

INDICADORES TECNOLÓGICOS: ESTRATÉGIA DE BUSCA DE DOCUMENTOS DE PATENTES RELACIONADOS À INSTRUMENTAÇÃO APLICADA AO AGRONEGÓCIO

Leandro Innocentini Lopes de Faria¹

Nayara Cristini Bessi²

Douglas Henrique Milanez³

RESUMO

Indicadores elaborados com base em documentos de patentes são úteis para medir e mapear aspectos do desenvolvimento de tecnologias. A qualidade desses indicadores depende fortemente da qualidade dos dados recuperados em bases de dados de patentes. A presente pesquisa tem por objetivo apresentar uma estratégia de busca de documentos de patentes relacionados a instrumentos utilizados no agronegócio e o processo de elaboração desta. Tal processo consistiu na escolha de uma base de dados e no estabelecimento de uma estratégia inicial para agronegócio e instrumentação, baseando-se em dados provenientes da literatura. Em seguida, aperfeiçoamentos na estratégia inicial foram aplicados em ambos os contextos (agronegócio e instrumentação), e as melhorias foram avaliadas por meio de cálculos de revocação e precisão. A estratégia de busca final mostra a complexidade para o estabelecimento de termos e códigos de classificação que descrevam suficientemente o ambiente tecnológico do agronegócio e da instrumentação. O procedimento desenvolvido pode também ser aplicado na elaboração de estratégias de busca em outras temáticas, para outras necessidades.

Termos para a indexação: busca de patentes, informação tecnológica, instrumentos agroindustriais.

TECHNOLOGICAL INDICATORS: A SEARCH STRATEGY FOR PATENT DOCUMENTS RELATED TO INSTRUMENTATION APPLIED TO AGRIBUSINESS

ABSTRACT

Indicators based on patent documents are useful for measuring and mapping aspects of technological development. The quality of these indicators relies heavily on the quality of

¹ Engenheiro de materiais, Doutor em Ciências e Engenharia dos Materiais, Doutor em Ciência da Informação e da Comunicação, professor adjunto da Universidade Federal de São Carlos (UFSCar), vinculado ao Departamento de Ciência da Informação e ao Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade, coordenador do Núcleo de Informação Tecnológica em Materiais da UFSCar (NIT/Materiais), Rodovia Washington Luís, Km 235, Caixa Postal 676, CEP 13565-905 São Carlos, SP. leandro@nit.ufscar.br

² Cientista da informação e da documentação, mestranda do Programa de Pós-Graduação em Ciência, Tecnologia e Sociedade da UFSCar, pesquisadora integrante do NIT/Materiais, atual bolsista Capes, Rodovia Washington Luís, Km 235, Caixa Postal 676, CEP 13565-905 São Carlos, SP. nayara@nit.ufscar.br

³ Engenheiro de materiais, doutorando do Programa de Pós-Graduação em Ciências e Engenharia dos Materiais pela UFSCar, pesquisador integrante do NIT/Materiais, atual bolsista CNPq, Rodovia Washington Luís, Km 235, Caixa Postal 676, CEP 13565-905 São Carlos, SP. douglas@nit.ufscar.br

data retrieved from patent databases. This research aims to present a search strategy for patent documents related to instrumentation applied to agribusiness and its drafting process. The developed process consisted in choosing a database and establishing an initial strategy for agribusiness and instrumentation, based on data from literature. Then, improvements in the initial strategy were implemented in both contexts (agribusiness and instrumentation) and were evaluated by calculation of recall and precision. The final search strategy shows the complexity in establishing terms and classification codes that sufficiently describe the technological environment of agribusiness and instrumentation. The developed process can also be applied in the development of search strategies in other topics, for other needs.

Index terms: patent search, technological information, agribusiness instruments.

INTRODUÇÃO

Indicadores de ciência e tecnologia (C&T) têm sido recorrentemente utilizados como forma de sintetizar, mensurar e avaliar a dinâmica de pesquisas científicas e desenvolvimentos tecnológicos, que são cada vez mais complexos e envoltos de grande quantidade de informação. Um indicador de C&T pode ser compreendido como dado estatístico que visa medir bens intangíveis, fornecendo panoramas, comparações e tendências sobre o avanço científico e tecnológico de uma dada temática para subsidiar a tomada de decisão e o planejamento em políticas públicas ou estratégias organizacionais (MUGNAINI et al., 2004). O documento de patente é um dos principais insumos para a construção de indicadores tecnológicos, pois, além de ser um título de propriedade, fornece informações de natureza técnica, legal e de negócio sobre a tecnologia patenteada. Quanto a isso, indicadores tecnológicos são elementos essenciais para analisar a concorrência, acompanhar as tendências tecnológicas, realizar o mapeamento de tecnologias e analisar outras perspectivas (ASHTON; KLAVANS, 1997).

Uma das variáveis que influenciam a qualidade do indicador é a estratégia de busca empregada na recuperação do conjunto de dados a ser analisado quantitativamente. Embora nenhuma busca seja exaustiva, a elaboração da estratégia de busca é um dos fatores críticos na construção de qualquer indicador, pois uma amostra ideal deve ser ampla para descrever suficientemente o assunto de interesse e, ao mesmo tempo, conter o mínimo de registros irrelevantes, que podem distorcer a análise final. A presente pesquisa teve por objetivo apresentar uma estratégia de busca – e o processo de elaboração desta – para a recuperação de documentos de patentes relacionados à instrumentação utilizada

no agronegócio, que fosse representativa e pudesse ser utilizada na elaboração de indicadores tecnológicos.

No Brasil, o agronegócio possui um papel preponderante para o desenvolvimento socioeconômico e ambiental, é responsável por parcela considerável do PIB brasileiro e é um dos principais fatores que fornecem vantagem competitiva ao país, havendo necessidade constante de estudos que aumentem as percepções sobre aspectos desse contexto. Este artigo traz contribuições de caráter metodológico em duas perspectivas: a) em relação ao procedimento adotado, pois ele pode ser utilizado para a elaboração de estratégias de busca de documentos de patentes em outras temáticas; e b) em relação à utilização da estratégia desenvolvida, pois ela pode ser aplicada por outras pesquisas que objetivarem compreender o avanço do conhecimento no que tange ao desenvolvimento de maquinários, dispositivos, sensores ou instrumentos utilizados em processos agroindustriais e na agropecuária.

Na seção “Indicadores tecnológicos e o agronegócio”, são expostas, de modo geral, relações teóricas sobre indicadores tecnológicos e pesquisas que os desenvolveram no contexto do agronegócio. A próxima seção, “Recuperação da informação de patentes”, menciona aspectos relacionados ao processo de busca e recuperação de informações providas de documentos de patentes. Por fim, nas seções “Procedimento desenvolvido para elaboração da estratégia de busca” e “Aplicação no contexto do agronegócio e instrumentação”, são apresentados, respectivamente, uma síntese do processo metodológico desenvolvido, seguida da aplicação deste no contexto temático proposto pela presente pesquisa.

INDICADORES TECNOLÓGICOS E O AGRONEGÓCIO

A sociedade globalizada está inserida num ambiente extremamente competitivo, dinâmico e complexo, cuja alta oferta de informação desafia a compreensão desse meio. Para Senge (2010), pensar sistemicamente é mais necessário do que nunca, e o desafio tange ao fato de a humanidade ter capacidade de criar muito mais informações do que é capaz de absorver. Quanto a isso, ferramentas e metodologias que buscam sintetizar a grande quantidade de informações disponíveis na forma de indicadores que forneçam uma visão holística sobre a realidade e deem suporte a decisões e ao planejamento sustentável das atividades humanas se tornam importantes.

Os indicadores podem ser compreendidos como parâmetros que fornecem indicações ou descrições sobre o estado de um fenômeno (OCDE, 2002). Desde o aparecimento da estatística como ciência, os indicadores têm sido importantes para o entendimento de sistemas complexos, pelo fato de representarem um compromisso entre a busca pela exatidão científica, proporcionada pelo cálculo matemático, e a informação concisa e clara (GRUPO DE ORDENAMENTO DO TERRITÓRIO, 2012). Entretanto, segundo Santos e Kobashi (2005, p. 5) “[...] os indicadores quantitativos não representam uma ‘verdade’ sobre o estado da ciência e da tecnologia (C&T), mas são aproximações da realidade ou uma expressão incompleta dela”. Dessa forma, os conjuntos de dados estatísticos devem ser interpretados com certa cautela, levando-se em consideração os aspectos sociais, econômicos e políticos da sociedade ou da área de conhecimento que eles contextualizam.

Indicadores que mensuram aspectos relacionados à tecnologia se utilizam de documentos técnicos como insumos em seu processo de elaboração – entre eles estão os documentos de patentes. Esses documentos são uma importante fonte de informação tecnológica que pode ser compreendida como “todo tipo de conhecimento sobre tecnologias de fabricação, de projeto e de gestão que favoreça a melhoria contínua da qualidade e a inovação no setor produtivo” (UNIVERSIDADE DE BRASÍLIA, 2012). Entre os documentos técnicos, os mais utilizados na elaboração de indicadores tecnológicos são os documentos de patentes. Segundo a OCDE (2009), patentes são um meio de proteger invenções desenvolvidas por empresas, instituições ou indivíduos. A carta patente é o documento que confere ao seu titular ou depositante o direito de impedir terceiros de reproduzir ou comercializar a tecnologia desenvolvida sem seu consentimento.

Informações sobre os depósitos de patentes estão disponíveis em bases de dados de escritórios nacionais ou em bases comerciais, sendo importante conhecer as ferramentas de recuperação da informação que cada uma delas oferece (BARROSO et al., 2009). A elaboração de indicadores baseia-se na bibliometria, que pode ser entendida como: “[...] uma técnica quantitativa e estatística para medir índices de produção e disseminação do conhecimento, bem como acompanhar o desenvolvimento de diversas áreas científicas e os padrões de autoria, publicação e uso dos resultados de investigação” (LOPES et al., 2012, p. 1). Atualmente é recorrente o uso de softwares bibliométricos ou

de mineração de dados em virtude do grande volume de da grande quantidade de informação disponível nas bases de dados. Um exemplo de software para mineração de dados é o Vantage Point, que tem capacidade para separar os dados brutos obtidos de base de dados e realizar sua contagem bibliométrica (VANTAGE POINT, 2012).

No contexto de indicadores tecnológicos, Ramos (2012) avaliou o patenteamento mundial dentro da temática nanotecnologia aplicada ao agronegócio de 2000 a 2009, utilizando dados da base Derwent Innovations Index. Para recuperar os registros de patentes, o autor aplicou, inicialmente, uma estratégia de busca em nanotecnologia e, posteriormente, no agronegócio, adaptadas da literatura (MILANEZ, 2011; OBSERVATOIRE DES SCIENCES ET DES TECHNIQUES, 2008; PORTER et. al, 2008), tendo realizado a interseção de ambas com o auxílio do software Vantage Point. Como resultado, foi verificado o aumento no total mundial de patentes depositadas até 2008, seguido de uma desaceleração na taxa de crescimento e uma queda brusca do número de documentos em 2009, possivelmente em decorrência da crise mundial de 2008. O autor ainda identificou como principal atuante no assunto estudado a China, que ocupou a liderança do patenteamento a partir de 2007, seguida dos Estados Unidos, Japão e Coreia do Sul. Embora as empresas tenham contribuído significativamente para o desenvolvimento tecnológico de nanotecnologia aplicada ao agronegócio, instituições de pesquisa destacaram-se, como a chinesa Universidade de Zhejiang, que liderou o patenteamento no período 2000–2009, sugerindo a dependência de pesquisas básicas para o avanço do assunto.

Alves et al. (2006) desenvolveram indicadores para analisar a indústria da avicultura, tendo utilizado 3.104 documentos de patentes de 1976 a 2002, obtidos das bases do Escritório Americano de Patentes e Marcas (USPTO: sigla em inglês), do Escritório Europeu de Patentes (EPO), do Escritório de Patentes do Japão (JPO) e do Instituto Nacional da Propriedade Intelectual (Braspat), no Brasil. A recuperação da informação considerou os principais segmentos e polos inovadores do setor da avicultura industrial, tais como máquinas e equipamentos, medicamentos e vacinação, nutrição e genética. Esses segmentos foram correlacionados à seção “A” da Classificação Internacional de Patentes, tendo-se selecionado 13 códigos IPC dessa seção. As patentes recuperadas por meio desses códigos foram filtradas por meio das palavras-chave “poultry, chicken, frango” e posteriormente analisadas levando-se em consideração o

título, a data de depósito, os nomes dos depositantes, os países de depósito, o estado da técnica e as citações das patentes. Esses procedimentos constataram que os Estados Unidos, a Holanda e o Japão possuem um alto grau de inovação e inventividade nesse setor. Identificaram o domínio norte-americano e europeu em áreas tecnológicas como a genética, e destacaram a atuação diversificada do Japão em áreas como química e saúde humana. Um grupo pequeno de empresas multinacionais detém significativo percentual dos depósitos em nível mundial, com destaque para as empresas holandesas do segmento de máquinas e equipamentos, que tiveram intensa participação na produção de tecnologias relacionadas ao abate e processamento de frangos e à vacinação, considerados estratégicos na avicultura. Em relação ao patenteamento no Brasil, ficou evidente a forte participação de empresas estrangeiras provenientes dos Estados Unidos e da Alemanha nos setores de máquinas e equipamentos e de nutrição avícola, sem indício de atividade de institutos de pesquisas ou universidades nacionais. Esse resultado sugere que poucas tecnologias são desenvolvidas e protegidas no âmbito nacional, embora o agronegócio tenha papel preponderante para o desenvolvimento do Brasil.

RECUPERAÇÃO DA INFORMAÇÃO DE PATENTES

A estratégia de busca é uma das variáveis preponderantes no processo de elaboração de indicadores, pois “[...] a confiabilidade dos dados usados para computação dos indicadores depende não só das características dos softwares utilizados, mas também do cuidado de como essas atividades manuais são realizadas” (MUGNAINI et al., 2004, p. 126). A estratégia de busca pode ser definida como uma técnica que possibilita o encontro entre um questionamento e uma resposta provida de informações armazenadas em uma base de dados (LOPES, 2002). Dependendo dos objetivos da pesquisa e da amplitude da análise, a estratégia de busca para recuperação de documentos de patente pode envolver vários critérios, entre os quais se destacam: as palavras-chave, os códigos da Classificação Internacional de Patente (CIP), as datas de depósito e prioridade, os números das patentes, os nomes dos depositantes, entre outros (WIPO, 2012a).

Segundo Nijhof (2007), com o avanço tecnológico, com as ferramentas de busca em bases de dados e com a grande quantidade de informação disponível para consulta, o operador da busca deve ser cauteloso ao definir o

escopo e a estratégia para recuperar as informações de que necessita. Em geral, quando se pretende analisar um campo tecnológico, as buscas recorrentemente baseiam-se em palavras-chave e códigos CIP, e o responsável pela busca deve ter domínio claro das terminologias acerca do assunto (NIJHOF, 2007, WIPO, 2012a). A CIP foi estabelecida em 1971 e tem por objetivo descrever o conteúdo da patente, independentemente das terminologias usadas na escrita do documento. As invenções são divididas em oito seções principais, sendo elas: A (necessidades humanas); B (operações de processamento e transporte); C (química e metalurgia); D (têxteis e papel); E (construções fixas); F (engenharia mecânica, iluminação e aquecimento, armas e fornos); G (física); e H (eletricidade). Com base nas seções, a classificação subdivide-se em outros quatro níveis hierárquicos, conforme exemplificado pela classificação A01B 1/24 (INPI, 2012):

- Seção A: necessidades humanas.
- Classe A01: agricultura; silvicultura; pecuária; caça; captura em armadilhas; pesca.
- Subclasse A01B: trabalho do solo em agricultura ou silvicultura; peças, detalhes ou acessórios de máquinas ou implementos agrícolas, em geral.
- Grupo principal A01B 1/00: implementos manuais.
- Subgrupo A01B 1/24: implementos manuais para tratamento de prados ou gramados.

A CIP é adotada por mais de 100 países e coordenada pela Organização Mundial da Propriedade Intelectual (WIPO: sigla em inglês). O uso de seus códigos em estratégias de busca é uma vantagem para a recuperação de documentos de patente, visto que a CIP abrange todos os campos tecnológicos e é constantemente revisada, considerando a emergência de novas tecnologias (WIPO, 2012b).

Em geral, o processo de elaboração de estratégia de busca parte de uma expressão piloto simples e pertinente à temática de interesse, que é moldada conforme novos termos e códigos CIP que se mostrem relevantes são identificados. Watts e Porter (1997) recomendam que a busca inicial seja aplicada em bases de dados distintas e que o operador faça refinamentos

na busca por meio de palavras-chave extraídas pela leitura dos títulos e resumos. Especialistas no assunto a ser estudado também podem contribuir no estabelecimento de palavras e expressões para uma busca – por exemplo, o refinamento que resultou em uma estratégia de busca modular para nanotecnologia, um campo considerado emergente, interdisciplinar e difuso. Em um trabalho, Porter et al. (2008) compilaram diversas estratégias de busca já propostas para nanotecnologia e consultaram especialistas da área quanto à pertinência dos termos e recomendações para adição, subtração ou adaptação dos termos. No entanto, consultar e agregar a opinião de diversos especialistas é um processo desafiador. Ao final do processo, o operador deve ter consciência que a estratégia de busca final pode ser simples, com apenas um termo de busca, ou sofisticada, com a combinação de diversos termos, operadores e campos de busca (LANCASTER, 1993).

Outros fatores que influenciam a estratégia de busca são a abrangência e cobertura da base de dados; suas políticas e regras de funcionamento; os campos bibliográficos disponíveis para realizar a busca; e opção de uso de operadores booleanos, de truncagem e coringas (NIJHOF, 2007; WIPO, 2012a). Geralmente, os escritórios de patente mantêm bases de dados com acesso gratuito às informações bibliográficas, ao andamento de processos, e mantêm até a cópia digital do documento de patente depositado, publicado e concedido naquele país em específico. São exemplos desse tipo de base: a Braspat⁴, mantida pelo Instituto Nacional da Propriedade Intelectual (Inpi), e as bases do Escritório Americano de Patentes e Marcas (USPTO⁵: sigla em inglês), do Escritório de Patente do Japão (JPO⁶: sigla em inglês) e do Escritório de Propriedade Intelectual da República Popular da China (SIPO⁷: sigla em inglês). Uma exceção é a Espacenet⁸, base mantida pelo Escritório Europeu de Patentes (EPO), que, além da opção de busca por patentes regional, possui a de abrangência mundial, agregando patentes de mais de 80 repositórios, entre os quais BRASPAT, USPTO, JPO e SIPO. Com exceção da Espacenet, as bases de dados supranacionais são de

⁴ <<http://formulario.inpi.gov.br/MarcaPatente/jsp/servimg/validamagic.jsp?BasePesquisa=Patentes>>.

⁵ <<http://appft1.uspto.gov/netathtml/PTO/search-bool.html>>.

⁶ <http://www.ipdl.inpit.go.jp/homepg_e.ipdl>.

⁷ <http://59.151.93.237/sipo_EN/search/tabSearch.do?method=init>.

⁸ <<http://worldwide.espacenet.com/>>.

domínio particular, entre as quais se destacam a Derwent Innovations Index⁹ e a Micropatent¹⁰. No processo de recuperação da informação, deve-se escolher a base de dados que melhor representa a temática do indicador a ser elaborado, considerando-se a cobertura limitada da base.

No âmbito da recuperação da informação, a estratégia de busca deve obter uma quantidade suficiente de documentos, e, ao mesmo tempo, esses documentos devem ser o mais pertinentes possível à temática do estudo (NIJHOF, 2007). Lancaster (1993) expõe dois coeficientes que contribuem para analisar uma estratégia de busca, intitulados coeficientes de revocação (*recall*) e precisão (*precision*). O coeficiente de revocação é útil para medir a capacidade da estratégia de recuperar documentos relevantes, e o coeficiente de precisão, para medir a capacidade de se evitarem documentos irrelevantes. Para Boccatto e Fujita (2006), a revocação pode ser calculada pela quantidade de documentos relevantes que foram recuperados pela estratégia dividida pelo número de documentos relevantes, para busca, existente na base de dados. Já a precisão é o número de documentos relevantes que foram recuperados dividido pelo número total de documentos recuperados pela estratégia.

$$\text{Revocação} = \frac{\text{N}^{\circ} \text{ de referências relevantes recuperadas}}{\text{N}^{\circ} \text{ de referências relevantes existentes no sistema}} \quad (1)$$

$$\text{Precisão} = \frac{\text{N}^{\circ} \text{ de referências relevantes recuperadas}}{\text{N}^{\circ} \text{ de referências recuperadas pelo sistema}} \quad (2)$$

Cardoso (2000, p. 2) expressou tais conceitos matematicamente, conforme apresentado pelas equações 3 e 4, em que N é o conjunto de resposta ideal de uma busca, $|N|$ é o número total de documentos do conjunto N , R é o que o autor denomina de vetor resultado recuperado (i.e., um conjunto de documentos recuperados), e $|R|$ é o número total de documentos do conjunto R .

$$\text{Revocação} = \frac{N \cap R}{|N|} \quad (3)$$

⁹ <http://apps.webofknowledge.com.ez31.periodicos.capes.gov.br/DIIDW_GeneralSearch_input.do?product=DIIDW&search_mode=GeneralSearch&SID=3CF9ONjrNBhdj7UBd8U&preferencesSaved=>>.

¹⁰ <<http://www.micropat.com/static>>.

$$Precisão = \frac{N \cap R}{|N|} \quad (4)$$

A fórmula do coeficiente de revocação (equação 1) expressa a intersecção entre o número de resposta ideal – ou seja, a quantidade de documentos relevantes existentes na base de dados – e os resultados recuperados considerados relevantes, sobre a quantidade total de documentos da base. Já a fórmula de coeficiente de precisão (equação 2) expressa a intersecção, novamente, entre a quantidade de documentos relevantes existentes na base de dados e a quantidade de documentos recuperados, sobre o resultado total de recuperados. O desafio de utilizar essas fórmulas está na necessidade de avaliar cada registro recuperado e não recuperado quanto à sua relevância. Em particular, para a determinação da revocação, é necessário avaliar a relevância dos documentos não recuperados pela busca. A solução adotada em estudos dessa natureza tem sido o uso de amostras ou conjuntos reduzidos de documentos, de tal forma que todos possam ser avaliados quanto à sua relevância (ARAÚJO, 1979). Adicionalmente, nenhuma busca é exaustiva a ponto de separar cada documento que é relevante para a busca, e “ruídos” sempre existirão nos indicadores finais, cabendo ao operador minimizá-lo.

PROCEDIMENTO DESENVOLVIDO PARA ELABORAÇÃO DA ESTRATÉGIA DE BUSCA

A elaboração da estratégia de busca se deu basicamente em cinco etapas principais: escolha da base de dados; desenvolvimento de uma estratégia de busca inicial; aperfeiçoamentos da estratégia de busca inicial por meio da incorporação de códigos CIP e palavras-chave, sendo estes aplicados tanto para agronegócio quanto para instrumentação; avaliação da pertinência dessas melhorias para o assunto de interesse, utilizando-se testes de revocação e precisão (R e P); e, finalmente, incorporação das melhorias consideradas pertinentes à estratégia de busca inicial, sendo gerada uma estratégia de busca final. Os procedimentos desenvolvidos estão sistematizados no fluxograma da Figura 1. A seguir as etapas desenvolvidas serão detalhadas.

O procedimento iniciou com a escolha da base de dados (etapa 1). Esta deve considerar os objetivos da pesquisa, a amplitude da análise (nacional ou supranacional) e as ferramentas disponíveis para recuperar e filtrar os dados

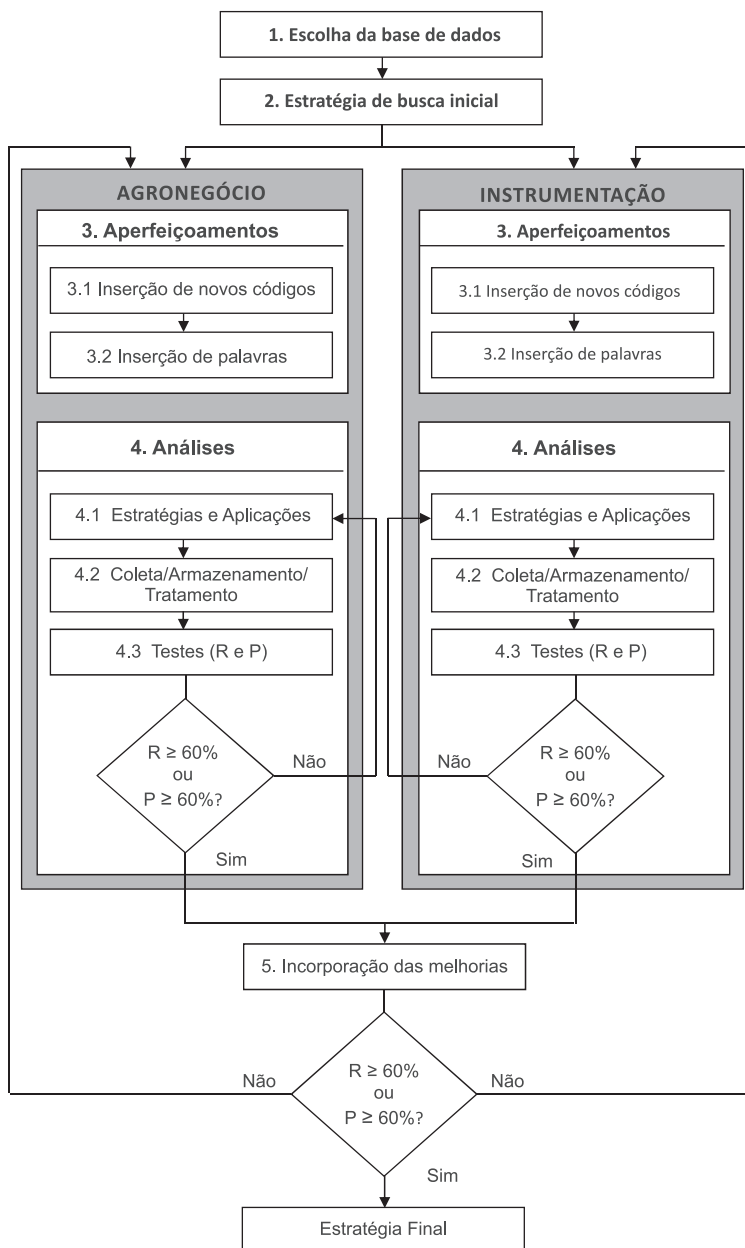


Figura 1. Procedimentos para elaboração da estratégia de busca.

e as limitações da base. Posteriormente, elaborou-se uma estratégia de busca inicial (etapa 2), que serviu de base para as etapas posteriores. No geral, uma estratégia de busca de patentes é composta de códigos e/ou palavras. Na presente pesquisa, a busca inicial foi elaborada por meio de agrupamentos de códigos da classificação internacional de patentes já desenvolvidos anteriormente na literatura, referentes às duas temáticas centrais que circundam o assunto de interesse – no caso, agronegócio e instrumentação (OBSERVATOIRE DES SCIENCES ET DES TECHNIQUES, 2008).

Na etapa 3, de aperfeiçoamentos, preocupou-se em obter melhorias na estratégia inicial inserindo-se novos códigos e palavras nos contextos de agronegócio e instrumentação, elaborando-se, para cada inserção, uma estratégia de busca. Com o intuito de decidir quais outros códigos inserir (etapa 3.1), registros de patentes de uma organização relacionada à temática da estratégia foram utilizados. Nesse caso, a organização escolhida foi a Embrapa Instrumentação, uma Unidade da Embrapa com grande representatividade no patenteamento do agronegócio brasileiro e com *core business* estritamente relacionado à temática da busca. Com base nas credenciais da Embrapa Instrumentação, considerou-se o conjunto de patentes da Unidade como relevante para o tema, o qual foi utilizado para a avaliação da revocação da busca.

Para obter informações sobre tais registros de patentes, uma busca foi feita na BRASPAT, base de dados de patentes do Inpi (Instituto Nacional de Propriedade Industrial). As Unidades da Embrapa não são discriminadas como depositantes das patentes indexadas no Inpi – somente a instituição de um modo geral (Embrapa). Portanto, para separar as informações da Unidade de interesse, foi cruzado o campo depositante – preenchido por “Embrapa” ou “Centro Nacional de Pesquisa e Desenvolvimento de Instrumentação Agropecuária” – com o campo inventor, preenchido conforme as informações sobre pesquisadores do corpo técnico e científico vinculados atualmente à Embrapa Instrumentação, disponíveis no site¹¹ da Embrapa Instrumentação.

Por meio dessa busca, foi possível coletar dados referentes ao número de depósito, códigos CIP e número total de registros de patentes da Embrapa Instrumentação, e considerá-los como informações sobre os depósitos efetuados

¹¹ Relação de pesquisadores da Embrapa Instrumentação: <http://www.cnpdia.embrapa.br/_tecnico.html>.

por essa Unidade. No total, 54 registros de patentes foram recuperados. Desse modo, foi possível fazer uma comparação dos códigos existentes na estratégia inicial, relacionados ao agronegócio e à instrumentação, com os códigos encontrados nos registros das patentes da Embrapa Instrumentação, inserindo-se somente aqueles que não eram repetidos e cuja descrição mencionada na Classificação Internacional de Patentes se relacionava de forma estreita ao agronegócio e à instrumentação.

Em relação às palavras inseridas (etapa 3.2), estas foram retiradas dos títulos dos registros de patentes recuperados por meio da estratégia inicial aplicada na base selecionada, e os registros foram armazenados em um arquivo dentro de um computador local destinado à pesquisa. Esse arquivo passou por um tratamento utilizando-se o software Vantage Point (versão 7.0). Por meio deste, utilizando-se a opção “further processing/ divide at/ whitespace”, as palavras dos títulos de todos os registros foram separadas de modo a criar lista de todas as palavras neles existentes. Dentre esse universo de palavras, foram separadas as que de certa forma se relacionavam ao agronegócio e à instrumentação; logo depois, os registros de patentes que continham tais palavras foram checados, por meio de sua leitura, tendo-se verificado se estes realmente correspondiam a documentos de patentes relacionados à instrumentação aplicada ao agronegócio, de modo a gerar uma lista com palavras consideradas pertinentes para cada contexto.

Posteriormente, analisou-se o comportamento das inserções de novos códigos e palavras-chave – mais especificamente, as melhorias propostas na etapa de aperfeiçoamento foram analisadas (etapa 4), relacionadas ao agronegócio e posteriormente à instrumentação, e foram submetidas a testes de revocação e precisão (R e P). Para cada melhoria, uma estratégia e uma aplicação na base de dados foi feita (etapa 4.1). No total foram aplicadas cinco estratégias de busca. A estratégia inicial foi designada como estratégia 1, aplicada com o intuito de analisar as condições iniciais, servindo de base para as análises posteriores; as estratégias 2 e 3 foram aplicadas para analisar as possíveis melhorias no primeiro contexto referente ao agronegócio, inserindo-se códigos e palavras respectivamente; e a 4 e a 5 tiveram a mesma finalidade que a das duas anteriores, porém, foram aplicadas ao contexto instrumentação. As buscas foram feitas na opção “Advanced search”, durante julho de 2012, e foi delimitado o período de busca pelo “timespan” de 2000 a 2012.

Para cada estratégia aplicada, os dados foram coletados, armazenados e tratados (etapa 4.2) separadamente em um computador local disponibilizado para o desenvolvimento da pesquisa. Os arquivos com os dados recuperados foram importados também separadamente e tratados bibliometricamente com auxílio do software Vantage Point (versão 7.0). Os dados coletados foram analisados (etapa 4.3) segundo os parâmetros de revocação e precisão (CARDOSO, 2000, p. 2). Para efetuar o cálculo de revocação, é necessário saber o número total de documentos existentes na base que são relevantes, e esse fator geralmente é desconhecido. Portanto, o total de registros de patentes da Embrapa Instrumentação existente na base de dados escolhida foi considerado como o total de documentos relevantes existentes, a ser recuperado. Cada estratégia aplicada recuperou um total de patentes da Embrapa Instrumentação – essa informação pode ser encontrada pelo campo depositante do registro de patente importado pelo software Vantage Point –, e essas inferências permitiram realizar cálculos matemáticos. A seguir ilustra-se o cálculo de revocação efetuado.

Considerando-se 20 o número total de documentos de patentes da Embrapa Instrumentação possíveis de serem encontrados na base selecionada ($|N|$), pode-se afirmar que, para a aplicação, todas essas patentes foram relevantes; portanto, 20 também foi considerado o total máximo de documentos relevantes existentes na base (N). Aplicando-se a estratégia de busca, esta retorna apenas cinco desses documentos, sendo esse o total de registros recuperados (R). Nessas condições, a revocação é de 25%.

$$\begin{array}{l} |N|=20 \\ N=20 \\ R=5 \end{array} \quad \text{Revocação} = \frac{20 \cap 5}{|20|} = \frac{5}{20} = 0,25 = 25\%$$

Em relação ao cálculo de precisão, este foi feito para cada estratégia separadamente, utilizando-se um gerador de números aleatórios online RANDOM¹². O cálculo foi aplicado de forma diferente para as estratégias que continham somente códigos e para aquelas que continham códigos e palavras. Para as que continham somente códigos, geraram-se 100 números dentro do intervalo entre 1 e o número de registros coletados em cada aplicação. O número gerado representava um registro de documento de patente. Desse modo, utilizando-se o Vantage Point 7, checkou-se cada registro. Por exemplo, o número

¹² <<http://the-rioblog.blogspot.com.br/2010/07/randomorg-gerador-de-numeros-aleatorios.html>>.

aleatório gerado 233 correspondia ao registro de patente 233; selecionando-se o registro, o software forneceu outras informações, como o título, o resumo e os códigos CIP. A relação com a temática de instrumentação em agronegócio foi feita pela leitura dos títulos e resumos.

Nas estratégias que continham palavras e códigos, de forma aleatória, 11 palavras de cada contexto foram selecionadas também se utilizando do gerador de números aleatórios. Todas as palavras das listas geradas na etapa de aperfeiçoamento foram enumeradas de 1 ao número de palavras listadas. Desse modo, se o número aleatório gerado fosse 23, a palavra correspondente seria a 23ª da lista. As palavras selecionadas para o cálculo no agronegócio foram: agrifood; goat; tomato; coffee; watermelon; acaricid; corn; seed; melon; husbandry; fertilizer. Para o cálculo em instrumentação, foram: transmitter; multivibrator; transponder; dehydrator; extractor; equator; bio-reactor; interferometer; propeller; circuitry; tools. Os registros de patentes que continham as palavras selecionadas foram checados. Para cada palavra, dez registros foram checados por meio de sua leitura – desses, cinco registros continham a palavra selecionada em seu título, e cinco no resumo. Verificou-se se as palavras escolhidas realmente faziam referência a documentos de patentes que fossem de interesse.

Optou-se por essa forma de cálculo, e não pela fórmula de Cardoso (2000, p. 2), pois dessa maneira se estabelece uma noção mais realista dos documentos, sendo possível analisar e conhecer a amostra de forma mais estreita. Em uma etapa final, estabeleceram-se parâmetros de incorporação (etapa 5) – as melhorias propostas na etapa de aperfeiçoamentos que obtiveram índices de revocação ou precisão superiores a aproximadamente 60% foram incorporadas e moldadas na estratégia final. Quando a estratégia não atingia esses índices, esta foi descartada da incorporação, e outra estratégia foi elaborada.

APLICAÇÃO NO CONTEXTO DO AGRONEGÓCIO E INSTRUMENTAÇÃO

Entre as bases de dados disponíveis, a Derwent Innovations Index (DII), base de dados da Thomson Reuters que está disponível no Portal Capes, mostrou-se a mais promissora em virtude da sua facilidade de manuseio, possibilidade de uso de estratégias de busca booleanas complexas, e cobertura supranacional, com informações bibliográficas de patentes de 40 principais

escritórios/repositórios do mundo, com um total de 22.764.581 registros de patentes, desde 1963 até fevereiro de 2013. Outra vantagem é que a base faz o agrupamento dos principais depositantes do mundo em códigos específicos, padronizando a análise final e evitando erros nos indicadores (THOMSON REUTERS, 2012).

Para a aplicação se fez também necessário conhecer as características operacionais de busca da base. No caso, a DII possui três modalidades de busca: geral, avançada e patentes citadas. Além disso, é possível acessar um histórico de pesquisas anteriores, e a função *timespan* permite delimitar um período para busca. A base possibilita o uso de combinação de operadores booleanos (AND, OR, NOT e SAME) e curingas (*, ? e \$), permitindo a elaboração de estratégias de busca complexas, atendendo às necessidades da presente pesquisa.

A estratégia inicial foi elaborada por meio do agrupamento dos documentos em domínios e subdomínios tecnológicos conforme procedimento proposto pelo Observatoire des Sciences et des Techniques (OST). O agrupamento relaciona o código CIP de um documento de patente a um domínio ou subdomínio específico, como apresentado nas Tabelas 1 e 2. Foi utilizado, na estratégia de busca inicial, o domínio instrumentação, que engloba os subdomínios ótica, análise-mensuração-controle, engenharia médica e técnicas nucleares, e os subdomínios produtos agrícolas e alimentares, e aparelhos agrícolas e alimentares que mais se assemelham ao contexto do agronegócio.

Tabela 1. Domínio instrumentação.

	Ótica	G02; G03B, G03C, G03D, G03F, G03G, G03H; H01S
Instrumentação	Análise-mensuração-controle	G01B, G01C, G01D, G01F, G01G, G01H, G01J, G01K, G01L, G01M, G01N, G01P, G01R, G01S, G01V, G01W; G04; G05B, G05D; G07; G08B, G08G; G09B, G09C, G09D; G12
	Engenharia médica	A61B, A61C, A61D, A61F, A61G, A61H, A61J, A61L, A61M, A61N
	Técnicas nucleares	G01T; G21; H05G, H05H

Fonte: Observatoire des Sciences et des Techniques (2008).

Tabela 2. Subdomínios produtos agrícolas e alimentares, e aparelhos agrícolas e alimentares.

Produtos agrícolas e alimentares	A01H; A21D; A23B, A23C, A23D, A23F, A23G, A23J, A23 K, A23L; C12C, C12F, C12G, C12H, C12 J; C13D, C13F, C13J, C13 K
Aparelhos agrícolas e alimentares	A01B, A01C, A01D, A01F, A01G, A01J, A01K, A01L, A01M; A21B, A21C; A22; A23N, A23P; B02B; C12L; C13C, C13G, C13H

Fonte: Observatoire des Sciences et des Techniques (2008).

Depois de sua elaboração, a estratégia inicial passou pelo processo de aperfeiçoamento e inserção das melhorias, inicialmente no contexto agronegócio e posteriormente no contexto instrumentação. A seguir serão expostas as estratégias adotadas em cada aplicação – cada melhoria está em destaque dentro das Tabelas. A Estratégia 1 (Tabela 3) refere-se à estratégia inicial elaborada por meio da intersecção dos códigos agrupados pelo Observatoire des Sciences et des Techniques (2008), conforme explicado acima.

Tabela 3. Estratégia 1: correspondente à estratégia inicial.

<u>Agronegócio</u>	<u>Instrumentação</u>
IP=(A01H* OR A21D* OR A23B* OR A23C* OR A23D* OR A23F* OR A23G* OR A23J* OR A23K* OR A23L* OR C12C* OR C12F* OR C12G* OR C12H* OR C12J* OR C13D* OR C13F* OR C13J* OR C13K* OR A01B* OR A01C* OR A01D* OR A01F* OR A01G* OR A01J* OR A01K* OR A01L* OR A01M* OR A21B* OR A21C* OR A22* OR A23N* OR A23P* OR B02B* OR C12L* OR C13C* OR C13G* OR C13H*)	IP=(G02* OR G03B* OR G03C* OR G03D* OR G03F* OR G03G* OR G03H* OR H01S* OR G01B* OR G01C* OR G01D* OR G01F* OR G01G* OR G01H* OR G01J* OR G01K* OR G01L* OR G01M* OR G01N* OR G01P* OR G01R* OR G01S* OR G01V* OR G01W* OR G04* OR G05B* OR G05D* OR G07* OR G08B* OR G08G* OR G09B* OR G09C* OR G09D* OR G12* OR A61B* OR A61C* OR A61D* OR A61F* OR A61G* OR A61H* OR A61J* OR A61L* OR A61M* OR A61N* OR G01T* OR G21* OR H05G* OR H05H*)
AND	

Na Estratégia 2, os códigos CIP referentes ao agronegócio, provenientes dos registros de patentes da Embrapa Instrumentação, foram inseridos (Tabela 4).

Tabela 4. Estratégia 2: melhoria na estratégia inicial pela adição de códigos CIP na temática agronegócio.

Agronegócio	Instrumentação
IP=(A01H* OR A21D* OR A23B* OR A23C* OR A23D* OR A23F* OR A23G* OR A23J* OR A23K* OR A23L* OR C12C* OR C12F* OR C12G* OR C12H* OR C12J* OR C13D* OR C13F* OR C13J* OR C13K* OR A01B* OR A01C* OR A01D* OR A01F* OR A01G* OR A01J* OR A01K* OR A01L* OR A01M* OR A21B* OR A21C* OR A22* OR A23N* OR A23P* OR B02B* OR C12L* OR C13C* OR C13G* OR C13H*) OR IP=(A01N* OR B02B* OR C05F* OR B08B* OR C08H*)	AND IP=(G02* OR G03B* OR G03C* OR G03D* OR G03F* OR G03G* OR G03H* OR H01S* OR G01B* OR G01C* OR G01D* OR G01F* OR G01G* OR G01H* OR G01J* OR G01K* OR G01L* OR G01M* OR G01N* OR G01P* OR G01R* OR G01S* OR G01V* OR G01W* OR G04* OR G05B* OR G05D* OR G07* OR G08B* OR G08G* OR G09B* OR G09C* OR G09D* OR G12* OR A61B* OR A61C* OR A61D* OR A61F* OR A61G* OR A61H* OR A61J* OR A61L* OR A61M* OR A61N* OR G01T* OR G21* OR H05G* OR H05H*)

A Estratégia 3 propõe a inserção de palavras-chave consideradas pertinentes ao contexto do agronegócio (Tabela 5).

Tabela 5. Estratégia 3: compreende melhoria na estratégia inicial pela adição de palavras-chave na temática agronegócio.

Agronegócio	Instrumentação
IP=(A01H* OR A21D* OR A23B* OR A23C* OR A23D* OR A23F* OR A23G* OR A23J* OR A23K* OR A23L* OR C12C* OR C12F* OR C12G* OR C12H* OR C12J* OR C13D* OR C13F* OR C13J* OR C13K* OR A01B* OR A01C* OR A01D* OR A01F* OR A01G* OR A01J* OR A01K* OR A01L* OR A01M* OR A21B* OR A21C* OR A22* OR A23N* OR A23P* OR B02B* OR C12L* OR C13C* OR C13G* OR C13H*) OR TS=(agricult* OR agroindustr* OR...	AND IP=(G02* OR G03B* OR G03C* OR G03D* OR G03F* OR G03G* OR G03H* OR H01S* OR G01B* OR G01C* OR G01D* OR G01F* OR G01G* OR G01H* OR G01J* OR G01K* OR G01L* OR G01M* OR G01N* OR G01P* OR G01R* OR G01S* OR G01V* OR G01W* OR G04* OR G05B* OR G05D* OR G07* OR G08B* OR G08G* OR G09B* OR G09C* OR G09D* OR G12* OR A61B* OR A61C* OR A61D* OR A61F* OR A61G* OR A61H* OR A61J* OR A61L* OR A61M* OR A61N* OR G01T* OR G21* OR H05G* OR H05H*)

A Estratégia 4 propõe a inserção dos códigos CIP, provindos dos registros de patentes da Embrapa Instrumentação, relacionados ao contexto instrumentação (Tabela 6).

Tabela 6. Estratégia 4: melhoria na estratégia inicial pela adição de códigos CIP na temática instrumentação.

Agronegócio	Instrumentação
IP=(A01H* OR A21D* OR A23B* OR A23C* OR A23D* OR A23F* OR A23G* OR A23J* OR A23K* OR A23L* OR C12C* OR C12F* OR C12G* OR C12H* OR C12J* OR C13D* OR C13F* OR C13J* OR C13K* OR A01B* OR A01C* OR A01D* OR A01F* OR A01G* OR A01J* OR A01K* OR A01L* OR A01M* OR A21B* OR A21C* OR A22* OR A23N* OR A23P* OR B02B* OR C12L* OR C13C* OR C13G* OR C13H*)	IP=(G02* OR G03B* OR G03C* OR G03D* OR G03F* OR G03G* OR G03H* OR H01S* OR G01B* OR G01C* OR G01D* OR G01F* OR G01G* OR G01H* OR G01J* OR G01K* OR G01L* OR G01M* OR G01N* OR G01P* OR G01R* OR G01S* OR G01V* OR G01W* OR G04* OR G05B* OR G05D* OR G07* OR G08B* OR G08G* OR G09B* OR G09C* OR G09D* OR G12* OR A61B* OR A61C* OR A61D* OR A61F* OR A61G* OR A61H* OR A61J* OR A61L* OR A61M* OR A61N* OR G01T* OR G21* OR H05G* OR H05H*) OR <i>IP=(A47J* OR B01J* OR B65B* OR B65D* OR C12M* OR D01D* OR E04H* OR F25B* OR G06G* OR H01J*)</i>

A Estratégia 5 propõe a inserção de palavras-chave consideradas pertinentes ao contexto instrumentação (Tabela 7).

Tabela 7. Estratégia 5: compreende melhoria na estratégia inicial pela adição de palavras-chave na temática instrumentação.

Agronegócio	Instrumentação
IP=(A01H* OR A21D* OR A23B* OR A23C* OR A23D* OR A23F* OR A23G* OR A23J* OR A23K* OR A23L* OR C12C* OR C12F* OR C12G* OR C12H* OR C12J* OR C13D* OR C13F* OR C13J* OR C13K* OR A01B* OR A01C* OR A01D* OR A01F* OR A01G* OR A01J* OR A01K* OR A01L* OR A01M* OR A21B* OR A21C* OR A22* OR A23N* OR A23P* OR B02B* OR C12L* OR C13C* OR C13G* OR C13H*)	IP=(G02* OR G03B* OR G03C* OR G03D* OR G03F* OR G03G* OR G03H* OR H01S* OR G01B* OR G01C* OR G01D* OR G01F* OR G01G* OR G01H* OR G01J* OR G01K* OR G01L* OR G01M* OR G01N* OR G01P* OR G01R* OR G01S* OR G01V* OR G01W* OR G04* OR G05B* OR G05D* OR G07* OR G08B* OR G08G* OR G09B* OR G09C* OR G09D* OR G12* OR A61B* OR A61C* OR A61D* OR A61F* OR A61G* OR A61H* OR A61J* OR A61L* OR A61M* OR A61N* OR G01T* OR G21* OR H05G* OR H05H*) OR <i>TS=(compartment* OR compressor* OR ...</i>

Posteriormente, as estratégias acima elaboradas foram aplicadas na base DII e passaram pelo processo de avaliação. A Tabela 8 expõe os resultados e as análises feitas em cada aplicação.

Tabela 8. Síntese dos resultados das buscas.

	Estratégia 1	Estratégia 2	Estratégia 3	Estratégia 4	Estratégia 5
Total de registros	18.417	26.539	> 100.000 ¹³	40.119	> 100.000
Total – Embrapa Instrumentação	1	1	5	1	8
Revocação	5%	5%	25%	5%	40%
Precisão	64%	50%	74%	73%	89%

Na base de dados DII, foi encontrado um total de 20 registros de patentes da Embrapa Instrumentação. Com aplicação da Estratégia 1, uma quantidade relevante de registros foi recuperada com um alto nível de precisão (64%); porém, dos 20 registros de patentes da Embrapa Instrumentação existentes na base DII, somente 1 foi recuperado por essa estratégia, tendo-se atingido baixa revocação (5%). Adotando-se a Estratégia 1 como parâmetro, pode-se inferir que, com a Estratégia 2, apesar do aumento na quantidade de registros recuperados, em geral, não houve melhora na revocação, e a precisão da amostra diminuiu de 64% para 50%. Ambos os parâmetros de revocação e precisão foram inferiores a 60%, mostrando a necessidade de prosseguir com o processo de melhoria, porém, sem a inserção dos novos códigos referentes ao agronegócio.

A Estratégia 3 retornou mais de 100.000 registros, sendo 5 deles os documentos de patentes da Embrapa Instrumentação. Em relação à Estratégia 1, verificou-se melhoria da revocação, que passou de 5% para 25%. A precisão, que inicialmente era de 64%, aumentou para 74%. Dessa forma, com essa terceira aplicação, houve um aumento tanto nos parâmetros revocação e precisão quanto no número de documentos recuperados; portanto, infere-se que as palavras referentes ao agronegócio contribuem significativamente na estratégia final, sendo importante sua incorporação numa estratégia de busca final.

¹³ A base DII limita a visualização dos resultados de uma busca em no máximo 100.000 registros; porém, é possível fragmentar as partes da busca e contabilizar a quantidade de registros recuperada em cada fragmento para depois somá-los. Nessa pesquisa, nos casos em que o número recuperado foi maior que 100.000, significa que contabilizar os registros por meio dessa fragmentação não foi possível, pois, mesmo fragmentando-se, um ou mais deles excederam 100.000, não sendo possível realizar a somatória.

Em relação ao contexto da instrumentação, a Estratégia 4 recuperou 40.119 registros no total. No que tange à revocação, observou-se resultado semelhante ao da Estratégia 2, sugerindo que a incorporação de códigos CIP para instrumentação também não seja relevante. Entretanto, embora a revocação tenha permanecido baixa (5%) e idêntica à da Estratégia 1, houve aumento na precisão, que passou para 73%, por causa da diminuição do número total de registros recuperados. Apesar do comportamento semelhante ao da Estratégia 2 em relação à revocação, a precisão encontrada foi superior a 60%; assim, os códigos utilizados nessa busca contribuem para uma amostra mais precisa, devendo ser incorporados à estratégia de busca final.

A Estratégia 5 recuperou mais de 100.000 registros, entre os quais 8 pertenciam à Embrapa Instrumentação, o que causou aumento na revocação (40%) e precisão (89%) em relação à Estratégia 1. Dessa forma, a inserção de palavras referentes à instrumentação foi a que mais contribuiu para o processo de aperfeiçoamento. Em síntese, tendo em vista as análises acima, a estratégia de busca final deve incorporar os códigos provindos dos domínios e subdomínios (OST) iniciais, as palavras-chave referentes ao agronegócio, os códigos CIP e as palavras referentes à instrumentação. A estratégia de busca final foi modularizada e está disponível na Tabela 9.

Tabela 9. Estratégia de busca final.

#1	IP=(A01H* OR A21D* OR A23B* OR A23C* OR A23D* OR A23F* OR A23G* OR A23J* OR A23K* OR A23L* OR C12C* OR C12F* OR C12G* OR C12H* OR C12J* OR C13D* OR C13F* OR C13J* OR C13K* OR A01B* OR A01C* OR A01D* OR A01F* OR A01G* OR A01J* OR A01K* OR A01L* OR A01M* OR A21B* OR A21C* OR A22* OR A23N* OR A23P* OR B02B* OR C12L* OR C13C* OR C13G* OR C13H*)
#2	TS=(agricult* OR agroindustr* OR agribusiness OR (“crop production”) OR tillage* OR husbandry OR (“agri-food”) OR agrifood* OR agrochemic* OR agronomic* OR aquacultur* OR cultivate* OR farm* OR food* OR forestry OR fruit* OR vegetable* OR harvest* OR horticultur* OR irrigat* OR legume* OR pulverization* OR silvicultur* OR veterinari* OR sow* OR livestock* OR soil* OR seed*)
#3	TS=(bean* OR citric* OR citru* OR grain* OR grape* OR melon* OR onion* OR orange* OR peanut* OR pepper* OR potato* OR rice* OR sugarcane* OR coffee* OR cotton* OR corn OR corns OR wheat* OR banana* OR lemon* OR apple* OR tomato* OR cassava* OR watermelon* OR cocoa*)
#4	TS=(biocid* OR fertilizer* OR fungicide* OR insecticid* OR fungicid* OR herbicid* OR molluscicid* OR nematocid* OR raticid* OR rodenticid* OR acaricid* OR pesticide*)
#5	TS=(bees OR bee OR bovine* OR canine* OR cow* OR dog* OR equine* OR fish* OR horse* OR pork* OR swine* OR chicken* OR poult* OR buffaloe* OR caprine* OR goat* OR sheep*)

Continua...

Tabela 9. Continuação.

#6	#5 OR #4 OR #3 OR #2 OR #1
#7	IP=(G02* OR* G03B* OR G03C* OR G03D* OR G03F* OR G03G* OR G03H* OR H01S* OR G01B* OR G01C* OR G01D* OR G01F* OR G01G* OR G01H* OR G01J* OR G01K* OR G01L* OR G01M* OR G01N* OR G01P* OR G01R* OR G01S* OR G01V* OR G01W* OR G04* OR G05B* OR G05D* OR G07* OR G08B* OR G08G* OR G09B* OR G09C* OR G09D* OR G12* OR A61B* OR A61C* OR A61D* OR A61F* OR A61G* OR A61H* OR A61J* OR A61L* OR A61M* OR A61N* OR G01T* OR G21* OR H05G* OR H05H*)
#8	IP=(A47J* OR B01J* OR B65B* OR B65D* OR C12M* OR D01D* OR E04H* OR F25B* OR G06G* OR H01J*)
#9	TS=(accelerator* OR adder* OR agitator* OR analyser* OR applicator* OR aspirator* OR barometer* OR battery* OR bioinstrumentation* OR (bio-reactor*) OR boiler* OR bottle* OR cabinet* OR cable* OR calliper* OR camera* OR capaciflector* OR capacitance* OR capacitor* OR catcher* OR catching OR circuitry* OR circulator* OR collector* OR colorimeter*)
#10	TS=(compartment* OR compressor* OR container* OR controller* OR transmitter* OR converter* OR (converter-sensor*) OR conveyor* OR cooler* OR cropper* OR cutter* OR cytometer* OR degasser* OR dehydrator* OR dendrometer* OR densimeter* OR depository* OR detector* OR determinator* OR device* OR diffuser* OR digger* OR digitizer* OR discharger* OR displayer* OR distributor* OR ditcher* OR dripper* OR drive* OR dryer* OR duster* OR ejector* OR electrometer* OR emulator* OR equalizer* OR equator* OR equipment*)
#11	TS=(exchanger* OR exerciser* OR extractor* OR feeder* OR feeler* OR filter* OR forager* OR fumigator* OR gathering OR goniometer* OR harvester* OR processor* OR reaper* OR hoe* OR holder* OR homogenizer* OR hydrometer* OR hydrophone* OR hygrometer* OR incubator* OR instrument* OR interferometer* OR interrogator* OR Ionizer* OR leafcutter* OR limiter* OR liquefier* OR localizer* OR machine* OR (machine-readable*) OR machinery OR (machine-specific*) OR manometer* OR (measure-controller*) OR measurer*)
#12	TS=(mechanism* OR microcomputer* OR (micro-computer*) OR microcontroller* OR micropump* OR (micro-sprayer*) OR (mini-lysimeter*) OR moistener* OR monitor* OR motor* OR multivibrator* OR object* OR ombrometer* OR orchidometer* OR pipe* OR lanimeter* OR planter* OR pot* OR processor* OR propeller* OR protractor* OR provider* OR pulsator*)
#13	TS=(pulveriser* OR pumper* OR puncher* OR pusher* OR pyrheliometer* OR reaper* OR reflector* OR retainer* OR saturator* OR saucer* OR scraper* OR sensor* OR sower* OR spectrophotometer* OR stroboscope* OR supporter* OR telescope* OR thermalizer* OR thermometer* OR thermosensor* OR tool* OR tractor* OR transmitter* OR receiver* OR transponder* OR truck* OR vibrator*)
#14	#13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7
#15	#14 AND #6

A estratégia final recuperou 772.069 registros de patentes, sendo 14 registros da Embrapa Instrumentação, o que resultou em um valor de revocação de 70% e de 68% para precisão. Consequentemente, pode-se aferir que, do total da amostra, 525.003 registros estariam associados à temática de interesse. Os demais 247.066 registros de documentos de patente podem até ser considerados ruídos, mas é importante ressaltar que eles possuem pelo menos um código CIP ou termo associado a agronegócio e instrumentação. Se algum

dos parâmetros de revocação e precisão fosse inferior a 60%, o processo de aperfeiçoamento deveria ser revisado. Pelos resultados da Tabela 8, quando a revocação calculada foi abaixo de 60%, os níveis de precisão foram altos (acima de 60%), com exceção somente da Estratégia 2. Portanto, atingir o equilíbrio dos dois parâmetros foi tido como suficiente até o presente momento. Em relação à estratégia inicial, o aumento no parâmetro “total de registros recuperados” foi de 98%, seguido de um aumento de 65% na revocação e de 4% na precisão.

CONSIDERAÇÕES FINAIS

A estratégia de busca desenvolvida pode auxiliar pesquisadores interessados na compreensão de aspectos relacionados ao desenvolvimento de maquinários, dispositivos e sensores utilizados na agroindústria. Também pode ser aplicada em outras bases de dados de patentes, sendo necessária apenas sua adaptação quanto aos campos, caracteres e operadores permitidos na base. Aperfeiçoamentos futuros podem ser desenvolvidos na estratégia de busca resultante da presente pesquisa, visto que há a dinâmica de desenvolvimento tecnológico, e novas tecnologias podem englobar outros termos-chave e a criação de novos códigos CIP de interesse à instrumentação aplicada ao agronegócio. O procedimento desenvolvido para elaboração de estratégia de busca mostrou-se eficiente na medida em que, embora seja aconselhável, nem sempre é possível a consulta a especialistas para delimitar uma estratégia de busca. Além disso, o procedimento considera cálculos de revocação e precisão como critérios para averiguar a melhoria da estratégia de busca com a inserção de novos elementos (no caso, códigos CIP e palavras-chave). Outras pesquisas podem se basear no procedimento para elaborar novas estratégias de busca tanto para instrumentação aplicada ao agronegócio quanto para outras temáticas de interesse.

Os próximos passos e contribuições da presente pesquisa buscarão elaborar e analisar indicadores tecnológicos com os dados recuperados pela estratégia proposta, para diagnosticar: quais são as instituições em nível mundial que depositam patentes relacionadas; quais são as que mais patenteiam; a quais outras áreas do conhecimento esses instrumentos se relacionam; como o interesse em patentear esses instrumentos vem se desenvolvendo ao longo do tempo, e como está o panorama atual; quais são os países que são potenciais mercados; e quais são os principais fornecedores, entre outras questões que podem ser utilizadas estrategicamente para diversos fins tanto políticos quanto econômicos e sociais.

AGRADECIMENTOS

Os autores agradecem à Universidade Federal de São Carlos (UFSCar), à Embrapa Instrumentação (CNPDIA), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

REFERÊNCIAS

- ALVES, J. M. de S.; MARTINELLI, O.; DEWES, H. Dinâmica inovativa no agronegócio: a inovação tecnológica na avicultura industrial por meio da análise de patentes. **Cadernos de Ciência & Tecnologia**, Brasília, DF, v. 23, n. 2/3, p. 207-233, 2006.
- ARAÚJO, E. E de. Revocação (Recall) e Precisão (Precision) no SDI/CIN/CNEN. **Ciência da Informação**, Brasília, DF, v. 8, n. 1, p. 47-50, 1979.
- ASHTON, W. B.; KLAVANS, R. A. **Keeping abreast of science and technology**: technical intelligence for business. Columbus: Batelle Press, 1997.
- BARROSO, W.; QUONIAM, L.; PACHECO, E. Patents as technological information in Latin America. **World Patent Index**, [S.l.], v. 31, n. 3, p. 207-215, 2009.
- BOCCATO, V. R. C.; FUJITA, M. S. L. Estudos de avaliação quantitativa e qualitativa de linguagens documentárias: uma síntese bibliográfica. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 2, p. 267-281, 2006.
- CARDOSO, O. N. P. Recuperação de informação. **InfoComp**, Lavras, v. 2, n. 1, p. 33-38, 2000.
- UNIVERSIDADE DE BRASÍLIA. Faculdade de Ciencia da Informação. **Glossário geral da Ciência da Informação**. Disponível em: <<http://www.cid.unb.br/publico/setores/100/123/sistema/m0039015.htm>>. Acesso em: 8 dez. 2012.
- GRUPO DE ORDENAMENTO DO TERRITÓRIO. **Conceito de indicador**. Disponível em: <http://www.igeo.pt/instituto/cegig/got/17_Planning/Files/indicadores/conceito_indicador.pdf>. Acesso em : 8 dez. 2013.
- INPI. Instituto Nacional da Propriedade Industrial. Disponível em: <<http://www.inpi.gov.br/portal/>>. Acesso em: 8 jan. 2012.
- LANCASTER, F. W. A. **Indexação e resumos**: teoria e prática. Trad. Antonio Agenor Briquet Lemos. Brasília, DF: Briquet de Lemos, 1993.
- LOPES, I. L. Estratégia de busca na recuperação da informação: revisão da literatura. **Ciência da Informação**, Rio de Janeiro, v. 31, n. 2, p. 60-71, 2002.
- LOPES, S.; COSTA, M. T.; FERNÁNDEZ LLIMÓS, F.; AMANTE, M. J.; LOPES, P. F. A bibliometria e a avaliação da produção científica: indicadores e ferramentas. **Actas do**

Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas, Lisboa, v. 11 2012.

MILANEZ, D. H. **Nanotecnologia**: indicadores tecnológicos sobre os avanços em materiais a partir da análise de documentos de patentes. 2011. 176 f. Dissertação (Mestrado em Ciência e Engenharia dos Materiais) - Centro de Ciências Exatas e de Tecnologia, Universidade Federal de São Carlos, São Carlos.

MUGNAINI, R.; JANNUZZI, P. de M.; QUONIAM, L. Indicadores bibliométricos da produção científica brasileira: uma análise a partir da base Pascal. **Ciência da Informação**, Brasília, DF, v. 33, n. 3, p. 123-131, 2004.

NIJHOF, E. Subject analysis and search strategies: has the searcher become the bottleneck in the search process. **World Patent Information**, Oxford, v. 29, n. 1, p. 20-25, 2007.

OCDE. **Rumo a um desenvolvimento sustentável**: indicadores ambientais. Paris: OCDE, 2002. p. 224. (Cadernos de referência ambiental, v. 9).

OCDE. **Patent statistics manual**. Paris: OCDE, 2009. p. 162.

OBSERVATOIRE DES SCIENCES ET DES TECHNIQUES. **Indicateurs de sciences et de technologies**. Paris: Economica, 2008. Disponível em: <http://www.obs-ost.fr/fileadmin/medias/PDF/R08_Complet_Liens.pdf>. Acesso em: 19 fev. 2012.

PORTER, A. L.; YOUTIE, J.; SHAPIRA, P.; SHOENECK, D. Refining search terms for nanotechnology. **Journal of Nanoparticle Research**, [Heidelberg], v. 10, n. 5, p. 715-728, 2008.

RAMOS, R. C. **Elaboração de indicadores de patentes sobre nanotecnologia aplicada ao agronegócio**. 2012. 112 f. Dissertação (Mestrado em Ciência Tecnologia e Sociedade) - Programa de Pós-graduação Ciência Tecnologia e Sociedade, Universidade Federal de São Carlos, São Carlos.

SANTOS, R. N. M. dos; KOBASHI, N. Y. Aspectos metodológicos da produção de indicadores em ciência e tecnologia. In: ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO, 6., 2005, Salvador. **Anais...** Salvador: Cinform, 2005.

SENGE, P. M. **A quinta disciplina**: a arte a prática da organização que aprende. 26. ed. Rio de Janeiro: Best Seller, 2010.

THOMSON REUTERS. **Derwent Innovations Index Help**: table of contents. Disponível em: <<http://images.webofknowledge.com.ez87.periodicos.capes.gov.br/WOKRS55B6/help/DII/contents.html>>. Acesso em: 20 ago. 2012.

VANTAGE POINT. **Products**. Disponível em: <<http://www.thevantagepoint.com/products.html>>. Acesso em: 8 fev. 2012.

L. I. L. de Faria et al.

WATTS, R. J.; PORTER, A. L. Innovation forecasting. **Technological Forecasting and Social Change**, New York, v. 56, n. 1, p. 25-47, 1997.

WIPO. **Guide to use patent information**. Disponível em: <http://www.wipo.int/freepublications/en/patents/434/wipo_pub_1434_03.pdf> Acesso em: 8 fev. 2012a.

WIPO. **World Intellectual Property Organization**. Disponível em: <<http://www.wipo.int/portal/index.html.en>>. Acesso em: 8 jan. 2012b.

Trabalho recebido em 18 de fevereiro de 2013 e aceito em 15 de abril de 2013.