



Universidade Federal de Campina Grande
Centro de Engenharia Elétrica e Informática
Departamento de Engenharia Elétrica

Leticia Rebecca Medeiros de Lucena

Uma Abordagem de Visual Analytics para Estudos de Estado da Arte: Desenvolvimento de um Dashboard Aplicado a Baterias de Segunda Vida

Campina Grande, Paraíba, Brasil
26 de março de 2026

Leticia Rebecca Medeiros de Lucena

Uma Abordagem de Visual Analytics para Estudos de Estado da Arte: Desenvolvimento de um Dashboard Aplicado a Baterias de Segunda Vida

Trabalho de Conclusão de Curso submetido à
Unidade Acadêmica de Engenharia Elétrica
da Universidade Federal de Campina Grande
como parte dos requisitos necessários para a
obtenção do grau de Bacharel em Ciências no
Domínio da Engenharia Elétrica.

Orientador: Prof. Gutemberg Gonçalves dos Santos Júnior, Dr.

Campina Grande, Paraíba, Brasil

26 de março de 2026

Leticia Rebecca Medeiros de Lucena

Uma Abordagem de Visual Analytics para Estudos de Estado da Arte: Desenvolvimento de um Dashboard Aplicado a Baterias de Segunda Vida

Trabalho de Conclusão de Curso submetido à Unidade Acadêmica de Engenharia Elétrica da Universidade Federal de Campina Grande como parte dos requisitos necessários para a obtenção do grau de Bacharel em Ciências no Domínio da Engenharia Elétrica.

Aprovado em: ____/____/____

Prof. Gutemberg Gonçalves dos Santos Júnior, Dr.
Orientador

Prof. Danilo Freire de Souza Santos, Dr.
Avaliador

Campina Grande, Paraíba, Brasil
26 de março de 2026

Dedico este trabalho aos meus pais, João Vidal (in memoriam) e Ednês

Agradecimentos

Agradeço primeiramente ao Senhor, por ter me capacitado, guiado meus passos e me dado força ao longo de toda essa jornada. O Senhor me sustentou quando eu não tinha forças para seguir e me concedeu a graça de finalizar a graduação, a Ele todo o louvor. "Porque dEle, e por Ele, e para Ele são todas as coisas; glória, pois, a Ele eternamente. Amém!"(Romanos 11:36).

Aos meus pais, pelo apoio incondicional, incentivo e compreensão em todos os momentos. Infelizmente meu pai não verá a filha oficialmente engenheira, mas se fez presente em vida trazendo motivação nos dias mais difíceis. Mainha, apesar de tantas circunstâncias desafiadoras, permanece como uma rocha para as filhas. Meus pais acreditaram em mim quando eu não acreditava.

Às minhas irmãs e meu cunhado, que considero um irmão mais velho. Sou muito grata por perceber que não somos unidas apenas por laços de sangue e, apesar da distância física, o amor segue o mesmo. Meu cunhado-irmão, fonte de inspiração e orgulho, trouxe conselhos valiosos desde a aprovação no vestibular até aqui. A engenharia parecia mais interessante por causa dele.

À Daniel Loula, o homem que me mostrou que amor é construção, suporte, carinho e compreensão. Apesar de não ter visto meu início de jornada na graduação, o fim dela não seria leve sem ele.

Ao professor Gutemberg Junior, pelos ensinamentos, orientação e contribuição fundamental para o meu desenvolvimento. Além de ter sido essencial como professor durante a jornada da graduação, aceitou o desafio de orientar a produção desse trabalho à distância.

"In God we trust, all others must bring data."

William Edwards Deming

Resumo

Este relatório descreve as atividades no âmbito do Trabalho de Conclusão de Curso (TCC) de Engenharia Elétrica. O trabalho abordou o desenvolvimento deconsistiu no desenvolvimento de um sistema de visualização de dados, na forma de um dashboard interativo, voltado à integração, tratamento e análise de dados provenientes de bases científicas e tecnológicas como IEEE Xplore, Scopus e USPTO, aliados à aplicação de técnicas de análise quantitativa e visual analytics. O desenvolvimento do dashboard foi orientado à facilitação da interpretação de grandes volumes de informação, contribuindo para a identificação de principais autores, instituições, países e evolução temporal das publicações.

Palavras-chaves: Estado da arte; visual analytics; dashboard; bibliometria; baterias.

Abstract

This report describes the activities carried out within the scope of the Final Course Project (TCC) of Electrical Engineering. The work consisted in the development of a data visualization system in the form of an interactive dashboard, aimed at the integration, processing, and analysis of data from scientific and technological databases such as IEEE Xplore, Scopus, and USPTO, combined with quantitative analysis and visual analytics techniques. The dashboard was designed to facilitate the interpretation of large volumes of information, contributing to the identification of leading authors, institutions, countries, and the temporal evolution of publications.

Key-words: State of the art; visual analytics; dashboard; bibliometrics; batteries.

Lista de abreviaturas e siglas

UAEE	Unidade Acadêmica de Engenharia Elétrica
CEEI	Centro de Engenharia Elétrica e Informática
UFCG	Universidade Federal de Campina Grande
SOH	<i>State of Health</i> - estado de saúde
SOC	<i>State of charge</i> - estado de carga
IEEE	Instituto de Engenheiros Eletricistas e Eletrônicos
USPTO	<i>United States Patent and Trademark Office</i>
CPC	<i>Cooperative Patent Classification</i>

Lista de ilustrações

Figura 1 – Pipeline da metodologia de construção do banco de dados. Fonte: Autoria própria.	11
Figura 2 – Distribuição de score utilizando apenas CPY. Fonte: autoria própria . .	21
Figura 3 – Distribuição de score utilizando CPY normalizado. Fonte: autoria própria	22
Figura 4 – Tela principal. Fonte: Autoria própria.	23
Figura 5 – Tela de patentes. Fonte: Autoria própria.	24
Figura 6 – Tela de artigos. Fonte: Autoria própria.	24
Figura 7 – Tela de distribuição geográfica. Fonte: Autoria própria.	25
Figura 8 – Resultado de implementação de filtros: Tela inicial. Fonte: Autoria própria.	26
Figura 9 – Resultado de implementação de filtros: Autoria de artigos. Fonte: Autoria própria.	26
Figura 10 – Resultado de implementação de filtros: Autoria de artigos. Fonte: Autoria própria.	27

Lista de tabelas

Tabela 1	–	Palavras-chave para identificação de instituições	14
Tabela 2	–	Listagem de elementos do <i>dashboard</i>	20

Sumário

1	INTRODUÇÃO	1
1.1	Objetivos	3
1.1.1	Objetivo Geral	3
1.1.2	Objetivos Específicos	3
2	FUNDAMENTAÇÃO TEÓRICA	4
2.1	Estado da arte	4
2.1.1	Revisão de Literatura e Estado da Arte	4
2.1.2	Visualização de Dados e Dashboards	5
2.2	Estudo de caso	7
2.2.1	Sistemas de Armazenamento de Energia com Baterias (BESS)	7
2.2.2	Baterias de Segunda Vida	8
3	METODOLOGIA	10
3.1	Natureza e objetivo da pesquisa	10
3.2	Bases de dados	11
3.3	Normalização e modelagem dos dados de artigos	13
3.3.1	Nome do autor	13
3.3.2	País de filiação do autor	14
3.3.3	Instituição de filiação do autor	14
3.3.4	Identificação de inconsistências	15
3.3.5	Identificação de artigos	15
3.4	Método de ranqueamento e seleção de artigos	16
3.5	Normalização e modelagem de patentes	18
3.6	Análise dos dados	19
3.7	Construção do <i>dashboard</i>	19
4	RESULTADOS	21
4.1	Definição de métricas	21
4.2	Normalização de dados	22
4.3	Telas construídas	22
4.4	Filtros	25
5	CONSIDERAÇÕES FINAIS	28
	Referências	29

1 Introdução

A crescente complexidade dos estudos de estado da arte em engenharia, especialmente em áreas multidisciplinares como sistemas de armazenamento de energia, tem demandado o desenvolvimento de estratégias mais eficientes para organização, análise e visualização de dados [1]. Nesse contexto, a motivação central deste trabalho não está associada exclusivamente a um tema específico de aplicação, mas à necessidade de estruturar uma abordagem que permita a condução de estudos de estado da arte de forma mais eficiente, sistemática e reprodutível.

Como estudo de caso para aplicação dessa proposta, foi selecionado o tema de baterias de íons de lítio de segunda vida. A escolha desse tema está diretamente relacionada ao contexto de atuação da autora em um instituto de pesquisa, no qual há a necessidade constante de análise de tendências tecnológicas e identificação de oportunidades no setor elétrico. Nesse cenário, o reaproveitamento de baterias provenientes de veículos elétricos se apresenta como uma alternativa promissora [2], especialmente quando associado a aplicações em sistemas de armazenamento de energia em larga escala.

O interesse crescente por esse tema está diretamente relacionado à expansão do mercado de veículos elétricos, impulsionado por políticas globais de descarbonização e transição energética. Esse movimento tem resultado em um aumento significativo da demanda por baterias de íons de lítio, ao mesmo tempo em que evidencia desafios relacionados à limitação de recursos naturais e ao gerenciamento adequado do ciclo de vida desses dispositivos. Nesse contexto, estratégias baseadas nos princípios de redução, reutilização e reciclagem (3R) tornam-se fundamentais para garantir a sustentabilidade do setor.

Além disso, iniciativas globais voltadas à mitigação das mudanças climáticas, como o Acordo de Paris, adotado durante a Conferência das Partes (COP21), reforçam a necessidade de desenvolvimento de tecnologias que contribuam para a transição energética [5], incluindo soluções de armazenamento de energia. Dessa forma, o estudo de baterias de segunda vida se insere em um contexto estratégico e de alta relevância tecnológica.

Apesar da importância desse tema, a condução de estudos de estado da arte apresenta desafios significativos. Um dos principais problemas identificados neste trabalho está relacionado à grande quantidade de dados disponíveis e ao volume de informações que precisam ser analisadas para que o estudo seja representativo em nível internacional. A compreensão adequada de uma tecnologia exige a análise de um conjunto amplo de publicações, permitindo identificar sua evolução, os principais autores, os países com maior produção científica e as tendências de desenvolvimento ao longo do tempo.

Entretanto, o tempo necessário para a realização dessa análise se apresenta como

uma limitação crítica, especialmente em contextos de pesquisa aplicada, nos quais é necessário obter rapidamente uma visão geral sobre determinado tema. O acesso às informações ocorre por meio de diferentes bases de dados, como IEEE Xplore¹, Scopus² e bases de patentes, cada uma com suas próprias estruturas, interfaces e metodologias de busca. Embora essas plataformas sejam eficientes individualmente, a integração das informações provenientes dessas fontes é complexa e demanda esforço significativo por parte do pesquisador.

Além disso, a ausência de padronização na organização e visualização dos dados dificulta a realização de análises comparativas e a extração de *insights* relevantes. Antes da adoção de uma abordagem estruturada, as análises baseavam-se predominantemente na leitura individual de artigos, o que, além de demandar tempo elevado, limitava a abrangência do estudo e dificultava a identificação de padrões e métricas relevantes.

Nesse contexto, observa-se uma lacuna na integração entre metodologias consolidadas de revisão da literatura e ferramentas de visualização interativa que permitam a análise dinâmica e orientada por dados. Embora existam metodologias estruturadas para revisões sistemáticas e ferramentas de análise bibliométrica, essas abordagens, em geral, não oferecem soluções integradas, acessíveis e de fácil implementação que permitam a atualização contínua dos dados e a exploração interativa das informações.

Diante dessa lacuna, este trabalho propõe a utilização de *dashboards* interativos como ferramenta para apoiar a condução de estudos de estado da arte. A escolha pela utilização do *Power BI* como plataforma de desenvolvimento está relacionada à sua capacidade de organizar grandes volumes de dados em estruturas visuais claras, bem como à possibilidade de interação dinâmica com as informações, permitindo a exploração de diferentes perspectivas em tempo real. Além disso, a ferramenta possibilita a integração com diferentes fontes de dados e a atualização automática das informações, contribuindo para a construção de uma base de análise contínua.

A relevância dessa proposta é evidenciada por sua aplicação em situações reais de pesquisa. Em contextos nos quais foi necessário avaliar rapidamente tecnologias ainda pouco conhecidas, a utilização da metodologia estruturada, aliada à visualização por meio de *dashboards*, permitiu a organização dos dados coletados e a construção de uma base sólida para a tomada de decisão. Da mesma forma, em atividades relacionadas à elaboração de estudos de anterioridade, a abordagem proposta possibilitou a análise estruturada de informações em um curto período de tempo, contribuindo para a fundamentação técnica de propostas desenvolvidas.

¹ Base de dados digital criada pelo IEEE com acesso a artigos, conferências e normas técnicas nas áreas de engenharia e computação.

² Base de dados bibliográfica multidisciplinar criada pela Elsevier, com acesso a artigos científicos, conferências e métricas de pesquisa.

Dessa forma, este trabalho contribui ao propor uma abordagem integrada que combina metodologia de análise, estruturação de dados e visualização interativa, permitindo tornar o processo de construção de estudos de estado da arte mais eficiente, acessível e orientado por métricas.

1.1 Objetivos

Este tópico apresenta os objetivos que orientam o desenvolvimento deste trabalho, definindo de forma concisa os resultados que se pretende alcançar. A delimitação dos objetivos tem função de guiar as etapas metodológicas e direcionar a análise dos dados, garantindo coerência entre proposta, execução e resultados obtidos.

1.1.1 Objetivo Geral

O trabalho em questão tem como objetivo central desenvolver um *dashboard* interativo, baseado em uma metodologia estruturada de coleta, tratamento e análise de dados, para apoiar a condução de estudos de estado da arte e de anterioridade, visando auxiliar pesquisadores e instituições na tomada de decisão em processos de pesquisa e desenvolvimento tecnológico.

1.1.2 Objetivos Específicos

- Definir uma metodologia estruturada para coleta e seleção de dados em bases científicas e tecnológicas;
- Desenvolver um processo de tratamento e organização dos dados que permita sua integração em uma base única;
- Estabelecer critérios relevantes para análise do estado da arte, como distribuição temporal, autores, países e áreas de pesquisa;
- Implementar um *dashboard* interativo utilizando a ferramenta *Power BI* para visualização dos dados;
- Avaliar a aplicabilidade da metodologia e da ferramenta proposta em contextos reais de pesquisa;
- Demonstrar o uso da abordagem por meio de um estudo de caso relacionado a baterias de segunda vida.

2 Fundamentação Teórica

Este capítulo tem como objetivo apresentar os principais conceitos, métodos e abordagens que sustentam o desenvolvimento deste trabalho. Inicialmente, são discutidos os fundamentos relacionados à análise do estado da arte, com ênfase na utilização de técnicas bibliométricas para organização e avaliação da produção científica. Em seguida, são abordados os conceitos de *Visual Analytics* e sua importância na interpretação de grandes volumes de dados. Por fim, são apresentados os principais elementos associados às bases de dados utilizadas e ao contexto tecnológico do estudo, fornecendo o embasamento necessário para a compreensão da metodologia proposta.

2.1 Estado da arte

A presente seção tem como objetivo apresentar uma visão consolidada do estado da arte acerca do tema investigado, a partir da análise da literatura científica e tecnológica disponível. Por meio da revisão de trabalhos relevantes, busca-se identificar as principais abordagens, tendências e lacunas existentes, fornecendo subsídios para a contextualização do problema e para a fundamentação das escolhas metodológicas adotadas neste trabalho.

2.1.1 Revisão de Literatura e Estado da Arte

A revisão de literatura constitui uma etapa fundamental no desenvolvimento de pesquisas científicas, especialmente em áreas tecnológicas, nas quais o conhecimento evolui de forma contínua e cumulativa. Nesse contexto, o conceito de estado da arte está diretamente relacionado à compreensão do nível atual de desenvolvimento de uma determinada tecnologia ou área de estudo [4], considerando o conjunto de conhecimentos já produzidos ao longo do tempo. O estado da arte pode ser entendido como a análise do conhecimento existente sobre um determinado tema, permitindo identificar como diferentes abordagens têm sido utilizadas para resolver problemas específicos, quais soluções já foram propostas e quais lacunas ainda permanecem em aberto. Dessa forma, essa análise possibilita não apenas a compreensão do funcionamento de uma tecnologia, mas também a identificação de tendências, limitações e oportunidades de avanço.

Na área de engenharia, a realização de estudos de estado da arte é particularmente relevante, uma vez que o desenvolvimento de novas soluções tecnológicas depende diretamente do entendimento prévio do que já foi investigado e consolidado na literatura. Esse processo permite evitar redundâncias, orientar a formulação de novas propostas e embasar tecnicamente decisões relacionadas à pesquisa e desenvolvimento.

Diferentes abordagens metodológicas podem ser utilizadas para a condução de revisões de literatura, destacando-se, entre elas, as revisões sistemáticas e as revisões exploratórias. As revisões sistemáticas são caracterizadas por um processo estruturado e rigoroso de busca, seleção e análise de publicações, seguindo critérios previamente definidos. Esse tipo de abordagem busca garantir a reprodutibilidade do estudo e minimizar vieses na seleção das fontes, sendo amplamente utilizado em pesquisas que exigem elevado nível de rigor metodológico.

Por outro lado, as revisões exploratórias apresentam maior flexibilidade, sendo utilizadas principalmente em estágios iniciais de investigação ou em contextos nos quais há necessidade de obtenção rápida de uma visão geral sobre o tema. Essa abordagem permite maior liberdade na seleção e análise das fontes, favorecendo a compreensão ampla do domínio estudado e a identificação preliminar de tendências e lacunas.

Apesar de sua relevância, ambas as abordagens apresentam limitações quando aplicadas a contextos que envolvem grandes volumes de dados e necessidade de análise ágil. As revisões sistemáticas, embora estruturadas, tendem a demandar um elevado tempo de execução e esforço analítico [6], o que pode dificultar sua aplicação em cenários de pesquisa aplicada. Já as revisões exploratórias, apesar da flexibilidade de estrutura, podem apresentar desafios relacionados à organização das informações e à padronização da análise [7], especialmente quando se busca extrair métricas quantitativas ou realizar comparações entre diferentes conjuntos de dados. Dessa forma, embora as metodologias de revisão de literatura sejam fundamentais para a construção do estado da arte, ainda existem desafios relacionados à eficiência na organização, análise e interpretação de grandes volumes de informações, especialmente em áreas tecnológicas com alta produção científica.

2.1.2 Visualização de Dados e Dashboards

A análise de grandes volumes de dados tem se tornado cada vez mais relevante em diferentes áreas do conhecimento, especialmente em contextos que envolvem a tomada de decisão baseada em evidências. Nesse cenário, destaca-se o conceito de *Visual Analytics*, definido como a integração entre técnicas de análise de dados e representações visuais interativas, com o objetivo de potencializar a compreensão de informações complexas e apoiar o processo decisório. De acordo com estudos clássicos da área de *Visual Analytics*, como os propostos por [3], a combinação entre processamento computacional e visualização interativa permite explorar grandes conjuntos de dados de forma mais eficiente, facilitando a identificação de padrões, tendências e relações que não seriam facilmente percebidos por meio de análises tradicionais. Nesse contexto, a visualização não atua apenas como um meio de apresentação, mas como parte integrante do processo analítico.

Uma das principais formas de aplicação dos conceitos de *Visual Analytics* é por meio de dashboards, que consistem em interfaces visuais projetadas para apresentar in-

formações de forma organizada, sintética e interativa. Diferentemente de planilhas tradicionais, nas quais os dados são apresentados de forma tabular e, muitas vezes, estática, os dashboards permitem a exploração dinâmica das informações, possibilitando ao usuário interagir com os dados por meio de filtros, segmentações e diferentes perspectivas de análise.

Essa característica interativa torna os dashboards ferramentas mais eficientes para a interpretação de dados, uma vez que facilitam a compreensão de cenários complexos e reduzem a necessidade de processamento manual das informações. Além disso, a organização visual dos dados contribui para uma leitura mais intuitiva, permitindo que o usuário identifique rapidamente padrões e insights relevantes. Nesse contexto, ferramentas de *Business Intelligence (BI)* têm sido amplamente utilizadas para o desenvolvimento de dashboards interativos. Entre essas ferramentas, destaca-se o *Microsoft Power BI*, uma plataforma que permite a integração de dados provenientes de diferentes fontes, bem como a construção de visualizações interativas e personalizadas.

Para que os dashboards apresentem informações confiáveis e consistentes, é necessário realizar previamente o tratamento e a normalização dos dados. Esse processo está fundamentado nos princípios da Modelagem de Dados, que visam organizar os dados de forma estruturada, eliminando redundâncias e inconsistências. No contexto de bases científicas e tecnológicas, é comum encontrar variações na forma de registro das informações, como nomes de instituições, autores e palavras-chave. Dessa forma, torna-se necessário aplicar técnicas de padronização e limpeza dos dados, de modo a garantir a integridade das análises realizadas.

Além disso, destaca-se a importância de distinguir entre a estrutura normativa ideal dos dados — por exemplo, “Departamento, Instituição, País” — e a estrutura observada nas bases reais, frequentemente caracterizada por inconsistências e falta de padronização. A normalização dos dados, nesse sentido, é essencial para viabilizar análises comparativas e a correta interpretação dos resultados. O Power BI possibilita a conexão com diversas bases de dados, a realização de transformações e modelagens dos dados e a criação de dashboards que podem ser atualizados de forma automática. Essa integração entre coleta, tratamento e visualização de dados em uma única plataforma contribui para a eficiência do processo analítico, especialmente em contextos que demandam agilidade e flexibilidade na análise de informações.

Assim, o uso de ferramentas baseadas em *Visual Analytics*, como dashboards interativos, representa um avanço significativo em relação às abordagens tradicionais de análise de dados, permitindo não apenas a organização das informações, mas também a sua exploração de forma dinâmica e orientada à geração de conhecimento.

2.2 Estudo de caso

A presente seção tem como objetivo contextualizar o estudo de caso adotado neste trabalho, abordando os principais aspectos relacionados aos Sistemas de Armazenamento de Energia com Baterias (BESS). A partir dessa contextualização, busca-se delimitar o domínio de aplicação utilizado para demonstrar a metodologia proposta, evidenciando sua relevância no cenário energético atual e sua adequação para análise por meio das ferramentas desenvolvidas. Destaca-se que, apesar do estabelecimento de um tema para demonstração da metodologia proposta, a aplicação é generalista e não está restrita à tecnologia estudada.

2.2.1 Sistemas de Armazenamento de Energia com Baterias (BESS)

Os Sistemas de Armazenamento de Energia em Baterias (BESS) consistem em sistemas capazes de armazenar energia elétrica sob a forma química e disponibilizá-la posteriormente conforme a demanda do sistema elétrico. De forma geral, o funcionamento de um BESS envolve a conversão de energia elétrica proveniente de uma fonte de geração, que pode estar em corrente alternada (CA) ou corrente contínua (CC), por meio de sistemas de eletrônica de potência que realizam a conversão adequada para o carregamento das baterias. Durante a etapa de descarga, o processo inverso ocorre, permitindo a reinjeção da energia armazenada no sistema elétrico.

A crescente adoção de BESS está diretamente associada às transformações recentes nos sistemas elétricos, especialmente devido à expansão das fontes renováveis de energia. Nesse contexto, destaca-se o fenômeno do *curtailment* (corte de geração), que ocorre quando a produção de energia — particularmente de fontes intermitentes, como solar e eólica — precisa ser reduzida ou interrompida devido a limitações operacionais do sistema, como restrições na transmissão ou desequilíbrios entre geração e demanda. Esse processo resulta em perdas econômicas significativas para os agentes geradores e representa um desafio para a eficiência do sistema elétrico.

Diante desse cenário, os sistemas BESS emergem como uma solução estratégica, permitindo armazenar a energia excedente que, de outra forma, seria descartada, e disponibilizá-la em momentos de maior demanda. Além disso, esses sistemas contribuem para a melhoria da confiabilidade e estabilidade da rede elétrica, desempenhando funções importantes no suporte à operação do sistema.

Entre as principais aplicações dos sistemas BESS, destacam-se:

- Arbitragem energética (*energy arbitrage*): armazenamento de energia em períodos de baixa demanda (fora de ponta) e despacho em períodos de alta demanda (ponta), otimizando custos;

- Suporte à estabilidade do sistema: controle de frequência e tensão, contribuindo para a regulação do sistema elétrico;
- Integração de fontes renováveis: mitigação da intermitência de fontes como solar e eólica;
- Redução de curtailment: armazenamento de energia que seria descartada devido a limitações operacionais;
- *Peak shaving*: redução de picos de demanda;
- Aplicações em microrredes: suporte a sistemas isolados ou parcialmente desconectados da rede principal;
- Black start e suporte emergencial: auxílio na recomposição do sistema elétrico após falhas.

No contexto brasileiro, a relevância dos BESS é ainda mais evidente devido às características da matriz elétrica, fortemente baseada em fontes renováveis e interligada por um sistema de transmissão de grande extensão. O crescimento da geração distribuída, especialmente a partir de sistemas fotovoltaicos, tem introduzido novos desafios operacionais, como variações de tensão e frequência, exigindo soluções que garantam a estabilidade do sistema. Nesse sentido, tecnologias mais avançadas, como os sistemas grid-forming, vêm sendo desenvolvidas. Esses sistemas são capazes de atuar como referência de rede, contribuindo para a formação de tensão e frequência mesmo na ausência de uma rede elétrica convencional, o que amplia as possibilidades de aplicação dos BESS em cenários com alta penetração de fontes renováveis.

Adicionalmente, observa-se uma tendência global de expansão do uso de sistemas BESS. Países como a Austrália têm investido fortemente em sistemas de armazenamento em larga escala para lidar com desafios relacionados à transmissão de energia em longas distâncias, enquanto a China se destaca pelo rápido crescimento na produção de baterias e pela implementação de soluções tanto em escala de rede quanto em aplicações distribuídas, como sistemas residenciais integrados a geração fotovoltaica.

2.2.2 Baterias de Segunda Vida

O conceito de baterias de segunda vida refere-se ao reaproveitamento de baterias que já foram utilizadas em sua aplicação original — predominantemente em veículos elétricos —, mas que ainda possuem capacidade remanescente suficiente para desempenhar funções em aplicações menos exigentes. Com o crescimento expressivo do mercado de veículos elétricos, impulsionado por políticas globais de descarbonização, observa-se um aumento significativo no volume de baterias de íons de lítio que atingem o fim de sua

vida útil automotiva. Nessas aplicações, as baterias são geralmente retiradas de operação quando sua capacidade atinge um determinado limiar, tipicamente em torno de 70% a 80% da capacidade original, devido a requisitos de desempenho e segurança.

Entretanto, mesmo após esse ponto, essas baterias ainda apresentam potencial para uso em aplicações estacionárias, especialmente em sistemas de armazenamento de energia. Nesse contexto, o reaproveitamento dessas baterias representa uma oportunidade para estender seu ciclo de vida útil, reduzindo custos e impactos ambientais associados à produção de novas baterias.

A utilização de baterias de segunda vida em sistemas BESS está diretamente alinhada com os princípios de sustentabilidade [9], particularmente os conceitos de redução, reutilização e reciclagem (3R). Além de contribuir para a diminuição do descarte prematuro de materiais, essa abordagem permite a maximização do valor econômico dos recursos já extraídos e processados. No entanto, a aplicação de baterias de segunda vida também apresenta desafios relevantes, como a variabilidade no estado de saúde (SOH), questões relacionadas à segurança, padronização dos sistemas e necessidade de técnicas avançadas de diagnóstico e gerenciamento.

Apesar desses desafios, o uso de baterias de segunda vida tem se consolidado como uma alternativa promissora no contexto dos sistemas de armazenamento de energia, especialmente em aplicações estacionárias, nas quais os requisitos operacionais são menos rigorosos do que no setor automotivo.

3 Metodologia

Este capítulo descreve os procedimentos metodológicos adotados para o desenvolvimento do trabalho, detalhando as etapas de coleta, tratamento e análise dos dados. Inicialmente, são apresentadas as estratégias de busca e seleção das informações nas bases científicas e tecnológicas utilizadas. Em seguida, são descritos os processos de padronização e organização dos dados, bem como os critérios adotados para filtragem e análise. Por fim, é apresentada a estrutura do *dashboard* desenvolvido, destacando sua organização e as funcionalidades implementadas para suporte à análise do estado da arte.

3.1 Natureza e objetivo da pesquisa

O presente trabalho propõe uma metodologia voltada à organização do estado da arte em tecnologias relacionadas a sistemas de armazenamento de energia, com foco na estruturação, análise e visualização de dados provenientes de diferentes bases de informação. Diferentemente de abordagens tradicionais, que frequentemente se concentram em análises específicas e aprofundadas de uma única tecnologia, a metodologia desenvolvida neste estudo tem como objetivo principal a construção de uma estrutura generalista, capaz de ser aplicada a diferentes temas e contextos.

Nesse sentido, a proposta não se limita à análise de uma tecnologia específica, mas sim à definição de um fluxo metodológico que permita a coleta, tratamento e análise de dados de forma integrada, visando a construção de um panorama inicial do estado da arte. Essa abordagem pode ser entendida como uma etapa preliminar à revisão sistemática, funcionando como um mecanismo de organização e preparação das informações que subsidiam análises mais aprofundadas.

A motivação para o desenvolvimento dessa metodologia está associada à dificuldade de integração entre diferentes etapas do processo de pesquisa, especialmente no que se refere à coleta de dados, ao tratamento das informações e à sua posterior análise e visualização. Observa-se que, em muitos casos, essas etapas são realizadas de forma isolada, o que pode comprometer a eficiência do processo analítico e dificultar a extração de insights relevantes. Assim, a metodologia proposta busca integrar essas etapas em um único fluxo estruturado, permitindo não apenas a organização dos dados, mas também a sua exploração de maneira dinâmica e orientada à identificação de padrões e tendências.

Quanto à natureza da pesquisa, este trabalho possui caráter predominantemente quantitativo, uma vez que se concentra na análise de métricas relacionadas à produção científica e tecnológica, como número de publicações, distribuição temporal, instituições

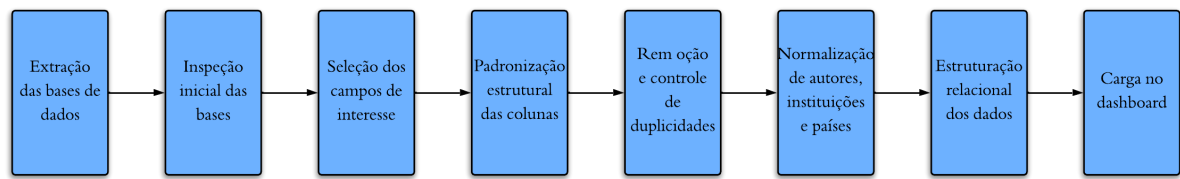


Figura 1 – Pipeline da metodologia de construção do banco de dados. Fonte: Autoria própria.

e países de origem. A escolha por uma abordagem quantitativa está relacionada à intenção de desenvolver uma ferramenta generalista, capaz de ser aplicada a diferentes temas sem a necessidade de interpretações qualitativas específicas, que poderiam restringir sua aplicabilidade.

A Figura 1 apresenta o *pipeline* metodológico adotado para o tratamento e integração dos dados. O fluxo foi estruturado de forma sequencial, contemplando as etapas de extração das bases, inspeção inicial, seleção dos campos de interesse e padronização estrutural das colunas. Em seguida, foram realizadas a remoção de duplicidades e a normalização dos dados de autores, instituições e países. Por fim, os dados foram organizados em uma estrutura relacional e, por fim, carregados no *dashboard*, possibilitando sua análise por meio de técnicas de *Visual Analytics*.

3.2 Bases de dados

Para a construção da base de dados utilizada neste estudo, foram selecionadas três plataformas amplamente reconhecidas no meio acadêmico e tecnológico: *IEEE Xplore*, *Scopus* e *United States Patent and Trademark Office (USPTO)*¹. A base *IEEE Xplore* foi utilizada por sua relevância na área de engenharia elétrica e eletrônica, reunindo um grande volume de publicações científicas de alto impacto, especialmente em temas relacionados a sistemas de potência, eletrônica e armazenamento de energia. A *Scopus* foi incorporada à análise por sua ampla cobertura multidisciplinar e por disponibilizar metadados estruturados, incluindo informações detalhadas sobre autores, afiliações, citações e periódicos. Essa base complementa a *IEEE Xplore* ao permitir uma visão mais abrangente da produção científica, ampliando a representatividade dos dados analisados. Por sua vez, o *USPTO* foi incluído com o objetivo de incorporar à análise dados relacionados à produção tecnológica, permitindo a avaliação de tendências também no âmbito das patentes e contribuindo para uma análise integrada entre pesquisa científica e inovação tecnológica.

A estratégia de busca adotada neste trabalho foi baseada na utilização de palavras-chave representativas do tema em análise. A definição dessas palavras-chave constitui uma etapa fundamental do processo, uma vez que está diretamente relacionada à qualidade

¹ Agência norte-americana responsável por patentes e marcas, instituída pelo *Patent Act* de 1790.

e relevância dos dados coletados. Para a seleção dos termos de busca, foi necessário um conhecimento prévio do tema a ser investigado, de modo a identificar conceitos centrais e terminologias frequentemente utilizadas na literatura. A partir dessas palavras-chave, foram construídas consultas utilizando operadores booleanos, como AND e OR, com o objetivo de refinar os resultados e direcionar a busca para conteúdos mais relevantes. A pesquisa realizada seguiu os seguintes parâmetros:

- Scopus: TITLE-ABS-KEY (second life battery) AND (LIMIT-TO (SUBJAREA , "ENGI") OR LIMIT-TO (SUBJAREA , "ENER") OR LIMIT-TO (SUBJAREA , "COMP") OR LIMIT-TO (SUBJAREA , "MATH") OR LIMIT-TO (SUBJAREA , "MATE") OR LIMIT-TO (SUBJAREA , "ENVI") OR LIMIT-TO (SUBJAREA , "CHEM") OR LIMIT-TO (SUBJAREA , "PHYS") OR LIMIT-TO (SUBJAREA , "CENG") OR LIMIT-TO (SUBJAREA , "BUSI") OR LIMIT-TO (SUBJAREA , "ECON"))
- ITEEM Xplore: "All Metadata": "second life"AND "batter*"
- USPTO: "second life battery"

A utilização conjunta dessas bases possibilita uma visão mais abrangente do estado da arte, contemplando tanto a produção científica quanto a tecnológica. Além disso, a escolha dessas plataformas está associada à sua confiabilidade, relevância e ampla utilização em estudos acadêmicos. Neste trabalho, foram considerados tanto artigos científicos quanto patentes, de forma a ampliar o escopo da análise e permitir a identificação de tendências complementares entre pesquisa acadêmica e desenvolvimento tecnológico. Destaca-se que não foi estabelecido um recorte temporal fixo para a análise dos dados. Essa decisão foi tomada com o objetivo de manter a generalidade da metodologia proposta, permitindo sua aplicação a diferentes temas sem restrições pré-definidas de período. Dessa forma, o intervalo temporal analisado é determinado dinamicamente a partir dos dados disponíveis para cada tema investigado.

Com o intuito de manter a consistência da metodologia, optou-se pela utilização das mesmas consultas nas diferentes bases de dados analisadas. Essa padronização permite maior comparabilidade entre os resultados obtidos e contribui para a uniformidade do processo de coleta de dados. Dessa forma, a estratégia de busca foi estruturada de maneira a equilibrar abrangência e especificidade, garantindo a obtenção de um conjunto de dados representativo do tema analisado, sem comprometer a relevância das informações coletadas.

3.3 Normalização e modelagem dos dados de artigos

Após a etapa de coleta dos dados nas diferentes bases selecionadas, foi realizada a etapa de tratamento e integração das informações, com o objetivo de consolidar uma base unificada e consistente para análise. Inicialmente, foi feita a exportação completa dos resultados obtidos em cada base de dados, garantindo a preservação de todos os campos disponíveis. Foram definidos como essenciais os campos relacionados à identificação dos documentos, autores, afiliações institucionais, palavras-chave e informações temporais.

A etapa de normalização e modelagem dos dados consistiu na preparação e padronização das informações provenientes de diferentes bases científicas, com destaque para as plataformas *IEEE Xplore* e *Scopus*. Esse processo foi necessário devido às diferenças estruturais e semânticas existentes entre as bases, que apresentam variações tanto na quantidade de colunas quanto na forma de representação dos dados. Para isso, foi realizada a análise de equivalência entre as colunas das bases, com o objetivo de identificar correspondências entre os campos e garantir a compatibilização dos dados para posterior integração. Observou-se que, embora ambas as bases apresentem informações semelhantes, como autores, afiliações e ano de publicação, a estrutura e a nomenclatura dos campos diferem significativamente, exigindo ajustes prévios.

A decomposição dos dados associada ao uso de identificadores únicos está diretamente relacionada aos princípios de normalização de bancos de dados, especialmente às três primeiras formas normais. Inicialmente, a separação de múltiplos autores em registros individuais atende aos requisitos da Primeira Forma Normal (1FN), que estabelece a atomicidade dos atributos e a eliminação de valores multivalorados em um mesmo campo. Em seguida, a organização das informações de autoria, instituição e país, associadas corretamente a cada registro, contribui para o atendimento da Segunda Forma Normal (2FN), ao reduzir dependências parciais. Por fim, a padronização de dados como nomes de países e instituições está alinhada com a Terceira Forma Normal (3FN), ao minimizar dependências transitivas e redundâncias. Dessa forma, a estrutura adotada garante maior consistência, integridade e confiabilidade dos dados utilizados nas análises.

3.3.1 Nome do autor

Um dos principais desafios encontrados foi a padronização dos nomes dos autores. Na base *IEEE Xplore*, os autores são apresentados no formato “N. Sobrenome”, enquanto no *Scopus* a formatação ocorre como “Sobrenome N.”, podendo ainda incluir múltiplos sobrenomes ou iniciais adicionais, como em “Sobrenome N. S. S.”. Diante disso, foi definido um padrão único de representação no formato “N. Sobrenome”, sendo realizada a transformação dos dados para garantir consistência e evitar duplicidades na identificação dos autores.

3.3.2 País de filiação do autor

A informação do país da instituição, na maioria dos casos, encontra-se posicionada ao final da string tanto na base IEEE quanto no *Scopus*. No entanto, foram identificadas inconsistências, especialmente em registros provenientes dos Estados Unidos, nos quais o campo final apresentava o estado (por exemplo, “CA”) em vez do país. Para resolver esse problema, foi realizada a padronização desses registros, convertendo-os para a nomenclatura unificada “USA”.

3.3.3 Instituição de filiação do autor

Em seguida, foi realizado o tratamento das informações de afiliação dos autores, etapa que apresentou maior complexidade. Embora exista uma estrutura normativa recomendada para a descrição de afiliações — geralmente organizada como “Departamento, Instituição, Cidade, País” —, verificou-se que nem todas as publicações seguem esse padrão, resultando em inconsistências nos dados. Para a identificação de instituições, as strings de afiliação foram inicialmente segmentadas utilizando a vírgula como delimitador, permitindo a separação dos diferentes níveis hierárquicos da informação. Em seguida, foi aplicada uma estratégia baseada na busca por palavras-chave (Tabela 1) com o objetivo de identificar o segmento correspondente à instituição de vínculo principal.

Palavras-chave
univ, politecnico, universidade, institute, company, institution, universiti, center, universiteit, college, foundation

Tabela 1 – Palavras-chave para identificação de instituições

Como exemplo, para uma afiliação no formato “Laboratório de Alta Tensão, Departamento de Engenharia Elétrica, Universidade Federal de Campina Grande, Brasil”, os elementos foram separados e analisados individualmente, sendo selecionado o trecho que continha o termo indicativo de instituição (“Universidade Federal de Campina Grande”). Esse procedimento permitiu isolar a informação institucional de forma mais consistente, mesmo diante da variabilidade dos registros.

Posteriormente, com o objetivo de compatibilizar e padronizar os nomes das instituições identificadas, utilizou-se uma base auxiliar contendo uma lista de universidades e instituições de referência em âmbito internacional. A partir dessa base, foi realizada uma comparação entre os nomes extraídos das afiliações e os nomes presentes na lista auxiliar, buscando identificar correspondências textuais que permitissem substituir diferentes variações de escrita por uma denominação única e normalizada. Essa etapa foi importante para reduzir problemas decorrentes de abreviações, diferenças ortográficas e formas alternativas de apresentação de uma mesma instituição.

Ao final desse processo, foram geradas duas colunas distintas: a primeira, resultante da identificação inicial com base em palavras-chave nas strings de afiliação; e a segunda, proveniente da compatibilização com a base auxiliar de instituições. Para consolidar os resultados, foi criada uma coluna final unificada, na qual se adotou como prioridade a informação obtida por meio da base de dados auxiliar. Nos casos em que essa segunda etapa não retornou correspondência, manteve-se o valor identificado pela estratégia inicial baseada em palavras-chave. Dessa forma, foi possível obter uma normalização mais abrangente e consistente das instituições, minimizando perdas de informação e aumentando a confiabilidade das análises posteriores.

3.3.4 Identificação de inconsistências

Após essa etapa, foi realizada uma filtragem dos registros que não puderam ser corretamente normalizados, incluindo casos em que a afiliação correspondia a empresas, ausência de informação de país ou inconsistências estruturais mais severas. Esses registros foram tratados individualmente ou excluídos da análise, quando necessário, de modo a garantir a integridade do conjunto de dados final. Por fim, foi conduzida a padronização dos nomes dos países, considerando variações como “USA”, “EUA” e abreviações de estados. Todas essas ocorrências foram unificadas em uma única nomenclatura, assegurando consistência nas análises posteriores, especialmente na geração de indicadores geográficos.

Esse processo de normalização foi fundamental para reduzir redundâncias, corrigir inconsistências e viabilizar a integração das bases de dados, conforme os princípios da modelagem de dados. Dessa forma, garantiu-se maior confiabilidade nas análises realizadas no *dashboard*, permitindo uma representação mais precisa das informações científicas e tecnológicas.

3.3.5 Identificação de artigos

Para o tratamento das informações relacionadas aos autores e suas respectivas afiliações, foi adotada uma abordagem baseada na decomposição dos dados. Considerando que um único artigo pode possuir múltiplos autores, vinculados a diferentes instituições e países, foi realizada a separação dessas informações em múltiplas linhas, associadas a um mesmo identificador de artigo. Essa estratégia permitiu uma análise mais detalhada da distribuição geográfica e institucional das publicações.

No entanto, para evitar a contagem indevida de um mesmo artigo múltiplas vezes, foi utilizada a função de contagem distinta (*distinct count*) no *Power BI*, garantindo que cada publicação fosse considerada de forma única nas análises quantitativas. Dessa forma, quando um artigo apresenta autores vinculados a diferentes países, ele é contabilizado uma vez para cada país envolvido, refletindo a colaboração internacional. Por outro

lado, evita-se a supercontagem em casos de múltiplos autores pertencentes ao mesmo país, assegurando a consistência dos resultados. Essa abordagem permite representar de maneira mais fiel a distribuição das contribuições científicas, considerando não apenas a origem do artigo, mas também a diversidade de instituições e países envolvidos na sua produção.

3.4 Método de ranqueamento e seleção de artigos

O processo de seleção e priorização dos documentos científicos constitui uma etapa fundamental na organização do estado da arte, especialmente em abordagens que envolvem grandes volumes de dados provenientes de diferentes bases. Nesse contexto, diversos métodos têm sido propostos na literatura com o objetivo de ranquear publicações de acordo com sua relevância, entre os quais se destaca o *Methodi Ordinatio* [8].

O *Methodi Ordinatio* propõe um índice de relevância baseado na combinação de três fatores principais: número de citações, ano de publicação e fator de impacto do periódico. No entanto, embora amplamente utilizado, esse método apresenta limitações quando aplicado ao contexto deste trabalho. A principal delas está relacionada à dependência do fator de impacto, uma métrica que varia ao longo do tempo e depende de fontes externas, podendo introduzir inconsistências na análise, especialmente em estudos que envolvem múltiplas bases de dados e diferentes áreas do conhecimento.

Diante dessas limitações, optou-se por desenvolver uma adaptação metodológica voltada à construção de um indicador de relevância mais alinhado com a natureza dos dados analisados e com os objetivos deste trabalho. Essa adaptação baseia-se em princípios consolidados da análise de dados e da cientometria, especialmente no que se refere ao tratamento de distribuições assimétricas e à normalização de indicadores bibliométricos. Inicialmente, foi adotado o indicador de citações por ano (*Citations per Year – CPY*), definido como:

$$CPY_i = C_i / (A_{ref} - A_i + 1) \quad (3.1)$$

em que C_i representa o número total de citações do artigo i , A_i corresponde ao ano de publicação e A_{ref} é o ano de referência da análise. A inclusão do termo “+1” evita a indeterminação para artigos recentes e permite a consideração de publicações do ano corrente.

O uso do CPY permite mitigar o viés temporal associado ao acúmulo de citações ao longo do tempo, possibilitando uma comparação mais equilibrada entre artigos de diferentes períodos. No entanto, mesmo após essa correção, a distribuição de citações permanece altamente assimétrica, caracterizada pela presença de poucos artigos com elevado

número de citações e uma grande quantidade de publicações com baixo impacto.

Esse comportamento é amplamente reconhecido na literatura de cientometria e análise de dados, sendo frequentemente modelado por distribuições do tipo *heavy-tailed*. Nesses casos, a utilização de valores brutos pode levar à supervalorização de *outliers* e comprometer a análise comparativa. Para lidar com esse problema, é comum a aplicação de transformações não lineares, como a transformação logarítmica, com o objetivo de reduzir a dispersão dos dados e tornar a distribuição mais adequada para análise. Dessa forma, foi definido o seguinte indicador de relevância:

$$Score_i = \ln(1 + CPY_i) \quad (3.2)$$

A aplicação da transformação $\ln(1 + x)$ é amplamente utilizada em análises de dados para variáveis não negativas e assimétricas, pois permite reduzir a influência de valores extremos sem eliminar a informação relativa entre os dados. Além disso, a adição de 1 garante que o indicador seja definido mesmo para valores nulos de CPY. Essa abordagem permite equilibrar dois aspectos fundamentais: a correção temporal das citações e a redução da influência de artigos extremamente citados, resultando em um ranqueamento mais estável e representativo do conjunto de dados analisado. Adicionalmente, para fins de visualização no *dashboard*, o *score* pode ser normalizado conforme a equação a seguir. A normalização não altera o ranqueamento dos documentos, sendo utilizada apenas para facilitar a interpretação visual dos resultados.

$$CPY_{norm} = \frac{\ln(1 + CPY)}{\max(\ln(1 + CPY))} \quad (3.3)$$

No que se refere às patentes, optou-se por não utilizar um indicador de ranqueamento baseado em citações ou métricas equivalentes. Essa decisão está relacionada à natureza distinta dos documentos de patente em relação aos artigos científicos, uma vez que métricas de impacto bibliométrico não são diretamente comparáveis nesse contexto. Assim, a seleção e organização das patentes foram realizadas com base em critérios de classificação tecnológica, utilizando códigos do sistema *Cooperative Patent Classification* (CPC). Inicialmente, foram identificados os códigos mais aderentes ao tema em análise, a partir de uma investigação prévia do domínio tecnológico. Em seguida, os documentos foram filtrados de acordo com esses códigos, garantindo maior precisão na delimitação do escopo da análise.

Essa abordagem está alinhada com práticas consolidadas em estudos de prospecção tecnológica, nos quais os sistemas de classificação patentária são utilizados como principal ferramenta para identificação de áreas tecnológicas, tendências de inovação e especialização industrial. Dessa forma, a metodologia adotada neste trabalho diferencia claramente o tratamento dos dados científicos e tecnológicos, utilizando métricas bibliométricas para

artigos e critérios de classificação para patentes, de modo a respeitar as especificidades de cada tipo de documento e garantir maior consistência na análise do estado da arte.

3.5 Normalização e modelagem de patentes

No que se refere aos dados de patentes obtidos a partir do *United States Patent and Trademark Office* (USPTO), não foi necessária a realização de um processo extensivo de limpeza e integração de bases, uma vez que os dados foram coletados a partir de uma única fonte. Ainda assim, foram realizadas etapas de tratamento e organização com o objetivo de adequar os dados à análise proposta. A principal transformação realizada consistiu na decomposição dos códigos de classificação tecnológica das patentes, especificamente os códigos CPC (Cooperative Patent Classification). Considerando que uma única patente pode estar associada a múltiplos códigos CPC, esses registros foram separados em múltiplas linhas, mantendo-se a associação com o identificador da patente. Em seguida, foram extraídos os quatro primeiros caracteres de cada código, correspondentes à classe principal da tecnologia, permitindo uma categorização mais geral e adequada à análise temática.

Com base na temática do estudo, foram aplicados filtros adicionais para remoção de códigos que não apresentavam aderência ao escopo da pesquisa, contribuindo para o refinamento dos resultados. Adicionalmente, observou-se uma limitação na construção das consultas na base do USPTO, na qual a busca por expressões compostas, como “second life battery”, não é interpretada como um termo exato, mas sim como a combinação independente das palavras “second”, “life” e “battery”. Essa característica resultou na recuperação de documentos não diretamente relacionados ao tema, tornando necessária a realização de uma filtragem manual complementar para exclusão de patentes irrelevantes.

Outro aspecto relevante no tratamento dos dados de patentes refere-se ao conceito de família de patentes. Uma família de patentes corresponde a um conjunto de documentos que protegem uma mesma invenção, podendo incluir depósitos realizados em diferentes países ou variações do mesmo pedido ao longo do tempo. Dessa forma, múltiplos registros podem representar, na prática, uma única inovação tecnológica. Considerar essa característica é fundamental para evitar a superestimação dos resultados, uma vez que a contagem de documentos individuais pode levar à duplicação de uma mesma tecnologia na análise. Dessa forma, o tratamento dos dados de patentes envolveu não apenas a organização dos códigos tecnológicos, mas também a compreensão da estrutura dos registros patentários, garantindo maior precisão na interpretação dos resultados e alinhamento com os objetivos do estudo.

3.6 Análise dos dados

A etapa de análise dos dados, no contexto deste trabalho, não foi conduzida de forma isolada em relação à etapa de visualização, mas sim integrada ao processo de construção do *dashboard*. Essa abordagem está alinhada aos princípios de *Visual Analytics*, nos quais a análise e a visualização dos dados ocorrem de maneira simultânea e interativa, permitindo uma exploração mais dinâmica das informações. Dessa forma, a análise foi realizada diretamente no ambiente do *Power BI*, por meio da construção de visualizações que possibilitam a identificação de padrões, tendências e relações relevantes no conjunto de dados. Essa integração entre análise e visualização permite que o processo analítico seja continuamente refinado à medida que diferentes perspectivas são exploradas.

As principais análises realizadas concentram-se em métricas quantitativas relacionadas à produção científica, incluindo a identificação dos autores, instituições e países com maior volume de publicações. Para isso, foram utilizados gráficos de barras que apresentam o ranking dos principais contribuintes, com foco nos dez elementos mais relevantes em cada categoria, estratégia que permite uma visualização mais clara e objetiva das informações. Além disso, foi incorporada uma análise geográfica por meio de um mapa-múndi, no qual a distribuição das publicações é representada visualmente com base na localização das afiliações dos autores. A utilização de escala de cores permite evidenciar a intensidade da produção científica em diferentes regiões, facilitando a identificação de polos de pesquisa e concentração de conhecimento.

Outro aspecto relevante da análise é a possibilidade de interação com os dados. O *dashboard* permite a aplicação de filtros, como a seleção de períodos específicos, possibilitando que a análise seja adaptada dinamicamente conforme o interesse do usuário. Essa funcionalidade torna possível a investigação de tendências temporais e a comparação entre diferentes intervalos de tempo. Adicionalmente, indicadores agregados, como o número total de publicações dentro dos filtros selecionados, são apresentados de forma destacada, contribuindo para uma visão geral do conjunto de dados analisado.

A partir dessas análises, torna-se possível identificar, para um determinado tema, os principais agentes envolvidos na produção científica, a distribuição geográfica do conhecimento e a concentração de publicações ao longo do tempo. Essas informações são fundamentais para a compreensão do estado da arte e para a identificação de tendências e lacunas de pesquisa.

3.7 Construção do *dashboard*

O *dashboard* desenvolvido neste trabalho foi construído utilizando a ferramenta Microsoft Power BI, escolhida devido à sua capacidade de integração de dados, modelagem

e criação de visualizações interativas em um único ambiente. A estrutura foi projetada com o objetivo de facilitar a exploração dos dados e permitir uma análise intuitiva por parte do usuário. Para isso, foram implementadas diferentes visualizações, cada uma voltada à análise de aspectos específicos da produção científica. Entre os principais elementos do *dashboard*, destacam-se a lista apresentada na Tabela 2.

Elemento	Objetivo
Mapa geográfico (mapa-múndi)	representar a distribuição das publicações por país, com aplicação de escala de cores para indicar a intensidade da produção científica
Gráficos de barras	apresentar o ranking dos principais autores, instituições e países, com foco nos dez maiores contribuidores em cada categoria
Indicadores agregados	apresentam o total de publicações considerando os filtros aplicados
Filtros interativos	Seleção de períodos de análise e explorar diferentes recortes dos dados

Tabela 2 – Listagem de elementos do *dashboard*

A escolha por limitar a visualização aos principais elementos (top 10) está relacionada à necessidade de evitar sobrecarga visual e facilitar a interpretação dos resultados, mantendo o foco nos agentes mais relevantes dentro do contexto analisado. O *dashboard* desenvolvido incorpora mecanismos de filtragem interativa que permitem ao usuário explorar os dados de forma dinâmica e personalizada. Entre os principais recursos implementados, destacam-se os filtros por período temporal e por temática, os quais possibilitam a adaptação da análise conforme o objetivo do estudo.

O filtro temporal permite restringir os resultados a intervalos específicos de anos de publicação, possibilitando a análise da evolução das pesquisas ao longo do tempo. Essa funcionalidade é particularmente relevante para identificar tendências, períodos de maior produção científica e o comportamento recente das publicações relacionadas ao tema investigado. Adicionalmente, foi implementado um filtro temático, baseado na classificação dos artigos a partir de palavras-chave identificadas nos títulos e resumos. Esse filtro permite ao usuário selecionar uma ou múltiplas categorias de interesse, como degradação, diagnóstico, aplicações e segunda vida de baterias, possibilitando uma análise mais direcionada e aprofundada de subconjuntos específicos da base de dados.

A combinação desses filtros proporciona maior flexibilidade na exploração das informações, permitindo que diferentes perspectivas analíticas sejam adotadas de acordo com o foco do usuário. Dessa forma, o *dashboard* não apenas apresenta os dados de forma estruturada, mas também atua como uma ferramenta interativa de apoio à análise do estado da arte, facilitando a identificação de padrões, lacunas e oportunidades de pesquisa.

4 Resultados

Neste capítulo são apresentados e discutidos os resultados obtidos a partir da aplicação da metodologia proposta. A análise é realizada com base nas visualizações e indicadores disponíveis no *dashboard* desenvolvido, permitindo a identificação de padrões, tendências e atores relevantes no contexto estudado.

4.1 Definição de métricas

Na etapa de análise dos indicadores bibliométricos, foi realizada uma comparação entre o índice de citações por ano (CPY) em sua forma original e o CPY após aplicação de normalização. Para isso, foram gerados dois gráficos de ranqueamento dos artigos, permitindo a análise visual do comportamento de cada métrica. Observou-se que o CPY em sua forma original (Figura 2) apresenta uma distribuição altamente desigual, caracterizada por grandes discrepâncias entre os artigos. Nesse caso, poucos trabalhos concentram valores significativamente elevados, enquanto a maioria apresenta valores substancialmente inferiores. Essa diferença acentuada resulta em uma ordenação pouco homogênea, dificultando a comparação entre os artigos, especialmente nos níveis intermediários do ranking.

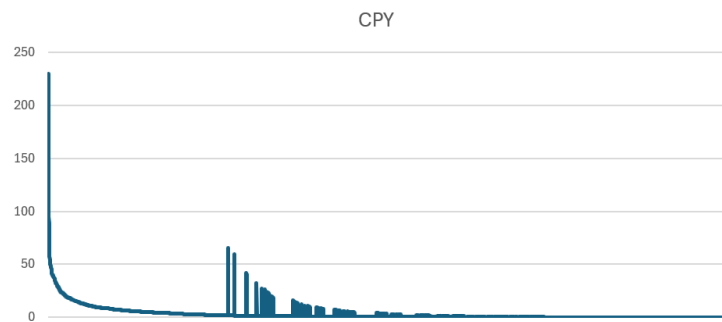


Figura 2 – Distribuição de score utilizando apenas CPY. Fonte: autoria própria

Por outro lado, ao aplicar a normalização do CPY, verificou-se uma distribuição mais equilibrada dos valores. O gráfico correspondente (Figura 3) apresenta uma curva decrescente mais contínua e suave, reduzindo o impacto de valores extremos e proporcionando uma melhor diferenciação entre os artigos ao longo de todo o ranking. Essa característica torna a análise mais interpretável e consistente, especialmente para fins de priorização e identificação de relevância relativa. Portanto, os resultados indicam que a utilização do CPY normalizado contribui para uma representação mais estável e informativa do impacto dos artigos, evitando distorções causadas por outliers e permitindo uma análise comparativa mais robusta dentro do conjunto de dados.

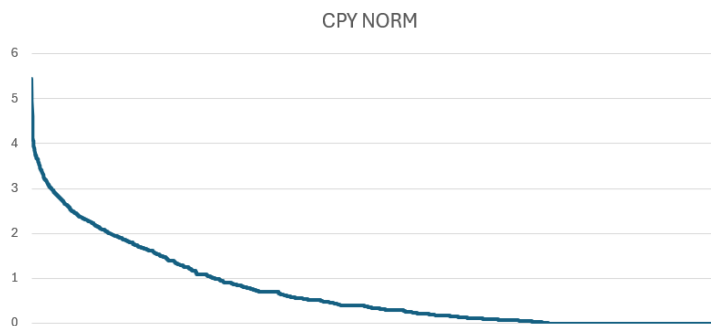


Figura 3 – Distribuição de score utilizando CPY normalizado. Fonte: autoria própria

4.2 Normalização de dados

No processo de consolidação das bases de dados, observou-se que a junção inicial resultou em um total de 4.155 registros. Entretanto, após a etapa de remoção de duplicidades, realizada com base no título e DOI do artigo, esse número foi reduzido para 3.103 registros únicos. Essa etapa foi fundamental para evitar a super-representação de determinados trabalhos e garantir maior confiabilidade nas análises subsequentes.

Já na etapa de normalização das instituições, verificou-se que nem todas as entradas apresentavam informações suficientes para a identificação do vínculo institucional dos autores, mesmo após a aplicação das estratégias de extração e compatibilização descritas anteriormente. Diante disso, optou-se pela remoção das linhas nas quais não foi possível identificar, de forma confiável, ao menos a instituição de afiliação.

Essa decisão foi adotada com o objetivo de preservar a consistência dos resultados, uma vez que a inclusão de registros incompletos poderia comprometer análises posteriores, especialmente aquelas relacionadas à distribuição geográfica, ranqueamento institucional e identificação de atores relevantes. Assim, ao garantir que todas as entradas consideradas possuam informação institucional válida, assegura-se maior robustez metodológica e maior confiabilidade na interpretação dos dados.

4.3 Telas construídas

O *dashboard* desenvolvido tem como objetivo principal facilitar a análise e interpretação dos dados coletados ao longo da pesquisa, permitindo uma visualização integrada das informações bibliométricas e tecnológicas. A construção das telas foi orientada pelos princípios de *Visual Analytics*, priorizando clareza, interatividade e suporte à tomada de decisão. Ressalta-se que as imagens apresentadas nesta seção correspondem a uma versão preliminar da base de dados, contendo apenas os registros provenientes da base *IEEE Xplore*, anteriores à incorporação dos dados da *Scopus*.

O *dashboard* desenvolvido foi estruturado em quatro telas principais, organizadas

de forma a facilitar a navegação e a exploração das informações pelo usuário. A tela inicial (Figura 4), denominada menu, tem como objetivo central servir como ponto de acesso às demais seções da ferramenta. Nela, o usuário pode visualizar uma visão geral da proposta do sistema, bem como acessar rapidamente as diferentes análises disponíveis. Essa organização contribui para uma navegação intuitiva e orientada, permitindo que o usuário selecione a área de interesse de maneira clara e eficiente.

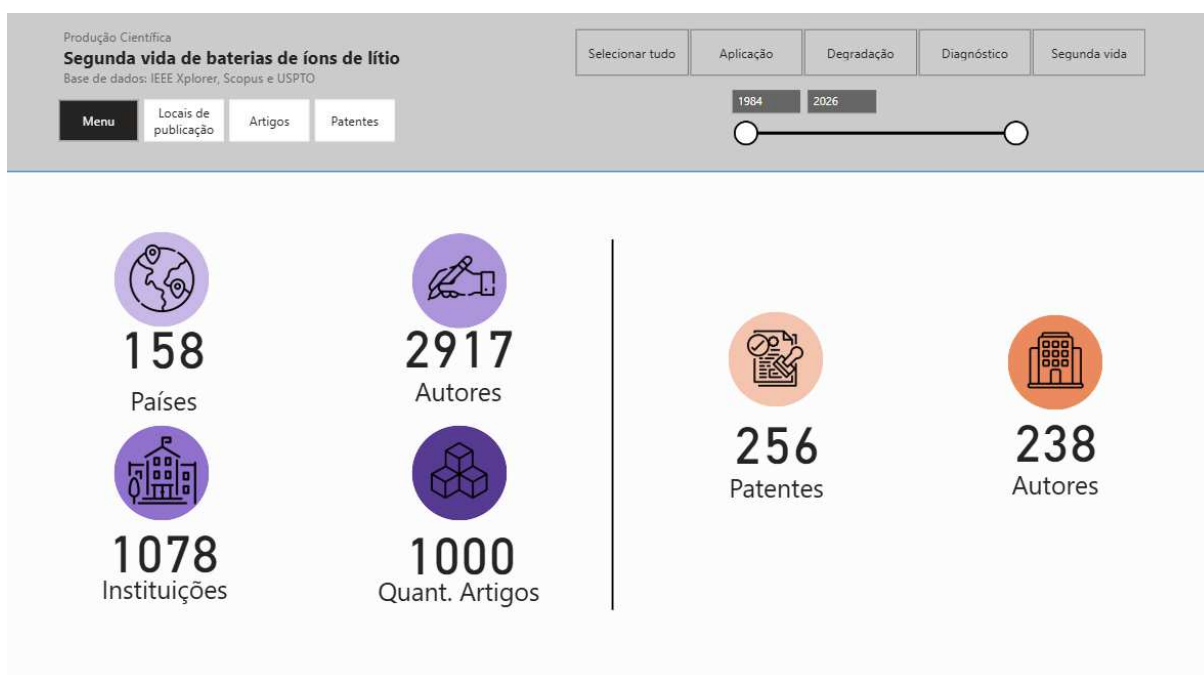


Figura 4 – Tela principal. Fonte: Autoria própria.

A tela de apresentação das patentes (Figura 5) foi concebida para exibir informações relacionadas às tecnologias protegidas por propriedade intelectual. Nessa seção, são apresentadas análises quantitativas e categóricas, como distribuição por códigos CPC e identificação de áreas tecnológicas predominantes. Essa visualização permite compreender o direcionamento tecnológico das inovações associadas ao tema estudado, contribuindo para a identificação de tendências e oportunidades no desenvolvimento de soluções.

A tela dedicada à apresentação dos artigos científicos (Figura 6) tem como objetivo principal disponibilizar, de forma estruturada, os dados dos artigos selecionados na base analisada. Nessa seção, são exibidas informações como título, ano de publicação e demais metadados relevantes, organizadas em formato tabular. Além disso, a tela incorpora os filtros previamente definidos, como período temporal e temática, permitindo que o usuário visualize subconjuntos específicos de interesse. Como diferencial, cada registro contém o link de acesso direto ao artigo, possibilitando que o usuário consulte a fonte original de maneira rápida e prática, o que torna a ferramenta não apenas analítica, mas também um apoio à navegação e aprofundamento na literatura científica.

Por fim, a tela de autoria dos artigos (Figura 7) tem como objetivo identificar os

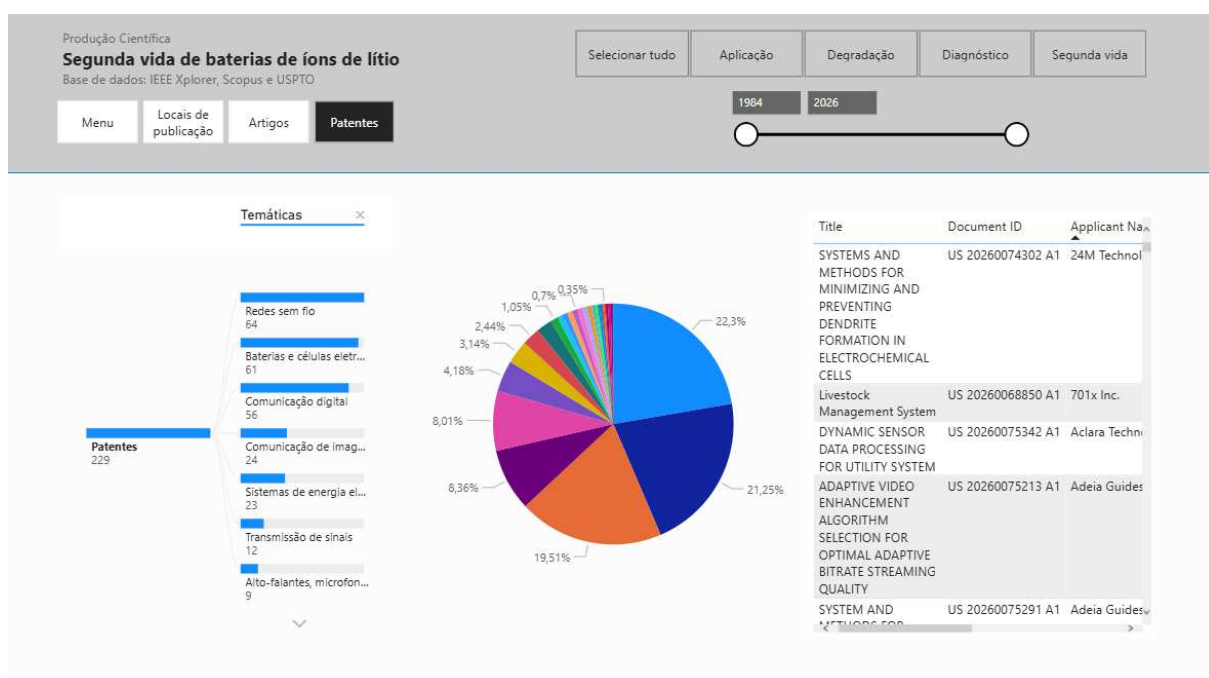


Figura 5 – Tela de patentes. Fonte: Autoria própria.

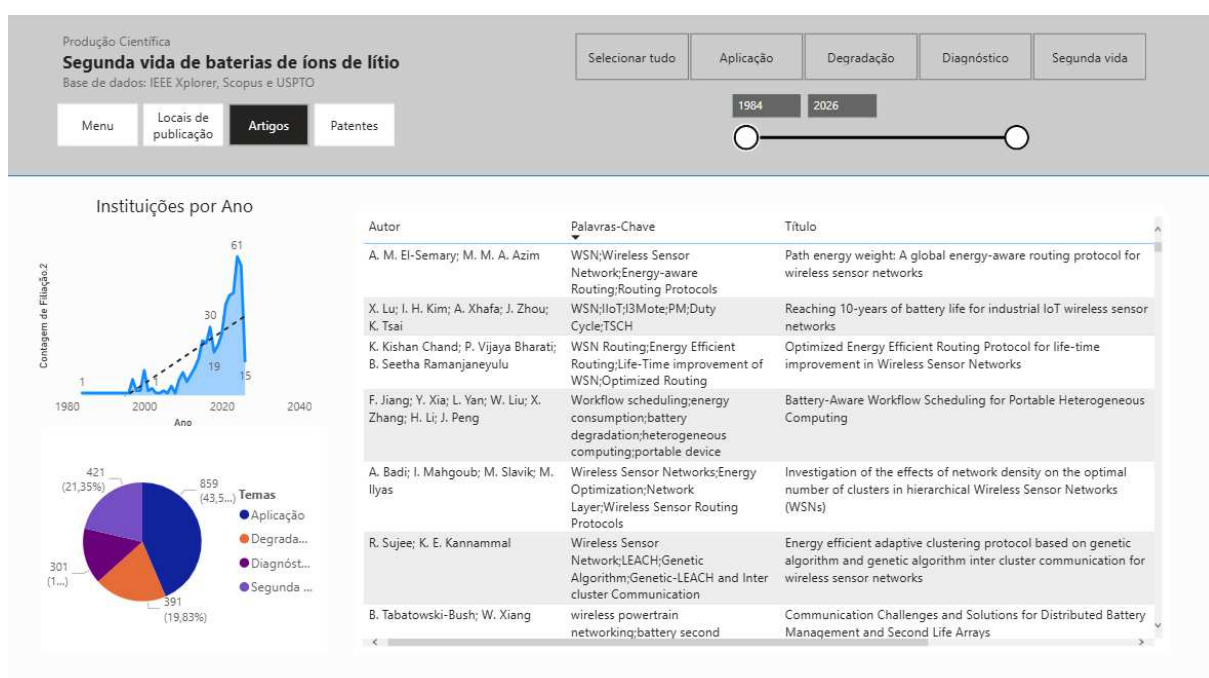


Figura 6 – Tela de artigos. Fonte: Autoria própria.

principais pesquisadores e instituições atuantes na área. Nessa seção, são apresentados rankings de autores com base na relevância de suas publicações, bem como análises relacionadas à distribuição geográfica e institucional das pesquisas. Essa abordagem permite compreender a dinâmica da produção científica, destacando os principais atores envolvidos e contribuindo para a identificação de referências relevantes no campo de estudo.

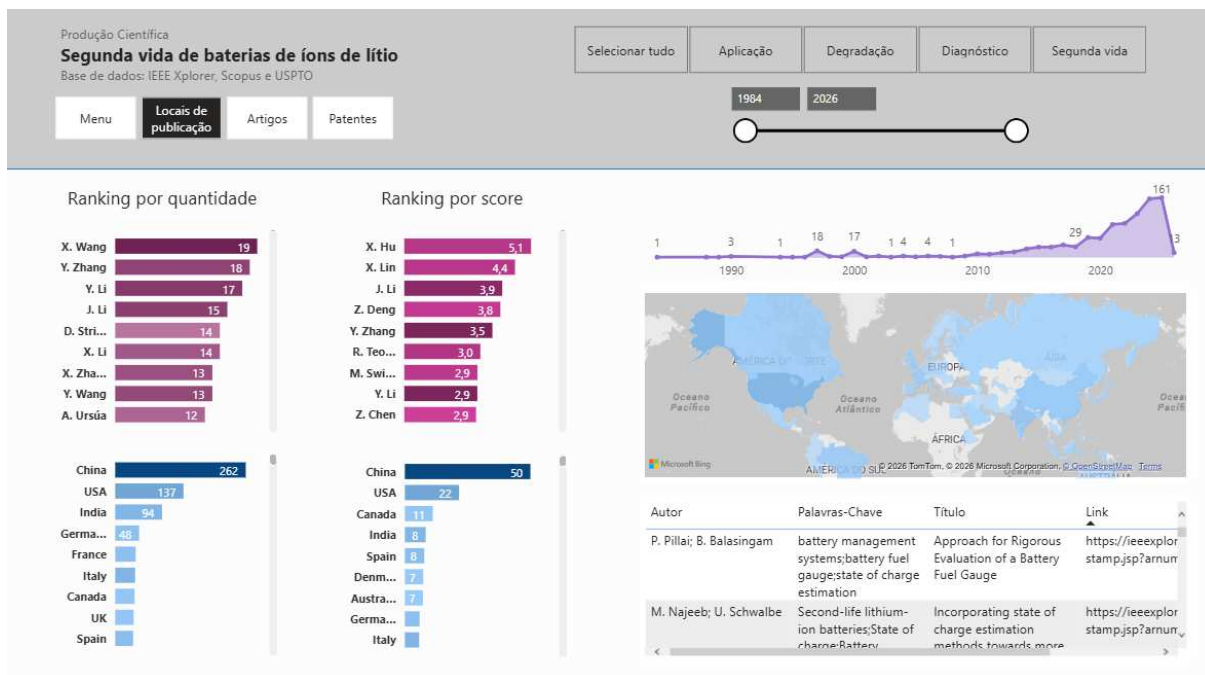


Figura 7 – Tela de distribuição geográfica. Fonte: Autoria própria.

4.4 Filtros

Como mencionado previamente, o *dashboard* desenvolvido incorpora filtros interativos que permitem ao usuário modificar dinamicamente os resultados apresentados em cada tela, proporcionando uma análise personalizada conforme o objetivo da investigação. Esses filtros atuam diretamente sobre a base de dados, de modo que qualquer alteração realizada pelo usuário reflete imediatamente nos gráficos, tabelas e indicadores exibidos. Dessa forma, a ferramenta possibilita diferentes perspectivas de análise a partir de um mesmo conjunto de dados, ampliando sua aplicabilidade no estudo do estado da arte.

Um exemplo de aplicação consiste na utilização do filtro temporal, no qual o usuário pode selecionar um intervalo específico de anos de publicação. Ao restringir a análise para o período de 2010 a 2020 (Figura 8 e 9), o *dashboard* passa a exibir apenas os artigos e patentes publicados dentro desse intervalo, ajustando automaticamente todos os resultados apresentados nas telas. Essa funcionalidade permite investigar, por exemplo, a evolução da produção científica em uma década específica, identificar tendências históricas e analisar o comportamento das publicações em períodos distintos. Os resultados dessa filtragem são apresentados nos prints das telas, evidenciando a alteração dos dados exibidos após a aplicação do filtro.

De forma complementar, o filtro temático possibilita a seleção de categorias específicas associadas aos conteúdos analisados, como degradação ou diagnóstico. Ao aplicar um tema específico (Figura 10), o *dashboard* exibe exclusivamente os registros relacionados àquela categoria, refinando a análise para um domínio mais direcionado. Essa funciona-

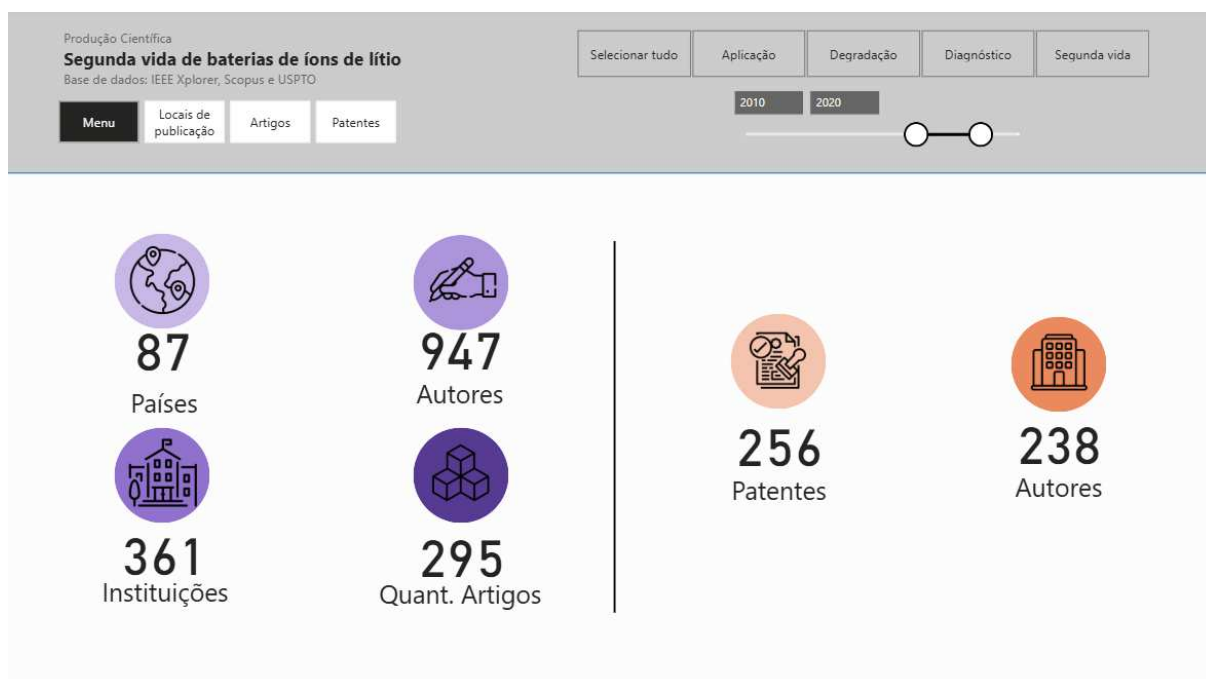


Figura 8 – Resultado de implementação de filtros: Tela inicial. Fonte: Autoria própria.

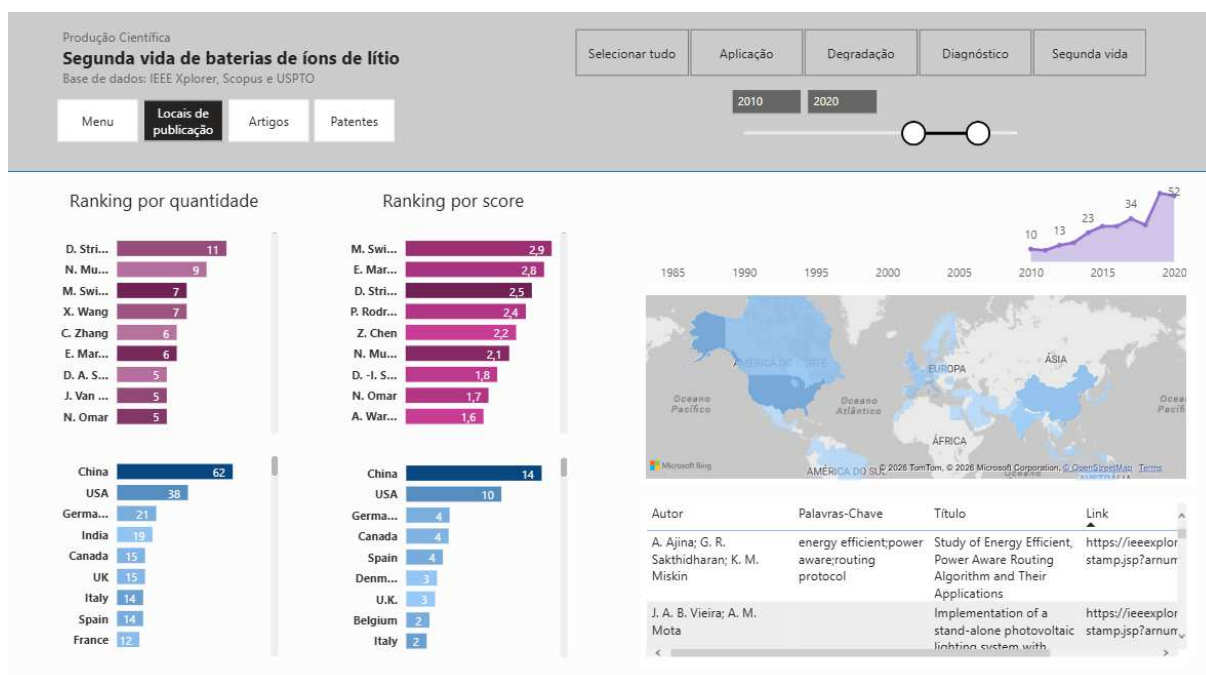


Figura 9 – Resultado de implementação de filtros: Autoria de artigos. Fonte: Autoria própria.

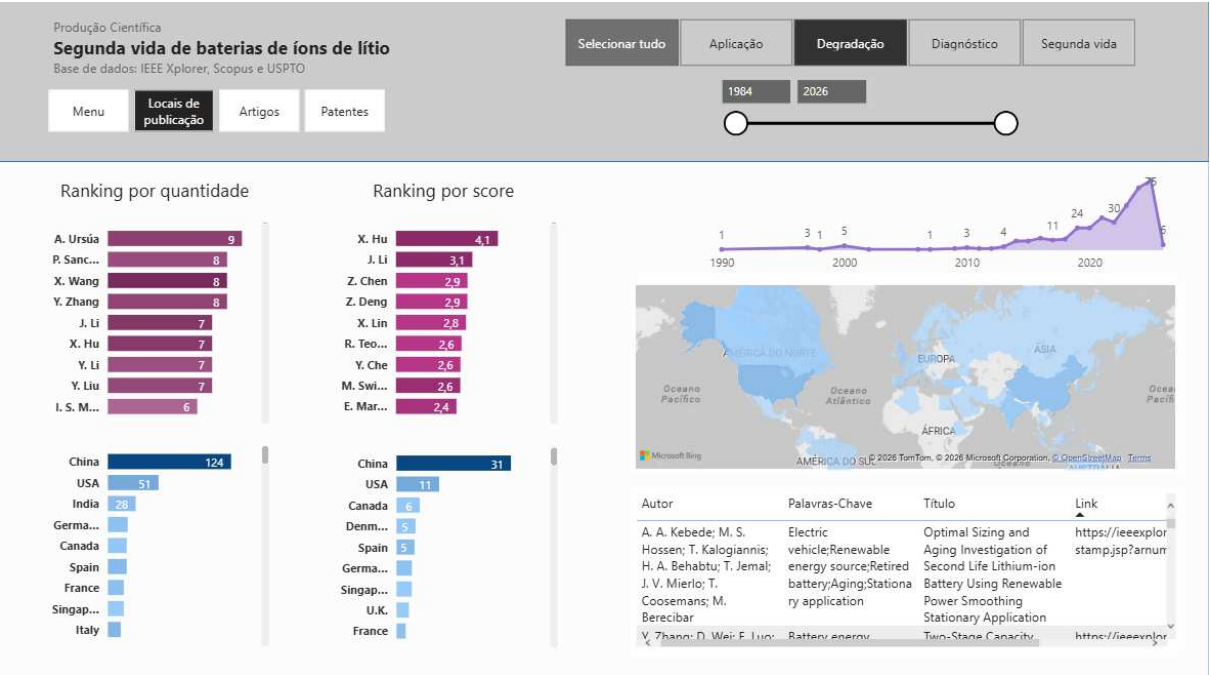


Figura 10 – Resultado de implementação de filtros: Autoria de artigos. Fonte: Autoria própria.

lidade é especialmente útil para investigações focadas, permitindo ao usuário explorar subconjuntos relevantes da base de dados e identificar padrões específicos dentro de uma temática. Assim como no caso do filtro temporal, os efeitos dessa seleção são apresentados nos prints das telas, demonstrando a capacidade da ferramenta de adaptar os resultados conforme o interesse do usuário.

Dessa maneira, a utilização de filtros interativos confere maior flexibilidade e profundidade à análise, permitindo que o *dashboard* funcione não apenas como um meio de visualização, mas também como uma ferramenta ativa de exploração de dados, na qual diferentes cenários podem ser avaliados de forma rápida e eficiente.

5 Considerações Finais

O presente trabalho apresentou o desenvolvimento de um *dashboard* interativo voltada à análise do estado da arte, integrando dados provenientes de bases científicas e tecnológicas. A utilização de técnicas de tratamento de dados e *Visual Analytics* permitiu a organização e interpretação eficiente de um grande volume de informações, contribuindo para a identificação de padrões, tendências e atores relevantes na área estudada.

A adoção de indicadores bibliométricos, como o CPY normalizado, possibilitou o ranqueamento dos artigos com base em critérios objetivos, enquanto a implementação de filtros interativos ampliou a flexibilidade da análise, permitindo diferentes abordagens conforme o interesse do usuário. Dessa forma, a ferramenta desenvolvida demonstrou-se eficaz como suporte à exploração e compreensão da literatura científica.

Por fim, destaca-se que o *dashboard* pode ser adaptado para outras áreas do conhecimento, configurando-se como uma solução versátil para análise de dados em estudos de estado da arte. Como trabalhos futuros, sugere-se a incorporação de novas métricas, integração com outras bases de dados e aprimoramento das funcionalidades interativas, visando ampliar ainda mais o potencial analítico da ferramenta.

Referências

- 1 Linnenluecke, M., Marrone, M., & Singh, A. (2019). Conducting systematic literature reviews and bibliometric analyses. *Australian Journal of Management*, 45, 175 - 194. <https://doi.org/10.1177/0312896219877678>.
- 2 Font, C., Siqueira, H., Neto, J., Santos, J., Stevan, S., Converti, A., & Corrêa, F. (2023). Second Life of Lithium-Ion Batteries of Electric Vehicles: A Short Review and P
- 3 KEIM, D. et al. Visual Analytics: Definition, Process, and Challenges. *Lecture Notes in Computer Science*, p. 154–175, 2008.
- 4 Barry, E., Merkebu, J., & Varpio, L. (2022). State-of-the-art literature review methodology: A six-step approach for knowledge synthesis. *Perspectives on Medical Education*, 11, 281 - 288. <https://doi.org/10.1007/s40037-022-00725-9>.
- 5 Kabeyi, M., & Olanrewaju, O. (2022). Sustainable Energy Transition for Renewable and Low Carbon Grid Electricity Generation and Supply. , 9. <https://doi.org/10.3389/fenrg.2021.743114>.
- 6 Pham, B., Bagheri, E., Rios, P., Pourmasoumi, A., Robson, R., Hwee, J., Isaranuwatthai, W., Darvesh, N., Page, M., & Tricco, A. (2018). Improving the conduct of systematic reviews: a process mining perspective.. *Journal of clinical epidemiology*, 103, 101-111 . <https://doi.org/10.1016/j.jclinepi.2018.06.011>.
- 7 Soufan, A., Ruthven, I., & Azzopardi, L. (2024). Uncharted Territory: Understanding Exploratory Search Behaviours in Literature Reviews. *Proceedings of the 2024 Conference on Human Information Interaction and Retrieval*. <https://doi.org/10.1145/3627508.3638334>. erspectives. Energies. <https://doi.org/10.3390/en16020953>.
- 8 PAGANI, R. N.; KOVALESKI, J. L.; RESENDE, L. M. Methodi Ordinatio: a proposed methodology to select and rank relevant scientific papers encompassing the impact factor, number of citation, and year of publication. *Scientometrics*, v. 105, n. 3, p. 2109–2135, 12 set. 2015.
- 9 Hu, X., Deng, X., Wang, F., Deng, Z., Lin, X., Teodorescu, R., & Pecht, M. (2022). A Review of Second-Life Lithium-Ion Batteries for Stationary Energy Storage Applications. *Proceedings of the IEEE*, 110, 735-753. <https://doi.org/10.1109/jproc.2022.3175614>.