre os requisitos apontados, diversas bases atendem aos critérios, que são de uso frequente na comunidade científica tanto para artigos quanto para patentes, como a *Scopus*, *Web of Science*, *Derwent*, *Espacenet*, *The Lens*, dentre outras.

A *Scopus* é uma base de dados que abrange 240 disciplinas, mais de 2.6 bilhões de citações desde o ano de 1970, mais de 25 milhões de documentos em acesso aberto, além de vários livros e pré-prints, acervo contabilizado em mais de 90 milhões de registros. Ela se destaca devido cobertura global e por indexar títulos de todas regiões geográficas em até 40 idiomas locais – desde que esses registros possuam resumo em inglês. Seu uso por pesquisadores, bibliotecários e pessoas no setor financeiro utilizam-a como ferramenta de inteligência, já que fornece métricas e perfis avaliam impacto e desempenho acadêmico. Dentre suas principais funcionalidades estão as Métricas de Artigo e Citação – como o *Field-Weighted Citation Impact* (FHCI) – Métricas Alternativas (*PlumX Metrics*) que captam pegadas online de uso, mídias sociais, menções e afins; Análise institucional e localização de autores; Análise por tópicos (*SciVal Topics*) e ODS. (Elsevier, 2023)

A base de dados *Web of Science* (WoS) é multidisciplinar e indexa periódicos (mais de 21 mil, atualmente), conferências e livros com maior número de citações em suas respectivas áreas de conhecimento. É também utilizada como índice de citações atribuídas e recebidas por

---

<sup>13</sup> Base de dados de espécies em ameaça CNCFlora. Disponível em: <https://ckan.jbrj.gov.br/dataset/?q=flora+ameacada+2021>. Acesso em: 11 fev. 2025.

trabalho, proprietária da *Clarivate Analytics*. (Brasil; MEC; CAPES, 2024) A WoS é considerada uma ferramenta essencial para estudos bibliométricos, dado suas funcionalidades de Pesquisa de Referência Citada (*Cited Reference Searching*) para análise de influências; Ferramenta de Análise (*Analyze Tool*) que auxiliam na descobertas de padrões ocultos; dentre outros aspectos (Thomson Reuters, 2011).

A *Derwent* (ou *Derwent Innovations Index - DII*), é uma aplicação proprietária da *Clarivate Analytics* (WIPO, 2025). Sua disponibilização está através da interface da WoS e conta com um acervo de referências e resumos de mais de 11 milhões de patentes, incluindo publicações de mais de 60 autoridades emissoras de registros e concessões de patentes nacional e internacionalmente (Brasil; MEC; CAPES, 2024). O núcleo da plataforma é o *Derwent World Patents Index* (DWPI) e seu uso para análise bibliométrica (ou inteligência de patentes) se dá principalmente com base em suas configurações. A base foi projetada justamente para fornecer dados e ferramentas confiáveis para monitorar tendências tecnológicas, cenários competitivos, embasar pareceres de liberdade de operação (FTO), processar patentes, monetizar ativos e apoiar atividades de litígio. A base conta ainda indexadores humanos que agrupam dados de famílias das patentes, além de obter recursos para análise de citação, padronização de titulares das patentes e previsões de *status* legal (WIPO, 2025). Contudo, a base é de acesso restrito e pode ser acessada pelo Portal de Periódicos CAPES.

A *Espacenet* é uma base de dados mantida pelo Escritório Europeu de Patentes (EPO) e oferta mais de 130 milhões de documentos de patentes de mais de 100 milhões de países (INPI, 2022). Em sua interface há a possibilidade de busca inteligente (*Smart search*) e busca avançada (*Advanced search*) que permite uso de recursos de truncagem, por exemplo (*Espacenet*, 2022). Dentre as funcionalidades com potencial aplicação visando análise bibliométrica estão a Análise de Citações (*REFI* e *Cited/Citing*) – para rastreabilidade de citações realizadas e recebidas; Famílias de Patentes (INPADOC) – possibilita o agrupamento de patentes por famílias (simples e estendidas) evitando duplicações; Filtros de resultados por níveis de famílias ou publicações – ideal para identificar oportunidades de países onde a proteção não foi solicitada; Filtros estatísticos e gráficos; Classificação CPC – extensão da IPC; Busca por símbolos de classificação IPC; ou Monitoramento Legal (*Legal Events*) (*Espacenet*, 2022; Brasil *et al.*, 2018; INPI, 2022)

A base de dados *The Lens* possui uma infraestrutura cibernética global e aberta, criada pela organização sem fins lucrativos Cambia (com sede na Austrália). Diferente de muitas bases comerciais, a base *The Lens* posiciona-se como um "bem público digital", oferecendo acesso gratuito e integrado a um dos maiores acervos mundiais de patentes e literatura acadêmica. A

plataforma contempla mais de 250 milhões de trabalhos acadêmicos e mais de 120 milhões de documentos de patentes. Na patentometria, seu uso considera ferramentas que integram em uma única interface o universo de publicações científicas e de patentes. Dentre as suas principais funcionalidades, tem-se *PatCite* – possibilitando explorar citações entre patentes e artigos; *Visualização e Dashboards* – gerando gráficos dinâmicos para visualização instantânea de indicadores – bem como a opção de Dados Abertos e APIs. (Cambia, 2025)

### 3.2.3.1 Busca exploratória inicial

Apresentado as bases de dados que atendem aos critérios pré-estabelecidos, aplicou-se uma busca exploratória com uma amostra do conjunto de expressões para busca visando identificar o volume de dados recuperados em cada uma das bases (Tabela 1).

**Tabela 1** - Busca exploratória e resultados obtidos em diversas bases de dados

Grupo taxonômico	Parte correspondente da expressão	Artigos (tópico)		Patentes			
		Scopus	WoS	Scopus (texto completo)	Derwent (Tópico)	Espacenet (texto completo)	The Lens (texto completo)
Grupo 1	Parte 1	27	26	0	0	1	1
	Parte 2	39	34	0	0	2	2
	Parte 3	12	7	1	0	1	1
	Parte 4	21	18	1	0	1	1
	Parte 5	23	18	1	0	1	1
	Parte 6	12	13	0	0	0	0
	Parte 7	14	10	0	0	0	0
	Parte 8						
	Parte 9	35	27	1	0	3	1
	Parte 10						
	Parte 11	45	38	1	1	3	1
<b>TOTAIS</b>		<b>228</b>	<b>191</b>	<b>5</b>	<b>1</b>	<b>12</b>	<b>8</b>

Fonte: A autora (2024).

Nota: Coleta realizada no primeiro semestre de 2024.

Portanto, embasado na pesquisa exploratória com parte das expressões de busca do Grupo Taxonômico 1, a base de dados que retornou um maior número de artigos foi a selecionada (*Scopus*) e nas patentes (*Espacenet*) respectivamente, considerando que esta seria capaz de entregar uma quantidade maior de dados para análise.

Com base na Tabela 1 pode-se verificar que a base de dados *Scopus* recuperou um número de registros de artigos superior à base *Web of Science*, conforme colunas 3 e 4 respectivamente. Com relação às patentes, a base com maior número de registros recuperados é a *Espacenet*, e percebe-se que a base *Scopus* redireciona os registros de patentes recuperados para esta mesma base. Nota-se também que, embora a *Scopus* redirecione seus resultados para

a *Espacenet*, ainda assim o número de registros recuperados na *Scopus* é inferior ao número da *Espacenet* utilizando a mesma expressão de busca.

A escolha da base de dados *Scopus* como fonte primária para os registros científicos é justificada pelo maior volume de documentos recuperados na busca exploratória. Este volume superior é estratégico para a pesquisa, pois permite a exploração e o mapeamento de um cenário científico mais abrangente e geral sobre as plantas endêmicas da Amazônia, em consonância com o objetivo de construir uma sistemática para monitoramento. Além disso, ao priorizar uma base com maior cobertura, esta pesquisa se estabelece como um ponto de partida robusto para análises futuras, que poderão, posteriormente, expandir o corpus de dados com a inclusão de outras fontes (como a *Web of Science* e repositórios regionais), visando um resultado analítico ainda mais consolidado.

A busca por registros tecnológicos (patentes) exige uma base de dados que ofereça não apenas ampla cobertura internacional, mas também funcionalidades especializadas de análise de propriedade intelectual. Por essa razão, a base *Espacenet*, mantida pelo Escritório Europeu de Patentes (EPO), foi selecionada para a coleta dos registros patentométricos. Sua escolha se fundamenta em sua vasta cobertura geográfica e, principalmente, na disponibilização de metadados essenciais para estudos métricos, tais como: a Classificação Internacional de Patentes (IPC), dados de aplicantes e inventores, e a possibilidade de pesquisa em texto integral. Tais funcionalidades são cruciais para a aplicação dos indicadores compostos propostos, permitindo a análise da dinâmica tecnológica e a intersecção com classificações de sustentabilidade (como o *IPC Green Inventory*, detalhado na Seção 2.4.4.5).

### 3.2.3.2 Limitações da *Scopus*

A *Scopus* – base escolhida para a presente pesquisa – embora tenha uma das maiores coberturas internacionais, pode não conter todos periódicos de nicho ou regionais que não atendam seus critérios de seleção editorial. (Burnham, 2006)

Limitações por tipos de documentos podem ser também mencionados, já que determinados tipos de documentos podem ser deliberadamente excluídos de sua indexação, no qual poderia atender interesses de algumas áreas específicas (como Humanidades ou Ciências Sociais). Exemplo desses documentos são resenhas de livros (*book reviews*) – sob a justificativa de não serem fontes originais e raramente serem citadas; ou resumos de encontros (*conference meeting abstracts*) – pois consideram como conteúdo científico limitado e muitas vezes preliminar. (Elsevier, 2026)

Por fim, as limitações referentes à profundidade temporal das referências citadas. A cobertura de referências citadas pela *Scopus* se mostra inferior à da *WoS* para periódicos mais antigos. Embora a *Scopus* possua documentos partindo de 1788, sua rede mais completa de citações para análise bibliométrica são posterior à 1970, porquanto na *WoS* são posterior à 1900. Esta característica pode limitar estudos historiográficos, por exemplo. (Burnham, 2006; Elsevier, 2026)

### 3.2.3.3 Limitações da *Espacenet*

Referente às bases de dados para patentes, observa-se um diferencial no número de documentos recuperados com base no mecanismo de busca utilizado. Portanto, buscas aplicadas por tópico (título, resumo, referências, descrição, citações etc.) ou pelo texto completo retornam diferentes resultados. No caso de busca em todo o documento (como no caso da *Espacenet*, base de dados de patentes escolhida para a presente pesquisa) apresentaram por resultados plantas que não necessariamente são os requisitos de invenção, podendo ser assim mencionada em alguma referência ou como um exemplo de planta que possua atribuições similares à invenção que está de fato sendo requerida no documento da patente. Por outro lado, as patentes recuperadas por tópico indicam que a planta mencionada potencialmente possui maior relevância para a inovação ou tecnologia requerida. Ambos casos vão ao encontro da colocação de Santos, Silva e Faria (2023) que apontam o fato de uma planta ser mencionada em uma patente não necessariamente como objeto de reivindicação de invenção.

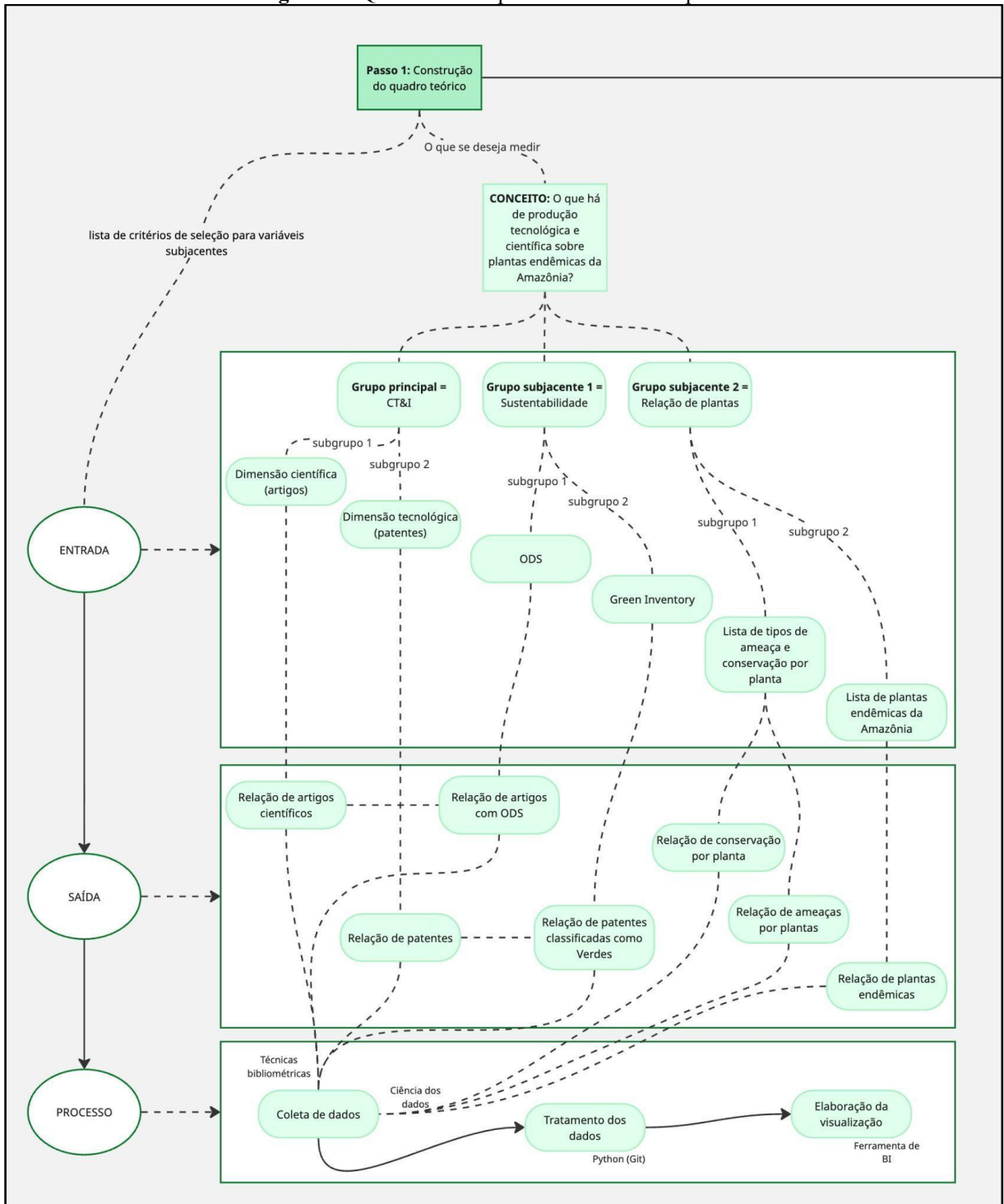
Outras limitações referem-se ao conteúdo e cobertura. Naturalmente, a base não abrange todo o estado da arte pesquisado apenas em patentes de um país. O número e qualidade dos campos acessíveis é outra ressalva, e devem ser consideradas em análises e conclusões de resultados de uma pesquisa (*Espacenet*, 2019)

Além disso, nem sempre um documento está completo. Ou seja, é possível que determinados registros bibliográficos de patentes estejam disponíveis em alguns países, entretanto o documento completo não. Limitações de idioma – priorizando o inglês – também é mencionado. Por fim, limitações técnicas e operacionais que abrangem restrições de número de arquivos recuperados, possibilidade de análises integradas, número de termos para busca, ou download de dados em massa (*Espacenet*, 2016, 2017, 2019)

### 3.2.4 Definição de indicadores

Para definição dos indicadores, optou-se previamente por seguir a recomendação do Manual para a Construção de Indicadores Compostos (OECD, 2008). Adotou-se, então, alguns dos passos sugeridos, sendo o primeiro para a construção do quadro teórico, onde deve-se basear no que é desejável medir – e não em quais indicadores estão disponíveis – considerando seus grupos, subgrupos, processo e variáveis afins. Contempla-se sua aplicação na Figura 4 à seguir.

**Figura 4 – Quadro teórico para indicadores compostos**

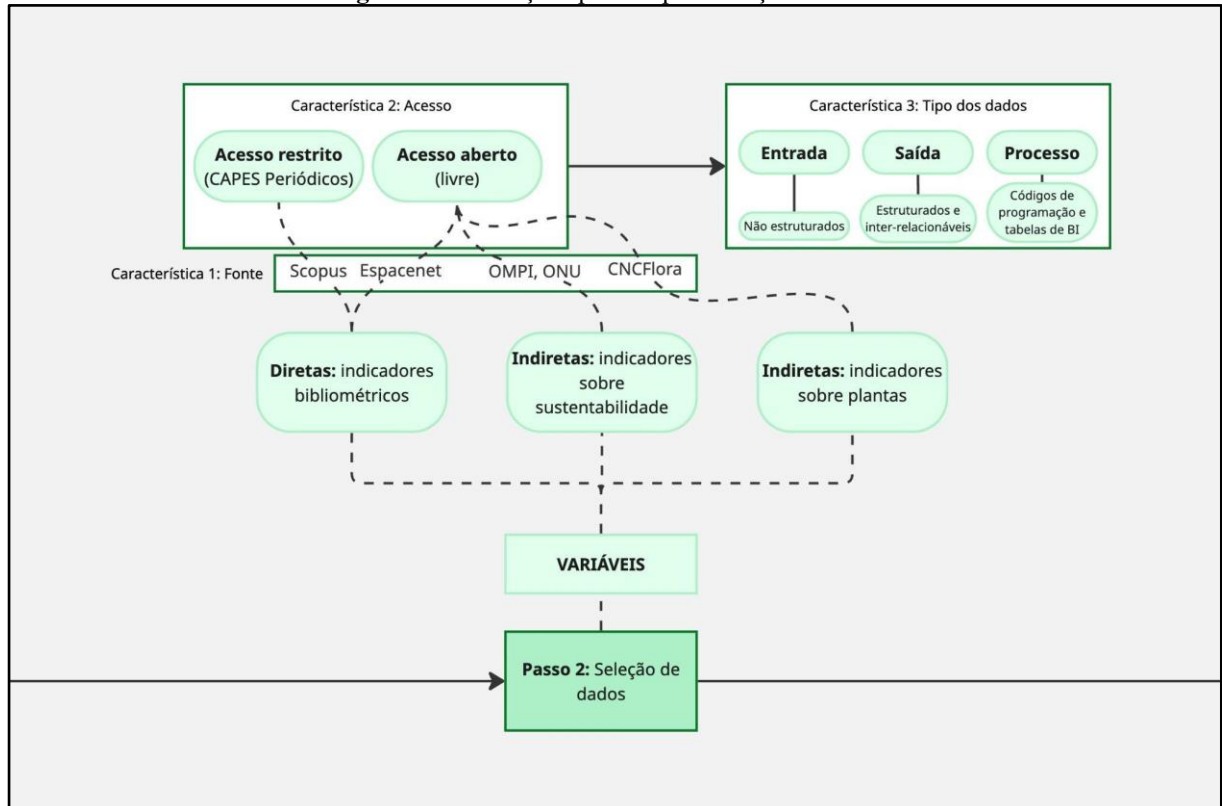


Fonte: A autora (2025).

Nota: Elaborado com base no manual OECD (2008), usando a ferramenta *Miro*.

E o segundo passo, que sugere elaborar as variáveis baseado na solidez analítica, mensurabilidade, cobertura e relevância dos indicadores para o fenômeno e a relação entre si. Assim, têm-se a Figura 5.

**Figura 5** – Definições prévias para seleção de dados



Fonte: A autora (2025).

Nota: Elaborado com base nas orientações do manual OECD (2008), usando a ferramenta *Miro*.

Com base na análise e na profunda contemplação do quadro teórico estabelecido, que serviu como pilar conceitual fundamental para a pesquisa, pôde-se não apenas compreender a complexidade do tema abordado, mas também elaborar e delinear a parte prática e processual do trabalho. Este processo de transposição do arcabouço teórico para a aplicação metodológica foi iniciado pela construção minuciosa e criteriosa dos indicadores.

Estes indicadores, essenciais para a mensuração e avaliação dos fenômenos estudados, foram desenvolvidos de forma a refletir as categorias e variáveis identificadas no referencial teórico. A sua formulação demandou um esforço de operacionalização conceitual, garantindo que fossem instrumentos válidos e confiáveis para a coleta e posterior análise dos dados. A etapa de construção dos indicadores configurou-se, portanto, como o primeiro passo crucial para a execução da pesquisa empírica e o avanço em direção aos objetivos propostos.

### 3.2.4.1 Indicadores diretos: bibliométricos de artigos e patentes

Posto definições prévias, com a fundamentação teórica e conceitual devidamente estabelecida, segue-se para a etapa crucial da pesquisa que consiste na delimitação específica e na operacionalização dos indicadores que serão utilizados posteriormente para a análise proposta. Este processo é essencial para garantir a precisão, a validade e a replicabilidade dos resultados. A identificação e a seleção criteriosa desses indicadores, apresentados detalhadamente no Quadro 8, refletem o esforço em traduzir os construtos abstratos definidos anteriormente em variáveis mensuráveis e passíveis de observação empírica, permitindo assim uma avaliação sistemática e rigorosa do objeto de estudo.

**Quadro 8** – Delimitação dos indicadores macro-bibliométricos descritivos de produção científica

Pergunta norteadora	Indicador	Variável	Objetivo
Quantos artigos versam sobre as plantas endêmicas da Amazônia Brasileira?	Número de publicações por tipos de documentos	Contagem de publicações de artigos (publicações científicas)	Obter medida quantitativa inicial do número de artigos dentro do escopo da pesquisa.
Quantos e quais são os autores e instituições (ou organizações) que publicaram sobre as plantas endêmicas da Amazônia Brasileira?	Número de publicações por tipos de autoria	Contagem de artigos por autor	Obter medida quantitativa inicial do número de autores que publicaram sobre a temática da pesquisa.
		Contagem de artigos por instituição	Obter medida quantitativa inicial do número de instituições que publicaram sobre a temática da pesquisa.
Quantos e quais são os países que publicam sobre as plantas endêmicas da Amazônia Brasileira?	Número de publicação por localização (distribuição geográfica)	Contagem de artigos por geolocalização (país/região)	Obter medida quantitativa inicial do número de publicações por país/região para comparativo entre cenário nacional (Brasil) e internacional.
Qual a distribuição das publicações ao longo dos anos sobre as plantas endêmicas da Amazônia Brasileira?	Número de publicação por ano	Contagem de artigos por ano de publicação	Obter medida quantitativa inicial do número de publicações por ano para comparativo ao longo do tempo.

Fonte: A autora (2025).

Nota: Adaptado de Okubo (1997); Prado e Castanha (2020); e de Sousa, Almeida e Bezerra (2024).

Neste sentido, a presente análise se aprofunda na delimitação dos indicadores macro-bibliométricos que serão empregados para a descrição detalhada da produção tecnológica. Tais indicadores, apresentados sinteticamente na Quadro 9 (detalhada posteriormente com suas

respectivas métricas e fontes), foram criteriosamente selecionados por sua capacidade de oferecer uma visão abrangente e descritiva dos resultados de P&D (Pesquisa e Desenvolvimento), atuando como um panorama quantitativo da atividade inventiva e inovadora. A utilização desses indicadores é fundamental para mapear a evolução, o volume e a distribuição geográfica e institucional da tecnologia analisada, permitindo assim uma compreensão robusta do ecossistema de inovação e dos agentes que o compõem.

**Quadro 9** – Delimitação dos indicadores macro-bibliométricos descritivos de produção tecnológica

Pergunta norteadora	Indicador	Variável	Objetivo
Quantas patentes versam sobre as plantas endêmicas da Amazônia Brasileira?	Número de publicações por tipos de documentos	Contagem de publicações de patentes	Obter medida quantitativa inicial do número de patentes dentro do escopo da pesquisa
Quanto e quais são os inventores e proprietários (aplicantes) que estão relacionados às patentes que versam sobre as plantas endêmicas da Amazônia Brasileira?	Número de publicações por tipos de autoria	Contagem de artigos por inventor	Obter medida quantitativa inicial do número de inventores das patentes
		Contagem de artigos por aplicante	Obter medida quantitativa inicial do número de aplicantes que detêm prioridade da inovação patenteadas
Quanto e quais são os países com patentes aplicadas ou criadas sobre as plantas endêmicas da Amazônia Brasileira?	Número de publicação por localização (distribuição geográfica)	Contagem de patentes por geolocalização (país/região)	Obter medida quantitativa inicial do número de publicações por país/região para comparativo entre cenário nacional (Brasil) e internacional, e aspectos de: <ul style="list-style-type: none"> <li>• difusão global da tecnologia;</li> <li>• propriedade transfronteiriça (onde está o aplicante vs. inventor) para entender fluxos de conhecimento global.</li> </ul>
Qual a distribuição das patentes ao longo dos anos sobre as plantas endêmicas da Amazônia Brasileira?	Número de publicação por ano (série temporal)	Contagem de patentes por ano de publicação	Obter medida quantitativa inicial do número de publicações por ano de prioridade e de publicação.
Qual a distribuição das patentes por áreas tecnológicas sobre as plantas endêmicas da Amazônia Brasileira?	Número de publicação por área tecnológica (IPC)	Contagem de patentes por IPC	Obter medida quantitativa inicial do número de publicações por classificação/especialização tecnológica

Fonte: A autora (2025).

Nota: Adaptado de Guellec e Potterie, 2005; Okubo (1997); Prado e Castanha (2020); e de Sousa, Almeida e Bezerra (2024).

Em uma perspectiva geral, os indicadores de Ciência, Tecnologia e Inovação (CT&I) não se limitam a ser meras estatísticas; eles visam, fundamentalmente, proporcionar uma visão panorâmica e compreensiva do cenário da CT&I em um determinado contexto, seja ele

nacional, setorial ou regional. Essa visão abrange aspectos cruciais como o volume de investimentos, a produção científica, a formação de recursos humanos qualificados, a capacidade de absorção e difusão de tecnologias e o impacto da inovação na economia e sociedade (Prado; Castanha, 2020)

#### 3.2.4.2 Indicadores indiretos sobre plantas e sobre sustentabilidade

A definição dos indicadores indiretos constituiu-se como uma etapa crucial e metodologicamente orientada no desenvolvimento desta pesquisa. Para a sua delimitação, foi imperativo estabelecer uma clara conexão com o objeto de medição pretendido, buscando na Ciência da Informação (CI) e na Ciência de Dados (CD) os métodos, abordagens e ferramentas capazes de nortear e efetivar a integração entre as diferentes naturezas de indicadores.

Tal qual se procedeu com a construção dos indicadores diretos, a estrutura dos indicadores indiretos também se baseia em uma pergunta norteadora (o que se deseja medir). Essa pergunta fundamentalmente define o escopo e a relevância de cada indicador.

A seguir, a construção e a apresentação desses indicadores seguem uma estrutura lógica e detalhada, conforme explicitado no Quadro 10:

1. Pergunta Norteadora: Localizada na primeira coluna da estrutura analítica, esta seção expressa de forma concisa o objetivo específico de medição.
2. Indicador: Esta coluna nomeia o indicador propriamente dito e, crucialmente, visa indicar onde este deve ser aplicado em conjunto com o indicador direto correspondente (geralmente de natureza bibliométrica). A combinação destes dois tipos de indicadores é o cerne da proposta de integração desta pesquisa.
3. Descrição da Definição: Na terceira coluna, é apresentada uma descrição concisa, mas essencialmente informativa, sobre o processo de definição do indicador. Detalha-se qual é o tipo de métrica que compõe este indicador indireto e, mais especificamente, qual relação ou fórmula mais específica foi utilizada para sua construção. Esta descrição estabelece a base empírica e conceitual da métrica.
4. Justificativa do Objetivo: Por fim, a última coluna dedica-se a justificar a pertinência e a relevância deste indicador indireto para os objetivos mais amplos da presente investigação. Explica-se de que forma a mensuração proposta contribui para a compreensão e análise dos fenômenos estudados, validando sua inclusão no modelo de avaliação.

Essa abordagem sistemática garante que os indicadores indiretos não sejam métricas isoladas, mas sim ferramentas construídas com rigor metodológico, capazes de fornecer uma perspectiva complementar e essencialmente integrada à análise dos dados bibliométricos diretos, avançando assim na capacidade de mensuração proposta por este trabalho.

**Quadro 10** – Delimitação dos indicadores indiretos sobre as plantas e sustentabilidades

Pergunta norteadora	Indicador	Como foi feito	Objetivo
Quantas e quais são as plantas endêmicas da Amazônia pesquisadas ou mencionadas em patentes?	Número de publicações (artigos e patentes) sobre plantas endêmicas	Contagem de artigos e patentes por nomenclatura binominal científica em latim da planta, localizada por busca dentro dos campos: título, resumo, palavra-chave ou resumo.	Obter medida quantitativa inicial do número de publicações sobre cada planta, e identificar as plantas com maiores ocorrências nas publicações.
Quantas e quais são as relações de publicações pela Lista Vermelha do CNCFlora? Estão ameaçadas? Estão em áreas protegidas?	Número de publicações (artigos e patentes) de acordo com a Classificação de Ameaça – Lista Vermelha	Contagem de artigos e patentes por nomenclatura binominal científica em latim da planta, localizada por busca dentro dos campos: título, resumo, palavra-chave ou resumo; cruzando-as com a Lista Vermelha do CNCFlora.	Obter medida quantitativa inicial do número de publicações sobre cada planta, e identificar as plantas com maiores ocorrências nas publicações de acordo com sua categoria de ameaça.
	Número de publicações (artigos e patentes) de acordo com a Classificação de Tipo de Proteção	Contagem de artigos e patentes por nomenclatura binominal científica em latim da planta, localizada por busca dentro dos campos: título, resumo, palavra-chave ou resumo; cruzando-as com a Lista de Classificação de Tipo de Proteção do CNCFlora	Obter medida quantitativa inicial do número de publicações sobre cada planta, e identificar as plantas com maiores ocorrências nas publicações de acordo com sua classificação de proteção.
Quantas publicações versam sobre sustentabilidade, e quais são seus focos principais?	Número de artigos com ODS	Contagem de artigos recuperados de acordo com a combinação de expressão de busca das plantas com a expressão de busca por cada ODS	Obter a medida quantitativa inicial do número de artigos que versam sobre sustentabilidade e estabelecer a relação das plantas com os ODS.
	Número de patentes com a classificação <i>Green Inventory</i>	Contagem de patentes cuja classificação internacional (IPC) esteja classificado como <i>Green Inventory</i>	Obter a medida quantitativa inicial do número de patentes com viés de sustentabilidade e estabelecer a relação das plantas com os tipos de sustentabilidade realizados.

Fonte: A autora (2025).

Este processo, orientado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), não apenas facilita a consolidação das informações e dos resultados

obtidos, mas também assegura a replicabilidade das metodologias e das boas práticas para outras jurisdições ou contextos. A adoção das diretrizes da OCDE, reconhecida globalmente por seu rigor metodológico e seu foco em políticas públicas baseadas em evidências, confere maior credibilidade e robustez ao arcabouço conceitual e operacional em questão. Isso permite que os *insights* gerados sejam aplicáveis e adaptáveis, promovendo a convergência de padrões internacionais e a otimização de recursos em escala global ou regional. A aderência a este padrão estabelecido pela OCDE é crucial para garantir a transparência, a comparabilidade dos dados e a eficácia das intervenções subsequentes, consolidando, assim, uma base sólida para a tomada de decisões estratégicas e para o avanço das políticas relacionadas ao tema central da pesquisa ou da iniciativa.

### 3.2.4.3 Relação final dos indicadores compostos

Os indicadores de pesquisa englobam as dimensões de Ciência, Tecnologia e Inovação (CT&I), juntamente com os indicadores de sustentabilidade. Adicionalmente, foi estabelecida a relação entre esses indicadores e a amostra desejada para análise dessa relação. Para facilitar o entendimento, essa informação está sistematizada em uma matriz apresentada no Quadro 11 a seguir.

**Quadro 11** - Visão geral dos indicadores simples e compostos finais

Indicador bibliométrico	Indicador classificatório de sustentabilidade	Indicadores simples	Indicadores compostos
Produção científica > Artigos científicos	- ODS - Lista Vermelha de Espécies Ameaçadas e Classificação de Conservação	<ul style="list-style-type: none"> <li>• Artigos que mencionam espécies endêmicas da Amazônia;</li> </ul>	<ul style="list-style-type: none"> <li>• Artigos que mencionam espécies endêmicas da Amazônia alinhados aos ODS;</li> <li>• Artigos que mencionam espécies endêmicas da Amazônia e respectivos tipos de ameaça das espécies listadas;</li> <li>• Artigos que mencionam espécies endêmicas da Amazônia e respectivas classificações de ações de conservação das espécies listadas;</li> </ul>
Propriedade intelectual / industrial > Patentes	- <i>Green Inventory</i> IPC - Lista Vermelha de Espécies Ameaçadas e Classificação de Conservação	<ul style="list-style-type: none"> <li>• A. Patentes que mencionam espécies endêmicas da Amazônia;</li> </ul>	<ul style="list-style-type: none"> <li>• Patentes que mencionam espécies endêmicas da Amazônia alinhadas ao conceito de patente verde pelo código <i>Green Inventory</i> IPC;</li> <li>• Patentes que mencionam espécies endêmicas da Amazônia e respectivos tipos de ameaça das espécies listadas;</li> </ul>

Indicador bibliométrico	Indicador classificatório de sustentabilidade	Indicadores simples	Indicadores compostos
			<ul style="list-style-type: none"> <li>Patentes que mencionam espécies endêmicas da Amazônia e respectivas classificações de ações de conservação das espécies listadas;</li> </ul>

Fonte: A autora (2025).

A metodologia para identificar a intersecção entre os dados contidos nos artigos científicos revisados e os Objetivos de Desenvolvimento Sustentável (ODS) da Agenda 2030 baseou-se em um procedimento rigoroso de busca e filtragem. Primeiramente, foi estabelecido um conjunto de expressões de busca principais, focado no tema central da pesquisa, que neste caso se concentrou em "plantas endêmicas". Este termo serviu como ponto de partida para a recuperação de todos os artigos relevantes.

Em uma etapa subsequente e crucial, para correlacionar o conteúdo dos artigos recuperados com cada um dos 17 ODS individualmente, foi utilizado um conjunto de expressões-chave específicas e distintas para cada ODS. Este conjunto de termos foi cuidadosamente selecionado para abranger os conceitos, metas e indicadores primários associados a cada objetivo (como, por exemplo, "erradicação da pobreza" para ODS 1, "fome zero" para ODS 2, "saúde e bem-estar" para ODS 3, e assim por diante até "vida terrestre" para ODS 15 e "parcerias" para ODS 17).

A intersecção de dados foi efetivamente alcançada por meio da combinação lógica dessas expressões principais (relacionadas a "plantas endêmicas") com o conjunto de expressões específicas de cada ODS. Essa combinação permitiu mapear quais artigos, que tratavam de plantas endêmicas, também abordavam, de forma direta ou indireta, as temáticas específicas de cada ODS. O procedimento detalhado, incluindo a lista completa e exaustiva das expressões de busca utilizadas para cada ODS, está integralmente documentado e apresentado de forma estruturada no Quadro 12 deste trabalho. A consulta a este Quadro é essencial para a completa compreensão da abrangência e precisão da metodologia empregada na associação dos dados.

**Quadro 12** - Exemplo de estratégia de busca para mapear ODS (Grupo 3 - ODS 11)

Expressão principal	Expressão ODS	Expressão final
TITLE-ABS KEY ("Taxithelium juruense" OR "Sphagnum amazonicum" OR "Sphagnum curicuriariense" OR "Sphagnum dimorphophyllum")	TITLE-ABS-KEY ( ( city OR cities OR {human settlement} OR {human settlements} OR urban OR metropoli* OR town* OR municipal* ) AND ( gentrification OR congestion OR transportation OR {public transport} OR housing OR slum* OR {sendai framework} OR {Disaster Risk Reduction} OR {DRR} OR {smart city} OR {smart cities} OR {resilient building} OR {resilient buildings} OR {sustainable building} OR {sustainable buildings} OR {building design} OR {buildings design} OR urbani?ation OR {zero energy building} OR {zero energy buildings} OR {zero-energy building} OR {zero-energy buildings} OR {basic service} OR {basic services} OR {governance} OR {citizen participation} OR {collaborative planning} OR {participatory planning} OR {inclusiveness} OR {cultural heritage} OR {natural heritage} OR {UNESCO} OR {disaster} OR {ecological footprint} OR {environmental footprint} OR {waste} OR {pollution} OR {pollutant*} OR {waste water} OR {recycling} OR {circular economy} OR {air quality} OR {green space} OR {green spaces} OR {nature inclusive} OR {nature inclusive building} OR {nature inclusive buildings} ) )	((TITLE-ABS-KEY ("Taxithelium juruense" OR "Sphagnum amazonicum" OR "Sphagnum curicuriariense" OR "Sphagnum dimorphophyllum")) AND ((TITLE-ABS-KEY ( ( city OR cities OR {human settlement} OR {human settlements} OR urban OR metropoli* OR town* OR municipal* ) AND ( gentrification OR congestion OR transportation OR {public transport} OR housing OR slum* OR {sendai framework} OR {Disaster Risk Reduction} OR {DRR} OR {smart city} OR {smart cities} OR {resilient building} OR {resilient buildings} OR {sustainable building} OR {sustainable buildings} OR {building design} OR {buildings design} OR urbani?ation OR {zero energy building} OR {zero energy buildings} OR {zero-energy building} OR {zero-energy buildings} OR {basic service} OR {basic services} OR {governance} OR {citizen participation} OR {collaborative planning} OR {participatory planning} OR {inclusiveness} OR {cultural heritage} OR {natural heritage} OR {UNESCO} OR {disaster} OR {ecological footprint} OR {environmental footprint} OR {waste} OR {pollution} OR {pollutant*} OR {waste water} OR {recycling} OR {circular economy} OR {air quality} OR {green space} OR {green spaces} OR {nature inclusive} OR {nature inclusive building} OR {nature inclusive buildings} ) )

Fonte: A autora (2024).

Nota: Elaborado no primeiro semestre de 2024.

Ainda no que tange à análise dos artigos científicos revisados, a construção dos indicadores propostos para este estudo foi estabelecida a partir da composição de duas variáveis cruciais: a relação de classificação das espécies vegetais por tipos de ameaças e os tipos de classificação de proteção aplicados a elas.

Esta relação não foi estabelecida de maneira arbitrária; ela exigiu um procedimento de cruzamento de dados rigoroso. Especificamente, o nome da planta citado em cada artigo (empregando-se a nomenclatura binominal em latim, garantindo a precisão e universalidade da identificação taxonômica) foi sistematicamente cruzado com suas classificações oficiais presentes na Lista Vermelha do CNCFlora (Centro Nacional de Conservação da Flora).

O CNCFlora é o órgão de referência nacional para a avaliação do estado de conservação da flora brasileira, seguindo os critérios e categorias da IUCN (União Internacional para a Conservação da Natureza). Portanto, esse cruzamento permitiu associar a menção da espécie no artigo (que pode tratar da sua ocorrência, uso ou ecologia) diretamente ao seu status de conservação mais atual e reconhecido (como "Em Perigo" (EN), "Vulnerável" (VU), etc.) e,

em seguida, verificar as medidas ou classificações de proteção (como "espécie-chave", "prioridade de conservação", inclusão em Unidades de Conservação, etc.) mencionadas pelos autores ou inferidas pela legislação aplicável. Essa metodologia garante que os indicadores sejam robustos, baseados em dados taxonômicos precisos e em classificações oficiais de ameaça, essenciais para uma análise aprofundada das estratégias de conservação.

Por fim, foram desenvolvidos e analisados indicadores compostos especificamente desenhados para o mapeamento de patentes com um viés explícito em sustentabilidade, focando em duas abordagens complementares.

Primeiramente, foi utilizado a relação do *Green Inventory* proposto pela Organização Mundial da Propriedade Intelectual (WIPO). Este inventário categoriza tecnologias que contribuem para a mitigação dos impactos das mudanças climáticas e para o desenvolvimento sustentável, permitindo a identificação de patentes que explicitamente se alinham com tecnologias "verdes". A aplicação do *Green Inventory* fornecerá uma base sólida e internacionalmente reconhecida para a classificação inicial das patentes.

Em seguida, foi empregada uma metodologia adaptada baseada na listagem do Centro Nacional de Conservação da Flora (CNCFlora). Embora esta lista seja frequentemente utilizada em artigos para mapear a ocorrência e conservação de espécies da flora brasileira, a adaptação proposta será a de cruzar esta listagem não com a localização geográfica ou a ocorrência das espécies per se, mas sim com a planta especificamente mencionada ou reivindicada no conteúdo das patentes. Essa abordagem visa identificar patentes que utilizam ou se baseiam em recursos da biodiversidade brasileira, particularmente aqueles sob algum nível de atenção ou risco, conforme classificado pelo CNCFlora, adicionando uma camada de análise de uso da biodiversidade no contexto da sustentabilidade e da bioprospecção.

### **3.2.5 Escolha da ferramenta de visualização: *Looker Studio***

O *Looker Studio* (LS) é uma ferramenta de *Business Intelligence* (BI) utilizada para transformar grandes volumes de dados brutos e planilhas em informações estratégicas e visuais, como para o caso de tomada de decisão por gestores públicos. Seu uso para criação de painéis dinâmicos e intuitivos permite elaborar gráficos de distribuição por áreas do conhecimento e mapas de localização. Sua gratuidade torna a ferramenta acessível, eficiente e intuitiva, mesmo para iniciantes (ou seja, reduz a demanda de curva de aprendizagem). (Dutra, 2024)

Esta ferramenta pode ser acessada por qualquer usuário com acesso à web e não precisa

ser cadastrado em nenhuma conta. Para criar e editar um painel, entretanto, precisa-se dos seguintes requisitos: estar conectado à uma conta do *Google*; estar em um dos países em que há suporte; e aceitar os Termos de Serviço e as políticas do LS. Esta ferramenta está disponível em 37 idiomas com formatos de data e hora de acordo com esses idiomas, no qual é incluído o português brasileiro. Seu acesso é gratuito, mas há também a disponibilidade de recursos avançados pagos por meio da assinatura paga<sup>14</sup> (*Google*, 2026)

Existem outras ferramentas de BI gratuitas, como a *Microsoft Power BI*, ou *Tableau Public*. Contudo, o LS contempla mais pontos positivos em termos de interface *user-friendly* já familiar aos usuários de produtos *Google* e pela facilidade de compartilhamento via link sem custos. (Patel, 2023)

Outro aspecto positivo para escolha do LS é a possibilidade de integração com contas "*Gmail*". O *Looker Studio* se mostra eficiente em diversos aspectos, entretanto possui limitações importantes de ressalva. A primeira limitação é que este depende fortemente de uma etapa prévia de tratamento de dados (como por *Extract, Transform, Load – ETL*) externa à ela. Ou seja, suas capacidades de limpeza nativa são limitadas. Além disso, trata-se de uma ferramenta proprietária, portanto é de suma importância que os dados originais estejam estruturados externamente ao LS, a fim de mitigar eventuais riscos (Dutra, 2024)

Considerando as características apontadas acima, o Quadro 13 congrega em uma matriz os pontos fortes, limitações e riscos e mitigações aplicáveis ao LS.

**Quadro 13** – Matriz de Análise Integrada: *Google Looker Studio & Workspace*

Dimensão	Pontos fortes (vantagens)	Limitações (técnicas e institucionais)	Riscos e Mitigações
<b>Custo e Acesso</b> (Conta pública sem vínculo direto à UFSCar)	<b>Gratuidade Total:</b> A ferramenta oferece seus recursos de forma gratuita, permitindo viabilidade econômica para projetos públicos e acadêmicos — <b>Compartilhamento Aberto:</b> Diferente do <i>Microsoft Power BI</i> (versão <i>Free</i> ), o <i>Looker</i> permite publicar relatórios via <i>link web</i> acessível a qualquer cidadão sem necessidade de <i>login</i> ou licença corporativa.	<b>Dependência de Conta <i>Google</i>:</b> O acesso para criação e gestão exige uma conta <i>Google</i> , embora a visualização possa ser pública.	<b>Risco:</b> Alteração na política de gratuidade da ferramenta proprietária. — <b>Mitigação:</b> Como o <i>Looker</i> é uma ferramenta " <i>web</i> " de baixo custo de entrada, a migração é menos custosa do que em <i>softwares</i> pagos. A gratuidade garante a sustentabilidade imediata do projeto sem onerar a universidade.
<b>Custo e Acesso</b>	<b>Gratuidade Garantida:</b> A UFSCar utiliza a edição	<b>Dependência de Vínculo:</b> O acesso pleno	<b>Risco:</b> Mudança na política de preços do fornecedor

<sup>14</sup> Mais informações em: <https://support.google.com/looker-studio/?hl=pt-BR#topic=6267740>. Acesso em 27 jan. 2025.

Dimensão	Pontos fortes (vantagens)	Limitações (técnicas e institucionais)	Riscos e Mitigações
(Conta vinculada à UFSCar)	<p><i>Education Fundamentals</i>, que é gratuita para instituições qualificadas</p> <p>—</p> <p><b>Acesso Universal:</b> Disponível para toda a comunidade (discentes, centros acadêmicos), eliminando custos de licença individual.</p>	<p>depende da manutenção do vínculo ativo com a UFSCar. Contas inativas (&gt;180 dias) podem ser excluídas, exigindo gestão constante.</p>	<p>proprietário.</p> <p>—</p> <p><b>Mitigação:</b> O contrato <i>Education Fundamentals</i> garante a gratuidade dos serviços essenciais (<i>Core Services</i>) sem anúncios, mitigando o risco de cobrança súbita para funcionalidades básicas.</p>
<p><b>Infraestrutura e Nuvem</b></p> <p>(Conta vinculada à UFSCar)</p>	<p><b>SaaS (<i>Software as a Service</i>):</b> Ferramentas 100% em nuvem, sem necessidade de instalação local ou servidores próprios da universidade.</p> <p>—</p> <p><b>Ecosistema Integrado:</b> O <i>Looker</i> conecta-se nativamente ao <i>Google Sheets</i> e <i>Drive</i> já usados na UFSCar, facilitando a colaboração em tempo real.</p>	<p><b>Cotas de Armazenamento:</b> A UFSCar implementou limites de cota (ex: 10GB para contas temporárias/eventos) devido às novas políticas do <i>Google</i>, o que exige "limpeza" frequente de dados brutos pesados.</p>	<p><b>Risco:</b> Perda de dados por falta de espaço ou encerramento de conta.</p> <p>—</p> <p><b>Mitigação:</b> A política de retenção da UFSCar é clara. Como mitigação técnica, recomenda-se manter <i>backups</i> dos dados brutos (CSV/Excel) em repositórios locais ou institucionais alternativos, não confiando apenas na nuvem para arquivamento morto.</p>
<p><b>Privacidade e Propriedade de Dados</b></p>	<p><b>Propriedade Institucional:</b> O <i>Google</i> declara explicitamente que "as escolas são donas de seus dados" e retêm direitos de propriedade intelectual. Os dados do <i>Core Services</i> (<i>Drive, Gmail, Docs</i>) não são usados para publicidade.</p> <p>—</p> <p><b>Conformidade:</b> Adesão à LGPD/GDPR e padrões globais de segurança.</p>	<p><b>Auditoria Externa:</b> Sendo software proprietário ("caixa preta"), a instituição depende das auditorias de terceiros contratadas pelo <i>Google</i> para verificar a segurança, sem acesso direto ao código.</p>	<p><b>Risco:</b> Uso indevido de dados de pesquisa por empresas privadas.</p> <p>—</p> <p><b>Mitigação:</b> A documentação oficial assegura contratualmente que os dados inseridos no Workspace for Education são isolados de mineração para anúncios. Além disso, o tratamento prévio dos dados (ETL externo) garante que apenas os dados finais de visualização subam para a nuvem.</p>
<p><b>Usabilidade e Curva de Aprendizado</b></p>	<p><b>Interface Familiar:</b> A interface segue o padrão do ecossistema <i>Google</i> (<i>Docs, Sheets</i>), sendo intuitiva e <i>user-friendly</i>, o que reduz o tempo de treinamento.</p> <p>—</p> <p><b>Suporte Institucional:</b> A SIn/UFSCar oferece tutoriais e suporte para o ambiente <i>Workspace</i>, reduzindo a curva de aprendizado da equipe.</p> <p>—</p> <p><b>Baseada na Nuvem:</b> Não requer instalação de <i>software</i></p>	<p><b>Recursos simplificados:</b> Possui menos opções de customização complexa e ferramentas de análise profunda se comparado a concorrentes como <i>Tableau</i> ou <i>Power BI</i>.</p> <p>—</p> <p><b>Recursos "Core" vs. Pagos:</b> Algumas funcionalidades avançadas de IA (<i>Gemini</i>) ou segurança (investigação avançada) podem estar restritas às</p>	<p><b>Risco 1:</b> Obsolescência da ferramenta gratuita.</p> <p>—</p> <p><b>Mitigação 1:</b> Uso de ferramentas com ampla base de suporte na comunidade (fóruns, tutoriais) e alta interoperabilidade (exportação para PDF/Excel), garantindo que os relatórios não fiquem "presos" na plataforma.</p> <p>—</p> <p><b>Risco 2:</b> Falta de recursos nativos para análises estatísticas complexas.</p>

Dimensão	Pontos fortes (vantagens)	Limitações (técnicas e institucionais)	Riscos e Mitigações
	(SaaS), facilitando o uso em qualquer computador institucional.	versões pagas ( <i>Education Plus</i> ), que a universidade pode não assinar.	— <b>Mitigação 2:</b> Utilizar o <i>Looker</i> apenas como "camada de apresentação". Toda a análise pesada deve ser feita previamente nas ferramentas de base ( <i>Python</i> ) antes da visualização
<b>Gestão e Processamento de Dados (ETL)</b>	<b>Integração Nativa:</b> Conecta-se facilmente com <i>Google Sheets</i> e arquivos <i>Excel</i> , comuns no ambiente acadêmico . — <b>Dinamismo:</b> Permite filtros interativos e atualização em tempo real para o usuário final.	<b>Baixa capacidade de ETL:</b> A ferramenta não é adequada para limpeza e tratamento de dados brutos ("sujos"); o tratamento deve ser feito antes da importação. — <b>Performance:</b> Pode apresentar lentidão ou erros com grandes volumes de dados não tratados	<b>Risco (Proprietário):</b> "Lock-in" dos dados dentro da plataforma. — <b>Mitigação (Crucial):</b> O modelo de funcionamento exige que os dados sejam tratados externamente ( <i>Python</i> ). Isso garante que os dados (ativo principal) permaneçam sob posse do pesquisador em arquivos locais ou institucionais, e não "presos" dentro do LS.

Fonte: A autora (2025).

Nota: Adaptado de Dutra, 2024; *Google*, 2026; Patel, 2023; Site da SIn e UFSCar, 2024.

Contudo, o LS mostra-se uma ferramenta capaz de atender critérios de usabilidade, infraestrutura, custo e acesso para a presente pesquisa. Entende-se que há limitações na ferramenta, principalmente referente ao fato de ser proprietária da *Google*, embora medidas de mitigação foram previstas para tal. Futuros trabalhos poderão adotar outras ferramentas abertas e gratuitas, até mesmo em efeito comparativo.

Neste contexto, a escolha estratégica por um painel (*dashboard*) dinâmico e interativo (alinhado ao Objetivo Específico - OE 4) representa uma solução robusta e altamente eficaz de Inteligência Competitiva (IC). Este tipo de painel não apenas organiza grandes volumes de dados de maneira superior, mas também viabiliza o monitoramento periódico e proativo (abordando o Desafio 1.1.3) dos desenvolvimentos e tendências emergentes em Ciência, Tecnologia e Inovação (CT&I). O dinamismo e a interatividade são a chave para transformar dados brutos em *insights* acionáveis e estratégicos.

Para maximizar a eficácia de um *dashboard* de indicadores compostos (OECD, 2008), que demanda clareza devido à sua complexidade, é crucial seguir princípios de design e usabilidade (Few, 2006). Um dos pilares é a Relevância (Foco e Priorização), onde o design deve ser hierárquico, focando em priorizar os ICs no topo da tela (above the fold) para guiar a atenção, e a informação deve ser sumarizada para uma visão geral estratégica e rápida. Outro

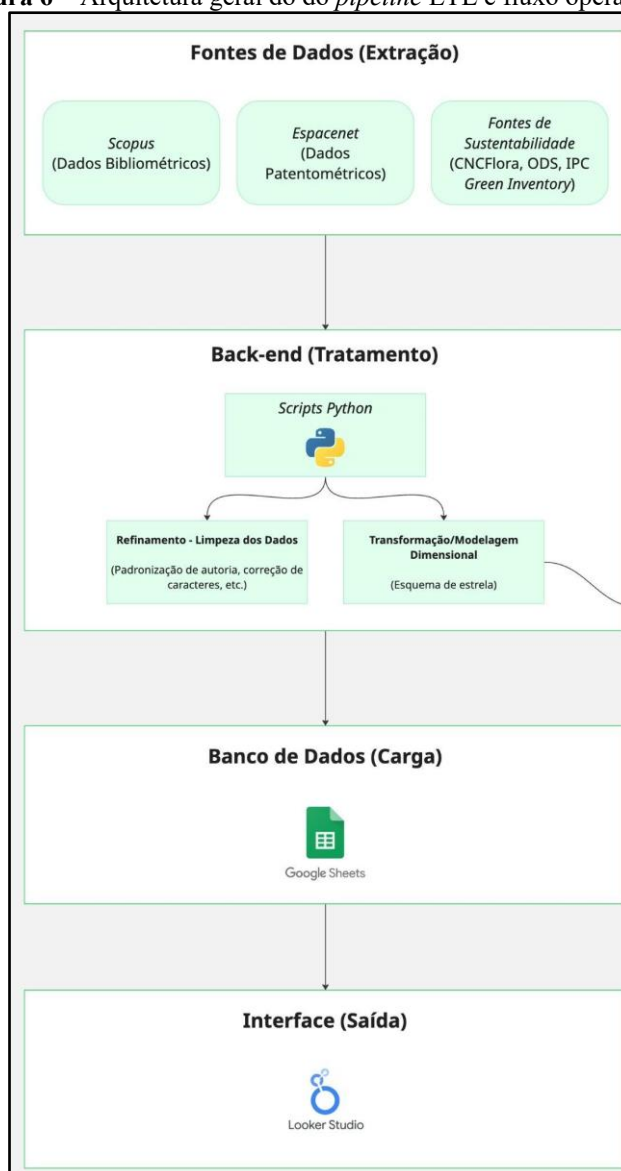
ponto fundamental é a Transparência (Confiança e Acessibilidade), visto que a confiança nos ICs exige transparência.

O sistema deve garantir que o dado bruto seja sempre rastreável e auditável, e funcionalidades como *drill-down*, filtros e tooltips interativos são essenciais para inspecionar a origem e o cálculo. A Comparabilidade (Contextualização e Evolução Temporal) também é chave, pois a evolução e contextualização dos dados superam o valor isolado; o painel deve usar séries temporais, filtros de período flexíveis e *benchmarking* para identificar padrões e avaliar o desempenho ao longo do tempo. Por fim, o Alinhamento Cognitivo (Usabilidade e Mapeamento Mental) garante que a interação espelhe a lógica de intersecção e o modelo de dados subjacente, o que significa que a funcionalidade, como a filtragem por um Objetivo de Desenvolvimento Sustentável (ODS), deve logicamente filtrar indicadores associados, reduzindo a curva de aprendizado e potencializando a capacidade analítica.

### 3.3 FASE 2: EXECUÇÃO DO PROTOCOLO E *PIPELINE* ETL DE MODELAGEM DE DADOS

Esta seção detalha o procedimento técnico de execução do protocolo de coleta e o subsequente processamento dos dados. Para garantir a reprodutibilidade e a transparência metodológica, todo o processo foi concebido e será descrito com base na arquitetura de dados adotada. O desafio de integrar e harmonizar dados heterogêneos provenientes de múltiplas fontes motivou a implementação de *pipelines* de dados automatizados, seguindo o modelo de Extração, Transformação e Carga (ETL – Extract, Transform, Load), conforme Figura 6.

**Figura 6** – Arquitetura geral do do *pipeline* ETL e fluxo operacional



Fonte: A autora (2025).

Nota: Elaborado utilizando a ferramenta *Miro*.

Este modelo não apenas garante a eficiência do processo, mas também é fundamental para a criação dos indicadores compostos. O *pipeline* de desenvolvimento e a infraestrutura tecnológica desta pesquisa foram criteriosamente concebidos para garantir eficiência, reprodutibilidade e facilidade de integração de dados. A seguir, detalhamos as ferramentas e tecnologias que compuseram este ambiente:

- linguagem de programação python (versão 3.12): foi escolhida como a linguagem principal devido à sua vasta aplicabilidade em ciência de dados, robustez e ao rico ecossistema de bibliotecas que suporta o desenvolvimento rápido e a prototipagem. A versão 3.12 assegura o uso de features e otimizações recentes da linguagem;

- b) bibliotecas e ecossistema de dados:
- i) *pandas*: essencial para todas as etapas de manipulação, limpeza e análise de dados. Sua estrutura de *DataFrames* permitiu o carregamento eficiente, a transformação e a estruturação dos conjuntos de dados utilizados no estudo;
  - ii) *Gspread* e *Gspread-dataframe*: estas bibliotecas foram cruciais para estabelecer uma integração fluida e programática com a API do *Google Sheets*. Essa integração possibilitou a leitura e a escrita direta de dados em planilhas colaborativas, facilitando a coleta de dados e a comunicação de resultados intermediários;
  - iii) outras dependências: um conjunto de outras bibliotecas e módulos auxiliares foi empregado para tarefas específicas (como tratamento de strings, visualização básica, etc.). A gestão dessas dependências foi realizada de forma rigorosa e documentada integralmente no arquivo “requirements.txt”, garantindo que o ambiente de execução possa ser replicado com precisão. Este arquivo se encontra disponível e versionado no repositório do projeto, conforme referenciado no “Apêndice D”;
- c) Ambiente de Desenvolvimento Integrado (IDE) – *GitHub Codespaces*: o desenvolvimento ocorreu majoritariamente neste ambiente em nuvem, o que proporcionou uma infraestrutura padronizada e acessível. Baseado no *Visual Studio Code*, o *Codespaces* permitiu a escrita e a execução do código diretamente em um ambiente pré-configurado, eliminando problemas de compatibilidade local e facilitando a colaboração;
- d) controle de versão e colaboração – *Git* e *GitHub*: Estas ferramentas foram a espinha dorsal do controle de versão do projeto. O *Git* foi utilizado localmente (dentro do *Codespaces*) para rastrear modificações no código-fonte e na documentação. O *GitHub* serviu como repositório remoto central, sendo essencial para: versionamento completo e histórico de todas as alterações; gerenciamento de *branches* para o desenvolvimento de *features* de forma isolada; facilitação da revisão de código e da colaboração; armazenamento seguro e acessível do código-fonte e da documentação técnica.

A automatização do processamento dos dados foi implementada através de um *pipeline* estruturado, composto por um conjunto de scripts em linguagem Python. Este *pipeline* está dividido em três etapas orquestradas, conforme ilustrado na Figura 7. O código-fonte

completo de todos os scripts está disponível no repositório digital no *GitHub*, acessível pelo DOI (<https://doi.org/10.5281/zenodo.16851246>).

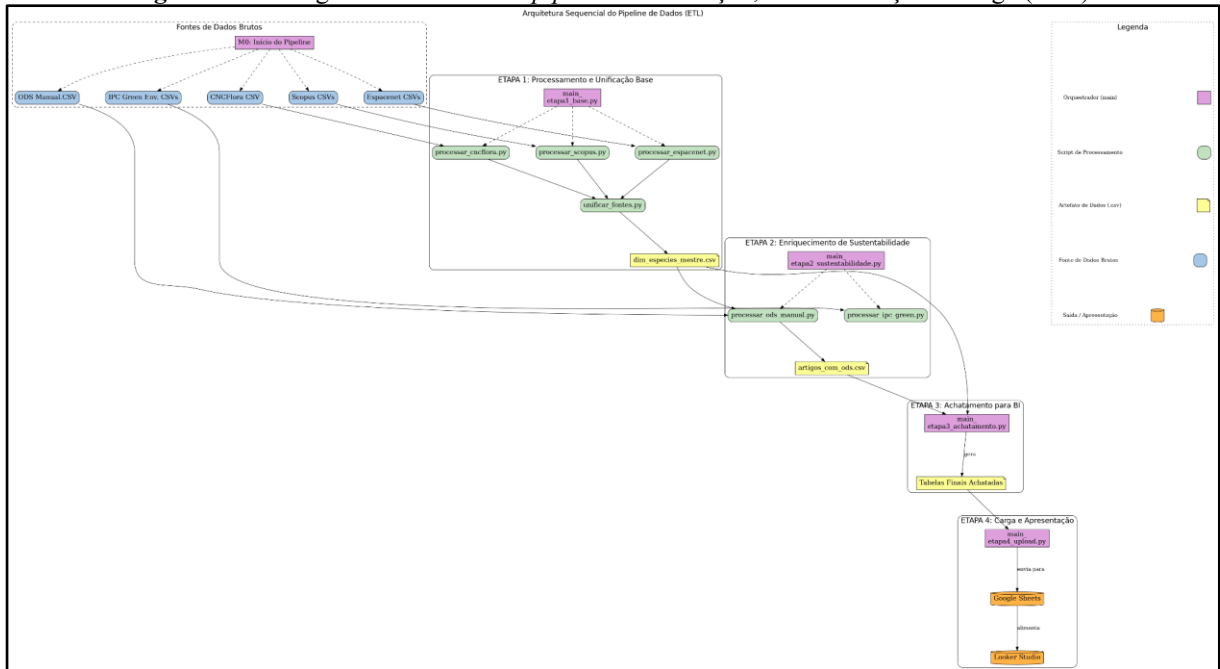
Portanto, os dados brutos exigiram um processo rigoroso de pré-processamento. A linguagem *Python* e o ecossistema de bibliotecas, notavelmente a biblioteca *Pandas* para manipulação de *DataFrames*, foram as ferramentas centrais. O Fluxograma detalhado do *pipeline* ETL (Figura 7), a seguir, ilustra a sequência de operações implementadas via *scripts*.

Acompanhando o fluxo operacional de ponta a ponta do *pipeline* ETL, inicia-se na fase de Extração (E), que abrange-se múltiplas fontes: as de bases primárias (*Scopus* e *Espacenet*) e as provenientes de fontes secundárias para alinhamento de sustentabilidade. O *pipeline* foi desenvolvido em linguagem de programação *Python* e é onde reside o núcleo da automação do sistema.

O bloco seguinte é o de Transformação (T), que representa o ponto de maior intervenção metodológica. Nesta etapa realiza-se o refinamento, ou seja, a limpeza dos dados brutos, processos tais como padronização de metadados, normalização de nomes e afins. Aqui também é criada a modelagem dimensional, através da esquematização em estrelas. É nesta etapa que garante-se que a heterogeneidade de dados seja harmonizada em uma estrutura unificada.

Posteriormente, o processo de Carga (L), em que se carrega os dados modelados para o *Google Sheets* preparando-os para o consumo imediato pela plataforma analítica *Looker Studio*. Este *pipeline* colabora para a sustentabilidade operacional do sistema, que pode ser replicado e atualizado periodicamente.

**Figura 7** – Fluxograma detalhado do *pipeline* de Extração, Transformação e Carga (ETL)



Fonte: A autora (2025).

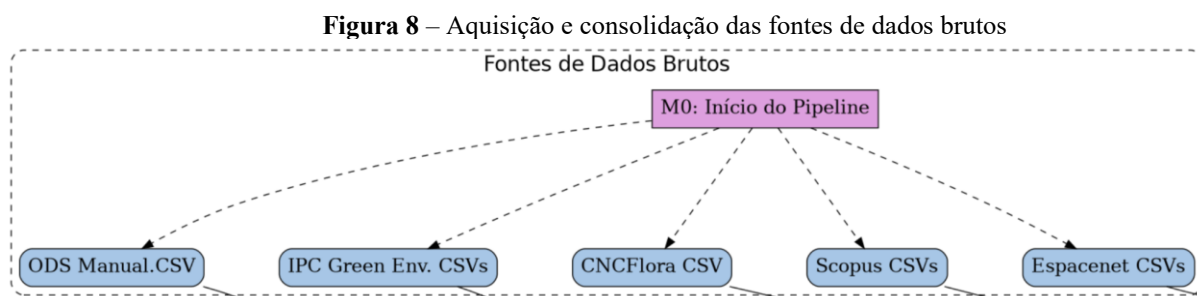
Nota: Elaborado com *python*, no segundo semestre de 2025.

O *pipeline* de processamento e unificação de dados é essencial para a gestão da informação e automação do monitoramento. É estruturado em três componentes principais: Módulos de Processamento por Fonte, que são scripts especializados (*Scopus*, *CNCFlora* e *Espacenet*) para extração, *parsing* e limpeza de dados de cada fonte heterogênea; Módulo de Unificação, que cria dimensões conformadas (mestras), como a “*dim\_especies\_mestre*”, para conectar e padronizar dados entre as fontes, fundamental para a arquitetura de *Data Warehouse*; e o Módulo Orquestrador (*main.py*), que gerencia o fluxo de trabalho, respeitando as dependências e permitindo a execução de todo o *pipeline* por um único comando. Este sistema mitiga erros humanos e garante a periodicidade e confiabilidade do monitoramento, alinhando-se com a necessidade de um sistema robusto e escalável.

### 3.3.1 Extração e Fontes de Dados Brutos

Seguindo o protocolo de busca e as expressões definidas na Fase 1 (Seção 3.2), a etapa de Extração (E) consistiu na coleta dos registros de artigos científicos na base *Scopus* e de patentes na *Espacenet*. Adicionalmente, dados de fontes de sustentabilidade (*CNCFlora*, *IPC Green Inventory* e referências para mapeamento ODS) foram adquiridos de suas respectivas plataformas. O volume de dados extraídos totalizou 800 artigos científicos e 153 patentes, os

quais, juntamente com os dados de sustentabilidade, constituíram o input bruto para a etapa subsequente.



Fonte: A autora (2025).

Nota: Elaborado utilizando *Python*.

O pré-processamento dos dados partem do input de 5 fontes coletadas e carregadas para o ambiente de processamento em formato “*csv*”. Cada fonte possui um formato diferente de dados, e para homogeneizar segue-se o *pipeline* para etapas subsequentes de tratamento.

### 3.3.2 Modelagem e Transformação dos dados

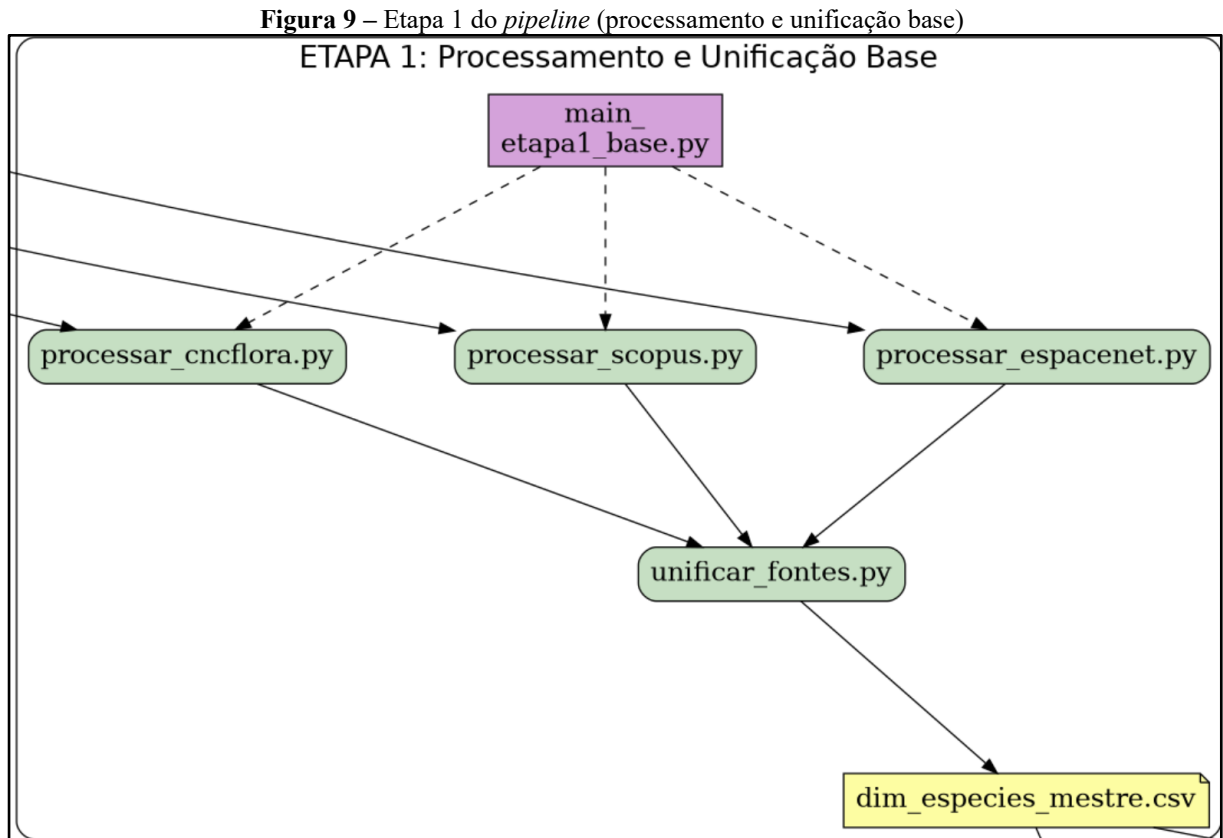
A Transformação (T) representa o núcleo metodológico desta pesquisa. Sua finalidade é converter o conjunto de dados heterogêneos e brutos em uma estrutura de informação padronizada e enriquecida, apta para a geração de indicadores. A Arquitetura geral do *pipeline* ETL e fluxo operacional fornece uma visão macro do processo (Apêndice D).

A execução do *pipeline* foi conduzida por rotinas de scripts em linguagem *Python* (versão 3.12), escolhida pela sua robustez em ciência de dados, o que assegura automação, rastreabilidade e facilidade para futuras atualizações do painel. A partir da arquitetura apresentada, a transformação foi sequencialmente organizada nas etapas seguintes.

#### 3.3.2.1 Lógica de limpeza e padronização dos dados

O módulo de processamento de metadados de grupos é responsável por uma série de transformações nos dados brutos, visando a uniformidade e a completude. Nesta etapa, o *pipeline* considera as particularidades dos dados brutos heterogêneos das fontes imputadas

(Figura 9). A rotina de limpeza utiliza *hashing*<sup>15</sup> (divisão) e *fuzzy matching*<sup>16</sup> (correspondência aproximada) para identificar e unificar dados padronizando-os.



Fonte: A autora (2025)

Nota: Elaborado utilizando *Python*.

Por processamento de dados base, entende-se a etapa inicial (limpeza e estruturação) dos dados brutos de suas respectivas fontes. O primeiro “Orquestrador<sup>17</sup>” (*main\_etapa1\_base.py*) foi responsável por executar *scripts* de tratamento primário dos dados, conforme Quadro 14.

**Quadro 14 – Estrutura de *scripts* composição do Orquestrador 1**

<i>Script</i>	Exemplo de entrada	Responsabilidade
<i>processar_scopus.py</i>	dados bibliométricos obtidos na <i>Scopus</i> ( <i>scopus_input.csv</i> )	Consolidar, padronizar e modelar dimensionalmente os dados dos artigos, ex.: Funções de limpeza e de criação de modelos

<sup>15</sup> **O que é Hashing.** Disponível em: <https://www.ime.usp.br/~pf/algoritmos/aulas/hash.html>. Acesso em: 25 jan. 2026.

<sup>16</sup> **Fuzzy matching.** Disponível em: <https://learn.microsoft.com/en-us/sql/relational-databases/fuzzy-string-match/overview?view=fabric-sqlldb>. Acesso em: 25 jan. 2026.

<sup>17</sup> Script criado para o processo coordenado do fluxo dos demais scripts garantindo que dependências sejam respeitadas e que o processo ocorra de modo automatizado e monitorável (Reis; Housley, 2022).

<i>Script</i>	<b>Exemplo de entrada</b>	<b>Responsabilidade</b>
		de autores/afiliações
processar_cncflora.py	dados obtidos no site CNCFlora (lista_vermelha_cnc_flora.csv E termos_plantas.txt)	Consolidar as informações de espécies, criando a dimensão mestre (dim_especies_mestre.csv)
unificar_fontes.py	dados de relação de artigos e espécies (scopus_dados_limpos_temp.csv E dim_especies_cncflora_temp.csv)	Ler os dados processados dos artigos, criar a dimensão de espécies unificadas e a tabela ponte artigo-espécie
processar_espacenet.py	dados obtidos na <i>Espacenet</i>	Consolidar, padronizar e modelar dimensionalmente os dados das patentes, ex.: Funções de limpeza e criação de modelos de países por publicação, por aplicante ou por inventor

Fonte: A autora (2025).

Nota: Elaborado no segundo semestre de 2025.

As principais rotinas implementadas incluem a Padronização de Nomes de Instituições e a Unificação de Formatos de Data. A padronização de nomes é um procedimento essencial para consolidar registros que se referem à mesma entidade, mas que foram registrados com variações textuais.

Por exemplo, variações como "Univ. de S. Paulo", "Universidade de S. Paulo - USP", "USP" e "Universidade de São Paulo" são mapeadas para uma única representação canônica ("Universidade de São Paulo"). O código mantém um dicionário de mapeamento institucional, que é atualizado periodicamente.

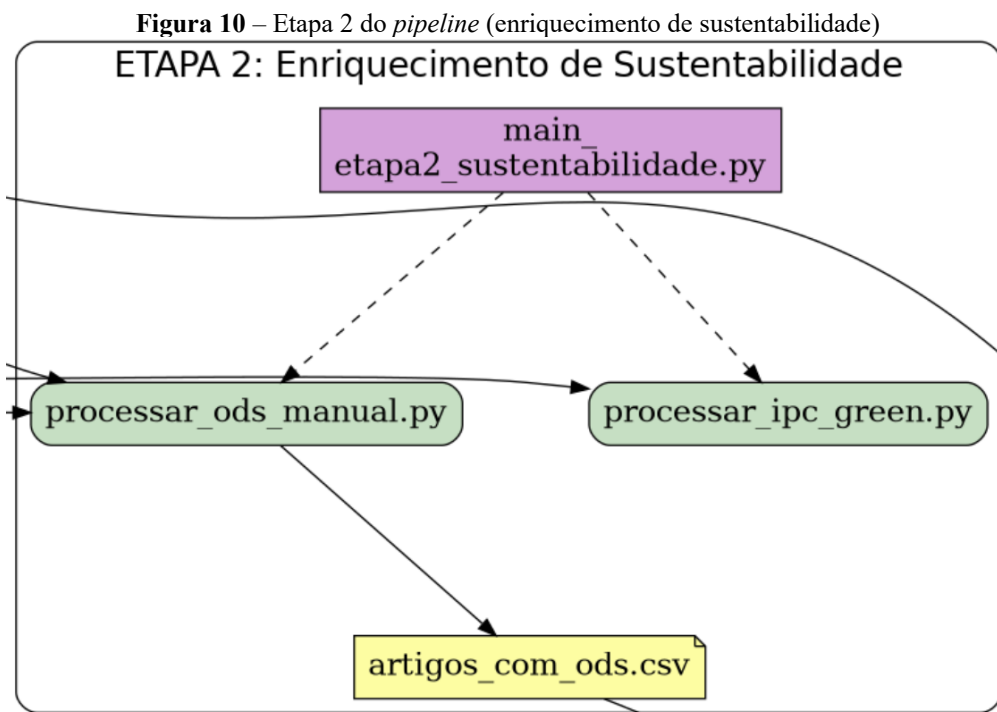
Além disso, a ausência de dados em campos importantes (*Missing Values*) é tratada de forma sistemática para evitar viés nas análises. As estratégias aplicadas incluem a Busca Manual e Imputação: para campos importantes com dados faltantes, foi realizada uma busca manual em fontes complementares para a recuperação do dado; caso a informação não fosse encontrada após a busca manual, o valor é substituído pela categoria explícita ("Não Informado" ou "Desconhecido"). A Remoção de registros com dados faltantes é realizada com cautela, somente quando a proporção de dados faltantes é baixa e o campo não é essencial para a análise principal, ou quando o registro está incompleto em múltiplas variáveis-chave.

O código é estruturado para operar em um fluxo sequencial: Extração → Limpeza/Padronização → Transformação → Carregamento (ETL). Após a execução do módulo de tratamento, é gerado um relatório de *log* que detalha a quantidade de registros afetados por cada regra de limpeza (ex.: número de instituições unificadas, número de missing

values imputados), permitindo a rastreabilidade e a validação da qualidade dos dados processados.

### 3.3.2.2 Mapeamento e enriquecimento de sustentabilidade

Esta etapa da Transformação (Figura 10) visa categorizar a produção científica e tecnológica em indicadores de alinhamento estratégico, utilizando os Objetivos de Desenvolvimento Sustentável (ODS) e a Classificação Internacional de Patentes para Tecnologia Verde (IPC *Green Inventory*). O mapeamento dos ODS foi aplicado aos artigos, enquanto o IPC GI foi aplicado aos registros de patentes.

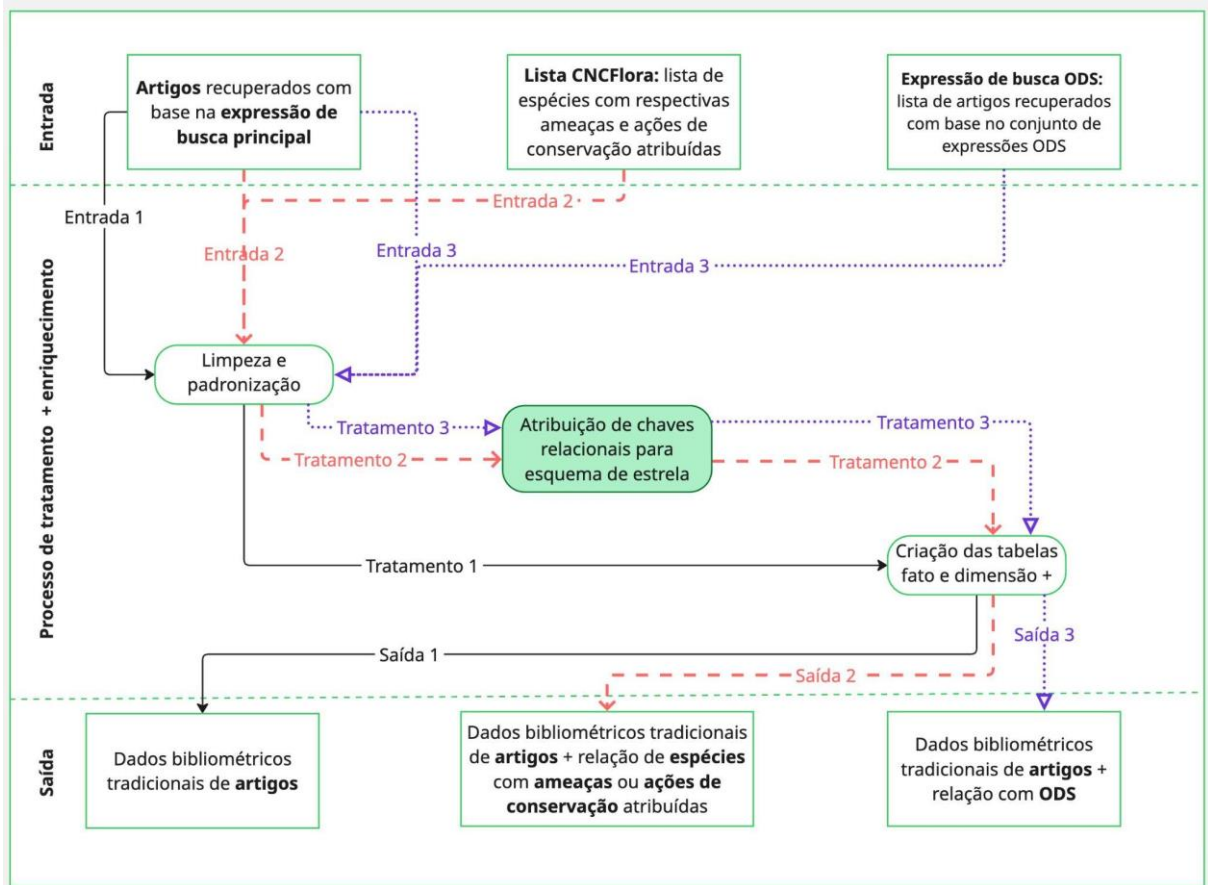


Fonte: A autora (2025).

Nota: Elaborado utilizando *Python*.

O processo de enriquecimento dos artigos científicos com as variáveis de sustentabilidade e a lógica de inter-relação das fontes está detalhado no fluxograma do relacionamento das fontes extraídas de artigos e sustentabilidade (Figura 11).

**Figura 11** – Lógica de enriquecimento de sustentabilidade para artigos



Fonte: A autora (2025).

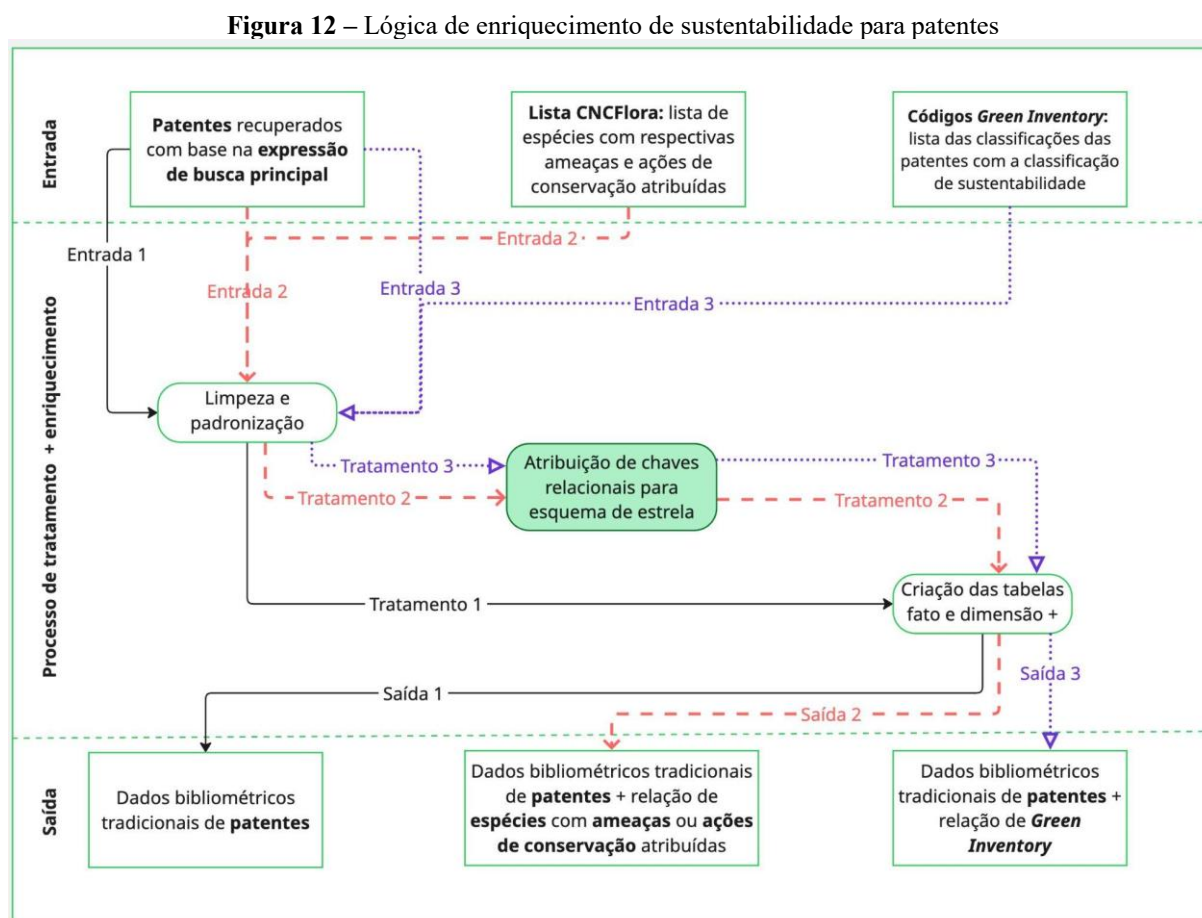
Nota: Elaborado usando a ferramenta *Miro*.

A lógica operacional dos scripts para mapeamento de ODS consistiu em:

- aplicação do protocolo de busca: os termos de busca específicos de cada ODS foram aplicados para identificar e classificar os registros de artigos e patentes alinhados a cada objetivo;
- geração de variável dimensional: foi gerada uma nova variável dimensional (*flag*) para cada registro, indicando a associação com um ou mais ODS (ex: ODS\_3, ODS\_15);
- criação de sub-amostra: a intersecção resultante permitiu a criação de uma sub-amostra de desenvolvimentos em CT&I sobre plantas endêmicas com relevância direta para a Agenda 2030, preparando o dado para a geração do indicador composto ODS.

Para os registros de patentes, o script realizou o cruzamento do código da Classificação Internacional de Patentes (IPC) com o *IPC Green Inventory*. Essa intersecção permitiu a criação de um indicador composto que sinaliza patentes alinhadas a tecnologias ambientalmente sustentáveis. A especificidade do enriquecimento das patentes e a lógica de suas fontes estão

ilustradas no fluxograma do relacionamento das fontes extraídas de patentes e sustentabilidade (Figura 12).



Fonte: A autora (2025).

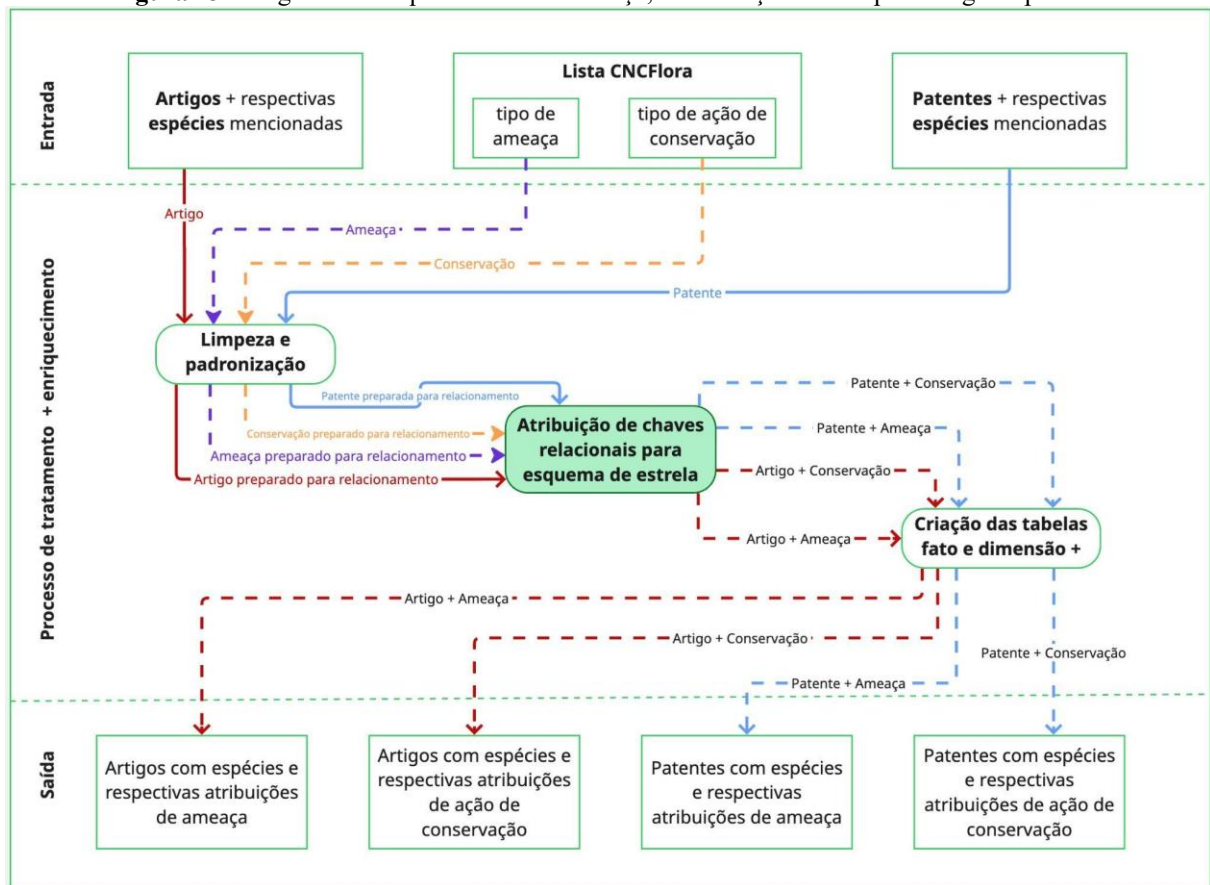
Nota: Elaborado usando a ferramenta *Miro*.

O Esquema Conceitual e Lógico para enriquecimento de patentes estabelece uma metodologia de sustentabilidade baseada na intersecção de dois eixos principais. O primeiro é a tecnologia verde, que, por meio do IPC *Green Inventory*, classifica as patentes alinhadas a tecnologias que mitigam as mudanças climáticas e promovem o desenvolvimento sustentável. O segundo eixo é o Uso da Biodiversidade e Risco, que utiliza o CNCFlora adaptada para avaliar se a inovação utiliza recursos da flora, especificamente plantas endêmicas, que estão sob risco ou atenção, garantindo a análise do uso responsável e sustentável da biodiversidade. Essa combinação de eixos cria indicadores compostos, que enriquecem a análise ao mensurar o esforço de inovação tecnológica juntamente com o uso responsável de recursos naturais, validando a estrutura proposta para informar políticas públicas.

Para o enriquecimento dos dados de conservação, risco e ameaça presentes nos artigos e patentes, e sua integração ao esquema estrela, a chave primária estabelecida foi a correlação

entre as dimensões por meio das espécies endêmicas citadas nos documentos e a listagem do CNCFlora. Conforme ilustrado na Figura 13, o processo de enriquecimento de dados de artigos e patentes baseado em plantas seguiu as etapas de limpeza e padronização dos dados, atribuição de chaves no esquema estrela para viabilizar a relação multidimensional e, por fim, a criação das tabelas Fato e Dimensão, que subsidiaram a construção dos gráficos e indicadores para análise no painel.

**Figura 13** – Lógica de enriquecimento de ameaça, conservação e risco para artigos e patentes



Fonte: A autora (2025).

Nota: Elaborado usando a ferramenta *Miro*.

Esta etapa é também composta por orquestrador (Quadro 15) responsável por integrar as classificações previamente realizadas de forma manual (ODS) e cruzar os códigos da IPC. Esta fase crítica de modelagem inclui o *script* que executa o processo de "achatamento" (*flattening*) das tabelas. O objetivo central deste processo é desnormalizar e unificar as tabelas de dimensão e quaisquer tabelas ponte (*bridge tables*) previamente criadas em um conjunto reduzido de tabelas finais. O resultado são tabelas otimizadas, com todas as chaves descritivas incluídas, prontas para serem consumidas diretamente por ferramentas de *Business Intelligence* e visualização, como o *Looker Studio*, minimizando a necessidade de múltiplas operações de

join no momento da consulta. Essa otimização visa garantir um desempenho rápido e uma experiência de usuário fluida na exploração dos dados.

**Quadro 15** – Estrutura de *scripts* de composição do Orquestrador 2

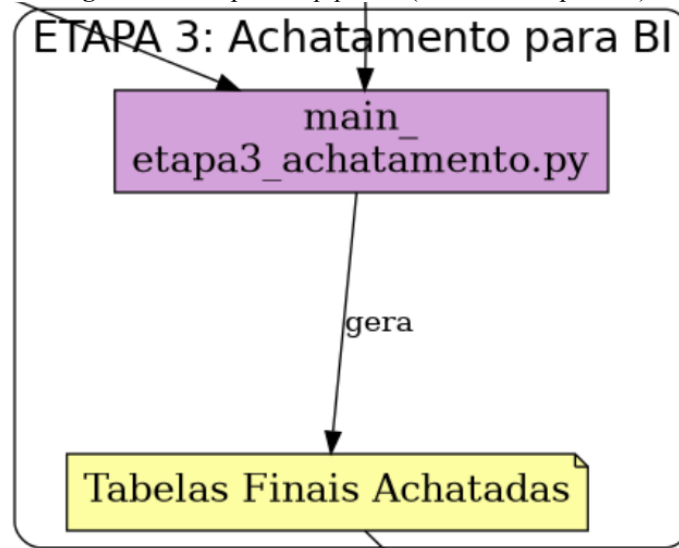
<i>Script</i>	Exemplo de entrada	Responsabilidade
processar_ods_manual.py	scopus_dados_limpos_temp.csv v scopus_ods_v1.csv artigos_com_ods.csv	Integrar a classificação manual de ODS aos dados de artigos, utilizando o identificador EID como chave de ligação
processar_ipc_green.py	espacenet/dim_ipc.csv espacenet/ipc_classificado_green.csv	Cruzar os códigos de Classificação Internacional de Patentes (IPC) de cada patente com a lista do IPC <i>Green Inventory</i> , adicionando um indicador booleano ( <i>is_green</i> ) e a descrição da tecnologia verde correspondente.

Fonte: A autora (2025).

Dessa forma, o aprimoramento futuro dos dados de sustentabilidade demandará um esforço cognitivo consideravelmente menor. Isso se deve à existência de um modelo conceitual e de uma aplicação prática já validados como eficientes e sustentáveis. Tal otimização de tempo permitirá avanços e melhorias capazes de gerar resultados e indicadores significativos sobre a sustentabilidade, especialmente em relação aos artigos e patentes que abordam plantas endêmicas da Amazônia brasileira.

### 3.3.3 Modelagem Dimensional

A terceira etapa (Figura 14) do *pipeline* realiza o achatamento de todos os dados já tratados, padronizados e enriquecidos nas etapas anteriores.

Figura 14 – Etapa 3 do *pipeline* (achatamento para BI)

Fonte: A autora (2025).

Nota: Elaborado utilizando *Python*.

O design do esquema foi orientado para a otimização de consultas e a clareza na representação dos dados, seguindo uma abordagem *star schema* simplificada. No final, criou-se Tabelas Fato e Tabelas Dimensão, permitindo que aconteça o carregamento na estrutura necessária para integração ao *Looker Studio* do *Google*. O orquestrador desta etapa (*main\_etapa3\_achatamento.py*) é o responsável por processar os scripts conforme Quadro 16.

Quadro 16 – Estrutura de *scripts* de composição do Orquestrador 3

<i>Script</i>	Subetapas do <i>script</i>	Responsabilidade
<i>criar_tabelas_finais.py</i>	1 - Carregamento robusto de todos os arquivos; 2 - Garantir consistência dos tipos de dados; 3 - Criação e salvamento das tabelas de relação (artigos); 4 - Criação e salvamento das tabelas de relação (patentes); 5 - Salvamento dos resultados.	1 - Definir as chaves de relacionamento de cada fonte; 2 - Garantir quais chaves se inter-relacionam especificamente; 3 - Criar as tabelas fato e dimensão dos dados de artigos; 4 - Criar as tabelas fato e dimensão dos dados de patentes; 5 - Criar e salvar arquivos finais prontos para carregamento à ferramenta de BI.
<i>criar_fato_cncflora</i>	1 - Configuração dos caminhos; 2 - Carregamentos de dados; 3 - Processamento dos dados; 4 - Criação da tabela fato final; 5 - Salvamento dos resultados.	1 - Definir pastas de entrada (Apêndice C) e saída; 2 - Definir as chaves de relacionamento (cada campo/coluna); 3 - Processar relacionamento dos dados de ameaça e de conservação; 4 - Criar a tabela fato final compilando dados das plantas; 5 - Criar e salvar arquivos finais prontos para carregamento à ferramenta de BI.

Fonte: A autora (2025).

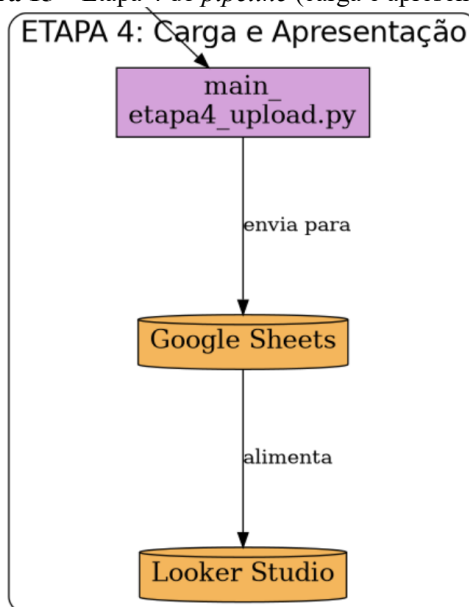
A criação de dimensões é essencial para a categorização e segmentação dos dados. As principais dimensões estabelecidas incluem: a dimensão ODS – armazena as classificações hierárquicas e descritivas dos ODSs, permitindo a agregação de dados por objetivo, meta ou sub-meta. As classificações foram padronizadas e pré-processadas para garantir a integridade referencial; ou a dimensão tempo – contempla a granularidade temporal necessária para a análise de tendências, incluindo ano, trimestre e mês de publicação dos documentos. Há também outras dimensões relevantes (ex: Dimensão Tipo de Documento, Dimensão Instituição, etc.), que também foram criadas para alocar classificações específicas.

A Tabela de Fato (central) foi construída mediante o levantamento e contagem de documentos (artigos científicos e patentes) que foram alinhados com as classificações de sustentabilidade. O processo envolveu, por exemplo, a contagem – foi realizada a contagem exata dos documentos base que se ligavam a cada combinação única de dimensões; o algoritmo de ligação taxonômica – a ligação entre os documentos e as dimensões de sustentabilidade (ODS, por exemplo) foi consolidada utilizando o algoritmo de ligação pela chave taxonômica previamente definida, garantindo a rastreabilidade e a precisão do alinhamento; ou o cálculo dos Indicadores Compostos – que são críticos para a análise, como a "Contagem de Patentes alinhadas ao ODS 9 (Indústria, Inovação e Infraestrutura)" ou "Total de Artigos Científicos relacionados ao ODS 13 (Ação Contra a Mudança Global do Clima)", e foram calculados diretamente no ambiente de *script* (ETL *script*) antes da Carga (*Loading*) final no *Data Warehouse*. Isso assegurou que as métricas complexas estivessem prontas para consumo imediato.

### 3.3.4 Estruturação e carregamento

Esta etapa contempla a última etapa do *script* e visa carregar automaticamente os dados trabalhados nas etapas anteriores para subsidiar a criação e publicação do protótipo do painel (Figura 15).

**Figura 15** – Etapa 4 do *pipeline* (carga e apresentação)



Fonte: A autora (2025).

Nota: Elaborado utilizando *Python*

Para o armazenamento das fontes de dados que alimentam o painel, a ferramenta *Google Sheets* foi mantida devido ao seu acesso gratuito e à integração nativa com o *Google Looker Studio*. Entretanto, a carga dessas planilhas foi automatizada por meio da utilização de uma Interface de Programação de Aplicações (API – *Application Programming Interface*) integrada ao código. Essa automação garante que, ao executar o código, as fontes de dados do *Google Sheets* sejam atualizadas, o que, por sua vez, atualiza automaticamente os dados no painel.

O Orquestrador da Etapa 4 contém o script final que utiliza a API para automatizar o carregamento das tabelas com dados estruturados e tratados para o *Google Sheets*. Dessa forma, é feita a conexão final do *pipeline*, onde a carga de dados realiza o upload definitivo dos dados modelados, servindo como fonte de dados em tempo real para o LS.

A estruturação e a aplicação de conceitos a esses dados culminam na possibilidade de criar uma visualização simplificada, composta por um *back-end* para o painel, portanto, o Esquema de Estrela.

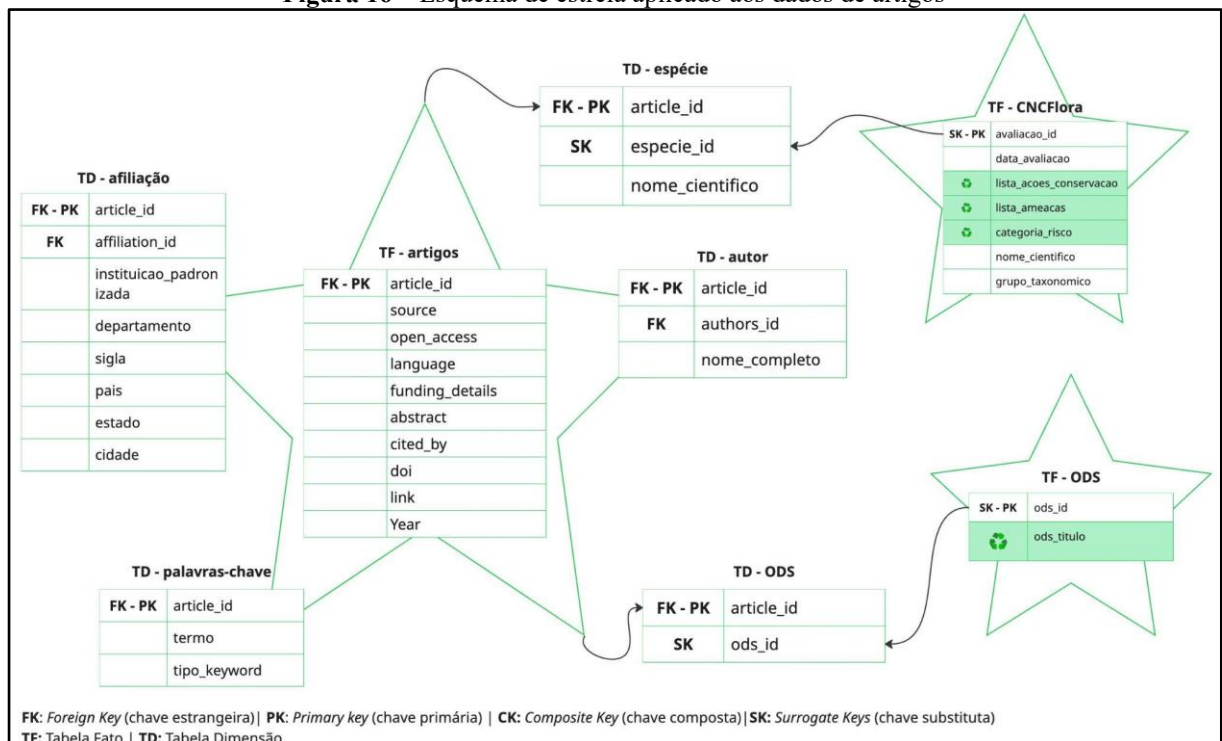
O Esquema Estrela permite a intersecção dos indicadores bibliométricos (fatos) com os indicadores de sustentabilidade (dimensões). Na Figura 16, é detalhada a arquitetura de dados para a dimensão científica (artigos). Esta modelagem dimensional é o resultado da fase de Transformação (o T do ETL), e é essencial para o Objetivo Específico 2 (Elaborar e integrar indicadores). No centro da primeira estrela (à esquerda da imagem) encontra-se a “Tabela de

Fato: Artigo” que possui a chave primária (PK), capaz de vincular as múltiplas tabelas de dimensão e assim fornecer o contexto analítico, conforme orientações de Kimball e Ross (2013).

O esquema de estrela e suas tabelas fato e dimensão, permitem a integração de dados heterogêneos (*Scopus, Espacenet* e *CNCFlora*) de forma padronizada, estrutura esta que garante a inserção de novos dados e atualizações sem quebrar a lógica dos indicadores, portanto a garantia de sustentabilidade de dados e indicadores do sistema. (Moraes *et al.*, 2024)

A chave de relacionamento que unifica o viés de sustentabilidade – proveniente do *CNCFlora* – aos artigos que versam sobre tipos de ações de conservação atribuído à uma planta, ou qual sua categoria de risco (tipo de ameaça), é através do campo de chave aleatória criada (SK) para cada espécie endêmica encontrada nos artigos. O viés de sustentabilidade com base em ODS são relacionados por meio da também chave primária criada usando o próprio número dos ODS (de 1 a 16).

**Figura 16** – Esquema de estrela aplicado aos dados de artigos



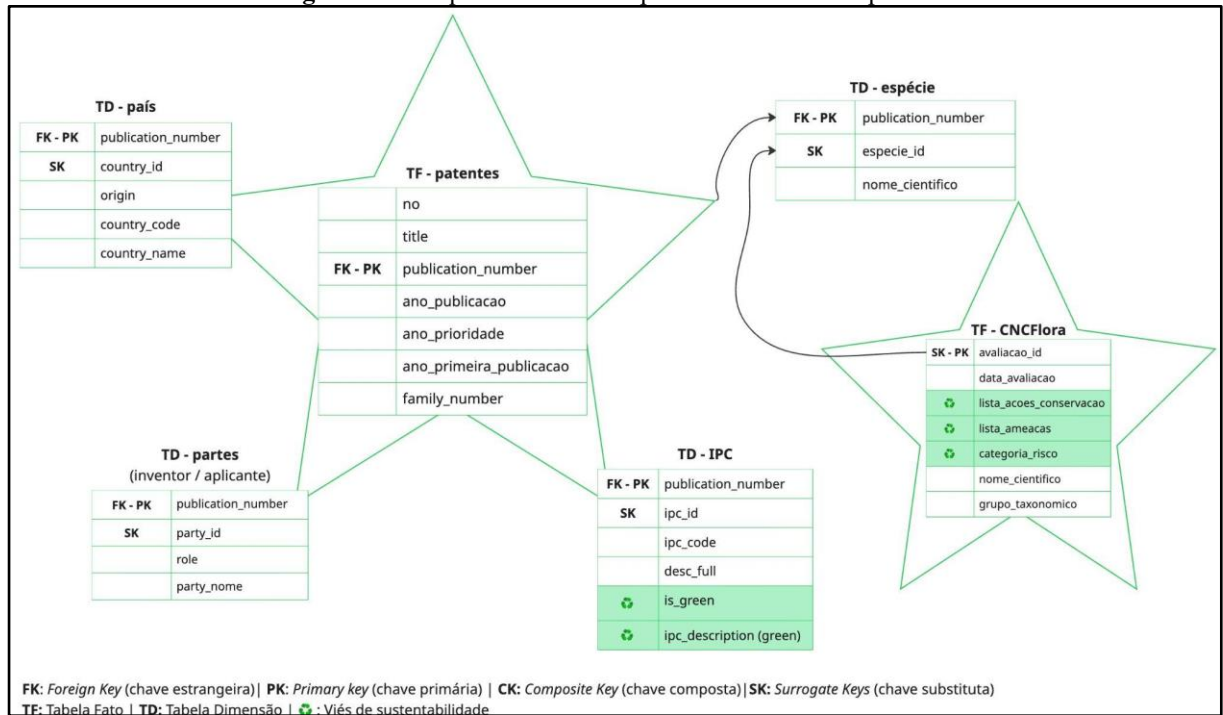
Fonte: A autora (2025).

Nota: Elaborado usando a ferramenta *Miro*.

A Figura 17 apresenta a modelagem da dimensão tecnológica (patentes). Assim como ocorre na estrutura de artigos, o centro é a “Tabela de Fato: Patente”. A figura permite ver que o principal desafio da heterogeneidade se dá pela inclusão das dimensões de sustentabilidade para patentes, onde é necessário cruzar a classificação *Green Inventory* com a classificação

geral IPC. É através deste cruzamento que pode-se mensurar o esforço de inovação tecnológica com potencial de impacto ambiental positivo.

**Figura 17 – Esquema de estrela aplicado aos dados de patente**



Fonte: A autora (2025).

Nota: Elaborado usando a ferramenta *Miro*.

A consistência da modelagem é assegurada pelo uso da dimensão de planta endêmica, sendo ele um elo de conexão imprescindível para ambos os esquemas de estrela. Esta chave comum valida o conceito de Indicadores Composto, uma vez que garante que a análise da produção científica e tecnológica possa ser realizada de forma integrada e sob o mesmo viés classificatório de sustentabilidade, no que tange a classificação de ameaças atribuídas às espécies ou tipos de ações de conservação.

## 4 RESULTADOS E DISCUSSÃO

O presente capítulo materializa os resultados do sistema de monitoramento automatizado, conforme detalhado no Capítulo 3 (Métodos), apresentando a análise dos Indicadores Compostos desenvolvidos e implementados. Diferente de análises bibliométricas tradicionais, esta seção se concentra na interpretação da dinâmica dos desenvolvimentos científicos e tecnológicos sobre plantas endêmicas da Amazônia a partir da intersecção dos dados de produção (*Scopus* e *Espacenet*) com as classificações de sustentabilidade (ODS, CNCFlora e *IPC Green Inventory*). Estes dividem-se nos subtópicos com foco nos artigos (4.1) e nas patentes (4.2).

A subseção “Indicadores Bibliométricos de Produção Científica” (4.1), foca na análise de publicações, autores, instituições e periódicos. A ênfase é dada à contribuição científica para a sustentabilidade, avaliada por meio dos Objetivos de Desenvolvimento Sustentável (ODS) e listas de ameaça. Detalha subseções sobre o viés científico das publicações por ameaças, ODS e oportunidades de pesquisas futuras. O segundo ponto trata dos “Indicadores Bibliométricos de Produção Tecnológica e Inovativa” (4.2), focado na análise de patentes, aplicantes, inventores e classificações IPC. A ênfase é na tecnologia alinhada à sustentabilidade, avaliada por meio da classificação *IPC Green Inventory* e listas de ameaça. Detalha uma subseção sobre o viés tecnológico das patentes.

A etapa de discussão apresenta além da descrição dos dados, propõe uma análise crítica das tendências e lacunas, e correlaciona o desempenho dos indicadores com as implicações para a gestão da informação ambiental.

Antes de detalhar a análise, é fundamental reconhecer as limitações metodológicas inerentes à construção do sistema e de seus indicadores. Conforme discutido na Fase 1 da Metodologia, a seleção das bases de dados (*Scopus* e *Espacenet*) implica uma restrição do universo analisado à produção indexada, podendo sub-representar pesquisas regionais. No que tange aos indicadores compostos, a quantificação da sustentabilidade por meio de classificadores textuais (ODS, *IPC Green Inventory* e listas do CNCFlora) atua como um filtro, permitindo a análise sistemática, mas exigindo interpretação crítica devido ao risco de viés inerente à delimitação por palavras-chave e mineração textual. Esses limites contextualizam o escopo de validade dos resultados apresentados.

A análise da produção científica (Seção 4.1) e tecnológica (Seção 4.2) é fundamentada na Bibliometria, que, segundo Alvarado (2008) e Araújo (2006), constitui a ferramenta legítima para a contagem e análise dos registros documentais de um campo. Os indicadores de produção,

como o número de publicações por autor (Seção 4.1.3), refletem a densidade e o protagonismo dos pesquisadores na área. No entanto, ao apresentar as métricas de autoria, é essencial reconhecer a limitação imposta por problemas de desambiguação e normalização, conforme discutido por De Sousa *et al.* (2024). O Painel mitiga parcialmente este desafio através de sua *pipeline* ETL, mas a contagem de produção deve ser entendida como um indicador de tendência, e não de impacto qualitativo absoluto.

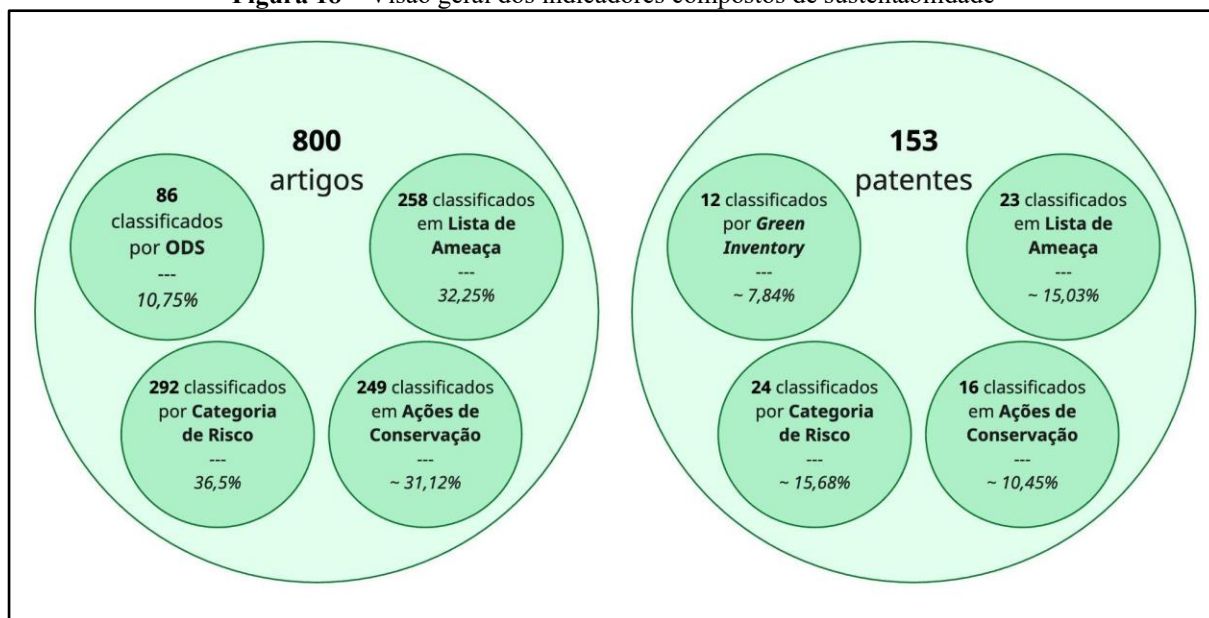
A partir da busca pelas plantas, realizada em 14 de dezembro de 2023, recuperou-se o total de 1975 espécies, sendo estas pertencentes a 25 gêneros, com 36 subespécies e 95 variedades. Estas foram organizadas conforme seu respectivo grupo taxonômico, conforme Tabela 2 abaixo. A busca pelos artigos foi realizada em abril de 2024 na *Scopus* e pelas patentes em agosto de 2025 na *Espacenet*.

**Tabela 2** - Relação quantitativa de espécies pesquisadas e registros bibliométricos recuperados

Grupo Taxonômico (número)	Grupo Taxonômico (nome)	Quantidade de espécies	Artigos científicos (título, resumo e palavras-chave)	Patentes (all)
Grupo 1	Angiospermas	1913	753	113
Grupo 2	Briófitas > Hepáticas	23	16	0
Grupo 3	Briófitas > Musgos	4	1	0
Grupo 4	Fungos > Strictu sensus	4	1	0
Grupo 5	Samambaias e Licófitas	31	29	2
<b>TOTAIS</b>		1975	800	115 (famílias)

Fonte: A autora (2025).

Além da análise bibliométrica quantitativa (Alvarado, 2008; Araújo, 2006), o estudo aprofunda a classificação de sustentabilidade por meio de dados complementares e cruciais, organizados em macros. Estes macros, essenciais para contextualizar o impacto real e a aderência à sustentabilidade, englobam três pilares principais: ODS, *Green Inventory* e listas do CNCFlora. Essa integração de dados macros permite classificar entidades/artigos e estabelecer um panorama abrangente para as conclusões da dissertação, visualizado na Figura 18.

**Figura 18** – Visão geral dos indicadores compostos de sustentabilidade

Fonte: A autora (2026)

Nota: Elaborado utilizando a ferramenta *Miro* em 2026.

A análise geral revela uma integração modesta da sustentabilidade em patentes, com a relevância do tema oscilando entre 7% e 17% do conteúdo. No âmbito científico, essa proporção é ligeiramente maior, variando de 10% a 37%, embora ainda permaneça abaixo de 50%. Tais resultados levantam o questionamento sobre a possibilidade de uma maior inclusão da sustentabilidade nas produções científicas e tecnológicas. Este questionamento é particularmente pertinente para as patentes, dado o seu potencial comercial e o foco em exploração, sugerindo um espaço significativo para aumentar a atenção a essa dimensão.

#### 4.1 INDICADORES BIBLIOMÉTRICOS DE PRODUÇÃO CIENTÍFICA

Em uma perspectiva macro, foi possível mapear 58 países (com base na afiliação dispostas nos registros bibliométricos dos artigos). Destes, somou-se 446 instituições, das quais predominaram as universidades, tanto nacionais quanto internacionais – além de institutos de pesquisa e empresas privadas, por exemplo. Afiados à esses países e instituições somaram-se 2.358 autores, sendo 1.665 destes vinculados ao Brasil (cerca de 69,60%). Além disso, os artigos foram publicados em 6 diferentes idiomas. Estes dados gerais permitem identificar um interesse global e de diversos atores no cenário acadêmico e científico pelas plantas endêmicas da Amazônia brasileira, refletindo assim a importância e relevância desta flora não somente para os brasileiros, mas também por instituições e autores de diversos países pelo mundo.

Das 1975 espécies endêmicas utilizadas como entrada para as buscas, identificou-se 437

plantas mencionadas nos artigos, ou seja, cerca de 22% das plantas endêmicas foram mapeadas no radar científico. Isto permite questionar se há algum tipo de negligência por parte dos atores de pesquisa e ciência frente aos outros 78% de espécies, que podem oferecer benefícios semelhantes (ou não) aos dos 22% estudados. Ou então um questionamento inverso, se estes 78% realmente não possuem potencial e, portanto, são arbitrariamente ignorados. Tais questionamentos poderão ser abordados em pesquisas futuras, tendo por subsídio os dados apresentados como ponto de partida.

No Brasil, 259 espécies foram mencionadas. Isto representa cerca de 59% das plantas estudadas, ou seja, há ao menos 178 plantas nativas que compõem pesquisas no exterior que não foram localizadas nas publicações brasileiras. Relacionar quais são essas plantas e seus focos de pesquisa poderão oferecer insumos para entender o porquê do interesse externo aos nossos recursos exclusivos. Além disso, é relevante verificar se tais pesquisas levam em consideração os aspectos de sustentabilidade, ou se possuem visão para extrativismo e comercialização, implicando em potencial biopirataria e aspectos econômicos de extrema importância de discussão e elucidação.

Um maior detalhamento dos indicadores obtidos foi elaborado à seguir.

#### **4.1.1 Número das publicações por afiliação e ano**

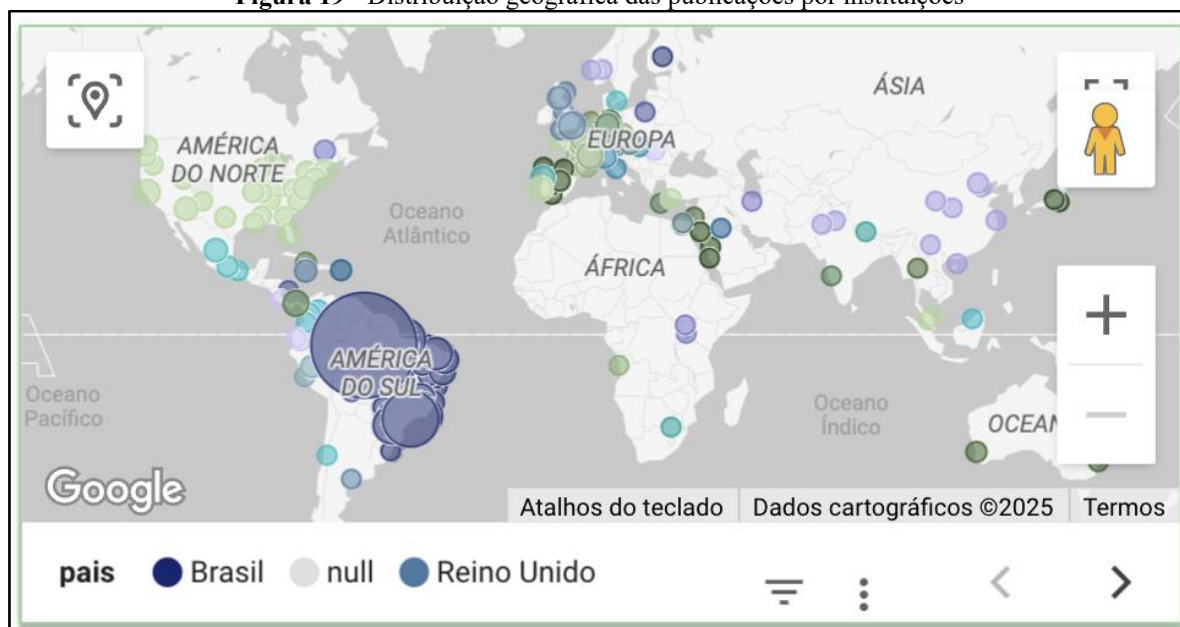
O mapa abaixo apresenta a distribuição geográfica das publicações com base nas instituições de afiliação. Na Figura 19 que segue, pode-se verificar esferas de diferentes tamanhos, posicionadas em demarcações geoespaciais das respectivas instituições. As cores representam o país, e a proporção (tamanho) das esferas o número de publicações.

Primeiramente, pode-se notar a diversidade de países (também agrupados por cores) presentes nas pesquisas. As esferas marcam presença em continentes tais como América do Norte e América do Sul, Europa, Ásia, África e Oceania. Através desta visualização já pode-se inferir que instituições ao redor do mundo demonstram interesse (interesse global) direto à plantas endêmicas da Amazônia brasileira, em uma perspectiva científica.

Em segundo instância, evidencia-se o tamanho das esferas, predominantemente maiores na região do Brasil. Isto significa que o número de publicações de instituições brasileiras (415 artigos – mais de 51% do todo) protagonizam o cenário científico, e possuem volume de artigos significativamente maiores que demais países. Potencialmente, o Brasil reconhece o valor de sua biodiversidade e investe em estudos destes recursos naturais, ao que indica a produção

científica.

**Figura 19** - Distribuição geográfica das publicações por instituições



Fonte: A autora (2025).

Nota: Captura de tela do LS, realizado no segundo semestre de 2025.

Na Figura 20 observa-se o enfoque no Brasil. Visualmente percebe-se que instituições em todas as regiões do país publicaram sobre plantas endêmicas. A lista de instituições considera a entidade mencionada como afiliação do autor no momento da publicação de cada artigo, refletindo o vínculo do autor e afiliação no período da publicação.

A maior esfera no mapa se posiciona em Manaus e trata-se do Instituto Nacional de Pesquisas da Amazônia (INPA), indicando ser a instituição responsável pelo maior número de publicações científicas frente à flora endêmica amazônica (69 artigos). Ainda na região norte e nordeste, destacam-se instituições como Universidade Federal do Amazonas (UFAM) com 35 publicações, Universidade Federal do Pará (UFPA) com 21 artigos, Universidade Federal Rural do Amazonas (UFRA) com 20 artigos e Universidade Federal do Ceará (UFC) com 11 publicações.

A segunda região de destaque em número de publicações concentra-se no sudeste brasileiro. O estado de São Paulo predomina com maior volume de artigos provenientes de instituições como Universidade de São Paulo (USP) com 29 artigos, Universidade Estadual de Campinas (UNICAMP) com 21 publicações, e Universidade Federal de São Carlos (UFSCar) com 6 pesquisas. No estado do Rio de Janeiro destaca-se a Fundação Oswaldo Cruz (Fiocruz) com 17 artigos, seguido por Universidade Federal do Rio de Janeiro (UFRJ) com 13 publicações, Instituto Militar de Engenharia (11 artigos) e Jardim Botânico do Rio de Janeiro

(JBRJ) com 10 artigos.

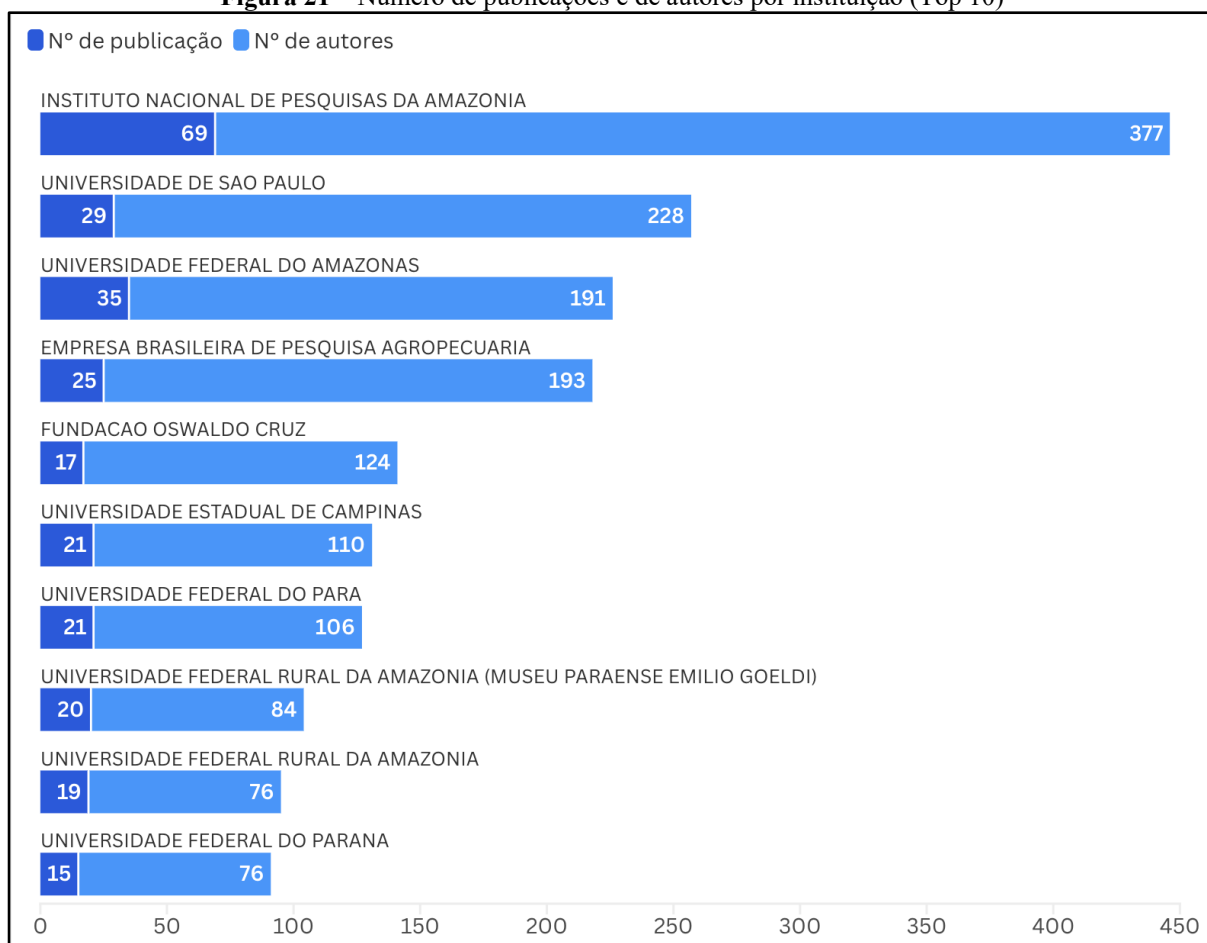
**Figura 20** - Distribuição geográfica das publicações por instituições brasileiras



Fonte: A autora (2025).

Nota: Captura de tela do LS, realizado no segundo semestre de 2025.

A proeminência e o engajamento do Brasil no campo de pesquisa em análise são incontestáveis, o que se evidencia de forma particular ao examinar a afiliação institucional dos autores e o volume de publicações. Ao analisar o ranking das instituições mais produtivas, constata-se um domínio absoluto de entidades nacionais. Conforme ilustrado pela Figura 21, a totalidade do top 10 — as dez primeiras posições — é ocupada por instituições brasileiras, tanto em termos de número absoluto de publicações quanto no que concerne ao número de autores a elas vinculados. Este dado não apenas corrobora a robustez da produção científica interna, mas também ressalta a concentração da expertise e da capacidade de pesquisa dentro do território nacional, com universidades e centros de pesquisa brasileiros liderando o cenário de geração de conhecimento na área. Essa concentração demonstra a relevância das políticas de investimento e fomento à pesquisa básica e aplicada implementadas pelas agências e órgãos de financiamento nacionais, bem como a consolidação de grupos de pesquisa de excelência no âmbito acadêmico do país.

**Figura 21** – Número de publicações e de autores por instituição (Top 10)

Fonte: A autora (2025)

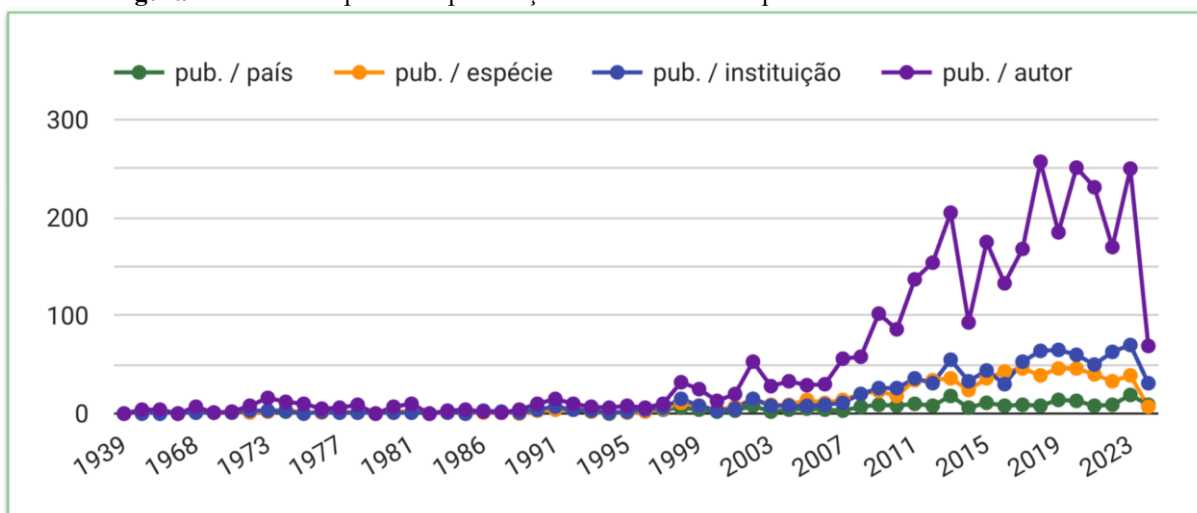
Nota: Elaborado utilizando a ferramenta *Flourish*.

O Instituto Nacional de Pesquisas da Amazônia (INPA) é a instituição líder na pesquisa sobre plantas endêmicas do bioma amazônico, destacando-se tanto pelo maior número de publicações (69) quanto pelo maior número de autores (377). Na sequência, a Universidade de São Paulo (USP) aparece com 29 publicações e um número expressivo de pesquisadores (228). A Universidade Federal do Amazonas (UFAM), por sua vez, supera a USP em publicações (35), mas conta com um número menor de autores (191). O ranking inclui ainda outras instituições de destaque, como a Embrapa e a Fiocruz, além de universidades das regiões Norte, Sudeste e Sul do país.

Com base no ano de publicação por artigos, a Figura 22 apresenta a distribuição temporal de publicações por países (linha verde), espécies (linha laranja), instituições (linha azul marinho) e autores (linha roxa). Há um crescente interesse na temática em todas as vertentes, sobretudo na proporção de autores diferentes. Os números se mantiveram tímidos de 1939 até cerca de 2001. Em 2002 o número de autores começa a despontar (saltando de 20 para 53), e apesar de oscilações cresce significativamente após 2007 e atinge o ápice em 2018 com

257 autores publicados.

**Figura 22** - Série temporal das publicações científicas sobre plantas endêmicas da Amazônia



Fonte: A autora (2025).

Nota: Captura de tela do LS, realizado no segundo semestre de 2025.

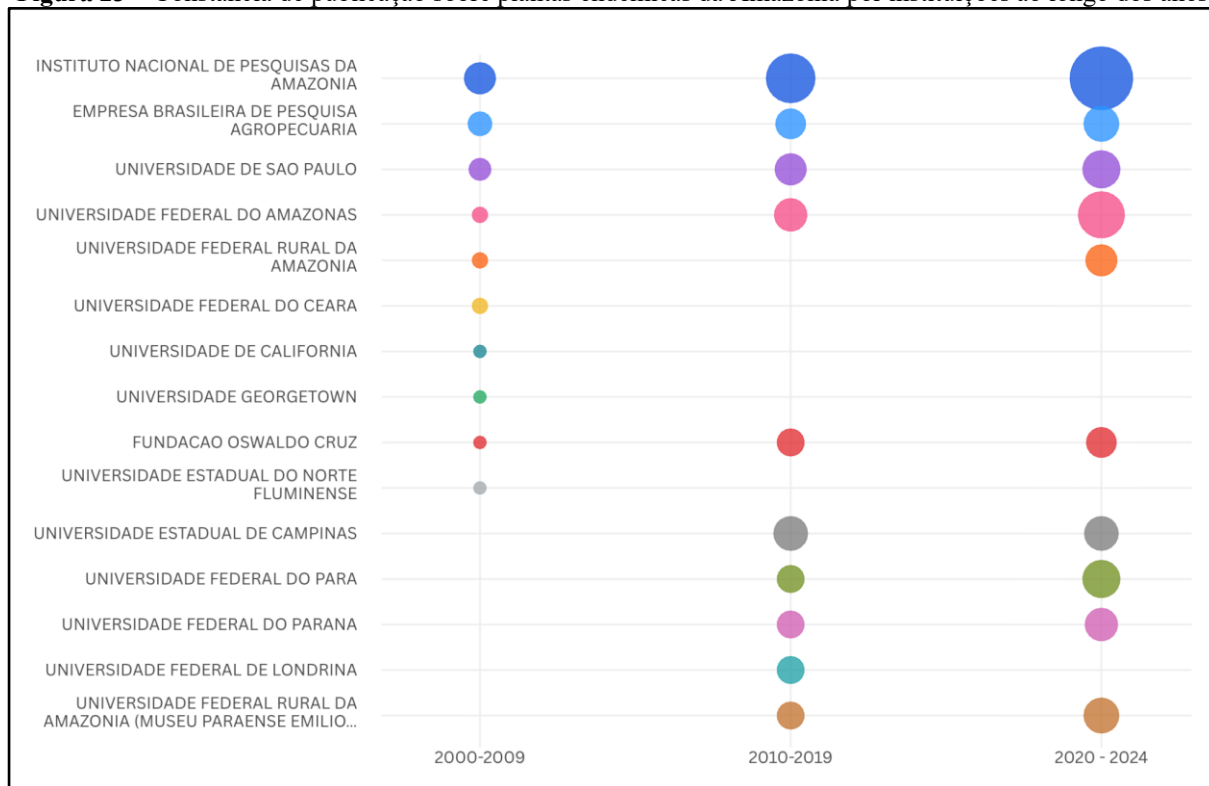
Referente às demais variáveis (espécie, instituição e país) percebe-se valores mais aproximados e constantes, indicando uma solidez pelo interesse destes ao longo dos anos. Nos últimos 10 anos (2014 à 2024) somaram-se 426 artigos, evidenciando a concentração de mais de 50% das publicações na última década. Este padrão de crescimento na documentação acadêmica é uma expectativa confirmada por Alvarado (2008) e Araújo (2006), que posicionam a bibliometria como ferramenta essencial para atestar a expansão de um campo de conhecimento. No entanto, ao se observar a dimensão tecnológica (Seção 4.2.2), o crescimento das patentes é um fenômeno que, segundo Grupp & Moge (2005), exige uma interpretação distinta, pois a tecnologia reflete a aplicação inovativa e não apenas o avanço do conhecimento básico.

A análise do ranqueamento de publicações por períodos revela a consistência de instituições como INPA, Embrapa, USP, UFAM e Fiocruz entre 2000 e 2024. Este período foi dividido em três frações: 2000-2009, 2010-2019 e 2020-2024 (conforme Figura 23).

No entanto, observa-se que algumas instituições presentes no primeiro período, como a Universidade Federal do Ceará (3 artigos), a Universidade da Califórnia, a Universidade Georgetown e a Universidade Federal do Norte Fluminense (2 artigos cada), não se mantiveram nos períodos subsequentes. Essa mudança deve-se à ascensão de outras instituições com um volume significativamente maior de publicações nos dois períodos mais recentes. A Unicamp é um destaque, com 14 publicações em cada um dos últimos dois períodos, totalizando 24 artigos. Outras quatro instituições que se destacaram recentemente são a Universidade Federal

do Pará, Universidade Federal do Paraná, Universidade Federal de Londrina e o Museu Paraense Emílio Goeldi (vinculado à UFRAM).

**Figura 23** – Constância de publicação sobre plantas endêmicas da Amazônia por instituições ao longo dos anos



Fonte: A autora (2026).

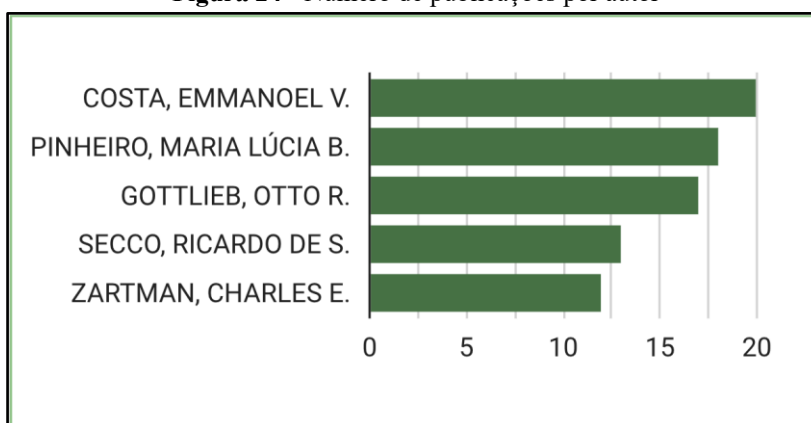
Nota: Elaborado utilizando a ferramenta *Flourish* no primeiro semestre de 2026.

Com isso, universidades brasileiras superaram as estrangeiras no ranqueamento, o que pode indicar um crescente reconhecimento e valorização dos recursos naturais brasileiros por parte da própria comunidade acadêmica nacional.

No próximo subtópico (4.2.2), apresenta-se a relação de publicações com base nos autores.

#### 4.1.2 Número de publicações por autor

Cinco autores se destacam entre os envolvidos nas publicações, conforme ilustra a Figura 24: Costa, Emmanoel V (com 20 artigos); Pinheiro Maria Lúcia B (com 18 publicações); Gottlieb, Otto R (com 17 artigos); Secco, Ricardo de S (13 artigos); e Zartman, Charles E (12 artigos).

**Figura 24** - Número de publicações por autor

Fonte: A autora (2025).

Nota: Captura de tela do LS, realizado no segundo semestre de 2025.

A análise da colaboração internacional e da produtividade dos autores revela padrões distintos de atuação no conjunto de dados estudado. Os três primeiros autores listados – Costa, Pinheiro e Gottlieb – demonstram um extenso e diversificado alcance geográfico em suas publicações, estabelecendo colaborações com instituições em uma ampla gama de países. Essa rede de pesquisa se estende por nações na América do Sul, como Peru e Colômbia; no Oriente Médio, como Paquistão e Irã; e na Europa, englobando Portugal e Reino Unido. Evidentemente, essas parcerias internacionais se somam às colaborações mantidas dentro do próprio Brasil, sugerindo uma perspectiva global em suas agendas de pesquisa.

Em contraste, o quarto autor, Secco, concentra suas parcerias de publicação estritamente no âmbito nacional, estabelecendo laços com instituições sediadas em diferentes regiões brasileiras, especificamente as regiões Norte e Sul. Essa concentração sugere um foco em questões de relevância regional ou um maior engajamento com a rede de pesquisa interna do país. Por fim, o quinto autor, Zartman, apresenta um padrão de colaboração intermediário, com parcerias internacionais relevantes que incluem o Canadá e a China, além da manutenção de publicações conjuntas com instituições brasileiras.

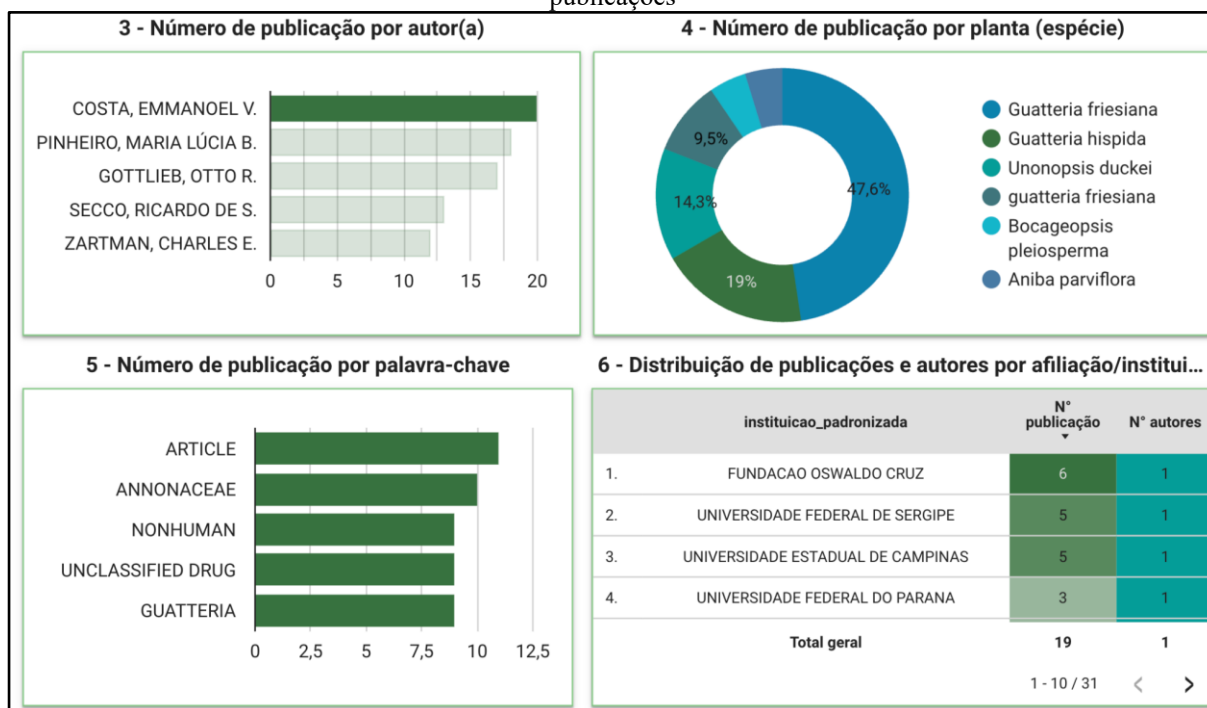
No que tange à produtividade, Emmanoel V. Costa se destaca como o autor mais prolífico no conjunto de dados em análise, sendo responsável pelo maior número de artigos publicados. A profundidade de sua pesquisa é evidenciada pela menção a seis espécies distintas em seus trabalhos (conforme detalhado na Figura 25). Adicionalmente, um exame mais detalhado do foco de pesquisa de Costa revela uma forte concentração na espécie *Guatteria friesiana*. Esta espécie é a mais mencionada em seus artigos, representando expressivos 47,6% do total de menções de espécies feitas por este autor. Tal dado sugere que *Guatteria friesiana*

pode ser um organismo modelo ou de particular interesse central na linha de pesquisa desenvolvida por Emmanoel V. Costa.

A análise detalhada da porção inferior esquerda da figura revela as palavras-chave que apresentam o maior índice de indexação nos artigos científicos de autoria do pesquisador em foco. Essas palavras-chave indicam as áreas temáticas centrais e os objetos de estudo mais recorrentes em sua produção acadêmica. As cinco palavras-chave de maior destaque são: "artigo", que denota a natureza do material indexado; as categorias botânicas "annonaceae" e "guatteria", sugerindo uma especialização em estudos de plantas pertencentes a essas famílias e gêneros, respectivamente; e os termos "nonhuman" e "unclassified drug", que apontam para a utilização de modelos não-humanos em experimentos e o foco em substâncias de origem natural ou ainda não formalmente classificadas como medicamentos, o que é frequentemente associado à pesquisa em produtos naturais e etnofarmacologia.

Complementarmente, no lado direito da figura, é apresentada a distribuição das instituições de afiliação associadas às publicações que contêm essas palavras-chave. Este dado é crucial para entender a rede de colaboração e o impacto institucional da pesquisa do autor. Entre as instituições parceiras, a Fundação Oswaldo Cruz (Fiocruz) se sobressai de forma notável, sendo associada a 6 (seis) ocorrências dessas publicações. Essa proeminência da Fiocruz sublinha uma forte colaboração ou afiliação institucional do autor com uma das mais importantes e tradicionais instituições de ciência e tecnologia em saúde do Brasil, reforçando o contexto de pesquisa em saúde, fármacos e bioprospecção, consistente com as palavras-chave identificadas.

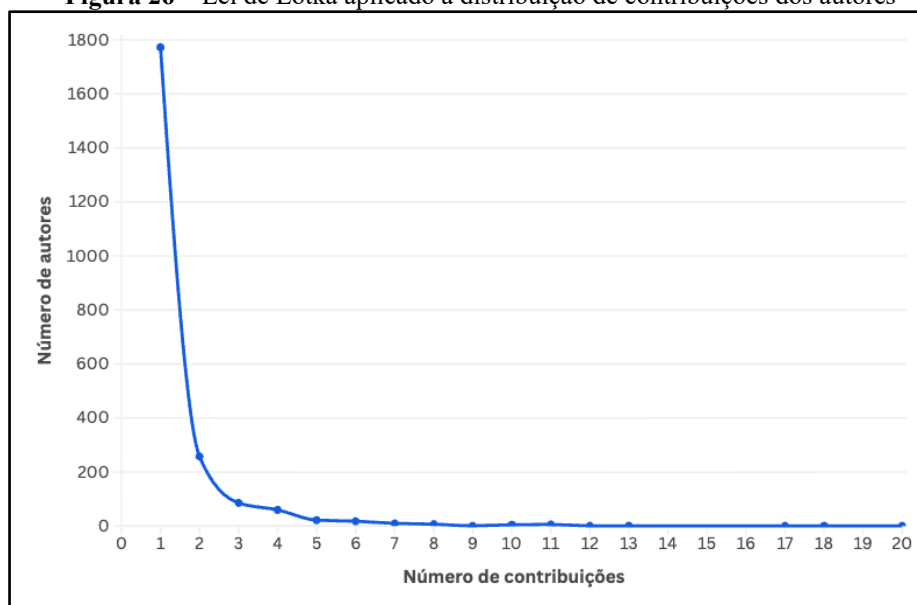
**Figura 25** - Gráficos do protótipo do painel com filtro aplicado referente ao autor com maior número de publicações



Fonte: A autora (2025).

Nota: Captura de tela do LS realizado no segundo semestre de 2025.

Com base na contagem de autores por publicação, a Lei de Lotka que descreve a produtividade de autores, se corrobora no presente conjunto de dados. Esta lei aponta que muitos autores produzem pouco artigos sobre um determinado tema, enquanto poucos autores produzem muito sobre a mesma temática. Narin e Breitzman (1995) especificam que para cada 100 cientistas que publicam um artigo, a distribuição de produtividade sugere que, aproximadamente, 25 ( $100/2^2$ ) produzirão dois artigos, 11 ( $100/3^2$ ) produzirão três, e assim por diante, com apenas um cientista nesse grupo responsável por dez artigos.

**Figura 26** – Lei de Lotka aplicado à distribuição de contribuições dos autores

Fonte: A autora (2025).

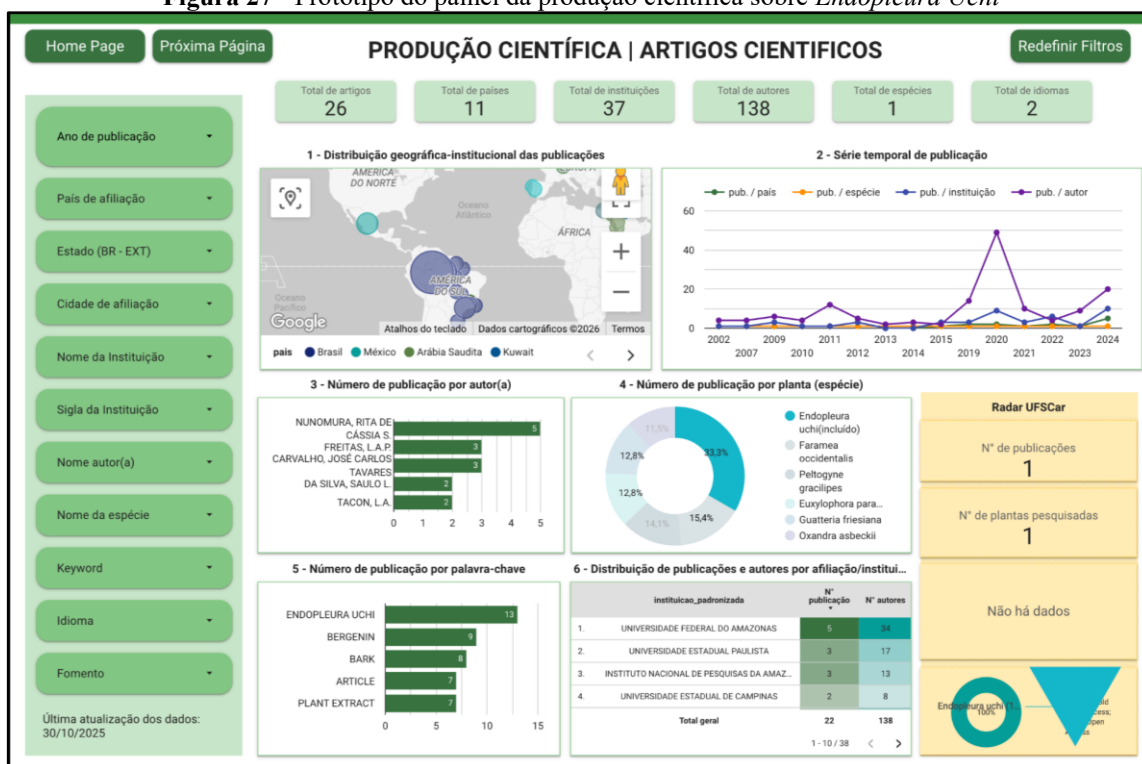
Nota: Elaborado utilizando a ferramenta *Flourish*, no segundo semestre de 2025.

Confirma-se que a distribuição de publicações sobre o tema é bastante desigual entre os autores: mais de 200 autores possuem poucas publicações (entre 1 e 4 cada), enquanto um número reduzido (menos de 100 autores) é responsável por um volume significativamente maior de artigos (acima de 10 cada).

#### 4.1.3 Número de publicação por planta

Duas plantas se destacaram com maior volume de publicações. A primeira delas: *Endopleura uchi*, ocorrendo em 33,3% dos artigos (Figura 26). Esta planta reflete o interesse de diversas instituições e países/regiões tais como Brasil, México, Arábia Saudita, Haiti, Portugal, Reino Unido, Alemanha, Holanda, Egito e Indonésia. Suas publicações iniciaram em 2002 e no último ano analisado (2024) pelo menos 20 autores afiliados à 10 instituições de 5 países a mencionaram em suas pesquisas.

Figura 27 - Protótipo do painel da produção científica sobre *Endopleura Uchi*



Fonte: A autora (2025).

Nota: Captura de tela do LS realizado no segundo semestre de 2025.

Ao que concerne o Brasil 32,2 % de suas publicações mencionam esta espécie. O destaque de instituições com relação à esta planta fica com a UFAM (5 publicações), Universidade Estadual Paulista (UNESP) e INPA (com 3 publicações cada). Cabe ressaltar ainda que no radar da UFSCar, esta planta (*Endopleura uchi*) é mencionada em ao menos 1 publicação.

Esta espécie tem seu potencial uso comercial associado ao seu fruto (Figura 27), sua madeira e óleo extraído. Os frutos com finalidade alimentícia, como para sorvete. Uso também mencionado para farmacopéia popular, como para tratamento uterinos ou para baixar níveis de colesterol plasmático, assimilados à efeitos anti-inflamatórios. Sua madeira também pode ser utilizada para construção civil ou naval (Cristina-Silva; Faria, 2024).

A diversidade de países e instituições pesquisando-a, sugere que suas aplicações são de interesse da comunidade científica além do Brasil, e estão no radar acadêmico desde 2002.

**Figura 28** – Frutos provenientes da árvore *Endopleura Uchi*



Fonte: Reprodução da *Wikipédia*.

Nota: Busca realizada em dezembro de 2025.

A segunda planta com maior número de menções em artigos, entretanto, não correspondem a publicações vinculadas às instituições brasileiras (Figura 28). Os países que a pesquisam, no conjunto de dados presente, são Cuba, Estados Unidos da América, França, Panamá e Brunei. Portanto questiona-se se tal planta não está no radar acadêmico-científico brasileiro, ou se trata-se de uma lacuna técnica – como pelo fato do conjunto de dados analisado ser da base de dados *Scopus* e não ser especificamente nacional, por exemplo. Pesquisas futuras poderão explorar melhor tais aspectos.

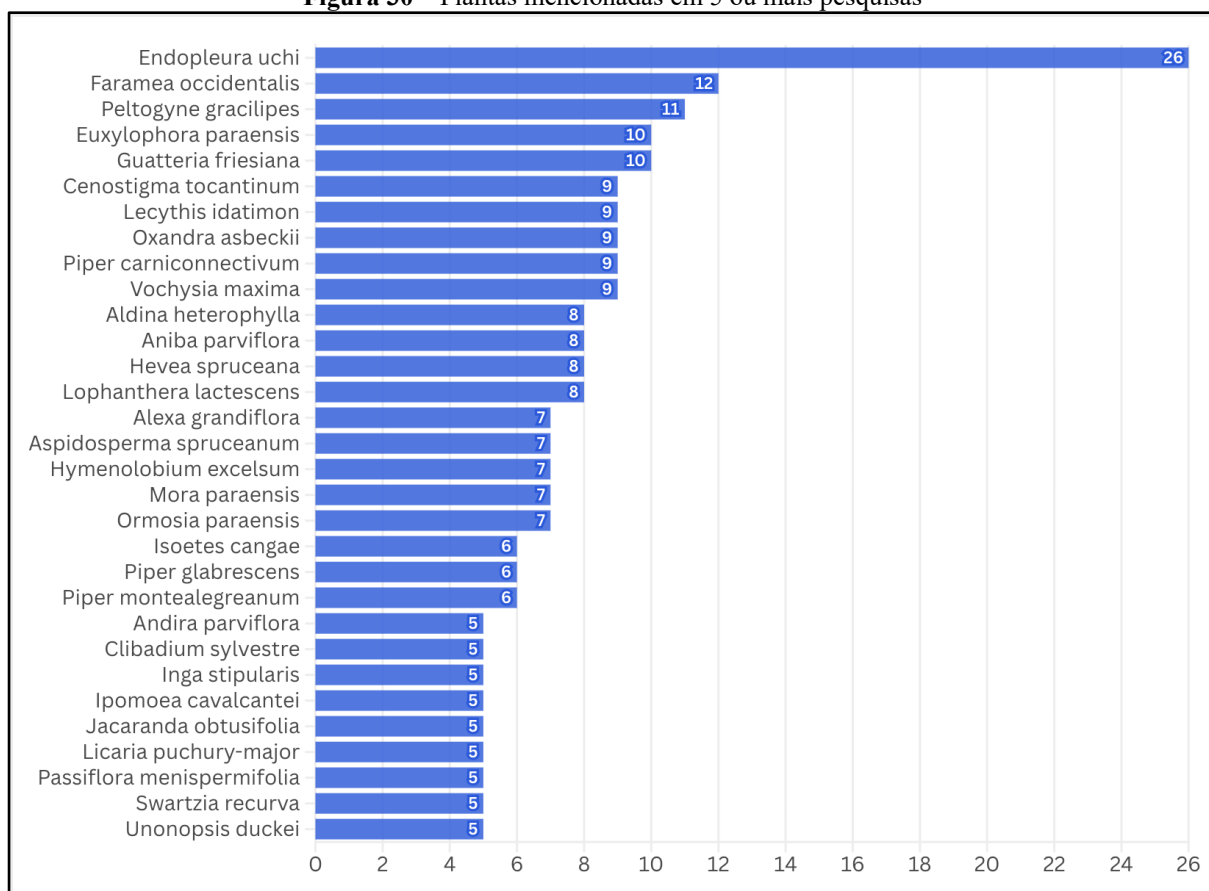
**Figura 29** - Relação de indicadores de publicações sobre a planta *Faramea occidentalis*



Fonte: A autora (2025).

Nota: Captura de tela do LS realizado no segundo semestre de 2025.

Explorar demais plantas (Figura 29) relacionando-as aos indicadores – tais como série temporal de publicações, por instituição, idioma e fomento – poderão oferecer *insights* consideráveis para eventuais análises do cenário científico.

**Figura 30** – Plantas mencionadas em 5 ou mais pesquisas

Fonte: A autora (2025).

Nota: Elaborado utilizando a ferramenta *Flourish* no primeiro semestre de 2026.

No próximo subtópico (4.2.4), apresenta-se a relação de publicações com base nas instituições de afiliação dos autores.

## 4.2 INDICADORES BIBLIOMÉTRICOS DE PRODUÇÃO TECNOLÓGICA E INOVATIVA

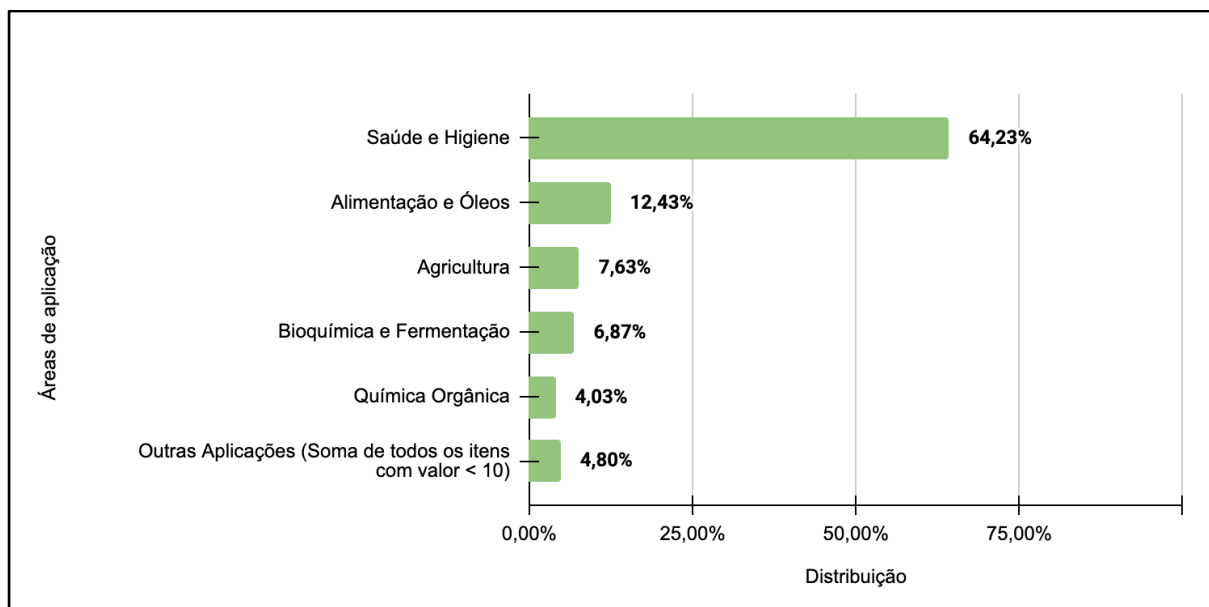
O uso de patentes mostrou-se crucial para análise da dimensão tecnológica e inovativa, para compreender além do reflexo científico (Grupp; Mogege, 2005). O conjunto de dados analisado, focado em invenções que exploram o potencial de espécies endêmicas da Amazônia brasileira, revela uma dimensão global de interesse. Este acervo engloba um total de 153 patentes distintas, que se organizam em 115 famílias de patentes. As publicações iniciam em 1992 e vão até 2025, com pico notável em 2017 (16 publicações).

#### 4.2.1 Áreas tecnológicas das patentes

A análise detalhada da aplicação tecnológica das patentes (Figura 33) revelou uma forte concentração no campo da saúde e higiene, que representa a maioria com 64,23% do total. Essa predominância sublinha a importância da tecnologia em pautas relacionadas ao bem-estar e à prevenção de doenças, abrangendo uma vasta gama de invenções, desde novos compostos e produtos de limpeza até dispositivos médicos e tecnologias de diagnóstico. Em seguida, a aplicação em alimentação e óleos surge como o segundo maior foco de inovação, correspondendo a 12,43% das patentes. Este campo engloba desenvolvimentos em processamento, conservação e melhoria nutricional de alimentos, bem como a extração e o uso de óleos para diversas finalidades, reforçando a relevância da tecnologia para a segurança alimentar e a indústria de ingredientes.

Outro setor de destaque é o da agricultura, com uma participação de 7,63%. As inovações aqui são cruciais para o manejo de culturas e o controle de pragas, abrangendo o desenvolvimento de biocidas mais eficazes e sustentáveis, repelentes ou atrativos de pragas e reguladores de crescimento de plantas, visando otimizar a produtividade e a saúde vegetal. Adicionalmente, os achados tecnológicos também se estendem para áreas da bioquímica (6,87%) e química orgânica (4,03%). Nesses campos, o foco está nos processos de extração de substâncias ativas, na criação de novas composições químicas e na fabricação de insumos essenciais, como fertilizantes, que sustentam tanto a agricultura quanto outras indústrias químicas.

Por fim, uma parcela minoritária, porém diversificada, das patentes abarca outras aplicações menos frequentes. Estas incluem inovações em engenharia mecânica (por exemplo, sistemas de iluminação e aquecimento otimizados para processos como a secagem), operações de processamento (especialmente no tratamento de materiais plásticos) ou no campo da física (com invenções focadas em medição, testes e em aplicações específicas de tecnologias de informação e comunicação - TIC), demonstrando potencial transversalidade da aplicação tecnológica em múltiplos setores industriais.



Fonte: A autora (2026).

A análise da distribuição das patentes revelou um foco predominante na Seção A (Necessidades Humanas), que concentra um total expressivo de 773 ocorrências. Dentro desta seção, a Classe A61 (Ciência médica ou veterinária; higiene) é a mais relevante, somando 589 ocorrências, o que sublinha a orientação das invenções para o setor da saúde e bem-estar. Ao aprofundar a investigação, a Subclasse A61K (Preparações para finalidades médicas, odontológicas ou de higiene pessoal) se destaca com 405 ocorrências, indicando que a maior parte da atividade patentária está direcionada para o desenvolvimento de formulações e composições. Este dado é crucial para compreender as aplicações práticas das inovações.

O Grupo A61K 36 emerge como o de maior interesse, referente a preparações medicinais contendo materiais de constituição indeterminada derivados de algas, líquens, fungos ou plantas, ou derivados dos mesmos, por exemplo, medicamentos tradicionais à base de ervas. A relevância deste grupo, que trata de produtos de origem natural, aponta para uma forte tendência de pesquisa e desenvolvimento baseada em recursos biológicos. Dentro do Grupo A61K 36, o subgrupo A61K 36/18, que abrange especificamente as *Magnoliophyta* (angiospermas), demonstra ser o mais explorado. Este achado sugere que as plantas com flores são a principal fonte de materiais para as composições medicinais e de higiene pessoal patenteadas.

Em suma, a predominância de patentes nas classes e grupos relacionados à saúde, higiene e, especificamente, a formulações à base de extratos vegetais (angiospermas), permite afirmar que mais de 60% das patentes analisadas se relacionam diretamente com potenciais aplicações voltadas à indústria farmacêutica ou cosmética. Esta concentração evidencia o forte

interesse em capitalizar o potencial terapêutico e cosmético de compostos naturais, alinhando a atividade de propriedade intelectual com as demandas e tendências de mercado para produtos baseados em recursos botânicos.

#### 4.2.2 Patente por autoria

Um achado particularmente notável e que levanta profundas questões é a origem dessas invenções: há uma participação comprovada de 17 países diferentes. Essa ampla diversidade de origens geográficas é alarmante, especialmente considerando a natureza inerentemente endêmica e nativa dessas plantas, cuja ocorrência natural está restrita, ou pelo menos majoritariamente concentrada, no bioma amazônico brasileiro.

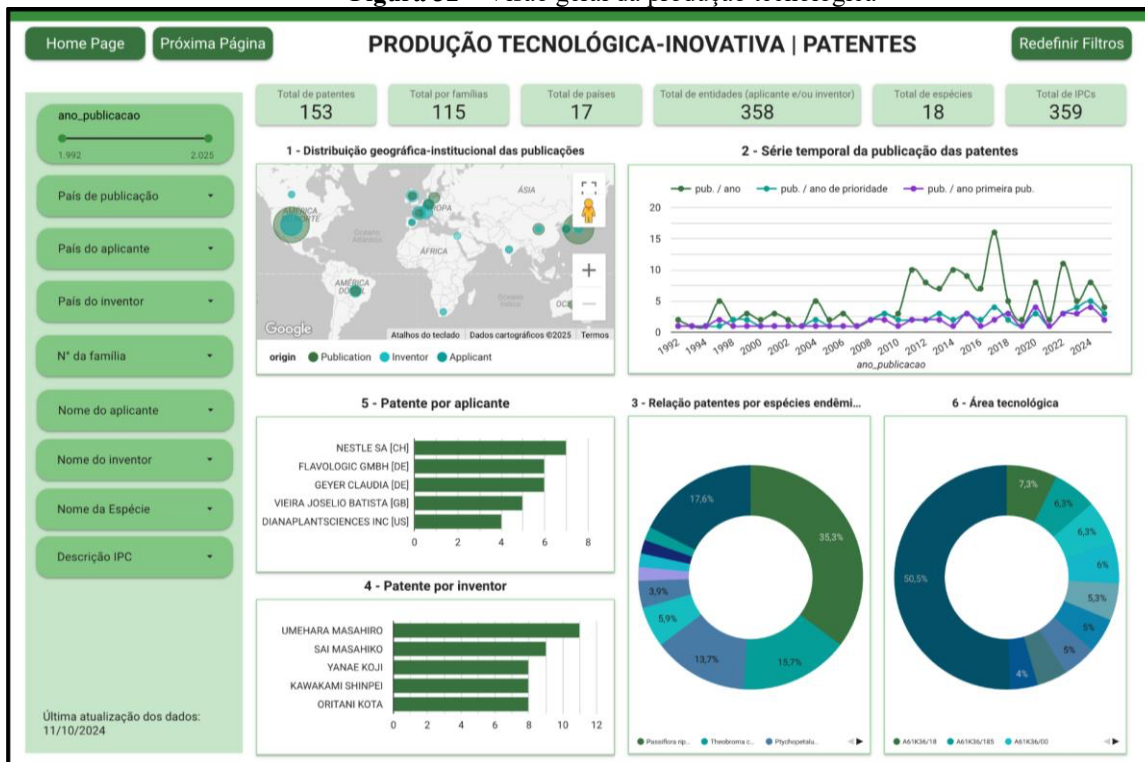
Essa dispersão geográfica das nações depositantes de patentes suscita preocupações legítimas sobre o real grau de interesse e a extensão da exploração econômica e científica dessas espécies por entidades estrangeiras. O volume e a variedade de patentes fora do Brasil sugerem um esforço significativo de pesquisa e desenvolvimento internacional capitalizando sobre a biodiversidade brasileira.

Consequentemente, torna-se imperativo e urgente um debate aprofundado sobre os impactos multifacetados desse cenário. É crucial discutir os riscos e a materialização de fenômenos como a biopirataria, caracterizada pela apropriação não autorizada de recursos biológicos; o extrativismo não regulamentado ou predatório dos recursos naturais; e a apropriação indevida dos conhecimentos tradicionais associados a essas plantas, desenvolvidos e preservados pelas comunidades locais e povos indígenas ao longo de gerações. Além disso, a discussão deve centrar-se na garantia de um mecanismo de repartição de benefícios justo e equitativo, assegurando que o valor gerado por essas patentes retorne de forma tangível e significativa para as comunidades detentoras desse conhecimento ancestral e dos próprios recursos.

Adicionalmente, a análise detalhada das patentes revela um ecossistema complexo de atores envolvidos. Percepções preliminares (as quais serão detalhadas na Figura 34) indicam uma significativa diversidade de empresas, instituições de pesquisa e inventores individuais que participam da criação dessas invenções que mencionam as espécies amazônicas. Este universo de inventores e cessionários é composto por um total impressionante de 358 entidades diferentes. Essa multiplicidade de *players* reforça a necessidade de monitoramento e regulação

eficazes, dada a escala e a complexidade dos interesses envolvidos na exploração da flora amazônica.

Figura 32 – Visão geral da produção tecnológica



Fonte: A autora (2025).

Nota: Captura de tela do LS realizado no segundo semestre de 2025.

A análise da relação de aplicantes revela uma predominância de entidades estrangeiras no cenário de patentes estudado. Os Estados Unidos da América destacam-se com 28 aplicantes, evidenciando uma forte presença de inovação e pesquisa proveniente deste país. O continente europeu também apresenta uma contribuição significativa, com países como a Alemanha (9 aplicantes) e a Suíça (14 aplicantes) demonstrando um interesse notável na área.

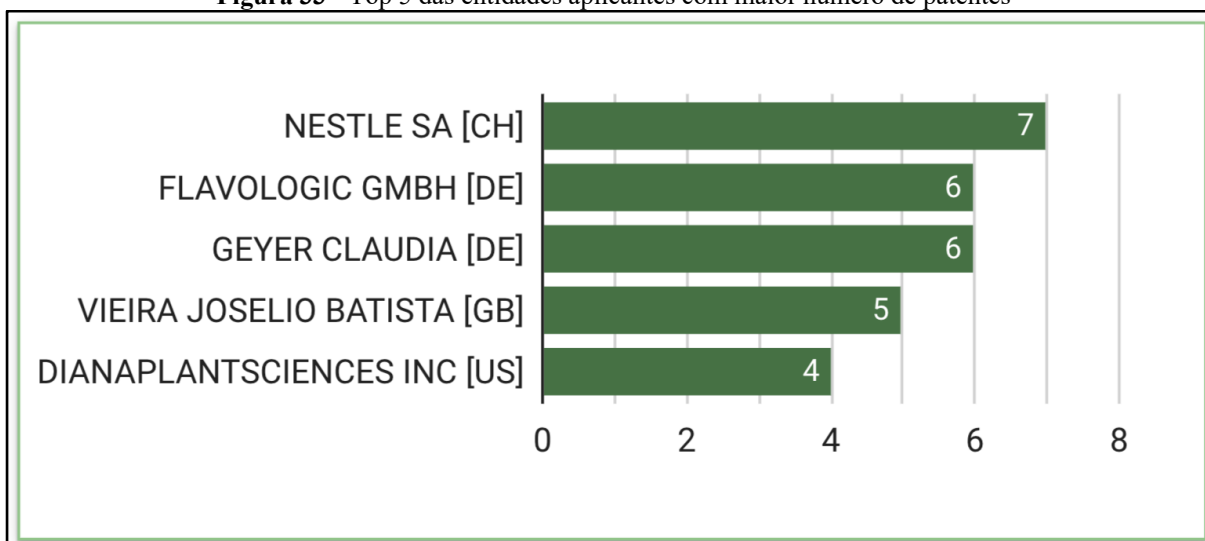
Ao examinar o Top 5 de aplicantes, observa-se que todas as entidades ranqueadas são de origem estrangeira, sublinhando o caráter internacional da liderança no registro de patentes neste campo (conforme ilustrado na Figura 35).

No topo da lista, a Nestlé SA, com sede na China, ocupa o primeiro lugar com um total de 7 patentes registradas, indicando um foco estratégico em pesquisa e desenvolvimento na região asiática. O segundo e terceiro lugares do pódio são compartilhados por empresas alemãs, ambas com 6 patentes cada: a Flavologic GmbH e a Geyer Claudia. Essa alta representatividade alemã no topo do ranking sugere um ambiente propício à inovação e um investimento substancial em tecnologia na Alemanha.

A quarta posição é novamente preenchida por uma entidade alemã, a Vieira Joselito Batista, que detém 5 patentes. Por fim, fechando o Top 5, encontra-se a Dianaplantsciences INC, dos Estados Unidos, com 4 patentes. A presença de empresas alemãs em três das cinco primeiras posições reforça a proeminência deste país no panorama de patentes, enquanto a inclusão de entidades dos EUA e da China ressalta a natureza global da competição e da colaboração em inovação neste setor específico.

A concentração de patentes nessas entidades não apenas mapeia os principais players no mercado, mas também pode indicar tendências de pesquisa e desenvolvimento que guiarão a inovação futura. A diversidade geográfica dos aplicantes líderes sugere uma distribuição heterogênea da excelência tecnológica, com a Europa e a América do Norte, junto com a emergente força da China, desempenhando papéis centrais.

**Figura 33** - Top 5 das entidades aplicantes com maior número de patentes



Fonte: A autora (2025).

Nota: Captura de tela do LS realizado no segundo semestre de 2025.

A análise detalhada da distribuição de patentes revela um panorama internacional no que concerne à autoria das invenções. É um dado notável a ausência de inventores de nacionalidade brasileira entre os mapeados neste levantamento específico. A liderança em número de patentes está com Geyer Claudia, proveniente da Alemanha, que se destaca com um total de seis patentes registradas. Logo em seguida, está Vieira Joselito Batista, da Suíça, com um portfólio de cinco patentes.

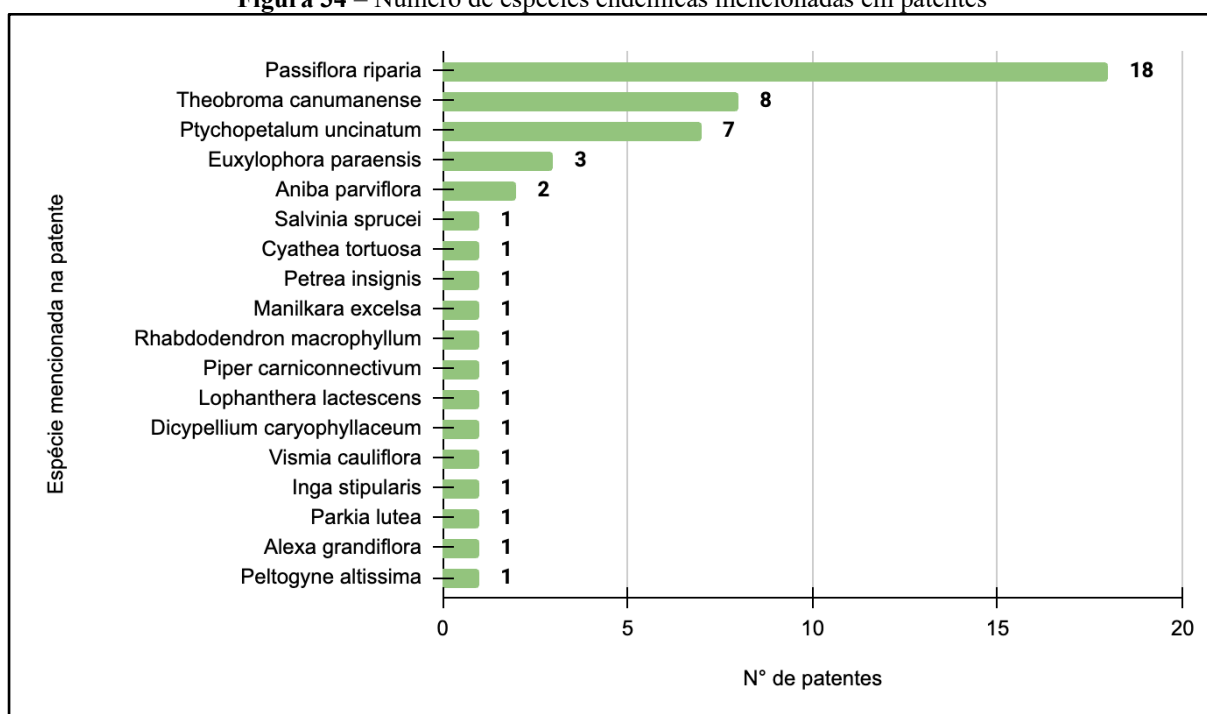
Embora o Brasil não tenha inventores na lista principal, o país demonstra participação significativa como depositante. Identificou-se um total de 13 patentes publicadas que possuem depositantes brasileiros, as quais estão agrupadas em sete famílias de patentes distintas. Entre

os depositantes brasileiros de destaque estão instituições e empresas de renome nacional, como a Universidade Federal de Minas Gerais (UFMG), uma das principais instituições de ensino e pesquisa do país, e a Pele Nova Biotecnologia S.A., uma empresa que atua na área de biotecnologia. Essa presença como depositante sublinha o esforço nacional em pesquisa e desenvolvimento, mesmo que a figura do inventor principal mapeado seja de origem estrangeira. A concentração de patentes em inventores estrangeiros sugere uma intensa atividade de pesquisa e inovação em outros países, indicando a necessidade de fortalecer a participação brasileira no polo de criação e registro de novas tecnologias.

#### 4.2.3 Patentes por espécies

A Figura 36 demonstra uma variedade de plantas de interesse tecnológico. No entanto, o número de espécies mencionadas (18) representa uma fração muito pequena (cerca de 0,9%) do total de 1975 plantas pesquisadas. Essa baixa representatividade sugere algumas possibilidades: (1) há muitas plantas com potencial tecnológico sendo negligenciadas; (2) a maioria das plantas de fato não possui interesse comercial; ou (3) em um cenário mais preocupante, a exploração dessas plantas pode ocorrer de forma não registrada, o que levanta a possibilidade de biopirataria.

**Figura 34** – Número de espécies endêmicas mencionadas em patentes



Fonte: A autora (2025)

A espécie com maior número de menções nas patentes é a *Passiflora riparia* (18 ocorrências). Esta planta (Figura 37) é subordinada ao grupo de angiosperma e é também conhecido como “maracujá-suspiro” (Jardim Botânico do Rio de Janeiro, 2019).

**Figura 35** – Espécie endêmica com maior número de menções nas patentes analisadas (*Passiflora riparia*)



Fonte: Reprodução do *site* do Jardim Botânico do Rio de Janeiro (2026)  
Nota: Autor da imagem Luiz Otávio Adão<sup>18</sup>

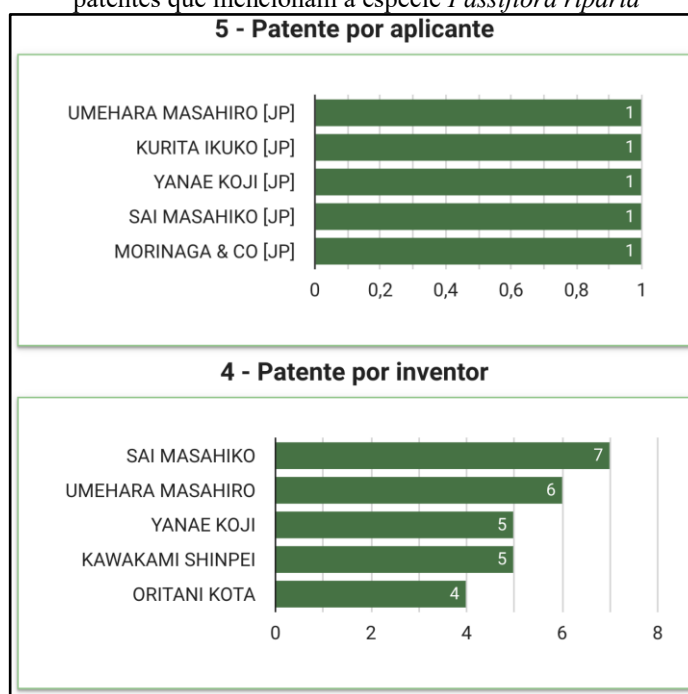
A análise do cenário de patentes relacionadas à planta em questão revela um foco geográfico e temporal bem definido. O Japão emerge como o país de referência, sendo o principal responsável pelas reivindicações de patentes e detentor dos principais inventores associados a esta espécie vegetal (Figura 38). O período de atividade de patentes abrange cerca de uma década, com as publicações iniciando em 2012 e se estendendo até 2020. A curva de depósito atingiu seu ápice no ano de 2016, com o registro de 5 novas patentes.

De um total de 18 patentes identificadas neste levantamento, uma fração significativa, precisamente 16,7% delas, possui a classificação internacional A61K31/05. Esta classificação, no sistema IPC (*International Patent Classification*), é altamente relevante e fornece um forte

<sup>18</sup> Imagem publicada por Ana Carolina Mezzonato-Pires, em 2019. Disponível em: <https://floradobrasil.jbrj.gov.br/FB24154>. Acesso em: 18 jan. 2026.

indício da natureza das aplicações pretendidas. A hierarquia desta classificação desdobra-se da seguinte forma: a Seção A (Necessidades Humanas) indica o campo geral de aplicação; a Subclasse A61 (Ciência médica ou veterinária, e higiene) foca em saúde e bem-estar; o Grupo A61K (Preparações para finalidades médicas, odontológicas ou de higiene pessoal, podendo ser misturas para uma composição ou processo de tratamento utilizando esta composição) especifica o uso de composições ou preparações; e o Subgrupo A61K 31/05 (Preparações medicinais contendo ingredientes ativos orgânicos > Fenóis) detalha a composição química dos ingredientes ativos, apontando para a presença de compostos fenólicos nesta planta como a base para a atividade reivindicada.

**Figura 36** – Principais aplicantes e inventores com relação às patentes que mencionam a espécie *Passiflora riparia*



Fonte: A autora (2026).

Nota: Captura de tela do protótipo do painel criado no *Looker Studio*.

A predominância desta classificação no conjunto de patentes, especialmente aquelas originárias do Japão, constitui um forte indicativo do potencial comercial da planta voltado especificamente a aplicações farmacêuticas. Embora o interesse do Japão seja notavelmente o mais proeminente, é importante ressaltar que há também publicações de patentes utilizando esta planta e detentoras desta mesma classificação (A61K31/05) originadas nos Estados Unidos, sugerindo um interesse internacional mais amplo, mas com a liderança tecnológica e de mercado claramente estabelecida no Japão.

Relações desta espécie também são observadas na área tecnológica A61K 8/34, que abrange o uso em cosméticos ou higiene pessoal de compostos orgânicos, como álcool. Essa classificação está presente em 12,1% das patentes analisadas, sendo a classe A61K a predominante entre as diversas classificações encontradas.

A segunda espécie com maior ocorrência nos registros é a *Theobroma canumanense* (15,7%). Neste cenário, há uma maior diversidade de países envolvidos como aplicantes da China (Nestlé com 4 publicações), Suíça (reivindicado por um inventor certamente pessoa física chamado Vieira Joselito Batista, com 3 publicações) e Estados Unidos (Mars Inc, com 2 patentes, contando inclusive com inventores de referência brasileira). As patentes relacionadas à esta planta datam de 2012 até 2024 e teve o maior número da série histórica em 2022 com 3 patentes.

De acordo com Silva, Venturieri e Figueira (2004), a *Theobroma canumanense* compõe o acervo de germoplasma amazônico, sendo uma espécie fundamental para a caracterização genética e potencial hibridação no melhoramento do cacauero. A análise da principal área tecnológica da planta confirma a abordagem dos autores, visto que mais de 62% dos registros de IPC (5 de 8) estão na classe A23G, que é especificamente relacionada ao cacau e seus substitutos.

A terceira espécie mais mencionada nos registros é a *Ptychopetalum uncinatum*, com um total de 7 ocorrências. Esta frequência significativa indica sua relevância no contexto estudado. A importância desta espécie é corroborada pela literatura científica, como demonstrado por Sá *et al.* (2017), que a identificaram como uma fonte rica e reconhecida de compostos fenólicos bioativos. Estes compostos fenólicos são os principais responsáveis pelas propriedades farmacológicas da *Ptychopetalum uncinatum*, o que justifica sua ampla utilização na composição de diversos medicamentos fitoterápicos comerciais. Dada a sua aplicação terapêutica, a caracterização química rigorosa desta espécie é um requisito essencial para garantir a qualidade e a eficácia dos produtos.

Dada a natureza quantitativa desta pesquisa, as demais espécies endêmicas mapeadas emergem como alvos prioritários para estudos futuros. É fundamental aprofundar o conhecimento sobre seus potenciais benefícios, investigando quem as mencionam em patentes, quais interesses impulsionam sua aplicação tecnológica e, crucialmente, como garantir a distribuição de benefícios para as comunidades locais de onde são originárias.

Outra linha de investigação essencial é a análise das políticas públicas e dos mecanismos de proteção e conservação associados a cada uma dessas espécies. Tal análise

permitirá avaliar o nível de exploração desses recursos naturais, exclusivos da Amazônia brasileira, e determinar a necessidade de ações mitigatórias e reparadoras no contexto atual.

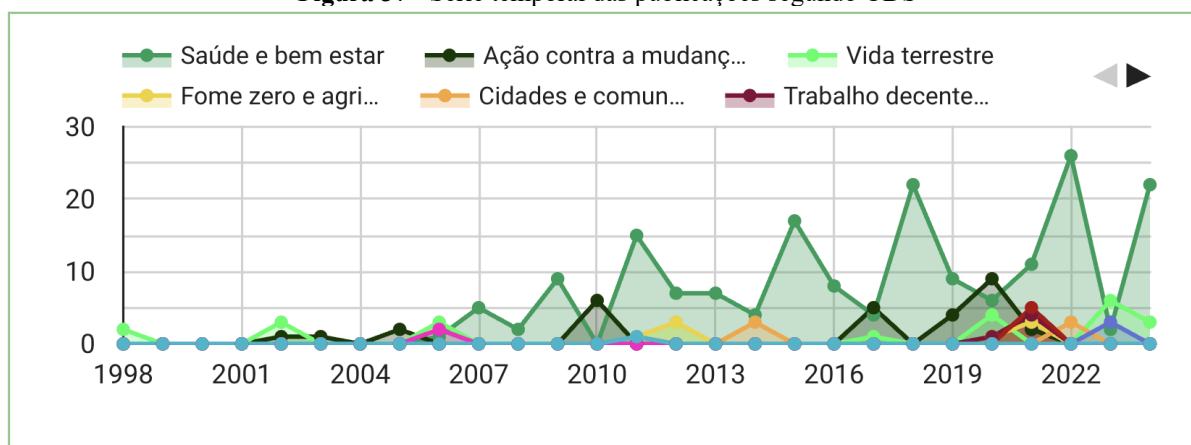
#### 4.3 INDICADORES DE SUSTENTABILIDADE

##### 4.3.1 Indicadores de sustentabilidade baseados na produção científica

Na dimensão científica, observa-se uma notável convergência de publicações com os ODS. De modo geral, a variedade de ODS abordados ressalta o potencial das plantas endêmicas amazônicas para contribuir com os objetivos de desenvolvimento sustentável em escala global.

Ao longo dos anos, as publicações científicas têm demonstrado uma crescente relação com os Objetivos de Desenvolvimento Sustentável (ODS). O ODS "Saúde e bem-estar" se destaca como o mais frequente na maioria das publicações, conforme ilustrado na Figura 30. No entanto, houve exceções em certos anos: em 2010, 2009, 2017 e 2020, o ODS "Ação contra a mudança global do clima" foi o predominante, e em 2023, o ODS "Vida terrestre" sobressaiu.

**Figura 37** - Série temporal das publicações segundo ODS

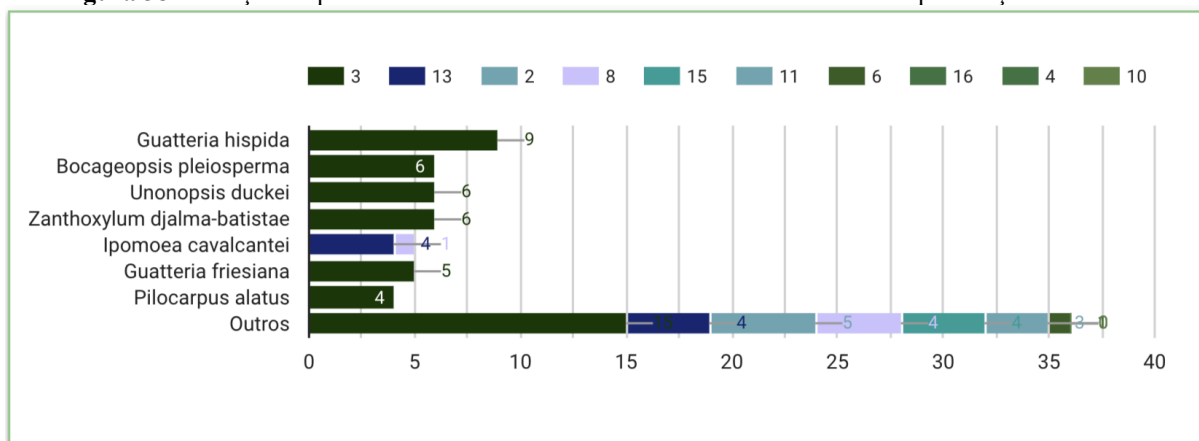


Fonte: A autora (2025).

Nota: Captura de tela do LS realizado no segundo semestre de 2025.

Adicionalmente, a relação direta de cada planta com cada ODS é evidente nas publicações científicas analisadas. Conforme ilustrado na Figura 31, há fortes indícios de que espécies como *Guatteria hispida*, *Bocageopsis pleiosperma* e *Unonopsis duckei*, entre outras, estão associadas ao ODS 3 (Saúde e bem-estar). Isso sugere que esses recursos naturais possuem um grande potencial para contribuir com o objetivo de garantir uma vida saudável e promover o bem-estar para todas as pessoas, em todas as faixas etárias.

**Figura 38** - Relação de plantas endêmicas da Amazônia alinhadas aos ODS nas publicações científicas



Fonte: A autora (2025).

Nota: Captura de tela do LS realizado no segundo semestre de 2025.

A classificação de risco de extinção das plantas, conforme os critérios internacionais da União Internacional para a Conservação da Natureza (IUCN), e mapeadas com base na Lista disponibilizado pelo CNCFlora, é fundamental para direcionar esforços de conservação, e os dados levantados nas publicações científicas revisadas revelaram a seguinte distribuição, em ordem decrescente de frequência de menção, conforme Figura 32.

A categoria de LC (Pouco Preocupante – *Least Concern*), que representa as espécies com menor risco de extinção, foi a mais frequentemente abordada, sendo mencionada em 137 publicações. Essa alta frequência pode refletir tanto uma maior abundância e distribuição dessas espécies quanto um foco de pesquisa em espécies mais comuns para estudos ecológicos gerais.

Em seguida, as espécies classificadas como EN (Em Perigo – *Endangered*), que enfrentam um risco muito alto de extinção na natureza, apareceram em 66 artigos, demonstrando um foco significativo da pesquisa em espécies que requerem atenção urgente de conservação. A investigação sobre estas espécies é crucial para entender as ameaças específicas e desenvolver estratégias de manejo eficazes.

As espécies classificadas como DD (Dados Insuficientes – *Data Deficient*) representam uma lacuna de conhecimento significativa, sendo mencionadas em 40 publicações por não possuírem informações suficientes para uma avaliação de risco precisa. Essa evidência sugere que pesquisas futuras devem priorizar a coleta de dados populacionais e de distribuição para viabilizar uma reclassificação de risco mais acurada. Adicionalmente, a aplicação de análise de conteúdo nos artigos que citam essas espécies pode ser uma estratégia útil para

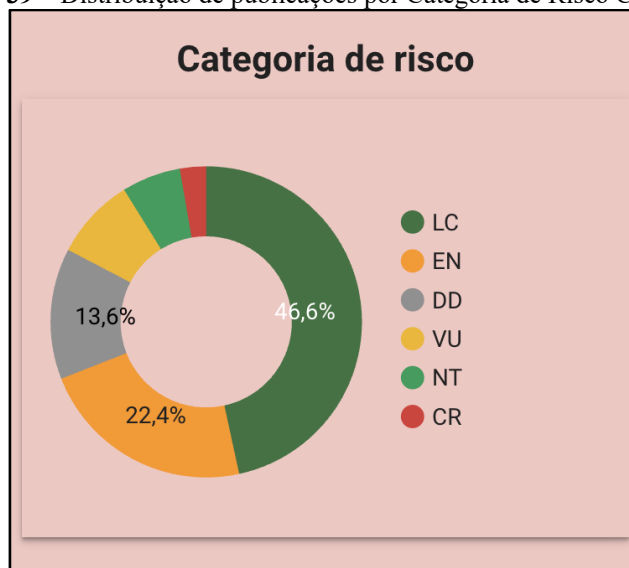
auxiliar o CNCFlora (Centro Nacional de Conservação da Flora) nas definições de risco e, assim, contribuir para a classificação dessas plantas e suprir a falta de dados.

As espécies na categoria VU (Vulnerável – *Vulnerable*), que enfrentam um alto risco de extinção a médio prazo na natureza, foram registradas em 25 publicações. O número, embora menor que as categorias EN e LC, ainda aponta para um grupo de espécies sob pressão, cujos estudos são importantes para prevenir um agravamento de seu status de conservação.

Em penúltimo lugar, as espécies NT (Quase Ameaçada – *Near Threatened*), que estão próximas de serem classificadas nas categorias de ameaça, foram identificadas em 18 publicações. Para estas, o monitoramento contínuo e a pesquisa preventiva são essenciais para evitar que passem para um status de maior risco.

Por fim, a categoria de maior risco, CR (Criticamente em Perigo – *Critically Endangered*), que indica um risco extremamente alto de extinção na natureza, foi a menos abordada, aparecendo em apenas 8 publicações. Este baixo número é particularmente preocupante, pois sugere que as espécies à beira da extinção estão recebendo menos atenção na literatura científica revisada, apesar da sua urgência de conservação ser a mais alta.

**Figura 39** – Distribuição de publicações por Categoria de Risco CNCFlora



Fonte: A autora (2025).

Nota: Captura de tela do LS realizado no segundo semestre de 2025.

A distribuição dessas categorias nas publicações demonstra um panorama da atenção de pesquisa dedicada ao risco de extinção das plantas. Enquanto as espécies LC e EN dominam o foco, o número significativo de espécies DD ressalta a necessidade de mais estudos básicos de taxonomia e ecologia. Além disso, o baixo número de publicações sobre espécies CR indica a urgência de concentrar esforços de pesquisa e conservação nas espécies mais ameaçadas.

Futuras pesquisas poderão utilizar um indicador baseado na comparação de publicações sobre plantas endêmicas, considerando os seguintes cruzamentos de critérios: o período de publicação (2014-2020 ou 2021-2024) e a categoria de ameaça da espécie (conforme classificação de 2014 ou 2021). Os quatro critérios de comparação são: publicações de 2014-2020 baseadas na ameaça de 2014 ou na ameaça de 2021, e publicações de 2021-2024 baseadas na ameaça de 2014 ou na ameaça de 2021.

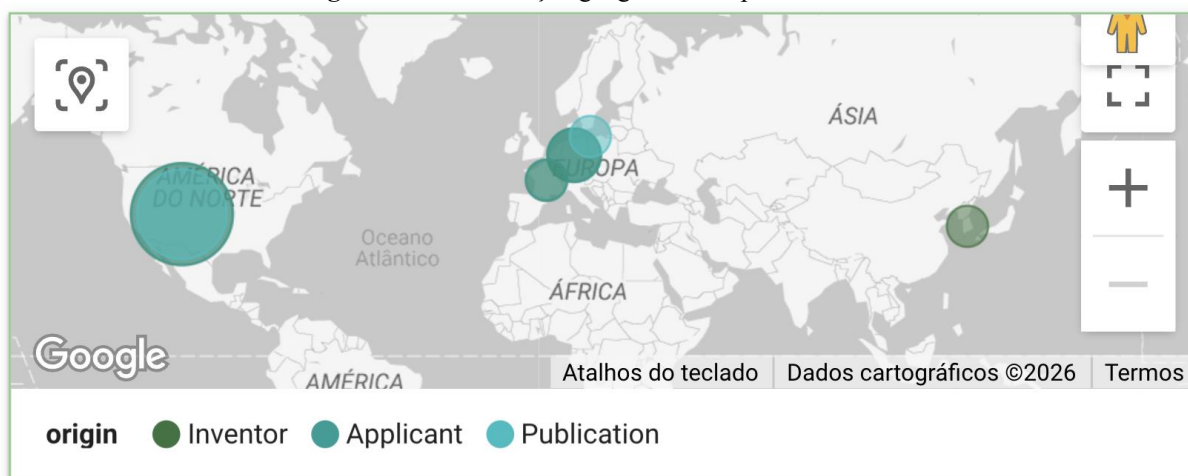
Este indicador tem como objetivo permitir a observação de inferências que relacionam as publicações que citam essas espécies de acordo com os períodos analisados, e verificar se a classificação de uma espécie em determinada categoria de ameaça influencia a produção científica a seu respeito.

O cruzamento dos dados bibliométricos com os classificadores de sustentabilidade (ODS e status de ameaça do CNCFlora) garante a sustentabilidade analítica do painel. Ao demonstrar que 65% dos artigos sobre a flora endêmica se alinham diretamente ao ODS 3 (Saúde e Bem-Estar) ou ODS 15 (Vida Terrestre), e que 30% das plantas mencionadas estão em alguma categoria de risco, o protótipo do painel mostrou-se eficiente no papel de uma ferramenta de diagnóstico estratégico. Essa capacidade de quantificar a produção científica e tecnológica sob a lente de uma agenda global (ODS) e de uma agenda de conservação (CNCFlora) estabelece a relevância contínua da pesquisa, transcendendo a mera contagem e consolidando a contribuição metodológica para a Ciência da Informação.

Um total de 12 patentes, pertencentes a 8 famílias distintas, receberam a classificação de "patente verde". Essa classificação indica que as tecnologias desenvolvidas contribuem para a sustentabilidade e mitigação das mudanças climáticas. O conjunto dessas patentes se distribui por seis países diferentes, o que demonstra uma dispersão geográfica na origem da proteção. Quanto às referências de publicação, as patentes estão registradas em quatro importantes regiões de propriedade intelectual: os Estados Unidos (EUA), a Alemanha, a Organização Europeia de Patentes (OPE/EPO) e a Organização Mundial da Propriedade Intelectual (OMPI/WIPO). Esta diversidade nas regiões de publicação reforça a relevância e o alcance internacional das invenções.

A origem dos depositantes (aplicantes) abrange três regiões: EUA, Alemanha e França. Em contrapartida, a origem dos inventores é um pouco mais diversificada, englobando quatro regiões: EUA, Alemanha, Coreia do Sul e França (Figura 39). A diferença entre a localização dos aplicantes e dos inventores sugere a existência de colaborações internacionais ou a transferência de direitos de propriedade intelectual entre as regiões mencionadas.

**Figura 40** – Distribuição geográfica das patentes verdes



Fonte: A autora (2026)

Nota: Captura de tela do *Looker Studio* realizado em janeiro de 2026

Um ponto de destaque na análise é a ausência de participação do Brasil no conjunto de referências das patentes verdes analisadas, seja como país de origem dos aplicantes, dos inventores ou de publicação principal. Este fato sinaliza uma potencial lacuna na contribuição brasileira para o desenvolvimento e registro de tecnologias classificadas como verdes no contexto da amostra estudada.

No recorte temporal analisado, que se estende de 1992 a 2025, a primeira correlação entre patentes depositadas e a temática da sustentabilidade é registrada em 2002, enquanto a última ocorre em 2020. Ao longo desse período, o ano que se destacou com o maior volume de publicações de patentes que incorporavam uma classificação ou menção à sustentabilidade foi 2017, com o registro de 3 publicações.

Este padrão de surgimento e concentração temporal sugere que a preocupação e a articulação formal de inventos relacionados a plantas endêmicas com os princípios da sustentabilidade é um fenômeno relativamente recente no panorama da propriedade intelectual. Tal percepção é reforçada pela análise do conjunto geral de dados de patentes examinadas, que possui registros desde 1992.

A existência de uma lacuna significativa de 10 anos (de 1992 até o primeiro registro relacionado à sustentabilidade em 2002) evidencia um hiato temporal considerável entre o início do registro de patentes sobre o tema central (plantas endêmicas, no contexto mais amplo) e o momento em que a variável "sustentabilidade" começou a ser formalmente integrada e reconhecida nesses documentos como um fator relevante para a inovação.

### 4.3.2 Indicadores de sustentabilidade baseados em patentes

A análise das áreas tecnológicas relacionadas à sustentabilidade, conforme o Quadro 17, revelou quatro classificações de patentes distintas, cada uma com sua própria hierarquia detalhada na Classificação Internacional de Patentes (IPC) da WIPO e sua correspondência no *Green Inventory* (GI).

A primeira área tecnológica identificada está concentrada na classe A01N 65/00, totalizando 4 patentes, e é intitulada "Alternativas para Pesticidas". Esta classificação pertence à Seção A: Necessidades Humanas (que abrange agricultura) e, na hierarquia da IPC, desdobra-se em: Classe A01 (Agricultura, silvicultura, pecuária, caça, armadilhas, pesca), Subclasse A01N (Conservação – biocidas, repelentes ou atraentes de pragas, ou reguladores de crescimento de plantas) e Grupo A01N 65/00 (Biocidas, repelentes ou atrativos de pragas, ou reguladores de crescimento vegetal contendo material de algas, líquenes, briófitas, fungos multicelulares ou plantas, ou extratos dos mesmos). No *Green Inventory* (GI), A01N 65/00 é descrita como "alternativas para pesticidas" e está subordinada ao campo da agricultura e silvicultura, indicando foco em soluções mais ecológicas e de base biológica para o controle de pragas. (WIPO, 2026)

A segunda área tecnológica sustentável, denominada "Separação de Componentes de Gás de Aterro", está vinculada à Seção B: Operações de Processamentos ou Transporte. Sua hierarquia na IPC é: Classe B01 (Processos ou aparelhos físicos ou químicos em geral), Subclasse B01D (Separação de sólidos), Grupo B01D 53 (Separação de gases ou vapores e afins) e Subgrupo B01D 53/04 (Com absorventes fixos). O enfoque no GI demonstra a relevância ambiental desta tecnologia, que se insere na cadeia de valor da energia limpa e gestão de resíduos, abrangendo: 1. Produção de energia alternativa, 2. Aproveitamento energético de resíduos produzidos pelo homem, 3. Gás de aterro, e 4. Separação de componentes. Este subgrupo refere-se a métodos e equipamentos que utilizam materiais absorventes fixos para a purificação e o aproveitamento do gás de aterro sanitário (biogás) para geração de energia.

A terceira área tecnológica, "Utilização da Energia Solar para Secagem", foca na otimização de processos de secagem através do uso de calor, sendo enquadrada na Seção F: Engenharia Mecânica, Iluminação, Aquecimento, Armas, Desmontagem. A hierarquia da IPC é: Classe F26 (Secagem), Subclasse F26B (Secagem de materiais ou objetos sólidos por remoção de líquido) e Grupo F26B 3/00 (Secagem de materiais ou objetos sólidos por processos que envolvem a aplicação de calor). A sustentabilidade reside na fonte de calor utilizada, com

o GI contextualizando o grupo na hierarquia: 1. Produção de energia alternativa, 2. Energia solar, e 3. Utilização da energia solar para secagem de materiais ou objetos, o que sugere tecnologias que integram coletores solares para reduzir o consumo de energia de fontes não renováveis em processos industriais e agrícolas.

A quarta e última área tecnológica sustentável, "Biocombustíveis a partir de Organismos Geneticamente Modificados", pertence à Seção C: Química e Metalurgia, indicando um foco em biotecnologia. A hierarquia da IPC é: Classe C12 (Bioquímica; microbiologia; engenharia genética), Subclasse C12N (Microrganismos ou enzimas; mutação ou engenharia genética) e Grupo CN12 5/00 (Células humanas, animais ou vegetais indiferenciadas, por exemplo, linhas celulares; tecidos; cultivo ou manutenção dos mesmos). A descrição no GI associa esta tecnologia diretamente à produção de energia limpa e biotecnologia, desdobrando-se em: 1. Produção alternativa de energia, 2. Biocombustíveis e 3. A partir de organismos geneticamente modificados. Este grupo engloba inovações que utilizam células e culturas de tecidos, frequentemente por meio de engenharia genética, para otimizar a produção de biomassa para biocombustíveis.

Análises de conteúdo das patentes são necessárias em estudos futuros, a fim de avaliar a real relevância de cada espécie vegetal para a inovação sustentável. Isso se deve ao fato de que a simples menção de uma planta na patente não garante que ela seja o insumo principal da reivindicação.

**Quadro 17** – Descrição IPC geral e sustentável das patentes

IPC	Descrição IPC geral	Descrição IPC <i>Green Inventory</i>	número de patentes
<b>A01N 65/00</b>	biocides, pest repellants or attractants, or plant growth regulators containing material from algae, lichens, bryophyta, multi-cellular fungi or plants, or extracts thereof	pesticide alternatives	4
<b>B01D 53/04</b>	with stationary adsorbents	separation of components	3
<b>F26B 3/00</b>	drying solid materials or objects by processes involving the application of heat	use of solar energy for drying materials or objects	3
<b>C12N 5/10</b>	undifferentiated human, animal or plant cells, e.g. cell lines; tissues; cultivation or maintenance thereof; culture media therefor (plant reproduction by tissue culture techniques)	from genetically engineered organisms	2

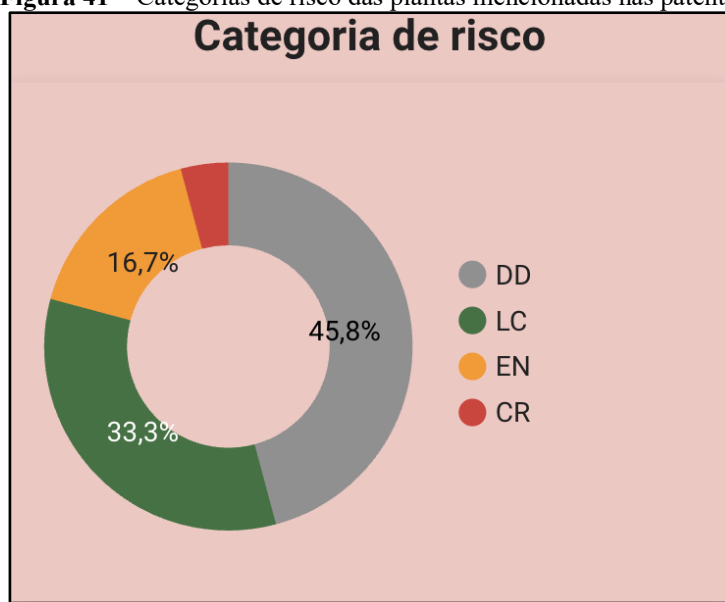
Fonte: A autora (2026).

A análise da produção tecnológica via IPC *Green Inventory* é um diferencial metodológico. A utilização desta classificação, conforme a metodologia validada por Favot *et al.* (2023a, 2023b), permite identificar explicitamente quais patentes de plantas endêmicas possuem um viés de tecnologia verde. A presença dessas patentes verdes, embora em volume ainda modesto, dialoga com os estudos de Cui *et al.* (2022) e D'amato *et al.* (2021) sobre o papel da inovação na sustentabilidade. Seus trabalhos sugerem que o crescimento de patentes verdes pode ser um reflexo da pressão regulatória e da demanda por soluções ambientais. A predominância de patentes estrangeiras com classificação verde indica que essa demanda está sendo atendida primariamente por atores externos, ressaltando a urgência de políticas nacionais de fomento à inovação sustentável brasileira.

#### **4.3.3 Lista Vermelha CNCFlora nas patentes**

Referente à Lista Vermelha do CNCFlora (Figura 40), achados apontam para um cenário em que 11 das patentes mencionam plantas endêmicas da categoria “dados insuficientes” (DD). 8 das patentes mencionam plantas classificadas em “menos preocupantes” (LC), seguido de 4 patentes com plantas classificadas como “em perigo” (EN) e por fim 1 patente com planta classificada como “criticamente em perigo” (CR).

Esta classificação é um indicativo importante de que plantas nativas e exclusivas no bioma amazônico brasileiro com riscos críticos, em perigo ou mesmo com dados insuficientes sobre seu nível de risco sofrem potencial exploração comercial.

**Figura 41** – Categorias de risco das plantas mencionadas nas patentes

Fonte: A autora (2025).

Nota: Captura de tela do LS realizada em 2025.

A análise detalhada (Quadro 18) da classificação DD (Deficiente em Dados) revela a *Theobroma canumanense* como a espécie de maior destaque, presente em um total de 8 patentes depositadas, indicando um foco significativo da pesquisa e desenvolvimento em torno desta planta. Além da *T. canumanense*, a categoria DD inclui outras três espécies, cada uma mencionada em apenas uma patente, o que sugere um interesse menos concentrado ou mais recente. Estas espécies são a *manilkara excelsa*, a *lophanthera lactescens* e a *parkia lutea*.

Passando para a categoria LC (Preocupação Mínima), observa-se que o interesse está focado em duas plantas. A espécie *Ptychopetalum uncinatum* se destaca claramente, sendo encontrada em 7 patentes, o que a posiciona como um importante objeto de estudo ou aplicação na categoria LC. A segunda planta desta categoria, a *Vismia cauliflora*, aparece em uma patente.

Na categoria EN (Em Perigo), foram identificadas também duas espécies. A *Euxylophora paraensis* é a mais representativa, com presença em 3 patentes, enquanto a *Peltogyne altissima* está registrada em 1 patente. Por fim, a categoria CR (Criticamente em Perigo) é a menos explorada no contexto das patentes analisadas, registrando-se apenas uma espécie, a *Dicypellium caryophyllaceum*, mencionada em uma patente.

**Quadro 18** – Relação de espécies por categoria de risco e número de ocorrências em patentes

Espécie	Categoria de risco	Nº de patentes
<i>Theobroma canumanense</i>	DD – Dados insuficientes	8
<i>Ptychopetalum uncinatum</i>	LC – Menos preocupante	7
<i>Euxylophora paraensis</i>	EN – Em perigo	3

<i>Dicypellium caryophyllaceum</i>	CR – Criticamente em perigo	1
<i>Manilkara excelsa</i>	DD – Dados insuficientes	1
<i>Lophanthera lactescens</i>	DD – Dados insuficientes	1
<i>Parkia lutea</i>	DD – Dados insuficientes	1
<i>Peltogyne altissima</i>	EN – Em perigo	1
<i>Vismia cauliflora</i>	LC – Menos preocupante	1

Fonte: A autora (2025).

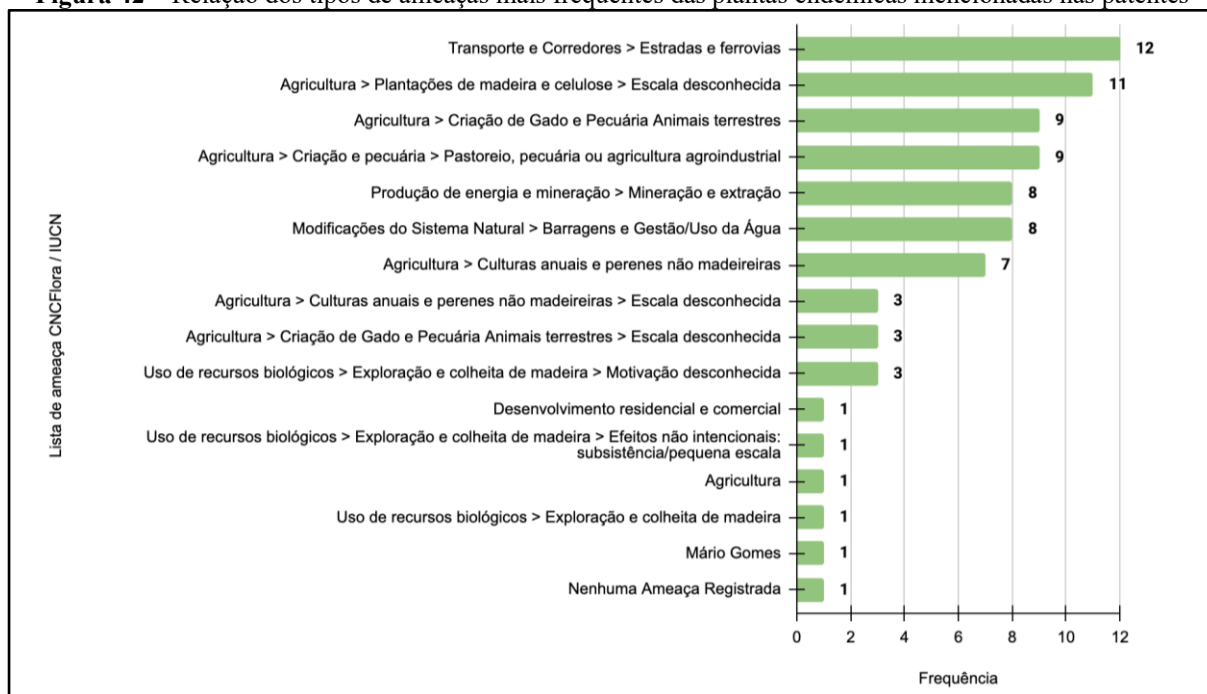
Em resumo, a distribuição das patentes por classificação de risco e por espécie revela padrões claros de interesse. A *Theobroma canumanense* (DD) e a *Ptychopetalum uncinatum* (LC) são, isoladamente, as espécies que mais atraem a atenção do setor, totalizando 15 patentes combinadas. As categorias DD e LC concentram a maior parte das patentes (18 no total), enquanto as categorias EN e CR apresentam um número substancialmente menor (4 no total). Essa distribuição pode refletir tanto o potencial econômico e biológico das espécies quanto os desafios e as restrições éticas ou legais associadas ao uso de espécies mais ameaçadas de extinção. Considerando, por exemplo, fatores de regulação, pesquisas futuras podem aprofundar a exploração desses aspectos relacionando-os às plantas e respectivas classificações de risco.

Na "Lista de ameaças", as ameaças diretas mais frequentes às plantas endêmicas citadas nas patentes são o Pastoreio, pecuária ou agricultura agroindustrial (2.3.3). Outra ameaça relevante é a de Plantações de Madeira e Celulose (2.2.3), que consiste em conjuntos de árvores (muitas vezes espécies não nativas) plantadas para madeira ou fibra fora de florestas naturais, cuja escala é desconhecida ou não registrada. Esta combinação de ameaças foi relacionada à planta *Theobroma canumanense*, mencionada em 8 patentes (IUCN, 2022)

A análise da distribuição e frequência das ameaças (Figura 41) revela que "Estradas e ferrovias" é a mais comum, com 12 ocorrências. Em seguida, a agricultura aparece como um fator significativo, destacando-se: plantações de madeira e celulose em escala desconhecida (11), agropecuária (9), agricultura industrial (9), cultura não madeireira (7), e escalas desconhecidas da cultura não madeireira (3) ou criação de gado e pecuária (3). Outras ameaças relevantes para as plantas incluem mineração e extração (8), barragens e gestão do uso da água por meio de modificação de sistemas naturais (8), e o uso de recursos biológicos para exploração e coleta de madeira com motivação desconhecida (3). A capacidade de correlacionar essas ameaças com as plantas em questão indica que o reconhecimento dessas ameaças está sendo considerado. No entanto, levanta-se uma questão crucial: se ameaças já estão mapeadas

para essas plantas, o seu subsequente aparecimento em patentes deveria funcionar como um alerta adicional para evitar a exploração dessas espécies?

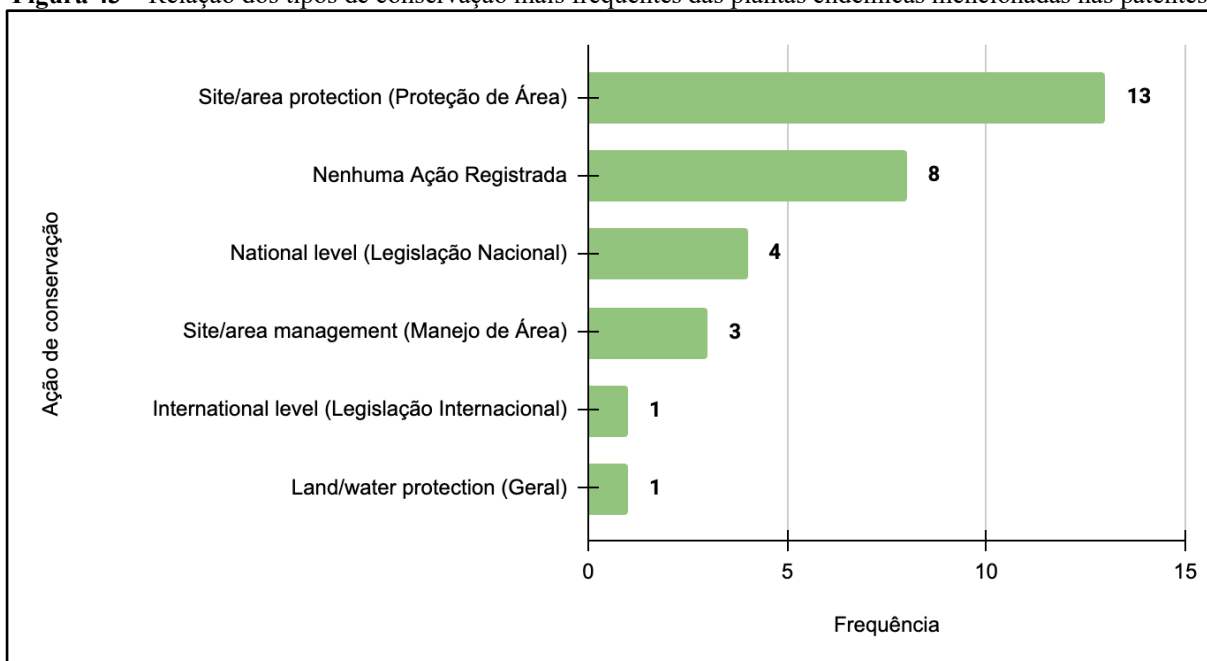
**Figura 42** – Relação dos tipos de ameaças mais frequentes das plantas endêmicas mencionadas nas patentes



Fonte: A autora (2025).

No aspecto de conservação (Figura 42), a análise das patentes sobre plantas demonstrou que a estratégia dominante é a proteção *in situ*, presente em nove (9) patentes, focada na preservação do local de ocorrência da espécie. Apesar da proteção existente, oito (8) subpopulações ou áreas de ocorrência carecem de ações de conservação em andamento, sendo prioritárias para futuras iniciativas.

Em termos legais, quatro (4) patentes envolvem espécies ou áreas com proteção legal em nível nacional, e uma (1) com proteção em nível global. Uma estratégia mais ativa, a proteção por manejo de área, foi identificada em três (3) patentes, englobando ações de gestão como monitoramento de habitats. Por fim, uma (1) patente se enquadra na categoria de proteção geral, combinando diferentes níveis de proteção *in situ* e *ex situ*.

**Figura 43** – Relação dos tipos de conservação mais frequentes das plantas endêmicas mencionadas nas patentes

Fonte: A autora (2025).

Com característica exploratória, os dados prévios de “Lista de ameaça” e “Lista de ações para conservação” do CNCFlora poderão trazer *insights* igualmente relevantes aos da “Lista de categoria de risco”, ao ser aprimorado em processos futuros, permitindo identificar a relação específica como exemplificado na matriz do Quadro 19.

**Quadro 19** – proposta de matriz para relacionamento de dimensões do CNCFlora

Classificações CNCFlora	Categoria de risco	Lista de ameaças	Lista de ações para conservação
<b>Categoria de risco</b>	-	patentes com plantas em categoria de risco e respectivos tipos de ameaças listadas	patentes com plantas em categoria de risco e respectivos tipos de ações de conservação listadas
<b>Lista de ameaças</b>	patentes com plantas em lista de ameaças e respectivos tipos de categorias de risco	-	patentes com plantas em lista de ameaças e respectivas ações de conservação listadas
<b>Lista de ações para conservação</b>	patentes com plantas em lista de ações para conservação e respectivas categorias de risco	patentes com plantas em lista de ações para conservação e respectivos tipos de ameaças listadas	-

Fonte: A autora (2026).

A análise por matriz fornece informações mais detalhadas e aprofundadas do que seria possível com uma análise simplificada e unidimensional.

#### 4.3.4 Intersecção entre Patentes Verdes e Lista Vermelha CNCFlora

Posto a viabilidade de intersecção pelo modelo criado, ao vincular dados de patentes com base no GI aos dados de patentes da Lista Vermelha, informações importantes são encontradas. Uma delas é o fato de que apenas 1 espécie está relacionada com a patente verde, sendo esta *Peltogyne altissima*, classificada como “Em perigo” (EN) subordinada à GI “Utilização da Energia Solar para Secagem” (F26B3/00). Esta patente foi publicada em 2013 com país de referência de publicação, de aplicante e de inventor dos Estados Unidos da América (Figura 43).

**Figura 44** – Mapa da publicação de “patente verde” mencionando a planta *Peltogyne altissima* categorizada como “em perigo”



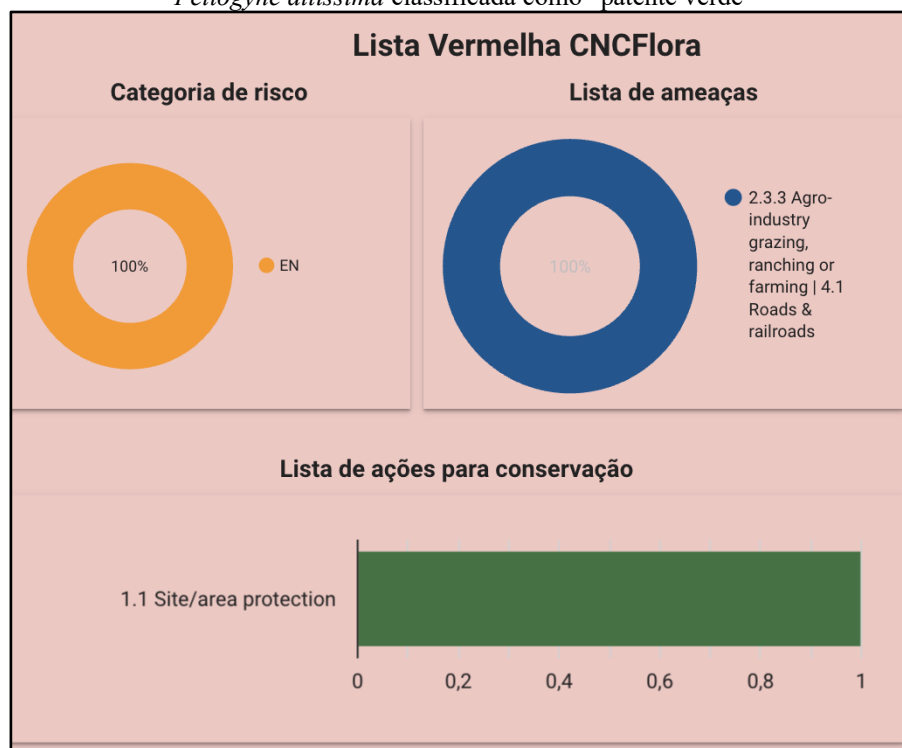
Fonte: A autora (2025).

Nota: Captura de tela do LS, realizado no segundo semestre de 2025.

Analisando a única patente que apresenta intersecção (Figura 44), identificam-se como principais ameaças à planta: Pastoreio, pecuária ou agricultura agroindustrial (2.3.3); e Rodovias e ferrovias (4.1). A primeira ameaça direta é causada pela agricultura e aquicultura, que resultam da expansão e intensificação agrícola, incluindo silvicultura, maricultura e aquicultura (e os impactos de cercamentos em áreas cultivadas), com foco em Pastoreio, Criação de Gado ou Agricultura Agroindustrial. A segunda ameaça direta está relacionada a "Corredores de Transporte e Serviços", que são corredores longos e estreitos e os veículos que os utilizam, gerando inclusive mortalidade da vida selvagem. Especificamente, trata-se de "Rodovias e Ferrovias: Transporte terrestre em rodovias e vias férreas dedicadas". Tais corredores provocam estresses à biodiversidade, notadamente a fragmentação de habitats, e

potencializam outras ameaças, como fazendas, espécies invasoras e caçadores furtivos (IUCN, 2022)

**Figura 45** – Compêndio de indicadores da Lista Vermelha CNCFlores da planta *Peltogyne altissima* classificada como “patente verde”



Fonte: A autora (2025).

Nota: Captura de tela do LS, realizado no segundo semestre de 2025.

Na "Lista de ações de conservação", esta planta está classificada como "Proteção do local/área" (1.1). De acordo com a IUCN (2012), esta classificação abrange todas as ações destinadas a proteger diretamente a biodiversidade através de parques, reservas, servidões e similares. Isso inclui, mais especificamente, o estabelecimento ou a expansão de parques, reservas e outras áreas protegidas, sejam elas públicas ou privadas.

Contudo, corrobora-se a relevância de aprofundar a exploração da relação direta e complexa entre os riscos, ameaças e as ações de conservação de espécies vegetais exclusivas da Amazônia brasileira mencionadas em documentos de patente, com particular atenção àquelas depositadas por instituições estrangeiras. Essa análise aprofundada é crucial, pois serve de base sólida e bem fundamentada para a elaboração, implementação e aperfeiçoamento de políticas públicas eficazes. Tais políticas devem ser desenhadas não apenas para garantir a proteção efetiva desses valiosos recursos genéticos e da biodiversidade amazônica como um todo, mas também para servir como um aparato robusto no combate à biopirataria e à apropriação indevida do Conhecimento Tradicional Associado (CTA). A identificação de

espécies-chave em patentes estrangeiras pode indicar áreas de maior pressão de exploração e, conseqüentemente, demandar esforços de monitoramento e fiscalização mais intensos.

Adicionalmente, é imperativo que qualquer eventual exploração comercial ou científica desses recursos seja regida por princípios éticos rigorosos e sustentáveis. É fundamental que, em situações onde a exploração se mostre viável (sem risco de extinção da espécie, degradação ambiental ou violação de direitos de comunidades tradicionais), haja um mecanismo de Repartição de Benefícios justo e equitativo. Esse mecanismo deve assegurar que os benefícios econômicos, sociais e tecnológicos gerados por essa exploração sejam revertidos, de forma concreta e substancial, para as comunidades locais detentoras do Conhecimento Tradicional Associado e que habitam as áreas de ocorrência dessas plantas. Isso fomenta o desenvolvimento local, valoriza a cultura e o conhecimento tradicional e incentiva a própria conservação. Da mesma forma, é vital criar incentivos e mecanismos de apoio para pesquisadores, cientistas, técnicos e outros profissionais brasileiros que estão direta ou indiretamente envolvidos com o estudo, a conservação, a catalogação e o manejo sustentável dessas plantas. O investimento em pesquisa nacional fortalece a soberania científica e tecnológica do país sobre seus recursos naturais.

Em suma, a intersecção entre a propriedade intelectual, a conservação da biodiversidade e os direitos das comunidades tradicionais na Amazônia exige uma abordagem multidisciplinar e estratégica para garantir que a riqueza biológica brasileira seja protegida e que seus benefícios sejam compartilhados de maneira justa.

#### 4.3 VALIDAÇÃO DA HIPÓTESE E CUMPRIMENTO DOS OBJETIVOS

A criação da sistemática e do protótipo do painel, ao viabilizar o monitoramento periódico da dinâmica da CT&I, comprova a hipótese central da pesquisa e atinge o objetivo geral proposto. Mais do que um mero acompanhamento, os produtos criados se estabelecem alinhado à visão de Hoffmann (2011), como um elemento de inteligência estratégica para programas de pesquisa e desenvolvimento. Ao oferecer filtros e cruzamentos dinâmicos (como a Matriz de ODS), o sistema permite que o usuário realize não apenas a função básica de monitoramento (o que está sendo feito?), mas também o planejamento estratégico (quais lacunas devem ser priorizadas?), transformando dados em insumos para a tomada de decisão em políticas públicas ou mesmo empresariais.

### 4.3.1 Intersecção e Fluxo de Conhecimento: Ciência, Tecnologia e Sustentabilidade

A assimetria observada na distribuição geográfica da produção — ciência majoritariamente brasileira (4.1.1) versus patentes predominantemente estrangeiras (4.2.1) — exige uma discussão pautada na geopolítica do conhecimento. Tal dinâmica reforça o alerta de Becker (2005) sobre a vulnerabilidade da Amazônia em relação ao seu Capital Natural e as preocupações levantadas por Chaves e Nogueira (2008) sobre a Propriedade Intelectual e o risco de apropriação indevida. A intersecção desses dados com o status de ameaça da CNCFlora (4.2.7.2) torna o painel uma ferramenta de alerta, pois evidencia que o interesse comercial estrangeiro (via patenteamento) pode estar incidindo sobre recursos biológicos de alto risco e de domínio exclusivo do bioma brasileiro. Essa constatação reitera a urgência do monitoramento contínuo para subsidiar a regulamentação.

O alinhamento majoritário da produção científica ao ODS 3 (Saúde e Bem-Estar), conforme a Matriz de ODS, corrobora a perspectiva do Paradigma do Desenvolvimento Sustentável de Freitas e Freitas (2016), que destaca a saúde humana como um eixo central para a valorização de bioativos naturais. Contudo, este foco deve ser analisado criticamente em face do risco. A evidência de que 30% das espécies pesquisadas encontram-se em alguma categoria de ameaça — somada à observação de que instituições estrangeiras estão ativas na pesquisa (4.1.1) e no patenteamento (4.2.1) — reforça o alerta de Becker (2005) e Chaves & Nogueira (2008). Esta dinâmica exige monitoramento contínuo para mitigar o potencial risco de apropriação indevida do capital natural brasileiro.

### 4.3.2 Validação da Arquitetura, Usabilidade e Sustentabilidade do Painel como MVP

A garantia da sustentabilidade técnica da sistematização para o painel reside na sua arquitetura de dados. A modelagem em Esquema Estrela, baseada nos princípios de *Data Warehouse* de Kimball & Ross (2013), foi fundamental para integrar os dados heterogêneos (*Scopus*, *Espacenet* e classificadores de sustentabilidade) em um formato escalável e coeso. Essa estrutura não apenas otimiza o desempenho das consultas, mas é o pilar que suporta a periodicidade das atualizações via *pipeline* ETL. Além disso, a escolha do *Looker Studio* e seu *design* visual buscam atender ao princípio de Usabilidade estabelecido por Few (2006): reduzir o esforço cognitivo do usuário para a obtenção de *insights* estratégicos. É fundamental salientar que o painel resultante desta pesquisa é um protótipo. Sua função primordial é servir como

ponto de partida para aprimoramentos futuros, abrangendo tanto o processo de tratamento de dados quanto os códigos utilizados e dos respectivos indicadores para análise.

A arquitetura de dados desenvolvida é o pilar da validade e sistematicidade dos indicadores subsequentes. A sustentabilidade e a natureza dinâmica do sistema são garantidas pelo *pipeline* ETL (Extração, Transformação e Carga), detalhado na Fase 2 da Metodologia, que automatiza o tratamento de dados heterogêneos. Essa automação, comprovada pela eficiência dos scripts na normalização dos 800 artigos e 153 patentes, transforma a visualização de CT&I em um monitoramento contínuo e replicável da dinâmica do campo. Em termos práticos, isso valida a hipótese de automação da pesquisa.

Metodologicamente, esta pesquisa se alinha à intersecção da Ciência da Informação (CI) com a Ciência de Dados, conforme o debate atual (Fernandes, 2020; Garoufallou e Gaitanou, 2021). Esses autores defendem a adoção de técnicas de Data Science pela CI para lidar com *Big Data* e dados heterogêneos. Assim, o desenvolvimento do *pipeline* ETL transcende um mero procedimento técnico, representando uma evolução metodológica na CI. Ele confere ao sistema a capacidade sistemática e escalável de manusear e enriquecer os dados bibliométricos.

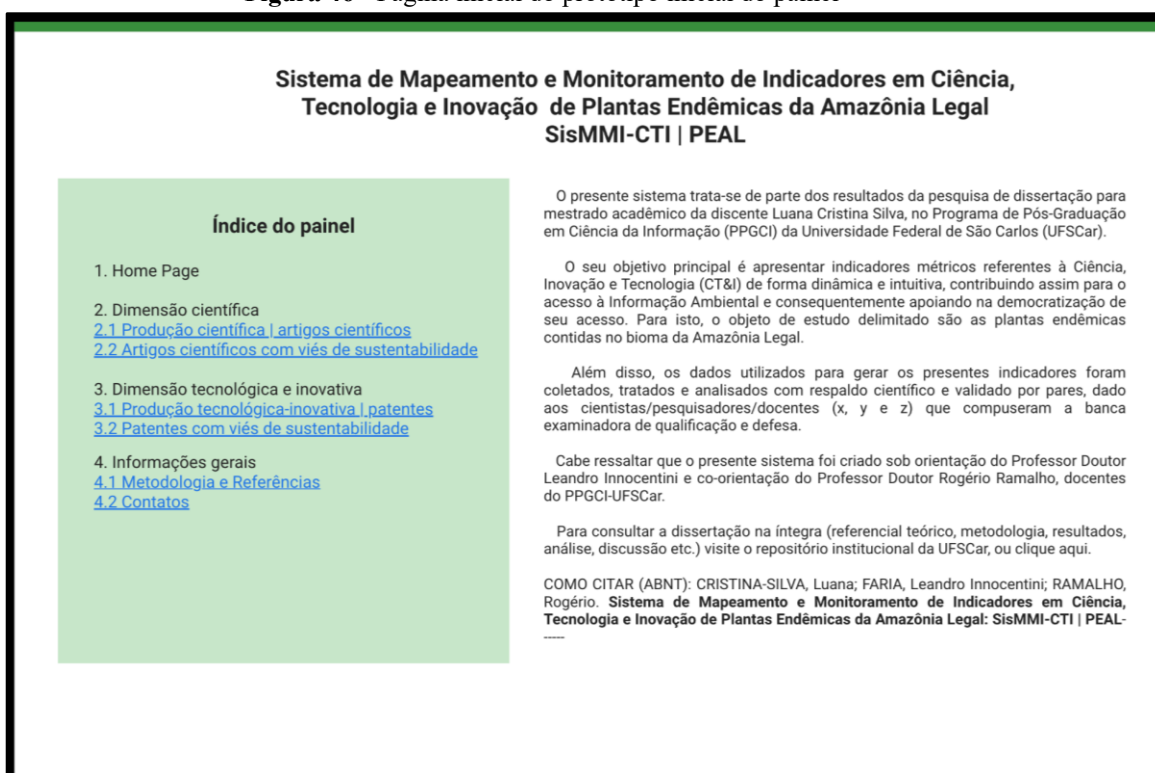
A garantia da longevidade e da eficácia do sistema é cimentada pela arquitetura de visualização adotada. A escolha do *Looker Studio* (LS) do *Google*, conforme discutido em Métodos (Seção 3.3.3.2), demonstrou ser crucial para a sustentabilidade da adoção. A interface *user-friendly* do LS, já familiar a usuários de ferramentas *Google*, facilita a redução do esforço cognitivo do usuário para a interpretação dos indicadores. Além disso, a facilidade de compartilhamento e o baixo custo operacional, combinados com a estrutura dimensional dos dados (Esquema Estrela), permitem que a ferramenta seja mantida e utilizada por uma audiência diversificada de gestores e pesquisadores, assegurando que o sistema não se torne obsoleto ou inacessível.

Com base nos dados modelados, foi possível criar um painel em formato de protótipo (versão inicial). Os usuários podem acessar e navegar neste painel gratuitamente por meio de um *link*<sup>19</sup> público. A Figura 45 ilustra a Home Page, onde constam o título do painel e informações introdutórias sobre seus objetivos e propósito.

---

<sup>19</sup> Link para acesso ao protótipo do painel (versão 2): <https://lookerstudio.google.com/s/ktTLx6ARlfc>

Figura 46 - Página inicial do protótipo inicial do painel



Fonte: A autora (2025).

Nota: Captura de tela do LS realizado no segundo semestre de 2025.

A estrutura de navegação no protótipo do painel adotada foi a hierárquica, contendo uma sequência de seções e páginas, onde o usuário poderá utilizar o recurso na barra lateral esquerda do para localização e navegação em nível macro no relatório, conforme descrito no Quadro 20.

Quadro 20 – Estrutura de navegação sugerida para o painel

Seção	Páginas	Conteúdo Principal
<b>1. Home Page (Página Inicial)</b>	-	Propósito da ferramenta e como citá-la.
<b>2. Dimensão Científica</b>	2.1	Produção científica (artigos científicos).
	2.2	Artigos científicos com viés de sustentabilidade (ODS   Lista Vermelha CNCFlora).
<b>3. Dimensão Tecnológica</b>	3.1	Produção tecnológica (patentes).
	3.2	Patentes com viés de sustentabilidade (IPC <i>Green Inventory</i>   Lista Vermelha CNCFlora).
<b>4. Informações Gerais</b>	4.1	Metodologia e Referências.
	4.3	Contatos.
<b>5. Relatórios Estáticos</b>	-	Relatórios de versões anteriores.

Fonte: A autora (2025).

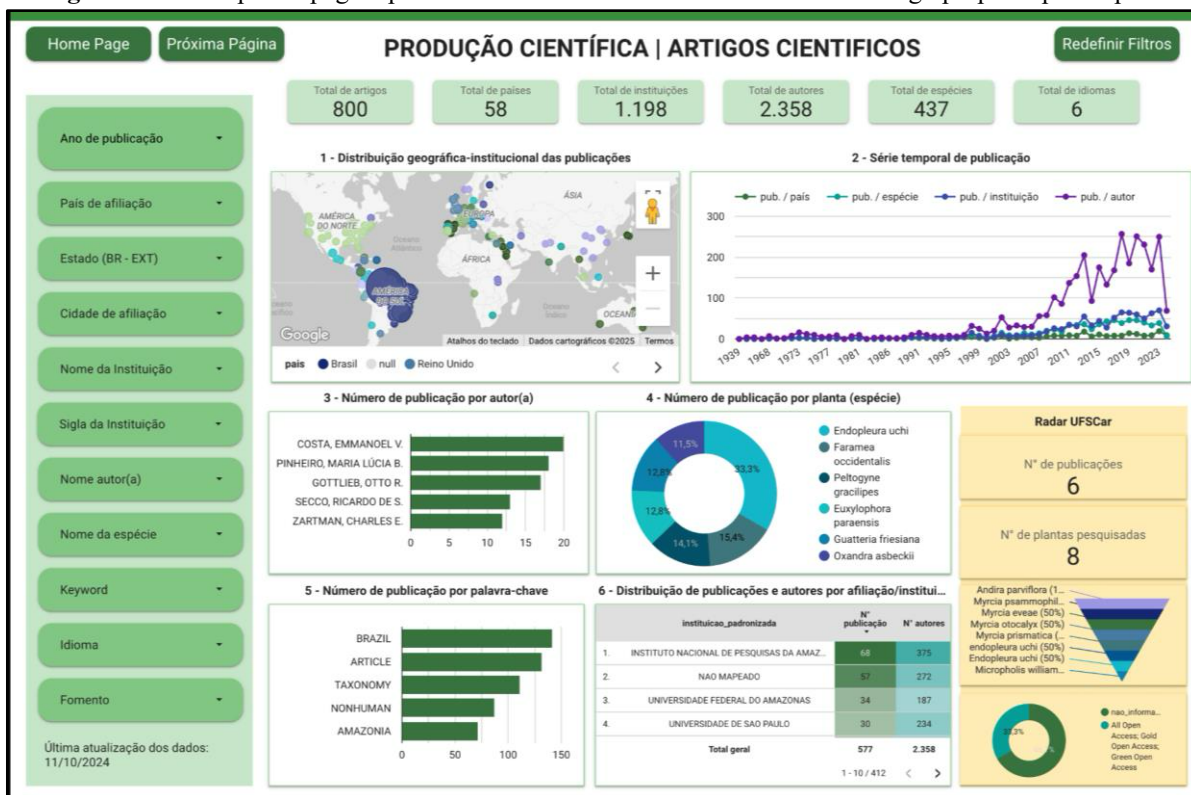
Nesta proposta de organização hierárquica, a divisão das páginas do relatório é estruturada de forma estratégica para agrupar indicadores por seções temáticas ou funcionais. Este arranjo visa facilitar significativamente a visualização e interpretação dos dados, apresentando a informação de maneira mais simplificada e segmentada. Essa segmentação não é meramente estética, mas sim funcional, pois alinha-se diretamente aos princípios de design de informação estabelecidos por especialistas como Stephen Few (2006).

Ao segmentar o conteúdo, o relatório busca reduzir a sobrecarga cognitiva do leitor. Apresentar um grande volume de indicadores em uma única página pode levar à confusão e dificultar a identificação de *insights* relevantes. A estrutura hierárquica, por sua vez, permite que o leitor se concentre em um conjunto menor e mais coeso de informações por vez, facilitando o processo cognitivo de absorção e análise.

Especificamente, a aderência aos princípios de Few (2006), que enfatizam a clareza, precisão e eficiência na comunicação visual de dados quantitativos, garante que a apresentação dos indicadores seja focada no que é essencial para a tomada de decisão. Busca-se, assim, garantir que cada elemento visual contribua efetivamente para a compreensão da informação. Portanto, essa arquitetura de informação não só organiza o relatório de forma lógica, mas também otimiza a experiência do usuário, transformando um potencial desafio de interpretação de dados complexos em um processo de visualização mais intuitivo, direcionado e eficiente.

Neste sentido, o protótipo do painel sugere o padrão exemplificado na Figura 46. Em linha com os princípios de design de informações propostos por Edward Tufte (2007), que preconizam a máxima razão entre dados e tinta (*data-ink ratio*), o design busca a clareza e a eficiência visual. A tarja lateral esquerda demarca o campo onde estão dispostos os filtros dinâmicos, que são a representação funcional da Modelagem Dimensional (Esquema Estrela) aplicada na fase de Tratamento e Enriquecimento de Dados (ETL). Esta arquitetura de dados subjacente permite ao usuário selecionar uma ou mais variáveis de interesse – como o grupo taxonômico da planta ou uma instituição de pesquisa específica – e observar a atualização dinâmica e imediata de todos os gráficos.

**Figura 47** - Exemplo de página padrão com os indicadores bibliométricos de artigo proposto para o painel



Fonte: A autora (2025).

Nota: Captura de tela do LS realizado no segundo semestre de 2025.

Essa capacidade de segmentação ágil é essencial para o objetivo da pesquisa, pois permite que o monitoramento vá além da simples contagem de produção (macro-bibliometria). Os gráficos apresentados, como a série temporal e a distribuição geográfica-institucional, demonstram a dinâmica da produção de CT&I, enquanto a inclusão de filtros de Indicadores de Sustentabilidade (ODS e CNCFlora) viabiliza a análise de dados em intersecção, consolidando o modelo analítico proposto pela pesquisa. A simplicidade visual do painel, portanto, é uma consequência direta do rigor metodológico aplicado no *pipeline* ETL. Sugere-se que a identidade visual, incluindo gráficos e outros elementos visuais, seja replicada em todas as páginas do painel, admitindo-se as adaptações que se façam necessárias.

#### 4.3.3 Comprovação da Hipótese e Alcance dos Objetivos da Pesquisa

A hipótese central desta pesquisa, que postula o desenvolvimento de um sistema de monitoramento automatizado capaz de integrar dados heterogêneos para viabilizar a análise periódica dos desenvolvimentos científicos e tecnológicos sobre plantas endêmicas da

Amazônia, é categoricamente confirmada pelos resultados e pela arquitetura do sistema demonstrada.

A viabilidade do monitoramento sistemático se baseou na concretização de três pilares: (1) a criação do *pipeline* ETL (Extração, Transformação e Carga), que demonstrou capacidade de harmonizar dados estruturados (*Scopus/Espacenet*) e não estruturados (fontes de sustentabilidade); (2) a geração do novo modelo analítico de intersecção de métricas, cruzando indicadores bibliométricos com indicadores de sustentabilidade, como ODS e CNCFlora (Seções 4.1.7 e 4.2.7); e (3) a arquitetura de visualização, que garante a aplicabilidade do monitoramento periódico e a redução do esforço cognitivo do usuário. Este modelo comprova que a incorporação de métricas externas é crucial para prover uma visão contextualizada e alinhada às demandas de desenvolvimento sustentável da Amazônia.

O sucesso na comprovação da hipótese é resultado direto do alcance integral dos Objetivos Específicos (OEs) propostos na Seção 1.2.2. A pesquisa seguiu o ciclo metodológico previsto, começando pela definição dos protocolos de busca e indicadores na Fase 1 (Seção 3.2), que estabeleceu a base para a coleta e o escopo da macro-bibliometria. Posteriormente, o desenvolvimento do *pipeline* ETL e a modelagem dimensional (Esquema Estrela) foram executados na Fase 2 (Seção 3.3.3), resultando no protótipo do painel de visualização (Seção 4.3.2). Por fim, a validação da funcionalidade do sistema e a proposição do protocolo de atualização (Seção 3.4.3) cumpriram o último OE, assegurando a perenidade da ferramenta e a governança dos dados, um requisito fundamental para a Ciência de Dados e para a Ciência da Informação.

Em suma, o sistema de monitoramento e o protótipo do painel de visualização desenvolvido representam uma solução metodológica para a lacuna de indicadores sistematizados para a Amazônia identificada na Introdução (Seção 1.1.1). Ao integrar de forma transparente os dados de Ciência, Tecnologia e Inovação (CT&I) com a perspectiva da sustentabilidade, o trabalho estabelece um novo paradigma analítico para a área, onde a Bibliometria se funde com a Arquitetura da Informação para prover subsídios estratégicos. Conclui-se, portanto, que a pesquisa não só atingiu seu Objetivo Geral, como também consolidou uma contribuição metodológica e prática para os estudos métricos e para o apoio à tomada de decisão sobre a conservação de recursos naturais exclusivos do bioma amazônico.

## 5 CONSIDERAÇÕES FINAIS

No decorrer histórico da ciência, teóricos e filósofos evidenciaram que conclusões de investigações científicas não se sustentam em princípios de auto-evidência, verdades absolutas ou provas conclusivas. Por outro lado, outros teóricos e filósofos defendem que a ciência não se constrói através do intuito de corroborar ou refutar teorias e verdades, mas sim com o objetivo de tornar o mundo inteligível, sistemático e seguro da realidade. Na contemporaneidade, portanto, a ciência se caracteriza pela reformulação e atualização constante, e se assegura através dos métodos (Koche, 2011)

A pesquisa comprova integralmente a hipótese ao desenvolver e validar um sistema de monitoramento automatizado, consubstanciado no protótipo de um painel de visualização dinâmico. Esta solução preenche a lacuna inicialmente identificada, ao fornecer um novo modelo analítico que integra, de forma sistemática e periódica, indicadores bibliométricos com métricas classificatórias de sustentabilidade (*Green Inventory*, ODS e CNCFlora). A arquitetura modular do *pipeline* ETL, aliada ao *Looker Studio*, não só otimiza o esforço de atualização e análise de dados heterogêneos, mas também consolida uma contribuição metodológica robusta para o campo da Ciência da Informação, particularmente na vertente de estudos métricos e Inteligência Competitiva.

O rigor da abordagem quantitativa se manifesta na extensão do corpus analisado, que incluiu 800 artigos da base *Scopus* e 153 patentes da *Espacenet*, a partir da prospecção de 1975 espécies endêmicas da flora amazônica. Este escopo demonstra a exaustividade do protocolo de busca e garante a validade dos indicadores gerados para a temática das plantas amazônicas, fornecendo uma base de dados estruturada e enriquecida para futuras investigações.

A pesquisa revela uma assimetria crítica que exige atenção estratégica: o Brasil demonstra uma clara liderança na produção científica sobre as plantas endêmicas, com a maioria das publicações alinhada aos Objetivos de Desenvolvimento Sustentável (ODS) e com atenção às Categorias de Risco CNCFlora. Essa aliança de indicadores demonstra que a comunidade científica nacional está ativamente engajada na pesquisa sob a lente da responsabilidade social e ambiental. Em contrapartida, a produção tecnológica (patentes) é notavelmente dominada por escritórios estrangeiros (EUA, Japão e Europa). Mais preocupante, a análise da intersecção entre o IPC *Green Inventory* (Patentes Verdes) e as Categorias da Lista Vermelha CNCFlora demonstra que essas patentes estrangeiras frequentemente mencionam espécies classificadas em risco de extinção.

Essa concentração de desenvolvimento tecnológico estrangeiro sobre o capital genético de espécies ameaçadas, conforme evidenciado pelos indicadores de sustentabilidade, reforça um alerta estratégico sobre a potencial biopirataria e a urgência de fortalecer a soberania nacional sobre o conhecimento e a inovação derivados da biodiversidade. Desse modo, o painel proposto é mais do que um sistema de métricas; é uma ferramenta de Inteligência Estratégica essencial para que tomadores de decisão e órgãos reguladores possam monitorar essas dinâmicas de forma preventiva e sistemática.

Um achado relevante é que uma parcela considerável de plantas endêmicas (78% nas produções científicas e 99% nas tecnológicas) não foi mencionada, o que sugere um potencial negligenciamento dessas espécies. Essa lacuna de conhecimento pode deixar de lado informações cruciais com potencial aplicação em áreas como saúde, bem-estar e alimentação, entre outras. Portanto, o fomento de mais pesquisas e desenvolvimento referentes aos bioinsumos – considerando suas particularidades de sustentabilidade, riscos, ameaças e proteções – são urgentes, tanto no sentido de promover recursos para a sociedade quanto para melhor proteger e regulamentar seu uso de maneira sustentável.

Apesar da robustez do *pipeline* ETL desenvolvido, esta pesquisa apresenta limitações inerentes à delimitação de seu escopo e às fontes de dados selecionadas. Em primeiro lugar, a cobertura da produção científica foi restrita à base de dados multidisciplinar *Scopus* e a tecnológica, à *Espacenet*. Embora estas sejam referências internacionais, a exclusão de bases regionais (como o INPI ou SciELO) e outras fontes de literatura cinzenta pode excluir desenvolvimentos localizados e relevantes para a realidade da Amazônia brasileira.

Em segundo lugar, a amostra de 1975 espécies endêmicas, embora abrangente e sistematicamente definida por cinco grupos taxonômicos principais, não representa a totalidade da flora endêmica do bioma, impondo um recorte fitogeográfico e taxonômico que se constitui como um limite à generalização absoluta dos achados. Por fim, a natureza da classificação de sustentabilidade, baseada em indicadores indiretos (mapeamento via ODS, *IPC Green Inventory* e Lista Vermelha do CNCFlora), é uma limitação que, embora necessária para a sistematicidade, impede a captura das nuances de sustentabilidade que poderiam ser identificadas apenas por meio de uma análise de conteúdo aprofundada dos textos completos de cada artigo ou patente. A agenda futura recomenda priorizar a Análise de Sensibilidade em Indicadores Compostos (sugerida pela OCDE) para testar a robustez do modelo, focando nas escalas métricas para um *benchmarking* mais preciso.

O sistema e o painel desenvolvidos estabelecem-se como uma ferramenta essencial de Ciência Aberta, fornecendo suporte técnico-científico e de Inteligência Estratégica. Um dos

principais focos é o de auxiliar os tomadores de decisão em políticas públicas e na conservação do capital natural amazônico. O painel, funcionando como um protótipo (MVP) e observatório, simplifica o acesso à informação e o processo cognitivo dos usuários, sendo útil, por exemplo, no monitoramento de atividades que orientem à indicativos de eventuais biopiratarrias.

Para o futuro da pesquisa, sugere-se a expansão e o aprimoramento contínuo deste sistema de monitoramento bibliométrico e de sustentabilidade. As principais direções para trabalhos futuros incluem a expansão da cobertura de dados, integrando outras bases bibliográficas relevantes (como *Web of Science* e SciELO) e bases patentométricas locais (como o INPI no Brasil) para obter uma visão mais completa da produção científica e tecnológica nacional e internacional. Outra direção é o aprofundamento da automação do *pipeline* ETL, visando evoluir a solução de um MVP para um sistema em produção, garantindo a atualização de dados em tempo real ou em ciclos mais curtos, com menor dependência de intervenção manual.

O desenvolvimento de indicadores diretos de sustentabilidade também é fundamental, aplicando técnicas de *text mining* e *machine learning* sobre resumos e/ou textos completos (sujeito à política de acesso) para gerar indicadores diretos de sustentabilidade, superando as limitações dos indicadores indiretos baseados em IPC e ODS, e refinando a classificação de risco das plantas endêmicas. Por fim, propõe-se a realização de estudos comparativos em outros biomas, utilizando o modelo de painel desenvolvido no *Looker Studio* como base para estudos comparativos em outros biomas de alta importância ecológica e econômica, como a Mata Atlântica ou o Cerrado, consolidando a metodologia como um padrão analítico na Ciência da Informação no viés da Informação Ambiental. Além disso, pesquisas futuras podem enriquecer as análises com a inclusão de outros indicadores bibliométricos, de sustentabilidade ou de políticas públicas.

Outro importante desdobramento para trabalhos futuros reside na elaboração e proposição de um manual de uso detalhado para o painel de visualização. Este manual é fundamental para garantir a máxima usabilidade e o pleno aproveitamento da ferramenta por parte de seus diversos públicos-alvo. Deve-se considerar a adaptação do conteúdo para atender às necessidades específicas de usuários como: pesquisadores (para aprofundamento das análises bibliométricas e de sustentabilidade); concessionários de patentes (para identificação de tendências tecnológicas e avaliação de inovações sustentáveis); tomadores de decisão da esfera governamental e elaboradores de políticas públicas (para subsidiar decisões estratégicas e a criação de políticas de desenvolvimento e conservação na Amazônia); sociedade civil e outros

interessados (para promover a transparência e a conscientização sobre o fluxo de conhecimento e o risco dos recursos naturais).

Em um contexto de crescente importância do desenvolvimento sustentável e da relevância global da Amazônia, este estudo destaca questões cruciais que demandam atenção em pesquisas futuras. É fundamental questionar: Quantas espécies nativas e endêmicas da Amazônia foram, estão ou serão exploradas sem o conhecimento adequado das comunidades científica, política ou civil? Quantas espécies foram, são ou ainda serão exploradas, mesmo com amparo legal, sem que suas particularidades de risco, ameaça e proteção sejam devidamente consideradas? Quantos recursos naturais do bioma foram, são ou serão explorados por entidades estrangeiras sem um rastreamento efetivo que garanta o retorno financeiro direcionado à sustentabilidade do ecossistema e das comunidades locais? Ou então, o conhecimento científico gerado no país está, de fato, sendo aproveitado no desenvolvimento de Pesquisa e Desenvolvimento (P&D) nacional na mesma medida em que pode estar sendo explorado em nível internacional? Essas questões emergem dos resultados obtidos pela metodologia aplicada e convidam a ciência a um constante aprofundamento e atualização, norteando futuras investigações.

## REFERÊNCIAS

- ALVARADO, R. U. A bibliometria: história, legitimação e estrutura. **Para entender a Ciência da Informação**. 1. ed. Salvador, BA: EDUFBA - Editora da Universidade Federal da Bahia, 2008. v. 1, ed. 1, p. 185–217.
- ARAÚJO, C. A. Bibliometria: evolução histórica e questões atuais. **Em Questão**, [S. l.], v. 12, n. 1, p. 11–32, 2006.
- ARAÚJO, C. A. Correntes teóricas da ciência da informação. Brasília, DF, v. 38, p. 192–204, 2009.
- ARAÚJO, C. A. Á. Teorias e tendências contemporâneas da ciência da informação. Fortaleza, CE, v. 2, n. 2, 2017.
- ASSUNÇÃO, S. S.; FONSECA, L. D. P. A. D.; CARVALHO, R. D. A.; REDIGOLO, F. M. Artigos científicos indexados na base BRAPCI sobre desenvolvimento sustentável na Amazônia no contexto da Ciência da Informação. **Informação@Profissões**, [S. l.], v. 11, n. 1, p. 1, 25 jun. 2022. DOI: 10.5433/2317-4390.2022v11n1p1. Disponível em: <https://www.uel.br/revistas/uel/index.php/infoprof/article/view/45014>. Acesso em: 12 dez. 2025.
- BATES, M. J.; MAACK, M. N (org.). Information Science. **Encyclopedia of Library and Information Sciences, Third Edition**. 0 ed. [S. l.]: CRC Press, 17 dez. 2009. ed. 0, p. 2570–2585. DOI: 10.1081/E-ELIS3-120043704. Disponível em: <https://www.taylorfrancis.com/books/9781000031805/chapters/10.1081/E-ELIS3-120043704>. Acesso em: 20 jan. 2026.
- BAUMAN, Z. **Globalização: as consequências humanas**. Rio de Janeiro: Jorge Zahar, 1999.
- BECKER, B. K. Geopolítica da Amazônia. **Estudos Avançados**, [S. l.], v. 19, n. 53, p. 71–86, abr. 2005. DOI: 10.1590/S0103-40142005000100005. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0103-40142005000100005&lng=pt&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40142005000100005&lng=pt&tlng=pt). Acesso em: 12 dez. 2025.
- BIOMA | MICHAELIS ON-LINE. *In*: Michaelis. [S. l.]: [S.n.], 2025 (Dicionário Brasileiro da Língua Portuguesa). Disponível em: <https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/bioma/>. Acesso em: 19 dez. 2025.
- BOCHI, F.; GABRIEL JUNIOR, R. F.; MOURA, A. M. M. Patentes nos estudos métricos da informação. *In*: GRÁCIO, M. C. C.; MARTÍNEZ-ÁVILA, D.; OLIVEIRA, E. F. T. de; SAMPAIO ROSAS, F (org.). **Tópicos da bibliometria para bibliotecas universitárias**. 1. ed. Marília, SP: Faculdade de Filosofia e Ciências, 2020. ed. 1, p. 264–286. DOI: 10.36311/2020.978-65-86546-91-0. Disponível em: [https://ebooks.marilia.unesp.br/index.php/lab\\_editorial/catalog/book/187](https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/book/187).
- BRASIL; CONSELHO NACIONAL DO MEIO AMBIENTE (CONAMA). **RESOLUÇÃO CONAMA No 012, de 04 de maio de 1994**. Brasília, DF: CONAMA, [S. d.]. Disponível em: <https://www.ibama.gov.br/sophia/cnia/legislacao/MMA/RE0012-040594.PDF>.

BRASIL; INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). Monitoramento do território: mudanças climáticas. 24 set. 2024. **Instituto Nacional de Pesquisas Espaciais**. Disponível em: <https://www.gov.br/inpe/pt-br/aceso-a-informacao/perguntas-frequentes/principais-produtos-e-servicos-do-inpe/monitoramento-do-territorio-mudancas-climaticas>. Acesso em: 19 jan. 2026.

BRASIL; MINISTÉRIO DA ECONOMIA (ME); INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL (INPI). **Manual Básico para Proteção por Patentes de Invenções, Modelos de Utilidade e Certificados de Adição**. Brasília, DF: INPI, [S. d.]. Disponível em: <https://www.gov.br/inpi/pt-br/servicos/patentes/guia-basico/ManualBsicodePatentes20210607b.pdf>.

BRASIL; MINISTÉRIO DA EDUCAÇÃO; COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR. Portal de Periódicos da CAPES. 2024. **Portal de Periódicos da CAPES**. Disponível em: <https://www.periodicos.capes.gov.br/>. Acesso em: 19 jan. 2026.

BRASIL; MINISTÉRIO DA GESTÃO E DA INOVAÇÃO EM SERVIÇOS PÚBLICOS (MGI). GitHub. 2023. **Ministério da Gestão e da Inovação em Serviços Públicos**. Disponível em: <https://www.gov.br/gestao/pt-br/assuntos/gestaoeinovacao/informacoes-sistemas-e-servicos-de-gestao/processo-eletronico-nacional/destaques/github/github>. Acesso em: 19 jan. 2026.

BRASIL; MINISTÉRIO DA INDÚSTRIA, COMÉRCIO EXTERIOR E SERVIÇOS; INSTITUTO NACIONAL DE PROPRIEDADE INDUSTRIAL (INPI); CEPIT; DIESP. **Guia simplificado para buscas em bases de patente gratuitas**. Rio de Janeiro: INPI, [S. d.].

BRASIL; MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS (MDIC); INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL (INPI). **Guia de Classificação Internacional de Patentes**. Brasília, DF: [S.n.], [S. d.].

BRASIL; MINISTÉRIO DO MEIO AMBIENTE E MUDANÇA DO CLIMA (MMA). Amazônia. 2021. Disponível em: <https://www.gov.br/mma/pt-br/assuntos/biodiversidade-e-biomas/biomas-e-ecossistemas/biomas/amazonia>. Acesso em: 19 dez. 2025.

BRASIL; MINISTÉRIO DO MEIO AMBIENTE (MMA). PORTARIA MMA Nº 148, DE 7 DE JUNHO DE 2022 - DOU - Imprensa Nacional. 8 jun. 2022. Disponível em: <https://www.in.gov.br/en/web/dou/-/portaria-mma-n-148-de-7-de-junho-de-2022-406272733>. Acesso em: 20 jan. 2026.

BUCKLAND, M. K. Information as thing. **Journal of the American Society for Information Science**, [S. l.], v. 42, n. 5, p. 351–360, 1991.

BURNHAM, J. F. Scopus database: a review. **Biomedical Digital Libraries**, [S. l.], v. 3, n. 1, p. 1, 8 mar. 2006. DOI: 10.1186/1742-5581-3-1. Disponível em: <https://bio-diglib.biomedcentral.com/articles/10.1186/1742-5581-3-1>. Acesso em: 12 dez. 2025.

CAMBIA. What is The Lens?: an open global cyberinfrastructure. 2025. **The Lens**. Disponível em: <https://about.lens.org/what/>. Acesso em: 19 jan. 2026.

CAPURRO, R.; HJORLAND, B. O Conceito de informação. **Perspectivas em Ciência da Informação**, [S. l.], v. 12, n. 1, p. 148–207, 2007.

CARVALHO, R. L. *et al.* Pervasive gaps in Amazonian ecological research. **Current Biology**, [S. l.], v. 33, n. 16, p. 3495-3504.e4, ago. 2023. DOI: 10.1016/j.cub.2023.06.077. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0960982223008631>. Acesso em: 12 dez. 2025.

CASTELLS, M. **The rise of the network society**. 2. ed. with a new preface, [reprint]. Malden, Mass.: Wiley-Blackwell, 2011. 597 p (The information age / Manuel Castells, 1).

CHAVES, M. do P. S. R.; NOGUEIRA, M. G. Propriedade intelectual, globalização e desenvolvimento: uma reflexão sobre os caminhos para o desenvolvimento sustentável da Amazônia. **Desenvolvimento e Meio Ambiente**, [S. l.], v. 18, p. 115–128, 2008.

CRISTINA-SILVA, L. **Pipeline ETL para alimentação do painel de indicadores bibliométricos sobre plantas endêmicas da Amazônia: subproduto da dissertação de mestrado em Ciência da Informação**. Versão 1.2. São Carlos: Zenodo, 13 ago. 2025. DOI: 10.5281/zenodo.16851246. Disponível em: <https://zenodo.org/records/16851246>. Acesso em: 9 fev. 2026.

CRISTINA-SILVA, L.; FARIA, L. I. L. de. Patenteamento e a Flora Endêmica no Bioma Amazônico. **Revista Eletrônica Competências Digitais para Agricultura Familiar**, [S. l.], v. 10, n. esp. 1, 2024.

CUI, J.; DAI, J.; WANG, Z.; ZHAO, X. Does Environmental Regulation Induce Green Innovation? A Panel Study of Chinese Listed Firms. **Technological Forecasting and Social Change**, [S. l.], v. 176, p. 121492, mar. 2022. DOI: 10.1016/j.techfore.2022.121492. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0040162522000245>. Acesso em: 12 dez. 2025.

D'AMATO, A.; MAZZANTI, M.; NICOLLI, F. Green technologies and environmental policies for sustainable development: Testing direct and indirect impacts. **Journal of Cleaner Production**, [S. l.], v. 309, p. 127060, ago. 2021. DOI: 10.1016/j.jclepro.2021.127060. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0959652621012798>. Acesso em: 12 dez. 2025.

DE SOUSA, M. N. A.; ALMEIDA, E. P. D. O.; BEZERRA, A. L. D. Bibliometrics: what is it? What is it used for? And how to do it? **Cuadernos de Educación y Desarrollo**, [S. l.], v. 16, n. 2, p. e3042, 9 fev. 2024. DOI: 10.55905/cuadv16n2-021. Disponível em: <https://ojs.cuadernoseducacion.com/ojs/index.php/ced/article/view/3042>. Acesso em: 19 fev. 2026.

DELFANTI, A.; PITRELLI, N. Ciência aberta: revolução ou continuidade? *In*: ALBAGLI, S.; MACIEL, M. L.; ABDO, A. H (org.). **Ciencia aberta, questões abertas**. Brasília, DF: IBICT, 2015.

DIAS, P. M.; DIAS, T. M. R.; MOITA, G. F. PRODUÇÃO CIENTÍFICA BRASILEIRA EM PERIÓDICOS DE ACESSO ABERTO: uma análise baseada em indicadores bibliométricos. [S. l.], v. 14, 2021.

DUTRA, R. D. Aplicação da ferramenta Google Looker Studio para análise e visualização de dados de um sistema de gestão de pesquisa ambiental. **Derbyana**, [S. l.], v. 45, n. 825, 27 dez.

2024. DOI: 10.69469/derb.v45.825. Disponível em:  
<https://revistaig.emnuvens.com.br/derbyana/article/view/825>. Acesso em: 20 jan. 2026.

ELSEVIER. Scopus: abstract and citation database. 2026. [www.elsevier.com](http://www.elsevier.com). Disponível em:  
<https://www.elsevier.com/products/scopus>. Acesso em: 19 jan. 2026.

ELSEVIER. **Scopus content coverage guide**. [S. l.]: [S.n.], [S. d.]. Disponível em:  
[https://assets.ctfassets.net/o78em1y1w4i4/EX1iy8VxBeQKf8aN2XzOp/c36f79db25484cb38a5972ad9a5472ec/Scopus\\_ContentCoverage\\_Guide\\_WEB.pdf](https://assets.ctfassets.net/o78em1y1w4i4/EX1iy8VxBeQKf8aN2XzOp/c36f79db25484cb38a5972ad9a5472ec/Scopus_ContentCoverage_Guide_WEB.pdf).

EMBRAPA AMAZÔNIA ORIENTAL; BRASIL; MINISTÉRIO DA AGRICULTURA E DO ABASTECIMENTO (MAPA). **Dendrogene: conservação genética em florestas manejadas na Amazônia**. Belém, PA: [S.n.], [S. d.]. Disponível em:  
<https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/373876/1/Dendrogene-conservacao-genetica...pdf>.

ESPAENET. Espacenet - Limitations. 2016. Disponível em:  
[https://pt.espacenet.com/help?locale=pt\\_PT&method=handleHelpTopic&topic=limitations](https://pt.espacenet.com/help?locale=pt_PT&method=handleHelpTopic&topic=limitations). Acesso em: 22 dez. 2025.

ESPAENET. **Espacenet – pocket guide**. [S. l.]: [S.n.], [S. d.]. Disponível em:  
<https://link.epo.org/web/technical/espacenet/espacenet-pocket-guide-en.pdf>.

ESPAENET. Espacenet - Questões Frequentes (FAQ) sobre pesquisas. 2017. **Espacenet**. Disponível em: [https://lp.espacenet.com/help?locale=pt\\_LP&method=handleHelpTopic&topic=faqsearching](https://lp.espacenet.com/help?locale=pt_LP&method=handleHelpTopic&topic=faqsearching). Acesso em: 22 dez. 2025.

ESPAENET. Introduction to Espacenet. 2019. **Espacenet**. Disponível em:  
[https://worldwide.espacenet.com/?locale=en\\_EP&view=intro](https://worldwide.espacenet.com/?locale=en_EP&view=intro). Acesso em: 22 dez. 2025.

ESPADA, A. L. V.; SOBRINHO, M. V. Manejo comunitário e governança ambiental para o desenvolvimento local: análise de uma experiência de uso sustentável de floresta na Amazônia. **Administração Pública e Gestão Social**, [S. l.], v. 7, n. 4, p. 169–177, 6 nov. 2015. DOI: 10.21118/apgs.v7i4.669. Disponível em:  
<https://periodicos.ufv.br/ojs/apgs/article/view/4606>. Acesso em: 12 dez. 2025.

FARIA, L. I. L. de; GREGOLIN, J. Â. R.; HOFFMANN, W. A. M.; QUONIAM, L. Análise da produção científica a partir de publicações em periódicos especializados. In: CRUZ, C. H. de B.; MASCARENHAS, M. da G.; ARANTES, J. T (ed.). **Indicadores de Ciência, Tecnologia e Inovação em São Paulo 2010**. São Paulo: FAPESP, 2011. v. 2. Disponível em:  
<https://fapesp.br/indicadores2010>.

FAVOT, M.; VESNIC, L.; PRIORE, R.; BINCOLETTO, A.; MOREA, F. Green patents and green codes: How different methodologies lead to different results. **Resources, Conservation & Recycling Advances**, [S. l.], v. 18, p. 200132, out. 2023. DOI: 10.1016/j.rcradv.2023.200132. Disponível em:  
<https://linkinghub.elsevier.com/retrieve/pii/S2667378923000044>. Acesso em: 12 dez. 2025.

FERNANDES, J. H. C. Interlocações bibliográficas e epistemológicas entre a ciência de dados e a ciência da informação. **Ciência da Informação**, [S. l.], v. 50, n. 3, p. 232–242, 2020.

FERNANDES, N. D. S.; DESOTI, V. C.; DIAS, A.; DA SILVA, Y. C.; DE AZEVEDO DOS SANTOS, A. P.; PASSARINI, G. M.; NAKAMURA, C. V.; DA VEIGA JUNIOR, V. F. Styrylpyrone, isolated from an Amazon plant, induces cell cycle arrest and autophagy in *Leishmania amazonensis*. **Natural Product Research**, [S. l.], v. 35, n. 22, p. 4729–4733, 17 nov. 2021. DOI: 10.1080/14786419.2020.1715395. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/14786419.2020.1715395>. Acesso em: 22 dez. 2025.

FITOGEOGRÁFICO. *In*: Michaelis On-line. [S. l.]: Editora Melhoramentos. Disponível em: <https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/fitogeogr%C3%A1fico/>. Acesso em: 19 jan. 2026.

FREITAS, M. de; FREITAS, M. C. da S. **A sustentabilidade como paradigma: cultura, ciência e cidadania**. 1. ed. São Paulo: Vozes, 2016. Acesso em: 2 out. 2024.

GAROUFALLOU, E.; GAITANOU, P. Big Data: Opportunities and Challenges in Libraries, a Systematic Literature Review. **College & Research Libraries**, [S. l.], v. 82, n. 3, p. 410, 2021. DOI: 10.5860/crl.82.3.410. Disponível em: <https://crl.acrl.org/index.php/crl/article/view/24918>. Acesso em: 12 dez. 2025.

GOOGLE. Looker Studio. 2026. **Google Cloud Documentation**. Disponível em: <https://docs.cloud.google.com/looker/docs/studio?hl=pt>. Acesso em: 20 jan. 2026.

GRUPP, H.; MOGEE, M. E. Indicators for National Science and Technology Policy. *In*: MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U (org.). **Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems**. Dordrecht: Springer Netherlands, 2005. p. 75–94. DOI: 10.1007/1-4020-2755-9. Disponível em: <http://link.springer.com/10.1007/1-4020-2755-9>.

GUELLEC, D.; POTTERIE, B. van P. de la. Measuring the internationalisation of the generation of knowledge: an approach based on patent data. *In*: MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U (org.). **Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems**. Dordrecht: Springer Netherlands, 2005. DOI: 10.1007/1-4020-2755-9. Disponível em: <http://link.springer.com/10.1007/1-4020-2755-9>. Acesso em: 22 dez. 2025.

HILÁRIO, C. M.; FREITAS, J. L. Indicadores de colaboração científica: aspectos éticos, práticos e formas de mensuração. *In*: GRÁCIO, M. C. C.; MARTÍNEZ-ÁVILA, D.; OLIVEIRA, E. F. T. de; SAMPAIO ROSAS, F (org.). **Tópicos da bibliometria para bibliotecas universitárias**. 1. ed. Marília, SP: Faculdade de Filosofia e Ciências, 2020. ed. 1, p. 71–92. DOI: 10.36311/2020.978-65-86546-91-0. Disponível em: [https://ebooks.marilia.unesp.br/index.php/lab\\_editorial/catalog/book/187](https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/book/187). Acesso em: 12 dez. 2025.

HINZE, S.; SCHMOCH, U. Opening the Black Box. *In*: MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U (org.). **Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems**. Dordrecht: Springer Netherlands, 2005. DOI: 10.1007/1-4020-2755-9. Disponível em: <http://link.springer.com/10.1007/1-4020-2755-9>. Acesso em: 22 dez. 2025.

HOFFMANN, W. A. M. Monitoramento da informação e inteligência competitiva: realidade organizacional. **InCID: Revista de Ciência da Informação e Documentação**, [S. l.], v. 2, n. 2, p. 125–144, 13 dez. 2011. DOI: 10.11606/issn.2178-2075.v2i2p125-144. Disponível em: <http://www.revistas.usp.br/incid/article/view/42356>. Acesso em: 21 jan. 2026.

INSTITUTO BRASILEIRO DE FLORESTAS (IBF). **Bioma Amazônico**. 2020. IBF. Disponível em: <https://www.ibflorestas.org.br/bioma-amazonico>. Acesso em: 19 dez. 2025.

INSTITUTO NACIONAL DE PROPRIEDADE INDUSTRIAL (INPI). **Base Espacenet**. Módulo 3. Rio de Janeiro: [S.n.], [S. d.]. Disponível em: [https://www.gov.br/inpi/pt-br/uso-estrategico-da-pi/estudos-e-informacao-tecnologica/Espacenet\\_Mdulo3\\_Tutorialdebusca\\_INPI\\_out2022.pdf](https://www.gov.br/inpi/pt-br/uso-estrategico-da-pi/estudos-e-informacao-tecnologica/Espacenet_Mdulo3_Tutorialdebusca_INPI_out2022.pdf).

JARDIM BOTÂNICO DO RIO DE JANEIRO. Flora e Funga do Brasil. 2019. **Reflora**. Disponível em: <https://floradobrasil.jbrj.gov.br/consulta/ficha.html?idDadosListaBrasil=24154>. Acesso em: 28 jan. 2026.

JESUS, J. B. D.; ROSA, C. N. D.; BARRETO, Í. D. D. C.; FERNANDES, M. M. Análise da incidência temporal, espacial e de tendência de fogo nos biomas e unidades de conservação do Brasil. **Ciência Florestal**, [S. l.], v. 30, n. 1, p. 176, 6 abr. 2020. DOI: 10.5902/1980509837696. Disponível em: <https://periodicos.ufsm.br/cienciaflorestal/article/view/37696>. Acesso em: 12 dez. 2025.

KIMBALL, R.; ROSS, M. **The Data Warehouse Toolkit: the definitive guide to dimensional modeling**. 3. ed. Indianapolis, IN: Wiley, 2013.

KOCHE, J. C. **Fundamentos de Metodologia Científica: teoria da ciência e iniciação à pesquisa**. Petrópolis, RJ: Vozes, 2011.

KUPFER, D.; TIGRE, P. Prospecção Tecnológica. **Modelo SENAI de Prospecção: Documento Metodológico**. [S. l.]: ORGANIZACIÓN INTERNACIONAL DEL TRABAJO - CINTERFOR, 2004. p. 1–19.

LAKATOS, E. M.; MARCONI, M. de A. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Atlas, 2003.

LAPOLA, D. M.; SILVA, J. M. C. D.; BRAGA, D. R.; CARPIGIANI, L.; OGAWA, F.; TORRES, R. R.; BARBOSA, L. C. F.; OMETTO, J. P. H. B.; JOLY, C. A. A climate-change vulnerability and adaptation assessment for Brazil's protected areas. **Conservation Biology**, [S. l.], v. 34, n. 2, p. 427–437, abr. 2020. DOI: 10.1111/cobi.13405. Disponível em: <https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.13405>. Acesso em: 19 dez. 2025.

LIU, S.; WANG, Y. Green innovation effect of pilot zones for green finance reform: Evidence of quasi natural experiment. **Technological Forecasting and Social Change**, [S. l.], v. 186, p. 122079, jan. 2023. DOI: 10.1016/j.techfore.2022.122079. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S004016252200600X>. Acesso em: 12 dez. 2025.

MARTINELLI, G.; MORAES, M. Á (org.). **Livro vermelho da flora do Brasil**. Rio de Janeiro: Andrea Jakobsson Estudio Editorial Ltda, 24 jul. 2013. 1100 p.

MEADOWS, D. Indicators and Information Systems for Sustainable Development. *In*: SATTERTHWAITE, D. **The Earthscan Reader in Sustainable Cities**. 1. ed. London: Routledge, 13 dez. 2021. ed. 1, p. 364–393. DOI: 10.4324/9781315800462-21. Disponível em: <https://www.taylorfrancis.com/books/9781315800462/chapters/10.4324/9781315800462-21>. Acesso em: 12 dez. 2025.

MENEZES, C.; SANTOS, S.; BORTOLI, R. Mapeamento de Tecnologias Ambientais: Um Estudo sobre Patentes Verdes no Brasil. **Revista de Gestão Ambiental e Sustentabilidade**, [S. l.], v. 5, n. 1, p. 110–127, 1 abr. 2016. DOI: 10.5585/geas.v5i1.369. Disponível em: <http://www.revistageas.org.br/ojs/index.php/geas/article/view/369>. Acesso em: 12 dez. 2025.

MERTON, R. K. **Ensaio de sociologia da ciência**. 1. ed. São Paulo: Editora 34, 2013. 304 p.

MICROSOFT. Conceitos básicos do Git e do GitHub para a documentação do Microsoft Learn. 2024. **Microsoft Learn**. Disponível em: <https://learn.microsoft.com/pt-br/contribute/content/git-github-fundamentals>. Acesso em: 19 jan. 2026.

MORAES, M. H. de; SILVA, J. M. da; ANTUNES NETO, J. M. F.; MENEGASSI, W. J. C. Data Warehouse e usabilidade do Modelo Dimensional Estrela: implicações organizacionais e o papel da gestão da tecnologia da informação. *Itapira*, v. 6, n. 1, p. 415–445, 2024. DOI: 10.5281/ZENODO.12726974. Disponível em: <https://zenodo.org/doi/10.5281/zenodo.12726974>. Acesso em: 21 jan. 2026.

MUELLER, S. P. M. Estudos métricos da informação em ciência e tecnologia no Brasil realizados sobre a unidade de análise artigos de periódicos. **Liinc em Revista**, [S. l.], v. 9, n. 1, p. 6–27, 2013.

MUELLER, S. P. M. Literatura científica, comunicação científica e ciência da informação. **Para entender a Ciência da Informação**. [S. l.]: [S.n.], 6 ago. 2021. p. 125–144.

NAÇÕES UNIDAS BRASIL. Objetivos de Desenvolvimento Sustentável. 2024. Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 19 jan. 2026.

NARIN, F.; BREITZMAN, A. Inventive productivity. **Research Policy**, [S. l.], v. 24, n. 4, p. 507–519, jul. 1995. DOI: 10.1016/S0048-7333(94)00780-2. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0048733394007802>. Acesso em: 12 dez. 2025.

NÚCLEO DE INFORMAÇÃO TECNOLÓGICA (NIT). **Manual de Inteligência Competitiva**. São Carlos: UFSCar - DEMa, 2004. 38 p. Disponível em: <https://www.nit.ufscar.br/arquivos/nit-materiais-manual-de-inteligencia-competitiva.pdf>.

OECD. **OECD Patent Statistics Manual**. [S. l.]: OECD, 5 fev. 2009. DOI: 10.1787/9789264056442-en. Disponível em: [https://www.oecd.org/en/publications/oecd-patent-statistics-manual\\_9789264056442-en.html](https://www.oecd.org/en/publications/oecd-patent-statistics-manual_9789264056442-en.html). Acesso em: 21 jan. 2026.

OECD. **Science, Technology and Innovation Indicators in a Changing World: Responding to Policy Needs**. [S. l.]: OECD, 20 set. 2007. DOI: 10.1787/9789264039667-en. Disponível em: [https://www.oecd.org/en/publications/science-technology-and-innovation-indicators-in-a-changing-world\\_9789264039667-en.html](https://www.oecd.org/en/publications/science-technology-and-innovation-indicators-in-a-changing-world_9789264039667-en.html). Acesso em: 19 jan. 2026.

OECD; EUROPEAN UNION; JOINT RESEARCH CENTRE - EUROPEAN COMMISSION. **Handbook on Constructing Composite Indicators: Methodology and**

**User Guide.** [S. l.]: OECD, 22 ago. 2008. DOI: 10.1787/9789264043466-en. Disponível em: [https://www.oecd.org/en/publications/handbook-on-constructing-composite-indicators-methodology-and-user-guide\\_9789264043466-en.html](https://www.oecd.org/en/publications/handbook-on-constructing-composite-indicators-methodology-and-user-guide_9789264043466-en.html). Acesso em: 22 dez. 2025.

OKUBO, Y. **Bibliometric Indicators and Analysis of Research Systems: Methods and Examples.** [S. l.]: [S.n.], [S. d.] (OECD Science, Technology and Industry Working Papers). DOI: 10.1787/208277770603. Disponível em: [https://www.oecd.org/en/publications/bibliometric-indicators-and-analysis-of-research-systems\\_208277770603.html](https://www.oecd.org/en/publications/bibliometric-indicators-and-analysis-of-research-systems_208277770603.html). Acesso em: 12 dez. 2025.

OLDHAM, P.; BARNES, C.; HALL, S. **A Review of UK Patent Activity for Genetic Resources and associated Traditional Knowledge.** Reino Unido: One World Analytics, [S. d.] (UK Patent Activity for Genetic Resources and Traditional Knowledge).

OLIVEIRA, C. E. D.; AVELLAR, A. P. M. D. Impacto da inovação tecnológica nas indústrias situadas no Brasil. **P2P E INOVAÇÃO**, [S. l.], v. 7, n. 1, p. 260–274, 19 set. 2020. DOI: 10.21721/p2p.2020v7n1.p260-274. Disponível em: <http://revista.ibict.br/p2p/article/view/5194>. Acesso em: 19 fev. 2026.

OLIVEIRA, S. S. de; VON DER WEID, I.; VILLA VERDE, F. R.; SANTOS, C. d’Urso de S. M.; COSTAL, G. C. S. Z.; STELET, A. G.; VALVERDE NETO, M. dos S. M. Mapeamento de tecnologias desenvolvidas a partir de bioinsumos da Amazônia. **Revista Letras Raras**, [S. l.], v. 8, n. 2, p. 217–221, 30 jun. 2019. DOI: 10.35572/rlr.v8i2.1482. Disponível em: <https://revistas.editora.ufcg.edu.br/index.php/RLR/article/view/1461>. Acesso em: 12 dez. 2025.

PATEL, D. B. Comparative Analysis of Data Visualization Tools: Tableau, Power BI & Looker. **The Eastasouth Journal of Information System and Computer Science**, [S. l.], v. 01, n. 01, p. 55–59, 2023.

PELSA, I.; BALINA, S. Development of economic theory – from theories of economic growth and economic development to the paradigm of sustainable development. **DIEM: Dubrovnik International Economic Meeting**, [S. l.], v. 7, n. 1, p. 91–101, fev. 2022. DOI: 10.17818/DIEM/2022/1.10. Disponível em: <https://hrcak.srce.hr/clanak/393858>. Acesso em: 19 dez. 2025.

PRADO, M. A. R. do; CASTANHA, R. C. G. Indicadores: conceitos fundamentais e importância em CT&I. In: GRÁCIO, M. C. C.; MARTÍNEZ-ÁVILA, D.; FRANCINA TANNURI DE OLIVEIRA, E.; SAMPAIO ROSAS, F (org.). **Tópicos da bibliometria para bibliotecas universitárias**. 1. ed. Marília, SP: Faculdade de Filosofia e Ciências, 2020. ed. 1, p. 49–70. <https://doi.org/10.36311/2020.978-65-86546-91-0>.

PUERTA-DÍAZ, M.; MARTÍ-LAHERA, Y.; MARTÍNEZ-ÁVILA, D. Altméria: métricas alternativas para bibliotecários. In: GRÁCIO, M. C. C.; MARTÍNEZ-ÁVILA, D.; OLIVEIRA, E. F. T. de; SAMPAIO ROSAS, F (org.). **Tópicos da bibliometria para bibliotecas universitárias**. 1. ed. Marília, SP: Faculdade de Filosofia e Ciências, 2020. ed. 1. DOI: 10.36311/2020.978-65-86546-91-0. Disponível em: [https://ebooks.marilia.unesp.br/index.php/lab\\_editorial/catalog/book/187](https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/book/187). Acesso em: 12 dez. 2025.

RAYWARD, W. B. The history and historiography of information science: Some reflections. **Information Processing & Management**, [S. l.], v. 32, n. 1, p. 3–17, jan. 1996. DOI: 10.1016/0306-4573(95)00046-J. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/030645739500046J>. Acesso em: 12 dez. 2025.

REIS, J.; HOUSLEY, M. **Fundamentals of data engineering: plan and build robust data systems**. First edition. Beijing Boston Farnham Sebastopol Tokyo: O'Reilly, 2022. 422 p.

REIS, M. de J.; CARVALHO, T. de. Ciência de Dados e Ciência da Informação: evolução e paradigmas da ciência. Fortaleza, CE, v. 7, p. 1–21, 2022. <https://doi.org/10.36517/2525-3468.ip.v7i00.2022.78490.1-21>.

SÁ, R. R.; MATOS, R. A.; SILVA, V. C.; DA CRUZ CALDAS, J.; DA SILVA SAUTHIER, M. C.; DOS SANTOS, W. N. L.; MAGALHÃES, H. I. F.; DE FREITAS SANTOS JÚNIOR, A. Determination of bioactive phenolics in herbal medicines containing *Cynara scolymus*, *Maytenus ilicifolia* Mart ex Reiss and *Ptychopetalum uncinatum* by HPLC-DAD. **Microchemical Journal**, [S. l.], v. 135, p. 10–15, nov. 2017. DOI: 10.1016/j.microc.2017.07.009. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0026265X17302680>. Acesso em: 28 jan. 2026.

SALDANHA, R. D. F.; BARCELLOS, C.; PEDROSO, M. D. M. Ciência de dados e big data: o que isso significa para estudos populacionais e da saúde? **Cadernos Saúde Coletiva**, [S. l.], v. 29, n. spe, p. 51–58, 2021. DOI: 10.1590/1414-462x202199010305. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1414-462X2021000600051&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1414-462X2021000600051&tlng=pt). Acesso em: 20 jan. 2026.

SANTOS, C. V. D.; SILVA, F. M. E.; FARIA, L. I. L. D. The Brazilian Atlantic Forest genetic resources in patents and the challenges to control the economic use of biodiversity. **World Patent Information**, [S. l.], v. 74, p. 102218, set. 2023. DOI: 10.1016/j.wpi.2023.102218. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0172219023000480>. Acesso em: 12 dez. 2025.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, [S. l.], v. 1, n. 1, p. 41–62, 1996.

SARACEVIC, T. Interdisciplinary nature of information science. [S. l.], v. 24, n. 1, 1995.

SCHMOCH, U. The Technological Output of Scientific Institutions. In: MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U (org.). **Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems**. Dordrecht: Springer Netherlands, 2005. DOI: 10.1007/1-4020-2755-9. Disponível em: <http://link.springer.com/10.1007/1-4020-2755-9>. Acesso em: 22 dez. 2025.

SCHONS, D. L.; PRADO FILHO, H. V.; GALDINO, J. F. Política Nacional de Inovação: uma questão de crescimento econômico, desenvolvimento e soberania nacional. **Coleção Meira Mattos**, Rio de Janeiro, v. 14, n. 49, 21 jan. 2020. DOI: 10.22491/cmm.a021. Disponível em: <http://ebrevistas.eb.mil.br/index.php/RMM/article/view/3063>. Acesso em: 20 jan. 2026.

SECRETARIA GERAL DE INFORMÁTICA (SIN); UNIVERSIDADE FEDERAL DE SÃO CARLOS (UFSCAR). Google Workspace for Education - E-mail. 2024. **Secretaria Geral de**

**Informática.** Disponível em: <https://www.sin.ufscar.br/servicos/e-mail-e-listas/gsuite>. Acesso em: 5 fev. 2026.

SEMENOVA, A.; SEMENOV, K.; STORCHEVOY, M. One, Two, Three: How Many Green Patents Start Bringing Financial Benefits for Small, Medium and Large Firms? **Economies**, [S. l.], v. 11, n. 5, p. 137, 4 maio 2023. DOI: 10.3390/economies11050137. Disponível em: <https://www.mdpi.com/2227-7099/11/5/137>. Acesso em: 12 dez. 2025.

SILVA, C. R. S.; VENTURIERI, G. A.; FIGUEIRA, A. Description of Amazonian Theobroma L. collections, species identification, and characterization of interspecific hybrids. **Acta Botanica Brasilica**, [S. l.], v. 18, n. 2, p. 333–341, jun. 2004. DOI: 10.1590/S0102-33062004000200012. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-33062004000200012&lng=en&tlng=en](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-33062004000200012&lng=en&tlng=en). Acesso em: 19 fev. 2026.

SILVA, D. W.; CLAUDINO, L. S.; OLIVEIRA, C. D.; MATEI, A. P.; KUBO, R. R. Extrativismo e desenvolvimento no contexto da Amazônia brasileira. **Desenvolvimento e Meio Ambiente**, [S. l.], v. 38, 31 ago. 2016. DOI: 10.5380/dma.v38i0.44455. Disponível em: <https://revistas.ufpr.br/made/article/view/44455>. Acesso em: 12 dez. 2025.

SILVA, M. B.; SANTANA, J. R. Propriedade intelectual e desempenho da indústria do Brasil. **RAM. Revista de Administração Mackenzie**, [S. l.], v. 23, n. 5, p. eRAMF220131, 2022. DOI: 10.1590/1678-6971/eramf220131.pt. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1678-69712022000500401&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1678-69712022000500401&tlng=pt). Acesso em: 19 dez. 2025.

SILVA, R. W. D. C.; PAULA, B. L. D. Causa do aquecimento global: antropogênica versus natural. **Terrae Didactica**, [S. l.], v. 5, n. 1, p. 42, 1 jul. 2015. DOI: 10.20396/td.v5i1.8637501. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/td/article/view/8637501>. Acesso em: 19 dez. 2025.

SOUZA, C. da M. **A Amazônia nas publicações científicas mapeando temáticas e atores.** 2018.

STEFANI CATIVELLI, A.; MAÑANA RODRIGUEZ, J.; LASCURAIN SANCHEZ, M. L. Brazilian green patents: behavior of inventions that receive the highest number of citations. **Em Questão**, [S. l.], p. 116098, 23 fev. 2022. DOI: 10.19132/1808-5245282.116098. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/116098>. Acesso em: 12 dez. 2025.

STEFANI CATIVELLI, A.; PINTO, A. L.; LASCURAIN SANCHEZ, M. L. Patent value index: Measuring Brazilian green patents based on family size, grant, and backward citation. **Iberoamerican Journal of Science Measurement and Communication**, [S. l.], v. 1, n. 1, p. 004, 13 nov. 2020. DOI: 10.47909/ijsmc.03. Disponível em: <https://ijsmc.pro-metrics.org/index.php/i/article/view/2>. Acesso em: 12 dez. 2025.

TEIXEIRA, A. V. OS PROGRAMAS DE PATENTES VERDES FRENTE AO DESENVOLVIMENTO SUSTENTÁVEL. **Revista de Direito, Inovação, Propriedade Intelectual e Concorrência**, [S. l.], v. 4, n. 1, p. 37–58, 21 ago. 2018. DOI: 10.26668/IndexLawJournals/2526-0014/2018.v4i1.4046. Disponível em: <https://indexlaw.org/index.php/revistadipic/article/view/4046>. Acesso em: 19 dez. 2025.

TEIXEIRA, A. V (org.). **Patentes verdes: tecnologias para o desenvolvimento sustentável**. 1. ed. Caxias do Sul, RS: Educs, 2020. Disponível em: <https://plataforma.bvirtual.com.br>. Acesso em: 2 out. 2024.

THOMSON REUTERS. **Web of Knowledge: a versatile workflow solution that fulfills the varying needs of your research community**. [S. l.]: Thomson Reuters, [S. d.]. Disponível em: [https://www.ub.fernuni-hagen.de/download/WoK5-2\\_web3.pdf](https://www.ub.fernuni-hagen.de/download/WoK5-2_web3.pdf).

TIJSSEN, R. J. W. Measuring and Evaluating Science–Technology Connections and Interactions. *In*: MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U (org.). **Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems**. Dordrecht: Springer Netherlands, 2005. p. 695–716. DOI: 10.1007/1-4020-2755-9. Disponível em: <http://link.springer.com/10.1007/1-4020-2755-9>.

TOLEDO, A. G. L.; DE CAMPOS, L. A. Patentes Verdes e o setor de Fabricação de Abrasivos no Brasil: Discutindo o Potencial Estratégico do Programa à Luz do Shared Value. **Revista de Gestão Ambiental e Sustentabilidade**, [S. l.], v. 7, n. 1, p. 146–161, 22 mar. 2018. DOI: 10.5585/geas.v7i1.313. Disponível em: <http://www.revistageas.org.br/ojs/index.php/geas/article/view/313>. Acesso em: 19 dez. 2025.

UNITED NATIONS ENVIRONMENT PROGRAMME (UNEP); INTERNATIONAL INSTITUTE FOR SUSTAINABLE DEVELOPMENT; COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION. **Sustainable Consumption and Production Indicators for the Future SDGs: UNEP Discussion Paper**. [S. l.]: UNEP, 23 mar. 2015. Disponível em: <https://wedocs.unep.org/handle/20.500.11822/8966>.

VAN BELLEN, H. M. **Indicadores de Sustentabilidade: Uma Análise Comparativa**. [S. l.]: FGV Editora, 2005. 264 p.

VAN LEEUWEN, T. Descriptive versus evaluative bibliometrics: monitoring and assessing of national R&D systems. *In*: MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U (org.). **Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems**. Dordrecht: Springer Netherlands, 2005. p. 373–388. DOI: 10.1007/1-4020-2755-9. Disponível em: <http://link.springer.com/10.1007/1-4020-2755-9>. Acesso em: 22 dez. 2025.

VAN RAAN, A. F. J. Measuring science: capita selecta of current main issues. *In*: MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U (org.). **Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems**. Dordrecht: Springer Netherlands, 2005. p. 19–50. DOI: 10.1007/1-4020-2755-9. Disponível em: <http://link.springer.com/10.1007/1-4020-2755-9>. Acesso em: 22 dez. 2025.

VOLENZO, T.; ODIYO, J. Integrating endemic medicinal plants into the global value chains: the ecological degradation challenges and opportunities. **Heliyon**, [S. l.], v. 6, n. 9, p. e04970, set. 2020. DOI: 10.1016/j.heliyon.2020.e04970. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S2405844020318132>. Acesso em: 19 dez. 2025.

WALZ, R.; PFAFF, M.; MARSCHEIDER-WEIDEMANN, F.; GLÖSER-CHAHOU, S. Innovations for reaching the green sustainable development goals –where will they come

from? **International Economics and Economic Policy**, Karlsruhe, Germany, v. 14, 2017.  
<https://doi.org/10.1007/s10368-017-0386-2>.

WILKINSON, M. D.; DUMONTIER, M.; AALBERSBERG, Ij. J.; APPLETON, G.; AXTON, M.; BAAK, A. Comment: The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, [S. l.], v. 3, n. 160018, 2016.  
<https://doi.org/10.1038/sdata.2016.18>.

WORLD INTELLECTUAL PROPERTY ORGANIZATION (WIPO). Derwent Innovation. 2025. **Inspire WIPO**. Disponível em: <https://inspire.wipo.int/derwent-innovation>. Acesso em: 19 dez. 2025.

WORLD INTELLECTUAL PROPERTY ORGANIZATION (WIPO). Publicação IPC. 2026. Disponível em:  
<https://ipcpub.wipo.int/?notion=scheme&version=20260101&symbol=A01N0065000000&menulang=en&lang=en&viewmode=f&fipcpc=no&showdeleted=yes&indexes=no&headings=yes&notes=yes&direction=o2n&initial=A&cwid=none&tree=no&searchmode=smart>. Acesso em: 22 jan. 2026.

YANAI, A. E. **Patentes de produtos naturais amazônicos: análise do impacto da inovação tecnológica mundial**. 2012.



## APÊNDICES

### APÊNDICE A – LISTA DE ESPÉCIES ENDÊMICAS DA AMAZÔNIA, RECUPERADAS NO SITE<sup>20</sup> REFLORA E A DIVISÃO EM GRUPOS PARA BUSCA NAS BASES DE DADOS

Grupo	String para busca
GRUPO 1 Angiospermas	<p>"Justicia birae" OR "Justicia carajensis" OR "Justicia cowanii" OR "Justicia divergens" OR "Justicia mcdadeana" OR "Justicia potamogeton" OR "Justicia pseudoamazonica" OR "Justicia ulei" OR "Justicia viridiflavescens" OR "Justicia zarucchii" OR "Lepidagathis callistachys" OR "Lepidagathis paraensis" OR "Mendoncia multiflora" OR "Mendoncia rizziniana" OR "Pachystachys gracilis" OR "Ruellia anamariae" OR "Ruellia hirsutissima" OR "Ruellia nitida" OR "Carpotroche froesiana" OR "Lindackeria paraensis" OR "Alstroemeria paraensis" OR "Urceolina castelnaeana" OR "Anacardium microsepalum" OR "Polygonanthus amazonicus" OR "Polygonanthus punctulatus" OR "Annona amazonica" OR "Annona angustifolia" OR "Annonaannonoides" OR "Annona calcarata" OR "Annona longipedicellata" OR "Annona paraensis" OR "Annona poeppigii" OR "Bocageopsis pleiosperma" OR "Cymbopetalum euneurum" OR "Cymbopetalum physaloides" OR "Diclinanona matogrossensis" OR "Duckeanthus grandiflorus" OR "Duguetia arenicola" OR "Duguetia aripuanae" OR "Duguetia elliptica" OR "Duguetia longicuspis" OR "Duguetia manausensis" OR "Duguetia megalocarpa" OR "Duguetia oblongifolia" OR "Duguetia subcordata" OR "Guatteria amapaensis" OR "Guatteria campinensis" OR "Guatteria cryandra" OR "Guatteria duckeana" OR "Guatteria elongata" OR "Guatteria friesiana" OR "Guatteria hispida" OR "Guatteria japurensis" OR "Guatteria myriocarpa" OR "Guatteria odorata" OR "Guatteria oriximinae" OR "Guatteria polyantha" OR "Guatteria sabuletorum" OR "Guatteria trichostemon" OR "Malmea manausensis" OR "Oxandra asbeckii" OR "Pseudoxandra borbensis" OR "Pseudoxandra duckei" OR "Pseudoxandra obscurinervis" OR "Pseudoxandra pilosa" OR "Pseudoxandra rionegrensis" OR "Unonopsis duckei" OR "Unonopsis heterotricha" OR "Xylopia annoniflora" OR "Xylopia calophylla" OR "Xylopia eglariana" OR "Xylopia longicuspis" OR "Xylopia polyantha" OR "Xylopia trichostemon" OR "Xylopia uniflora" OR "Xylopia xylantha" OR "Ambelania duckei" OR "Aspidosperma auriculatum" OR "Aspidosperma huberianum" OR "Aspidosperma inundatum" OR "Aspidosperma obscurinervium" OR "Aspidosperma salgadense" OR "Aspidosperma spruceanum" OR "Aspidosperma williamii" OR "Forsteronia amazonica" OR "Forsteronia manausana" OR "Forsteronia paludosa" OR "Forsteronia prancei" OR "Galactophora pulchella" OR "Geissospermum urceolatum" OR "Hylaea leptoloba" OR "Lacmellea densifoliata" OR "Lacmellea floribunda" OR "Lacmellea gracilis" OR "Laubertia brasiliensis" OR "Malouetia amazonica" OR "Malouetia amplexicaulis" OR "Malouetia aquatica" OR "Malouetia gracillima" OR "Malouetia lata" OR "Malouetia nitida" OR "Malouetia pumila" OR "Mandevilla arenicola" OR "Mandevilla krukovii" OR "Matelea brasiliensis" OR "Matelea friesii" OR "Matelea microphylla" OR "Matelea quindecimlobata" OR "Molongum zschokkeiforme" OR "Nephradenia reflexa" OR "Pacouria paraensis" OR "Parahancornia amara" OR "Parahancornia krukovii" OR "Prestonia amazonica" OR "Prestonia macroneura" OR "Prestonia plumierifolia" OR "Rauvolfia pentaphylla" OR "Rhodocalyx rotundifolius" OR "Rhytidostemma fontellanum" OR "Ruehssia amylycea" OR "Ruehssia bergii" OR "Ruehssia thomasii" OR "Secondatia duckei" OR "Spongiosperma longilobum" OR "Tabernaemontana angulata" OR "Tabernaemontana brasiliensis" OR "Tabernaemontana cumata" OR "Tabernaemontana grandiflora" OR "Tabernaemontana muricata" OR "Tassadia guanchezii" OR "Cathedra paraensis" OR "Ilex petiolaris" OR "Adelonema erytropus" OR "Adelonema wendlandii" OR "Alloschemone inopinata" OR "Anaphyllopsis cururuana" OR "Anthurium curicuriariense" OR "Anthurium icanense" OR "Anthurium krukovii" OR "Anthurium mayoanum" OR "Anthurium pranceanum" OR "Anthurium ptarianum" OR "Anthurium sagittatum" OR "Anthurium vittariifolium" OR "Bognera recondita" OR "Caladium amazonicum" OR "Caladium intermedium" OR "Dieffenbachia parvifolia" OR "Dracontium laetum" OR "Dracontium narae" OR "Dracontium prancei" OR "Heteropsis duckeana" OR "Heteropsis reticulata" OR "Heteropsis spruceana" OR "Homalomena crinipes" OR "Lorenzia umbrosa" OR "Philodendron caracaraiense" OR "Philodendron carajasense" OR "Philodendron cardosoi" OR "Philodendron hopkinsianum" OR "Philodendron joaosilvae" OR "Philodendron lupinum" OR</p>

<sup>20</sup> Disponível em: [https://reflora.jbrj.gov.br/consulta/?grupo=6&familia=null&genero=&especie=&autor=&nomeVernaculo=&nomeCompleto=&formaVida=null&substrato=null&ocorreBrasil=SIM&ocorrencia=SOMENTE\\_OCORRE&endemismo=ENDEMICA\\_BRASIL&origem=TODO&regiao=QUALQUER&ilhaOceanica=32767&estado=QUALQUER&domFitogeograficos=AMAZONIA&vegetacao=TODO&mostrarAte=ESPECIES&opcoesBusca=NOME\\_ACEITO&loginUsuario=Visitante&senhaUsuario=&contexto=consulta-publica&pagina=1#CondicaoTaxonCP](https://reflora.jbrj.gov.br/consulta/?grupo=6&familia=null&genero=&especie=&autor=&nomeVernaculo=&nomeCompleto=&formaVida=null&substrato=null&ocorreBrasil=SIM&ocorrencia=SOMENTE_OCORRE&endemismo=ENDEMICA_BRASIL&origem=TODO&regiao=QUALQUER&ilhaOceanica=32767&estado=QUALQUER&domFitogeograficos=AMAZONIA&vegetacao=TODO&mostrarAte=ESPECIES&opcoesBusca=NOME_ACEITO&loginUsuario=Visitante&senhaUsuario=&contexto=consulta-publica&pagina=1#CondicaoTaxonCP). Acesso em: 14 dez. 2023

Grupo	String para busca
	<p>"Philodendron melloi" OR "Philodendron nullinervium" OR "Philodendron pinnatilobum" OR "Philodendron prancei" OR "Philodendron rio-pretense" OR "Philodendron squamiferum" OR "Philodendron toshibai" OR "Philodendron traunii" OR "Rhodospatha moritziana" OR "Spathiphyllum globulispadix" OR "Xanthosoma appendiculatum" OR "Xanthosoma lucens" OR "Xanthosoma pulchrum" OR "Xanthosoma rubrispathum" OR "Xanthosoma striolatum" OR "Xanthosoma ulei" OR "Zomicarpella amazonica" OR "Crepinella chimantensis" OR "Crepinella gracilis" OR "Crepinella pallens" OR "Dendropanax simplicifolius" OR "Didymopanax ciliatifolius" OR "Didymopanax plurifolius" OR "Didymopanax pubicarpus" OR "Didymopanax umbrosus" OR "Astrocaryum giganteum" OR "Astrocaryum sociale" OR "Attalea guianensis" OR "Bactris syagroides" OR "Bactris tefensis" OR "Barcella odora" OR "Geonoma acaulis" OR "Syagrus vermicularis" OR "Aristolochia lorenae" OR "Cavalcantia glomerata" OR "Guayania penninervata" OR "Mikania longiacuminata" OR "Mikania stygia" OR "Monogereion carajensis" OR "Parapiqueria cavalcantei" OR "Sciadocephala gracielliae" OR "Acmella kalelii" OR "Acmella marajoensis" OR "Aspilia cavalcantei" OR "Clibadium sylvestre" OR "Wedelia rudis" OR "Unxia kubitzkii" OR "Stiffia uniflora" OR "Lepidaploa cleocalderonae" OR "Lepidaploa grisea" OR "Lepidaploa paraensis" OR "Lepidaploa silvae" OR "Piptocarpha prancei" OR "Vernonanthura piresii" OR "Stenopadus aracaensis" OR "Begonia matogrossensis" OR "Adenocalymma adenophorum" OR "Adenocalymma apurense" OR "Adenocalymma calcareum" OR "Adenocalymma cidii" OR "Adenocalymma graciellae" OR "Adenocalymma magnificum" OR "Adenocalymma molle" OR "Adenocalymma velutinum" OR "Amphilophium campinae" OR "Amphilophium lohmanniae" OR "Amphilophium racemosum" OR "Anemopaegma robustum" OR "Cuspidaria cinerea" OR "Cuspidaria monophylla" OR "Fridericia egensis" OR "Fridericia lauta" OR "Fridericia oxycarpa" OR "Jacaranda bullata" OR "Jacaranda campinae" OR "Jacaranda carajasensis" OR "Jacaranda egleri" OR "Jacaranda macrocarpa" OR "Jacaranda obtusifolia" OR "Lundia laevis" OR "Pachyptera aromatica" OR "Pachyptera incarnata" OR "Pleonotoma bracteata" OR "Pleonotoma fissicalyx" OR "Pleonotoma longiflora" OR "Pyrostegia millingtonioides" OR "Tanaecium decorticans" OR "Tynanthus densiflorus" OR "Aechmea egleriana" OR "Aechmea fernandae" OR "Aechmea murcae" OR "Aechmea polyantha" OR "Aechmea rodriguesiana" OR "Aechmea vallerandii" OR "Aechmea xinguana" OR "Billbergia dasilvae" OR "Brocchinia amazonica" OR "Bromelia araujoi" OR "Bromelia epiphytica" OR "Bromelia gracilisepala" OR "Bromelia morreniana" OR "Bromelia oliveirae" OR "Bromelia rondoniana" OR "Dyckia paraensis" OR "Dyckia sickii" OR "Dyckia silvae" OR "Fosterella batistiana" OR "Guzmania nubicola" OR "Guzmania vittata" OR "Lindmania maguirei" OR "Lindmania piresii" OR "Mezobromelia pleiosticha" OR "Navia corrugata" OR "Navia eldorado" OR "Navia piresii" OR "Navia tenuifolia" OR "Neoregelia margaretae" OR "Pitcairnia anomala" OR "Pitcairnia buscalionii" OR "Pitcairnia crinita" OR "Pitcairnia egleri" OR "Pitcairnia frequens" OR "Pitcairnia neeana" OR "Pitcairnia rondonicola" OR "Werauhia hylaeana" OR "Gymnosiphon saccatus" OR "Dacryodes edilsonii" OR "Dacryodes froesiana" OR "Dacryodes spatulata" OR "Protium cordatum" OR "Protium herisonii" OR "Protium pedicellatum" OR "Protium pilosellum" OR "Protium tonyanum" OR "Trattinnickia dalyana" OR "Trattinnickia zickeliana" OR "Cabomba schwartzii" OR "Melocactus estevesii" OR "Caraipa aracaensis" OR "Caraipa balbinensis" OR "Caraipa foveolata" OR "Caraipa glabra" OR "Caraipa iracemensis" OR "Caraipa longisepala" OR "Caraipa minor" OR "Caraipa multinervia" OR "Caraipa odorata" OR "Caraipa rodriguesii" OR "Clusiella impressinervis" OR "Haploclathra grandiflora" OR "Haploclathra verticillata" OR "Kielmeyera oreophila" OR "Celtis schippii" OR "Cheiloclinium puberulum" OR "Monteverdia laurina" OR "Peritassa manaora" OR "Peritassa petiolata" OR "Acioa edulis" OR "Acioa longipendula" OR "Couepia cataractae" OR "Couepia cidiana" OR "Couepia eriantha" OR "Couepia froesii" OR "Couepia glabra" OR "Couepia marleneae" OR "Couepia morii" OR "Couepia reflexa" OR "Couepia robusta" OR "Couepia spicata" OR "Couepia stipularis" OR "Gaulettia amaraliae" OR "Hirtella arenosa" OR "Hirtella barnebyi" OR "Hirtella conduplicata" OR "Hirtella dorvalii" OR "Hirtella fasciculata" OR "Hirtella kuhlmannii" OR "Hirtella lancifolia" OR "Hirtella longifolia" OR "Hirtella myrmecophila" OR "Hirtella paraensis" OR "Hirtella radamii" OR "Hirtella scaberula" OR "Hymenopus adolphoduckei" OR "Hymenopus conferruminatus" OR "Hymenopus miltonii" OR "Hymenopus sothersiae" OR "Leptobalanus joseamosii" OR "Licania apiculata" OR "Licania aracaensis" OR "Licania bellingtonii" OR "Licania ferreirae" OR "Licania impressa" OR "Licania marleneae" OR "Licania maxima" OR "Licania nelsonii" OR "Licania teixeirae" OR "Licania tocantina" OR "Moquilea anaeae" OR "Moquilea fritschii" OR "Moquilea maranhensis" OR "Parinari cardiophylla" OR "Clusia lutea" OR "Clusia nigrolineata" OR "Clusia nitida" OR "Clusia polysepala" OR "Garcinia albuquerquei" OR "Lorostemon coelhoi" OR "Lorostemon negrense" OR "Moronobea pulchra" OR "Tovomita cornuta" OR "Tovomita duckei" OR "Tovomita hopkinsii" OR "Tovomita longirostrata" OR "Tovomita manauara" OR "Tovomita speciosa" OR "Tovomita vismiifolia" OR "Combretum rotundifolium" OR "Terminalia congesta" OR "Terminalia duckei" OR "Terminalia fanshawei" OR "Terminalia macrophylla" OR "Terminalia</p>

Grupo	String para busca
	<p>ochroprumna" OR "Terminalia pallidovirens" OR "Terminalia pulcherrima" OR "Terminalia suaveolens" OR "Terminalia viridiflora" OR "Cnestidium froesii" OR "Connarus angustifolius" OR "Connarus manausensis" OR "Connarus marleneae" OR "Connarus ramiflorus" OR "Rourea duckei" OR "Rourea kappleri" OR "Rourea ligulata" OR "Rourea paraensis" OR "Rourea pubescens" OR "Bonamia ferruginea" OR "Dicranostyles falconiana" OR "Dicranostyles solimoesensis" OR "Dicranostyles yrypoana" OR "Ipomoea cavalcantei" OR "Ipomoea marabaensis" OR "Ipomoea spruceana" OR "Maripa elongata" OR "Cordia decipiens" OR "Cordia ripicola" OR "Cordia ulei" OR "Chamaecostus fragilis" OR "Chamaecostus fusiformis" OR "Costus fortalezae" OR "Costus juruanus" OR "Costus sprucei" OR "Cayaponia duckei" OR "Gurania malacophylla" OR "Gurania spruceana" OR "Gurania velutina" OR "Asplundia brasiliensis" OR "Asplundia krukoffii" OR "Asplundia latifrons" OR "Bisboeckelera longifolia" OR "Bulbostylis cangae" OR "Bulbostylis graminifolia" OR "Cyperus bampsii" OR "Eleocharis ayacuchensis" OR "Eleocharis pedroviana" OR "Eleocharis tenuiculmis" OR "Rhynchospora acanthoma" OR "Rhynchospora paraensis" OR "Rhynchospora ruiziana" OR "Rhynchospora seccoi" OR "Scleria pusilla" OR "Scleria scandens" OR "Dichapetalum coelhoi" OR "Dichapetalum pauper" OR "Tapura lanceolata" OR "Tapura singularis" OR "Davilla cuspidulata" OR "Davilla lanosa" OR "Davilla neei" OR "Davilla pedicellaris" OR "Davilla strigosa" OR "Doliocarpus aracaensis" OR "Doliocarpus areolatus" OR "Doliocarpus aureobaccus" OR "Doliocarpus humboldtianus" OR "Doliocarpus prancei" OR "Doliocarpus pruskii" OR "Dioscorea lacerdae" OR "Dioscorea neblinensis" OR "Dioscorea traillii" OR "Drosera amazonica" OR "Diospyros akaraiensis" OR "Diospyros arupaj" OR "Diospyros bullata" OR "Diospyros cachimboensis" OR "Diospyros guatterioides" OR "Diospyros kondor" OR "Diospyros krukovii" OR "Diospyros landii" OR "Diospyros longistyla" OR "Diospyros manausensis" OR "Diospyros santaremnensis" OR "Diospyros tarim" OR "Diospyros trombetensis" OR "Diospyros uaupensis" OR "Rotula pohlii" OR "Sloanea dubia" OR "Sloanea erismoides" OR "Sloanea froesii" OR "Sloanea heteroneura" OR "Sloanea parva" OR "Sloanea pseudovorticillata" OR "Gaylussacia amazonica" OR "Vaccinium pipolyi" OR "Eriocaulon carajense" OR "Eriocaulon herzogii" OR "Eriocaulon milhoense" OR "Eriocaulon multiscapum" OR "Paepalanthus cururensis" OR "Paepalanthus fasciculatus" OR "Paepalanthus piresii" OR "Paepalanthus septentrionalis" OR "Paepalanthus singularis" OR "Syngonanthus cachimboensis" OR "Syngonanthus discretifolius" OR "Syngonanthus egleri" OR "Syngonanthus heteropepoides" OR "Syngonanthus lanatus" OR "Syngonanthus sickii" OR "Syngonanthus xinguensis" OR "Heisteria amazonica" OR "Heisteria amphoricarpa" OR "Heisteria huberiana" OR "Heisteria laxiflora" OR "Erythroxyllum campinense" OR "Erythroxyllum carajasense" OR "Erythroxyllum cordato-ovatum" OR "Erythroxyllum nelson-rosae" OR "Erythroxyllum timothei" OR "Acalypha acuminata" OR "Angostylis longifolia" OR "Anomalocalyx uleanus" OR "Astrococcus cornutus" OR "Caryodendron amazonicum" OR "Chicomendes rubiginosus" OR "Conceveiba prealta" OR "Croton abonari" OR "Croton arirambae" OR "Croton ascendens" OR "Croton asperimus" OR "Croton borbensis" OR "Croton campinarenis" OR "Croton caryophyllus" OR "Croton diasii" OR "Croton dissectistipulatus" OR "Croton faroensis" OR "Croton gigantifolius" OR "Croton paraensis" OR "Croton sacaquinha" OR "Croton spruceanus" OR "Croton subasperimus" OR "Croton trombetensis" OR "Croton viroleoides" OR "Dalechampia gentryi" OR "Dalechampia hastata" OR "Dalechampia juruana" OR "Dalechampia olympiana" OR "Dendrothrix wurdackii" OR "Euphorbia duckei" OR "Hevea camargoana" OR "Hevea camporum" OR "Hevea spruceana" OR "Joannesia heveoides" OR "Mabea ovata" OR "Mabea uleana" OR "Manihot leptophylla" OR "Manihot marajoara" OR "Micrandra lopezii" OR "Micrandropsis scleroxylon" OR "Pausandra macropetala" OR "Pausandra trianae" OR "Sagotia heterocalyx" OR "Sapium pallidum" OR "Senefeldera trianae" OR "Tetrorchidium duckei" OR "Abarema acreana" OR "Abarema campestris" OR "Abarema cochleata" OR "Abarema piresii" OR "Albizia decandra" OR "Aldina heterophylla" OR "Aldina microphylla" OR "Aldina occidentalis" OR "Aldina polyphylla" OR "Alexa bauhiniiflora" OR "Alexa confusa" OR "Alexa grandiflora" OR "Andira macrothyrsa" OR "Andira micrantha" OR "Andira multistipula" OR "Andira parviflora" OR "Andira praecox" OR "Andira unifoliolata" OR "Androcalymma glabrifolium" OR "Bauhinia aureopunctata" OR "Bauhinia brachycalyx" OR "Bauhinia grandifolia" OR "Bauhinia longipedicellata" OR "Bauhinia piresii" OR "Calliandra jariensis" OR "Campsiandra robclarkiana" OR "Canavalia obidensis" OR "Cassia rubriflora" OR "Cenostigma tocantinum" OR "Centrosema carajasense" OR "Chamaecrista aiarana" OR "Chamaecrista didyma" OR "Chamaecrista egleri" OR "Chamaecrista ipanorensis" OR "Chamaecrista scleroxylon" OR "Chamaecrista xinguensis" OR "Clitoria cavalcantei" OR "Clitoria obidensis" OR "Clitoria tunuhiensis" OR "Copaifera glycyarpa" OR "Copaifera piresii" OR "Cynometra cerebriiformis" OR "Cynometra duckei" OR "Cynometra longifolia" OR "Cynometra macrocarpa" OR "Cynometra marleneae" OR "Cynometra stenopetala" OR "Dalbergia grandistipula" OR "Dalbergia guttembergii" OR "Dalbergia monophylla" OR "Deguelia decorticans" OR "Deguelia duckeana" OR "Deguelia glaucifolia" OR "Deguelia occidentalis" OR "Deguelia tenuiflora" OR</p>

Grupo	String para busca
	<p>"Dialium rondoniense" OR "Dicymbe amazonica" OR "Dicymbe arenicola" OR "Dicymbe heteroxylon" OR "Dicymbe puncticulosa" OR "Dimorphandra campinarum" OR "Dimorphandra caudata" OR "Dimorphandra mediocris" OR "Dimorphandra parviflora" OR "Dimorphandra urubuensis" OR "Dioclea fimbriata" OR "Diploctropis rodriguesii" OR "Discolobium tocaninum" OR "Elizabetha bicolor" OR "Elizabetha durissima" OR "Elizabetha paraensis" OR "Elizabetha speciosa" OR "Entada simplicata" OR "Eperua duckeana" OR "Eperua glabriflora" OR "Eperua oleifera" OR "Eperua praesagata" OR "Heterostemon impar" OR "Hydrochorea uaupensis" OR "Hymenaea jeaniana" OR "Hymenaea reticulata" OR "Hymenolobium excelsum" OR "Hymenolobium grazielanum" OR "Hymenolobium heterocarpum" OR "Hymenolobium modestum" OR "Hymenolobium sericeum" OR "Inga bicoloriflora" OR "Inga bullatorugosa" OR "Inga calantha" OR "Inga caudata" OR "Inga chartacea" OR "Inga exilis" OR "Inga leptocarpa" OR "Inga microcalyx" OR "Inga obtusata" OR "Inga panurensis" OR "Inga retinocarpa" OR "Inga salicifoliola" OR "Inga santaremnensis" OR "Inga stipularis" OR "Inga suberosa" OR "Inga xinguensis" OR "Jacqueshuberia purpurea" OR "Jacqueshuberia quinquangulata" OR "Luetzelburgia amazonica" OR "Machaerium aureiflorum" OR "Machaerium castaneiflorum" OR "Machaerium caudatum" OR "Machaerium complanatum" OR "Machaerium costulatum" OR "Machaerium froesii" OR "Machaerium hoehneanum" OR "Machaerium piresii" OR "Machaerium trifoliolatum" OR "Macrolobium aracaense" OR "Macrolobium brevense" OR "Macrolobium cidii" OR "Macrolobium conjunctum" OR "Macrolobium cowanii" OR "Macrolobium duckeanum" OR "Macrolobium froesii" OR "Macrolobium furcatum" OR "Macrolobium longipedicellatum" OR "Macrolobium palustre" OR "Macrolobium parvifolium" OR "Macrolobium prancei" OR "Macrolobium urupaense" OR "Macropsychanthus duckei" OR "Macropsychanthus flexuosus" OR "Macrosamanea duckei" OR "Macrosamanea macrocalyx" OR "Macrosamanea prancei" OR "Mimosa calliandroides" OR "Mimosa dasilvae" OR "Monopteryx angustifolia" OR "Mora paraensis" OR "Muellera denudata" OR "Muellera tozziana" OR "Nissolia rondonensis" OR "Ormosia froesii" OR "Ormosia paraensis" OR "Ormosia santaremnensis" OR "Paloue induta" OR "Panurea bowdichioides" OR "Paramachaerium krukovii" OR "Parkia lutea" OR "Parkia paraensis" OR "Peltogyne altissima" OR "Peltogyne campestris" OR "Peltogyne catinae" OR "Peltogyne excelsa" OR "Peltogyne gracilipes" OR "Peltogyne lecointei" OR "Peltogyne micrantha" OR "Peltogyne paradoxa" OR "Petaladenium urceoliferum" OR "Platymiscium filipes" OR "Recordoxylon stenopetalum" OR "Schnella alata" OR "Schnella altiscandens" OR "Schnella erythrantha" OR "Schnella longiseta" OR "Schnella platycalyx" OR "Schnella sprucei" OR "Schnella stenopetala" OR "Senna cornigera" OR "Staminodianthus rosae" OR "Stryphnodendron conicum" OR "Stryphnodendron procerum" OR "Stryphnodendron riparium" OR "Swartzia amazonica" OR "Swartzia arumateuana" OR "Swartzia conferta" OR "Swartzia coriaceifolia" OR "Swartzia corrugata" OR "Swartzia discocarpa" OR "Swartzia duckei" OR "Swartzia emarginata" OR "Swartzia fimbriata" OR "Swartzia fraterna" OR "Swartzia ingens" OR "Swartzia ingifolia" OR "Swartzia krukovii" OR "Swartzia lamellata" OR "Swartzia lanata" OR "Swartzia laurifolia" OR "Swartzia longistipitata" OR "Swartzia lucida" OR "Swartzia macrocarpa" OR "Swartzia manausensis" OR "Swartzia martii" OR "Swartzia obscura" OR "Swartzia oriximinaensis" OR "Swartzia permitida" OR "Swartzia prolata" OR "Swartzia recurva" OR "Swartzia rugosa" OR "Swartzia tomentifera" OR "Swartzia ulei" OR "Swartzia velutina" OR "Tachigali argyrophylla" OR "Tachigali candelabrum" OR "Tachigali catinae" OR "Tachigali fusca" OR "Tachigali goeldiana" OR "Tachigali grandistipulata" OR "Tachigali hypoleuca" OR "Tachigali leiocalyx" OR "Tachigali macropetala" OR "Tachigali macrostachya" OR "Tachigali melanocarpa" OR "Tachigali multijuga" OR "Tachigali prancei" OR "Tachigali schultesiana" OR "Trischidium racemosum" OR "Vataireopsis iglesiasii" OR "Vouacapoua pallidior" OR "Zygia transamazonica" OR "Zygia trunciflora" OR "Aripuana cullmaniorum" OR "Irlbachia nemorosa" OR "Macrocarpaea piresii" OR "Roraimaea aurantiaca" OR "Tachia grandiflora" OR "Tachia lancisepala" OR "Tachia siwertii" OR "Tachia smithii" OR "Voyria alvesiana" OR "Yanomamua araca" OR "Chrysothemis kuhlmannii" OR "Chrysothemis rupestris" OR "Lesia tepuiensis" OR "Nautilocalyx decumbens" OR "Nautilocalyx mulfordii" OR "Sinningia minima" OR "Heliconia adeliana" OR "Duckesia verrucosa" OR "Endopleura uchi" OR "Humiriastrium purusensis" OR "Hylocarpa heterocarpa" OR "Sacoglottis ceratocarpa" OR "Schistostemon macrophyllum" OR "Vantanea aracaensis" OR "Vantanea deniseae" OR "Vantanea macrocarpa" OR "Vantanea micrantha" OR "Vantanea tuberculata" OR "Vismia cauliflora" OR "Vismia cavalcantei" OR "Casimirella ampla" OR "Casimirella rupestris" OR "Aegiphila duckei" OR "Aegiphila exiguiflora" OR "Aegiphila froesii" OR "Aegiphila goeldiana" OR "Gymneia moniliformis" OR "Hyptidendron dorothyianum" OR "Hyptis hispida" OR "Hyptis tumidicalyx" OR "Scutellaria alborosea" OR "Scutellaria leucantha" OR "Vitex brevilabiata" OR "Vitex duckei" OR "Vitex excelsa" OR "Vitex froesii" OR "Vitex krukovii" OR "Vitex maranhana" OR "Vitex odorata" OR "Vitex pseudolea" OR "Vitex snethlagiana" OR "Vitex spongiocarpa" OR "Vitex sprucei" OR "Aniba burchellii" OR "Aniba parviflora" OR "Aniba rosiodora" OR "Aniba santalodora" OR</p>

Grupo	String para busca
	<p>"Dicypellium caryophyllaceum" OR "Dicypellium manausense" OR "Endlicheria bullata" OR "Endlicheria coriacea" OR "Endlicheria directonervia" OR "Endlicheria goeldiana" OR "Endlicheria krukovii" OR "Endlicheria longicaudata" OR "Endlicheria robusta" OR "Licaria hirsuta" OR "Licaria puchury-major" OR "Licaria rodriguesii" OR "Mezilaurus caatingae" OR "Mezilaurus decurrens" OR "Mezilaurus duckei" OR "Mezilaurus manausensis" OR "Mezilaurus micrantha" OR "Nectandra grisea" OR "Nectandra impressa" OR "Ocotea albescens" OR "Ocotea bilocellata" OR "Ocotea delicata" OR "Ocotea douradensis" OR "Ocotea duplocolorata" OR "Ocotea fragrantissima" OR "Ocotea ligulata" OR "Ocotea minor" OR "Ocotea nigrescens" OR "Ocotea obliqua" OR "Ocotea silvae" OR "Paraia bracteata" OR "Persea benthamiana" OR "Persea jariensis" OR "Pleurothyrium prancei" OR "Pleurothyrium undulatum" OR "Rhodostemonodaphne curicuriarensis" OR "Rhodostemonodaphne negrensis" OR "Rhodostemonodaphne parvifolia" OR "Rhodostemonodaphne peneia" OR "Rhodostemonodaphne recurva" OR "Allantoma integrifolia" OR "Allantoma kuhlmannii" OR "Allantoma pachyantha" OR "Allantoma pauciramosa" OR "Allantoma uaupensis" OR "Cariniana penduliflora" OR "Corytophora alta" OR "Couratari longipedicellata" OR "Couratari prancei" OR "Couratari tauari" OR "Eschweilera amazonica" OR "Eschweilera amazoniciformis" OR "Eschweilera atropetiolata" OR "Eschweilera carinata" OR "Eschweilera cyathiformis" OR "Eschweilera grandiflora" OR "Eschweilera obversa" OR "Eschweilera pseudodecolorans" OR "Eschweilera rabeliana" OR "Eschweilera rankinae" OR "Eschweilera rhododendrifolia" OR "Eschweilera romeu-cardosoi" OR "Eschweilera subcordata" OR "Eschweilera truncata" OR "Gustavia elliptica" OR "Gustavia erythrocarpa" OR "Gustavia longepetiolata" OR "Gustavia longifolia" OR "Lecythis barnebyi" OR "Lecythis graciana" OR "Lecythis idatimon" OR "Lecythis parvifructa" OR "Lecythis prancei" OR "Lecythis retusa" OR "Lecythis serrata" OR "Genlisea multiflora" OR "Utricularia ariramba" OR "Utricularia jaramacaru" OR "Utricularia physoceras" OR "Bonyunia magnifica" OR "Bonyunia venusta" OR "Spigelia megapotamica" OR "Spigelia spruceana" OR "Strychnos froesii" OR "Strychnos macrophylla" OR "Strychnos malacosperma" OR "Strychnos neglecta" OR "Strychnos pachycarpa" OR "Passovia bracteata" OR "Passovia brasiliana" OR "Psittacanthus atrolineatus" OR "Psittacanthus dentatus" OR "Psittacanthus ovatus" OR "Struthanthus gracilis" OR "Struthanthus prancei" OR "Cuphea alatosperma" OR "Cuphea carajasensis" OR "Cuphea egleri" OR "Cuphea exilis" OR "Cuphea rionegrensis" OR "Cuphea sabulosa" OR "Acmathera cowanii" OR "Acmathera duckei" OR "Acmathera latifolia" OR "Acmathera longifolia" OR "Acmathera minima" OR "Acmathera parviflora" OR "Alicia macrodisca" OR "Banisteriopsis schwannioides" OR "Blepharandra cachimbensis" OR "Blepharandra intermedia" OR "Bronwenia brevipedicellata" OR "Burdachia duckei" OR "Byrsonima delicatula" OR "Byrsonima euryphylla" OR "Byrsonima frondosa" OR "Byrsonima hirsuta" OR "Byrsonima piresii" OR "Byrsonima rodriguesii" OR "Byrsonima souzae" OR "Christianella paludicola" OR "Diacidia aracaensis" OR "Diplopterys cachimbensis" OR "Diplopterys cururuensis" OR "Diplopterys rondoniense" OR "Glandonia prancei" OR "Heteropterys catoptera" OR "Heteropterys ciliata" OR "Heteropterys marleneae" OR "Heteropterys mathewsana" OR "Heteropterys murcapiresii" OR "Hiraea holmgreniorum" OR "Hiraea silvae" OR "Hiraea silvicola" OR "Janusia paraensis" OR "Lophanthera lactescens" OR "Lophanthera pendula" OR "Lophanthera spruceana" OR "Malpighiodes bracteosa" OR "Malpighiodes leucanthele" OR "Mascagnia riparia" OR "Mascagnia tucuruensis" OR "Mezia sericea" OR "Pterandra arborea" OR "Pterandra egleri" OR "Pterandra evansii" OR "Stigmaphyllon stylopogon" OR "Tetrapterys cordifolia" OR "Verrucularina piresii" OR "Aguaria excelsa" OR "Apeiba trombetensis" OR "Bastardiopsis grewiiifolia" OR "Catostemma albuquerquei" OR "Catostemma cavalcantei" OR "Catostemma grazielae" OR "Catostemma lanceolatum" OR "Catostemma milanezii" OR "Catostemma sclerophyllum" OR "Eriotheca longipedicellata" OR "Gaya scopulorum" OR "Helicteres lenta" OR "Helicteres urupensis" OR "Herrania nitida" OR "Hibiscus amazonicus" OR "Lueheopsis burretiana" OR "Matisia bicolor" OR "Matisia lecythicarpa" OR "Mollia cuneata" OR "Mollia longifolia" OR "Mollia macrophylla" OR "Mollia nitida" OR "Mollia paraensis" OR "Pachira manausensis" OR "Pachira tocantina" OR "Patinoa paraensis" OR "Patinoa sphaerocarpa" OR "Pavonia paraensis" OR "Pavonia piptocalyx" OR "Quararibea amazonica" OR "Quararibea spatulata" OR "Scleronema spruceana" OR "Sida marabaensis" OR "Sterculia albidiflora" OR "Sterculia apeibophylla" OR "Sterculia megalocarpa" OR "Sterculia pendula" OR "Sterculia rigidifolia" OR "Sterculia stipulifera" OR "Theobroma canumanense" OR "Wissadula cuspidata" OR "Goepertia crocata" OR "Goepertia lagoagriana" OR "Goepertia loeseneri" OR "Ischnosiphon crassispicus" OR "Ischnosiphon paryrizinho" OR "Maranta sophiana" OR "Monotagma breviscapum" OR "Monotagma congestum" OR "Monotagma floribundum" OR "Monotagma humile" OR "Monotagma lilacinum" OR "Marcgravia punctifolia" OR "Aciotis ferreirana" OR "Aciotis wurdackiana" OR "Bellucia egensis" OR "Bellucia nigricans" OR "Bellucia riparia" OR "Bellucia wurdackiana" OR "Brasilianthus carajensis" OR "Clidemia ferox" OR "Comolia amazonica" OR "Comolia berberifolia" OR "Graffenrieda maturaca" OR "Graffenrieda moensis" OR</p>

Grupo	String para busca
	<p>"Leandra coadunata" OR "Macairea cuieirasii" OR "Macairea theresiae" OR "Macrocentrum aurimontium" OR "Meriania aracaensis" OR "Meriania neblinensis" OR "Miconia comptifolia" OR "Miconia cowanii" OR "Miconia forzzae" OR "Miconia manauara" OR "Miconia mattogrossensis" OR "Miconia mazanana" OR "Miconia navioensis" OR "Miconia papillosperma" OR "Miconia rondoniensis" OR "Miconia santaremensis" OR "Miconia schwackei" OR "Miconia secundiflora" OR "Miconia suberosa" OR "Miconia subsimplex" OR "Miconia umbrosa" OR "Miconia waimiri-atroari" OR "Microlicia sickii" OR "Mouriri brachyanthera" OR "Mouriri dimorphandra" OR "Mouriri ducleanoides" OR "Mouriri eugeniifolia" OR "Mouriri froesii" OR "Mouriri lunatanthera" OR "Mouriri micradenia" OR "Mouriri monopora" OR "Mouriri obtusiloba" OR "Mouriri spruceana" OR "Mouriri torquata" OR "Mouriri trunciflora" OR "Pleroma carajasense" OR "Poteranthera windischii" OR "Pterolepis picorondonica" OR "Siphanthera todziae" OR "Tibouchina edmundoi" OR "Votomita orbinaxia" OR "Votomita plerocarpa" OR "Votomita roraimensis" OR "Guarea crispata" OR "Guarea humaitensis" OR "Guarea sprucei" OR "Trichilia amapaensis" OR "Trichilia bullata" OR "Trichilia fasciculata" OR "Trichilia minima" OR "Trichilia septentrionalis" OR "Caryomene glaucescens" OR "Caryomene olivascens" OR "Odontocarya froesii" OR "Odontocarya krukoviana" OR "Odontocarya membranacea" OR "Odontocarya syncretica" OR "Sciadotenia javariensis" OR "Sciadotenia pachnococca" OR "Sciadotenia paraensis" OR "Telitoxicum negroensis" OR "Telitoxicum rodriguesii" OR "Emmotum acuminatum" OR "Emmotum floribundum" OR "Mollinedia grazielae" OR "Brosimum longifolium" OR "Castilla ulei" OR "Ficus blepharophylla" OR "Ficus christianii" OR "Naucleopsis insculptula" OR "Naucleopsis jamariensis" OR "Naucleopsis pauciflora" OR "Naucleopsis stipularis" OR "Naucleopsis tubulata" OR "Compsonera racemosa" OR "Iryanthera campinae" OR "Iryanthera dialyandra" OR "Virola crebrinervia" OR "Virola guggenheimii" OR "Calycolpus andersonii" OR "Campomanesia cucullata" OR "Eugenia agathopoda" OR "Eugenia belemitana" OR "Eugenia breviracemosa" OR "Eugenia caducibracteata" OR "Eugenia calva" OR "Eugenia campina" OR "Eugenia denigrata" OR "Eugenia dentata" OR "Eugenia fortuita" OR "Eugenia froesii" OR "Eugenia illepida" OR "Eugenia lasiothecia" OR "Eugenia longiracemosa" OR "Eugenia manausensis" OR "Eugenia pallidopunctata" OR "Eugenia percincta" OR "Eugenia stenosepaloides" OR "Myrcia aliena" OR "Myrcia amanana" OR "Myrcia amapensis" OR "Myrcia amoena" OR "Myrcia breviflora" OR "Myrcia caloneura" OR "Myrcia cantana" OR "Myrcia castanea" OR "Myrcia corticosa" OR "Myrcia cuprea" OR "Myrcia dispar" OR "Myrcia divisoria" OR "Myrcia elevata" OR "Myrcia eveae" OR "Myrcia hoffmannseggii" OR "Myrcia incompleta" OR "Myrcia iranduba" OR "Myrcia irregularis" OR "Myrcia javariana" OR "Myrcia laxa" OR "Myrcia longiramea" OR "Myrcia macaca" OR "Myrcia magna" OR "Myrcia manausensis" OR "Myrcia maraana" OR "Myrcia melanophylla" OR "Myrcia neoschomburgkiana" OR "Myrcia oreophila" OR "Myrcia otocalyx" OR "Myrcia piptocalyx" OR "Myrcia prismatica" OR "Myrcia psammophila" OR "Myrcia rufotomentosa" OR "Myrcia scoparia" OR "Myrcia sticta" OR "Myrcia summa" OR "Myrcia symmetrica" OR "Myrcia tarauacana" OR "Myrcia telephylla" OR "Myrcia tridymantha" OR "Myrcia uaioai" OR "Myrciaria cordata" OR "Plinia clausa" OR "Plinia humaitana" OR "Plinia rufiflora" OR "Plinia tapuruquarana" OR "Neea campanulata" OR "Neea rubescens" OR "Neea uleana" OR "Nymphaea paganuccii" OR "Nymphaea potamophila" OR "Nymphaea rapinii" OR "Blastemanthus grandiflorus" OR "Ouratea acuminata" OR "Ouratea barrae" OR "Ouratea cassinifolia" OR "Ouratea chrysopetala" OR "Ouratea cidiana" OR "Ouratea coccinea" OR "Ouratea duckei" OR "Ouratea inundata" OR "Ouratea javariensis" OR "Ouratea patens" OR "Ouratea pendulosepala" OR "Ouratea subcaudata" OR "Ouratea thyrsoides" OR "Ouratea verruculosa" OR "Ouratea vieirae" OR "Sauvagesia laciniata" OR "Wallacea multiflora" OR "Ptychopetalum uncinatum" OR "Acianthera saracataquerensis" OR "Acianthera silvae" OR "Anathallis amazonica" OR "Anathallis fastigiata" OR "Anathallis manausensis" OR "Anathallis roseopapillosa" OR "Anathallis taracuana" OR "Brassavola fasciculata" OR "Caluera tavaresii" OR "Campylocentrum amazonicum" OR "Campylocentrum tenue" OR "Catasetum aculeatum" OR "Catasetum alatum" OR "Catasetum albobovirens" OR "Catasetum albuquerquei" OR "Catasetum ×altaflorestense" OR "Catasetum ×apolloi" OR "Catasetum aripuanense" OR "Catasetum ariquemense" OR "Catasetum bifidum" OR "Catasetum ×canaense" OR "Catasetum carolinianum" OR "Catasetum carrenhianum" OR "Catasetum caxarariense" OR "Catasetum colidense" OR "Catasetum cucullatum" OR "Catasetum dalastranum" OR "Catasetum ×dasilvae" OR "Catasetum denticulatum" OR "Catasetum ×freitasii" OR "Catasetum galeritum" OR "Catasetum garnettianum" OR "Catasetum gnomus" OR "Catasetum hopkinsonianum" OR "Catasetum ×issanense" OR "Catasetum ivanae" OR "Catasetum kleberianum" OR "Catasetum kraenzlinianum" OR "Catasetum lanceatum" OR "Catasetum longipes" OR "Catasetum ×macedoi" OR "Catasetum meeae" OR "Catasetum ×mesquiae" OR "Catasetum mojuense" OR "Catasetum multifidum" OR "Catasetum ×nogueirae" OR "Catasetum oriximinaense" OR "Catasetum osakadianum" OR "Catasetum palmeirinhense" OR "Catasetum paranaitense" OR "Catasetum ×perazolianum" OR "Catasetum pulchrum" OR "Catasetum rigidum" OR "Catasetum rivularium" OR</p>

Grupo	String para busca
	<p>"Catasetum rolfeanum" OR "Catasetum schmidtianum" OR "Catasetum seccoii" OR "Catasetum semicirculatum" OR "Catasetum sophiae" OR "Catasetum teixeiranum" OR "Catasetum telespirense" OR "Catasetum tigrinum" OR "Catasetum tomasellii" OR "Catasetum ×valdisonianum" OR "Cattleya araguaiensis" OR "Coryanthes cavalcantei" OR "Coryanthes dasilvae" OR "Coryanthes elegantium" OR "Coryanthes elianae" OR "Coryanthes minima" OR "Coryanthes miuaensis" OR "Coryanthes pacaraimensis" OR "Cycnoches manoelae" OR "Dichaea cornuta" OR "Dichaea integrilabia" OR "Dichaea picta" OR "Dichaea pumila" OR "Dichaea rodriguesii" OR "Dryadella cardosoi" OR "Dryadella osmariniana" OR "Eltroplectris guimaraesii" OR "Encyclia yauaperyensis" OR "Epidendrum acreense" OR "Epidendrum pareciense" OR "Epidendrum rondoniense" OR "Galeandra santarenensis" OR "Gongora atropurpurea" OR "Gongora jauariensis" OR "Gongora juruaensis" OR "Habenaria achroantha" OR "Habenaria sylvicultrix" OR "Ionopsis burchellii" OR "Lepanthes brasiliensis" OR "Lockhartia ivainae" OR "Macroclinium bragae" OR "Macroclinium brasiliense" OR "Macroclinium paraense" OR "Macroclinium saraca-taquerense" OR "Macroclinium subroseum" OR "Maxillaria equitans" OR "Maxillaria yauaperyensis" OR "Mormodes aurea" OR "Mormodes buccinator" OR "Mormodes castroi" OR "Mormodes cucumerina" OR "Mormodes dasilvae" OR "Mormodes densiflora" OR "Mormodes elegans" OR "Mormodes gurupiensis" OR "Mormodes hoehnei" OR "Mormodes issanensis" OR "Mormodes jamanxinensis" OR "Mormodes kleberiana" OR "Mormodes matogrossensis" OR "Mormodes paraensis" OR "Mormodes rodriguesiana" OR "Mormodes tapoayensis" OR "Mormodes tigrina" OR "Mormodes vernixioidea" OR "Notylia angustifolia" OR "Notylia bispala" OR "Notylia flexuosa" OR "Notylia platyglossa" OR "Notylia yauaperyensis" OR "Octomeria napoleon" OR "Octomeria pygmaea" OR "Octomeria serpens" OR "Orleanesia cuneipetala" OR "Palmorchis caxiuanaensis" OR "Palmorchis duckei" OR "Palmorchis sobralioides" OR "Paphinia grandiflora" OR "Peristeria serroniana" OR "Platystele edmundoi" OR "Platystele paraensis" OR "Plectrophora schmidtii" OR "Rodríguezia carnea" OR "Rodríguezia huebneri" OR "Rodríguezia joesiana" OR "Rodríguezia negrensis" OR "Selenipedium isabelianum" OR "Sobralia cardosoi" OR "Sobralia imavieirae" OR "Sobralia malmiana" OR "Sobralia margaritae" OR "Sobralia pumila" OR "Trichosalpinx cryptantha" OR "Uleiorchis longipedicellata" OR "Vanilla hartii" OR "Vanilla karen-christianae" OR "Vanilla labellopapillata" OR "Xerorchis trichorhiza" OR "Buchnera carajasensis" OR "Buchnera nordestina" OR "Biophytum dormiens" OR "Biophytum somnians" OR "Dilkea retusa" OR "Passiflora araujoi" OR "Passiflora cordistipula" OR "Passiflora cryptopetala" OR "Passiflora echinasteris" OR "Passiflora ernestii" OR "Passiflora faroana" OR "Passiflora fissurosa" OR "Passiflora hexagonocarpa" OR "Passiflora lorenziana" OR "Passiflora menispermiifolia" OR "Passiflora quadrifaria" OR "Passiflora riparia" OR "Passiflora tholozanii" OR "Ternstroemia arcaeae" OR "Ternstroemia dehiscens" OR "Ternstroemia krukoffiana" OR "Ternstroemia prancei" OR "Ternstroemia subcaudata" OR "Ternstroemia urophora" OR "Pera coccinea" OR "Pera eiteniorum" OR "Pera manausensis" OR "Pera membranacea" OR "Pera pulchrifolia" OR "Amanoa almerindae" OR "Amanoa congesta" OR "Amanoa cupatensis" OR "Amanoa gracillima" OR "Amanoa marapiensis" OR "Amanoa sinuosa" OR "Chonocentrum cyathophorum" OR "Phyllanthus aracaensis" OR "Phyllanthus biantherifer" OR "Phyllanthus madeirensis" OR "Phyllanthus martii" OR "Phyllanthus spruceanus" OR "Picramnia ferrea" OR "Peperomia acreana" OR "Peperomia albopilosa" OR "Peperomia alegrensis" OR "Peperomia blackii" OR "Peperomia cruzeirensis" OR "Peperomia eglei" OR "Peperomia fluviatilis" OR "Peperomia huberi" OR "Peperomia lindmaniana" OR "Peperomia pellucidoides" OR "Peperomia pseudoserratifraxis" OR "Peperomia ripicola" OR "Peperomia simulans" OR "Peperomia spruceana" OR "Peperomia sulcata" OR "Peperomia sumidoriana" OR "Peperomia tenuilimba" OR "Piper acutilimbium" OR "Piper addisonii" OR "Piper albopilosum" OR "Piper aramanum" OR "Piper auriculifolium" OR "Piper belterraense" OR "Piper brachypetiolum" OR "Piper brevesanum" OR "Piper burenii" OR "Piper capitarianum" OR "Piper carniconnectivum" OR "Piper cavalcantei" OR "Piper chumboense" OR "Piper coariense" OR "Piper cowanii" OR "Piper crassipedunculum" OR "Piper curtistilum" OR "Piper emmerichianum" OR "Piper erectipilum" OR "Piper faroense" OR "Piper froesii" OR "Piper fulgidum" OR "Piper glabrescens" OR "Piper glandulosissimum" OR "Piper goeldii" OR "Piper humillimum" OR "Piper limosum" OR "Piper macapaense" OR "Piper manausense" OR "Piper mastersianum" OR "Piper millegranum" OR "Piper moense" OR "Piper montealegreanum" OR "Piper monteverdeanum" OR "Piper moreiranum" OR "Piper negroense" OR "Piper nigribaccum" OR "Piper ovatilimbium" OR "Piper paranum" OR "Piper pilgeri" OR "Piper plurinervosum" OR "Piper renitens" OR "Piper retropilosum" OR "Piper scandenticaule" OR "Piper snethlagei" OR "Piper solutidrupum" OR "Piper striatipetiolum" OR "Piper subfalcatum" OR "Piper subglabrifolium" OR "Piper sublepidotum" OR "Piper sublignosum" OR "Piper substilosum" OR "Piper suffrutescens" OR "Piper taperinhanum" OR "Piper uaupesense" OR "Piper uleanum" OR "Piper utinganum" OR "Aristida amazonensis" OR "Arthrostylidium fimbrinodum" OR "Arthrostylidium grandifolium" OR "Chusquea diversiglumis" OR "Eremocaulon</p>

Grupo	String para busca
	<p>amazonicum" OR "Guadua leonardoana" OR "Merostachys cachimboensis" OR "Merostachys fimbriata" OR "Merostachys rondoniensis" OR "Olyra davidseana" OR "Olyra tamanquareana" OR "Pariana concinna" OR "Pariana distans" OR "Pariana ecuadorensis" OR "Pariana imberbis" OR "Pariana intermedia" OR "Pariana ligulata" OR "Pariana modesta" OR "Pariana nervata" OR "Pariana ovalifolia" OR "Pariana sociata" OR "Raddiella lunata" OR "Raddiella minima" OR "Sporobolus multiramis" OR "Sporobolus temomairemensis" OR "Andropogon indetonsus" OR "Axonopus amapaensis" OR "Axonopus carajasensis" OR "Ichnanthus leptophyllus" OR "Ichnanthus tarumanensis" OR "Paspalum campinarum" OR "Paspalum cangarum" OR "Paspalum carajasense" OR "Paspalum lacustre" OR "Paspalum parvulum" OR "Paspalum seminudum" OR "Apinagia aripecuruensis" OR "Apinagia batrachifolia" OR "Apinagia tenuifolia" OR "Castelnavia fluitans" OR "Castelnavia monandra" OR "Lophogyne aripuanensis" OR "Lophogyne goeldiana" OR "Lophogyne paraensis" OR "Mourera monadelphica" OR "Oserya biceps" OR "Rhyncholacis minor" OR "Rhyncholacis unguifera" OR "Rhyncholacis varians" OR "Caamembeca amazonensis" OR "Diclidanthera octandra" OR "Moutabea angustifolia" OR "Moutabea chodatiana" OR "Moutabea victoriana" OR "Securidaca calophylla" OR "Securidaca marajoara" OR "Securidaca retusa" OR "Coccoloba gigantifolia" OR "Triplaris physocalyx" OR "Clavija elliptica" OR "Ctenardisia speciosa" OR "Cybianthus hoehnei" OR "Myrsine luae" OR "Stylogyne racemiflora" OR "Roupala nonscripta" OR "Roupala psilocarpa" OR "Lacunaria grandifolia" OR "Lacunaria oppositifolia" OR "Lacunaria sampaioi" OR "Quiina amazonica" OR "Quiina berryi" OR "Quiina cidiana" OR "Quiina maguirei" OR "Quiina negrensis" OR "Quiina piresii" OR "Touroulia amazonica" OR "Cephalostemon gracilis" OR "Monotrema bracteatum" OR "Rapatea undulata" OR "Saxofridericia brasiliensis" OR "Stegolepis piresii" OR "Rhabdodendron macrophyllum" OR "Gouania pyrifolia" OR "Sterigmata petalum plumbeum" OR "Alibertia curviflora" OR "Alibertia duckei" OR "Alibertia obidensis" OR "Alibertia sorbilis" OR "Alseis labatioides" OR "Amaioua contracta" OR "Amaioua monteiroi" OR "Appunia aurantiaca" OR "Appunia odontocalyx" OR "Appunia triphylla" OR "Bathysa bathysoides" OR "Borreria amapaensis" OR "Borreria carajasensis" OR "Borreria delicatula" OR "Borreria egleri" OR "Borreria elaiosulcata" OR "Borreria heteranthera" OR "Borreria paraensis" OR "Borreria petraea" OR "Borreria semiamplexicaule" OR "Calycophyllum tefense" OR "Carajasia cangae" OR "Carapichea fimbriiflora" OR "Carapichea panurensis" OR "Chimarrhis barbata" OR "Chimarrhis duckeana" OR "Coussarea duckei" OR "Coussarea insignis" OR "Coussarea japurana" OR "Coussarea krukovii" OR "Coussarea locuples" OR "Coussarea micrococca" OR "Coussarea panurensis" OR "Dendrosipanea prancei" OR "Duroia hirsutissima" OR "Duroia palustris" OR "Duroia prancei" OR "Duroia velutina" OR "Faramea occidentalis" OR "Faramea platyneura" OR "Faramea stenomeris" OR "Ferdinandusa cordata" OR "Ferdinandusa duckei" OR "Ferdinandusa lanceolata" OR "Ferdinandusa leucantha" OR "Ferdinandusa nitida" OR "Ferdinandusa papioriensis" OR "Ferdinandusa scandens" OR "Geophila tenuis" OR "Gleasonia cururuensis" OR "Gleasonia macrocalyx" OR "Gleasonia prancei" OR "Gleasonia uaupensis" OR "Guettarda duckei" OR "Guettarda hoffmannseggii" OR "Isertia longifolia" OR "Isertia psammophila" OR "Ixora amapaensis" OR "Ixora cowanii" OR "Ixora duckei" OR "Ixora faroensis" OR "Ixora martinsii" OR "Ixora pubescens" OR "Ixora rufa" OR "Ixora sandwithiana" OR "Kerianthera preclara" OR "Kutchubaea duckei" OR "Ladenbergia paraensis" OR "Malanea duckei" OR "Malanea egleri" OR "Malanea subtruncata" OR "Mitracarpus carajasensis" OR "Mitracarpus froesii" OR "Neobertiera montedouradensis" OR "Pagamea aracaensis" OR "Pagamea pilosa" OR "Pagamea puberula" OR "Pagamea spruceana" OR "Palicourea adderleyi" OR "Palicourea anisoloba" OR "Palicourea barraensis" OR "Palicourea canescens" OR "Palicourea decipiens" OR "Palicourea formosissima" OR "Palicourea kuhlmannii" OR "Palicourea megacephala" OR "Palicourea piresii" OR "Palicourea quadribractea" OR "Palicourea quinata" OR "Perama carajensis" OR "Perama galioides" OR "Platycarpum acreanum" OR "Platycarpum duckei" OR "Platycarpum egleri" OR "Platycarpum froesii" OR "Psychotria alemquerensis" OR "Psychotria anisocephala" OR "Psychotria arirambana" OR "Psychotria calocardia" OR "Psychotria duckei" OR "Psychotria gabrielis" OR "Psychotria hirtinervia" OR "Psychotria krukovii" OR "Psychotria kuhlmannii" OR "Psychotria maranhana" OR "Psychotria melanotricha" OR "Psychotria muelleriana" OR "Psychotria multiplex" OR "Psychotria neurothrix" OR "Psychotria phaneroplexa" OR "Psychotria rigescens" OR "Psychotria semifissa" OR "Psychotria trichotoma" OR "Ptylocarpus campinorum" OR "Ptylocarpus cururuensis" OR "Ptylocarpus ptylocarpoides" OR "Remijia asperula" OR "Remijia densiflora" OR "Remijia duckei" OR "Remijia firmula" OR "Remijia glomerata" OR "Remijia grazielae" OR "Remijia hirsuta" OR "Remijia hubbardiorum" OR "Remijia kuhlmannii" OR "Remijia physophora" OR "Remijia tenuiflora" OR "Retiniphyllum cataractae" OR "Retiniphyllum glabrum" OR "Retiniphyllum longiflorum" OR "Retiniphyllum parvifolium" OR "Ronabea isanae" OR "Rudgea cordata" OR "Rudgea lorentensis" OR "Rudgea palmarum" OR "Rustia thibaudioides" OR "Sabicea burchellii" OR "Sabicea mollissima" OR "Sabicea trailii" OR "Schradera brasiliensis" OR "Simira paraensis" OR "Sipanea glabrata" OR "Sipanea prancei" OR "Sipanea saxicola" OR</p>

Grupo	String para busca
	<p>"Sipanea wilson-brownei" OR "Sipaneopsis duckei" OR "Sphinctanthus acutilobus" OR "Angostura quinquefolia" OR "Euxylophora paraensis" OR "Galipea congestiflora" OR "Galipea grandifolia" OR "Hortia longifolia" OR "Hortia neblinensis" OR "Hortia nudipetala" OR "Hortia superba" OR "Pilocarpus alatus" OR "Pilocarpus carajaensis" OR "Pilocarpus trifoliolatus" OR "Raputia amazonica" OR "Raputia praetermissa" OR "Raputia ulei" OR "Rauia prancei" OR "Ravenia polygalaealix" OR "Ravenia pseudalterna" OR "Raveniopsis aracaensis" OR "Raveniopsis campincola" OR "Raveniopsis necopinata" OR "Spiranthera speciosa" OR "Ticorea froesii" OR "Zanthoxylum djalma-batistae" OR "Casearia manausensis" OR "Casearia murceana" OR "Casearia neblinae" OR "Ryania canescens" OR "Ryania riedeliana" OR "Antidaphne amazonensis" OR "Phoradendron bicarinatum" OR "Phoradendron oliveirae" OR "Cupania latifolia" OR "Matayba longipes" OR "Matayba robusta" OR "Paullinia alsmithii" OR "Paullinia anodonta" OR "Paullinia bilobulata" OR "Paullinia bracteosa" OR "Paullinia cidii" OR "Paullinia connaracea" OR "Paullinia firma" OR "Paullinia interrupta" OR "Paullinia largifolia" OR "Paullinia martinellii" OR "Paullinia prevostiana" OR "Paullinia reticulata" OR "Paullinia scaberula" OR "Paullinia setosa" OR "Porocystis acuminata" OR "Talisia douradensis" OR "Talisia ghilleana" OR "Talisia granulosa" OR "Talisia marleneana" OR "Talisia parviflora" OR "Thinouia trifoliata" OR "Toulicia bullata" OR "Toulicia petiolulata" OR "Chrysophyllum acreanum" OR "Chrysophyllum durifructum" OR "Chrysophyllum superbum" OR "Chrysophyllum wilsonii" OR "Ecclinusa campinae" OR "Ecclinusa lancifolia" OR "Manilkara excelsa" OR "Manilkara paraensis" OR "Micropholis caudata" OR "Micropholis grandiflora" OR "Micropholis resinifera" OR "Micropholis retusa" OR "Micropholis submarginalis" OR "Micropholis williamii" OR "Pouteria amapaensis" OR "Pouteria brevensis" OR "Pouteria cicatricata" OR "Pouteria decussata" OR "Pouteria ericoides" OR "Pouteria erythrochrysa" OR "Pouteria exstaminodia" OR "Pouteria flavilata" OR "Pouteria freitasii" OR "Pouteria fulva" OR "Pouteria kossmanniae" OR "Pouteria latianthera" OR "Pouteria lucens" OR "Pouteria maxima" OR "Pouteria minima" OR "Pouteria obscura" OR "Pouteria oppositifolia" OR "Pouteria pachyphylla" OR "Pouteria pallens" OR "Pouteria pariry" OR "Pouteria pentamera" OR "Pouteria petiolata" OR "Pouteria resinosa" OR "Pouteria stipulifera" OR "Pouteria stylifera" OR "Pouteria tarumanensis" OR "Pouteria tenuisepala" OR "Pradosia decipiens" OR "Pradosia granulosa" OR "Pradosia grisebachii" OR "Pradosia lahoziana" OR "Pradosia subverticillata" OR "Sarcaulus inflexus" OR "Sarcaulus vestitus" OR "Heliamphora ceracea" OR "Schlegelia aurea" OR "Homalolepis cavalcantei" OR "Siparuna petasiformis" OR "Smilax krukovii" OR "Duckeodendron cestroides" OR "Juanulloa parviflora" OR "Lycianthes amazonica" OR "Styrax prancei" OR "Thismia calcarata" OR "Thismia petasiformis" OR "Thismia pseudomelanomitra" OR "Thismia ribeiroi" OR "Thismia variabilis" OR "Daphnopsis filipedunculata" OR "Funifera ericiflora" OR "Lophostoma calophylloides" OR "Lophostoma dinizii" OR "Harperocallis paniculata" OR "Trigonia kerrii" OR "Sciaphila oligantha" OR "Turnera amapaensis" OR "Turnera amazonica" OR "Turnera discors" OR "Turnera kuhlmanniana" OR "Turnera laciniata" OR "Turnera tapajoensis" OR "Turnera urbanii" OR "Cecropia purpurascens" OR "Cecropia ulei" OR "Coussapoa arachnoidea" OR "Coussapoa cupularis" OR "Coussapoa scabra" OR "Coussapoa sprucei" OR "Pourouma cuspidata" OR "Pourouma formicarum" OR "Pourouma ovata" OR "Pourouma persecta" OR "Petrea brevicalyx" OR "Petrea campinae" OR "Petrea insignis" OR "Rinorea bicornuta" OR "Rinorea longistipulata" OR "Rinorea villosiflora" OR "Cissus acreensis" OR "Cissus amapaensis" OR "Erisma bracteosum" OR "Erisma fuscum" OR "Erisma gracile" OR "Erisma lanceolatum" OR "Erisma silvae" OR "Qualea brasiliana" OR "Qualea brevipedicellata" OR "Qualea cyanea" OR "Qualea decorticans" OR "Qualea labouriauana" OR "Qualea macropetala" OR "Qualea rupicola" OR "Qualea sprucei" OR "Qualea suprema" OR "Qualea themistoclesii" OR "Ruizterania belemnensis" OR "Ruizterania clavata" OR "Ruizterania trichanthera" OR "Ruizterania urceolata" OR "Ruizterania wittrockii" OR "Vochysia assua" OR "Vochysia calamana" OR "Vochysia eximia" OR "Vochysia expansa" OR "Vochysia floribunda" OR "Vochysia fontellae" OR "Vochysia hannekesaskiarum" OR "Vochysia ingens" OR "Vochysia inundata" OR "Vochysia ledouxii" OR "Vochysia mapuerae" OR "Vochysia mariziana" OR "Vochysia maxima" OR "Vochysia obidensis" OR "Vochysia pachyantha" OR "Vochysia parviflora" OR "Vochysia pinkusii" OR "Vochysia punctata" OR "Vochysia revoluta" OR "Vochysia rufescens" OR "Vochysia venulosa" OR "Douradoa consimilis" OR "Abolboda abbreviata" OR "Xyris brachyfolia" OR "Xyris brachysepala" OR "Xyris calderonii" OR "Xyris cylindrostachya" OR "Xyris downsiana" OR "Xyris eglei" OR "Xyris ferreirae" OR "Xyris mima" OR "Xyris pallidula" OR "Xyris pectinata" OR "Xyris piresiana" OR "Xyris tomentosa" OR "Renealmia matogrossensis"</p>
GRUPO 2 Briófitas > Hepáticas	<p>"Cephalantholejeunea temnanthoides" OR "Ceratolejeunea maranhensis" OR "Ceratolejeunea ocirii" OR "Ceratolejeunea semicornua" OR "Ceratolejeunea temnantha" OR "Cheilolejeunea amazonica" OR "Cheilolejeunea cuspidifera" OR "Cheilolejeunea papulosa" OR "Cheilolejeunea subcrenulata" OR</p>

Grupo	String para busca
	"Cololejeunea planiuscula" OR "Cololejeunea schusteri" OR "Diplasiolejeunea lanciloba" OR "Drepanolejeunea integribracteata" OR "Drepanolejeunea robinsonii" OR "Haplolejeunea amazonica" OR "Lejeunea combuensis" OR "Lepidolejeunea cordifissa" OR "Otolejeunea schnelli" OR "Prionolejeunea recurvula" OR "Pycnolejeunea gradsteinii" OR "Pycnolejeunea remotistipula" OR "Schusterolejeunea saxorum" OR "Verdoornianthus marsupiiifolius"
GRUPO 3 Briófitas > Musgos	"Taxithelium juruense" OR "Sphagnum amazonicum" OR "Sphagnum curicuriariense" OR "Sphagnum dimorphophyllum"
GRUPO 4 Fungos > Strictu sensus	"Astrocystis amazonensis" OR "Thamnomycetes fuciformis" OR "Clavulina connata" OR "Clavulina sprucei"
GRUPO 5 Samambaias e Licófitas	"Anemia dardanoi" OR "Blechnum areolatum" OR "Blechnum longipilosum" OR "Cyathea leucolepismata" OR "Cyathea tortuosa" OR "Hypolepis forzzae" OR "×Cyclobotrya telespirensis" OR "Mickelia bernoullii" OR "Dicranopteris seminuda" OR "Hymenophyllum prionema" OR "Isoetes amazonica" OR "Isoetes cangae" OR "Isoetes serracarajensis" OR "Danaea trifoliata" OR "Danaea ulei" OR "Metaxya lanosa" OR "Metaxya scalaris" OR "Cochlidium linearifolium" OR "Lellingeria hirsuta" OR "Moranopteris plicata" OR "Saccoloma membranaceum" OR "Salvinia sprucei" OR "Schizaea amazonica" OR "Selaginella deltoides" OR "Selaginella gynostachya" OR "Selaginella stomatoloma" OR "Selaginella surucucusensis" OR "Selaginella terezoana" OR "Selaginella ventricosa" OR "Selaginella zartmanii" OR "Goniopteris indusiata"

## APÊNDICE B – EXPRESSÕES DE BUSCAS SCOPUS PARA ODS

ODS (Meta nº)	Expressão de busca
1	TITLE-ABS-KEY ( ( {extreme poverty} OR {poverty alleviation} OR {poverty eradication} OR {poverty reduction} OR {international poverty line} OR ( {financial aid} AND {poverty} ) OR ( {financial aid} AND {poor} ) OR ( {financial aid} AND {north-south divide} ) OR ( {financial development} AND {poverty} ) OR {financial empowerment} OR {distributional effect} OR {distributional effects} OR {child labor} OR {child labour} OR {development aid} OR {social protection} OR {social protection system} OR ( {social protection} AND access ) OR microfinanc* OR micro-financ* OR {resilience of the poor} OR ( {safety net} AND {poor} OR {vulnerable} ) OR ( {economic resource} AND access ) OR ( {economic resources} AND access ) OR {food bank} OR {food banks} ) )
2	TITLE-ABS-KEY ( ( {land tenure rights} OR ( smallholder AND ( farm OR forestry OR pastoral OR agriculture OR fishery OR {food producer} OR {food producers} ) ) OR malnourish* OR malnutrition OR undernourish* OR {undernutrition} OR {agricultural production} OR {agricultural productivity} OR {agricultural practices} OR {agricultural management} OR {food production} OR {food productivity} OR {food security} OR {food insecurity} OR {land right} OR {land rights} OR {land reform} OR {land reforms} OR {resilient agricultural practices} OR ( agriculture AND potassium ) OR fertili?er OR {food nutrition improvement} OR {hidden hunger} OR {genetically modified food} OR ( gmo AND food ) OR {agroforestry practices} OR {agroforestry management} OR {agricultural innovation} OR ( {food security} AND {genetic diversity} ) OR ( {food market} AND ( restriction OR tariff OR access OR {north south divide} OR {development governance} ) ) OR {food governance} OR {food supply chain} OR {food value chain} OR {food commodity market} AND NOT {disease} ) )
3	TITLE-ABS-KEY ( ( ( human AND ( health* OR disease* OR illness* OR medicine* OR mortality ) ) OR {battered child syndrome} OR {cardiovascular disease} OR {cardiovascular diseases} OR {chagas} OR {child abuse} OR {child neglect} OR {child well-being index} OR {youth well-being index} OR {child wellbeing index} OR {youth wellbeing index} OR {water-borne disease} OR {water-borne diseases} OR {water borne disease} OR {water borne diseases} OR {tropical disease} OR {tropical diseases} OR {chronic respiratory disease} OR {chronic respiratory diseases} OR {infectious disease} OR {infectious diseases} OR {sexually-transmitted disease} OR {sexually transmitted disease} OR {sexually-transmitted diseases} OR {sexually transmitted diseases} OR {communicable disease} OR {communicable diseases} OR aids OR hiv OR {human immunodeficiency virus} OR tuberculosis OR malaria OR hepatitis OR polio* OR vaccin* OR cancer* OR diabet* OR {maternal mortality} OR {child mortality} OR {childbirth complications} OR {neonatal mortality} OR {neo-natal mortality} OR {premature mortality} OR {infant mortality} OR {quality adjusted life year} OR {maternal health} OR {preventable death} OR {preventable deaths} OR {tobacco control} OR {substance abuse} OR {drug abuse} OR {tobacco use} OR {alcohol use} OR {substance addiction} OR {drug addiction} OR {tobacco addiction} OR alcoholism OR suicid* OR {postnatal depression} OR {post-natal depression} OR {zika virus} OR dengue OR schistosomiasis OR {sleeping sickness} OR ebola OR {mental health} OR {mental disorder} OR {mental illness} OR {mental illnesses} OR {measles} OR {neglected disease} OR {neglected diseases} OR diarrhea OR diarrhoea OR cholera OR dysentery OR {typhoid fever} OR {traffic accident} OR {traffic accidents} OR {healthy lifestyle} OR {life expectancy} OR {life expectancies} OR {health policy} OR ( {health system} AND ( access OR accessible ) ) OR {health risk} OR {health risks} OR {inclusive health} OR obesity OR {social determinants of health} OR {psychological harm} OR {psychological wellbeing} OR {psychological well-being} OR {psychological well being} OR {public health} ) )
4	TITLE-ABS-KEY ( ( school OR education OR educational ) AND ( {school attendance} OR {school enrollment} OR {school enrolment} OR {inclusive education} OR {educational inequality} OR {education quality} OR {educational enrolment} OR {educational enrollment} OR {adult literacy} OR {numeracy rate}

ODS (Meta nº)	Expressão de busca
	<p>OR {educational environment} OR {educational access} OR ( {development aid} AND {teacher training} ) OR {early childhood education} OR {basic education} OR {affordable education} OR {educational financial aid} OR {school safety} OR {safety in school} OR ( {learning opportunities} AND ( {gender disparities} OR empowerment ) ) OR ( {learning opportunity} AND ( {gender disparities} OR empowerment ) ) OR {youth empowerment} OR {women empowerment} OR {equal opportunities} OR {child labour} OR {child labor} OR {discriminatory} OR {educational inequality} OR {educational gap} OR ( {poverty trap} AND {schooling} ) OR {special education needs} OR {inclusive education system} OR ( {schooling} AND ( {gender disparities} OR {ethnic disparities} OR {racial disparities} ) ) OR {education exclusion} OR {education dropouts} OR {global citizenship} OR {sustainable development education} OR {environmental education} OR {education policy} OR {educational policies} OR {international education} OR {education reform} OR ( {educational reform} AND {developing countries} ) OR {educational governance} OR ( {developing countries} AND {school effects} ) OR {education expenditure} OR {foreign aid} OR ( {teacher training} AND {developing countries} ) OR {teacher attrition} ) AND NOT {health literacy} )</p>
5	<p>TITLE-ABS-KEY ( ( {gender inequality} OR {gender equality} OR {employment equity} OR {gender wage gap} OR {female labor force participation} OR {female labour force participation} OR {women labor force participation} OR {women labour force participation} OR {womens' employment} OR {female employment} OR {women's unemployment} OR {female unemployment} OR ( access AND {family planning services} ) OR {forced marriage} OR {child marriage} OR {forced marriages} OR {child marriages} OR {occupational segregation} OR {women's empowerment} OR {girls' empowerment} OR {female empowerment} OR {female genital mutilation} OR {female genital cutting} OR {domestic violence} OR {women AND violence} OR {girl* AND violence} OR {sexual violence} OR ( {unpaid work} AND {gender inequality} ) OR ( {unpaid care work} AND {gender inequality} ) OR {women's political participation} OR {female political participation} OR {female managers} OR {women in leadership} OR {female leadership} OR {intra-household allocation} OR ( access AND {reproductive healthcare} ) OR {honour killing} OR {honor killing} OR {honour killings} OR {honor killings} OR {antiwomen} OR {anti-women} OR {feminism} OR {misogyny} OR {female infanticide} OR {female infanticides} OR {human trafficking} OR {forced prostitution} OR ( equality AND ( {sexual rights} OR {reproductive rights} OR {divorce rights} ) ) OR {women's rights} OR {gender injustice} OR {gender injustices} OR {gender discrimination} OR {gender disparities} OR {gender gap} OR {female exploitation} OR {household equity} OR {female political participation} OR {women's underrepresentation} OR {female entrepreneurship} OR {female ownership} OR {women's economic development} OR {women's power} OR {gender-responsive budgeting} OR {gender quota} OR ( {foreign aid} AND {women's empowerment} ) OR {gender segregation} OR {gender-based violence} OR {gender participation} OR {female politician} OR {female leader} OR {contraceptive behaviour} OR {women's autonomy} OR {agrarian feminism} OR {microfinance} OR {women's livelihood} OR {women's ownership} OR {female smallholder} OR {gender mainstreaming} ) )</p>
6	<p>TITLE-ABS-KEY ( ( ( ( {Safe} AND ( {water access} OR {drinking water} ) ) OR ( {clean} AND ( {drinking water} OR {water source} ) ) OR ( {water} AND ( {sanitation and hygiene} OR {sanitation &amp; hygiene} OR {quality} OR {resource} ) ) AND ( {water availability} OR {water-use efficiency} OR {water supply} OR {water supplies} OR {clean water} OR {hygienic toilet} OR {hygienic toilets} OR {antifouling membrane} OR {antifouling membranes} OR {anti-fouling membrane} OR {anti-fouling membranes} OR {water management} OR {aquatic toxicology} OR {water toxicology} OR {aquatic ecotoxicology} OR {water ecotoxicology} ) ) OR ( ( {freshwater} OR {fresh water} ) AND ( {water quality} ) AND ( {pollutant} OR {pollution} OR {contamina*} ) ) OR ( {freshwater} AND ( {water security} OR {water shortage} OR ( {waste water} AND "treatment" ) OR ( {wastewater} AND "treatment" ) OR {water conservation} OR {water footprint} OR {water infrastructure} OR {water pollution} OR {water purification} OR {water use} OR {water uses} OR sanit* OR sewer* ) ) OR ( ( {water} AND ( {ecosystem} OR {eco-system} ) ) AND (</p>

ODS (Meta n°)	Expressão de busca
	<p>{protection of} OR {endocrine disruptor} OR {endocrine disruptors} ) ) AND NOT {marine} ) OR ( {water} AND {water management} AND ( {pollution remediation} OR {pollutant removal} )) OR ( ( {groundwater} OR {ground water} OR {ground-water} ) AND {freshwater} ) OR ( ( {water pollution} OR {water pollutant} ) AND ( {waste water} AND “treatment” ) OR ( {wastewater} AND “treatment” ) ) OR {freshwater availability} OR {fresh water availability} OR {water scarcity} OR {open defecation} OR {blue water} OR {green water} OR {grey water} OR {black water} ) ) AND NOT {global burden of disease study} ) )</p>
7	<p>TITLE-ABS-KEY ( ( {energy efficiency} OR {energy consumption} OR {energy transition} OR {clean energy technology} OR {energy equity} OR {energy justice} OR {energy poverty} OR {energy policy} OR renewable* OR {2000 Watt society} OR {smart micro-grid} OR {smart grid} OR {smart microgrid} OR {smart micro-grids} OR {smart grids} OR {smart microgrids} OR {smart meter} OR {smart meters} OR {affordable electricity} OR {electricity consumption} OR {reliable electricity} OR {clean fuel} OR {clean cooking fuel} OR {fuel poverty} OR energiewende OR {life-cycle assessment} OR {life cycle assessment} OR {life-cycle assessments} OR {life cycle assessments} OR ( {photochemistry} AND {renewable energy} ) ) OR photovoltaic OR {photocatalytic water splitting} OR {hydrogen production} OR {water splitting} OR {lithium-ion batteries} OR {lithium-ion battery} OR {heat network} OR {district heat} OR {district heating} OR {residential energy consumption} OR {domestic energy consumption} OR {energy security} OR {rural electrification} OR {energy ladder} OR {energy access} OR {energy conservation} OR {low-carbon society} OR {hybrid renewable energy system} OR {hybrid renewable energy systems} OR {fuel switching} OR ( {foreign development aid} AND {renewable energy} ) OR {energy governance} OR ( {official development assistance} AND {electricity} ) OR ( {energy development} AND {developing countries} ) ) AND NOT ( {wireless sensor network} OR {wireless sensor networks} ) )</p>
8	<p>TITLE-ABS-KEY ( ( {economic growth} OR {economic development policy} OR {employment policy} OR {inclusive economic growth} OR {sustainable growth} OR {economic development} OR {economic globalization} OR {economic globalisation} OR {economic productivity} OR {low-carbon economy} OR {inclusive growth} OR microfinanc* OR micro-financ* OR micro-credit* OR microcredit* OR {equal income} OR {equal wages} OR {decent job} OR {decent jobs} OR {quality job} OR {quality jobs} OR {job creation} OR {full employment} OR {employment protection} OR {informal employment} OR {precarious employment} OR {unemployment} OR {precarious job} OR {precarious jobs} OR microenterprise* OR micro-enterprise* OR {small enterprise} OR {medium enterprise} OR {small enterprises} OR {medium enterprises} OR {small entrepreneur} OR {starting entrepreneur} OR {medium entrepreneur} OR {small entrepreneurs} OR {medium entrepreneurs} OR {starting entrepreneurs} OR {social entrepreneurship} OR {safe working environment} OR {labor market institution} OR {labor market institutions} OR {labour market institution} OR {labour market institutions} OR {forced labour} OR {forced labor} OR {child labour} OR {child labor} OR {labour right} OR {labor right} OR {labour rights} OR {labor rights} OR {modern slavery} OR {human trafficking} OR {child soldier} OR {child soldiers} OR {global jobs} OR {living wage} OR {minimum wage} OR {circular economy} OR {inclusive economy} OR {rural economy} OR {Foreign Development Investment} OR {Aid for Trade} OR {trade unions} OR {trade union} OR {working poor} OR {Not in Education, Employment, or Training} OR {carbon offset} OR {carbon offsetting} OR {carbon offsets} OR {offset project} OR {offset projects} OR {economic diversification} OR {material footprint} OR {resource efficiency} OR ( {cradle to cradle} AND {economy} ) OR {economic decoupling} OR {labour market disparities} OR {sustainable tourism} OR {ecotourism} OR {community-based tourism} OR {tourism employment} OR {sustainable tourism policy} OR {financial access} OR {financial inclusion} OR {access to banking} ) AND NOT {health} )</p>
9	<p>TITLE-ABS-KEY ( ( {industrial growth} OR {industrial diversification} OR {infrastructural development} OR {infrastructural investment} OR {infrastructure investment} OR {public infrastructure} OR {resilient infrastructure} OR {transborder infrastructure} OR {public infrastructures} OR {resilient infrastructures} OR</p>

ODS (Meta n°)	Expressão de busca
	<p>{transborder infrastructures} OR ( {industrial emissions} AND mitigation ) OR {industrial waste management} OR {industrial waste treatment} OR {traffic congestion} OR microenterprise* OR micro-enterprise* OR {small enterprise} OR {medium enterprise} OR {small enterprises} OR {medium enterprises} OR {small entrepreneur} OR {medium entrepreneur} OR {small entrepreneurs} OR {medium entrepreneurs} OR {value chain management} OR ( {broadband access} AND {developing countries} ) OR {manufacturing innovation} OR {manufacturing investment} OR {sustainable transportation} OR {accessible transportation} OR {transportation services} OR {inclusive transportation} OR {R&amp;D investment} OR {green product} OR {green products} OR {sustainable manufacturing} OR ( {cradle to cradle} AND industry ) OR {closed loop supply chain} OR ( industrial AND innovation ) OR {process innovation} OR {product innovation} OR {inclusive innovation} ) )</p>
10	<p>TITLE-ABS-KEY ( ( equality AND ( economic OR financial OR socio-economic ) ) OR ( inequality AND ( economic OR financial OR socio-economic ) ) OR {economic reform policy} OR {economic reform policies} OR {political inclusion} OR {social protection policy} OR {social protection policies} OR ( immigration AND NOT ( chemistry OR disease OR biodiversity ) ) OR ( emigration AND NOT ( chemistry OR disease OR biodiversity ) ) OR {foreign direct investment} OR {development gap} OR {development gaps} OR {migrant remittance} OR {responsible migration} OR {migration policy} OR {migration policies} OR {north-south divide} OR ( developing AND ( {tariffs} OR {tariff} OR {zero-tariff} OR {duty-free access} ) ) OR {social exclusion} OR {economic marginalization} OR {income inequality} OR {discriminatory law*} OR {discriminatory policies} OR {discriminatory policy} OR {economic empowerment} OR {economic transformation} OR ( {global market} AND {empowerment} ) ) )</p>
11	<p>TITLE-ABS-KEY ( ( city OR cities OR {human settlement} OR {human settlements} OR urban OR metropoli* OR town* OR municipal* ) AND ( gentrification OR congestion OR transportation OR {public transport} OR housing OR slum* OR {sendai framework} OR {Disaster Risk Reduction} OR {DRR} OR {smart city} OR {smart cities} OR {resilient building} OR {resilient buildings} OR {sustainable building} OR {sustainable buildings} OR {building design} OR {buildings design} OR urbani?ation OR {zero energy building} OR {zero energy buildings} OR {zero-energy building} OR {zero-energy buildings} OR {basic service} OR {basic services} OR {governance} OR {citizen participation} OR {collaborative planning} OR {participatory planning} OR {inclusiveness} OR {cultural heritage} OR {natural heritage} OR {UNESCO} OR {disaster} OR {ecological footprint} OR {environmental footprint} OR {waste} OR {pollution} OR {pollutant*} OR {waste water} OR {recycling} OR {circular economy} OR {air quality} OR {green space} OR {green spaces} OR {nature inclusive} OR {nature inclusive building} OR {nature inclusive buildings} ) )</p>
12	<p>TITLE-ABS-KEY ( {environmental pollution} OR {hazardous waste} OR {hazardous chemical} OR {hazardous chemicals} OR {toxic chemical} OR {toxic chemicals} OR {chemical pollution} OR {ozone depletion} OR {pesticide pollution} OR {pesticide stress} OR {pesticide reduction} OR {life cycle assessment} OR {life cycle analysis} OR {life cycle analyses} OR {life-cycle analysis} OR {life-cycle analyses} OR {low carbon economy} OR {low-carbon economy} OR {environmental footprint} OR {material footprint} OR {harvest efficiency} OR {solid waste} OR {waste generation} OR {corporate social responsibility} OR {corporate sustainability} OR {consumer behavior} OR {consumer behaviors} OR {consumer behaviour} OR {consumer behaviours} OR {waste recycling} OR {resource recycling} OR {resource reuse} OR {biobased economy} OR {zero waste} OR {sustainability label} OR {sustainability labelling} OR {global resource extraction} OR {material flow accounting} OR {societal metabolism} OR {food spill} OR {resource spill} OR {resource efficiency} OR {sustainable food consumption} OR {green consumption} OR {sustainable supply chain} OR {circular economy} OR {cradle to cradle} OR {sustainable procurement} OR {sustainable tourism} OR {fossil-fuel subsidies} OR {fossil-fuel expenditure} OR {sustainability label} OR {sustainability labelling} OR ( consumption AND ( {resource use} OR spill ) ) OR</p>

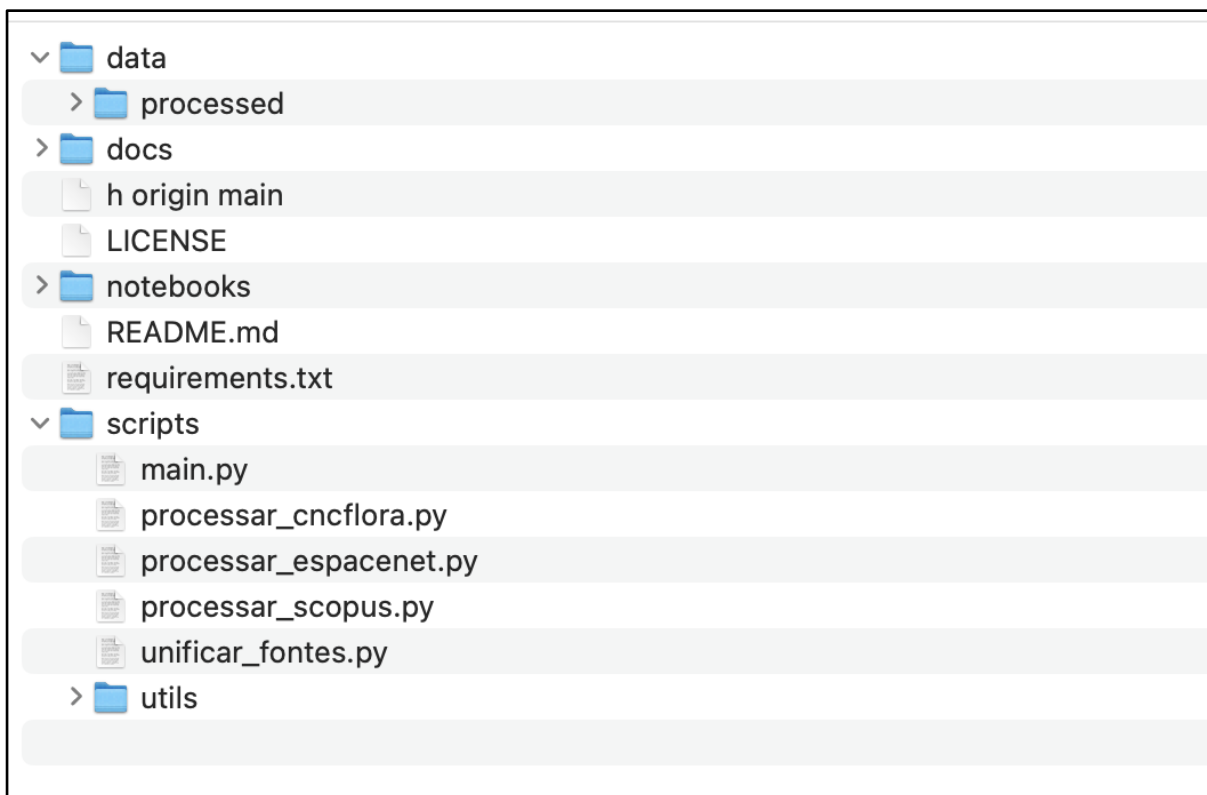
ODS (Meta nº)	Expressão de busca
	( production AND ( {resource use} OR spill ) ) AND NOT ( {wireless sensor network} OR {wireless sensor networks} OR {wireless network} OR {wireless networks} OR {wireless} OR {disease} OR {astrophysics} ) )
13	TITLE-ABS-KEY ( ( {climate action} OR {climate adaptation} OR {climate change} OR {climate capitalism} OR ipcc OR {climate effect} OR {climate equity} OR {climate feedback} OR {climate finance} OR {climate change financing} OR {climate forcing} OR {climate governance} OR {climate impact} OR {climate investment} OR {climate justice} OR {climate mitigation} OR {climate model} OR {climate models} OR {climate modeling} OR {climate modelling} OR {climate policy} OR {climate policies} OR {climate risk} OR {climate risks} OR {climate services} OR {climate service} OR {climate prediction} OR {climate predictions} OR {climate signal} OR {climate signals} OR {climate tipping point} OR {climate variation} OR {climate variations} OR ecoclimatology OR eco-climatology OR {Green Climate Fund} OR {regional climate} OR {regional climates} OR {urban climate} OR {urban climates} OR ( climate AND ( {adaptive management} OR awareness OR bioeconomy OR carbon OR {decision-making} OR {disaster risk reduction} OR {environmental education} OR {sustainable development education} OR {energy conservation} OR emission* OR extreme OR {food chain} OR {food chains} OR framework OR hazard* OR island* OR {land use} OR megacit* OR consumption OR production OR {small island developing states} OR anthropocene OR atmospher* OR {clean development mechanism} OR {glacier retreat} OR warming OR greenhouse OR {ice-ocean interaction} OR {ice-ocean interactions} OR {nitrogen cycle} OR {nitrogen cycles} OR {ocean acidification} OR {radiative forcing} OR {sea ice} OR {sea level} OR {sea levels} OR {thermal expansion} OR unfccc OR ozone ) ) ) AND NOT ( {drug} OR {geomorphology} ) )
14	TITLE-ABS-KEY ( ( marine OR ocean OR oceans OR sea OR seas OR coast* OR mangrove ) AND ( {water cycle} OR {water cycles} OR {biogeochemical cycle} OR {biogeochemical cycles} OR {oceanic circulation model} OR {oceanic circulation models} OR {oceanic circulation modelling} OR {oceanic circulation modeling} OR {ice-ocean} OR eutrophicat* OR marine OR {coral bleach} OR {coral bleaching} OR {coastal management} OR {coastal habitat} OR {coastal habitats} OR {marine debris} OR {ocean acidification} OR ( acidification AND seawater ) OR {fishery} OR {fisheries} OR {overfishing} OR {sustainable yield} OR {marine protected area} OR {marine protected areas} OR {marine conservation} OR {ecotourism} OR {community based conservation} OR {community-based conservation} OR {marine land slide} OR {marine pollution} OR {nutrient runoff} OR {coastal ecotourism} OR {destructive fishing} OR {local fisheries} OR {artisanal fishers} OR {fisheries rights} OR {species richness} OR {traditional ecological knowledge} OR {small Island development states} OR {marine quota} OR {marine economy} OR {marine policy} ) ) AND NOT ( {paleoclimate} OR {paleoceanography} OR {radiocarbon} OR {genetics} OR {medicine} OR {drug} OR {engineering} OR {aerosol} ) )
15	TITLE-ABS-KEY ( ( terrestrial OR land OR inland OR freshwater ) AND ( biodivers* OR {species richness} OR bioeconom* OR bio-econom* OR {biological production} OR deforest* OR desertif* OR {earth system} OR {ecological resilience} OR ecosystem* OR eco-system* OR {trophic cascade} OR {trophic level} OR {trophic web} OR {threatened species} OR {endangered species} OR {extinction risk} OR {extinction risks} OR poach* OR {wildlife product} OR {wildlife products} OR {wildlife traffic} OR {wildlife market} OR {wildlife markets} OR {wildlife trafficking} OR {invasive species} OR {alien species} OR {land uses} OR {land use} OR {land uses} OR {land degradation} OR {soil degradation} OR {LULUCF} OR *forest* OR {land conservation} OR wetland* OR mountain* OR dryland* OR {mountainous cover} OR {protected area} OR {protected areas} OR {REDD} OR {forest management} OR {silviculture} OR {timber harvest} OR {illegal logging} OR {slash-and-burn} OR {fire-fallow cultivation} OR {tree cover} OR {soil restoration} OR {land restoration} OR {drought} OR {sustainable land management} OR {mountain vegetation} OR {habitat restoration} OR {Red List species} OR {Red List Index} OR {extinction wave} OR {habitat fragmentation} OR {habitat loss} OR {Nagoya Protocol on Access to Genetic Resources} OR {genetic

ODS (Meta nº)	Expressão de busca
	resources} OR {biological invasion} OR {biodiversity-inclusive} OR {forest stewardship council} OR {rainforest alliance} OR {forest certification} OR {forest auditing} OR {ecotourism} OR {community-based conservation} OR {community based conservation} OR {human-wildlife conflict} ) )
16	TITLE-ABS-KEY ( ( {actual innocence} OR {false confession} OR {armed conflict} OR {armed conflicts} OR {civil conflict} OR {civil conflicts} OR ( war AND ( conflict OR warfare OR democracy OR {Geneva Convention} OR treaty OR peace ) ) OR {peacekeeping} OR ( corruption AND ( {institution} OR {public official} OR {government} OR {bribery} OR {conflict} ) ) OR crime OR crimes OR criminal OR {democratic deficit} OR ( democrati?ation AND ( institutional OR conflict OR decision-making OR society OR politics OR {financial aid} ) ) OR {ethnic conflict} OR {ethnic conflicts} OR exoneration OR genocid* OR homicid* OR murder* OR {human trafficking} OR {criminal justice system} OR {justice system} OR {arbitrary justice} OR refugee* OR terroris* OR violence OR torture OR {effective rule of law} OR {arms flow} OR {transparent institution} OR {transparent institutions} OR {good governance} OR {legal identity for all} OR {freedom of information} OR {human rights institution} OR {human rights activists} OR {fundamental freedom} OR {fundamental freedoms} OR {violent conflict} OR {violent conflicts} OR {peaceful society} OR {effective institution} OR {effective institutions} OR {accountable institution} OR {accountable institutions} OR {inclusive institution} OR {inclusive institutions} OR {child abuse} OR {arbitrary detention} OR {unsentenced detention} OR {judicial system} OR {criminal tribunal} OR {inclusive society} OR {inclusive societies} OR {responsive institution} OR {responsive institutions} OR {fair society} OR {fair societies} OR {legal remedy} OR {legal remedies} OR {independence of judiciary} OR {independent judiciary} OR {separation of powers} OR extremism OR {war crime} OR {peaceful society} OR {organized crime} OR {illicit transfer} OR {illicit money} OR {arms trafficking} OR {cybercrime} OR {insurgence} OR {democratic institution} OR {political instability} OR ( {political decision-making} AND ( responsive OR inclusive OR participatory OR representative ) ) OR {Aarhus Convention} OR {press freedom} OR {freedom of speech} ) AND NOT ( {disease} OR {genetics} ) )
17	Não incluso / não elaborado pela base até o momento da pesquisa

**APÊNDICE C – LISTA DE ARQUIVOS RAÍZ UTILIZADOS COMO INPUT PARA O  
PIPELINE DE ETL**

<b>Fonte (coletado de)</b>	<b>Nome do arquivo e formato</b>	<b>Localização no código (pasta em que está)</b>
<i>Site CNCFlora</i>	<ul style="list-style-type: none"> <li>● lista_vermelha_cncflora.csv</li> <li>● termos_plantas.txt</li> </ul>	data > raw > cncflora
<i>Base Espacenet</i>	<ul style="list-style-type: none"> <li>● espacenet_angiosperma.csv</li> <li>● espacenet_samambaias_e_licofitas.csv</li> </ul>	data > raw > espacenet_input
<i>Site IPC (WIPO)</i>	<ul style="list-style-type: none"> <li>● consolidated_ipc_level_1.csv</li> <li>● consolidated_ipc_level_2.csv</li> <li>● consolidated_ipc_level_3.csv</li> <li>● consolidated_ipc_level_4.csv</li> <li>● consolidated_ipc_level_5.csv</li> </ul> <hr/> <ul style="list-style-type: none"> <li>● processed_ipc_gi_level_9.csv</li> </ul> <hr/> <ul style="list-style-type: none"> <li>● espacenet_resumo_plantas.csv</li> </ul>	data > raw > ipc > ipc_full  <hr/> data > raw > ipc > ipc_gi  <hr/> data > raw
<i>Base Scopus</i>	<ul style="list-style-type: none"> <li>● scopus_angiosperma.csv</li> <li>● scopus_briófitas_hepáticas.csv</li> <li>● scopus_briófitas_musgos.csv</li> <li>● scopus_fungos_strictusensus.csv</li> <li>● scopus_samambaias_e_licófitas.csv</li> </ul> <hr/> <ul style="list-style-type: none"> <li>● scopus_ods_v1.csv</li> <li>● dicionário_afiliacoes.csv</li> </ul>	data > raw > scopus_input          <hr/> data > raw

**APÊNDICE D – ESTRUTURA GERAL DO PIPELINE<sup>21</sup> CRIADO E DISPONIBILIZADO PARA ETL DOS DADOS, VERSÃO 1.2**



<sup>21</sup> Cristina-Silva, Luana. *Pipeline ETL para alimentação do painel de indicadores bibliométricos sobre plantas endêmicas da Amazônia*: subproduto da dissertação de mestrado em Ciência da Informação. Zenodo, v\_1.2, 2025. Disponível em: <https://doi.org/10.5281/zenodo.16851246>. Acesso em: 9 fev. 2026.



# APÊNDICE E – TELAS DO PROTÓTIPO DO PAINEL DE VISUALIZAÇÃO EM LOOKER STUDIO

## Sistema de Mapeamento e Monitoramento de Indicadores em Ciência, Tecnologia e Inovação de Plantas Endêmicas da Amazônia Legal SisMMI-CTI | PEAL

### Índice do painel

1. Home Page
  2. Dimensão científica
    - [2.1 Produção científica | artigos científicos](#)
    - [2.2 Artigos científicos com viés de sustentabilidade](#)
  3. Dimensão tecnológica e inovativa
    - [3.1 Produção tecnológica-inovativa | patentes](#)
    - [3.2 Patentes com viés de sustentabilidade](#)
  4. Informações gerais
    - [4.1 Metodologia e Referências](#)
    - [4.2 Contatos](#)
- [Relatórios Estáticos](#)

O presente painel trata-se de parte dos resultados da pesquisa de dissertação para mestrado acadêmico da discente Luana Cristina Silva, no Programa de Pós-Graduação em Ciência da Informação (PPGCI) da Universidade Federal de São Carlos (UFSCar).

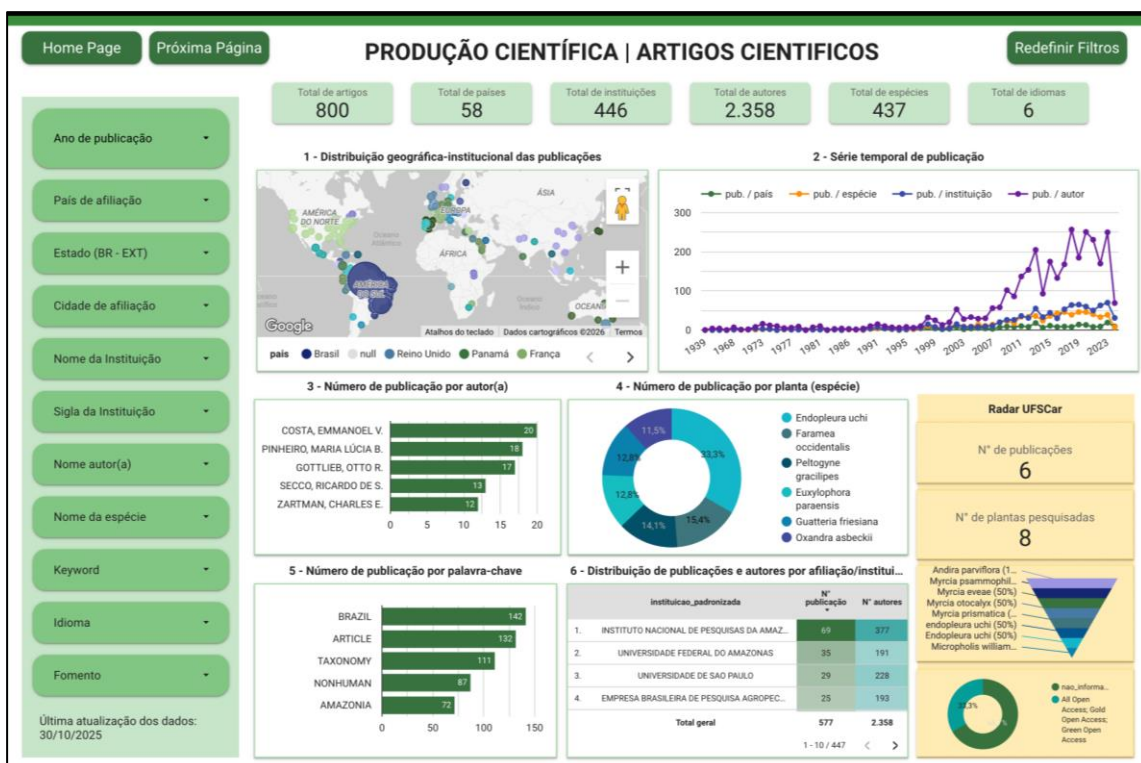
O seu objetivo principal é apresentar indicadores bibliométricos referentes à Ciência, Inovação e Tecnologia (CT&I) de forma dinâmica e intuitiva, contribuindo assim para o acesso à Informação Ambiental e consequentemente apoiando na democratização de seu acesso. Para isto, o objeto de estudo delimitado são as plantas endêmicas contidas no bioma da Amazônia Legal.

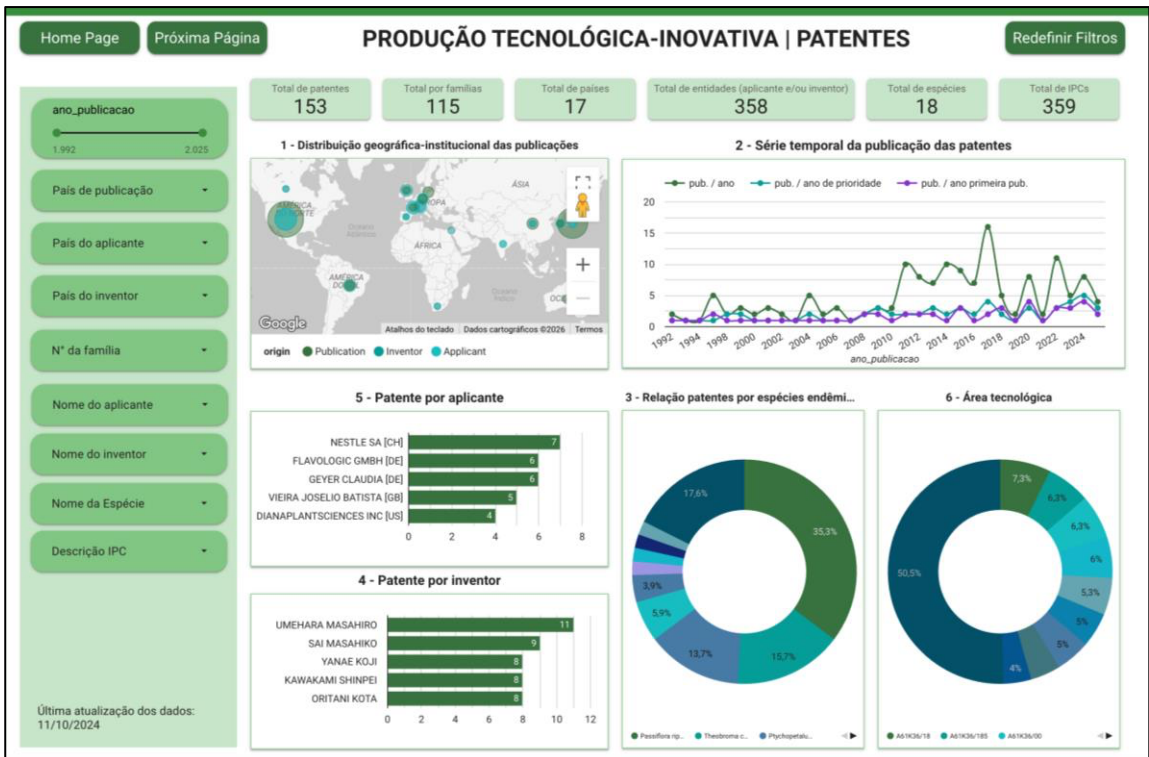
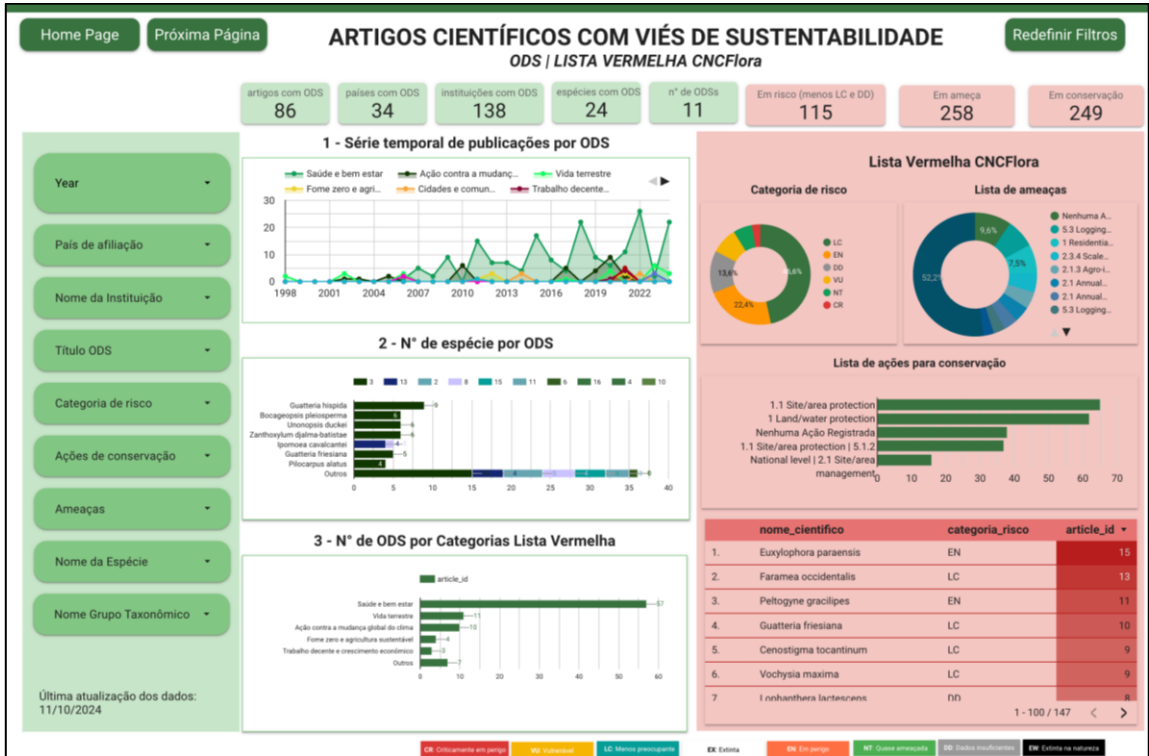
Além disso, os dados utilizados para gerar os presentes indicadores foram coletados, tratados e analisados com respaldo científico e validado por pares, dado aos cientistas/pesquisadores/docentes (x, y e z) que compuseram a banca examinadora de qualificação e defesa.

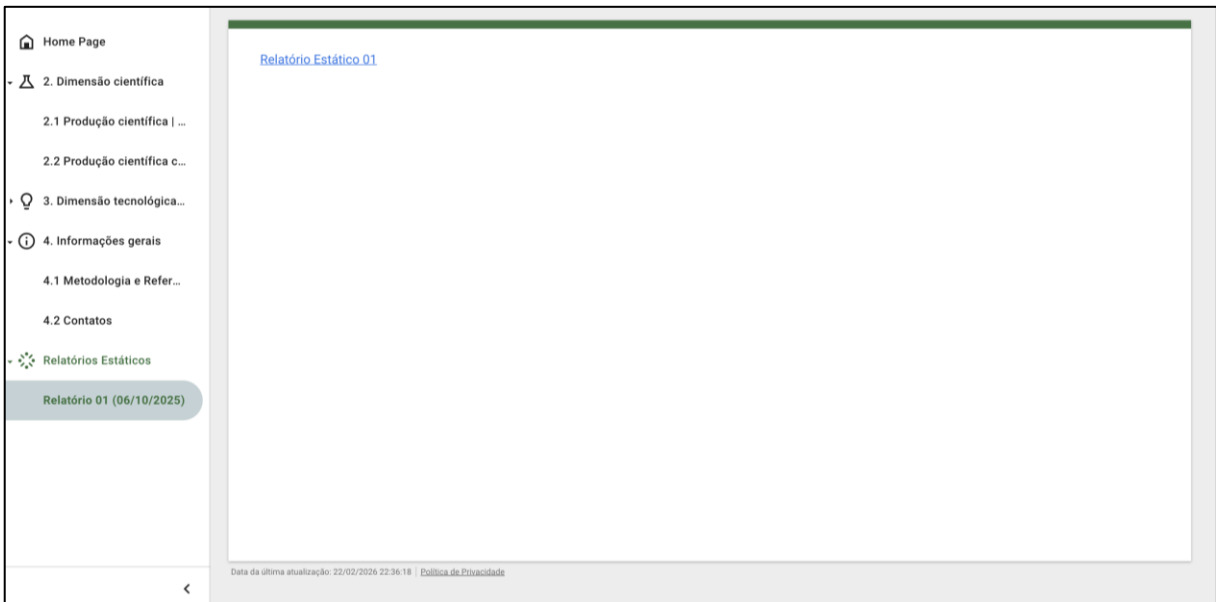
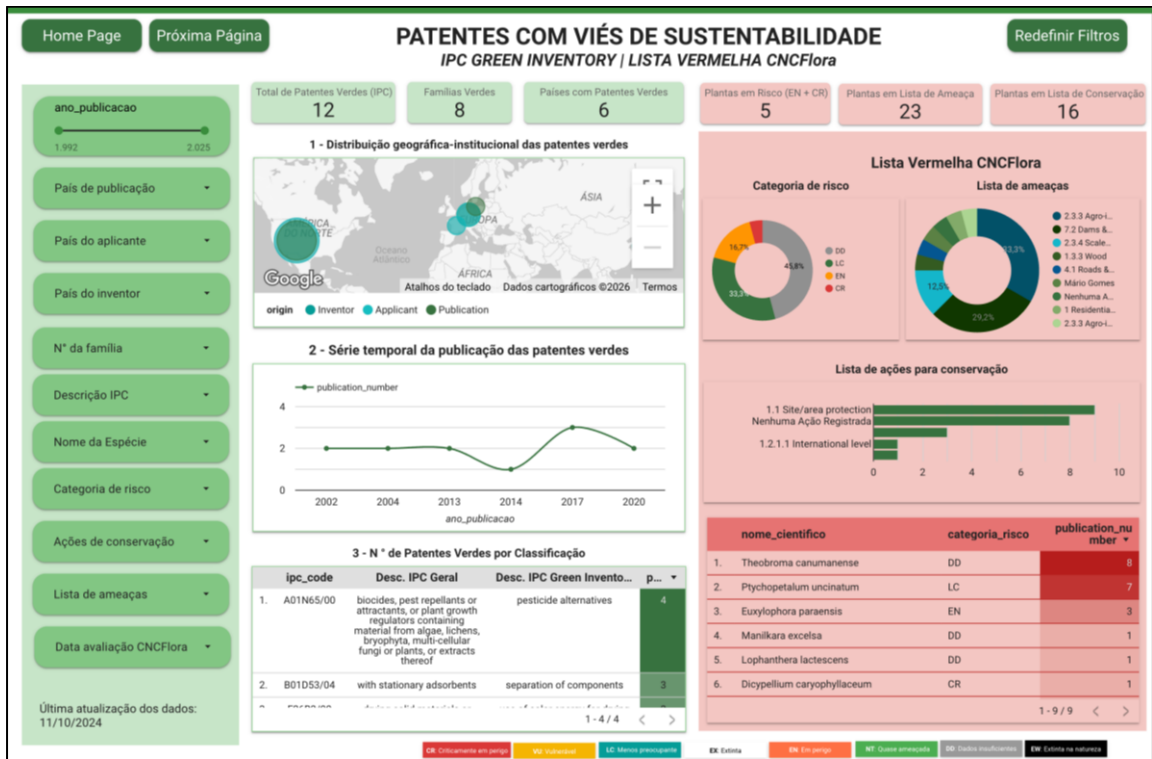
Cabe ressaltar que o presente sistema foi criado sob orientação do Professor Doutor Leandro Innocentini e co-orientação do Professor Doutor Rogério Ramalho, docentes do PPGCI-UFSCar.

Para consultar a dissertação na íntegra (referencial teórico, metodologia, resultados, análise, discussão etc.) visite o repositório institucional da UFSCar, ou clique aqui.

COMO CITAR (ABNT): CRISTINA-SILVA, Luana; FARIA, Leandro Innocentini; RAMALHO, Rogério. **Sistema de Mapeamento e Monitoramento de Indicadores em Ciência, Tecnologia e Inovação de Plantas Endêmicas da Amazônia Legal: SisMMI-CTI | PEAL**. — DOI: [10.5281/zenodo.17484946](https://doi.org/10.5281/zenodo.17484946). Disponível em: <https://lookerstudio.google.com/s/ldPec3CStDk>.







## APÊNDICE F – PLANO DE GESTÃO DE DADOS (PGD)

What data will you collect or create?

### **SOBRE OS DADOS**

#### **Tipo:**

Os dados coletados e gerados neste projeto são majoritariamente dados tabulares estruturados e textuais. Eles se enquadram nas seguintes categorias:

- **Dados Bibliométricos:** Metadados de publicações científicas, incluindo títulos, resumos, autores, afiliações, palavras-chave, contagem de citações e informações de periódicos.
- **Dados Tecnológicos:** Metadados de patentes, incluindo números de publicação, datas de prioridade e publicação, inventores, depositantes e classificações técnicas (IPC).
- **Dados de Biodiversidade:** Informações taxonômicas e de status de conservação de espécies da flora, como nome científico, família e categoria de ameaça.
- **Dados de Sustentabilidade:** Classificações categóricas que vinculam os documentos aos Objetivos de Desenvolvimento Sustentável (ODS) da ONU e à classificação de tecnologias verdes (IPC Green Inventory).
- **Dados Geográficos Derivados:** Nomes de países, estados e cidades extraídos e padronizados a partir dos dados de afiliação.

---

#### **Formato:**

- **Dados Brutos (Entrada):** Os dados foram coletados e armazenados no formato CSV (Comma-Separated Values). Alguns arquivos de origem utilizaram ponto e vírgula (;) como delimitador.
- **Dados Processados (Saída):** Todos os dados gerados pelo pipeline de processamento são salvos no formato CSV, utilizando vírgula (,) como delimitador, codificação UTF-8 e aspas duplas (") envolvendo todos os campos para garantir a integridade estrutural.

---

#### **Volume:**

O volume de dados é considerado de pequeno a médio, gerenciável em um ambiente de computação padrão. O volume estimado é:

- **Dados de Entrada:** Aproximadamente [Ex: 10 a 20] arquivos CSV, com um volume total inferior a [Ex: 50 MB].
- **Dados de Saída:** Geração de aproximadamente [Ex: 20 a 30] tabelas processadas (fato, dimensão e ponte), com um volume total estimado em [Ex: 100 a 200 MB]. O número total de registros processados é da ordem de milhares (Ex: ~800 artigos e ~150 registros de patentes expandidos).

-----

### **SOBRE OS FORMATOS E SOFTWARES**

#### **Formato:**

O formato CSV é um padrão aberto, universalmente aceito e não-proprietário. Ele pode ser lido por praticamente qualquer software de análise de dados, planilha eletrônica ou editor de texto (ex: Microsoft Excel, Google Sheets, R, Python, etc.), garantindo que os dados permaneçam acessíveis e utilizáveis por muitos anos, independentemente de mudanças tecnológicas. A codificação UTF-8 garante a compatibilidade com caracteres especiais de diversas línguas.

#### **Software:**

O pipeline de processamento foi desenvolvido integralmente com software de código aberto (open source), primariamente a linguagem de programação Python e suas bibliotecas (pandas). O código-fonte está versionado e publicamente disponível em um repositório Git (na plataforma GitHub). Isso garante que toda a metodologia de processamento seja transparente, auditável e possa ser replicada por qualquer pesquisador no futuro, sem a necessidade de licenças de software proprietário.

#### **Compartilhamento:**

O compartilhamento é garantido em múltiplos níveis:

- **Código-Fonte:** O repositório GitHub permite que toda a lógica de transformação dos dados seja compartilhada e reutilizada.
- **Dados Processados:** Os arquivos CSV finais podem ser facilmente compartilhados e depositados em repositórios de dados de pesquisa (como Zenodo, Figshare, ou o repositório institucional da universidade).
- **Visualização:** O uso do Google Sheets como camada intermediária e do Google Looker Studio para o painel permite o compartilhamento dos resultados finais através de um link público, facilitando o acesso aos indicadores por uma audiência ampla.

## ----- SOBRE A REUTILIZAÇÃO DOS DADOS

### Dados Reutilizáveis:

- **Bases de Dados Científicas:** O projeto reutiliza massivamente os dados provenientes das bases Scopus e Espacenet, que são fontes consolidadas e mantidas por instituições de renome.
- **Listas de Classificação e Autoridade:** O projeto reutiliza dados públicos e mantidos por organizações especializadas, incluindo:
  - **A Lista Vermelha do CNCFlora**, que representa o esforço de especialistas para avaliar o status de conservação da flora.
  - **O IPC Green Inventory da OMPI**, que é um padrão global para identificar patentes de tecnologias verdes.
  - **Dicionários de Códigos de Países.**
  - **Dicionários e Tesouros Manuais:** O projeto reutiliza dicionários de padronização de nomes de instituições (ex: "Lista de Autoridades UFSCar") e constrói seus próprios dicionários a partir da análise dos dados, que por sua vez se tornam um novo ativo de dados que pode ser reutilizado em projetos futuros.

How will the data be collected or created?

### Padrões ou metodologias

O projeto emprega uma metodologia de pipeline de dados de Extração, Transformação e Carga (ETL), implementado através de scripts em linguagem Python. A modelagem dos dados processados segue os princípios de um Modelo Estrela Híbrido, separando os dados em tabelas Fato (eventos mensuráveis, como publicações e patentes) e Dimensão (atributos descritivos, como autores, afiliações, espécies). Para o tratamento de dados de afiliação, foi adotada uma abordagem de curadoria semi-supervisionada, combinando a extração automática de entidades com a aplicação de um dicionário de autoridade (tesauro) mantido manualmente para garantir a padronização e o enriquecimento dos nomes de instituições.

### Estrutura e nome das pastas e arquivos

A organização dos dados e do código segue uma estrutura de projeto padronizada para garantir a clareza e a reprodutibilidade. A hierarquia de pastas principal é a seguinte:

- **/data/:** Contém todos os dados do projeto.
  - **/data/raw/:** Armazena os dados brutos e inalterados, exatamente como foram baixados das fontes originais (Scopus, Espacenet, CNCFlora, dicionários). Cada fonte possui sua própria subpasta (ex: /data/raw/scopus\_input/).
  - **/data/processed/:** Armazena os dados intermediários gerados pela Etapa 1 e 2 do pipeline (arquivos limpos, tabelas de dimensão e ponte). A estrutura de subpastas espelha a origem dos dados (ex: /data/processed/scopus/).
  - **/looker\_studio\_output/:** Armazena os arquivos CSV finais, otimizados e prontos para serem carregados na camada de visualização.

A nomenclatura dos arquivos é padronizada e descritiva, utilizando prefixos para indicar o tipo de tabela (dim\_, pon\_, fato\_) e sufixos que indicam o conteúdo (ex: dim\_autores\_scopus.csv, looker\_fato\_artigos.csv).

### Controle de versão

O controle de versão é gerenciado em dois níveis:

- **Código-Fonte:** O sistema de controle de versão Git é utilizado para gerenciar todas as alterações no código-fonte, nos scripts e na documentação. Todo o histórico de desenvolvimento, com mensagens de commit descritivas, é mantido na plataforma GitHub. Isso garante um registro completo e rastreável de toda a evolução do projeto.
- **Dados:** A versão dos dados não é controlada por nomes de arquivo (ex: v1, v2). Em vez disso, o princípio da reprodutibilidade é aplicado: os dados de saída são considerados efêmeros e podem ser completamente regenerados a qualquer momento executando o pipeline de scripts sobre os dados brutos. A versão dos dados está, portanto, diretamente ligada à versão do código no Git que os gerou. O conjunto de dados final, ao ser depositado em um repositório como o Zenodo, receberá um número de versão formal e um DOI (Digital Object Identifier), garantindo a citação e o acesso permanente a uma versão específica e estável.

### Processos de garantia de qualidade

A garantia de qualidade é um processo contínuo integrado ao pipeline:

- **Princípios FAIR:** Todo o fluxo de trabalho foi projetado para seguir os princípios FAIR. Os dados são Encontráveis (através do repositório e futuro DOI), Acessíveis (usando formatos abertos como CSV), Interoperáveis (com estrutura de dados bem definida) e Reutilizáveis (graças à documentação clara e licenças abertas).
- **Validação de Dados no Código:** Os scripts incluem verificações para garantir a consistência dos dados, como a conversão forçada de tipos de dados (.astype(str)) antes de operações de junção (merge) para prevenir erros de correspondência de chaves.
- **Limpeza e Padronização Automatizadas:** Funções específicas são utilizadas para normalizar dados textuais (remoção de acentos, padronização para caixa alta), garantindo a consistência na criação de chaves e na aplicação de dicionários.
- **Curadoria Semi-Supervisionada (Ciclo de Feedback):** Para dados complexos como afiliações, o processo de limpeza é interativo. Um notebook de prototipagem é utilizado para aplicar regras de padronização, e o resultado é validado contra um dicionário de autoridade. A saída do script gera listas de "não conformidades", que servem como feedback para o pesquisador aprimorar o dicionário de forma iterativa, melhorando a qualidade dos dados a cada ciclo.
- **Revisão por Pares do Código:** A disponibilização do código no *GitHub* permite uma potencial revisão por pares da metodologia implementada.

### Documentation and Metadata

What documentation and metadata will accompany the data?

1. **Informação necessária para que os dados sejam lidos e interpretados no futuro:** Para garantir a futura legibilidade e interpretação dos dados, uma combinação de informações contextuais, metodológicas e técnicas é essencial. A informação necessária inclui:
  - a. **Contexto do Projeto:** O objetivo da pesquisa, as questões que os dados visam responder (a relação entre CT&I e sustentabilidade na Amazônia), e o período de tempo coberto pela coleta.
  - b. **Proveniência dos Dados:** Origem de cada fonte de dados (ex: *Scopus*, *Espacenet*, *CNCFlora*), incluindo as datas de extração e os critérios de busca exatos (queries) utilizados.
  - c. **Metodologia de Processamento:** Uma descrição completa do *pipeline* de ETL, explicando as etapas de limpeza, transformação, enriquecimento e modelagem. Isso deve incluir a lógica de padronização de nomes, a separação de instituição/departamento, e como as classificações (ODS, *IPC Green*) foram aplicadas.
  - d. **Estrutura dos Dados e Dicionário de Dados:** Uma descrição clara do modelo de dados final (tabelas fato, dimensão e ponte). Para cada arquivo CSV gerado, um dicionário de dados é crucial, detalhando cada coluna (nome, descrição, tipo de dado [ex: texto, inteiro], e exemplos de valores).
  - e. **Dependências de Software:** A versão do *Python* utilizada e uma lista de todas as bibliotecas necessárias com suas respectivas versões, para garantir que o ambiente computacional possa ser replicado.
  - f. **Informações de Contato e Autoria:** Nome do(s) pesquisador(es) responsável(is), afiliação institucional (universidade, programa de pós-graduação) e informações de contato.

- g. **Licença de Uso:** Termos claros sob os quais os dados e o código podem ser acessados, reutilizados e distribuídos.

## 1. Captura e criação da documentação e metadados

A captura e criação de documentação e metadados serão um processo integrado ao desenvolvimento do projeto, utilizando uma abordagem multifacetada:

- **Documentação Incorporada ao Código (*Code-as-Documentation*):**
  - Comentários e Docstrings: Cada *scriptPython* (.py) e função principal é documentado internamente com "docstrings" que explicam seu propósito, entradas e saídas. Comentários são usados para esclarecer trechos de código complexos.
  - Nomenclatura Clara: As variáveis, funções e arquivos são nomeados de forma descritiva (ex: *processar\_scopus.py*, *dim\_afiliacoes\_FINAL.csv*) para que sua finalidade seja autoexplicativa.
- **Documentação Centralizada no Repositório:**
  - Arquivo README.md: O arquivo principal na raiz do repositório *Git* serve como o portal central de documentação. Ele contém a descrição do projeto, a arquitetura do pipeline (incluindo o fluxograma), a estrutura de pastas e, mais importante, as instruções detalhadas de como instalar as dependências e executar os *scripts*.
  - Dicionário de Dados (*dicionario\_de\_dados.md*): Um arquivo *Markdown* separado dentro do repositório será criado para documentar a estrutura de todas as tabelas finais geradas.
- **Metadados Gerados Automaticamente:**
  - requirements.txt: Este arquivo é gerado automaticamente pelo comando `pip freeze`, capturando a lista exata de todas as dependências de software e suas versões, garantindo a reprodutibilidade técnica.
  - Histórico do Git: O histórico de commits do *Git* serve como um registro cronológico detalhado de todas as alterações feitas no código e na documentação, incluindo quem fez a alteração, quando e por quê.

## 2. Padrões de metadados e justificativa

- Embora um esquema de metadados formal e complexo não seja necessário para a escala deste projeto, os princípios de padrões reconhecidos serão adotados para garantir a conformidade com as boas práticas de repositórios de dados.
  - **Padrão Utilizado:** A documentação seguirá uma estrutura inspirada no Dublin Core Metadata Initiative (DCMI), que é um padrão amplamente utilizado e flexível, focado na descrição de recursos digitais. Ele cobre os elementos essenciais para a descoberta e interpretação de dados.
  - **Por quê?** O *Dublin Core* é ideal porque é simples, interoperável e cobre os 15 elementos centrais da descrição de um recurso, que são diretamente aplicáveis ao nosso conjunto de dados. Ele é o padrão de fato para a maioria dos repositórios de dados acadêmicos (como *DSpace*, *Zenodo*, etc.).

---

1. Elementos do Dublin Core que serão preenchidos (no README.md e/ou ao depositar os dados):
  1. **Title:** Observatório de Indicadores Bibliométricos em CT&I e Sustentabilidade na Amazônia.
  2. **Creator:** Luana Cristina-Silva.
  3. **Subject:** Bibliometria, Indicadores de CT&I, Sustentabilidade, Amazônia, Patentes, ODS.
  4. **Description:** Resumo do projeto, descrevendo o pipeline de dados e os resultados gerados.
  5. **Publisher:** Universidade Federal de São Carlos / Programa de Pós-Graduação em Ciência da Informação.
  6. **Date:** 2025.
  7. **Type:** Dataset, Software.
  8. **Format:** CSV, Python.
  9. **Identifier:** <https://doi.org/10.5281/zenodo.16851246>
  10. **Source:** *Scopus*, *Espacenet*, CNCFlora, WIPO.
  11. **Language:** pt-BR (para os dados gerados), en-US (para alguns dados brutos).
  12. **Rights:** *Creative Commons Attribution* 4.0 (CC-BY 4.0), MIT License.

#### 4. Tipos de documentação que acompanharão os dados

Os dados serão acompanhados por três tipos principais de documentação para garantir que usuários secundários possam compreendê-los e reutilizá-los plenamente:

- **README.md:** Este é o documento principal e mais abrangente. Ele conterá:
  - **Informações Básicas:**
    - Título do Projeto: Observatório de Indicadores Bibliométricos em CT&I e Sustentabilidade na Amazônia.
    - Autor/Criador: [Seu Nome Completo], [Sua Afiliação].
    - Data de Criação: [Intervalo de datas do desenvolvimento, ex: 2024-2025].
    - Condições de Acesso: O código-fonte e a documentação serão disponibilizados publicamente no GitHub sob uma [Licença MIT], e os dados processados serão depositados em um repositório de acesso aberto sob uma licença [Creative Commons (CC-BY 4.0)].
  - **Informações Metodológicas:**
    - Descrição detalhada do pipeline de ETL, incluindo o fluxograma visual.
    - Descrição das fontes de dados, incluindo as queries de busca utilizadas.
    - Explicação da lógica de limpeza e padronização, incluindo a separação de instituição/departamento e a classificação de sustentabilidade.
  - **Informações Técnicas:**
    - Instruções claras sobre como configurar o ambiente (requirements.txt) e executar os scripts orquestradores (main\_etapa1\_base.py, etc.) para replicar todo o processo.

---

- **dicionario\_de\_dados.md:** Um arquivo específico que funcionará como um manual para os dados finais. Para cada arquivo CSV na pasta `looker_studio_output/`, ele detalhará:
  - Nome do Arquivo: Ex: `looker_artigo_autor.csv`.
  - Descrição: Ex: "Tabela de relação que conecta cada artigo (pelo `article_id`) ao nome completo de um de seus autores."
  - Detalhes das Colunas:
    - `article_id`: (Texto) Identificador único do artigo, correspondente ao `eid` da Scopus.
    - `authors_id`: (Texto) Identificador único do autor na Scopus.
    - `nome_completo`: (Texto) Nome completo do autor, padronizado.

---

- **Comentários no Código:** Os próprios *scriptsPython* servirão como uma forma de documentação técnica. As *docstrings* no início de cada arquivo e função explicarão o propósito de cada passo lógico do *pipeline*, permitindo que um usuário com conhecimento técnico entenda a implementação em detalhes.

#### Ethics and Legal Compliance

How will you manage any ethical issues?

Dada a natureza dos dados utilizados neste projeto, as questões éticas são mínimas, mas foram consideradas e gerenciadas da seguinte forma:

**Natureza dos Dados:** O projeto utiliza exclusivamente dados públicos ou publicamente disponíveis, derivados de publicações científicas e registros de patentes. Não há coleta de dados diretamente de participantes humanos, nem o manuseio de informações pessoais sensíveis (como dados médicos, financeiros ou privados). Os dados analisados (nomes de autores, afiliações institucionais) já são públicos por natureza, divulgados pelos próprios pesquisadores no ato da publicação.

**Consentimento para Preservação e Compartilhamento:** O consentimento é implícito e inerente ao ato de publicar uma pesquisa científica ou registrar uma patente. Autores e inventores publicam seus trabalhos com o

objetivo de disseminar o conhecimento, e as bases de dados como *Scopus* e *Espacenet* existem para agregar e facilitar o acesso a essa informação pública. Portanto, não é necessário obter um consentimento adicional para analisar esses metadados publicamente disponíveis.

**Proteção da Identidade dos Participantes:** Não há "participantes" no sentido tradicional de pesquisa com seres humanos. Os indivíduos mencionados nos dados são autores e inventores, figuras públicas no contexto de suas próprias publicações. Não há necessidade de anonimização, pois a atribuição correta do trabalho a seus criadores é um pilar ético fundamental da própria ciência. O objetivo do projeto não é analisar o comportamento privado dos pesquisadores, mas sim a produção científica e tecnológica agregada.

**Manuseio de Dados Sensíveis:** O projeto não lida com dados sensíveis. Todos os dados são públicos. O armazenamento e a transferência são gerenciados através de plataformas seguras (*GitHub Codespaces*, *GitHub*), mas não há requisitos especiais de segurança além das boas práticas padrão, pois não há risco de violação de privacidade de dados confidenciais.

Em resumo, o projeto adere aos princípios éticos ao utilizar apenas dados de domínio público para os fins para os quais foram publicados (análise da produção científica e tecnológica), sem infringir a privacidade de indivíduos.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

A gestão de direitos autorais e PI foi cuidadosamente planejada para respeitar as fontes originais e promover a reutilização aberta dos resultados gerados.

#### **Propriedade dos Dados:**

- **Dados de Terceiros:** Os dados brutos extraídos das bases *Scopus*, *Espacenet* e CNCFlora são de propriedade de suas respectivas entidades (*Elsevier*, EPO, Jardim Botânico do Rio de Janeiro, etc.). O projeto utiliza esses dados sob as políticas de "uso justo" (fair use) para fins de pesquisa acadêmica e não comercial. A análise se concentra nos metadados e fatos, que geralmente não são passíveis de direitos autorais, em vez de no conteúdo completo dos artigos ou patentes.
- **Dados Criados:** A propriedade intelectual dos dados derivados, do código-fonte e da estrutura do *pipeline* de ETL criados neste projeto pertence à autora, Luana Cristina-Silva, como parte de seu trabalho de dissertação de mestrado na Universidade Federal de São Carlos.

#### **Licenciamento para Reutilização:**

- Para maximizar o impacto e a aderência aos princípios da ciência aberta, os produtos gerados serão licenciados da seguinte forma:
  - **Código-Fonte:** Todo o código-fonte desenvolvido (os scripts *Python*) será disponibilizado no GitHub sob a Licença MIT. Esta é uma licença permissiva que permite a qualquer pessoa reutilizar, modificar e distribuir o código para qualquer fim (comercial ou não), exigindo apenas a manutenção do aviso de direitos autorais e da licença originais.
  - **Dados Processados e Derivados:** O conjunto de dados final e limpo, que será depositado em um repositório de dados de pesquisa, será licenciado sob a *Creative Commons Attribution 4.0 International* (CC-BY 4.0). Esta licença permite que outros compartilhem e adaptem os dados para qualquer fim, desde que deem o crédito apropriado à autora original, forneçam um link para a licença e indiquem se foram feitas alterações.

#### **Restrições sobre Dados de Terceiros:**

O projeto não redistribuirá os dados brutos extraídos da *Scopus* ou de outras bases comerciais. Apenas o código para processar os dados e os dados derivados e agregados (os resultados finais da análise) serão compartilhados. Isso respeita os termos de serviço das plataformas de origem, que geralmente permitem o uso para pesquisa, mas não a redistribuição em massa dos seus dados brutos. Os dados do CNCFlora e do IPC *Green Inventory*, por serem de fontes governamentais ou abertas, geralmente têm políticas de reutilização mais permissivas, que são compatíveis com os objetivos do projeto.

#### **Adiamento ou Restrição do Compartilhamento:**

Não haverá adiamento ou embargo nos dados. O objetivo do projeto é a disseminação pública do conhecimento. Não há intenção de buscar patentes ou de obter vantagem comercial a partir dos resultados, portanto, o código e os dados derivados serão compartilhados abertamente assim que a dissertação for defendida e publicada, ou conforme as políticas do programa de pós-graduação.

### Storage and Backup

How will the data be stored and backed up during the research?

A estratégia de armazenamento e backup foi projetada para garantir a integridade, segurança e disponibilidade dos dados e do código durante todo o ciclo de vida do projeto, utilizando uma abordagem de múltiplas camadas.

#### Armazenamento e Custo:

O volume de dados do projeto (estimado em menos de 500 MB) é relativamente pequeno. Não há necessidade de armazenamento adicional pago. Todo o armazenamento é coberto pelos serviços gratuitos e institucionais já disponíveis:

- **Armazenamento Primário (de Trabalho):** O ambiente de desenvolvimento em nuvem *GitHub Codespaces* serve como o local primário de armazenamento e processamento. Ele mantém uma cópia completa do código e dos dados em um contêiner seguro na nuvem da *Microsoft/GitHub*.
- **Armazenamento Secundário (Backup de Código):** O repositório *GitHub* ([github.com/silvaluah/sismm-cti-amazon](https://github.com/silvaluah/sismm-cti-amazon)) funciona como o principal sistema de backup para todo o código-fonte e documentação.
- **Armazenamento Terciário (Backup de Dados):** Uma cópia dos dados brutos (*/data/raw/*) e dos principais dados processados (*/looker\_studio\_output/*) é mantida em um serviço de armazenamento em nuvem pessoal e seguro (*Google Drive*), sincronizado com um computador local.

#### Processo de Backup:

O processo de backup é tanto automático quanto manual, garantindo redundância:

- **Backup Contínuo do Código:** Após cada sessão de trabalho significativa, as alterações no código-fonte e na documentação são salvas (*commit*) e sincronizadas (*push*) com o repositório remoto no *GitHub*. Este processo é manual, mas realizado com alta frequência (diariamente ou ao final de cada etapa de desenvolvimento).
- **Backup Automático do Ambiente:** A plataforma *GitHub Codespaces* gerencia automaticamente o estado do ambiente de desenvolvimento, preservando o trabalho entre as sessões.
- **Backup Periódico dos Dados:** Ao final de cada uma das principais etapas do pipeline (Etapa 1, 2, 3), os arquivos CSV gerados nas pastas */data/processed/* e */looker\_studio\_output/* são copiados manualmente para a pasta de backup no *Google Drive*.

#### Responsabilidade e Recuperação:

- **Responsabilidade:** A autora principal da pesquisa, Luana Cristina-Silva, é a única responsável por executar e gerenciar os processos de backup e recuperação, em primeira instância. Em casos futuros, o processo pode ser redirecionado à eventuais pesquisadores integrantes do NIT/Materiais UFSCar.
- **Recuperação em Caso de Incidente:** A estratégia de múltiplas camadas permite uma recuperação robusta:
  - **Falha no *Codespace*:** Se o ambiente do *Codespaces* for corrompido ou perdido, uma nova instância pode ser criada instantaneamente. O código-fonte mais recente pode ser clonado (*git clone*) diretamente do repositório *GitHub*. Os dados brutos podem ser recuperados do backup no *Google Drive* e reenviados para a pasta */data/raw/*, permitindo que todo o pipeline seja re-executado para regenerar os dados processados.
  - **Perda de Acesso ao *GitHub*:** No caso improvável de perda de acesso ao repositório, a cópia de trabalho no *Codespaces* (ou em um clone local) contém o histórico completo do *Git*.
  - **Perda de Dados Locais:** O repositório *GitHub* serve como o *backup* mestre do código, e o *Google Drive* como o *backup* mestre dos dados brutos.

How will you manage access and security?

A gestão de acesso e segurança foi planejada considerando a natureza pública e não-sensível dos dados, priorizando a integridade do trabalho e a segurança das credenciais de serviço.

#### Riscos de Segurança e Gestão:

O principal risco de segurança para este projeto não é a exposição de dados (pois são públicos), mas sim:

- **Corrupção ou Perda de Dados/Código:** Gerenciado pela estratégia de backup em três camadas (*Codespaces*, *GitHub*, *Google Drive*) e pelo controle de versão Git, que permite reverter para versões anteriores.
- **Exposição de Credenciais de API:** A chave da API do *Google* (arquivo JSON da Conta de Serviço) é o único dado sensível do projeto. Este risco é gerenciado da seguinte forma:
  - O arquivo de credenciais é armazenado em uma pasta */config/* separada.
  - Esta pasta foi explicitamente adicionada ao arquivo *.gitignore* do projeto. Isso impede que o *Git* rastreie ou envie o arquivo de credenciais para o repositório público no *GitHub*, mantendo-o seguro e local ao ambiente de desenvolvimento.

#### Controle de Acesso:

- **Ambiente de Desenvolvimento:** O acesso ao *GitHub Codespaces* é controlado pela conta pessoal e autenticação de dois fatores da autora no *GitHub*.
- **Repositório de Código:** O repositório no *GitHub* (*silvaluah/sisimm-cti-amazon*) é público para leitura, promovendo a transparência, mas o acesso de escrita (*push*) é restrito apenas à autora.
- **Dados de Saída (Google Sheets):** A planilha *INPUT\_V2* no *Google Sheets* é privada. O acesso de edição é concedido apenas à autora e à Conta de Serviço específica do projeto, cujo acesso é controlado pelo arquivo de credenciais JSON.

#### Acesso Seguro para Colaboradores:

No momento, o projeto não possui colaboradores externos. Se um colaborador fosse adicionado no futuro, o acesso seria concedido de forma granular:

- **Código:** Seria adicionado como colaborador no repositório *GitHub*.
- **Dados de Saída:** Seria adicionado com permissão de "Leitor" ou "Comentador" na Planilha *Google*, a menos que a permissão de "Editor" fosse estritamente necessária. A chave da API da Conta de Serviço não seria compartilhada.

#### Transferência Segura de Dados:

Não há coleta de dados em campo. Todas as transferências de dados ocorrem através de protocolos seguros padrão:

- O download dos dados brutos das plataformas (*Scopus*, *Espacenet*) é feito via HTTPS.
- A sincronização de código com o *GitHub* é feita via HTTPS ou SSH.
- A comunicação com as APIs do *Google* (*Sheets*) é feita via HTTPS, criptografada ponta a ponta.

#### Dados Confidenciais:

Conforme mencionado na seção de ética, o projeto não coleta, armazena ou processa dados confidenciais. Portanto, não são necessárias medidas de segurança adicionais, como criptografia de disco ou conformidade com padrões como a ISO 27001.

#### Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

Nem todos os dados gerados durante a pesquisa têm o mesmo valor de longo prazo. A estratégia de retenção e compartilhamento foi definida para maximizar a utilidade, a reprodutibilidade e o impacto do projeto.

#### Dados a Serem Retidos e Compartilhados (Alto Valor):

- **Código-Fonte Completo:** Todos os *scriptsPython* (.py) e *notebooksJupyter* (.ipynb) de prototipagem são de altíssimo valor, pois documentam a metodologia completa e permitem a replicação total dos resultados.

- **Dados Processados Finais (Camada de Apresentação):** Os arquivos CSV localizados na pasta /looker\_studio\_output/ (ex: art\_afiliacao.csv, art\_fato.csv, looker\_fato\_cncflora.csv, etc.). Estes são os "produtos finais" do pipeline, limpos, enriquecidos e prontos para análise direta por outros pesquisadores ou para importação em ferramentas de BI.
- **Dicionários de Curadoria Manual:** Os arquivos CSV criados manualmente para padronização (ex: dicionario\_afiliacoes\_final.csv, scopus\_ods\_v1.csv). Estes representam um esforço intelectual significativo e são um ativo de dados valioso por si só, podendo ser reutilizados em outros projetos de bibliometria.
- **Documentação:** Os arquivos README.md e dicionario\_de\_dados.md, que fornecem o contexto e as definições necessárias para entender e reutilizar os dados e o código.

#### Dados a Serem Retidos, mas Não Necessariamente Compartilhados Publicamente:

- **Dados Brutos (/data/raw/):** Embora essenciais para a reprodutibilidade, os dados brutos extraídos de bases comerciais como a Scopus não serão redistribuídos publicamente para respeitar os termos de serviço das plataformas. Eles serão preservados em um arquivo privado, mas a documentação (README.md) incluirá as queries e os métodos exatos para que outros pesquisadores possam obter os mesmos dados brutos da fonte original.

#### Dados a Serem Descartados (Baixo Valor):

- **Dados Intermediários (/data/processed/):** Arquivos como dim\_autores\_scopus.csv ou pon\_artigo\_autores\_scopus.csv são importantes para a execução do pipeline, mas não têm valor de longo prazo para um usuário final, pois seu conteúdo já está representado de forma mais útil nos arquivos finais da camada de apresentação. Como eles podem ser totalmente regenerados a partir do código e dos dados brutos, sua preservação de longo prazo não é necessária.

#### Usos Futuros Previsíveis:

Os dados de alto valor podem ser reutilizados para:

- **Validação:** Outros pesquisadores podem usar os dados e o código para validar e replicar os resultados desta dissertação.
- **Novos Estudos:** O conjunto de dados limpo pode ser a base para novas pesquisas, como análises de redes de colaboração mais profundas, estudos de impacto de financiamento (cruzando com a coluna funding\_details), ou análises longitudinais da evolução de temas de pesquisa.
- **Ensino:** O projeto como um todo (código + dados) pode ser usado como um estudo de caso em disciplinas de metodologia de pesquisa, ciência de dados ou bibliometria.

#### Período de Retenção:

Os dados de alto valor (código, dados finais, dicionários) serão preservados por um **período mínimo de 10 anos** após a conclusão do projeto, ou indefinidamente, se o repositório de dados escolhido garantir a preservação perpétua. Não há requisitos contratuais, legais ou regulatórios que exijam a destruição dos dados.

What is the long-term preservation plan for the dataset?

O plano de preservação de longo prazo foi projetado para garantir que os resultados da pesquisa permaneçam acessíveis, citáveis e utilizáveis para a comunidade científica além da duração deste projeto de mestrado.

#### Repositório de Dados:

Os dados de alto valor (o código-fonte completo e os arquivos CSV finais da pasta /looker\_studio\_output/) serão depositados no *Zenodo* (<https://zenodo.org>).

- **Justificativa da Escolha:** O *Zenodo* é um repositório de acesso aberto, multidisciplinar e de propósito geral, operado pelo CERN e *OpenAIRE*. Ele foi escolhido por várias razões:
  - Atribuição de DOI: O *Zenodo* atribui automaticamente um DOI (*Digital Object Identifier*) a cada *upload*, o que torna o conjunto de dados um produto de pesquisa formalmente citável.

- Integração com o *GitHub*: Ele possui uma integração direta com o *GitHub* que permite arquivar automaticamente uma versão ("*release*") do repositório, garantindo que o código e os dados associados àquela versão sejam preservados juntos.
- Preservação de Longo Prazo: Sendo mantido pelo CERN, o *Zenodo* tem um compromisso institucional com a preservação de longo prazo dos dados.
- Sem Custos: O serviço é gratuito para conjuntos de dados de até 50 GB, o que é mais do que suficiente para as necessidades deste projeto.

#### **Custos:**

Não haverá custos associados ao depósito e preservação dos dados no *Zenodo*.

#### **Preparação dos Dados para Preservação:**

Um tempo específico será alocado após a defesa da dissertação para preparar o conjunto de dados para o depósito. O esforço estimado é de aproximadamente [48 horas] de trabalho e incluirá as seguintes tarefas:

- **Limpeza Final do Código:** Revisar todos os scripts e notebooks, garantindo que os comentários estejam claros e que informações sensíveis (como chaves de API) tenham sido removidas.
- **Finalização da Documentação:** Realizar uma revisão final e completa do README.md e do dicionario\_de\_dados.md, garantindo que toda a metodologia, estrutura de arquivos e definições de variáveis estejam perfeitamente documentadas para um usuário externo.
- **Criação do Pacote de Depósito:** Criar um arquivo .zip contendo:
  - Todo o repositório de código.
  - A pasta /looker\_studio\_output/ com os arquivos de dados finais.
  - O arquivo dicionario\_de\_dados.md.
  - O arquivo README.md.
- **Preenchimento de Metadados:** Preencher o formulário de depósito no Zenodo com os metadados baseados no padrão Dublin Core (título, autor, resumo, palavras-chave, licença, etc.) para garantir a máxima "encontrabilidade" do conjunto de dados.
- **Criação de um "Release" no GitHub:** Criar uma versão oficial no GitHub (ex: v1.0.0) e conectar ao *Zenodo* para arquivamento automático.

Este plano garante que os resultados da pesquisa não se percam ao final do projeto, mas sim se tornem um recurso duradouro e valioso para a comunidade acadêmica.

#### **Data Sharing**

How will you share the data?

A estratégia de compartilhamento foi planejada para maximizar a visibilidade, o acesso e a reutilização dos resultados da pesquisa, em conformidade com os princípios da ciência aberta.

#### **Como os Usuários Potenciais Descobrirão os Dados?**

A descoberta dos dados será facilitada por múltiplos canais:

- **Publicação Acadêmica:** A própria dissertação, ao ser publicada nos repositórios da universidade, conterá um link direto para o conjunto de dados, referenciado pelo seu DOI (*Digital Object Identifier*).
- Artigos científicos futuros derivados desta pesquisa também citarão e incluirão o *link* para os dados.
- **Repositórios de Dados:** O conjunto de dados será indexado pelo *Zenodo*, tornando-o pesquisável em plataformas de busca de dados acadêmicos, como o *Google Dataset Search*.
- **Repositório de Código:** O repositório público no *GitHub* (github.com/silvaluah/sisimm-cti-amazon) servirá como um portal de acesso tanto ao código quanto à documentação, que por sua vez apontará para o conjunto de dados formalmente depositado no *Zenodo*.

#### **Com Quem os Dados Serão Compartilhados e Sob Quais Condições?**

- **Público:** Os dados processados finais e o código-fonte serão compartilhados publicamente com toda a comunidade (pesquisadores, estudantes, gestores públicos, jornalistas de dados, etc.), sem necessidade de registro ou solicitação.
- **Condições:** O compartilhamento será regido por licenças abertas. O código será compartilhado sob a Licença MIT, e os dados sob a licença *Creative Commons Attribution 4.0 (CC-BY 4.0)*. A única condição para reutilização é a atribuição, ou seja, a citação adequada do trabalho original (a dissertação/artigo) e do conjunto de dados (através do seu DOI).

#### **Mecanismo de Compartilhamento:**

O compartilhamento será feito primariamente através de um repositório de dados estabelecido (*Zenodo*). Esta abordagem foi escolhida em detrimento do tratamento de solicitações diretas por ser mais eficiente, sustentável e por garantir a atribuição de um identificador persistente. O *GitHub* servirá como um ponto de acesso secundário ao código e à documentação, com links claros para o conjunto de dados no *Zenodo*.

#### **Disponibilidade dos Dados:**

Os dados serão disponibilizados publicamente imediatamente após a defesa e aprovação da dissertação de mestrado, sem nenhum período de embargo.

#### **Identificador Persistente (DOI):**

Sim, um dos principais motivos para a escolha do *Zenodo* como repositório é que ele atribui automaticamente um DOI a cada versão do conjunto de dados. Isso garante que o dataset seja um objeto de pesquisa estável, citável e permanentemente localizável, mesmo que os links da web mudem no futuro. A forma de citação recomendada será fornecida diretamente na página do dataset no *Zenodo*.

Are any restrictions on data sharing required?

Não, não são previstas restrições significativas no compartilhamento dos dados criados e processados por este projeto, uma vez que o objetivo principal é a disseminação do conhecimento. No entanto, uma distinção importante é feita em relação aos dados de origem.

#### **Restrição sobre Dados Brutos de Terceiros:**

- **Restrição:** Conforme os termos de serviço de bases de dados comerciais como a Scopus, os dados brutos extraídos diretamente da plataforma não serão redistribuídos publicamente.
- **Ação para Minimizar a Restrição:** Para garantir a reprodutibilidade sem violar os direitos de terceiros, o projeto adotará a seguinte medida: a documentação (README.md) e os scripts incluirão as queries de busca exatas e o método de exportação utilizado. Isso permitirá que qualquer pesquisador com acesso à Scopus possa replicar a coleta de dados brutos de forma independente e, em seguida, utilizar o código-fonte fornecido para reproduzir todo o pipeline de processamento e chegar aos mesmos resultados finais.

#### **Período de Uso Exclusivo:**

Não haverá período de uso exclusivo (embargo) para os dados gerados. A autora não tem interesse comercial ou de patenteamento derivado deste conjunto de dados específico. Os dados serão liberados para a comunidade assim que o trabalho acadêmico associado (a dissertação) for finalizado e aprovado.

#### **Acordo de Compartilhamento de Dados (Data Sharing Agreement):**

Não será necessário um Acordo de Compartilhamento de Dados. As licenças abertas escolhidas (MIT para o código e CC-BY 4.0 para os dados) já servem como um acordo legal claro e padronizado, definindo os termos de uso e reutilização para qualquer pessoa, sem a necessidade de acordos individuais.

#### **Responsibilities and Resources**

Who will be responsible for data management?

A gestão de dados para este projeto é centralizada e claramente definida, com todas as responsabilidades recaindo sobre a pesquisadora principal.

#### **Responsabilidade Geral (Implementação e Revisão do PGD):**

Luana Cristina-Silva, como autora principal e pesquisadora do projeto de dissertação de mestrado, é a única responsável pela implementação, monitoramento, revisão e atualização deste Plano de Gestão de Dados (PGD).

O plano será revisado em marcos importantes do projeto (ex: finalização da coleta, conclusão do *pipeline*) para garantir que ele permaneça relevante e preciso.

#### **Responsabilidade por Atividade:**

Como este é um projeto individual, Luana Cristina-Silva é responsável por todas as atividades de gestão de dados, que incluem:

- **Coleta de Dados:** Executar as buscas nas bases de dados, exportar os dados brutos e armazená-los de forma segura.
- **Desenvolvimento do Pipeline:** Escrever, testar e depurar todos os scripts Python para o processamento e enriquecimento dos dados.
- **Produção de Metadados e Documentação:** Manter o arquivo README.md atualizado, criar e preencher o `dicionario_de_dados.md`, e garantir que o código esteja devidamente comentado.
- **Garantia de Qualidade:** Executar os notebooks de prototipagem, validar os resultados de cada etapa do pipeline e realizar a curadoria manual dos dicionários de autoridade.
- **Armazenamento e Backup:** Gerenciar as cópias de segurança do código no GitHub e dos dados no Google Drive, conforme descrito no plano.
- **Arquivamento e Compartilhamento de Longo Prazo:** Preparar o pacote de dados final, preencher os metadados e depositar o código e os dados no repositório Zenodo após a conclusão da pesquisa.
- **Conformidade:** Garantir que todas as políticas institucionais, de direitos autorais (IPR) e de ética em pesquisa sejam respeitadas ao longo do projeto.

#### **Divisão de Responsabilidades em Colaborações:**

Este projeto não envolve parceiros externos ou um consórcio. Todas as responsabilidades de gestão de dados são assumidas pela autora principal.

#### **Propriedade dos Dados e Acordos:**

A propriedade intelectual do código e dos dados gerados pertence à autora, conforme as políticas de propriedade intelectual da Universidade Federal de São Carlos. Não há acordos de consórcio ou contratos com parceiros, pois trata-se de um trabalho individual.

What resources will you require to deliver your plan?

O projeto foi planejado para ser executado utilizando recursos de baixo custo, amplamente disponíveis no ambiente acadêmico e de código aberto, não exigindo investimentos financeiros significativos.

#### **Expertise e Treinamento:**

- A autora possui a expertise técnica necessária em programação *Python* e na biblioteca *Pandas* para desenvolver o *pipeline* de dados.
- A familiaridade com o controle de versão *Git/GitHub*, ambientes de desenvolvimento como o *VS Code*, e plataformas de BI como o *Looker Studio* já existe.
- Nenhum treinamento adicional ou contratação de especialista é necessário para a execução deste plano.

#### **Hardware e Software:**

- **Hardware:** O desenvolvimento é realizado inteiramente no GitHub Codespaces, um ambiente de desenvolvimento em nuvem que elimina a necessidade de um computador local de alta performance. Os recursos computacionais fornecidos pelo nível gratuito do GitHub Codespaces são suficientes para o volume de dados e a complexidade do processamento. Nenhum *hardware* adicional precisa ser adquirido.
- **Software:** Todo o software utilizado no projeto é gratuito e/ou de código aberto:
  - Linguagem: *Python*.
  - Bibliotecas: *Pandas*, *Gspread*, *Geopy*, etc (gerenciadas via *pip*).
  - Ambiente de Desenvolvimento: *VS Code* no *GitHub Codespaces*.
  - Controle de Versão: *Git* e *GitHub*.
  - Armazenamento em Nuvem: *Google Drive* (nível gratuito).
  - Visualização: *Google Looker Studio* (gratuito).

- Não há necessidade de aquisição de licenças de *software*.

**Custos de Repositório de Dados:**

O repositório de dados de longo prazo escolhido, *Zenodo*, é um serviço gratuito para o arquivamento de conjuntos de dados de pesquisa, incluindo a atribuição de DOI. Não haverá cobranças para o depósito ou a preservação dos dados.

Em resumo, o plano de gestão de dados será entregue utilizando a infraestrutura e os recursos já disponíveis para a pesquisadora através de sua conta acadêmica e de serviços gratuitos de código aberto, não havendo necessidade de orçamento ou recursos adicionais. O principal recurso alocado é o tempo da pesquisadora para executar as tarefas de curadoria, desenvolvimento e documentação descritas neste plano.