

Raspagem de Dados no INPI: um recurso para a análise de informações de patentes

Web Scraping at INPI: a resource for patent information analysis

Ricardo Pimentel Nogueira¹, Antonio Roberto Giriboni Monteiro¹

¹Universidade Estadual de Maringá, Maringá, PR, Brasil

Resumo

A recuperação manual de grandes volumes de informações representa desafios consideráveis em termos de tempo e precisão, aumentando o risco de falhas. Nesse contexto, técnicas automatizadas como a raspagem de dados surgem como alternativas eficientes de apoio à prospecção tecnológica, permitindo a extração precisa e rápida de informações em bases de dados *on-line* de acesso aberto. Este estudo, que é recorte de pesquisa de mestrado profissional, apresenta uma metodologia construída por meio da técnica de raspagem de dados desenvolvida em linguagem de programação *Python*, aplicada ao portal de busca de patentes do Instituto Nacional da Propriedade Industrial (INPI) com apoio do *Power BI* institucional. O método aplicado demonstrou sua eficácia ao organizar e estruturar um grande volume de dados de patentes, proporcionando uma base sólida para futuras pesquisas em prospecção tecnológica. A técnica mostrou-se especialmente relevante para estudos patentométricos, acelerando a recuperação de informações para análises em áreas emergentes e de importância global.

Palavras-chave: Prospecção Tecnológica; Raspagem de Dados; Patentes.

Áreas Tecnológicas: Prospecção Tecnológica. Programação.

Abstract

The manual retrieval of large volumes of information presents significant challenges in terms of time and accuracy, increasing the risk of errors. In this context, automated techniques such as data scraping emerge as efficient alternatives to support technological prospecting, enabling precise and rapid extraction of information from open-access online databases. This study, derived from a professional master's degree research project, presents a methodology built using a data scraping technique developed in Python programming language, applied to the patent search portal of the National Institute of Industrial Property (INPI), with the institutional support of Power BI. The applied method demonstrated its effectiveness in organizing and structuring large volumes of patent data, providing a solid foundation for future research in technological prospecting. The technique proved especially relevant for patentometric studies, expediting the retrieval of information for analyses in emerging areas of global importance.

Keywords: Technology Foresight; Web Scraping; Patents.

1 Introdução

A coleta de um grande volume de informações apresenta desafios significativos em termos de tempo e precisão, especialmente quando realizada manualmente dado a dado. Esse processo, além de demandar um esforço considerável, aumenta exponencialmente o risco de falhas e de inconsistências afetando a confiabilidade dos dados coletados devido à repetitividade e à complexidade envolvidas. Diante disso, técnicas de automação como a raspagem de dados (*web scraping*) surgem como alternativas promissoras, permitindo a extração eficiente e precisa de grandes quantidades de dados diretamente de fontes *on-line*, reduzindo o tempo de coleta e minimizando falhas.

A prospecção tecnológica constitui estratégia essencial para acompanhar o avanço de inovações e direcionar esforços de pesquisa e desenvolvimento em áreas emergentes. No contexto de estudo de patentes, ela assume um papel ainda mais relevante, visto que alguns tipos de inovações, em especial as patentes verdes, são fundamentais para enfrentar desafios globais como mudanças climáticas, eficiência energética e conservação de recursos naturais. No entanto, um dos principais obstáculos para a realização de estudos aprofundados sobre patentes é a dificuldade de acessar e de organizar grandes volumes de dados de forma eficiente.

Nesse cenário, o presente estudo é um recorte de uma pesquisa de mestrado profissional e objetiva demonstrar a concepção de um método para a prospecção de dados que otimize e acelere a recuperação e análise de informações específicas disponíveis no portal eletrônico do Instituto Nacional da Propriedade Industrial (INPI). Por meio da técnica de raspagem de dados, foi possível extrair as informações do vasto repositório de registros de patentes verdes disponibilizadas pela ferramenta *Power BI* do INPI, na qual estão disponíveis dados estruturados de patentes de forma interativa, mas que necessitavam de uma abordagem automatizada para coleta de dados adicionais em larga escala.

Ademais, a técnica de raspagem de dados mostrou-se promissora ao lidar com o volume e a complexidade dos dados, acelerando significativamente o processo de obtenção de informações em detrimento dos tradicionais métodos manuais de coleta e prospecção de dados. Essa eficiência possibilita inclusive que estudos patentométricos possam ser conduzidos com maior agilidade e precisão. Assim, a raspagem de dados torna-se não apenas uma solução viável, mas uma ferramenta essencial para o avanço da pesquisa em prospecção tecnológica, especialmente em áreas de crescente importância como as patentes verdes.

Ao aplicar com sucesso a raspagem de dados para a coleta de registros de patentes verdes do INPI, pôde-se demonstrar o impacto dessa técnica na aceleração de estudos patentométricos, organizando de forma estruturada

um grande volume de informações, facilitando o tratamento dos dados e fornecendo uma base sólida para futuras investigações que dependam de grandes volumes de dados.

2 Metodologia

Devido às suas vantagens, entre elas, a capacidade de analisar grandes conjuntos de dados de forma eficiente, a abordagem quantitativa foi escolhida para este estudo devido às suas vantagens. Tal abordagem fundamenta-se na aplicação de técnicas de interpretação de informações com o objetivo de identificar padrões, relações e tendências, fornecendo uma base sólida para a pesquisa científica, orientando a tomada de decisões baseadas em evidências (Hair *et al.*, 2010; Kline, 2016).

A abordagem quantitativa possui, entretanto, limitações como a dificuldade de capturar a complexidade e a subjetividade de certos fenômenos e a necessidade de pressupor que os dados sigam distribuições específicas, o que pode não ser observado em todas as situações práticas. Em suma, a metodologia baseada em análise quantitativa oferece uma abordagem robusta e precisa para compreender certos fenômenos, auxiliando na obtenção de dados relevantes para análises.

A presente pesquisa é dispensada de submissão ao Comitê de Ética em Pesquisa, com base no artigo 1º, parágrafo único, incisos II, III, VI e VII, da Resolução n. 510/2016, da Comissão Nacional de Ética em Pesquisa (CONEP) (Brasil, 2016).

2.1 Utilização de Dados do *Power BI* do INPI

A utilização de ferramentas de análise de dados tem se tornado fundamental para a obtenção de dados significativos em um cenário cada vez mais orientado pela informação. Nesse contexto, o *Power BI* representa uma solução abrangente, permitindo a visualização e a análise de dados de maneira eficaz (Leão *et al.*, 2023). No Instituto Nacional da Propriedade Industrial (INPI), o *Power BI* foi adotado como uma ferramenta para disponibilizar dados sobre o trâmite prioritário de patentes, possibilitando a realização de análises estratégicas.

Embora seja uma ferramenta valiosa para análise de dados, pode não abranger todas as funcionalidades desejadas, especialmente em relação à facilidade de *download* de dados e flexibilidade de análise. O processo de extração de dados foi realizado durante o ano de 2023, com a consulta de trâmites prioritários de patentes verdes conduzida por meio de etapas específicas. A metodologia consiste na entrada no *site* do INPI e no acesso à página em que o *Power BI* está disponibilizado, conforme apresentado na Figura 1.

Os dados foram acessados por meio do *Power BI*, no qual constam informações sobre o exame técnico dos pedidos de patente prioritários (INPI, 2016). As etapas subsequentes incluem a filtragem das patentes em um intervalo temporal específico, seleção de modalidades de trâmite prioritário (INPI, 2020) (Passo 2) e filtragem por tecnologia verde (Passo 3), conforme apresentado na Figura 2.

O processo prosseguiu com seleção dos dados por data crescente de depósito. A seleção e a cópia dos dados durante essa etapa foram realizadas manualmente, pois,

para executar o código de raspagem de dados, é necessário possuir os números de registros de patentes a serem pesquisadas.

2.2 O Desenvolvimento do Programa

O código desenvolvido constituiu um programa abrangente para realizar a raspagem de dados do *site* “busca.inpi.gov.br,” (Figura 3 e Figura 4) mantido pelo Instituto Nacional da Propriedade Industrial (INPI, 2022).

Figura 1 – *Power BI* do INPI: recursos para filtragem de dados



Fonte: INPI (2022)

Figura 2 – Passos da filtragem de dados obtidos via *Power BI*

A captura de tela mostra a tabela de dados filtrados no *Power BI* do INPI. A tabela é intitulada "Dados do exame técnico do exame dos pedidos de patente prioritários" e contém as seguintes colunas: Protocolo, Modalidade, Data do requerimento, Situação do requerimento, Processo de patente, Data de depósito, Área tecnológica, Campo tecnológico, Situação do processo, e tempo entre o requerimento e a decisão. A tabela apresenta uma lista de pedidos de patente, com informações sobre sua data de depósito, situação, área e campo tecnológico, e o tempo decorrido desde o requerimento até a decisão.

Protocolo	Modalidade	Data do requerimento	Situação do requerimento	Processo de patente	Data de depósito	Área tecnológica	Campo tecnológico	Situação do processo	tempo entre o requerimento e a decisão
870230004038	Tecnologia verde	24/01/23	Admitido	10202013939	09/08/22	Química	Biotecnologia	Concedido	1807
870230004659	Tecnologia verde	24/01/23	Admitido	10202025259	16/11/22	Engenharia Mecânica	Motores, bombas e turbinas	Concedido	949
870230004447	Tecnologia verde	17/01/23	Inadmitido	11202023750	18/06/20	Engenharia Mecânica	Motores, bombas e turbinas	Formal	
870230004191	Tecnologia verde	16/01/23	Inadmitido	102019019346	17/09/19	Instrumentos	Tecnologia aplicada a medicina	Concedido	140
870230002311	Tecnologia verde	12/01/23	Admitido	102021000817	08/01/21	Química	Química de materiais básicos	Negado	934
870230002787	Tecnologia verde	04/01/23	Admitido	102020239392	21/11/22	Química	Engenharia química	Concedido	228
870230001065	Tecnologia verde	04/01/23	Admitido	11202012815	03/05/22	Química	Tecnologia aplicada ao meio ambiente	Concedido	890
870230000528	Tecnologia verde	03/01/23	Admitido	10202018065	09/09/22	Engenharia Mecânica	Outras máquinas especiais	Concedido	489
870220118452	Tecnologia verde	16/12/22	Inadmitido	102020251594	09/12/22	Engenharia Mecânica	Motores, bombas e turbinas	Exame	
870220117812	Tecnologia verde	16/12/22	Admitido	102017016418	01/07/17	Química	Biotecnologia	Concedido	575
870220116759	Tecnologia verde	16/12/22	Inadmitido	102017025946	07/11/17	Química	Tecnologia aplicada ao meio ambiente	Negado	272
870220116618	Tecnologia verde	13/12/22	Inadmitido	102020219134	15/06/22	Química	Produtos farmacêuticos	Exame	
870220116545	Tecnologia verde	16/12/22	Admitido	102020217145	26/08/22	Química	Tecnologia aplicada ao meio ambiente	Concedido	203
870220116890	Tecnologia verde	13/12/22	Admitido	112020148899	19/02/20	Química	Química macromolecular e polímeros	2ª instância	195
870220115974	Tecnologia verde	12/12/22	Inadmitido	11202015108	05/07/21	Engenharia Mecânica	Outras máquinas especiais	Exame	
870220115987	Tecnologia verde	12/12/22	Inadmitido	112020201315	29/06/21	Engenharia Mecânica	Outras máquinas especiais	Exame	
870220115874	Tecnologia verde	12/12/22	Admitido	202020209662	11/04/22	Engenharia Mecânica	Maneja	Concedido	421
870220115988	Tecnologia verde	09/12/22	Admitido	102019026000	11/12/19	Química	Química de materiais básicos	2ª instância	158
870220114401	Tecnologia verde	08/12/22	Admitido	102019026021	16/12/19	Química	Biotecnologia	Negado	222
870220114408	Tecnologia verde	08/12/22	Admitido	102020201841	24/01/20	Engenharia Mecânica	Outras máquinas especiais	Concedido	201
870220113935	Tecnologia verde	05/12/22	Admitido	102021006555	01/04/21	Química	Tecnologia aplicada ao meio ambiente	2ª instância	238
870220112567	Tecnologia verde	05/12/22	Inadmitido	102020217865	06/09/22	Engenharia Mecânica	Outras máquinas especiais	Exame	

Fonte: INPI (2022)

Figura 3 – Tela Inicial do Mecanismo de busca do INPI

Fonte: INPI (2022)

Figura 4 – Tela do mecanismo de busca de patentes completa

Fonte: INPI (2022)

O programa foi desenvolvido em linguagem de programação *Python* e se organiza em uma série de funções interconectadas, cada uma desempenhando um papel específico dentro do processo de coleta de informações. A primeira seção do código estabelece informações básicas, como o URL base do *site* e os cabeçalhos HTTP a serem utilizados, aprimorando a simulação de requisições de navegador por meio do campo “User-Agent.” Isso é seguido pela definição da função *login_request*, que realiza uma solicitação GET para obter os *cookies* de autenticação necessários para acessar áreas restritas do *site*. Essa função é projetada para enfrentar problemas de conectividade por meio de tentativas repetidas e atrasos.

O programa prossegue com duas funções principais: *url_scraping* e *data_scraping*. A primeira, *url_scraping*, extrai URLs associadas a registros de patentes, lendo um arquivo de entrada que contém números de protocolo das patentes a serem consultadas. A função chama a função *result_url_request*, que envia solicitações POST para obter as URLs de patentes correspondentes, usando os *cookies* obtidos anteriormente. Essas URLs são então escritas em um arquivo de saída no formato CSV (*Character-separated Values* – valores separados por vírgulas), que é um tipo de arquivo de texto fundamental para a transferência de informações, que possibilita o salvamento de dados em um formato estruturado de tabela.

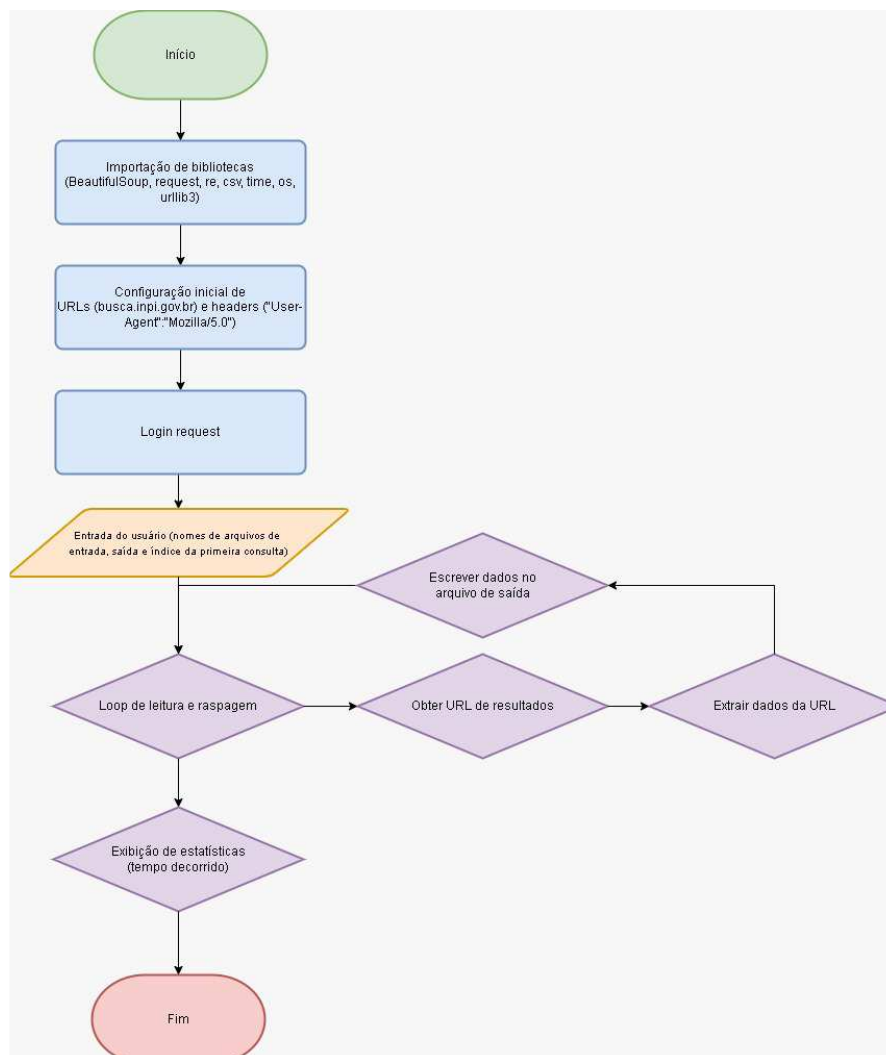
Por sua vez, a função `data_scraping` concentra-se na extração detalhada de informações de patentes individuais. Essa função lê um arquivo de entrada contendo as URLs das patentes a serem consultadas e, para cada URL, chama a função `result_data_request`, que extrai os dados detalhados da respectiva patente. Esses dados extraídos são organizados e escritos em um arquivo de saída no formato CSV, novamente usando o caractere de tabulação como delimitador.

A função principal “*main*” coordena o fluxo de trabalho, iniciando pelo processo de autenticação por meio da função `login_request`. Em seguida, ela solicita ao usuário os nomes dos arquivos de entrada e saída, bem como a operação a ser realizada (extração de URLs ou de dados). Com base na escolha do usuário, a função direciona o programa para a operação apropriada, usando a estrutura *match*. Após a conclusão de todas as operações, uma mensagem é exibida, sinalizando a finalização do processo de raspagem de dados.

Para melhor entendimento da execução do programa, foi elaborado um fluxograma simplificado apresentado na Figura 5.

Diante de um contexto de prospecção de dados, o código utilizado representa uma implementação eficaz de raspagem de dados em um ambiente controlado, demonstrando a habilidade de interagir com um *site*, autenticar-se, extrair informações de diferentes tipos de páginas e organizá-las em um formato estruturado, como um arquivo do tipo CSV. A estrutura modular do código, a utilização de bibliotecas externas como Beautiful Soup e o uso de estruturas de controle claras contribuem para a legibilidade, a manutenção e a eficiência do programa como um todo. Os dados extraídos foram posteriormente exportados e armazenados em um arquivo CSV, que, por sua vez, foi importado para uma planilha no formato Excel, como mostra a Figura 6.

Figura 5 – Fluxograma de execução do programa



Fonte: Elaborada pelos autores deste artigo

Figura 6 – Dados extraídos do INPI organizados em planilha do Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
	dataDeposito	dataPublicacao	dataConcessao	puPaís	publumer	puData	siglaIPC	descIPC	siglaCPC	descCPC	titulo	resumo	nomeInv	
1	PI 0406712-6 B1	23/01/2004	20/12/2005	06/07/2011	ESTADOS UNIDOS	10/350,4 24/01/2003	"F03D 7/04", "F03D 9/00", "F03C 03/00"	"F03D 7/04 (1968.09)	JCo	-	TURBINA EOLICA CONECTADA A GERADOR A TURBINA EOLICA COM UM WILHELM C			
2	PI 0409750-5 A2	08/04/2004	09/05/2006	-	ORGANIZACAO EUROPEIA DE PATENTES	03/00829 09/04/2003	"H02J 3/38", "F03D 9/00"	"H02J 3/38 (1968.09)	JDi	-	FAZENDA DE VENTO E METODO FAZENDA DE VENTO E METODO PARA O HANS HER			
3	PI 0403798-9 B1	02/09/2004	07/06/2005	14/03/2017	ESTADOS UNIDOS	10/655,5 03/09/2003	"F03D 9/00", "H02P 9/04"	"F03D 9/00 (1968.09)	JAd	-	Sistema de controle de tensão "CONTROLE DA TENSÃO PARA GERADOR Einar V. L. C			
4	PI 0405546-2 B1	10/12/2004	05/09/2006	29/09/2015	-	-	"H01M 8/0252", "F03D 13/10"	"H01M 8/0252 (2016.01)"	[H01M 8]	[H01M 8]	CONJUNTO DE ESTRUTURAS PARA CONJUNTO DE ESTRUTURAS PARA MAI Bento Ma T			
5	PI 0512825-0 A2	27/06/2005	08/04/2008	-	ESTADOS UNIDOS	10/880,9 30/06/2004	"C08F 2/32", "C08L 91/00", "C08F 2/32 (1974.07)"	"C08F 2/32 (1974.07)"	JPr	-	METODO DE TRATAMENTO DE POLIMERO DE EMULSAO INVERSA, E MI Michael C			
6	CI 0405546-2 F1	08/07/2005	09/09/2008	22/03/2016	-	-	"F03D 11/00"	"F03D 11/00 (1968.09)"	JD	-	DESENVOLVIMENTO EM CONJOL DESENVOLVIMENTO EM CONJUNTO DE Bento Ma T			
7	PI 0503924-0 B1	26/08/2005	24/04/2007	15/09/2015	-	-	"B01F 7/18"	"B01F 7/18 (1968.09)"	JMi	-	EQUIPAMENTO MISTURADOR E EQUIPAMENTO DESLIZANTE PARA FABI Hiroaki Sui			
8	PI 0603613-2 B1	14/03/2006	20/11/2007	11/08/2015	-	-	"F03D 1/06"	"F03D 1/06 (1968.09)"	JMc	-	PA MULTIELEMENTOS COM PE PA MULTIELEMENTOS COM PERIFERICO Bento Ma T			
9	PI 0609722-7 B1	23/03/2006	20/04/2010	22/04/2020	ESTADOS UNIDOS	60/750,2 13/12/2005	"B02C 18/14", "B02C 18/00"	"B02C 18/14 (1968.09)"	JD	-	METODO DE FAZER MATERIAL F METODO DE FAZER MATERIAL FIBROSO Marshall X			
10	PI 0602708-3 B1	19/06/2006	06/02/2008	26/06/2018	-	-	"B60K 13/00"	"B60K 13/00 (1968.09)"	JD	-	DISPOSICAO CONSTRUTIVA AP DISPOSICAO CONSTRUTIVA APLICADA E Teruaki M			
11	PI 0602764-4 B1	04/07/2006	26/02/2008	22/03/2016	-	-	"B65D 88/12", "B65D 90/00"	"B65D 88/12 (1980.01)"	JC	-	METODO E EMBALAGEM PARA METODO E EMBALAGEM PARA TRANSPHILLIPS A T			
12	PI 0603299-0 B1	28/07/2006	18/03/2008	01/07/2014	-	-	"E03B 11/02"	"E03B 11/02 (1968.09)"	JD	-	SISTEMA DE COBERTURA COM SISTEMA DE COBERTURA COM BARREIRO Guillermo C			
13	PI 0614940-5 B1	17/08/2006	02/01/2013	14/02/2018	CHINA	2005100 22/08/2005	"H04L 12/12", "H04M 11/06", "H04L 12/12 (1990.01)"	"H04L 12/12 (1990.01)"	JR	[H04L 12]	DISPOSITIVO DE COMUNICACAO DISPOSITIVO DE COMUNICACAO E METU JULIE XIA			
14	PI 0700517-2 B1	15/02/2007	30/09/2008	16/02/2016	-	-	"B01D 46/50", "B01D 53/00", "B01D 46/50 (1974.07)"	"B01D 46/50 (1974.07)"	JF	[B01D 46]	PROCESSO DEGRADADOR MOL PROCESSO DEGRADADOR MOLECULAR Ivan Dom F			
15	PI 0708440-4 A2	05/03/2007	07/06/2011	-	REINO UNIDO	0604241 03/03/2006	"H01M 8/04"	"H01M 8/04 (1974.07)"	JC	-	CONJUNTO DE CELULA DE COM CONJUNTO DE CELULA DE COMBUSTIVEL Paul Alan			
16	PI 0705710-5 B1	12/06/2007	06/09/2011	02/07/2019	-	-	"G01F 1/00", "G06N 3/02"	"G01F 1/00 (1968.09)"	JMe	-	SISTEMA PARA DETECÇÃO DE V SISTEMA PARA DETECÇÃO DE VAZEMEN Julio Rob			
17	PI 0713417-7 B1	12/06/2007	27/03/2012	02/04/2019	ESTADOS UNIDOS	11/453,9 15/06/2006	"B02C 19/00"	"B02C 19/00 (1968.09)"	JD	-	METODO PARA A FABRICACAO COMPOSICAO E MATERIAIS FIBROSOS Marshall X			
18	PI 0703180-7 B1	20/07/2007	10/03/2009	03/11/2015	-	-	"A01C 1/00", "A01C 14/00"	"A01C 1/00 (1968.09)"	JAp	-	PROCESSO PARA INDUCCAO DA PROCESSO DE GERMINACAO E PRODUCC Sergio Yo L			
19	MU 8703068-3 Y1	21/08/2007	18/03/2008	10/02/2016	-	-	"C02F 11/04"	"C02F 11/04 (1980.01)"	JTr	-	PROCESSO PARA CAPTACAO DE GASES Toshiro N			
20	PI 0715922-2 B1	01/10/2007	25/03/2014	12/09/2017	ORGANIZACAO EUROPEIA DE PATENTES	0612163 02/10/2006	"C12N 1/18", "C12P 7/08"	"C12N 1/18 (1980.01)"	JMi	-	CELULA DE SACCHAROMYCES E PROCESSO PARA PRODUCCAO DE ETANO Jacobus T C			
21	MU 8702870-0 U2	10/10/2007	07/07/2009	-	-	-	"F28F 9/02", "F28F 25/00"	"F28F 9/02 (1968.09)"	JW	-	COLETOR ECOLOGICO COLETOR ECOLOGICO. Patente de moc José Carl			
22	PI 0718314-3 A2	26/10/2007	26/11/2013	-	ESTADOS UNIDOS	60/881,8 23/01/2007	"C08H 8/00", "C08L 97/02", "C08H 8/00 (2010.01)"	"C08H 8/00 (2010.01)"	JCo	-	PROCESSAMENTO DE BIOMASSA Marshall			
23	PI 0704292-2 B1	23/11/2007	21/07/2009	25/09/2018	-	-	"C02F 3/10", "C02F 3/00"	"C02F 3/10 (1980.01)"	JTra	-	SISTEMA NATURAL DE TRATAM SISTEMA NATURAL DE TRATAMENTO DE Galdino S C			
24	PI 0704744-4 C8	30/11/2007	28/07/2009	08/05/2018	-	-	"C02F 3/465"	"C02F 3/465 (1990.01)"	JTr	-	Eletrólitos para tratar Eletrolitos/Eletrólitos para tratar ENIO LEAF			
25	PI 0810488-3 A2	18/04/2008	16/06/2015	-	JAPÃO	2007-115 27/04/2007	"C10L 5/44"	"C10L 5/44 (1968.09)"	JCo	-	APARELHO PARA PRODUCCAO C APARELHO PARA PRODUCCAO DE BICO Yoshimas			
26	PI 0810249-0 A2	22/04/2008	18/11/2014	-	JAPÃO	2007-115 27/04/2007	"C10L 5/44"	"C10L 5/44 (1968.09)"	JCo	-	APARELHO E PROCESSO PARA MATODO PARA CONTROLAR DEPOSITO Yoshimas			
27	PI 0802355-9 B1	04/06/2008	22/04/2009	16/05/2017	ARGENTINA	P 200701 04/06/2007	"B01J 20/24", "B01J 17/022"	"B01J 20/24 (1980.01)"	JC	-	UM PRODUTO SEQUESTRANTE Um produto sequestrante de hidroxido Ricardo F			
28	PI 0812983-5 A2	09/06/2008	16/12/2014	-	ESTADOS UNIDOS	11/772,6 02/07/2007	"C11D 17/00", "C11D 3/33", "C11D 17/00 (1968.09)"	"C11D 17/00 (1968.09)"	JM	-	MATRIZ DE SOLIDIFICACAO INC MATRIZ DE SOLIDIFICACAO Michael F			
29	PI 0802977-6 B1	17/07/2008	30/10/2010	08/03/2016	-	-	"B01D 36/00", "F24D 17/00"	"B01D 36/00 (1985.01)"	JC	-	SISTEMA AUTOMATIZADO DE R SISTEMA AUTOMATIZADO DE RECUPERAR MARCOS R			
30	PI 0804291-8 B1	30/09/2008	13/07/2010	15/12/2015	-	-	"F23D 1/02", "F23R 5/00"	"F23D 1/02 (1968.09)"	JDu	-	APERFECCIONAMENTO EM QUEM APERFECCIONAMENTO EM QUEIMADOR DI ROBERT S			
31	PI 0806139-4 A2	07/10/2008	21/09/2010	-	-	-	"F23C 1/06"	"F23C 1/06 (1968.09)"	JAp	-	PROCESSO DE GERACAO DE VA PROCESSO DE GERACAO DE VAPORE E DANIEL CI			
32	PI 0809055-6 A2	07/10/2008	19/10/2010	-	-	-	"C05F 5/00", "C12F 3/10", "F23G 01/00"	"C05F 5/00 (1968.09)"	JFer	-	PROCESSO DE GERACAO DE VA PROCESSO DE GERACAO DE VAPORE ENO DANIEL CI			
33	PI 0814534-3 B1	17/10/2008	09/10/2018	12/11/2019	FRANÇA	CT/R007 19/10/2007	"E02B 8/06", "E02B 9/06"	"E02B 8/06 (1968.09)"	JDe	[E02B 8]	COMPORTA FUSIVEL PARA UM COMPORTA FUSIVEL PARA UMA SEBASTIEN F			
34	PI 0814548-5 B1	23/10/2008	29/10/2014	01/03/2016	ESTADOS UNIDOS	60/987,5 13/11/2007	"C12P 1/04", "C12P 7/08", "E04H 01/00"	"C12P 1/04 (1980.01)"	JPr	-	METODO PARA CAPTURAR CA MATODO PARA CAPTURAR CARBONO P Sean Sim L			
35	PI 0820556-6 B1	13/11/2008	04/11/2014	22/03/2016	ESTADOS UNIDOS	60/987,7 13/11/2007	"C12N 1/20", "C12P 7/08"	"C12N 1/20 (1980.01)"	JMi	-	BACTERIA E METODOS PARA US BACTERIAS EMATODOS PARA USO DAS Richard L R			
36	PI 0817553-5 A2	28/11/2008	31/03/2015	-	SUECIA	0702648 29/11/2007	"F25B 17/08", "F25B 30/06", "F25B 17/08 (1968.09)"	"F25B 17/08 (1968.09)"	JM	[F25B 20]	BOMBA DE CALOR QUIMICA BOMBA DE CALOR QUIMICA. RYD OLSSO			
37	PI 0823414-0 A2	19/12/2008	06/10/2015	-	-	-	"B65D 61/00", "B65D 85/62", "B65D 61/00 (1968.09)"	"B65D 61/00 (1968.09)"	JM	-	METODO PARA EMBALAR E SIS Um sistema de embalagem de trãs d Ismael Ri			

Fonte: Elaborada pelos autores deste artigo

Essa abordagem foi adotada para viabilizar a análise e a exploração mais eficazes e convenientes das informações obtidas. A transposição dos dados para um ambiente de planilha eletrônica se apresenta como uma estratégia que permite a aplicação de diversas ferramentas de análise e manipulação de dados, facilitando a compreensão e a interpretação dos resultados.

Os dados extraídos correspondem às informações contidas na folha de rosto da carta patente, que são: dados bibliográficos essenciais, depositante ou inventor, título da invenção, país de origem, classificação internacional, nomes do inventor e do depositante, número da patente, data de depósito e de publicação e um sucinto resumo da matéria a ser protegida.

A implementação da raspagem de dados do INPI revelou uma relativa facilidade de acesso às informações, uma vez que o referido *site* não apresenta mecanismos robustos para mitigar a prática da raspagem de dados. Apesar da intrínseca instabilidade inerente às solicitações *web* e da possibilidade de ocorrência de erros de conexão, o processo de acesso aos recursos do INPI e a obtenção de dados em grandes volumes não demonstraram complexidade substancial.

A análise da acessibilidade do *site* do INPI à raspagem de dados ressalta a importância de um escopo de pesquisa mais abrangente, que considere também os aspectos de segurança e de privacidade associados à prática de coleta de informações *on-line*.

O processo de raspagem de dados foi realizado em um notebook Dell com processador Intel® Core™ i7-8550U, frequência de 1,80 GHz a 2,00 GHz, 8 GB de memória RAM e sistema operacional Windows 10. A conexão

de internet foi residencial, via cabo, com velocidade de 250 Mbps. O tempo de execução em condições normais, sem perda de conexão com o servidor do INPI ou outros problemas causados por excesso de acessos ao servidor, bem como o equipamento utilizado para execução do programa e a velocidade da rede doméstica estão descritos no Quadro 1.

Quadro 1 – Equipamento, sistema e rede utilizados na execução do programa

Equipamento: Notebook Dell, processador Intel® Core™ i7-8550U CPU 80GHz 100GHz
Memória RAM: 8gb
Sistema Operacional: Windows 10
Rede de Internet: Conexão residencial via cabo de 250 Mbps (Megabits por segundo)
Tempo médio de execução do programa: 1 hora

Fonte: Elaborado pelos autores deste artigo

3 Referencial Teórico

De acordo com Jungmann (2010), o conhecimento e a tecnologia desempenham um papel fundamental na promoção do desenvolvimento econômico. Dessa forma, a inovação quando analisada e explorada de maneira estratégica torna-se um fator-chave para a criação de riquezas e a melhoria da qualidade de vida da população.

As patentes representam uma fonte valiosa e estratégica para o monitoramento do estado da arte em tecnologia, oferecendo acesso direto a inovações que estão na fronteira do conhecimento em diversos setores. Tais documentos não apenas descrevem avanços específicos e soluções

tecnológicas, mas também fornecem informações essenciais que apoiam processos de inteligência competitiva.

A análise criteriosa de patentes permite acompanhar as atividades de pesquisa e desenvolvimento de concorrentes e mapear tecnologias que impactam o cenário competitivo. Dessa forma, as patentes constituem uma ferramenta robusta para empresas e instituições que buscam monitorar seus concorrentes e fortalecer sua posição no mercado (Teixeira; Souza, 2013).

A propriedade industrial concede o monopólio das inovações aos seus inventores, que são protegidos pelas informações que estão descritas dentro do documento de patentes, sendo assegurada pela Lei n. 9.279/96 (Brasil, 1996). A referida lei também fundamenta o Programa Patentes Verdes, resultado de fins do conjunto jurídico de proteção ambiental, inovação e demais aspectos legais. Assim, a lei proporciona segurança para o inventor divulgar suas pesquisas, com garantias legais, mesmo que por um tempo determinado, permitindo a rápida utilização, pela sociedade, de inovações tecnológicas benéficas ao meio ambiente e à sustentabilidade.

A linguagem de programação *Python* foi criada por Guido van Rossum e lançada pela primeira vez em 1991 (McKinney, 2018). Ela se destaca por sua simplicidade e legibilidade, sendo uma excelente escolha para iniciantes em programação. Além disso, *Python* é uma linguagem de alto nível, o que significa que seu código é mais próximo da linguagem humana, facilitando a compreensão e a escrita de programas. Sua sintaxe clara e concisa, aliada a uma ampla variedade de bibliotecas e *frameworks*, contribui para a popularidade e a versatilidade da linguagem.

Python é uma linguagem interpretada, o que significa que o código pode ser executado diretamente sem a necessidade de compilação, agilizando o processo de desenvolvimento. Outra vantagem é a sua ampla biblioteca padrão, que oferece uma vasta gama de módulos e funções prontos para uso, facilitando o desenvolvimento de aplicações. *Python* também é conhecido por sua versatilidade, sendo utilizado em uma variedade de áreas, desde desenvolvimento *web* até análise de dados e Inteligência Artificial.

O cenário atual aponta para um maior crescimento do uso de *Python* em diversas aplicações devido à sua facilidade de uso e à disponibilidade de bibliotecas e *frameworks* especializados. Além disso, a comunidade de desenvolvedores *Python* está sempre ativa, criando ferramentas e aprimorando as existentes, o que contribui para a contínua evolução e relevância da linguagem. Portanto, espera-se que *Python* continue a ser uma das linguagens de programação mais populares e influentes no cenário tecnológico nos próximos anos.

A prospecção tecnológica pode ser definida como um meio sistemático de mapear desenvolvimentos científicos e tecnológicos futuros capazes de influenciar de forma significativa uma indústria, a economia ou a sociedade como um todo. O estudo de prospecção tecnológica, juntamente com o acompanhamento de novas tecnologias no mercado, contribui na tomada de decisões em diversas esferas, constituindo uma estratégia de inteligência competitiva (Amparo; Ribeiro; Guarieiro, 2012; Paranhos; Ribeiro, 2018).

De acordo com Kupfer e Tigre (2004), a prática da prospecção, como definida, é realizada por empresas e organizações de diversas áreas. Essas investigações podem ser realizadas quando são analisadas as patentes, identificando o estado da técnica de uma inovação e mapeando cenários futuros de uma determinada tecnologia.

Estudos de prospecção realizados por meio de patentes possibilitam identificar se uma tecnologia está patenteada; qual seu estado da técnica; quais empresas e/ou países estão interessados em determinadas tecnologias; e quem são seus concorrentes (Quintella *et al.*, 2011), entre outros aspectos.

Indicadores de patentes têm aplicações de níveis estratégicos e são usados em estudos de diversos setores da indústria, com capacidade de investigar e monitorar tecnologias nacionais e regionais. A análise de patentes permite identificar os indicadores quantitativos de tecnologia de P&D, servindo como uma útil ferramenta para auxiliar investimentos em inovação (Narin; Breitzman; Thomas, 2005).

Informações tecnológicas presentes em patentes destacam-se tanto pelo seu volume quanto pela sua abrangência em todos os campos do conhecimento tecnológico, sendo acessíveis em meio eletrônico e organizadas de forma sistemática, com critérios objetivos que facilitam buscas e coleta de dados. Nesse contexto, o desenvolvimento e a utilização de rotinas e ferramentas voltadas para a prospecção e a inteligência tecnológica tornam-se essenciais para explorar essas informações de maneira eficaz.

Assim, conclui-se que a análise de qualquer tecnologia é incompleta sem o acesso e o uso das bases de dados estruturadas especificamente para reunir e disponibilizar esse vasto acervo informacional (Baltazar *et al.*, 2017).

4 Análise e Discussão dos Resultados

A capacidade de coletar dados de forma automatizada acelera a análise de patentes, representando, assim, uma vantagem competitiva na prospecção tecnológica, podendo ser utilizada como recurso para a estratégia de negócios

por meio do monitoramento de atividades da concorrência. Além disso, no âmbito da prospecção tecnológica, a prática permite a identificação de oportunidades de inovação, aumentando a competitividade (Guzmán Sánchez, 1999).

Ao permitir o monitoramento detalhado das atividades da concorrência, a automação oferece dados valiosos sobre tendências de mercado, lacunas tecnológicas e áreas promissoras para investimentos. Dessa forma, as organizações podem alinhar suas estratégias de negócios com maior precisão, aproveitando oportunidades emergentes para se posicionarem de maneira competitiva no mercado.

Além disso, no contexto da prospecção tecnológica, a automação da análise de patentes vai além de seu simples monitoramento. Essa prática permite a identificação precoce de áreas tecnológicas em ascensão, o mapeamento de redes de colaboração entre empresas e instituições de pesquisa e a avaliação de dinâmicas de inovação em setores estratégicos. Sendo assim, organizações ganham a capacidade de prever mudanças no cenário competitivo, orientar seus esforços de pesquisa e desenvolvimento e estabelecer parcerias com base em dados concretos e confiáveis.

A integração de técnicas automatizadas para a patentometria amplia, ainda, o alcance da inovação ao permitir que empresas e/ou pesquisadores identifiquem novas oportunidades em escala global por meio de dados disponibilizados gratuitamente e que possam ser passíveis de raspagem. Assim, além de aumentar a competitividade individual, essa prática contribui para o avanço tecnológico coletivo, promovendo um ambiente mais dinâmico e inovador ao tornar mais eficiente a coleta de informações.

Conforme destacado por Guzmán Sánchez (1999), a aplicação eficiente da automação nesse campo pode ser um divisor de águas para empresas e diversas instituições nos mais variados ramos de atividades que buscam não apenas se adaptar às mudanças tecnológicas, mas também liderá-las. A competitividade requer especial atenção e planejamento estratégico para que indústrias possam investir assertivamente seus recursos, tornando seus projetos economicamente viáveis e rentáveis, garantindo o sucesso e o alcance de seus objetivos empresariais.

Nesse sentido, a raspagem de dados proporciona a coleta automatizada de informações de patentes e tendências tecnológicas. Essa técnica tem a capacidade de automatizar a coleta de dados a partir de uma ampla variedade de fontes (Mitchell, 2018), ao contrário de metodologias executadas manualmente, como a de Santos e Santos (2018, p. 20) “Após a busca manual e individual de cada documento de patente, os dados foram extraídos, identificados e tabulados”.

Em consonância, a raspagem de dados é uma técnica amplamente utilizada para a extração automatizada de

informações disponíveis em *websites*. Essa prática consiste em utilizar ferramentas e *scripts* para capturar dados estruturados ou semiestruturados da *web*, que podem posteriormente ser armazenados, processados e analisados para diversas finalidades (Lawson, 2021; McKinney, 2018).

A raspagem de dados, também conhecida como *web scraping*, é uma técnica amplamente empregada para a extração automatizada de informações disponíveis em *websites*. Essa prática envolve o uso de ferramentas e *scripts* específicos para capturar dados estruturados ou semiestruturados, como tabelas, listas e textos em formato HTML (*Hypertext Markup Language*), transformando-os em informações utilizáveis para uma ampla gama de finalidades (Lawson, 2021; McKinney, 2018).

Os dados coletados podem ser armazenados em bancos de dados, processados para análises ou até mesmo integrados a sistemas de inteligência artificial e modelos preditivos de aprendizado de máquina para uso estratégico e tomada de decisões. Assim, a raspagem de dados não é apenas uma ferramenta técnica, mas também um componente estratégico em um mundo cada vez mais orientado por dados, fornecendo uma base sólida para a inovação e o desenvolvimento em diversas áreas.

Para o presente estudo, após o processo de raspagem das folhas de rosto de patentes verdes depositadas entre 2007 e 2022, as informações contidas nas referidas folhas de rosto foram cuidadosamente organizadas em um arquivo no formato CSV, que, em seguida, foi convertido para acesso em formato de planilha Excel para facilitar a manipulação e análise dos dados de 1.251 pedidos de depósitos de patentes verdes.

O desenvolvimento do código para raspagem de dados, demandou aproximadamente 15 horas de trabalho. Esse período abrangeu a implementação inicial do *script*, a organização e a estruturação dos dados extraídos, bem como a realização de testes para validar sua funcionalidade (Broucke; Baesens, 2020).

Durante os testes, foram identificados e corrigidos possíveis erros, garantindo a robustez e a confiabilidade do código desenvolvido. O tempo investido também considerou ajustes necessários para adaptar o *script* às especificidades do *site*, e eventuais variabilidades no carregamento das páginas, assegurando a precisão e a eficiência do processo.

Essa organização permitiu uma visualização mais estruturada das informações e contribuiu para o processamento eficiente dos dados coletados, viabilizando análises mais detalhadas sobre as tendências e inovações tecnológicas em tecnologias verdes ao longo do período estudado.

Por meio da raspagem de dados desenvolvida e implementada neste estudo, foi possível extrair as informações detalhadas de um total de 1.251 pedidos de

depósitos de patentes verdes. A raspagem aplicada ao mecanismo de busca de patentes do *site* do INPI mostrou-se eficaz e estável, podendo fazer a mineração dos dados predefinidos no programa.

Os dados coletados incluíram informações bibliográficas essenciais, como nome do depositante e do inventor, título da invenção, país de origem, classificação internacional, número da patente, além de datas de depósito e publicação. Adicionalmente, também foi obtido um resumo sucinto da matéria a ser protegida, permitindo uma visão clara e organizada do conteúdo analisado.

Os dados raspados foram sistematicamente estruturados conforme as definições e parâmetros estabelecidos no programa executado, garantindo a consistência e a integridade da informação para análises futuras.

O tempo médio para execução da raspagem de dados foi de aproximadamente uma hora, sendo realizadas diversas tentativas de execução do programa que, em muitas vezes, foram interrompidas pela instabilidade no acesso ao servidor do INPI. Esse tempo de processamento reflete tanto a adequação do equipamento quanto a velocidade da rede, sendo fundamental para cumprir as demandas de eficiência e precisão do estudo (Broucke; Baesens, 2020).

Em uma análise preliminar, sugere-se que a elevada quantidade de acessos ao servidor, a velocidade de conexão disponível, a fragilidade do *software* e até mesmo a obsolescência do *hardware* podem estar diretamente relacionados aos problemas de acesso às páginas e de interrupção de conexão enfrentados durante as tentativas de execução e acesso ao servidor do INPI.

A capacidade do servidor em lidar com um grande volume de requisições simultâneas parece ser insuficiente, resultando em instabilidades frequentes nas tentativas de acesso. Adicionalmente, a velocidade de conexão, tanto do usuário quanto do servidor, influencia na eficácia da raspagem de dados, já que as conexões lentas ou instáveis podem interromper o processo.

A possível fragilidade do *software* do INPI é um fator a ser considerado, pois as falhas e as interrupções apresentadas durante a navegação e extração de dados evidenciam a necessidade urgente de aprimoramentos na robustez e na resiliência das aplicações utilizadas. Esses problemas técnicos não apenas impactam a experiência do usuário, mas também limitam a eficiência dos processos de coleta das informações ali contidas para posterior análises.

Além disso, a possível obsolescência do *hardware* subjacente também pode comprometer significativamente a capacidade de processamento e a velocidade de resposta do servidor, intensificando as dificuldades de desempenho e reduzindo a disponibilidade dos dados para acesso dos usuários.

Investimentos para a modernização tanto do *software* quanto do *hardware* são, portanto, essenciais para garantir uma infraestrutura estável e eficiente, capaz de atender às demandas crescentes por acesso ágil e confiável aos registros de patentes disponibilizados pelo INPI.

Portanto, ao se constatar instabilidades no acesso ao servidor do INPI, se faz necessário investigar os fatores relacionados à fragilidade do *software* e à possível obsolescência do *hardware* no sistema do INPI, com o objetivo de propor melhorias estruturais que elevem a estabilidade e a eficiência dos serviços prestados ao público que tenta acessar as informações de patentes.

Essas melhorias, além de promoverem um ambiente tecnológico mais robusto e confiável, potencializariam a capacidade de suporte à prospecção tecnológica, permitindo que pesquisadores, empresas e demais usuários possam ter acesso ágil, contínuo e estável às informações de patentes, consideradas essenciais para a análise e o desenvolvimento das inovações.

Ao tornar os dados mais acessíveis e confiáveis, o INPI também ampliaria seu papel estratégico no fortalecimento do sistema nacional de inovação, fomentando avanços na pesquisa científica, no desenvolvimento de produtos e na competitividade tecnológica do Brasil em âmbito global.

Dessa forma, investir na modernização da infraestrutura tecnológica do INPI é mais do que uma necessidade operacional, trata-se de uma necessidade fundamental no sentido de impulsionar o crescimento econômico e o progresso tecnológico do país, fortalecendo e aperfeiçoando as ferramentas de prospecção tecnológica que dependem tanto de investimentos em infraestrutura quanto da modernização dos métodos de busca e recuperação de dados, promovendo, assim, o avanço no estado da arte de estudos de prospecção tecnológica.

5 Considerações Finais

A prospecção tecnológica em documentos de patentes requer estratégias bem elaboradas de busca e recuperação de registros que atendam às necessidades informacionais dos demandantes. Diante disso, o processo aplicado permitiu explorar e compreender os dados coletados disponíveis no *site* do INPI de maneira satisfatória, muito embora alguns desafios enfrentados ressaltem a importância contínua do aprimoramento das ferramentas de análise de dados e infraestrutura de acesso.

Embora a concepção da metodologia empregada no presente trabalho tenha se utilizado em momentos pontuais de técnicas manuais de pesquisa, ela pôde ser executada com êxito e agilidade devido ao emprego da raspagem de dados. Utilizar a raspagem de dados desenvolveu por meio da linguagem de programação *Python* na prospecção

tecnológica de patentes de tecnologias verdes no sítio eletrônico do INPI revelou-se uma abordagem satisfatória e inovadora, representando um avanço significativo em relação às metodologias tradicionais de busca e recuperação de informações manuais e sequenciais, permitindo a extração rápida e automatizada de grandes volumes de informações.

O processo de raspagem de dados foi cuidadosamente planejado e executado, garantindo a confiabilidade e a eficiência dos procedimentos, resultando em um método satisfatório, escalável e adaptável a outros mecanismos de busca em bases de dados de patentes, tanto nacionais quanto internacionais.

Apesar do grande volume de dados extraído, o tempo médio de execução mostrou-se satisfatório, evidenciando o equilíbrio entre os recursos de *hardware*, *software* e conexão à internet utilizados ao longo de toda a operação. Essa integração eficiente reforça o potencial do método para aplicações mais amplas e complexas, contribuindo para a análise e prospecção tecnológica em diferentes contextos.

Sua principal contribuição reside na capacidade de automatizar a extração, organização e estruturação de dados. Essa abordagem não apenas otimiza o tempo e os recursos necessários para a coleta de informações, mas também amplia as possibilidades de análise, permitindo que os dados sejam utilizados de forma direcionada para atender a diferentes objetivos estratégicos, sejam eles acadêmicos, corporativos ou tecnológicos. Assim, a raspagem de dados reafirma sua relevância em um cenário cada vez mais orientado por dados, sendo indispensável para transformar informações brutas em conhecimentos acionáveis.

A aplicação bem-sucedida desse método não apenas validou sua eficácia, mas também ressaltou seu potencial para acelerar estudos patentométricos. A organização estruturada das informações fornece uma base sólida para investigações futuras, otimizando o tratamento de grandes volumes de dados e facilitando análises em áreas emergentes.

Conclui-se que a raspagem de dados aplicada à prospecção tecnológica em documentos de patentes representa um avanço importante em relação aos métodos tradicionais, contudo, para consolidar e expandir essa abordagem, é imprescindível que novos estudos sejam realizados. A partir da técnica empregada e dos resultados obtidos, torna-se evidente que a raspagem de dados se configura como uma ferramenta altamente eficiente para apoiar a análise de informações em grandes volumes.

6 Perspectivas Futuras

Diante desse contexto, um dos principais desafios identificados é a limitação técnica e estrutural do *site* do

INPI, como a instabilidade de conexão e as limitações na capacidade de processamento de solicitações simultâneas. É essencial que o INPI considere melhorias em sua infraestrutura para suportar um volume mais elevado de acessos, por meio da implementação de servidores mais robustos e escaláveis. Adicionalmente, torna-se necessário repensar os mecanismos de busca e patentes no *site* do INPI, proporcionando uma solução prática para facilitar a extração de dados, reduzindo a dependência de técnicas de raspagem de dados, que, embora eficazes, acabam sobrecarregando o sistema.

Investigações futuras devem focar no aprimoramento das ferramentas e técnicas de raspagem de dados, bem como na infraestrutura de acesso a bases de dados, especialmente em contextos que envolvam volumes massivos de registros. É recomendável o desenvolvimento de metodologias que combinem processos automatizados e técnicas analíticas avançadas para que complementem a prospecção tecnológica. Tais esforços certamente contribuirão para o estado da arte, aprimorando a capacidade de recuperação e análise de informações.

Como perspectivas futuras, espera-se que investimentos sejam feitos na área de tecnologia da informação do INPI de modo a atender, sem interrupções do serviço, à demanda da sociedade civil e empresarial por informações estratégicas valiosas. É preocupante constatar que os sistemas de informação do INPI não são capazes de suportar o fluxo de requisições de acesso às informações.

Referências

- AMPARO, K. K. D. S.; RIBEIRO, M. D. C. O.; GUARIEIRO, L. L. N. Estudo de caso utilizando mapeamento de prospecção tecnológica como principal ferramenta de busca científica. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 17, n. 4, p. 195-209, out.-dez. 2012.
- BALTAZAR, L. F. *et al.* Patentes como fonte de informação tecnológica para subsídio à pesquisa: uma análise amostral da Universidade Federal do ABC. **Cadernos de Prospecção**, Salvador, v. 10, n. 4, p. 681-681, 2017.
- BRASIL. **Lei n. 9.279, de 14 de maio de 1996**. Regula direitos e obrigações relativos a propriedade industrial. *Diário Oficial da União*, Brasília, DF, 15 de maio 1996.
- BRASIL. Ministério da Saúde. Conselho Nacional de Saúde. Resolução n. 510, de 7 de abril de 2016. Trata sobre as diretrizes e normas regulamentadoras de pesquisa em ciências humanas e sociais. **Diário Oficial da União**, Brasília, DF, seção 1, n. 98, p. 44, 24 de maio de 2016.

BROUCKE, S. V.; BAESENS, B. **Practical web scraping for data science: best practices and examples with python**. Birmingham: Packt Publishing, 2020.

GUZMÁN SÁNCHEZ, M. V. **Patentometría: herramienta para el análisis de oportunidades tecnológicas**. 1999. 130f. Tese (Doutorado em Gerência de Informação Tecnológica) – Facultad de Economía, Universidad de La Habana, Havana, 1999.

HAIR JR., J. F. *et al.* **Multivariate data analysis**. 7. ed. London: Pearson, 2010.

INPI – INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL. **Resolução n. 175, de 5 de novembro de 2016**. Disciplina o exame prioritário de pedidos de “Patente Verde”. Rio de Janeiro: INPI, 2016.

INPI – INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL. **Busca de patentes**. 2022. Disponível em: <https://busca.inpi.gov.br/pePI/>. Acesso em: 11 jan. 2025.

INPI – INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL. **Portaria INPI n. 247/2020**. 2020. Disciplina o trâmite prioritário de processos de patentes no âmbito do INPI.

JUNGSMANN, D. M. **A caminho da inovação: proteção e negócios com bens de propriedade intelectual – guia para o empresário**. Brasília, DF: IEL, 2010.

KLINE, R. B. **Principles and practice of structural equation modeling**. 4. ed. New York: Guilford Press, 2016.

KUPFER, D.; TIGRE, P. B. Prospecção tecnológica. In: CARUSO, L. A.; TIGRE, P. B. (org.). **Modelo SENAI de prospecção: documento metodológico**. Montevideo: OIT/Cinterfor, 2004. p. 17-35.

LAWSON, R. **Web Scraping with Python: scrape data from any website with the power of Python**. Packt: Birmingham, 2021.

LEÃO, A. P. S. *et al.* Power BI para tomada de decisões estratégicas: análise de indicadores-chave de desempenho (KPIS). **Revista Foco**, Curitiba, v. 16, n. 7, e2472, p. 1-28, 2023. Disponível em: <https://ojs.focopublicacoes.com.br/foco/article/view/2472>. Acesso em: 11 jan. 2025.

MCKINNEY, W. **Python for data analysis**. O'Reilly Media: Sebastopol, 2018.

MITCHELL, R. **Web scraping com Python: coletando mais dados da web moderna**. O'Reilly Media: Sebastopol, 2018.

NARIN, F.; BREITZMAN, A.; THOMAS, P. Using patent citation indicators to manage a stock portfolio. In: MOED, H. F. *et al.* **Handbook of Quantitative Science and Technology Research: the use of publication and patent statistics in studies of S&T systems**. 2005. p. 553-568.

PARANHOS, R. D. C. S.; RIBEIRO, N. M. Importância da prospecção tecnológica em base em patentes e seus objetivos da busca. **Cadernos de Prospecção**, Salvador, v. 5, n. 1, p. 1274, 2018.

QUINTELLA, C. M. *et al.* Prospecção tecnológica como uma ferramenta aplicada em ciência e tecnologia para se chegar à inovação. **Revista Virtual de Química**, Niterói, v. 3, n. 5, p. 406-415, dez. 2011.

SANTOS, N. J. B.; SANTOS, M. J. C. Mapeamento do desenvolvimento tecnológico de patentes verdes relacionadas ao gerenciamento de resíduos. **Cadernos de Prospecção**, Salvador, v. 11, n. 1, p. 17-25, jan.-mar. 2018.

TEIXEIRA, R. C.; SOUZA, R. R. O uso das informações contidas em documentos de patentes nas práticas de inteligência competitiva: apresentação de um estudo das patentes da UFMG. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 18, n. 1, p. 106-125, jan.-mar. 2013.

Sobre os Autores

Ricardo Pimentel Nogueira

E-mail: ricardo_p_nogueira@yahoo.com.br

ORCID: <https://orcid.org/0000-0001-5969-9336>

Mestre em Propriedade Intelectual e Transferência de Tecnologia para Inovação – PROFNIT, pela Universidade Estadual de Maringá em 2024.

Endereço profissional: Centro de Ciências Sociais Aplicadas, C23, Vila Esperança, Maringá, PR. CEP: 87035-510.

Antonio Roberto Giriboni Monteiro

E-mail: argmonteiro@uem.br

ORCID: <https://orcid.org/0000-0003-1894-0765>

Doutor em Engenharia de Produção pela Universidade Federal de São Carlos em 2008.

Endereço profissional: Universidade Estadual de Maringá, Departamento de Engenharia de Alimentos, Rua Dez de Maio, Zona 7, Maringá, PR. CEP: 87030-130.