

Sistemas de Busca de Patentes: análise comparativa entre Espacenet, Patentscope, Google Patents, Lens, Derwent Innovation Index e Orbit Intelligence

Patent Search Systems: Comparative Analysis Between Espacenet, Patentscope, Google Patents, Lens, Derwent Innovation Index and Orbit Intelligence

Edilson Araújo Pires¹

Nubia Moura Ribeiro²

Cristina M. Quintella³

¹Universidade Federal do Recôncavo da Bahia, Salvador, BA, Brasil

²Instituto Federal de Educação, Ciência e Tecnologia da Bahia, Salvador, BA, Brasil

³Universidade Federal da Bahia, Cruz das Almas, BA, Brasil

Resumo

Os sistemas de busca de patentes são cruciais para analisar as tecnologias inovadoras, o estado da técnica e para monitorar concorrentes e identificar parceiros de transferência de tecnologia, entre outros. Aqui se apresenta a análise comparativa de seis sistemas de busca de patentes: Espacenet, Patentscope, Google Patents, Lens, Derwent Innovation Index e Orbit Intelligence. Foram considerados 13 critérios de tipos de acesso, mecanismos e ferramentas de busca, cobertura da base e ferramentas de análise estatística. Seguidamente, foram comparadas oito classificações internacionais. Concluiu-se que, em busca de dados de patentes utilizando códigos de classificações internacionais, os sistemas gratuitos, como o Lens, têm conseguido resultados muito próximos aos de sistemas pagos. Observou-se que o Orbit se destaca em ferramentas de análises estatísticas e de exportação de dados de patentes, abrangência de documentos, ferramentas de interpretação dos dados e facilidades de exportação, constituindo o mais adequado para as demandas de uso em programas de pós-graduação.

Palavras-chave: Bases de Dados de Patentes. Sistemas de Busca de Patentes. Informação Tecnológica.

Abstract

Patent search systems are crucial for analyze of innovative technologies, state of the art, monitoring competitors and identifying technology transfer partners, among others. Here is presented a comparative analysis of six patent search systems: Espacenet, Patentscope, Google Patents, Lens, Derwent Innovation Index and Orbit Intelligence. Thirteen criteria of access type, search engines and tools, base coverage and statistical analysis tools were considered. Then, eight international patent classifications (IPCs) were compared. In finding patent data using international classification codes, free systems, such as Lens, have achieved very close results to paid systems. It was noted that Orbit excels in statistical analysis and patent data export tools, document coverage, data interpretation tools and export facilities, making it the most suitable for the necessities of postgraduate programs.

Keyword: Patent Data Basis. Patent Search Systems. Technological Information.

Área Tecnológica: Ciência da Informação.



1 Introdução

A propriedade intelectual tem assumido uma função importante na economia baseada no conhecimento. Cada vez mais organizações públicas e privadas se dedicam a produzir conhecimentos que possam ser transformados em produtos, processos e serviços inovadores e, conseqüentemente, sejam capazes de alavancar o desenvolvimento econômico, científico e tecnológico.

A patente é uma das modalidades de propriedade intelectual mais utilizada estando, presente em todos os domínios tecnológicos, dado que titulares com produtos ou processos que envolvem tecnologias inovadora e patenteada podem obter vantagens competitivas no mercado. Em contrapartida, quando se requer uma patente, o requerente precisa apresentar as características técnicas da invenção para a qual deseja a proteção. Essas informações são disponibilizadas nos bancos de dados de patentes, geralmente 18 meses (período de sigilo) após o depósito da patente e, desse modo, permitem a disseminação do conhecimento tecnológico. Assim, os documentos de patentes se destacam como uma fonte primária de informação tecnológica, com acesso mais rápido do conhecimento das tecnologias e inovações fundamentais para a indústria, a partir da descrição do invento (FRANÇA, 1997).

No entanto, as informações contidas em um pedido de patente vão além da descrição da invenção. É possível identificar diversos outros dados dos recursos humanos envolvidos na criação da tecnologia, das organizações titulares daquela tecnologia, datas em que as patentes foram requeridas, os países que são foco da proteção (potenciais mercados), a anterioridade do estado da arte de cada invenção, entre outros. Assim, as informações contidas em documentos de patentes podem ser utilizadas para diversas finalidades, como identificação de conhecimento associada a um determinado domínio tecnológico, verificação da patenteabilidade de uma nova invenção, identificação de profissionais capacitados em determinada tecnologia, busca de tecnologias específicas para aquisição ou identificação de empresas interessadas em adquirir certa tecnologia, entre outras.

Ao destacar a relevância das patentes em atividades de Inteligência Competitiva, Teixeira e Souza (2013) afirmam que as patentes, como fontes privilegiadas de informações precisas e atuais, se caracterizam como um instrumento eficaz para obtenção de vantagem competitiva, minimizando os riscos na tomada de decisões e nos investimentos em pesquisa e desenvolvimento. Ademais, os dados de patentes têm sido constantemente utilizados por economistas para analisar os impactos das patentes na economia de países, regiões e de organizações (GRILICHES, 1998).

Carrara e Russo (2017) destacam que as razões para realizar uma busca de patentes podem ser resumidas em três categorias principais:

- a) Patentes como: fonte de informação para transformar dados em conhecimento técnico, empresariais e jurídicos; instrumento para o acompanhamento ou a pesquisa de uma tecnologia ou produto específico; mecanismo para monitorar as patentes dos concorrentes, identificar tendências tecnológicas e tecnologias emergentes; e para transferência de tecnologia.
- b) Pesquisas sobre a anterioridade de uma invenção para: avaliar se esta atende aos requi-

sitos de patenteabilidade; escrever um pedido de patente de uma invenção nova; evitar infrações de outras patentes; preparar medidas legais para proteger sua nova patente; e acompanhar o tempo de proteção de uma determinada invenção.

- c) Pesquisas para tomar decisões sobre um investimento financeiro ou sobre a transferência ou a aquisição de uma tecnologia.

Esses e outros fatores explicam seu papel indispensável para pesquisadores, inventores, gestores e diversos profissionais que necessitam de informações nos sistemas de propriedade intelectual, mais, particularmente, nos sistemas de divulgação de documentos de patentes.

Um sistema mundial de informação de patentes para ser ideal deve consistir em todas as patentes de invenção e modelo de utilidade, requeridas e concedidas, de todos os escritórios de patentes nacional, regional e multinacional (STOCK; STOCK, 2006). Os autores ressaltam que o sistema ideal de informações de patentes deveria conter todas as informações bibliográficas, a página da folha de rosto do documento de patente, o texto completo e os desenhos disponíveis. Além disso, deveria disponibilizar os códigos de vários sistemas de classificação (por exemplo: Classificação Internacional de Patentes – *International Patent Classification (IPC)* –, Classificação Europeia de Patentes – *Cooperative Patent Classification (CPC)* –, indexação do *Chemical Abstract System* (para patentes na área de química), códigos do sistema Derwent e outros sistemas de indexação de escritórios nacionais de patentes para permitir pesquisa por temas, *status* legal da patentes e oferta de conteúdo não patentário tecnologicamente relevante. Segundo a WIPO (2018b), a IPC se configura como um meio para obter uma classificação internacionalmente uniforme dos documentos de patentes e pode ser utilizada como uma ferramenta efetiva para a busca de documentos de patentes. Além de facilitar o acesso à informação tecnológica, a IPC serve como instrumento de arranjo ordenado de documentos de patentes, facilita a pesquisa do estado da arte de determinado campo tecnológico e serve de base para elaboração de estatísticas de propriedade industrial, permitindo a avaliação do desenvolvimento tecnológico em diversas áreas (WIPO, 2018b).

São conhecidos diversos sistemas de busca de patentes. Inicialmente podem ser citados os sistemas vinculados aos escritórios nacionais de propriedade industrial, encontrados em quase todos os países que adotam um sistema de proteção da propriedade intelectual. Por exemplo, nos Estados Unidos da América, existe o sistema de busca de patentes do *United States Patent and Trademark Office (USPTO)*; no Japão, o sistema do *Japan Patent Office (JPO)*; na China, o sistema do *State Intellectual Property Office of the People's Republic of China (SIPO)*; na Coreia do Sul, o sistema *Korean Intellectual Property Office (KIPO)*. Esses sistemas disponibilizam apenas documentos de patentes que foram requeridos exclusivamente em seus escritórios, de modo que é muito comum os escritórios nacionais dedicarem-se aos documentos locais, como ocorre com o Instituto Nacional da Propriedade Industrial (INPI) no Brasil. Uma exceção à regra é o *Depatisnet*, sistema do Escritório Alemão de Patentes e Marcas (*German Patent and Trade Mark Office*), que possibilita buscar documentos de patentes de mais de 100 autoridades de patentes.

Outros sistemas de abrangência internacional são o Espacenet e o Patentscope. O primeiro é um sistema de busca de patentes mantido pelo Escritório Europeu de Patentes – *European Patent Office (EPO)* – e possui mais de 100 milhões de documentos de patentes, de mais de 100 países (EPO, 2018). O segundo é o sistema oficial de busca de documentos de patentes requeridas no Tratado de Cooperação em termos de Patentes – *Patent Cooperation Treaty (PCT)*

–, mantido pela Organização Mundial de Propriedade Intelectual – *World Intellectual Property Organization (WIPO)* –, que também pesquisa patente de diversos escritórios nacionais e regionais, abrangendo mais de 67 milhões de documentos (WIPO, 2018a).

Existem outros sistemas que não estão ligados diretamente a um escritório nacional ou regional de patentes, por exemplo, Google Patents, Lens, Free Patents Online/Sumobrain e RPXInsigh, etc., todos com algum módulo de acesso parcial gratuito. Também são identificados provedores privados de informação em Propriedade Intelectual, por exemplo, STN, Proquest Dialog, Totalpatent, Patbase, Patseer, PatSnap's, Wips Global, Orbit, Derwent Innovation Index, Global Patent Index etc.

De fato, são várias as fontes que podem ser utilizadas para obter informações de patentes. Mas essas fontes de informação se diferenciam em diversos aspectos, entre os quais os mais importantes estão relacionados aos mecanismos de busca disponibilizados para encontrar o que se procura, por exemplo, os tipos de classificações de patentes que podem ser utilizadas na pesquisa, a quantidade de documentos ou famílias de patentes e a cobertura em termos de número de autoridades de patentes, os mecanismos de exportação de resultados de buscas e de análise de estatísticas de patentes.

Alguns estudos têm focado características e diferenças de sistemas de busca de patentes e analisado e comparado diversos desses provedores de informação de patentes. Kulp (1984) descreveu e comparou as funcionalidades dos bancos de dados Questel, Dialog, SDC, Pergamum e IMPADOC. Stock e Stock (2005) fizeram uma análise e descrição detalhada do Questel-Orbit para apresentar uma análise SWOT – *Strengths, Weaknesses, Opportunities e Threats* – do provedor. Em outro estudo, Stock e Stock (2006) analisaram o conteúdo e a funcionalidade do sistema dos principais provedores de informação em propriedade intelectual: Thomson DIALOG, Espacenet, Questel-Orbit, STN International.

Jürgens e Herrero-Solana fizeram a comparação entre Espacenet, Patentscope e Depatisnet, analisando aspectos relacionados à cobertura de dados, funcionalidades de pesquisa, lista de resultados, visualização de informações bibliográficas das patentes e opções de exportação dos dados (JÜRGENS; HERRERO-SOLANA, 2015). Também investigaram um caso específico, patentes relacionadas à nanotecnologia, mapeando a cobertura da busca na base Latipat – uma seção da Espacenet que compreende documentos de patentes de 20 países da América Latina e Espanha (OEP) – no que tange à nanotecnologia, usando três diferentes sistemas (Espacenet, Patentscope e Invenes), concluindo que os resultados foram diferentes para os três sistemas, (JÜRGENS; HERRERO-SOLANA, 2016).

Uma comparação das características de cobertura de dados, possibilidades de busca e saídas de dados de Patentscope e Global Patent Index (GPI) foi realizada por Cerny (2016), concluindo que o GPI fornece maior cobertura de dados e maior diversidade de saída de dados em relação ao Patentscope. Por outro lado, o autor aponta que o Patentscope é uma fonte de dados mais completa para documentos PCT, porém com cobertura mais pobre de autoridades de patentes.

Jürgens e Clarke (2018) compararam o conjunto de características de seis bases de busca de patentes, com abrangência de diversos países e compreendendo apenas sua interface de acesso gratuito: Google Patents, Lens, Patent Inspiration, Free Patents Online, SureChEMBL

e Octimine. Eles concluíram que cada um tem suas limitações, sendo alguns mais adequados para pesquisas do estado da arte ou análise de patentes, e outros para pesquisa de patentes em campos específicos como biotecnologia e fármacos.

Guerreiro *et al.* (2018) compararam as bases de dados do Orbit e do Espacenet no que tange a patentes que referem copaiá, concluindo que o Orbit estava mais atualizado.

A maioria dos trabalhos da literatura tem comparado os sistemas de busca de patentes gratuitos, sendo que a comparação entre sistemas gratuitos e privados tem sido pouco utilizada. Estudos que compararam de alguma forma sistemas gratuitos e privados – por exemplo, Stock e Stock (2006) – já têm mais de dez anos e, desse modo, não refletem os avanços realizados nos sistemas, principalmente nos gratuitos. Não somente os sistemas gratuitos, como todos os provedores de informações de patentes estão constantemente em aperfeiçoamento e atualização da sua cobertura e dos mecanismos de análise de estatísticas de patentes, portanto, este cenário sugere que há a necessidade de atualizar as pesquisas comparativas desses provedores. Desse modo, este trabalho analisa dois conjuntos diferentes de sistemas de informação: as bases gratuitas (Espacenet, Patentscope, Google Patents e Lens) e os sistemas de informações de patentes pertencentes a provedores privados (Derwent Innovation Index e Orbit Intelligence).

A pesquisa compara as características gerais dos sistemas de busca de patentes associadas ao período de cobertura e ao número de autoridades de patentes coberta, número de documentos e famílias de patentes, quantidade de campos aplicados à pesquisa, tipos de classificação de patentes pesquisáveis, idioma dos termos pesquisáveis, literatura não patentária e ferramentas de análise estatística de patentes. Adicionalmente, são comparados os resultados recuperados em cada sistema, utilizando um código IPC relevante de cada seção. Portanto, busca-se apresentar não somente uma visão geral dos sistemas, mas também identificar os mais eficientes na procura de tecnologias patenteadas nos IPCs mais relevantes.

2 Metodologia

Para realizar a comparação dos sistemas de busca de patentes, de acesso gratuito ou pago, que contemplem em seu banco de dados informações multinacionais (de mais de uma autoridade de patentes) de patentes e que seja caracterizado por prospectar documentos de patentes de diferentes áreas tecnológicas, os seguintes sistemas de busca foram identificados e considerados para comparação:

- a) Patentscope;
- b) Espacenet;
- c) Google Patents;
- d) Lens;
- e) Derwent Innovation Index; e
- f) Orbit Intelligence.

Os sistemas foram comparados pelas suas características gerais e pelo desempenho nos resultados de uma pesquisa para identificar documentos de patentes de tecnologias específicas. Em características gerais, foram analisados:

- a) tipo de acesso;
- b) número de autoridades de patentes coberta;
- c) período de cobertura;
- d) tamanho da base;
- e) ano de início da cobertura;
- f) número de campos de busca;
- g) tipos de classificação de patentes que pode ser utilizada na busca;
- h) exportação de dados;
- i) formato de exportação de arquivos;
- j) idioma dos termos pesquisáveis;
- k) literatura não patentária;
- l) ferramenta de análise rápida de estatísticas de patentes; e
- m) ferramenta de análise aprofundada de estatísticas de patentes.

Para comparar os resultados de pesquisa dos sistemas investigados, utilizou-se uma tecnologia específica de cada uma das seções de A a H da IPC. A seleção das IPCs considerou em cada seção apenas um, entre os primeiros códigos, que seja composto de um subgrupo que não tenha subgrupos secundários, de modo a contornar a dificuldade de algumas bases de patentes, por exemplo, Patentscope, apresentarem os resultados de pesquisa por IPC com documentos classificados nos subgrupos principais e secundários, evitando, assim, que resultados fossem enganadores quando comparados com outros sistemas de busca que consideram apenas o código específico inserido como os termos de busca. Os códigos IPCs selecionados para pesquisa foram:

- A01B3/02 – Arados de tração humana com relhas fixas;
- B01B1/02 – Processos ou aparelhos físicos ou químicos para prevenção de espumas;
- C01B5/02 – Água pesada; Preparação pela reação química de isótopos de hidrogênio ou seus compostos, por exemplo, $4ND3+7O2 \rightarrow 4NO2+6D2O$, $2D2+O2 \rightarrow 2D2O$;
- D01B1/02 – Separação mecânica de fibras vegetais das sementes;
- E01B5/16 – Distanciadores de trilhos;
- F01B1/01 – Máquinas ou motores de pistão de movimento alternado, caracterizados por serem constituídos por um único pistão;
- G01B3/11 – Correntes para medição de comprimentos;
- H01B5/06 – Condutores ou corpos condutores não isolados, caracterizados pela forma de tubo simples.

Os códigos IPCs foram utilizados, individualmente, para pesquisar em cada sistema de patentes, a fim de comparar o número de famílias e/ou documentos de patentes resgatados em cada um dos sistemas. A pesquisa de cada código foi realizada em todos os sistemas, no mesmo dia, para evitar variações nos resultados, ocasionadas por atualizações nos bancos de dados dos sistemas investigados.

3 Resultados e Discussão

São apresentadas as principais características dos seis sistemas de busca de patentes investigadas nesta pesquisa. Em seguida, são discutidos os resultados de uma pesquisa comparativa realizada com um código IPC de cada seção da classificação. A última seção dos resultados compara os mecanismos de análise estatística de patente das bases que possuem alguma ferramenta de análise.

3.1 Características Gerais dos Sistemas de Busca de Patentes

a) Patentscope (<http://patentscope.wipo.int/>)

O Patentscope é um sistema gratuito de pesquisa de patentes lançado em 2005 pela Organização Mundial da Propriedade Intelectual (WIPO). Nesse sistema, em 21 fevereiro de 2018, já era possível acessar 3.241.552 documentos do Tratado de Cooperação em termos de Patentes – *Patent Cooperation Treaty* (PCT) –, depositados a partir de 1976, e 64.070.406 documentos de outras 53 autoridades nacionais e regionais de patentes, com informações bibliográficas de documentos a partir de 1782 (WIPO, 2018a). A interface do sistema é disponibilizada em dez idiomas, facilitando a utilização do mecanismo de busca. O sistema de busca das patentes possui quatro modalidades de pesquisa: Pesquisa Simples, Pesquisa Avançada, Combinação de Campos, Expansão Multilíngue e Compostos Químicos. Possui 58 campos de busca e os termos ou palavras-chave podem ser pesquisados em 12 línguas. A pesquisa por classificação de patentes só é possível pela IPC. Depois do usuário criar uma conta, é possível baixar um arquivo no formato XLS com informações bibliográficas (número de publicação, data de publicação, título, data de prioridade, códigos IPC, depositantes e inventores) de até 10 mil documentos de patentes. Na página de visualização dos resultados, há apenas a possibilidade de análise rápida de estatísticas de patentes, com recursos que mostram no formato de tabela ou gráfico os principais países, códigos IPC, inventores, depositantes e data de publicação. No entanto, os dados estatísticos não estão disponíveis para exportação.

b) Espacenet (<http://worldwide.espacenet.com/>)

O Espacenet é uma plataforma *on-line* de acesso livre que permite pesquisar em mais 107,498 milhões de documentos de patentes, com informações bibliográficas a partir de 1782, de 105 autoridades de patentes de todo o mundo (EPO, 2018). O Espacenet foi lançado em 1998 e é disponibilizado pelo Escritório Europeu de Patentes (European Patent Office, EPO), juntamente com a Comissão Europeia e vários escritórios nacionais e regionais de patentes. São encontradas duas possibilidades de pesquisa: Pesquisa Simples e Pesquisa Avançada. Ao todo são 12 campos de busca e a pesquisa por classificação de patentes pode ser feita tanto pela IPC como pela CPC. É possível exportar os dados bibliográficos de até 500 documentos

de patentes, nos formatos XLS e CSV. Não há nenhuma ferramenta de análise estatística de patentes. Os documentos encontrados em pesquisas do Espacenet possuem informação do IMPADOC – banco de dados da EPO com informação do *status* jurídico e famílias de documentos de patentes de 80 autoridades de patentes do mundo (EPO, 2018). Além das patentes, é encontrada literatura não patentária, mostrando informações sobre os documentos que foram citados ou estão citando o documento analisado. Ademais, é possível pesquisar em texto completo em Inglês, Francês e Alemão.

c) Google Patents (<https://patents.google.com/>)

O sistema de busca de patentes do Google foi lançado pela empresa em 2006. Contém mais de 87 milhões de documentos de patentes de 17 autoridades de patentes com informações bibliográficas a partir de 1790, além de documentos técnicos e livros indexados no Google Scholar e Google Books (literatura não patentária). Na pesquisa simples e pesquisa avançada, os termos de busca podem ser inseridos em até 11 campos, com possibilidade de utilizar as classificações IPC, CPC e Classificação Americana de Patentes – *United States Patent Classification (USPC)*. Os resultados de busca apresentam apenas o primeiro documento de uma família simples de patentes, ou seja, os outros documentos de uma mesma família de patentes são removidos da lista de resultados. A maioria dos documentos (16 autoridades de patentes) tem seus textos traduzidos para o inglês e a descrição completa de um documento de patente é apresentada em HTML no próprio navegador, facilitando a leitura das informações e a visualização dos desenhos mostrados como imagem. Informações bibliográficas dos primeiros 1.000 resultados podem ser transferidos no formato CSV, e esse sistema possui uma ferramenta muito limitada de análise de resultados com apenas três categorias de análises que consideram apenas os 1.000 principais resultados (*Top 1000 results*) com base em data do pedido, depositante, inventor e código CPC.

d) Lens (<https://www.lens.org>)

O Lens foi lançado nos anos 2000 como uma iniciativa da Cambia, organização não governamental sediada em Camberra, Austrália, em parceria com a Queensland University of Technology. Em 25 de fevereiro de 2018, havia na sua base de dados 58,002 milhões de famílias de patentes e 106,89 milhões de documentos individuais, de 104 jurisdições de patentes ao redor do mundo. Seus dados de patentes são provenientes do Escritório Europeu de Patentes, do Escritório Americano de Patentes, do Escritório Australiano de Patentes e dos documentos relacionados ao Tratado de Cooperação em termos de Patentes da WIPO, com informações bibliográficas a partir de 1907. Também é encontrada literatura não patentária no Lens, com dados provenientes de PubMed e Crossref. Com uma interface em cinco idiomas, são disponibilizadas quatro modalidades de pesquisa – pesquisa simples, pesquisa estruturada, pesquisa pela classificação CPC, IPC ou USPC, pesquisa Patcite (procura publicações científicas citadas por patentes ou patentes citadas na literatura acadêmica) e pesquisa QUT In4M (que apresenta dados das 200 principais instituições de pesquisa que influenciam na indústria e na inovação). Esse sistema permite pesquisar documentos em inglês, francês, chinês, espanhol e russo. Os resultados podem ser baixados em quatro formatos (CSV, RIS, BibTeX, JSON), mas o Lens só permite baixar informações bibliográficas de até 1.000 documentos. Possui uma ferramenta de análise rápida dos dados, com 13 gráficos de estatísticas de patentes predefinidos. No entanto, esses dados estatísticos não podem ser exportados para análise posterior.

e) Derwent Innovation Index (<http://apps.webofknowledge.com>)

A Derwent Innovation Index é um provedor privado de Propriedade Intelectual fornecido pela Clarivate Analytics. Alimentado pela base *Derwent World Patents Index (DWPI)*, são disponibilizados 35 milhões de famílias de patentes e 74 milhões de documentos de patentes, com data de publicação a partir de 1963 e provenientes de 52 autoridades de patentes. Há três modalidades de pesquisa: pesquisa simples, pesquisa de patente citada e pesquisa avançada, e a pesquisa pode ser realizada em 20 diferentes campos. A pesquisa por classificação de patentes pode ser realizada pela IPC ou pela classificação Derwent (uma reclassificação das patentes feita pelos especialistas da Derwent, distribuindo os documentos em 20 áreas do conhecimento, nas seções de Química, Engenharia e Elétrica e Eletrônica) (CLARIVATE ANALYTICS, 2018). Os resumos das patentes de 30 idiomas diferentes são traduzidos para o inglês. Também é realizado aprimoramento no título e são elaborados resumos tecnológicos em inglês. Informações sobre o *status* legal são provenientes da DWPI e da INPADOC. Podem ser exportados até 500 documentos por vez, nos formatos TXT e CSV. A literatura não patentária disponível neste sistema é proveniente de *Web of Science, Institution of Engineering and Technology*. Como possibilidade de visualização e análise estatística de patentes, são disponibilizados mapas *ThemeScape*, agrupamento de texto em conjuntos relacionados, gráficos predefinidos e personalizados e mapas de citações de patentes.

f) Orbit Intelligence (<https://www.orbit.com>)

O Orbit Intelligence é uma plataforma privada fornecida pela Questel Co. Em 25 de fevereiro de 2018, o sistema já contava com 58,185 milhões de famílias de patentes e mais de 107 milhões de documentos individuais de patentes, fornecidas por 107 jurisdições de patentes. Existem informações bibliográficas a partir de 1827. A pesquisa pode ser realizada em mais de 100 campos, formados em várias modalidades: pesquisa fácil, pesquisa avançada, pesquisa semântica, pesquisa por requerente e pesquisa por número da patente. Quanto às classificações pesquisáveis, podem ser utilizadas a IPC, CPC, USPC, ECLA, o sistema japonês de classificação de patentes (*File Index*, FI) e a classificação somente por computador – *Computer Only Classification (ICO)*. Os termos podem ser pesquisados em quatro idiomas: inglês, francês, alemão e espanhol. Também apresenta literatura não patentária, mostrando documentos técnicos não patenteados que citam ou foram citados por patentes. Possui uma ferramenta de análise estatística de patentes, com gráficos predefinidos e uma diversidade de possibilidades de configuração de dados para análise, limitando-se a analisar até 2 milhões de documentos. Os dados podem ser extraídos em diversos formatos (TXT, PDF, CSV, XLS, XLSX, XML), abrangendo até 20 mil famílias de patentes por vez. Outros pontos a serem destacados é a cobertura de textos integrais disponibilizados pelo Orbit, ou seja, é possível realizar a pesquisa nos documentos completos e não só no título e no resumo; a tradução para inglês de múltiplos idiomas (com especial destaque para documentos em chinês e japonês, muitos dos quais não têm depósitos em outros países e, portanto, não acessíveis mediante família de patentes); a capacidade de análise e a presença de dados complementares sobre litígio (sobretudo por que atualmente os dados da DartsIP estão disponíveis no Orbit). Essas características tornam o Orbit adequado para uso empresarial, governamental e acadêmico, inclusive sendo o mais adequado para programas de pós-graduação.

Um resumo comparativo desses sistemas foi elaborado para identificar suas principais características, conforme mostra o Quadro 1.

Quadro 1 – Comparação das características dos principais sistemas de busca de patentes compreendendo tipo de acesso, cobertura, tamanho da base, início da cobertura temporal, campos de busca, buscas por classificações de patentes, exportação de dados, formato de exportação de dados, língua de utilização, língua dos documentos de patentes, literatura não patentária, ferramentas de análise rápida dos resultados e ferramentas de análise aprofundada dos resultados

DESCRIÇÃO	PATENTSCOPE	ESPCENET	GOOGLE PATENT	LENS	DERWENT INNOVATION INDEX	ORBIT
ACESSO	Gratuito	Gratuito	Gratuito	Gratuito	Pago	Pago
COBERTURA	54	105	17	104	52	107
TAMANHO DA BASE	3,241 milhões de PCTs; Mais de 64 milhões de publicações individuais	Mais de 107,498 milhões de publicação individuais	Cerca de 87 milhões de publicações individuais	58,002 milhões de famílias de patentes; 106,89 milhões de publicações individuais	35 milhões de famílias de patentes; 74 milhões de publicações individuais	58,185 milhões de famílias de patentes; mais de 107 milhões de publicações individuais
INÍCIO DA COBERTURA TEMPORAL	1782	1782	1790	1907	1963	1782
CAMPOS DE BUSCA	58	12	11	20	20	Mais de 100
BUSCAS POR CLASSIFICAÇÕES DE PATENTES	IPC	IPC, CPC	IPC, CPC, USPC	IPC, CPC, USPC	IPC e Código Derwent	IPC, CPC, ECLA, JP (FI), USPC
EXPORTAÇÃO DE DADOS	Limitado aos primeiros 10.000 documentos	Limitado aos primeiros 500 documentos	Limitado aos primeiros 1.000 documentos	Limitado aos primeiros 1.000 documentos	Limitado a 500 famílias por vez	Limitado a 20.000 famílias por vez
FORMATO DE EXPORTAÇÃO	XLS	XLS, CSV	CSV	CSV, RIS, BibTeX, JSON	TXT, CSV	TXT, PDF, CSV, XLS, XLSX, XML,
LÍNGUA DE USO	Chinês, Japonês, Alemão, Coreano, Inglês, Português, Frances, Russo, Espanhol, Holandês, italiano e Suíço	Inglês	Inglês	Inglês, Francês, Russo, Chinês e Espanhol	Inglês	Inglês, Francês, Alemão e Espanhol
LÍNGUA DOS DOCUMENTOS	Língua original	Língua original	Todos os documentos traduzidos para Inglês com o Google Translate	Língua original	Língua original	Todos os documentos traduzidos para Inglês

DESCRIÇÃO	PATENTSCOPE	ESPACENET	GOOGLE PATENT	LENS	DERWENT INNOVATION INDEX	ORBIT
LITERATURA NÃO PATENTÁRIA (NPL)	Não	Sim (apenas citações)	Sim	Sim	Sim (apenas citações)	Sim (apenas citações)
FERRAMENTAS DE ANÁLISE RÁPIDA DOS RESULTADOS	Sim	Não	Sim	Sim	Sim	Sim
FERRAMENTAS DE ANÁLISE APROFUNDADA DOS RESULTADOS	Não	Não	Não	Não	Não	Sim

Fonte: Elaborado pelos autores deste artigo (2019)

Observa-se que, entre os sistemas gratuitos de informações de patentes, o Patentscope é o que tem mais campos de pesquisa e sua interface é mais acessível pela quantidade de idiomas que é disponibilizada e também pela utilização da busca por expansão multilíngue, que identifica o termo inserido em uma pesquisa em 12 idiomas diferentes. Outro aspecto positivo do Patentscope é o número de documentos que podem ter informações estatísticas de patentes transferidas em Excel, entretanto, esse grande número de documentos pode ser reflexo do fato de que o Patentscope não remove automaticamente os documentos que compõem famílias de patentes, desse modo contaminando os resultados, requerendo limpezas extras de dados e levando a potenciais erros nas conclusões.

O Espacenet se destaca pelo número de autoridades de patentes (105) e pelo número de documentos de patentes (aproximadamente 107,5 milhões) (EPO, 2018). Outro fator importante no Espacenet (também no Patentscope que utiliza dados da EPO) é a data de início da cobertura dos dados bibliográficos de patentes, que começa em 1782 com dados de patentes do Reino Unido, portanto, a mais antiga de todas as bases gratuitas investigadas. Vale ressaltar que a base de dados do Espacenet é frequentemente utilizada em outros sistemas de busca gratuitos e privados, por exemplo, todas as bases que estão sendo investigadas nesta pesquisa.

O Google Patents se destaca pela facilidade de pesquisar documentos anteriores ao estado da arte, a partir do *click* em um botão (*Find Prior Art*) na tela de visualização de uma patente, de modo que são identificadas tanto patentes anteriores requeridas e literatura não patentária indexadas no *Google Books* e *Google Scholar* e classificadas por máquina, usando a classificação CPC. Outra vantagem é a tradução do texto completo de patentes de outros idiomas para o inglês, tornando possível pesquisar em texto completo. Esse sistema destaca-se também por apresentar os resultados considerando apenas um documento de uma família de patentes simples, evitando duplicações.

O Lens se destaca como uma importante ferramenta que contempla os documentos individuais ou agrupados por famílias de patentes, permitindo ao usuário escolher como os resultados são apresentados. Com a possibilidade de pesquisar em mais de 30 milhões de dados de literatura não patentária, no Lens, é possível ter uma noção dos documentos não patenteados que mais influenciam no desenvolvimento de novos produtos. Outra característica desse sistema de busca de patentes é sua ferramenta de análise estatística de dados de patentes, com 13 gráficos prede-

finidos, o que o classifica como o mais completo entre todos os sistemas gratuitos investigados nesta pesquisa para análises rápidas de estatísticas de patentes. Entretanto, é uma ferramenta com limitações em termos de facilidade de utilização de campos de busca, a atualização de documentos é menos frequente do que a da maioria das ferramentas pagas, além das restrições em termos de exportação de dados também em relação às ferramentas pagas. Ainda sobre o Lens, utilizando o PatSeq, encontra-se um conjunto de ferramentas de sequências biológicas divulgadas em documentos de patentes de diversas autoridades do mundo. Caracterizado como o maior banco de dados público com métricas internacionais de sequências biológicas, Lens possui cinco ferramentas para analisar as sequências biológicas encontradas em documentos de patentes: PatSeq Data; PatSeq Explorer; PatSeq Analyzer; PatSeq Finder e; PatSeq Text Search.

A Derwent Innovation Index tem como características principais o trabalho realizado para aprimorar os títulos das patentes e os resumos de muitos documentos, criando um resumo tecnológico, o que facilita a identificação e análise de tecnologias específicas. Outra característica da Derwent é a identificação das patentes pela sua classificação específica (classificação Derwent), podendo ser utilizada, pelos usuários familiarizados com esta classificação, para identificar mais facilmente as patentes de uma determinada área do conhecimento ou da tecnologia. No entanto, hoje em dia, busca-se a internacionalização e, assim, deve-se utilizar a IPC ou a CPC e não uma classificação específica para cada base.

O Orbit Intelligence foi o sistema com o maior número de autoridades de patentes (107) e a maior quantidade de família de patentes (58,185 milhões). O sistema se destaca pela diversidade de campos pesquisáveis que podem ser combinados numa busca de patente. Também é o que apresenta maior diversidade de classificação de patentes (IPC, CPC, ECLA, JP-FI, USPC). Além disso, é possível pesquisar por domínio tecnológico que distribui as patentes em 35 campos tecnológicos, de acordo com a classificação da tabela de concordância de tecnologia da IPC (WIPO IPC *Technology Concordance Table*). O Orbit destaca-se ainda pelos recursos de análise rápida, com mais de 60 gráficos predefinidos, e de análise personalizada para apresentação de estatísticas de patentes, podendo analisar até 2 milhões de famílias de patentes por vez. Outro fator importante é a diversidade de formatos de arquivos que podem ser baixados e o fato de poder extrair informações de até 20 mil famílias de patentes por vez.

3.2 Comparação dos Sistemas de Busca de Patentes

As patentes podem ser consultadas obtendo cada depósito de patente em cada país, ou buscando famílias de patentes (EPO, 2017b) que consistem em depósitos de patentes de um mesmo conteúdo tecnológico, apenas com a diferença das solicitações terem sido em países diversos.

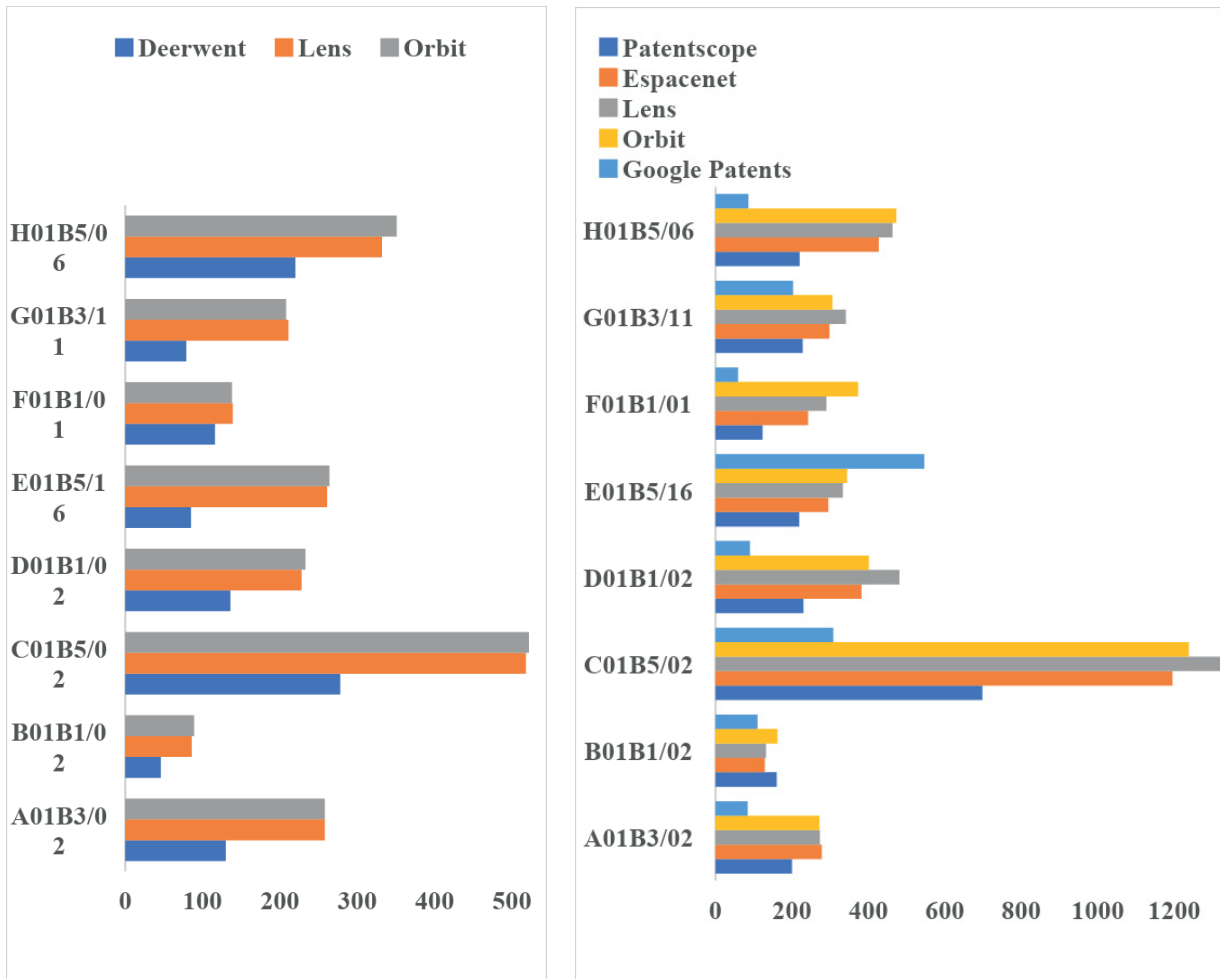
Desse modo, foram realizadas duas buscas, uma por documentos de famílias de patentes, que removem automaticamente as duplicidades, feita por Derwent, Lens e Orbit (Gráfico 1, figura da esquerda), e outra das que fornecem os documentos de patentes individuais, feita por Espacenet, Google Patents, Lens, Orbit e Patentscope (Gráfico 1, figura da direita). A busca em todos os sistemas foi realizada no mesmo dia e utilizando o mesmo critério de pesquisa, um código IPC de cada seção.

Na figura à esquerda do Gráfico 1, são apresentados os resultados da busca de Famílias de patente nos sistemas da Derwent, Lens e Orbit. Quando analisados os resultados pelas famílias de patentes, observa-se que Lens obteve o maior número de documentos em dois códigos (F01B1/01 e G01B3/11). O Orbit conseguiu o maior número de documentos em três códigos (C01B5/02, D01B1/02 e H01B5/06). No código A01B3/02, Orbit e Lens ficaram empatados com o melhor resultado de recuperação de documentos. A diferença entre os resultados de Lens e Orbit é muito pequena, de modo que em sete códigos a variação não foi maior que cinco famílias de patente e apenas no código H01B5/06 a diferença foi maior (19 famílias de patente). Se somarmos os resultados de cada sistema, será possível observar que Orbit e Lens obtiveram os melhores desempenhos, com 2.063 e 2.033 famílias de patentes, respectivamente. A Derwent resgatou quase 50% menos famílias de patentes que as outras duas bases (1.090).

A comparação dos sistemas que apresentam nos resultados da busca documentos individuais de patente é apresentada na figura à direita do Gráfico 1. Nos resultados por documentos individuais de patentes, o Orbit teve melhor desempenho em três códigos (B01B1/02, F01B1/01 e H01B5/06). Já Lens foi mais eficiente em três códigos (C01B5/02, D01B1/02 e G01B3/11). Por fim, o Espacenet obteve o maior número de documentos no código A01B3/02 e Google Patent no código E01B5/16. Observa-se que o Patentscope só não obteve o resultado mais baixo no código B01B1/02. O leitor se sentiria tentado a somar os resultados e concluir que o sistema Lens seria o mais eficiente, seguido pelo Orbit (3.569), no entanto, os resultados são contaminados, pois misturam dados patentários com literatura não patentária, o que explica a discrepância no número de documentos e leva à necessidade de análise caso a caso. Observa-se que ambos têm campos de pesquisa dedicados especialmente à pesquisa de literatura não patenteada, ficando com resultados semelhantes. Espacenet e Patentscope apresentam recuperação de documentos inferior, com 3.246 e 2.078, respectivamente.

De modo geral, pode-se afirmar que no que se refere aos buscadores gratuitos, com as limitações que são próprias das ferramentas gratuitas, o Lens se apresenta como uma boa opção para realização de pesquisas, tanto pelo número de documentos identificados, quanto pela possibilidade de apresentar os resultados em famílias de patentes ou documentos individuais de patentes, bem como literatura não patentária. Mas, comparando os buscadores pagos, o Orbit obteve melhores resultados em relação ao seu acesso gratuito e ao Derwent. Comparando com a Derwent, há ainda a vantagem de exportar os resultados como documentos de patentes individuais ou aglomerados por famílias de patentes, enquanto a Derwent somente possibilita a aglomeração por famílias de patentes.

Gráfico 1 – Comparação da busca de oito tecnologias, a partir da Classificação Internacional de Patentes (IPC), nos sistemas Derwent, Lens, Orbit, Espacenet e Patentscope; Esquerda: Famílias de patentes; Direita: Patentes



Fonte: Elaborado pelos autores deste artigo (2019)

Considerando os resultados obtidos por todos os sistemas investigados, o Lens se destaca entre as ferramentas gratuitas. No entanto, Orbit apresentou os melhores resultados entre todos os analisados por este artigo, tendo ainda como atrativos a diversidade de campos pesquisáveis e possíveis de serem combinados, a diversidade de classificação de patentes (IPC, CPC, ECLA, JP-FI, USPC), seus recursos de análise rápida e de análise personalizada, o número de famílias de patentes que podem ser analisadas por vez e a diversidade de formatos de arquivos que podem ser baixados por vez.

4 Considerações Finais

Os sistemas de busca de patentes, mesmo os sistemas de acesso gratuito, têm passado por melhorias e inserido ferramentas de busca e análise de informações de patentes cada vez mais relevantes para sua utilização pelos usuários, ao comparar a bibliografia anterior sobre o tema. Essas melhorias permitem aos diversos profissionais identificar, analisar e comparar informações contidas em documentos de patentes e em documentos técnicos citados por patentes. Dessa forma, tem se tornado frequente a utilização dessas informações para criar indicadores

econômicos e tecnológicos e também para desenvolver conhecimento científico e tecnológico cada vez mais aprimorado.

De fato, os resultados mostraram que muitas das otimizações identificadas por Stock e Stock (2006) como melhorias nos sistemas de busca de patentes têm sido realizadas. Por exemplo, a expansão nos dados de cobertura de documentos de patentes e de autoridades de patentes, a disponibilização de informações relacionadas a citações de patentes e a inclusão cada vez maior de literatura não patentária. Por outro lado, avanços relacionados a ferramentas de visualização gráfica de estatísticas de patentes e a possibilidade de exportar dados de pesquisas com quantidades, em geral maiores que 1.000 documentos de patentes, precisam ser aprimoradas em quase todos os sistemas investigados.

Os resultados da comparação de seis sistemas de busca de patentes e com características similares demonstram que a distância entre sistemas de patentes gratuitos e pagos está diminuindo, apesar de os sistemas pagos ainda apresentarem as facilidades de busca, de tratamento e de exportação de dados muito superiores às das ferramentas gratuitas.

A análise das características gerais dos sistemas investigados ressalta o aprimoramento de ferramentas de pesquisa que facilitam quebrar a barreira do idioma e identificar documentos de patentes em diversas línguas, por exemplo, a pesquisa por Expansão Multilíngue do Patentscope e a tradução para inglês realizada pelo Orbit em documentos em diversos idiomas, o que é especialmente relevante para documentos em chinês e japonês, muitos dos quais não têm depósitos em outros países e, portanto, não acessíveis mediante família de patentes. Outro exemplo é a capacidade de a ferramenta de pesquisa do Orbit e do Google Patents pesquisar na língua inglesa pelo fato de os documentos terem sido traduzidos por máquina para o inglês, no entanto, o Google Patents somente busca em 16 autoridades de patentes e o Orbit cobre todas as autoridades da base do EPO.

Assim, o Orbit foi o que apresentou melhor desempenho em todos os quesitos analisados, maior amplitude de base de dados e mais ferramentas para processamento dos dados, permitindo exportar em figura ou em dados brutos e sendo compatível com outros sistemas. Essas características fazem dele o *software* mais adequado para os setores empresarial, governamental e acadêmico, sendo especialmente adequado aos programas de pós-graduação por sua flexibilidade e amplitude.

Referências

CARRARA, P.; RUSSO, D. Patent searches opinion: How to minimize the risk when reviewing patent applications. **World Patent Information**, [S.l.], v. 49, p. 43-51, 2017. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0172219016301016>. Acesso em: 10 fev. 2018.

CERNY, J. Open Access Intellectual Property Systems: A Comparison to Commercial Solutions in Competitive Intelligence. **Journal of Systems Integration**, [S.l.], v. 7, n. 3, p. 16, 2016. Disponível em: <https://search.proquest.com/openview/d1fad08f1d4925719733103114e0c939/1?pq-origsite=gscholar&cbl=2032453>. Acesso em: 10 fev. 2018.

CLARIVATE ANALYTICS. **Derwent Innovation Index Ajuda**. 2018. Disponível em: https://images.webofknowledge.com/WOKRS512B4/help/pt_BR/DII/hs_international_patent_classification.html. Acesso em: 27 fev. 2018.

- EPO – EUROPEAN PATENT OFFICE. **Office Contents and coverage of the DOCDB bibliographic file**. 2018. Disponível em: <https://www.epo.org/searching-for-patents/data/coverage.html>. Acesso em: 27 fev. 2018.
- EPO – EUROPEAN PATENT OFFICE. **Patents families**. 2017a. Disponível em: <https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/patent-families.html>. Acesso em: 25 March 2019.
- EPO – EUROPEAN PATENT OFFICE. **Espacenet Resource Book 2.1**, 2017b. Disponível em: <https://www.epo.org/searching-for-patents/technical/espacenet.html#tab-1>. Acesso em: 28 jan. 2018.
- FRANÇA, R. O. Patente como fonte de informação tecnológica. **Perspectivas em Ciência da Informação**, [S.l.], v. 2, n. 2, p. 235-264, 1997. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/636>. Acesso em: 1º jun. 2018.
- GRILICHES, Z. Patent statistics as economic indicators: a survey. In: GRILICHES, Z. **R&D and productivity: the econometric evidence**. University of Chicago Press, 1998. p. 287-343. Disponível em: <http://www.nber.org/chapters/c8351.pdf>. Acesso em: 13 dez. 2017.
- GUERREIRO, E. S. *et al.* Análise de documentos de patentes sobre copaíba: uma comparação entre fontes de dados. **Cadernos de Prospecção**, Salvador, v. 11, n.1, p. 26-40, 2018. Disponível em: <https://portalseer.ufba.br/index.php/nit/article/view/22725>. Acesso em: 13 dez. 2017.
- JÜRGENS, B.; CLARKE, N. Study and comparison of the unique selling propositions (USPs) of free-to-use multinational patent search systems. **World Patent Information**, [S.l.], v. 52, p. 9-16, 2018. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0172219017300856>. Acesso em: 25 abr. 2018.
- JÜRGENS, B.; HERRERO-SOLANA, V. Espacenet, Patentscope and Depatisnet: a comparison approach. **World Patent Information**, [S.l.], v. 42, p. 4-12, 2015. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0172219015000460>. Acesso em: 22 fev. 2018.
- JÜRGENS, B.; HERRERO-SOLANA, V. Patents from Latin America and Spain with Latipat: Country coverage and ability to search for emerging topics like nanotechnology. **World Patent Information**, [S.l.], v. 46, p. 1-8, 2016. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0172219016300345>. Acesso em: 22 fev. 2018.
- KULP, C. S. Patent Databases. A Survey of What Is Available from DIALOG, Questel, SDC, Pergamon and INPADOC. **Database**, [S.l.], v. 7, n. 3, p. 56-72, 1984. Disponível em: <https://eric.ed.gov/?id=EJ304514>. Acesso em: 20 jan. 2018.
- STOCK, M.; STOCK, W. G. Intellectual property information: A comparative analysis of main information providers. **Journal of the Association for Information Science and Technology**, [S.l.], v. 57, n. 13, p. 1.794-1.803, 2006. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/asi.20498>. Acesso em: 27 jan. 2018.
- STOCK, M.; STOCK, W. G. Intellectual property information. A case study of Questel-Orbit. **Information Services & Use**, [S.l.], v. 25, n. 3-4, p. 163-180, 2005. Disponível em: <https://content.iospress.com/articles/information-services-and-use/isu469>. Acesso em: 27 jan. 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**, [S.l.], v. 18, n. 1, p. 106-125, mar. 2013. ISSN 19815344. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/844>. Acesso em: 2 jun. 2018.

WIPO – WORLD INTELLECTUAL PROPERTY ORGANIZATION. **Coleção nacional:** dados completos, 2018a. Disponível em: https://patentscope.wipo.int/search/pt/help/data_coverage.jsf. Acesso em: 28 jan. 2018.

WIPO – WORLD INTELLECTUAL PROPERTY ORGANIZATION. **Guide to the International Patent Classification**, 2018b. Disponível em: www.wipo.int/export/sites/www/classifications/ipc/en/guide/guide_ipc.pdf. Acesso em: 28 jan. 2018.

Sobre os Autores

Edilson Araújo Pires

E-mail: edilson@ufrb.edu.br

Doutorado em Ciência da Propriedade Intelectual pela Universidade Federal de Sergipe (UFS).

Endereço profissional: Universidade Federal do Recôncavo da Bahia, Rua Rui Barbosa, n. 710, Centro, Cruz das Almas, BA. CEP: 44380-000.

Nubia Moura Ribeiro

E-mail: nubiamouraribeiro@gmail.com

Doutorado em Química pela Universidade Federal do Rio de Janeiro (2004).

Endereço profissional: Instituto Federal da Bahia, Departamento de Química do Campus Simões Filho, Rua Emidio dos Santos, s/n, Barbalho, Salvador, BA. CEP: 40301-015.

Cristina M. Quintella

E-mail: cris5000tina@gmail.com

Doutorado interdisciplinar em Ciências Moleculares pela University of Sussex, UK (1993).

Endereço profissional: Universidade Federal da Bahia, Instituto de Química, Departamento de Química Geral e Inorgânica, Campus de Ondina, Ondina, Salvador, BA. CEP: 40170-290.