

Ministério da Saúde

**FIOCRUZ**  
**Fundação Oswaldo Cruz**



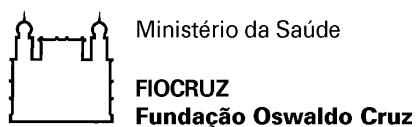
# **PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMAÇÃO E COMUNICAÇÃO EM SAÚDE – PPGICS ICICT/FIOCRUZ**

**Leonardo Silva Leite**

**Informação para prospecção: um estudo exploratório na  
área da saúde**

**Rio de Janeiro, RJ**

**2011**



## **PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMAÇÃO E COMUNICAÇÃO EM SAÚDE – PPGICS ICICT/FIOCRUZ**

**Leonardo Silva Leite**

### **Informação para prospecção: um estudo exploratório na área da saúde**

**Dissertação de Mestrado apresentada ao Programa de Pós-Graduação Stricto Sensu em Informação e Comunicação em Saúde (PPGICS) do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz (Icict/Fiocruz), como requisito parcial à obtenção do título de Mestre em Ciências.**

**Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Cícera Henrique da Silva**

**Rio de Janeiro, RJ**

**2011**

Ficha catalográfica elaborada pela  
Biblioteca de Ciências Biomédicas/ ICICT / FIOCRUZ - RJ

L533 Leite, Leonardo Silva.

Informação para prospecção: um estudo exploratório na área da saúde. / Leonardo Silva Leite. – Rio de Janeiro, 2011.  
xvi, 102 f. : il. ; 30 cm.

Dissertação (mestrado) – Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Pós-Graduação em Informação e Comunicação em Saúde, 2011.

Bibliografia: f. 96-102

1. Informação. 2. Comunicação e saúde. 3. Informação científica e tecnológica. 4. Inovação em saúde. 5. Prospecção tecnológica. 6. Recuperação da informação. 7. Estratégia de busca. 8. Bibliometria. 9. Patente. I. Título.

CDD 302.23

**Programa de Pós-Graduação em Informação e Comunicação em Saúde - PPGICS**

**ATA DE DEFESA DE DISSERTAÇÃO INTITULADA INFORMAÇÃO PARA PROSPECÇÃO: UM ESTUDO EXPLORATÓRIO NA ÁREA DA SAÚDE, DO MESTRANDO LEONARDO SILVA LEITE**

As 14 horas do dia 25 de agosto de 2011, na sala 710, no Prédio da Expansão do Campus da Fundação Oswaldo Cruz, foi realizada a defesa de dissertação intitulada Informação para Prospecção: Um Estudo Exploratório na Área da Saúde, como parte dos requisitos para obtenção do Título de Mestre em Ciências, no Curso de Mestrado em Informação e Comunicação em Saúde – PPGICS, sob a orientação da Professora DR.<sup>a</sup> Cícera Henrique da Silva (ICICT - FIOCRUZ). A Banca Examinadora foi composta pelos professores: DR.<sup>a</sup> Denise Nacif Pimenta (CDTS - VPPIS-FIOCRUZ), Dr.<sup>a</sup> Maria Cristina Soares Guimarães (ICICT - FIOCRUZ), Dr.<sup>a</sup> Cícera Henrique da Silva (ICICT - FIOCRUZ). A orientadora abriu os trabalhos e passou a palavra ao candidato para apresentar o trabalho. Finda a apresentação, o candidato foi arguido pelos examinadores, sendo suas respostas consideradas  satisfatórias  insatisfatória. A Banca examinadora deu por encerrada a sessão, declarando a dissertação do mestrando Leonardo Silva Leite:

Aprovada;  Aprovada com restrições;  Reprovada;

Observações:

*Resalta a importância do tema e o resgate de uma área de pesquisa que se torna cada vez mais importante na atualidade. A banca orienta a publicação enfatizando a importância das políticas públicas de informação em saúde.*

A presente ata foi firmada pelos três examinadores:

Assinatura: *Denise M. Pimenta* (DR.<sup>a</sup> Denise Nacif Pimenta (CDTS - VPPIS-FIOCRUZ))

Assinatura: *MCS Guimarães* (Dr.<sup>a</sup> Maria Cristina Soares Guimarães / ICICT - FIOCRUZ)

Assinatura: *CSilva* (DR.<sup>a</sup> Cícera Henrique da Silva / ICICT - FIOCRUZ)

Aos meus pais, Ana e José, e esposa, Thaís  
Fernanda Bette, pelo apoio e confiança.

## **AGRADECIMENTOS**

À Fundação Oswaldo Cruz, Coordenação de Gestão Tecnológica e Instituto de Comunicação e Informação Científica e Tecnológica em Saúde por terem oferecido as condições necessárias para o desenvolvimento deste trabalho.

À orientadora Dr<sup>a</sup> Cícera Henrique da Silva pelos ensinamentos, incentivo e paciência na orientação deste trabalho.

À Coordenadora da Gestão Tecnológica Maria Celeste Emerick, pela oportunidade de crescimento em conhecimentos e ao incentivo e apoio na realização deste mestrado.

À chefe imediata e amiga, Leila Costa Duarte Longa, que sempre acreditou no meu potencial, me motivando tanto na vida profissional quanto pessoal.

Às amigas da área de patentes da Coordenação de Gestão Tecnológica, Adriana Campos Moreira Britto, Fabrícia Pires Pimenta e Gisele Cruz de Mendonça pelo auxílio técnico específico da área e incentivo para a conclusão desta etapa profissional.

Aos demais colegas e amigos da Coordenação de Gestão Tecnológica que sempre me impulsionaram com uma palavra de conforto e sugestões para a conclusão deste trabalho.

À Dr<sup>a</sup> Maria Cristina Soares Guimarães, coordenadora, professora, orientadora no curso de Especialização em Informação Científica e Tecnológica em Saúde, em 2007, pelo conhecimento e experiência principalmente no campo da Ciência da Informação proporcionadas no curso que foi determinante na aquisição de conhecimentos para a entrada e conclusão deste mestrado.

À coordenadora e professora do mestrado, Dr<sup>a</sup> Inesita Soares de Araújo, que em conjunto com a Dr<sup>a</sup> Cícera Henrique da Silva são responsáveis pela disciplina

Portfolio que foi de suma importância para uma visão externa do tema debatido na dissertação.

Aos demais professores do Programa de Pós-graduação em Informação e Comunicação em Saúde (PPGICS).

Aos membros da banca examinadora do mestrado, agradeço o apoio e participação na defesa.

À Dr<sup>a</sup> Simone Alencar pela presteza e contribuição nas discussões e utilização da ferramenta de mineração de texto *Vantage Point*.

Aos profissionais de secretaria de ensino pelo apoio no decorrer do curso.

Aos amigos de longa data que apesar das reclamações entendiam os motivos da ausência temporária.

Aos mais novos amigos do mestrado pelo incentivo, união, força e amizade nas horas mais precisas do curso. Em especial ao amigo Claudio Machado que nesses últimos dois anos esteve presente como um irmão nos momentos alegres e não menos nos mais difíceis e angustiantes ao longo desta jornada.

À minha querida avó materna Eunice que se foi ao longo desta trajetória e que com certeza onde estiver está orgulhosa e muito feliz com meu crescimento intelectual.

À minha amada esposa Thaís Fernanda Bette, a maior conquista nesses últimos dois anos, por ser a fonte de inspiração e a motivação da minha vida.

Aos meus pais pela dedicação, educação e ensinamentos dos princípios da dignidade, honestidade e a razão da existência da família.

## RESUMO

A competitividade cada vez mais acirrada tem feito com que empresas busquem formas de obter e manter vantagens sobre seus concorrentes. Uma das formas é através do monitoramento ambiental ou monitoramento tecnológico. Para que o monitoramento de informação seja eficaz é necessário planejamento de busca que enfoque a melhor base para a realização da busca; as palavras-chave adequadas; a fórmula lógica da estratégia. O planejamento da busca necessita de um conhecimento minucioso das técnicas de estratégia de busca, dos prós e contras da estratégia no texto completo dos documentos, do uso das linguagens natural e controlada na recuperação da informação. Neste sentido, esta dissertação apresenta como questão norteadora a forma como se deve elaborar uma estratégia de busca de informação em documentos de patente para monitoramento tecnológico. Mais especificamente, o trabalho apresenta as características das principais bases de patentes existentes; avaliação do uso da linguagem natural e controlada; do uso de truncagem nas palavras-chave, da localização dos termos selecionados e quanto ao escopo de proteção do documento de patente. Como foco da estratégia de busca foi definido o diagnóstico da malária pela importância da doença em nível mundial, considerando que um diagnóstico preciso influencia no tratamento mais adequado e pode-se alcançar reduções nos custos das despesas de saúde em todo o mundo. A estratégia de busca foi elaborada, na base de dados de patentes Derwent, pelo cruzamento de palavras representativas do conceito malária e diagnóstico, associados à classificação internacional de patentes no intervalo 2005 a 2009. Os dados foram tratados e classificados em software de mineração de texto, em dois níveis: primeiramente quanto à relevância para a malária, em segundo quanto ao escopo de proteção. Os resultados mostram que maior parte dos documentos de patente não são específicos para malária, nem tão pouco para o diagnóstico, a busca por termos apenas no título limita o retorno dos resultados já que a maior parte dos conceitos foram encontrados no resumo; o uso da linguagem natural e controlada na mesma estratégia de busca é fundamental para se obter retorno mais preciso da informação; a truncagem de termos é importante na estratégia de busca, porém é necessário conhecimento das combinações dos termos para que não sejam recuperadas informações irrelevantes. Como conclusão pode-se dizer que não há



estratégia de busca que apresente 100% de relevância quanto aos documentos recuperados, no que diz respeito à busca de patentes. A meta da estratégia de busca ideal seria uma revocação elevada e uma alta precisão; no entanto, revocação e precisão tendem a se correlacionar inversamente, ou seja, uma maior precisão leva a uma menor revocação e vice-versa, o que significa que o pesquisador de patentes tem de encontrar o equilíbrio adequado entre estes polos de acordo com o tipo de busca escolhida.

**PALAVRAS-CHAVE:** informação, comunicação e saúde – informação científica e tecnológica – inovação em saúde –prospecção tecnológica – recuperação da informação – estratégia de busca - bibliometria – patente

## ABSTRACT

The increasingly fierce competition between companies in the market has made these companies look for ways to obtain and maintain an advantage over their competitors. One way is through environmental monitoring or tracking technology. For monitoring information is needed is an effective strategic planning search: better basis for decision of the context for and the decision of the keywords and the decision of the logical formula of the strategy. More specifically the strategic planning of search requires a thorough knowledge of the techniques of search strategy, the pros and cons of the strategy in full-text documents, the use of controlled natural languages in information retrieval. Thus, this paper presents as main question is how to devise a strategy for finding information in patent documents for tracking technology. More specifically, the paper presents the characteristics of the main bases of existing patents; evaluation of the use of natural language and controlled, the use of truncation in the keywords, the location of selected terms and the scope of protection of the patent document. The basic patent was selected by the *Derwent Innovation Index* reindex the title and abstract of the patent in addition to allowing the download of documents for use in data mining software. As a focus of the search strategy was defined by the importance of malaria diagnosis of the disease worldwide and diagnostic segment as an accurate diagnosis influences the most appropriate treatment and can achieve reductions in costs of health expenditures worldwide. The search strategy was developed by crossing concepts with the diagnosis of malaria associated with the international patent classification within the range 2005 to 2009. The documents were classified into two levels: first as to the relevance to malaria and second on the scope of protection. The results show that the search terms titles only limits the return of results, since most of the concepts found in the summary, the use of natural language and controlled in the same search strategy is crucial to get return more accurate information, the truncation is important in terms search strategy, but you need knowledge of the combinations of terms that are not retrieved irrelevant information. In conclusion we can say that there is no search strategy to produce 100% of the retrieved documents as relevant, with regard to the patent search. The goal of any search strategy would be revoked and a high precision, ie the recovery of virtually all relevant documents. However, recall and precision tend to correlate inversely, ie, a higher precision leads to lower recall and

vice versa, which means that the patent researcher has to find the right balance between these poles according to the type chosen search.

Keywords: information, communication and health – scientific and technological information – innovation in health – technological forecasting – information retrieval – search strategy – bibliometrics – patent

## LISTA DE ILUSTRAÇÕES

Figura 1 - Monitoramento ambiental e previsão tecnológica no processo de planejamento estratégico. ....	9
Figura 2: Processo de Monitoramento Ambiental.....	12
Figura 5 - 1ª Lei de Zipf.....	29
Figura 6 - 2ª Lei de Zipf: Lei Zipf-Booth.....	29
Figura 7: Distribuição de Vocabulário Controlado. ....	30
Figura 8 - Fêmea adulta de <i>Anopheles gambiae</i> em posição de alimentação de sangue na pele humana.....	61
Figura 9 - Países incidentes e áreas de risco da malária no mundo em 2010. ....	62
Figura 10 - Distribuição espacial do risco de transmissão da doença no Brasil em 2008. ....	63
Figura 11: Tela de busca básica na base Derwent. ....	71
Figura 12: Tela de busca por patente citada. ....	72
Figura 13: Tela de busca avançada da base Derwent. ....	72
Figura 14: Resultado da etapa 01 da busca.....	75
Figura 15: Fluxograma representativo das etapas 01 e 02 .....	77
Figura 16: Resultado da etapa 02 da busca.....	78
Gráfico 1: Documentos de patente selecionados pelo ano de prioridade entre 2005 e 2009. ....	81
Gráfico 2: Avaliação dos documentos de patente quanto à relevância para malária	82
Gráfico 3: Avaliação dos documentos de patente quanto ao escopo de proteção com base na malária.....	84
Gráfico 4: Linguagem Natural X Linguagem Controlada .....	85
Gráfico 5: Número de documentos com as palavras representativas do conceito malária .....	87
Gráfico 6: Número de documentos com as palavras representativas do conceito diagnóstico .....	87
Gráfico 7: Localização das palavras-chave nos documentos de patente .....	91
Quadro 1: Linguagem controlada e natural: vantagens e desvantagens .....	23
Quadro 2: Comparação das aplicações dos distintos métodos quantitativos.....	35

Quadro 3: Principais Tratados Internacionais relativos à Propriedade Industrial .....	39
Quadro 4: Abordagens do uso da informação contida em documentos de patentes	47
Quadro 5: Códigos INID .....	50
Quadro 6: Seções da CIP .....	54
Quadro 7: Classes da CIP: a G01 .....	54
Quadro 8: Subclasses da CIP referente à classe G01 .....	55
Quadro 9: Grupos e Subgrupos da CIP referente à classe G01 .....	57
Quadro 10: Exemplos de bases públicas de patentes .....	58
Quadro 11: Exemplos de bases privadas de patentes .....	59
Quadro 13: Falhas da Ciência, mercado e sistemas de saúde relacionadas às alternativas e soluções propostas .....	64
Quadro 13: Códigos de "tags" da Derwent.....	73
Quadro 14: Estratégia de Busca formulada .....	77
Quadro 15: Ruídos na recuperação da informação e suas características .....	82
Quadro 16: Análise da ocorrência de palavras utilizadas na estratégia de busca.....	86
Quadro 17: Ocorrência de palavras recuperadas a partir da truncagem.....	92

**LISTA DE TABELAS**

Tabela 1: Análise preliminar da estratégia de busca.....	85
--	----

## LISTA DE ABREVIATURAS E SIGLAS

CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CIP	Classificação Internacional de Patentes
CIPO	<i>Canadian Intellectual Property Office</i>
CUP	Convenção da União de Paris
EPO	<i>European Patent Office</i>
Fiocruz	Fundação Oswaldo Cruz
Gestec	Coordenação de Gestão Tecnológica
INID	<i>International Agreed Numbers for the Identification of Data</i>
INPADOC	<i>International Patent documentation Center</i>
INPI	Instituto Nacional da Propriedade Industrial
IPA	Incidência Parasitária Anual
ISI	<i>Institute of Scientific Information</i>
LC	Linguagem Controlada
LN	Linguagem Natural
OMC	Organização Mundial do Comércio
P&D	Pesquisa e Desenvolvimento
<i>P. falciparum</i>	<i>Plasmodium falciparum</i>
<i>P. knowlesi</i>	<i>Plasmodium knowlesi</i>
<i>P. malariae</i>	<i>Plasmodium malariae</i>
<i>P. ovale</i>	<i>Plasmodium ovale</i>
<i>P. vivax</i>	<i>Plasmodium. vivax</i>
PCT	<i>Patent Cooperation Treat</i>
SCI	<i>Science Citation Index</i>
TRIPS	<i>Trade Related Aspects on Intellectual Property Rights, including Counterfeiting of Goods</i>
USPTO	<i>United States Patent and Trademark Office</i>
VPPIS	Vice-Presidência de Produção e Inovação em Saúde
VT	Vigilância Tecnológica
WHO	<i>World Health Organization</i>
WIPO	<i>World Intellectual Property</i>
WWW	<i>World Wide Web</i>

## SUMÁRIO

AGRADECIMENTOS .....	v
RESUMO.....	vii
ABSTRACT .....	ix
LISTA DE ILUSTRAÇÕES .....	xi
LISTA DE TABELAS .....	xiii
LISTA DE ABREVIATURAS E SIGLAS .....	xiv
1. INTRODUÇÃO.....	1
2. PREVISÃO/MONITORAMENTO .....	5
2.1. Prospecção Tecnológica .....	5
2.2. Monitoramento Ambiental .....	8
3. BUSCA DE INFORMAÇÃO .....	15
3.1. Estratégia de Busca de informação.....	15
3.2. A linguagem do sistema de recuperação da informação.....	20
3.3. A mineração de texto para recuperação da informação.....	24
4. METRIAS DE INFORMAÇÃO.....	27
4.1. Bibliometria .....	27
4.2. Cienciometria .....	32
4.3. Informetria .....	33
4.4. Webometria .....	34
5. PATENTES.....	37
5.1. Patente como fonte de informação tecnológica .....	42
5.2. A estrutura do documento de patente .....	49
5.3. Classificação Internacional de Patente (CIP).....	52
5.3.1. A estrutura da CIP .....	53
5.4. As bases de patentes públicas e privadas .....	58
6. MALÁRIA.....	61
6.1. Epidemiologia da Malária .....	62
6.2. Investimentos em Malária .....	64
6.3. Diagnóstico da malária.....	66
7. OBJETIVO GERAL.....	68
7.1. OBJETIVOS ESPECÍFICOS .....	68
8. CAMINHO DA PESQUISA.....	69



8.1. Derwent Innovation Index.....	69
8.2. Formulação da estratégia de busca e coleta.....	74
8.2.1. Etapa 01 .....	74
8.2.2. Etapa 02 .....	75
8.3. Tratamento.....	78
8.4. Análise .....	79
9. RESULTADOS E DISCUSSÃO .....	81
10. CONCLUSÃO.....	93
REFERÊNCIAS.....	96

## 1. INTRODUÇÃO

O início do século XXI marca para muitas empresas, organizações e instituições um tempo de grandes incertezas, ameaças e oportunidades geradas pela proliferação de um ambiente cada vez mais competitivo. Tal competitividade “obriga” as empresas, organizações e instituições a estarem constantemente em busca de formas de se manterem ou obterem vantagens sobre seus concorrentes.

Uma das formas de obter vantagens consiste na utilização de informações que vem do ambiente através do estudo de abordagens e processos de natureza prospectiva ou estudos do futuro para subsidiar a tomada de decisões e a formulação de políticas.

Dentre as atividades prospectivas que podem ser divididas em famílias de técnicas está o monitoramento ambiental também conhecido como monitoramento tecnológico. O monitoramento provê o pano de fundo necessário no qual a prospecção se baseia e pode ser usado para buscar fontes de informação e produzir um rico e variado conjunto de dados.

Portanto, é fundamental que as empresas, organizações e instituições mantenham um processo de monitoramento contínuo de seu ambiente competitivo, identificando e selecionando fontes de informações úteis e confiáveis para a tomada de decisão estratégica que aumentem as possibilidades de sobrevivência e o crescimento organizacional nos mercados em que atuam ao longo do tempo.

Para realizar o monitoramento de informação é necessário identificar qual a melhor tipologia de informação que representa o que se deseja monitorar. A tipologia da informação que melhor representa a competitividade e a obtenção de vantagens de uma instituição/empresa sobre seus concorrentes é a patente.

A patente, considerada a mais rica fonte de informação tecnológica, revela de acordo com sua estrutura a informação mais recente em um dado setor da técnica, o estado da arte. Permite tanto a identificação de técnicas alternativas para o atendimento às necessidades da indústria, como identificação de empresas que atuam em determinado segmento tecnológico entre outras utilidades. Portanto, verifica-se a relação entre a informação patentária, monitoramento e prospecção tecnológica.

Para monitorar a informação em documentos de patente faz-se uso de metrias de informação, mais especificamente, a cienciometria ou cientometria,

definida por Macias-Chapula (1998) como o estudo dos aspectos quantitativos da ciência enquanto uma disciplina ou atividade econômica. Para Callon *et al.* citado por Vanti (2002), as empresas fazem uso dos métodos cientiométricos para conseguir determinar qual a estratégia tecnológica seguida por seus competidores, mais especificamente, utilizando os documentos de patente como fonte sobre os principais temas de investigação e atores significativos em um segmento de seu interesse. Portanto, a informação se constitui o meio pelo qual as empresas reduzem as incertezas, é a base da tomada de decisão.

Para que o monitoramento de informação seja eficaz, além da tipologia da informação é necessário o uso de termos específicos, combinados com a linguagem da Classificação Internacional de Patentes (CIP) que possam expressar de uma forma específica todo o conceito que se deseja monitorar. Contudo é fundamental o estabelecimento de uma estratégia de busca de informação.

Uma das etapas decisivas para o êxito de uma estratégia de busca é a de planejamento da estratégia. Para Oldroyd e Citroen (1977) o processo de planejamento estratégico da busca é realizado por três etapas decisórias: decisão da melhor base para o contexto da busca; decisão das palavras-chave; decisão da fórmula lógica da estratégia.

Kirkbride (1991) é mais específico ao afirmar da necessidade de um conhecimento minucioso das técnicas de estratégia de busca, dos prós e contras da estratégia no texto completo dos documentos, do uso das linguagens natural e controlada na recuperação da informação.

Com base nas informações apresentadas, este trabalho possui como questão norteadora identificar como se deve elaborar uma estratégia de busca em documentos de patente para monitoramento. Mais especificamente, visa verificar o papel do título e da classificação na recuperação de informação de patentes.

Como estudo de caso para a realização da estratégia de busca foi utilizado como base o diagnóstico da malária. Também conhecida como paludismo, a malária é uma doença infecciosa aguda ou crônica causada por protozoários parasitas do gênero *Plasmodium* que são transmitidos para o ser humano através da picada da fêmea do mosquito *Anopheles*.

A escolha da malária justifica-se por ser reconhecida como grave problema de saúde pública no mundo. Segundo Brasil (2010), a malária ocorre em quase 50% da população, em mais de 109 países e territórios. Sua estimativa é de 300 milhões de

novos casos e 1 milhão de mortes por ano, principalmente em crianças menores de 5 anos e mulheres grávidas do continente africano.

Quanto ao diagnóstico, justifica-se principalmente pela necessidade de identificar o movimento de patenteamento e consequente entrada no mercado de diagnósticos para detecção da malária, uma vez que devido às necessidades de controle da malária, o diagnóstico parasitológico é de grande importância para avaliar a eficiência e eficácia dos programas ou medidas de controle, bem como no estabelecimento da terapêutica e evolução das manifestações clínicas, principalmente em áreas rurais e remotas, de alta endemicidade. Um diagnóstico preciso influencia no tratamento mais adequado e, segundo Vitzthum *et al.* (2005), pode-se alcançar reduções nos custos das despesas de saúde em todo o mundo.

Portanto confirma-se a necessidade de um monitoramento contínuo de informação que, além de estratégico para o reposicionamento competitivo das mesmas, cria barreiras para a entrada de novas empresas no mercado.

Não se objetiva aqui identificar quais tecnologias ou quais atores estão neste mercado, nem quais os mercados promissores, por não se tratar de um estudo prospectivo ou de panorama de uma tecnologia específica.

Concretamente, esta dissertação tem como meta auxiliar os intermediários no processo da formulação de estratégias de busca em documentos de patentes para a tomada de decisão em empresas e instituições de ciência e tecnologia. Em especial, espera-se que seus resultados sejam apropriados pela Fundação Oswaldo Cruz (Fiocruz), casa de origem desta pesquisa que desde o final da década de 80 tem manifestado em sua estrutura organizacional, por meio da Coordenação de Gestão Tecnológica (Gestec), a preocupação com a proteção do patrimônio intelectual gerado na instituição.

A Gestec, vinculada a Vice-Presidência de Produção e Inovação em Saúde (VPPIS), é o núcleo central de inovação tecnológica na Instituição que, em resposta ao estudo de Emerick (2004), passou por um processo de reestruturação como forma de propiciar um incremento da capacidade adaptativa da Instituição aos novos desafios e oportunidades presentes nos ambientes interno e externo.

Neste sentido foi criada a área de Informação Tecnológica que, segundo Longa (2007), tem por finalidade dar subsídios às áreas de Patentes, através da proteção às pesquisas; Transferência de Tecnologia, através das negociações e

prospecções de parcerias, e Núcleos de Inovação das Unidades Técnico-Científica, através do fortalecimento de pesquisas/projetos da Fundação Oswaldo Cruz.

A dissertação está organizada em 8 capítulos, além desta introdução. No capítulo 2 apresenta-se a parte do referencial teórico que diz respeito à previsão e monitoramento, a sua importância para empresas/instituições/organizações na obtenção de vantagem sobre seus concorrentes na conquista ou consolidação de mercados.

As questões teóricas referentes à metrias de informação e seu uso para avaliar a ciência e os fluxos da informação em estudos prospectivos como forma de auxiliar a tomada de decisão são apresentadas no capítulo 3.

No capítulo 4 são apresentados os principais fundamentos teóricos de busca de informação, estratégia de busca, uso da linguagem natural e controlada e recursos de recuperação da informação em bases de dados para auxiliar no processo de monitoramento para tomada de decisão.

A patente, suas características, a legislação nacional, sua estrutura como um documento fonte de informação, principais bases de dados públicas e privadas disponíveis é apresentada no capítulo 5.

No capítulo 6 é apresentada a malária como um problema grave de saúde pública no mundo, seu mecanismo de transmissão e a importância do diagnóstico para malária como um estudo de caso selecionado para este trabalho.

No capítulo 7 apresenta-se a metodologia, identifica-se as principais características da base de patentes *Derwent Innovation Index*, base escolhida para a coleta de informação para este estudo, os passos preliminares para a formulação da estratégia de busca utilizando a estratégia conceituada com “*snow ball*”; o uso de minerador de texto para tratamento dos dados e os critérios de análise.

Os resultados e discussões são apresentados no capítulo 8 que revelam, por exemplo, que a maior parte dos documentos de patente não são específicos para malária, nem tão pouco para o diagnóstico e que, o maior retorno do uso dos termos utilizados na estratégia de busca foi no resumo e não no título.

As conclusões do trabalho são apresentadas no capítulo 9. Com base nos resultados, dentre os pontos apresentados, verificou-se que para uma otimização da estratégia de busca faz-se necessário o uso combinado da linguagem natural com a linguagem controlada.

## 2. PREVISÃO/MONITORAMENTO

Para muitas organizações, o século XXI chega com grandes incertezas, ameaças e oportunidades. Como afirmado por Braga (2008), o ambiente traz diversas incertezas geradas por inúmeras transformações ocorridas, inerentes ao ambiente, e para conviver com estas incertezas a instituição deve se organizar, usando informações que vem do ambiente para manter ou transformar seus processos e estruturas.

No ambiente estão localizadas informações que podem representar várias vantagens e subsidiar decisões e trazer importantes oportunidades. Willian R. Dill, um dos precursores do tema, identifica o ambiente como uma importante fonte de informação:

Deve-se tratá-lo como informação que se torna disponível para a organização ou que através de uma atividade formal de pesquisa, a organização pode ter acesso. O importante é a informação que o ambiente torna disponível para organização ser estudada sob a perspectiva de seus objetivos, são as condições sob as quais a organização irá trabalhar (DILL, *apud* BRAGA, 2008, p.6)

Para Nadler, Gerstein e Shaw (1993 *apud* OLIVEIRA, PAULA NETO OLIVEIRA, 2008), o ambiente pode afetar o funcionamento da organização de três maneiras distintas: (i) por meio da exigência de novos produtos ou serviços; (ii) por meio de limitações à ação organizacional como escassez de capital ou de tecnologia; (iii) por meio da oferta de oportunidades a serem exploradas pela organização.

De acordo com o contexto, inovar nos métodos e processos de gestão é, portanto, um dos desafios face às exigências de um mercado globalizado, com concorrência acirrada, elevado grau de incertezas e abundância de informação. (CANONGIA *et al.*, 2002, p. 156).

### 2.1. PROSPECÇÃO TECNOLÓGICA

As grandes incertezas, ameaças e oportunidades para muitas organizações no século XXI têm gerado a proliferação do ambiente competitivo, no qual somente

os mais adaptados terão maiores chances de sobreviver e de crescer nesse ambiente em constante e crescente evolução.

Para Branício, Peixoto e Carpinetti (2001), a competitividade relacionada com o produto e a disputa de segmentos comerciais, entre outros, fazem com que as empresas busquem formas de obter vantagem sobre seus concorrentes na conquista ou consolidação de mercados.

Neste contexto, Santos *et al.* (2004) dizem que as mudanças profundas no cenário internacional, particularmente no que tange à globalização da economia e à aceleração das mudanças tecnológicas propiciou o estudo de abordagens e processos de natureza prospectiva ou estudos do futuro para subsidiar a tomada de decisões e a formulação de políticas.

Para Santos *et al.* (2004), a capacidade de antecipar vem se tornando uma qualidade de extrema importância para assegurar a competitividade de empresas e países. As abordagens e processos de natureza prospectiva buscam entender as forças que orientam o futuro, visam promover transformações, negociar espaços e dar direção e foco às mudanças. (SANTOS *et al.*, 2004, p. 189). Ainda segundo os autores:

Estudos prospectivos são conduzidos de modo a "construir conhecimento", ou seja, buscam agregar valor as informações do presente, transformando-as em conhecimento de modo a subsidiar os tomadores de decisão e os formuladores de políticas na construção de suas estratégias, e identificar rumos e oportunidades futuras para os diversos atores sociais (SANTOS *et al.*, 2004, p.189).

Os autores mencionam também que no âmbito de sistemas de Ciência, Tecnologia e Inovação (CT&I), os exercícios prospectivos são fundamentais para promover a criação da capacidade de organizar sistemas de inovação que respondam aos interesses da sociedade. Fazer prospecção significa identificar quais são as oportunidades e necessidades mais importantes para a pesquisa e desenvolvimento (P&D) no futuro. (SANTOS *et al.*, 2004, p.189)

Segundo Kupfer e Tigre (2004), a prospecção tecnológica pode ser definida como:

... um meio sistemático de mapear desenvolvimentos científicos e tecnológicos futuros capazes de influenciar de forma significativa uma indústria, a economia ou a sociedade como um todo. Diferentemente das atividades de previsão clássica, que se dedicam a antecipar um futuro

suposto como único, os exercícios de prospecção são construídos a partir da premissa de que são vários os futuros possíveis. Esses são tipicamente os casos em que as ações presentes alteram o futuro, como ocorre com a inovação tecnológica. Os exercícios de prospecção funcionam como meio de atingir dois objetivos: O primeiro é preparar os atores na indústria para aproveitar ou enfrentar oportunidades ou ameaças futuras. O segundo objetivo é desencadear um processo de construção de um futuro desejável. (KUPFER; TIGRE, 2004, p.1)

Com relação aos benefícios dos exercícios de prospecção em ciência, tecnologia e inovação, Santos *et al.* (2004) consideram que podem servir para:

- promoção de canais e linguagens comuns para a circulação de informação e conhecimento de caráter estratégico para a inovação;
- mais inteligência antecipatória inserida no processo de tomada de decisão em ciência, tecnologia e inovação;
- incorporação crescente de visões de futuro no pensamento dos atores sociais envolvidos no processo de tomada de decisão e de criação de redes;
- apoio a decisões relativas ao estabelecimento de prioridades para P&D, gestão dos riscos das inovações tecnológicas, melhoria da competitividade tecnológica de produtos, processos e serviços.

Segundo Kupfer e Tigre (2004), existem três visões diferentes para o problema de prospectar o futuro que são consagradas e aceitas no meio especializado. A visão mais básica está baseada em inferência, onde acredita-se que o futuro busca reproduzir, de certa forma, situações que aconteceram no passado; este modelo não possui a característica de identificar rupturas ou descontinuidades na evolução dos objetos sob análise. A segunda visão está relacionada à geração sistemática de trajetórias alternativas, constrói cenários para representar as possíveis variações do futuro. A terceira e última visão é a que orienta o futuro por consenso, utiliza como base opiniões coletadas através do processo cognitivo e intuitivo de um grupo de especialistas.

As metodologias e as técnicas preferencialmente adotadas para prospecção segundo Santos *et al.* (2004) contemplam:

- A convergência de esforços para gerar orientações e recomendações;



- Um processo interativo de comunicação e articulação de atores para maximizar a disseminação de informações estratégicas;
- A promoção da criatividade e da busca permanente de novas oportunidades.

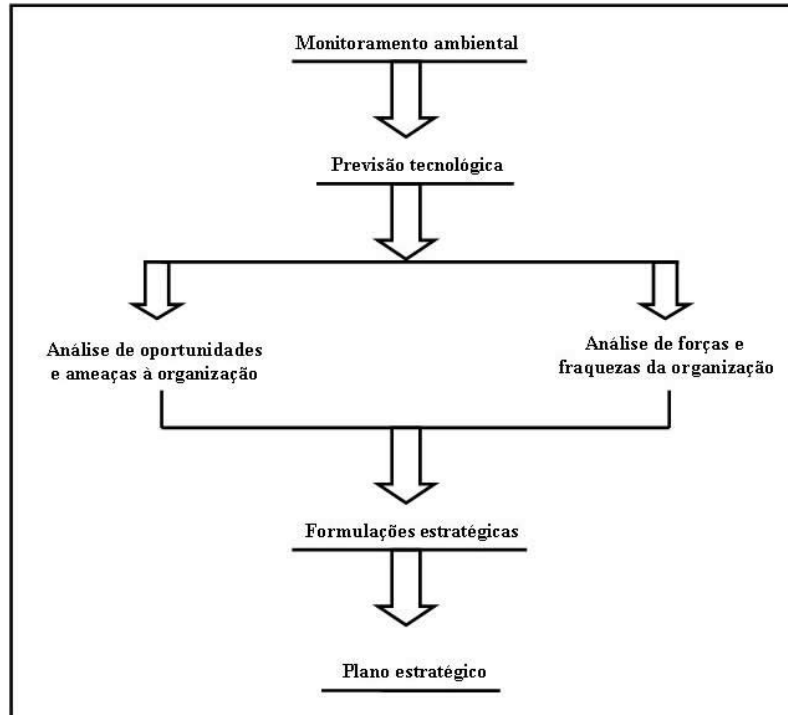
Na literatura da área, são encontrados estudos que propuseram a classificação dos métodos e técnicas existentes e em uso nas atividades prospectivas (Porter *et al.* (1991 e 2004), Skumanich e Sibernagel (1997) e Coelho (2003).

Mais recentemente, Porter *et al.* (2004, *apud* SANTOS *et al.*, 2004, p.197) propuseram a divisão das atividades prospectivas em famílias de técnicas que são: Criatividade, Métodos Descritivos e Matrizes, Métodos Estatísticos, Opinião de Especialistas, Monitoramento e Sistemas de Inteligência, Modelagem e Simulação, Cenários, Análises de Tendências, e Sistemas de Avaliação e Decisão.

Segundo Porter *et al.* (2004), o monitoramento provê o pano de fundo necessário no qual a prospecção se baseia e pode ser usado para buscar todas as fontes de informação e produzir um rico e variado conjunto de dados. As principais fontes em que se baseia são as de natureza técnica, tais como revistas, patentes, catálogos, artigos científicos, entre outros. Além disso, podem ser feitas entrevistas com especialistas e outras informações não literárias podem ser coletadas.

## 2.2. MONITORAMENTO AMBIENTAL

O monitoramento ambiental tem sido considerado uma das técnicas de previsão tecnológica pelos livros textos (MARTINO, 1983; JONES, 1978, *apud* GOODRICH, 1987, p. 5) e também é englobada em uma série de outros títulos tais como *enviromental scanning* (THOMAS, 1980; FAHEY, KING, NARAYANAN, 1981, *apud* GOODRICH, 1987, p. 5). A figura 1 a seguir ilustra o papel fundamental de monitoramento ambiental e previsão tecnológica no processo de planejamento estratégico.



**Figura 1 - Monitoramento ambiental e previsão tecnológica no processo de planejamento estratégico. Goodrich (1987)**

Em seu contexto histórico, Goodrich (1987) explica que o monitoramento ambiental, também conhecido como *environmental scanning*, iniciou-se em meados da década de 60, em países desenvolvidos, com o objetivo de atender demandas e necessidades de melhor acompanhar as rápidas e abruptas mudanças que acontecem no ambiente externo, em áreas políticas, econômicas, tecnológicas ou sociais que sejam de importância nacional ou internacional.

Em termos conceituais, Goodrich (1987, p. 6) afirma que o conceito é simples:

Identificar, acompanhar e analisar sinais de alarme precoce no ambiente. Estes sinais são os precursores de tendências e eventos emergentes que possam ter relevância futura no desenvolvimento dos negócios da organização. Como tal, precisam ser selecionados cuidadosamente dentre a abundância de informação bruta existente e analisada quanto a sua potencial relevância, antes que se executem previsões detalhadas para caracterizar as tendências e os eventos emergentes, e para especular sobre suas prováveis consequências para a organização.

Coates *et al.* (1986), citado por Porter *et al.* (1991), define que monitorar significa observar, checar e atualizar-se em relação aos desenvolvimentos numa área de interesse bem definida para uma finalidade bem específica. Porter *et al.* (1991) corroboram esta afirmação e consideram monitoramento como exploração do

ambiente apropriado para informação pertinente, que pode dizer respeito a uma tecnologia em particular, monitoramento tecnológico ou no ambiente em que esta se desenvolve – monitoramento contextual ou *issues management*.

Na definição de Coates (1986) “*issues management*” é considerado como:

*“The organized activity of identifying emerging trends, concerns, or issues likely to affect an organization in the next few years and developing a wider and more positive range of organizational responses toward that future. Issues management can help companies devise positive, anticipatory (rather than merely reactive) responses to new technologies, possible governmental constraints, or potential confrontations. Issues management emphasizes contextual monitoring (law-making, social controversies, etc.), but it can include technological monitoring, too.”* (COATES, 1986 apud SILVA, 2002, p. 67).

Na visão de muitos autores (AGUILAR, 1967; AUSTER; CHOO, 1993, *apud* OLIVEIRA, PAULA NETO e OLIVEIRA, 2008) a atividade de monitoramento ambiental (*environmental scanning*) normalmente é entendida na literatura como a aquisição e o uso da informação sobre eventos, tendências e relacionamentos que acontecem no ambiente externo da organização.

Em relação às organizações, Aguilar (1967, *apud* OLIVEIRA, PAULA NETO e OLIVEIRA, 2008) argumenta que o monitoramento é realizado em seus ambientes com a finalidade de entender as mudanças das forças externas a organização que afetam o seu funcionamento para que seja possível desenvolver respostas rápidas e precisas de forma a manter posições privilegiadas nos mercados em que atuam ao longo do tempo.

Nesta mesma perspectiva (OLIVEIRA *et al.*, 2006; MINTZBERG, 2003 *apud* OLIVEIRA, PAULA NETO e OLIVEIRA, 2008) a organização deve buscar a informação para alimentar suas estratégias que promovam um ajuste eficiente e eficaz entre os seus componentes e as variáveis ambientais.

Atualmente, os fatores imprevisibilidade e complexidade ambientais têm demandado posturas mais flexíveis, criativas e inovadoras por parte dos gestores e das organizações como um todo e, para tanto, Oliveira, Paula Neto e Oliveira (2008) afirmam que é fundamental que as organizações mantenham um processo de monitoramento contínuo de seu ambiente competitivo, identificando e selecionando fontes de informações úteis e confiáveis para a tomada de decisão estratégica que

aumentem as possibilidades de sobrevivência e o crescimento organizacional nos mercados em que atuam ao longo do tempo.

Assim como Oliveira, Paula Neto e Oliveira (2008), Sutton (1988) defende a necessidade de se manter um processo de monitoramento ambiental contínuo para ajudar os gestores a evitar surpresas, identificar ameaças e prover novas oportunidades de negócios, além de proporcionar a sustentabilidade das vantagens competitivas mediante atividades de planejamento de curto e longo prazos. Na visão de Auster e Choo (1993), a capacidade de adaptação organizacional é variável dependente das mudanças que acontecem em seu ambiente externo e está diretamente relacionada à sua base de conhecimentos sobre esse ambiente e de como tais mudanças são interpretadas, constituindo-se, segundo o autor, em um modo primário de aprendizagem organizacional.

Bright (1973), citado por Braga (2008), foi um dos primeiros a conceituar Monitoramento Ambiental para um ambiente empresarial. Em seus estudos, ele estabelece quatro atividades do processo de monitoramento que são:

- Busca no ambiente por sinais que possam ser antecipadores de uma significativa mudança tecnológica;
- Identificação das possíveis conseqüências destes sinais, se eles não são espúrios e se as tendências por eles sugeridas terão continuidade;
- Escolha dos parâmetros políticos, eventos e decisões que deveriam ser seguidos ordenadamente para verificar a verdadeira velocidade e direção da tecnologia e os efeitos do seu emprego;
- Apresentação dos dados a partir dos passos precedentes em tempo hábil e forma apropriados para que os gerentes possam usá-los nas decisões sobre a reação da organização.

Nesta mesma linha de pensamento, Santos *et al.* (2004) aponta alguns objetivos possíveis de monitoramento:

- Identificar eventos científicos, técnicos ou sócio-econômicos importantes.
- Definir ameaças potenciais, implícitas nesses eventos.
- Identificar oportunidades envolvidas nas mudanças no ambiente.

- Alertar os decisores sobre tendências que estão convergindo, divergindo, ampliando, diminuindo ou interagindo.

Palop e Vicente Gomilla (1999) utilizam o termo vigilância tecnológica (VT), afirmando que a vigilância é a forma organizada, seletiva e permanente de captar informação externa, analisá-la e convertê-la em conhecimento para tomar decisões com menor risco, permitindo antecipar-se a mudanças.

Palop e Vicente Gomilla (1999) lembram que a correta interpretação e difusão das informações melhoram a capacidade de visão e de antecipação da empresa, sem que se precise recorrer a práticas pouco éticas de obtenção de informação sobre competidores, estratégias, etc. SANTOS JR. (1996) citado por Palop e Vicente Gomilla (1999) reforça a necessidade e importância da vigilância, notando que médias e grandes empresas “precisarão de informações sobre as tendências mundiais, sobre concorrência, mercados consumidores potenciais e inovações tecnológicas, para enfrentar os novos desafios após a globalização da economia”.

Voltando as afirmações de Palop e Vicente Gomilla (1999), a VT é apontada como um instrumento de inovação para o desenvolvimento de produto e como toda atividade a ser desenvolvida na empresa, a VT precisa ter objetivos claros e razões para ser implementada.

Para Goodrich (1987) o processo de monitoramento ambiental é formado por inputs e outputs como pode ser visto na figura 2

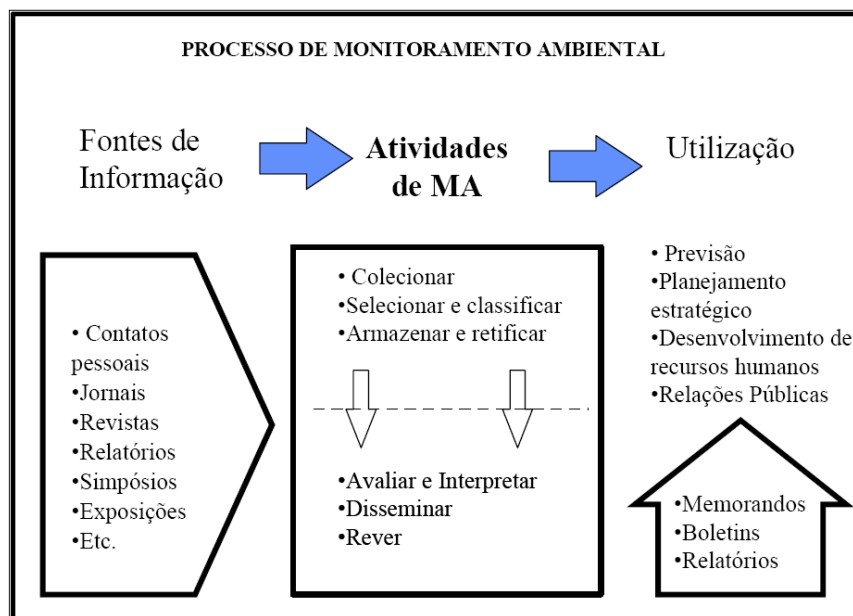


Figura 2: Processo de Monitoramento Ambiental. Goodrich (1987) adaptado por Braga (2008)

Ainda segundo Goodrich (1987) a série de atividades no centro da figura 2 são considerados elementos básicos no processo de monitoramento ambiental. A parte superior da série constitui a aquisição e o armazenamento da base de informação necessária para a alimentação da fase analítica do processo, na porção inferior da figura.

O autor destaca ainda que o processo de monitoramento ambiental apresentado em sua forma idealizada é designado a ser uma ajuda aos planejadores e tomadores de decisão e não um substituto para o pensamento criativo e a análise crítica.

A maior parte dos trabalhos contextualizam suas pesquisas na área do setor privado, mas Braga (2008) sinaliza a utilização destas técnicas por parte do setor público, com o intuito de reavaliar programas de trabalho, tendo como base a aquisição de informações ambientais. O setor público, mais especificamente o da saúde tem o compromisso de desenvolver instrumentos que possibilitem combater a marca da desigualdade no campo da saúde (aumentar os padrões de equidade do sistema de saúde). De acordo com Brasil (2007) a promoção da equidade requer que se assegure a todas as pessoas, independentemente de sua formação, etnia, gênero, local de moradia, raça ou posição social, a proteção adequada contra os fatores de adoecimento; o acesso a conhecimentos, produtos e serviços que as habilitem a reduzir os fatores de risco e a obter aconselhamento e tratamento; e a certeza de que não serão impedidas, por falta de recursos ou por outros obstáculos, de utilizarem o que está disponível para que alcancem e mantenham a boa saúde e otimizem o desenvolvimento pessoal.

Braga (2008, p. 3) é categórica ao afirmar que: "A sociedade anseia e espera que os órgãos do setor público tenham as mesmas preocupações e responsabilidades em melhorar os seus serviços e atendimentos, tanto quanto tem as organizações privadas".

Ainda segundo a autora para que o monitoramento ambiental ocorra de forma plena, os administradores públicos terão de desenvolver habilidades e ferramentas que os auxiliem na avaliação de seus desempenhos. Isto é, os gerentes devem empregar esforços no desenvolvimento de metodologias para obter informações, as quais serão utilizadas nas tomadas de decisão. Sua habilidade para responder às pressões da sociedade por uma prestação de serviço eficiente e eficaz dependerá

da sua habilidade em obter as informações de que precisam, no momento em que dela precisarem.

A questão que se coloca neste trabalho é que para se fazer prospecção é preciso monitorar as fontes e para monitorá-las é necessário selecioná-las e buscar a informação disponível, de forma adequada. Quais os requisitos importantes no planejamento de uma estratégia de busca em uma fonte formal, como as bases de dados de patentes?

### 3. BUSCA DE INFORMAÇÃO

Informação é o redutor de incertezas, é a base da tomada de decisão. Para Faibisoff e Ely (1978) informação pode existir como dados em livros, computadores, pessoas, arquivos e milhares de outras fontes. Essas fontes são consideradas como dados crus até que sejam usadas para resolver incertezas.

A informação bibliográfica, considerada apenas uma etapa na busca da informação propriamente dita, se torna a cada dia mais imprescindível para o usuário, principalmente em áreas especializadas do conhecimento. A explosão bibliográfica ocorrida no final da década de setenta e início da década de oitenta mudou consideravelmente o cenário referente à falta de informação que havia no passado. Kremer (1985) afirma que a explosão bibliográfica ocorrida foi tão grave a ponto dos cientistas, às vezes, alegarem que levam menos tempo fazendo um experimento, do que tentando descobrir se ele já foi realizado. Para remediar o problema fez-se necessário a criação e a utilização de mecanismos eficientes de seleção de documentos pertinentes para o usuário.

As bases de dados bibliográficos são elementos imprescindíveis para o controle bibliográfico e para a recuperação da informação (KREMER, 1985, p. 192). As bases e bancos de dados se tornaram um marco para a qualidade das buscas bibliográficas, já que de acordo com Lopes (2002) proporcionam diversificados pontos de acesso à informação.

Quando automatizadas, oferecem a possibilidade de se poder localizar e selecionar eficientemente os documentos que melhor atendam às necessidades dos usuários. Para Kremer (1985) a eficiência depende da qualidade e dos recursos oferecidos pelas bases de dados, mas também depende da qualidade das estratégias de busca utilizadas.

#### 3.1. ESTRATÉGIA DE BUSCA DE INFORMAÇÃO

Para o entendimento da questão relacionada com o processo de planejamento das estratégias de busca, faz-se necessário a apresentação dos conceitos Estratégia e Busca que são fundamentais.

Segundo Tarapanoff (2002), estratégia é definida como um conjunto de ações e atitudes que buscam, ao longo prazo, alcançar objetivos para o sucesso de um



empreendimento e que envolvem diversos níveis de atuação: decisão, avaliação, pessoal, necessidades informacionais, etc. De acordo com o Dicionário Aurélio *online* (1996) estratégia é: 1) arte militar de planejar e executar movimentos e operações de tropas, navios e/ou aviões, visando a alcançar ou manter posições relativas e potenciais bélicos favoráveis a futuras ações táticas sobre determinados objetivos; 2) arte militar de escolher onde, quando e com que travar um combate ou uma batalha; 3) arte de aplicar os meios disponíveis com vista à consecução de objetivos específicos. Já busca é definida como: 1) ato ou efeito de buscar; 2) procura com o fim de encontrar alguma coisa; 3) investigação cuidadosa: pesquisa, exame; 4) procura minuciosa: revista, exame; 5) movimento íntimo para alcançar um fim.

Hawkins (1981) define estratégia de busca como o meio pelo qual o pesquisador dialoga com o sistema, e é muitas vezes a chave para uma busca bem sucedida. Para Kremer (1985), estratégia de busca é o procedimento através do qual se procuram documentos determinados ou dados sobre um assunto. Lopes (2002) define estratégia de busca no âmbito da recuperação da informação como uma técnica ou conjunto de regras para tornar possível o encontro entre uma pergunta formulada e a informação armazenada em uma base de dados.

A busca em sistemas de recuperação de informação é:

“Um processo de alta complexidade envolvendo numerosos fatores e variáveis além de decisões e o entrelaçamento dos subprocessos inter-relacionados com a busca. Uma das chaves para um dos subprocessos é a seleção de termos para a estratégia de busca que por sua vez é influenciada por outros fatores, particularmente os relacionados com os resultados. Sintetizam essa questão com uma pergunta para as investigações: - Que termos de busca devem ser selecionados para um determinado tema que represente efetivamente o problema de informação do usuário?” (Spink e Saracevic, 1993, p. 63).

É interessante enfatizar que, quanto maior for a sofisticação de recursos existentes em um sistema de recuperação de informação, maior será a importância das estratégias utilizadas (KREMER, 1985, p. 199).

Portanto para a execução de uma estratégia de busca é necessário conhecimento e perfeito domínio de todos os recursos de busca disponíveis. Além de planejamento é necessário fazer escolhas, por exemplo, a base/banco de dados mais adequada; a abrangência do assunto; tipos de documentos indexados; campos de busca disponibilizados e a linguagem empregada. Para Lopes (2002) a estratégia

de busca requer um constante julgamento na sua utilização, pois o intermediário tem que tomar a decisão sobre o melhor momento de implementá-la durante o processo de busca. Essas escolhas podem determinar a eficácia da busca de informações independente do tipo de solicitação. Para atingir a eficácia, faz-se necessário ora restringir os resultados alcançados, ora ampliá-los para obter informações mais relevantes, conforme o pedido de busca demandado.

Dirnberger (2011) cita Van der Drift (1991) e Ziman (1969) afirmam que para avaliar a qualidade da estratégia de busca na recuperação da informação são utilizadas as medidas de precisão e a de revocação. Ainda segundo Dirnberger (2011), de acordo com a definição acima a meta de qualquer estratégia de busca seria uma revogação elevada e uma alta precisão, isto é, a recuperação de praticamente todos os documentos relevantes. No entanto, revocação e precisão tendem a se correlacionar inversamente, ou seja, uma maior precisão leva a uma menor revocação e vice-versa, o que significa que o pesquisador de patentes tem de encontrar o equilíbrio adequado entre estes polos de acordo com o tipo de pesquisa escolhida.

Kirkbride (1991) afirmou ser necessário um conhecimento minucioso das técnicas de estratégia de busca, dos prós e contras da estratégia no texto completo dos documentos, da busca em linguagem controlada, da política de indexação, da cobertura de assunto das bases de dados e das características das linguagens de recuperação dos diversos sistemas disponíveis. Em algumas ocasiões, um tema de pesquisa ou um tipo de formato solicitado exigem uma pré-seleção da base e do banco de dados, antes mesmo de se planejar qual técnica será usada na formulação da estratégia de busca.

Com relação às etapas decisórias no processo de planejamento estratégico da busca, Oldroyd e Citroen (1977) identificaram três etapas: decisão da melhor base para o contexto da busca; decisão das palavras-chave; decisão da fórmula lógica da estratégia.

Para Adams (1979) citado por Lopes (2002), a estratégia de busca deve ser iniciada com a etapa de entrevista, na qual o usuário deve explicar ao intermediário da busca, o seu tema de interesse e o funcionamento da lógica da pesquisa. Knox e Hlava (1979) demonstraram que a definição clara e precisa da necessidade de informação do usuário é fundamental para o planejamento e execução da estratégia de busca.

Na mesma linha de pensamento, Kremer (1985) afirma que buscas de informação bem sucedidas dependem, em primeiro lugar, de uma eficiente determinação das necessidades do usuário. Ainda segundo a autora, nem sempre as solicitações refletem as suas reais necessidades, muitas vezes o próprio usuário não é capaz de identificá-las com precisão.

O problema da diferença entre demanda e necessidade também costuma ocorrer por causa de ruídos na comunicação entre usuários e os serviços de informação, causados por deficiência ou incorreções no uso da terminologia técnico-científica, ou porque as pessoas tendem a pedir somente aquilo que pensam poder obter, sem conhecer realmente todos os recursos informacionais que se encontram à disposição. (KREMER, 1985, p. 193)

Outro problema apontado pela autora se refere ao escopo da solicitação, ou apresentam apenas uma parcela do que é necessário para o planejamento da estratégia ou não há especificação da real necessidade fazendo a solicitação de forma ampla e geral, quando apenas uma parte dessa demanda atenderia à necessidade o que, geralmente, pode gerar desperdício de capital intelectual humano, recursos financeiros e demora ao atendimento da solicitação. Cabe lembrar que uma solicitação de busca para tomada de decisão requer rapidez e agilidade na resposta à solicitação.

Kremer (1985) defende a idéia de que o ideal, nem sempre alcançado, seria que os usuários fossem capazes de realizar suas próprias buscas de informação, pois eles, melhor do que ninguém podem conhecer suas verdadeiras necessidades. Outra vantagem é que eles conhecem, ou deveriam conhecer a terminologia em suas áreas de especialização melhor do que os especialistas. Para tanto, é importante que os usuários sejam treinados para o uso das fontes de informação disponíveis e dos recursos de busca existentes nos sistemas de recuperação da informação. Mesmo que os usuários não realizem as buscas, o treinamento continua pertinente para conscientização da necessidade de uma explicação mais detalhada nas solicitações de busca. Se falhar a comunicação entre os usuários e o serviço de informação, será impossível alcançar a eficiência no seu atendimento (KREMER, 1985, p. 208).

Lancaster (1979) aponta a existência de fatores que influenciam no sucesso da interação usuário-sistema na solicitação de informação, são eles:

- A habilidade do usuário em definir sua necessidade para si mesmo;

- A habilidade do usuário em se expressar;
- As expectativas do usuário a respeito das capacidades do sistema. Há uma forte tendência para o usuário pedir, não o que realmente quer, mas aquilo que ele pensa que o sistema poderá lhe dar; e
- A quantidade e tipo de assistência que lhe é prestada pelo sistema.

Neste contexto, a preparação adequada é indispensável tanto para a busca computadorizada quanto para a busca manual. (LOPES, 2002, p. 68)

Para a preparação da estratégia de busca, o usuário deve fornecer preferencialmente, em um formulário específico, os seguintes dados: escrever um título sucinto; escrever uma pequena definição do problema; listar os termos mais apropriados para o tópico de interesse; listar os termos que não são desejados. A partir desses dados, o intermediário poderá executar a busca de informação de uma forma relativamente simples. (LOPES, 2002, p. 68)

Lopes (2002) apresenta orientações organizadas em etapas para auxiliar na definição do problema contribuindo para acentuar a qualidade da busca. Na primeira etapa, denominada discussão do tópico geral da pesquisa, há uma interação entre o usuário e o intermediário para a solução das eventuais dúvidas decorrentes do formulário específico. A interação pode ser através de uma entrevista, chamada telefônica ou até mesmo por correio eletrônico. É útil perguntar como os resultados da busca irão ser aplicados, porque a resposta pode mudar a direção ou a ênfase da busca.

A segunda etapa, definida como conhecimentos básicos sobre os instrumentos de busca, o usuário, caso tenha algum conhecimento dos instrumentos, deve informar as palavras-chave, já que o mesmo tem a expertise no assunto; o intermediário deve ajudar, mas não deve definir o assunto, porque, na maioria das vezes, a definição obtida para o tema difere completamente do pesquisador.

Na terceira etapa, formulação “provisória” da estratégia de busca, o intermediário da busca formula uma estratégia de busca, inicial, com base nas informações coletadas no formulário e/ou em consulta ao usuário. Caso o intermediário seja capaz de assegurar a recuperação de todas as citações para vários termos, a busca está definida; caso contrário o intermediário deve ajudar ou orientar o usuário a compor um outro conjunto de termos até alcançar a resposta

solicitada. É desejável que a informação solicitada ocorra em ambos os conjuntos ao mesmo tempo.

Na quarta etapa, compreensão da lógica dos conjuntos de termos, o intermediário deve orientar o usuário sobre as propriedades básicas da teoria dos conjuntos. O uso da interseção de mais de dois conjuntos de termos deve ser evitado, porque, embora os resultados possam ser bem precisos, eles são muito limitados e podem provocar uma possível exclusão de informações relevantes.

Na quinta etapa, interdisciplinaridade, os intermediários precisam conhecer os campos e como responder às buscas interdisciplinares. O intermediário deve propor a expansão da busca para campos específicos, aumentando, conseqüentemente, as possibilidades de documentos de interesse virem a ser recuperados.

A sexta etapa é a eliminação de termos indesejados. Há diferentes formas de eliminação desses termos, dentre elas está a falta de interesse sobre determinado conceito. A eliminação de termos indesejados pode gerar um problema no resultado final, porque termos solicitados na busca podem estar presentes junto na citação. Neste contexto, juntos, usuário e intermediário examinarão a necessidade de exclusão dos termos. A decisão para excluir termos nem sempre é fácil e, visualmente, depende da especificação do tópico.

A sétima e última etapa é a especificação dos parâmetros relevantes para a execução da busca. Todos os parâmetros relevantes devem ser considerados para se determinarem os limites da busca, dentre eles estão os recursos financeiros para a realização da busca; limitação da busca nos anos mais recentes; a abrangência da base de dados entre outros.

### 3.2. A LINGUAGEM DO SISTEMA DE RECUPERAÇÃO DA INFORMAÇÃO

Segundo Kremer (1985) os sistemas de recuperação da informação exercem várias atividades, que podem ser classificadas conforme suas funções de *input* (entrada) e *output* (saída). Na função de *input*, por exemplo:

Os documentos são organizados e controlados de forma que possam ser identificados e localizados em resposta às demandas dos usuários. As atividades de organização e controle dos documentos incluem sua catalogação, classificação, indexação e resumo. O processo de indexação envolve dois passos distintos: a análise do conteúdo temático dos documentos e a tradução do resultado dessa análise para o vocabulário do sistema de recuperação da informação. (KREMER, 1985, p. 188)

Lancaster (1993) ressalta que a qualidade da estratégia de busca e o vocabulário são fatores importantes para a recuperação da informação. A maioria dos sistemas de recuperação da informação utiliza o vocabulário controlado ou linguagem controlada para a indexação dos documentos ao invés do vocabulário não controlado ou linguagem natural. Como forma de aprimorar os sistemas de recuperação da informação, foram verificados alguns estudos sobre o uso das linguagens na formulação das estratégias de busca: a linguagem controlada (LC) e natural (LN).

Para Lopes (2002, p.41), a LC caracteriza-se como a utilizada apenas nos campos de descritor, termos de indexação e identificadores, sendo que a LN abrange os termos do título e do resumo dos documentos referenciados.

Nos artigos científicos, a principal representação do conteúdo do artigo está no título. Para Pessoa Jr. (2007) a importância do título reside em propiciar, juntamente com as palavras-chave, a devida indexação e, assim, uma forma fácil e correta de recuperar a informação desejada. Um artigo cujo título não se apresenta de forma clara e objetiva quanto ao conteúdo corre o risco de nunca ser recuperado/identificado pelo público ao qual foi destinado, portanto o título em artigos científicos deve ser o mais preciso e específico possível.

Segundo Ferreira e Abreu (2007) o título não deve conter caracteres indesejáveis como pontuação de qualquer natureza (ponto, dois pontos, vírgula, aspas, ponto e vírgula, interrogação, exclamação), nem sinais (parêntesis. +, -, =, X, x, \* , &, /, #) e muito menos chamadas. O título também não deve conter informações de local e quantidade (numerais) e ainda expressões tais como considerações sobre, estudo de, utilização de, estimação de, estimativa de, análise de, e tantas outra que não agregam informações importantes ao entendimento do conteúdo do artigo.

Lancaster (1993) define a LN como a linguagem do discurso técnico-científico, e no contexto da recuperação da informação a expressão normalmente se refere às palavras que ocorrem em textos impressos, considerando-se como seu sinônimo a expressão "texto livre". Já a LC, também conceituada como vocabulário controlado, é definida como um conjunto limitado de termos autorizados para uso na indexação e busca de documentos.

Lopes (2002) relata que Carrow e Nugent (1997) compararam os métodos de busca utilizando a LN e LC e revelaram que os dois métodos de busca apresentaram a mesma precisão no desempenho, porém as buscas com termos da linguagem controlada produziram uma significativa e melhor recuperação. Como conclusão do estudo, os autores apontam como complementares as duas linguagens e afirmaram que o melhor desempenho da estratégia de busca seria aquele que utilizasse os dois métodos concomitantemente.

Outro estudo comparativo foi realizado por Markey, Atherton e Newton (1980) que concluíram que a busca com termos da LN pode, frequentemente, ser a melhor opção quando se deseja alto índice de retorno, todavia o uso combinado com a LC ofereceu os melhores resultados. Para Harter (1986) em Lopes (2002) a LC é “rígida”, inflexível, mas precisa e a LN é altamente expressiva, flexível, mas potencialmente ambígua. Para Kremer (1985) o uso da LN pode ser complicado, pois exige muitas vezes a identificação e inclusão de sinônimos dos termos utilizados enquanto o uso de vocabulários controlados complica e atrasa o processo de indexação.

“A LC e a LN não podem mais ser tratadas como técnicas de busca separadas, mas devem sempre ser tratadas em conjunto, como uma combinação ideal para ampliar os resultados das buscas de informação.”  
(Muddamalle, 1998, p. 887)

Kaback (1992) testando as hipóteses sobre o uso das linguagens em bases de patentes concluiu que devem ser usadas ambas as linguagens na estratégia de busca, incluindo-se, ainda, a Classificação Internacional de Patentes. Com isso o “analista da busca” deve possuir grande habilidade em traduzir a necessidade de informação do usuário tanto para a linguagem de busca do sistema, quanto para as características de cada base. (LOPES, 2002, p. 43)

Lopes (2002) aponta algumas problemáticas no uso das LC e LN, por exemplo a indexação, isto porque os termos preferidos pelos indexadores frequentemente não são os termos utilizados pelos usuários em situações específicas de busca. Essas discrepâncias entre os termos assinalados pelos indexadores e os termos utilizados pelos usuários no momento de busca não podem ser considerados genericamente erros, porque, na realidade, o processo de

indexação ocorre em outro contexto, o de análise de conteúdo do documento e a tradução desse conteúdo para o vocabulário controlado da base.

Segundo Lopes (2002) uma base de dados que utilize um vocabulário controlado possibilita ao intermediário no planejamento da estratégia de busca, a recuperação, no campo específico de descritor, apenas daquelas palavras-chave listadas no *thesaurus* e/ou vocabulário controlado da base de dados.

No quadro 1, a seguir, estão listadas as principais vantagens e desvantagens da linguagem controlada e da linguagem natural.

**Quadro 1: Linguagem controlada e natural: vantagens e desvantagens**

LINGUAGEM CONTROLADA		LINGUAGEM NATURAL	
Vantagens	Desvantagens	Vantagens	Desvantagens
Controle total do vocabulário de indexação, minimizando os problemas de comunicação entre indexadores e usuário	Custos: a produção e manutenção da base de dados terá despesas maiores com a equipe de indexadores. Será necessário ainda manter pessoal especializado na atualização do <i>thesaurus</i> .	Permite o imediato registro da informação em uma base de dados, sem necessidade de consulta a uma linguagem de controle.	Os usuários da informação, no processo de busca, precisam fazer um esforço intelectual maior para identificar os sinônimos, as grafias alternativas, os homônimos etc.
Com o uso de um <i>thesaurus</i> e suas respectivas notas de escopo, os indexadores podem assinalar mais corretamente os conceitos dos documentos.	O vocabulário controlado poderá não refletir adequadamente os objetivos do produtor da base, caso esteja desatualizado.	Processo de busca é facilitado com a ausência de treinamentos específicos no uso de uma linguagem de controle.	Haverá alta incidência de respostas negativas ou de relações incorretas entre os termos usados na busca (por ausência de padronização).
Se bem constituído, o vocabulário controlado poderá oferecer alta recuperação e relevância e, também, ampliar a confiança do usuário diante de um possível resultado negativo.	Um vocabulário controlado poderá se distanciar dos conceitos adequados para a representação das necessidades de informação dos usuários.	Termos de entrada de dados são extraídos diretamente dos documentos que vão constituir a base de dados.	Custos de acesso tendem a aumentar com a entrada de termos de busca aleatórios.
As relações hierárquicas e as remissivas do vocabulário controlado auxiliam tanto o indexador, quanto o usuário na identificação de conceitos relacionados.	Necessidade de treinamento no uso dos vocabulários controlados tanto para os intermediários, quanto para os usuários finais.	Temas específicos citados nos documentos podem ser encontrados.	Uma estratégia de busca que arrole todos os principais conceitos e seus sinônimos deve ser elaborada para cada base de dados (ex: nomes comerciais de substâncias químicas não ocorrem no <i>Chemical Abstracts</i> ).
Redução no tempo de consulta à base, pois	Desatualização do vocabulário controlado	Elimina os conflitos de comunicação entre os	Perda de confiança do usuário em uma



a estratégia de busca será mais bem elaborada com o uso do <i>thesaurus</i> .	poderá conduzir a falsos resultados.	indexadores e os usuários, pois ambos terão acesso aos mesmos termos.	possível resposta negativa.
---	--------------------------------------	---	-----------------------------

Fonte: Adaptado de Lopes (2002).

Segundo Lopes (2002), para busca em LN duas técnicas especialmente úteis são a truncagem de termos e a busca com operadores de proximidade, conhecidos como operadores booleanos.

A truncagem de termos permite ao intermediário que operacionaliza a busca usar a raiz do termo sem especificar todas as possíveis variações desse termo (prefixos e/ou sufixos). Já a técnica de busca com operadores de proximidade ou de adjacência permite especificar, na estratégia, a posição relativa de dois ou mais termos entre eles próprios (LOPES, 2002, p. 50)

Ainda segundo a autora, nas buscas relativas ao levantamento das últimas tecnologias em uma determinada área, ou a um novo assunto, ou a um novo produto e, ainda, na busca em documentos de patentes, o uso da estratégia em LN pode ser o melhor caminho para o encontro da informação desejada. A terminologia pode ser muito atual, as aplicações ainda não são tão significativas para serem indexadas, portanto os novos termos não foram incorporados a nenhuma lista autorizada de LC.

Lopes (2002) com base em Harter (1986), Rowley (1994), Lancaster (1993) e outros estudiosos do assunto sugerem que a prática deve nortear a decisão da escolha dos termos a serem utilizados. Muitos usuários unificam as linguagens na formulação da estratégia de busca, já que a maioria das bases possui campos de busca simultânea combinados entre si. Nesta forma de busca, os campos de resumo, de títulos, de identificadores, de descritores ou cabeçalhos de assunto e de códigos de classificação podem ser amplamente utilizados para a obtenção de um resultado mais satisfatório, independentemente da verificação, no momento de operação da busca, de qual dessas linguagens terá melhor desempenho. O foco, portanto, está na obtenção de resultados satisfatórios, e não no instrumento utilizado para alcançar esses resultados.

### 3.3.A MINERAÇÃO DE TEXTO PARA RECUPERAÇÃO DA INFORMAÇÃO

Apesar dos avanços ocorridos no processamento da LN por computador, o “entendimento” de textos pelo computador ainda é muito limitada. As técnicas de

recuperação existentes tentam combinar palavras de consultas com as palavras dos documentos o que Deerwester *et al.* (1990) consideram um problema. O problema é que os usuários querem recuperar a informação com base no conteúdo conceitual, e palavras individuais fornecem evidências duvidosas sobre o assunto ou significado conceitual de um documento, isto se evidencia em palavras polissêmicas, ou seja, uma mesma palavra apresenta mais de um significado nos múltiplos contextos em que aparece, por exemplo, banco (instituição comercial financeira, assento); manga (parte da roupa, fruta) e cabo (posto militar, acidente geográfico, cabo da vassoura, da faca). Desta forma o resultado pode apresentar documentos que não são de interesse para o usuário.

Para Lancaster (1993), é possível construir instrumentos auxiliares morfológicos, sintáticos e semânticos que ajudem o computador a interpretar textos. A mineração de textos originalmente conhecida como *text mining* apresenta-se, neste contexto, como uma ferramenta de ponta que se baseia em análise quantitativa de ocorrência de palavras em um texto.

As ferramentas de *text mining* podem ser definidas como a aplicação de técnicas de tratamento automático de linguagem natural, de classificação automática e de representação gráfica do conteúdo cognitivo e factual dos dados bibliográficos. (SANTOS *et al.*, 2004, p. 209)

Para Feldman e Hirsh (1997), a mineração de textos é um meio de recuperação, filtragem, manipulação e resumo do conhecimento contido em grandes volumes de informações textuais, para apresentá-lo em forma de gráficos, listas ou tabelas. Corroborando com a afirmação, Araújo Junior e Tarapanoff (2006) apontam que a grande potencialidade da ferramenta está na capacidade de sumarizar grandes conjuntos de documentos em agrupamentos, apresentando-os sob a forma de listas de palavras que mais ocorrem por documento, ou por resultado de pesquisas (conjuntos de documentos), e em alguns casos com gráficos indicativos das relações semânticas entre os termos.

Araújo Junior e Tarapanoff (2006) apresentaram um estudo comparativo entre indexação manual e ferramenta de mineração de textos, por meio da análise do índice de precisão de resposta no processo de busca e recuperação da informação. Mais precisamente o estudo consistia em avaliar se a mineração de textos trazia ganho no índice de precisão em relação à lista de palavras-chave utilizadas na

indexação manual no processo de busca e recuperação da informação, com a finalidade de contribuir efetivamente no âmbito da ciência da informação para o estabelecimento de estratégias de uso da mineração de textos na melhoria contínua da resposta nestes sistemas em relação à questão da precisão amplamente discutida desde a década de 70.

Com base nos resultados os autores concluíram que:

- O uso da ferramenta de mineração de textos na busca e recuperação da informação traz como resposta uma maior quantidade de itens bibliográficos do que a lista de palavras-chave utilizadas na indexação manual.
- O uso da mineração de textos pode, em associação com o processo de indexação manual, trazer ganhos ao índice de precisão no processo de busca e recuperação da informação.
- O emprego da mineração de textos pode ser visto como instrumento de enriquecimento da lista de palavras-chave e/ou construção de um vocabulário controlado, utilizando, para tanto, a lista de palavras mais frequentes em cada documento recuperada em pesquisas realizadas no protótipo, além da lista de palavras mais frequentes do resultado total da pesquisa, também realizada no protótipo. Cabe ressaltar que a grande potencialidade da ferramenta é a captura de qualquer termo em qualquer parte do texto completo. Este fato se configura como um instrumento útil no aprimoramento contínuo do processo de indexação.

Concluem que a mineração de textos apoia o processo de indexação manual para aumento do índice de precisão de resposta no processo de busca e recuperação da informação, mas apesar de toda a discussão, os resultados só poderão ser validados pela avaliação dos usuários. “Os sistemas de recuperação da informação, além de buscar atender às demandas informacionais dos usuários, dependem destes para que a qualidade dos seus serviços seja reconhecida.” (OLIVEIRA JUNIOR; TARAPANOFF, 2006, p. 247).

## 4. METRIAS DE INFORMAÇÃO

Nas últimas décadas, com a evolução da ciência e tecnologia, tornou-se cada vez mais necessário avaliar estes avanços e determinar os desenvolvimentos alcançados pelas diversas disciplinas do conhecimento.

A palavra avaliar, de acordo com o Dicionário Michaelis online, significa calcular ou determinar o valor, o preço ou o merecimento. Segundo Vanti (2002) a avaliação, dentro de um determinado ramo do conhecimento, permite tornar digno o saber quando métodos confiáveis e sistemáticos são utilizados para mostrar à sociedade como tal saber vem sendo desenvolvido e de que forma tem contribuído para resolver os problemas que se apresentam dentro de sua área de abrangência.

Para Guimarães (1992) o conceito de avaliação vem sempre associado com a promessa de se obter informações que auxiliem a tomada de decisão. Para a autora, a práxis da avaliação é normalmente justificada pela necessidade do conhecimento na visão instrumental, e para isto são utilizadas uma série de técnicas, em diferentes abordagens, como forma de satisfazer esta necessidade.

Existem diversas formas de medição voltadas para avaliar a ciência e os fluxos da informação. Dentre estas, cabe citar: 1) a bibliometria; 2) a cienciometria; 3) a informetria e 4) a webometria. Segundo Vanti (2002), todas têm funções semelhantes, mas, ao mesmo tempo, cada uma delas propõe medir a difusão do conhecimento científico e o fluxo da informação sob enfoques diversos. O objetivo de estudo da primeira são livros, documentos, revistas, artigos, autores e usuários; da segunda, disciplinas, assuntos, áreas e campos; e da terceira, palavras, documentos e bases de dados (MACIAS-CHAPULA, 1998, p. 135) e a quarta são sítios na *word wide web* (VANTI, 2002, p. 160).

### 4.1. BIBLIOMETRIA

“A bibliometria [...] surge no início do século como sintoma da necessidade do estudo e da avaliação das atividades de produção e comunicação científica”. (ARAÚJO, 2006, p. 12). O termo foi utilizado inicialmente em 1934 por Paul Otlet em sua obra intitulada “*Traité de documentation*”, embora alguns artigos de revisão clássicos atribuam a Alan Pritchard, que na verdade popularizou o uso da palavra ‘bibliometria’, quando sugeriu que esta deveria substituir o termo ‘bibliografia

estatística', que vinha sendo utilizado desde a menção feita em 1922 por Edward Wyndham Hulme em uma conferência na Universidade de Cambridge (Fonseca, 1993).

De acordo com as palavras de Tague-Sutcliffe, traduzidas por Macias-Chapula pode-se definir a bibliometria como:

“[...] o estudo dos aspectos quantitativos da produção, disseminação e uso da informação registrada. A bibliometria desenvolve padrões e modelos matemáticos para medir esses processos, usando seus resultados para elaborar previsões e apoiar tomadas de decisões” (MACIAS-CHAPULA, 1998, p. 134).

Na mesma linha de pensamento, Araújo (2006) define a bibliometria como uma técnica quantitativa e estatística de medição dos índices de produção e disseminação do conhecimento científico.

Wormell (1998) afirma que os profissionais da informação podem ampliar consideravelmente os horizontes do seu trabalho se utilizar as metodologias clássicas da análise bibliométrica junto com os modernos mecanismos de busca *on-line*. Podem obter acesso não apenas a documentos completos ou a partes de documentos, mas também traçar as tendências e os desenvolvimentos na região. Por meio da análise quantitativa, podem mapear novos caminhos em áreas específicas da pesquisa, produção e consumo, ou identificar pequenas unidades de informação de grande valor em meio aos enormes volumes de informação eletrônica.

Para Araújo (2006) entre os principais marcos do desenvolvimento da bibliometria estão o método de medição da produtividade de cientistas de Lotka (1926), a lei de dispersão do conhecimento científico de Bradford de 1934 e o modelo de distribuição e frequência de palavras em um texto de Zipf de 1949.

Dentre estas, a que nos interessa é a terceira lei bibliométrica clássica, a Lei de Zipf, formulada em 1949 e que de acordo com Araújo (2006) descreve a relação entre palavras em um determinado texto relativamente grande e a ordem de série destas palavras (contagem de palavras em largas amostragens). Ainda segundo o autor, Zipf analisou a obra *Ulisses* de James Joyce e encontrou uma relação entre o número de palavras diferentes e a frequência de seu uso e concluiu que existe uma regularidade na seleção e uso das palavras e que um pequeno número de palavras é usado muito mais frequentemente. Zipf propôs que ao listar palavras que ocorrem

em um texto decrescente de frequência, a posição de uma palavra na lista multiplicada por sua frequência é igual a uma constante. Desta forma a equação que representa a Lei de Zipf é:

**1ª Lei de Zipf**

**$R \times F = K$**

**Legenda**  
R: ordem das palavras  
F: frequência das palavras  
x: constante

**Figura 3 - 1ª Lei de Zipf. Elaboração própria com base em Tarapanoff (2002)**

Zipf formulou o princípio do menor esforço (ARAÚJO, 2006, p. 16), esta segunda lei de Zipf é mais conhecida como lei de Zipf-Booth (TARAPANOFF, 2002, p. 93), estabelece a ocorrência de palavras de baixa frequência em um texto, onde as palavras menos frequentes ocorrem de modo que várias palavras acabam por apresentar a mesma frequência. A lei de Zipf-Booth apresenta a seguinte fórmula:

**2ª Lei de Zipf**  
**(Lei Zipf-Booth)**

**$\frac{l_n}{l_1} = \frac{2}{n \times (n - 1)}$**

**Legenda**  
 $l_n$ : ordem das palavras  
 $l_1$ : frequência das palavras  
2: constante

**Figura 4 - 2ª Lei de Zipf: Lei Zipf-Booth. Elaboração própria com base em Tarapanoff (2002)**

Araújo (2006) afirma que como a lei de Lotka e Bradford, a lei de Zipf foi reformulada, por exemplo, por Kendall e Brookes. O método foi sendo aperfeiçoado principalmente com estudos de frequência e co-ocorrência de descritores.

A lei de Zipf, particularmente, tem sido utilizada por muitos autores da linha francesa de monitoramento/prospecção, que afirmam que a lógica da busca para monitoramento tecnológico é de assumir uma busca “imperfeita”, pois se deve trabalhar com o risco de se recuperar documentos na frequência baixa que tanto

pode concentrar documentos relacionados à inovação como aqueles que são considerados ruídos estatísticos, por não apresentarem qualquer relação com o que se procura (QUONIAM, 1995; ROSTAING, 1996 *apud* SILVA, 2002).

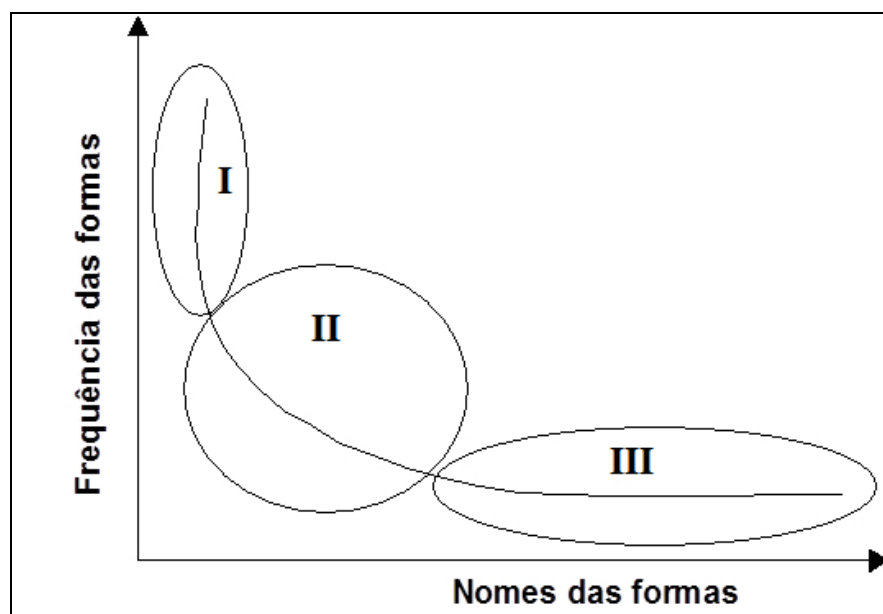


Figura 5: Distribuição de Vocabulário Controlado. Quoniam (1995)

Enquanto a maior parte dos autores utiliza o corte da distribuição bibliométrica em duas zonas: núcleo e dispersão, Quoniam (1995) propõe que o mesmo seja feito em três zonas, quando os elementos bibliográficos analisados forem oriundos de campos bibliográficos que tenham um vocabulário controlado. (Figura 7)

Zona I: Informação trivial – zona de alta frequência é a que define os assuntos centrais da área investigada;

Zona II: Informação interessante – que mostra ou os assuntos periféricos ou a informação potencialmente inovadora;

Zona III: Ruído – zona de baixa frequência – pode conter o ruído estatístico ou os conceitos inovadores ainda não emergentes.

Braga (1973) aponta que as pesquisas no campo da bibliometria investigam o comportamento do conhecimento e da literatura como parte dos processos de comunicação; entre elas estão os estudos relacionados a citações, desenvolvidos principalmente por Solla Price, Kessler, Martyn e Garfield. A área mais importante da bibliometria é a análise de citações (ARAUJO, 2006, p. 18).

A autora ainda afirma que o hábito de fazer referência a outros trabalhos nasceu junto com os periódicos científicos, no século 17, e tornou-se uma segunda natureza do cientista que tinha por costume trocar correspondência para fins científicos, antes da chamada "explosão da informação", originada pelos periódicos.

Kaplan (1965), citado por Braga (1973) afirma que a citação é uma forma de conciliar o desejo que têm os autores de comunicar suas pesquisas e descobertas, como contribuição à comunidade científica e ao mesmo tempo de proteger seus direitos de propriedade. Já Solla Price (1973), citado por Foresti (1990) afirma que as citações contribuem para o desenvolvimento da ciência, proveem o necessário reconhecimento de um cientista por seus colegas, estabelecem os direitos de propriedade e prioridade da contribuição científica de um autor, constituem importantes fontes de informação, ajudam a julgar os hábitos de uso da informação e mostram a literatura que é indispensável para o trabalho dos cientistas.

Para Araújo (2006), a análise de citações permite a identificação e descrição de uma série de padrões na produção do conhecimento científico.

Apesar de serem muitas vezes incompletas e inadequadamente empregadas, as citações são importantes e possuem padrões de comportamento que obedecem a determinadas leis. O primeiro índice de citações, o *Science Citation Index* (SCI), considerado um marco na época, foi criado por Eugene Garfield, fundador do *Institute of Scientific Information* (ISI) (ARAÚJO, 2006). "O desenvolvimento dos índices de citações (*citation indexes*) [...] proporcionou diferentes pontos de acesso à informação registrada". (BRAGA, 1973, p. 10).

Um conceito extremamente relevante na análise de citações é o fator de impacto. O fator de impacto é a divisão do número de citações recebidas por um autor dividido pelo número de trabalhos de receberam pelo menos uma citação. Segundo Araújo (2006) o índice é utilizado para identificar autores de pouca produção, porém significativa, isto é, que receberam muitas citações, em oposição a autores que podem ter tido muitas citações porque publicaram muitos trabalhos, mas cada um desses trabalhos isoladamente com pouca relevância no campo científico.

Outra abordagem além da contagem de palavras e frequência de publicações e de citações é a contribuição de Kessler que também se preocupou com a relação entre artigos que geram e recebem citações porém em uma escala multidimensional. Araújo (2006) afirma que Kessler desenvolveu as teorias do acoplamento



bibliográfico, que dizem respeito à força de associação entre dois ou mais documentos.

Completando o último conjunto de leis bibliométricas estão a obsolescência e vida média da literatura. Foram desenvolvidas para descrever a queda da validade ou utilidade de informações com o decorrer do tempo.

Segundo Strehl (2005), o conceito de vida média foi utilizado na área de ciência da informação pela primeira vez por Burton e Kebler em 1960, para expressar o período em que uma pesquisa alcança a metade de sua vida útil. A vida média é concebida como um indicador da influência dos periódicos que mede a obsolescência.

Sancho (1990) define a obsolescência como a diminuição da utilização da informação no decorrer do tempo, podendo ocorrer por alguma das seguintes causas: a) a informação é válida, mas foi substituída por outra mais moderna; b) a informação é válida, mas em um campo científico de interesse decrescente; c) a informação não é mais considerada válida.

Com relação à vida média de uma tecnologia patenteada Chen *et al.* (2009) apresentaram um estudo com o objetivo de desenvolver uma nova forma de medir a performance bibliométrica da patente e a sua vida média e concluíram que a vida média de uma patente é de cinco anos.

Segundo Araújo (2006), ao longo da década de 80 houve uma queda no interesse da bibliometria pela área de biblioteconomia, tanto no Brasil quanto no exterior. Esta queda é resultado das possibilidades do uso do computador para exploração de metodologias quantitativas.

#### 4.2. CIENCIOMETRIA

O termo cienciometria ou cientometria segundo Vanti (2002) surgiu na antiga União das Repúblicas Socialistas Soviéticas (URSS) e Europa Oriental e foi empregado especialmente na Hungria. Segundo Mikhilov *et al. apud* Spinak (1996), as primeiras definições consideravam a cienciometria como “a medição do processo informático”, onde o termo “informático” significava a disciplina do conhecimento que estuda a estrutura e as propriedades da informação científica e as leis do processo de comunicação.

De acordo com as palavras de Tague-Sutckiffe, traduzidas por Macias-Chapula, pode-se definir a cienciometria como:

“Cienciometria é o estudo dos aspectos quantitativos da ciência enquanto uma disciplina ou atividade econômica. A cienciometria é um segmento da sociologia da ciência, sendo aplicada no desenvolvimento de políticas científicas. Envolve estudos quantitativos das atividades científicas, incluindo a publicação e, portanto, sobrepondo-se à bibliometria”. (MACIAS-CHAPULA, 1998, p. 134)

Ainda segundo o autor, a cienciometria estuda, por meio de indicadores quantitativos uma determinada disciplina da ciência. Estes indicadores quantitativos são utilizados dentro de uma área do conhecimento, por exemplo, mediante a análise de publicações, com aplicação no desenvolvimento de políticas científicas. Tenta medir os incrementos de produção e produtividade de uma disciplina, de um grupo de pesquisadores de uma área, a fim de delinear o crescimento de determinado ramo do conhecimento.

Para Van Raan (1997), a cienciometria se dedica a realizar estudos quantitativos em ciência e tecnologia e a descobrir os laços existentes entre ambas, visando ao avanço do conhecimento e buscando relacionar este com questões sociais e de políticas públicas. A cienciometria teria, portanto, um caráter multidisciplinar no que diz respeito aos métodos que utiliza.

Para Callon *et al.* citado por Vanti (2002), a cienciometria se aplica, principalmente, ao tratamento e gerenciamento das informações formais provenientes de bases de dados científicas ou técnicas. Inclusive, segundo estes autores, as empresas utilizariam os métodos cienciométricos para conseguir determinar qual a estratégia tecnológica seguida por seus competidores. As empresas, desta forma, podem analisar as patentes, por exemplo, identificando os principais temas de investigação e atores significativos em um campo científico de seu interesse. Podem, igualmente, detectar as especialidades científicas que servem de base para suas tecnologias-chave.

#### 4.3. INFORMETRIA

O termo informetria, de acordo com Brookes (1990), foi proposto pela primeira vez por Otto Nacke, diretor do *Institut für Informetrie*, em Bielferd, Alemanha, 1979

em seus estudos tendo como objetivo a análise de todos os aspectos relacionados à armazenagem e recuperação de informações.

De acordo com as palavras de Tague-Sutcliffe, traduzidas por Macias-Chapula, pode-se definir a informetria como:

“Informetria é o estudo dos aspectos quantitativos da informação em qualquer formato, e não apenas registros catalográficos ou bibliografias, referente a qualquer grupo social, e não apenas aos cientistas. A informetria pode incorporar, utilizar e ampliar os muitos estudos de avaliação da informação que estão fora dos limites da bibliometria e cienciometria”. (MACIAS-CHAPULA, 1998, p. 134)

De acordo com Wormell (1999), a informetria é um subcampo emergente da ciência da informação, baseada na combinação de técnicas avançadas de recuperação da informação com estudos quantitativos dos fluxos da informação”.

Para Tague-Sutcliffe citado por Vanti (2002), a informetria se distinguiria claramente da cienciometria e da bibliometria no que diz respeito ao universo de objetos e sujeitos que estuda, não se limitando apenas à informação registrada, dado que pode analisar também os processos de comunicação informal, inclusive falada, e dedicar-se a pesquisar os usos e necessidades de informação dos grupos sociais desfavorecidos, e não só das elites intelectuais.

#### 4.4. WEBOMETRIA

Segundo Almind e Ingwersen (1997), a webmetria ou webometrics consiste na aplicação de métodos informétricos à World Wide Web (www). Para Cronin e McKim (1996), a web é um importante meio de comunicação para a ciência e a academia, pelo qual é lógico que os estudos quantitativos se estendam também a este ambiente.

Segundo Vanti (2002), no campo da webometria podem ser realizadas medições como a frequência de distribuição das páginas no *cyberespaço*. Pode-se realizar, também, uma mesma medição em tempos diferentes para comparar, assim, a evolução da presença de uma determinada instituição ou país na rede. Da mesma forma, é possível quantificar o crescimento ou perda de importância relativa de um tema ou matéria, o que aproxima, neste caso, a webometria à cienciometria, segundo as definições e aplicações já citadas.

Segundo Smith (1999) os instrumentos fundamentais para a realização de estudos webométricos têm sido os motores de busca, que permitem trabalhar com grandes volumes de informação.

**Quadro 2: Comparação das aplicações dos distintos métodos quantitativos**

Tipologia/ Subcampo	Bibliometria	Cienciometria	Informetria	Webometria
<b>Objeto de estudo</b>	Livros, documentos, revistas, artigos, autores, usuários.	Disciplinas, assuntos, áreas e campos científicos e tecnológicos. Patentes, dissertações e teses.	Palavras, documentos, bases de dados, comunicações informais (inclusive em âmbitos não científicos), <i>home pages</i> na <i>web</i> .	Sítios na <i>www</i> ( <i>URL</i> , título, tipo, domínio, tamanho e <i>links</i> ), motores de busca.
<b>Variáveis</b>	Número de empréstimos (circulação) e de citações, frequência de extensão de frases.	Fatores que diferenciam as subdisciplinas. Como os cientistas se comunicam.	Difere da cienciometria no propósito das variáveis, por exemplo, medir a recuperação, a relevância, a revocação.	Número de páginas por sítio, nº de <i>links</i> por sítio, nº de <i>links</i> que remetem a um mesmo sítio, nº de sítios recuperados.
<b>Métodos</b>	Ranking, frequência, distribuição.	Análise de conjunto e de correspondência, concorrência de termos, expressões, palavras-chave, etc.	Modelo vetor-espço, modelos booleanos de recuperação, modelos probabilísticos; linguagem de processamento, abordagens baseadas no conhecimento, tesouros.	Fator de Impacto da Web (FIW), densidade dos links, "situações", estratégias de busca.
<b>Objetivos</b>	Alocar recursos: pessoas, tempo, dinheiro, etc.	Identificar domínios de interesse. Onde os assuntos estão concentrados. Compreender como e quanto os cientistas se comunicam.	Melhorar a eficiência da recuperação da informação, identificar estruturas e relações dentro dos diversos sistemas de informação.	Avaliar o sucesso de determinados sítios, detectar a presença de países, instituições e pesquisadores na rede e melhorar a eficiência dos motores de busca na recuperação das informações.

Fonte: Vanti (2002) adaptado de McGrath citado por Macias-Chapula, (1998).

Tentou-se neste capítulo apresentar as diferenças conceituais entre as métricas de informação, entretanto, o que se vê na prática é que os termos são utilizados indistintamente, à exceção da webometria que é sempre mencionada individualmente.

Vale a pena lembrar que as metrias de informação são a base da fundamentação teórica da metodologia a ser utilizada para a busca de informação.

## 5. PATENTES

Até o primeiro quartel do século XVII, reis e governantes concediam a seus pares exclusividade para exercer um determinado comércio. Tais monopólios comerciais visavam tão-somente conceder favores ao invés de recompensar quaisquer possíveis esforços dispendidos pelos nobres que trouxessem um benefício social.

Macedo e Barbosa (2000) consideram que a ideia de incentivar as invenções mediante a concessão do monopólio, a patente, surgiu na República de Veneza, em 1477. Segundo França (2007) esta primeira lei sobre patentes beneficiou os fabricantes independentes de vidro de Murano através de uma carta-aberta, *litterae-patente*, um compromisso entre o Estado e um cidadão. Neste compromisso, o governo manteria o monopólio de manufatura em seu território para o cidadão e, em contrapartida, os procedimentos de fabricação seriam divulgados, tornando-se públicos.

Macedo e Barbosa (2000) afirmam que, após um século e meio, a concessão do monopólio ressurgiu pelo Estatuto dos Monopólios, promulgado pela Coroa Britânica em 1623; que, para França (2007) é a base do sistema contemporâneo de patentes. Ainda segundo as autoras, este estatuto gerou vários conflitos sobre os direitos de novidade, já que considerava qualquer produto ou processo de manufatura ainda não conhecido no Reino Unido, independentemente de já ser conhecido ou usado em outros países. Para Macedo e Barbosa (2000), foi a partir do Estatuto dos Monopólios que se difundiu a concessão do monopólio pela Europa, chegando à América no fim do século XVIII.

No transcorrer do século XIX, inúmeros países tinham suas leis nacionais de patentes. O Brasil iniciou o processo de concessão patentária às invenções a partir de 1830. Até fins do século XIX, as leis nacionais somente conferiam proteção aos inventores do próprio país, inexistindo a possibilidade de proteção de inventores estrangeiros (MACEDO E BARBOSA, 2000, P. 17).

A patente pode ser conceituada, inicialmente, tendo por base os princípios do “Contrato Social” de Rousseau, como um acordo entre o inventor e a sociedade.

O Estado concede o monopólio da invenção, isto é, a sua propriedade inerentemente caracterizada pelo uso exclusivo de um novo processo produtivo ou a fabricação de um produto novo vigente por um determinado

prazo temporal e, em troca, o inventor divulga a sua invenção, permitindo à sociedade o livre acesso ao conhecimento desta matéria objeto da patente (MACEDO E BARBOSA, 2000, p. 18)

Em virtude da necessidade de ampliar a proteção além das fronteiras nacionais, de forma a induzir o crescimento e consolidação do comércio internacional e a evitar que os produtos pudessem ser copiados em outros países que não o de origem da invenção, foi realizado, em 1873, o Congresso de Viena. A partir deste Congresso, iniciou-se o entendimento sobre a abrangência internacional da proteção legal. Continuando em Paris, em 1878 e 1880, o movimento finalizou pela Convenção da União de Paris para a Proteção da Propriedade Industrial ocorrida em 1883.

A Convenção da União de Paris (CUP), também conhecida como Convenção de Paris, foi a primeira tentativa de uma harmonização internacional dos diferentes sistemas jurídicos nacionais relativos à propriedade industrial (INPI, 2011). Ainda segundo o Instituto Nacional da Propriedade Industrial (INPI) (2011), a CUP foi elaborada de modo a permitir razoável grau de flexibilidade às legislações nacionais, desde que fossem respeitados alguns princípios fundamentais. Tais princípios são de observância obrigatória pelos países signatários. Cabe ressaltar que o Brasil foi um dos quatorze países signatários originais e que, segundo o *World Intellectual Property* (WIPO) (2011), atualmente 173 países fazem parte desta Convenção.

Sobre os princípios que regem a CUP há divergência entre autores, uns defendem a existência de três princípios, outros defendem a existência de quatro princípios. O INPI (2011) defende a existência de quatro princípios, são eles: Tratamento nacional; Prioridade unionista; Interdependência dos direitos e Territorialidade. Porém, a França (2007) defende a existência de apenas três princípios: Tratamento nacional; Prioridade unionista e Interdependência legal (dos direitos). O princípio discordante – territorialidade - de acordo com o INPI é relacionado ao princípio da interdependência dos direitos. França (2007) associa o princípio da Territorialidade ao da Interdependência legal (dos direitos), formando apenas três princípios.

França (2007) contextualiza os três princípios supracitados como: (i) O primeiro, que se refere ao tratamento nacional, aborda o tratamento igualitário aos inventores nacionais e estrangeiros, pelos países membros, na tramitação dos depósitos de patente, desde que os estrangeiros sejam de outros países-membros

da União; (ii) O segundo, que diz respeito à prioridade unionista, afirma que fica garantido ao depositante de uma patente original, em qualquer um dos países signatários, o direito de salvaguarda da prioridade de depósito em todos os países signatários pelo prazo de doze meses. E o terceiro princípio da interdependência legal ressalta a territorialidade, ou seja, a patente só tem validade no país que concedeu o direito de propriedade.

Macedo e Barbosa (2000) e Chaves *et al.* (2007) informam que, com a assinatura da Convenção de Paris, foi criado o Sistema Internacional de Patentes; com o objetivo de incentivar globalmente a inovação, proporcionando a proteção patentária em um espaço econômico maior por inserção de mercados externos.

Com o passar dos anos, o Sistema Internacional de Patentes foi evoluindo, foram criados novos tratados internacionais, regionais e bilaterais versando sobre as mais diversas matérias e assuntos específicos. A seguir são apresentados os principais tratados internacionais relativos à propriedade industrial.

**Quadro 3: Principais Tratados Internacionais relativos à Propriedade Industrial**

Tratados	Origem
Convenção de Paris	Tratado multilateral básico, firmado em 1883 na cidade de Paris; tendo, por objeto, os institutos de propriedade industrial, patentes em geral, marcas em geral, indicações de procedência e a proteção à concorrência desleal.
Tratado de Cooperação em Matéria de Patentes (PCT – <i>Patent Cooperation Treat</i> )	Tratado multilateral, firmado em 1970 na cidade de Washington, cujo principal objetivo é facilitar e reduzir os custos iniciais nos procedimentos de pedidos de patentes nos países membros. Este tratado prevê: (I) Único depósito; (II) Único idioma; (II) Relatório de Busca Preliminar; (IV) Tempo para avaliar a pertinência de entrada nas Fases nacionais e, em caso positivo, em quais destas Fases depositar.
Acordo de Estrasburgo - Classificação Internacional de Patentes (CIP)	Firmado em 1971 na cidade de Estrasburgo. Este acordo estabelece para os países membros um sistema de classificação das patentes por ramo da técnica, sendo amplamente adotado por todos os países desenvolvidos e pela maioria dos países em desenvolvimento.
Tratado de Budapeste sobre Depósitos de Microrganismos para a Finalidade de Proteção por Patente	Firmado em 1977, em Budapeste, estabelece - para os países membros - procedimentos e exigências para o depósito e guarda de microrganismos com fins de proteção patentária. Também estabelece normas para o fornecimento de amostras de microrganismos armazenados.
Acordo sobre Aspectos Comerciais de Direitos de Propriedade Intelectual, incluindo a Contrafação de Bens (TRIPS – <i>Trade Related Aspects on Intellectual Property Rights, including</i>	Firmado em 1994 na cidade de Marrakesh, como parte de um tratado maior que criou a Organização Mundial do Comércio (OMC) (antigo GATT). O Acordo TRIPS constitui-se fundamentalmente de parâmetros mínimos de proteção; embora presente, a regra de tratamento nacional é subsidiária em face do patamar uniforme de proteção. Seu objetivo é fazer com que a proteção e a aplicação de normas



<i>Counterfeiting of Goods</i>	<p>de proteção dos direitos de propriedade contribuam para a promoção da inovação tecnológica e para a transferência e difusão de tecnologia, em benefício mútuo de produtores e usuários de conhecimento tecnológico e de uma forma conducente ao bem-estar social e econômico e a um equilíbrio entre direitos e obrigações.</p> <p>O acordo estabelece padrões mínimos no âmbito do direito internacional relacionados às patentes, incluindo aquelas de medicamentos. Países membros da OMC – que hoje já são mais de 150 – concordaram com certos padrões comuns na forma de elaboração e implementação de legislações patentárias. Estes padrões incluem, entre outros, que as patentes devem ser concedidas durante um período mínimo de vinte anos, que as patentes podem ser concedidas para produtos e processos e que informações de testes de medicamentos podem ser protegidas contra o “uso comercial desleal”.</p>
--------------------------------	---

Fonte: Adaptado de Macedo e Barbosa (2000).

“Cada país tem sua legislação relativa à Propriedade Industrial, a qual é, evidentemente, aplicável somente aos atos ocorridos ou cometidos dentro dos seus limites territoriais” (MACEDO; MÜLLER; MOREIRA, 2001, p. 18). No Brasil, a proteção ao direito de propriedade industrial é regulamentada pela Lei nº 9279/96, lei da propriedade industrial, de 14 de maio de 1996. De acordo com França (2007), esta Lei, que regula direitos e obrigações relativos à propriedade industrial, foi promulgada para atender as exigências do acordo TRIPS.

No artigo 2 desta Lei, Brasil (1996), a proteção dos direitos relativos à propriedade industrial, efetua-se mediante a concessão de patentes de invenção e de modelo de utilidade; a concessão de registro de desenho industrial; a concessão de registro de marca; a repressão às falsas indicações geográficas e a repressão à concorrência desleal.

A responsabilidade pelo registro e fiscalização da propriedade industrial fica a cargo do INPI, uma autarquia federal sediada no Estado do Rio de Janeiro, encarregada do processamento e exame dos pedidos, além das concessões de patentes e registros. O INPI define a patente como:

Um título de propriedade temporária sobre uma invenção ou modelo de utilidade, outorgados pelo Estado aos inventores ou autores ou outras pessoas físicas ou jurídicas detentoras de direitos sobre a criação. Em contrapartida, o inventor se obriga a revelar detalhadamente todo o conteúdo técnico da matéria protegida pela patente. (INPI, 2011)

No caso específico da patente, a lei determina que a invenção atenda obrigatoriamente aos requisitos de novidade, atividade inventiva e aplicação

industrial. Macedo e Barbosa (2000) definem novidade quando um conhecimento técnico, para o qual se requer a proteção por patente não esteja disponível ao público (estado da arte) sob qualquer forma de divulgação anterior até a data de depósito da patente, ou seja, a invenção deve ser algo inédito. É previsto nesta lei, em seu artigo 12, que não é considerada como estado da técnica a divulgação de invenção ou modelo de utilidade, quando ocorrida durante os 12 (doze) meses que precederem a data de depósito ou a da prioridade do pedido de patente. A atividade inventiva é entendida como não obviedade, ou seja, não pode ser uma simples substituição de materiais ou de meios conhecidos para outras finalidades sem que haja um efeito técnico novo e inesperado. A aplicação industrial significa que a invenção deve ter finalidade de uso na produção econômica, produção em escala industrial.

A lei não considera patenteáveis as descobertas ou teorias científicas e métodos matemáticos; as concepções puramente abstratas; esquemas, planos, princípios ou métodos comerciais, contábeis, financeiros, educativos, publicitários, de sorteio e de fiscalização; as obras literárias, arquitetônicas, artísticas e científicas ou qualquer criação estética; programas de computador em si; apresentação de informações; regras de jogo; técnicas e métodos operatórios ou cirúrgicos, bem como métodos terapêuticos ou de diagnóstico, para aplicação no corpo humano ou animal; o todo ou parte de seres vivos naturais e materiais biológicos encontrados na natureza, ou ainda que dela isolados, inclusive o genoma ou germoplasma de qualquer ser vivo natural e os processos biológicos naturais.

O pedido de patente é mantido em sigilo durante dezoito meses contados da data de depósito ou da prioridade mais antiga, quando houver, à exceção do caso previsto no art. 75 (objeto de interesse à defesa nacional – sigilo e proibida a publicação) conforme previsto em lei. Cabe ressaltar que o período de sigilo é utilizado como um período de aprimoramento da invenção, caso haja um aprimoramento, é realizado o depósito de um complemento denominado certificado de adição de invenção; o período permite que o depositante requeira o depósito internacional do pedido via PCT para que após o prazo de 30 meses entre nas fases nacionais dos países, desta forma ampliando o escopo de proteção, por isso sua importância do sigilo.

O tempo de vigência de uma patente, de acordo com o artigo 40, é de vinte anos; enquanto a de modelo de utilidade é de quinze anos contados da data de

depósito. Entretanto, consta em lei que o prazo de vigência não poderá ser inferior a dez anos para a patente de invenção, e a sete anos para a patente de modelo de utilidade, a contar da data de concessão, ressalvada a hipótese do INPI estar impedido de proceder ao exame de mérito do pedido, por pendência judicial comprovada ou por motivo de força maior.

De acordo com Macedo, Müller e Moreira (2001), o sistema de patentes recompensa o inventor de uma maneira engenhosa. O benefício financeiro da patente é percebido somente quando os direitos de excluir terceiros são exercidos, tanto por meio de licenciamento (em troca de *royalties*<sup>1</sup>) para outros praticarem a invenção, como pela comercialização, pelo titular da patente, do produto ou do processo patentado. Essa recompensa é suscetível de promover a criatividade e a encorajar inventores e empresas a continuar o desenvolvimento de novas tecnologias, para torná-las comercializáveis, úteis e favoráveis ao interesse público.

Neste contexto, verifica-se que a patente é considerada um grande incentivo ao desenvolvimento tecnológico por ser um documento oficial de proteção à invenção; por ser exclusivo do titular, o qual, em caso de uso da patente por terceiros, recebe *royalties* como uma espécie de estímulo à criação de novas invenções; e por ser um documento considerado como a mais rica fonte de informação tecnológica no mundo.

### 5.1. PATENTE COMO FONTE DE INFORMAÇÃO TECNOLÓGICA

Cabral (2005) afirma que trabalhar com documentos de patente, pela sua especificidade, exige algum conhecimento mínimo de propriedade industrial para entender o que é esse documento e sua terminologia específica. Para Longa (2007), a maioria das instituições de pesquisa pública e universidades não utilizam a informação em patentes como instrumento capaz de subsidiar o desenvolvimento de pesquisas/projetos, buscar parcerias, licenciamento e transferência de tecnologia. Na verdade, o que estas instituições – e quando assim procedem - executam na área da busca em bases de patente é somente a utilização da informação para aferição da patenteabilidade de suas pesquisas.

---

<sup>1</sup> *Royalties*: Valor que o titular de objeto protegido recebe pela exploração da matéria protegida.

Segundo Araújo (1981), um estudo realizado para a Comissão de Comunidades Europeias em cinco países - Dinamarca, Alemanha, França, Irlanda e Holanda - revelou que somente 5% dos usuários potenciais da informação contida nos documentos de patente têm consciência de suas possibilidades de uso.

A autora ainda afirma que há vários fatores que acarretaram o baixo uso da informação contida em patentes:

- Cerca de 85% das patentes em vigor são de propriedade de empresas (a maioria delas sendo grandes corporações), que não têm interesse na livre divulgação da informação por si; já que a divulgação do conhecimento contido no documento de patente pode gerar uma inovação incremental ou radical por outras empresas acarretando na “perda” do monopólio.
- Existe um desconhecimento quase generalizado por parte dos profissionais, inclusive os da informação, de como obter documentos de patentes e de como utilizar seu aspecto informacional, quer técnico quer econômico;
- O estilo em que a patente é escrita torna-a de difícil leitura e, mais ainda, o desconhecimento de como a patente está estruturada cerceia a rápida recuperação da informação necessária;
- A diversidade de sistemas de classificação existentes, que são de difícil aprendizagem e uso; a diversidade de sistemas nacionais de propriedade industrial, o que conduz ao surgimento de problemas práticos quando da realização de buscas e da obtenção e processamento de documentos em nível internacional.

Saber quais as bases de dados que devem ser acessadas, além dos procedimentos que devem ser empregados na realização da busca em literatura de patentes, que traga resultados satisfatórios, muitas vezes necessita de orientação e treinamento, que devem ser administrados por um técnico especializado no assunto. Com a instrução, o profissional da informação poderá identificar que caminho o leva aos resultados correlatos à sua área de atuação científica, possibilitando ampliar seu conhecimento frente aos concorrentes, que poderão se tornar possíveis parceiros no desenvolvimento da tecnologia desejada. (LONGA. 2007, p. 1)

Até o momento, a maneira pela qual estava sendo enfocada a informação contida em documentos de patente era basicamente de cunho tecnológico (MACEDO; BARBOSA, 2000, p. 60). Ainda segundo estes autores, a informação patentária é muito mais ampla.

A proteção patentária é um instituto jurídico destinado a conferir aos titulares dos direitos o monopólio do uso na produção econômica da invenção, isto é, de uma informação técnico-produtiva inexistente no estado das artes úteis até a criação. Portanto, os documentos de patente têm três tipos de informação: jurídica, econômica e técnica. (MACEDO; BARBOSA, 2000, p. 60)

Considerando os três tipos de informação contida em documentos de patentes, existe uma ampla gama de usuários que podem fazer uso desses documentos, destacam-se:

- Empresas;
- Instituições de pesquisa e desenvolvimento, inclusive universidades;
- Autoridades governamentais em geral;
- Agentes da propriedade industrial;
- Inventores individuais; e
- Universitários e estudantes de nível técnico.

Toda a economia gerencial de uma empresa poderá ser decidida, estruturada e planejada se essa ferramenta, que é a informação tecnológica contida em documentação de patentes, for previamente utilizada. (CABRAL, p. 257). Ainda segundo a autora, as informações contidas na documentação de patente ajudam o empresário não apenas na fase de decisão, mas também desde o início das pesquisas, quando já há algum valor.

Longa (2007) afirma que, no cenário empresarial, o uso da informação tecnológica torna-se mais ativo devido à implantação de sistemas de inteligência competitiva, que utilizam patentes, analisando-as através de *softwares* especializados e realizando estudos de prospecção tecnológica com indicadores de mercado frente à determinada tecnologia pesquisada.

“O documento de patente é, em tese, a mais importante fonte primária de informação tecnológica” (FRANÇA, 2007, p. 168). Estudos revelam que 70% das informações tecnológicas contidas nestes documentos não estão disponíveis em qualquer outro tipo de fonte de informação (INPI, 2011).

Macedo e Barbosa (2000) afirmam que os usuários do sistema de informação patentária encontram diversos motivos para buscar a informação desejada em patentes, dentre eles estão:

- Conhecer os efetivos direitos da proteção patentária vigentes em determinado país, particularmente para evitar violações da patente;
- Definir o estado da técnica em determinado campo da técnica, inclusive conhecendo seus últimos avanços e aperfeiçoamentos;
- Avaliação técnica das diversas invenções em um campo da técnica, por exemplo, para fins de definir futuros parceiros de pesquisa ou licenciadores;
- Conhecer as potenciais alternativas técnicas;
- Definir potenciais rotas para aperfeiçoamentos em produtos e processos existentes;
- Encontrar soluções técnicas para um problema específico;
- Avaliação de rotas de pesquisa, para desprezar aquelas cujo potencial de fracasso ou de risco é elevado, a fim de evitar desperdício de recursos;
- Monitoração das atividades de empresas concorrentes; e
- Prospecção tecnológica.

Outros autores também apontam o uso da patente como fonte de informação tecnológica. França (2007) afirma que as patentes recém-publicadas podem atuar como indicadores do estado da arte, apresentando a informação mais recente em um dado setor da técnica, pois o pedido de patente deve necessariamente demonstrar o que preexistia e o que está sendo reivindicado como novidade.

Ainda segundo o autor, a patente pode ser útil em uma negociação de transferência de tecnologia, já que permite tanto a identificação de alternativas técnicas para o atendimento das necessidades da indústria, como busca de empresas que atuam em determinado segmento tecnológico.

Outra possibilidade é a análise de um setor industrial utilizando um conjunto de patentes depositadas ao longo de um determinado período de tempo, informando a evolução do setor e indicando novos caminhos de desenvolvimento.

Para Araújo (1981), o uso da patente como fonte de informação tecnológica pode ser destacado na identificação de tecnologias emergentes em um campo

específico da técnica e dentro de certo período de tempo, não somente refletindo a atividade inventiva e a "produção" de novo conhecimento técnico em um país, mas também possibilitando a identificação de atividades industriais vindouras, indicando assim novas tendências tecnológicas e novos desenvolvimentos, muito antes que seus efeitos sejam sentidos no mercado.

A identificação de tecnologias alternativas disponíveis quando do processo de negociação de uma dada tecnologia é outro ponto abordado pela autora. A análise do mercado internacional de tecnologia patenteada é da máxima importância para os países em desenvolvimento, cuja indústria utiliza, em grande parte, a tecnologia estrangeira, através de contratos de licença.

A identificação dos atores de uma dada tecnologia é outro ponto importante informado em documentos de patente. Os dados bibliográficos provêm informação, tanto sobre quem inventou, quanto sobre quem possui uma dada tecnologia. A identificação dos atores pode ser utilizada, como por exemplo, para a identificação de clientes ou potenciais competidores por parte de uma empresa; para agências governamentais de controle, como elementos de interesse para verificar se há uma concentração de empresas em um dado ramo industrial; para a geração de um cadastro de inventores independentes em uma dada tecnologia.

A patente pode ser utilizada para a ordenação dos fluxos tecnológicos com o exterior. As patentes estrangeiras solicitadas no país podem indicar em que área o capital externo pretende atuar. Isso possibilita, aos organismos governamentais ou privados, a formulação de políticas ou a implementação de ações que necessitam ser tomadas caso a área seja de interesse prioritário para o país.

Para Araújo (1981), ainda há outras possibilidades de uso como: (i) a formulação de políticas setoriais, de Ciência e Tecnologia, e industrial; (ii) o apoio ao setor produtivo, através da possibilidade de desenvolvimento de tecnologias mais adequadas às matérias-primas locais; e (iii) a melhoria da capacidade de tomada de decisão, tanto técnica quanto estratégica, tanto por parte do governo, como por parte das empresas e das instituições de Pesquisa e Desenvolvimento (P&D).

A análise da visão dos autores citados para o uso da informação contida em documentos de patente mostra que depende da abordagem feita pelo usuário, embora todos concordem da grande importância da informação em patentes. No quadro 4, a seguir são sumarizadas as abordagens apresentadas:

**Quadro 4: Abordagens do uso da informação contida em documentos de patentes**

<b>Abordagens</b>	<b>Informação</b>
Busca do estado da técnica	Estágio de P&D em que se encontra uma determinada tecnologia. Também permite a identificação de tecnologias alternativas, além da avaliação de uma tecnologia específica que está sendo oferecida ou considerada para licença.
Busca de novidade	Existência, ou não, da característica de novidade nos resultados de uma dada pesquisa. Permite contestar um pedido de patente ou uma patente já concedida, servindo de base para a anulação da patente no caso de falta de novidade.
Busca de patenteabilidade ou validade	A avaliação de, não apenas a novidade, mas também da atividade inventiva, ou seja, se a alegada invenção é ou não óbvia em relação à tecnologia similar conhecida. Normalmente, é a busca realizada pelas repartições nacionais de patentes durante o processo de pedido de patente.
Busca por nome	Informações envolvendo companhias específicas ou indivíduos, tais como depositantes, cedentes de tecnologia, titulares ou inventores.
Busca da atividade tecnológica	Identificação das companhias e/ou inventores que estão em atividade em um campo tecnológico específico. Também permite localizar os países em que uma dada tecnologia está sendo patenteada.
Busca de violação de direitos	Localização de patentes ou pedidos de patente publicados que possam ser violados por uma dada atividade industrial.
Busca de uma família de patentes	Informações sobre a seleção dos países nos quais um dado pedido de patente foi depositado (com relação aos Estados Unidos, a informação refere-se à patente já concedida), localização de um pedido de patente que esteja escrito em uma língua conhecida pelo usuário, entre outra. É também um bom índice de avaliação da importância que o depositante dá àquela invenção. Depósitos de uma mesma invenção em muitos países normalmente significam que o depositante acredita no elevado potencial de comercialização do produto/processo/uso.
Busca de status legal	Dados sobre a validade de uma patente ou de um pedido de patente com relação à legislação de patente aplicável em um ou mais países.

Fonte: Macedo, Müller e Moreira (2001).

Macedo e Barbosa (2000) apontam inúmeras razões para que o sistema de informação tecnológica contido em documentos de patentes seja superior a outros sistemas de informação:

- Tecnologia por excelência: Único sistema de informação precipuamente configurado para a finalidade de armazenar conhecimentos tecnológicos, isto é, destinados à produção de mercadorias.



- Classificação tecnológica: A CIP é o único sistema de classificação exclusivamente configurado para ordenar informações técnicas de produção, ou seja, é restrita e especializada para atender à área da produção econômica.
- Complementaridade: A CIP possui a característica de Interconexão com as mais variadas fontes e sistemas de informação, dada a gama de informação contida nos documentos de patente que facilitam e propiciam a entrada em outros sistemas de informação.
- Originalidade: A invenção, para ser concedida como patente, deve ter novidade, ou seja, ser original.
- Atualidade: O documento de patente apresenta a mais atualizada informação tecnológica existente, ou seja, o estado da técnica.
- Competitividade técnica e econômica: Permite inferir uma prospectiva dos ramos de atividade para os quais caminha a indústria, a agricultura etc. Conhecendo-se as mais recentes invenções no ramo de atividade de determinada empresa, é possível detectar os caminhos de pesquisa de empresas concorrentes, o estágio de avanço etc. Desta forma, é possível a tomada de novas rotas, em face do melhor conhecimento da trilha seguida pelos concorrentes.
- Padronização e uniformidade: As informações são uniformes e padronizadas, ou seja, há um padrão na formatação do documento de patente, o que facilita a busca por informação em documento de patente em outros países.
- Quebra da barreira linguística: A patente, por ser um documento territorial, é obrigatoriamente redigida na língua onde houve o depósito. A seleção dos territórios nos quais a proteção será requerida é influenciada pela existência de mercado consumidor e/ou de potenciais parceiros para reproduzir o objeto da invenção. O conjunto de documentos de patente referentes à invenção – isto é, as patentes solicitadas e/ou concedidas em diversos países, semelhantes à patente do país de origem (país do primeiro pedido de invenção) - é conhecido como Família de Patentes.

Para se afirmar que a patente é uma fonte de informação tecnológica faz-se necessário entender como está estruturado o documento de patente.

## 5.2. A ESTRUTURA DO DOCUMENTO DE PATENTE

Macedo e Barbosa (2000) afirmam que o documento de patente apresenta seis agregados de informações, a saber:

- Folha de rosto;
- Relatório Descritivo;
- Reivindicações;
- Listagem de sequências, se for o caso;
- Desenho(s) e/ou Figura(s), se for o caso; e
- Resumo da invenção.

“A folha de rosto é a porta de acesso ao sistema de informação patentária, sendo elaborada pela autoridade governamental depositária do pedido de patente” (MACEDO; BARBOSA, 2000, p. 63). A folha de rosto contém os principais dados bibliográficos dos demais documentos e outras informações relacionadas, identificação do inventor, do titular, endereços etc. Os demais documentos (Relatório Descritivo, Reivindicações, Listagem de sequências Desenho(s) e/ou Figura(s) e Resumo) oferecem a informação técnica da invenção.

No relatório descritivo é apresentada uma descrição detalhada dos procedimentos tecnológicos associados à invenção (OLIVEIRA *et al.*, 2005, S38). Destaca-se o estado da técnica associado à invenção, evidenciando os problemas técnicos existentes e ressaltando as vantagens da invenção em relação ao dito estado da técnica. Adicionalmente, descreve-se detalhadamente a invenção, de forma que um técnico no assunto possa reproduzi-la.

As reivindicações indicam o âmbito de abrangência da matéria reivindicada e, conseqüentemente, delimitam a proteção assegurada por ela. (OLIVEIRA *et al.*, 2005, S38). As reivindicações são fundamentadas no relatório descritivo, e o número de reivindicações define o objeto da invenção lembrando que não há um número máximo de reivindicações.

A Listagem de sequências, assim como os desenho(s) e/ou figura(s) complementam a suficiência descritiva exigida em um documento de patente.

No resumo da patente é apresentado o relatório descritivo de forma concisa. Este resumo é de grande importância para a publicação oficial do pedido de patente, pois serve como referência e orientação dos usuários na busca do estado da técnica.

Macedo e Barbosa (2000) afirmam que uma das principais vantagens do sistema de informação patentária é a sua padronização internacional. A padronização é feita por um código numérico denominado como “*International Agreed Numbers for the Identification of Data*” (INID). Segundo Longa (2007), o código identifica, na primeira página do documento, denominada “folha de rosto do documento”, os campos com dados bibliográficos específicos, presentes em todos os documentos de patente, independente do idioma. Isto significa que independente de onde a patente seja depositada, independente da língua, pelo código pré-estabelecido, qualquer pessoa consegue identificar o conteúdo de cada campo na folha de rosto.

A codificação INID, segundo Macedo e Barbosa (2000), é formada por oito grupos classificados de 10 a 80, contendo subdivisões. A seguir apresentam-se os códigos INDI com suas respectivas identificações.

**Quadro 5: Códigos INID**

INID	Significado
(10) IDENTIFICAÇÃO DO DOCUMENTO.	(11) Número do documento publicado. É sua verdadeira identificação. Essa informação é importante para se requerer, por exemplo, a cópia do documento. (12) Designação do tipo de documento (pedido de patente, patente concedida etc.). (13) Tipo de código de documento conforme o padrão OMPI ST. 16. (19) Padrão OMPI ST. 3, ou outra identificação da autoridade nacional editora.
(20) DADOS DE PEDIDOS NACIONAIS.	(21) Número(s) designado(s) para (o) pedido(s) nacional(is). Corresponde ao número que o pedido de patente recebeu no ato do depósito. (22) Data(s) de depósito do(s) pedido(s). (23) Outras(s) data(s), incluindo a de depósito das especificações completas, após especificações provisórias, ou data de exibição pública.
(30) DADOS DE PRIORIDADE.	(31) Número(s) designado(s) para a prioridade unionista. (32) Data(s) de depósito de pedido(s) de prioridade unionista. (33) Padrão OMPI ST.3, identificando a autoridade governamental nacional outorgante do número do pedido de prioridade ou a autoridade regional outorgante do número do pedido de prioridade regional. Para os pedidos internacionais sob o PCT, deve ser usado o código “WO”.
(40) DATA(S) DE ACESSO AO PÚBLICO.	(41) Data de acessibilidade ao público do documento por vista, ou cópia, por solicitação, de um documento ainda não examinado e sem concessão na data ou antes da data mencionada. (42) Data de acessibilidade ao público por vista, ou cópia por solicitação, de um documento examinado e com concessão na data ou antes da data mencionada.

(43)	Data de publicação impressa, ou processo similar, de um documento não examinado sem concessão antes ou até a data mencionada. Data em que ocorreu a divulgação da invenção.
(44)	Data de publicação impressa, ou processo similar, de um documento não examinado sem concessão antes ou até a data mencionada.
(45)	Data da publicação impressa, ou por processo similar, de um documento examinado e com concessão ou somente uma concessão provisória na data ou antes da data mencionada.
(47)	Data da acessibilidade ao público por vista, ou cópia por requisição, de um documento em que a concessão foi na data ou antes da data mencionada.
<b>(50) INFORMAÇÃO TÉCNICA.</b>	
(51)	Classificação Internacional de Patente. Fornece, com precisão, as divisões e as subdivisões da Classificação Internacional de Patentes em que estão arquivadas as informações relativas à invenção.
(52)	Classificação nacional.
(53)	Classificação Decimal Universal (CDU).
(54)	Título da Invenção.
(55)	Palavras-chave.
(56)	Relação de documentos sobre o estado da técnica anterior, se separados do relatório descritivo.
(57)	Resumo ou reivindicações.
(58)	Campo de busca.
<b>(60) REFERÊNCIA A OUTROS DOCUMENTOS DE PATENTES NACIONAIS LEGAIS OU PROCESSUALMENTE RELACIONADOS, INCLUINDO PEDIDOS NÃO PUBLICADOS.</b>	
(61)	Número e, se possível, data de depósito de pedidos anteriores, ou número de publicações anteriores, ou número de patentes anteriormente concedidas, certificados de inventor, modelo de utilidade ou similares em relação ao(s) qual(is) o presente documento é aditivo.
(62)	Número e, se possível, data de depósito de um pedido anterior em relação ao qual o presente documento é uma parte dividida.
(63)	Número e data do depósito de um pedido anterior em relação ao qual o pedido está relacionado, p. ex., pedido divisional.
(64)	Número de uma publicação anterior que é "reeditado".
<b>(70) IDENTIFICAÇÃO DAS PARTES RELACIONADAS COM O DOCUMENTO</b>	
(71)	Nome(s) do(s) depositante(s).
(72)	Nome do(s) inventor(es).
(73)	Nome(s) do(s) outorgado(s) titular(es).
(74)	Nome do procurador. Informa quem foi o representante legal do depositante.
<b>(80) IDENTIFICAÇÃO DE DADOS RELACIONADOS ÀS CONVENÇÕES INTERNACIONAIS ALÉM DA CONVENÇÃO DE PARIS.</b>	
(81)	Países designados. Identifica os Estados nos quais a proteção por patente foi solicitada.
(85)	Data do início da fase nacional.
(86)	Número, idioma e data do depósito internacional.
(87)	Número, idioma e data da publicação internacional.

Fonte: elaboração própria a partir de Macedo e Barbosa (2000); Macedo, Müller e Moreira (2001); INPI (2011).

Segundo Macedo e Barbosa (2000), e conforme o código INID, toda a designação da folha de rosto relativa à determinação do local em que está classificada a informação situa-se no grupo (50) Informação Técnica. E, neste grupo, os subgrupos (51) Classificação Internacional de Patentes; (54) Título da Invenção; (56) Relação de documentos sobre o estado da técnica anterior, se separados do relatório descritivo e (57) Resumo ou reivindicações são os mais apresentados na folha de rosto do documento de patente.

“O campo (51) da CIP é [...] o mais importante, pois [...] possibilita a existência do sistema internacional de informação técnico-produtivo, bem como permite a rápida recuperação das informações nele contidas” (MACEDO, BARBOSA, 2000, p. 69).

### 5.3. CLASSIFICAÇÃO INTERNACIONAL DE PATENTE (CIP)

Nos anos anteriores à década de 70, surgiram os primeiros esforços da Organização Mundial da Propriedade Intelectual (OMPI) para uniformizar e padronizar os diversos sistemas nacionais de classificação de documentos de patente, com o objetivo de difundir internacionalmente a informação contida nos mesmos.

Segundo Macedo e Barbosa (2000), em 1967 iniciaram-se as negociações entre a WIPO e o Conselho da Europa com base na Convenção Européia sobre a Classificação Internacional de Patentes de Invenção, firmada entre os países do Conselho em 1954. Em 1971, os países membros da União Internacional para a Proteção da Propriedade Intelectual firmaram o Acordo de Estrasburgo para a adoção de uma classificação internacional de patentes, classificação esta que entrou em vigor em 1975.

Segundo o guia da CIP fornecido pela WIPO, oitava edição (2006), esta classificação é comum para patentes de invenção, incluindo os pedidos de patentes publicados, certificados de inventores, modelos de utilidades e certificados de adição (de agora em diante designados "documentos de patentes").

França (2007) afirma que a CIP, criada em 1971 pelo Acordo de Estrasburgo, teve por objetivo a necessidade de recuperação da informação em patentes. Já o guia da CIP, fornecido pela WIPO e na sua oitava edição (2006), considera a CIP como um meio internacionalmente usado para se obter uma classificação uniforme de documentos de patentes, e que tem, por finalidade principal:

Criar uma ferramenta de busca eficaz para a recuperação de documentos de patentes pelos escritórios de propriedade intelectual e demais usuários, a fim de instituir a novidade e avaliar a etapa inventiva ou não obviedade (avaliando, inclusive, o avanço técnico e os resultados úteis ou sua utilidade) das características técnicas dos pedidos de patentes. (WIPO, 2006, p. 7)

Além disso, de acordo com o mesmo guia, a Classificação tem outras finalidades importantes como, por exemplo, servir de:

- Instrumento para disposições organizadas dos documentos de patente, a fim de facilitar o acesso às informações tecnológicas e legais contidas nos mesmos;
- Base de disseminação seletiva de informações a todos os usuários das informações de patentes;
- Base para investigar o estado da técnica em determinados campos da tecnologia;
- Base para preparar estatísticas sobre propriedade industrial que permitam a avaliação do desenvolvimento tecnológico em diversas áreas.

Com o intuito de avaliar a CIP e manter o sistema atualizado, considerando, principalmente, os avanços e o desenvolvimento tecnológico, a classificação é revisada por uma comissão de peritos que se reúne periodicamente, geralmente a cada 5 anos. Atualmente, a CIP está em sua oitava edição, a qual passou a vigorar desde janeiro de 2006.

### **5.3.1. A estrutura da CIP**

A estrutura da CIP pode ser entendida como uma chave de identificação que oferece alternativas em cada destinação e a escolha de uma das alternativas determina a etapa seguinte. A CIP divide a tecnologia em oito seções, com aproximadamente 70.000 subdivisões. Cada subdivisão tem um símbolo que consiste em algarismos arábicos e letras do alfabeto latino ou romano. Os símbolos da CIP são indicados em cada documento de patente e são distribuídos pelo escritório de propriedade industrial nacional ou regional que publica o documento de patente. Para documentos PCT, símbolos da CIP são atribuídos pela autoridade responsável pela pesquisa internacional.

Mais especificamente, a CIP é dividida em seções, subseções, classes, subclasses, grupos e subgrupos. Para a WIPO (2006), o total de 8 seções representa todo o conhecimento que possa ser considerado apropriado ao campo de patentes de invenção. As seções são o nível mais alto da hierarquia da classificação (WIPO, 2006, p. 10). Cada seção é subdividida em classes, as quais são o segundo nível da hierarquia da classificação, cada classe abrange uma ou mais subclasses,

as quais são o terceiro nível hierárquico da classificação. Cada subclasse é desdobrada em subdivisões, denominadas "grupos", que podem ser grupos principais (i.e. o quarto nível hierárquico da classificação) ou subgrupos (i.e. e níveis hierárquicos inferiores que dependem do nível do grupo principal da classificação). A seguir é especificado o que representa cada campo da CIP.

#### Nível 1: Seção

**Símbolo:** É identificada por uma letra maiúscula, de A até H.

**Título:** Deve ser considerado uma indicação bem ampla do conteúdo da seção.

**Conteúdo:** É seguido por um resumo dos títulos de suas principais subdivisões.

**Subseção:** São títulos sem símbolos de classificação

**Quadro 6: Seções da CIP**

<b>Seção</b>	<b>Título</b>
A	Necessidades Humanas
B	Operações de Processamento; Transporte
C	Química; Metalurgia
D	Têxteis; Papel
E	Construções Fixas
F	Engenharia Mecânica; Iluminação; Aquecimento; Armas; Explosão
G	Física
H	Eletricidade

Fonte: Elaboração própria.

#### Nível II: Classes

**Símbolo** - É constituído pelo símbolo da seção seguido de um número com dois dígitos. Ex.: G01

**Título** - Indica o conteúdo da classe. Ex.: Medição; Teste (G01)

**Índice** - Algumas classes têm um índice que serve apenas de resumo informativo, fornecendo um levantamento geral do conteúdo da classe.

**Quadro 7: Classes da CIP: a G01**

<b>Seção: Física</b>	
<b>Subseção: Instrumentos</b>	
<b>Classes</b>	
G01	Medição; Teste
G02	Óptica
G03	Fotografia; Cinematografia; Técnicas semelhantes utilizando ondas outras que não ondas ópticas; Eletrografia; Holografia

G04	Horologia
G05	Controle; Regulagem
G06	Cômputo; Cálculo; Contagem
G07	Dispositivos de Teste
G08	Sinalização
G09	Educação; Criptografia; Apresentação visual; Anúncios; Logotipos
G10	Instrumentos Musicais; Acústica
G11	Armazenamento de informações
G12	Detalhes de instrumentos
G21	Física Nuclear; Engenharia Nuclear
G99	Matéria não incluída em outro local desta seção

Fonte: Elaboração própria.

### Nível III: Subclasses

**Símbolo** - É constituído pelo símbolo da classe, seguido de uma letra maiúscula. Ex.: G01N

**Título** - Indica o conteúdo da subclasse. Ex.: Investigação ou análise dos materiais pela determinação de suas propriedades químicas ou físicas (G01N).

**Índice** - Algumas subclasses têm um índice que serve apenas como resumo informativo, oferecendo um levantamento geral do conteúdo da subclasse.

**Cabeçalho Guia** - Quando uma grande parte da subclasse se refere à matéria em comum, um cabeçalho guia indicando tal matéria pode ser fornecido no início de tal parte.

**Quadro 8: Subclasses da CIP referente à classe G01**

<b>Seção: Física</b>	
<b>Subseção: Instrumentos</b>	
<b>Classe: G01 Medição; Teste</b>	
<b>Subclasses</b>	
G01B	Medição de comprimentos, espessuras ou outras dimensões lineares semelhantes; medição de ângulos; medição de áreas; medição de irregularidades de superfícies ou contornos
G01C	Medição de distâncias, níveis ou rumos; topografia; navegação; instrumentos giroscópicos; fotogrametria ou videogrametria (medição do nível dos líquidos G01F; rádio navegação, determinando distância ou velocidade pelo uso de efeitos de propagação, por exemplo, efeito Doppler, tempo de propagação, de ondas de rádio, disposições similares usando outros tipos de ondas G01S)
G01D	Medição não especialmente adaptadas para uma variável específica; disposições para medir duas ou mais variáveis não abrangidos por uma única outra subclasse; aparelhos para medir tarifas; medição ou teste não incluídas em outro local
G01F	Medição de volumes, débitos volumétricos, débitos de massa, ou níveis de líquidos; medição por volume (dispositivos sensores do fluxo de leite em máquinas ou dispositivos para produção de leite A01J 5/01; medição ou registro do fluxo sanguíneo A61B 5/02, A61B 8/06; introdução de material no corpo humano A61M 5/168; buretas ou pipetas B01L 3/02; disposições de medidores de volume ou de medidores de débitos volumétricos em aparelhos de escoamento líquido, por ex., para vendas a varejo. B67D 7/16; bombas, motores a fluido, detalhes construtivos comuns aos dispositivos de contagem ou de medição e à bombas ou a motores a fluido. F01-F04; determinação da posição, direção ou velocidade utilizando a reflexão ou a reirradiação de ondas de rádio ou de outras ondas, em geral G01S; sistemas para controle de proporcionalidade G05D 11/00
G01G	Pesagem



G01H	Medição de vibrações mecânicas ou de ondas ultrassônicas, sônicas ou infrassônicas
G01J	Medição da intensidade, velocidade, conteúdo do espectro, polarização, fase ou pulsos característicos da luz infravermelha visível ou ultravioleta; colorimetria; pirometria das radiações.
G01K	Medição de temperatura; medição da quantidade de calor; elementos termosensíveis não incluídos em outro local (pirometria das radiações G01J 5/00)
G01L	Medição da força, tensão, torque, trabalho, potência mecânica, eficiência mecânica, ou pressão dos fluidos (pesagem G01G)
G01M	Teste do equilíbrio estático ou dinâmico de máquinas ou estruturas; teste de estruturas ou de aparelhos não incluídos em outro local
G01N	Investigação ou análise dos materiais pela determinação de suas propriedades químicas ou físicas (separação de componentes de materiais em geral B01D,B01J, B03, B07; aparelhos totalmente abrangidos por uma outra única subclasse, ver a subclasse apropriada, por ex., B01L; processos de medição ou teste, outros que não ensaios imunológicos, envolvendo enzimas ou micro-organismos C12M, C12Q; investigação do solo para fundações in situ E02D 1/00; dispositivos para monitoramento ou diagnóstico para aparelho de tratamento de exaustão de gás F01N 11/00; indicação das variações de umidade para compensação de medições de outras variáveis ou para compensação de leituras de instrumentos devido a variações da umidade, ver G01D ou a subclasse apropriada para a variável medida; teste ou determinação das propriedades das estruturas G01M; medição ou teste das propriedades elétricas ou magnéticas dos materiais G01R; sistemas em geral para determinar distância, velocidade ou presença, utilizando a recepção ou emissão de ondas de rádio ou outras ondas baseada nos efeitos da propagação, por ex., efeito Doppler, tempo de propagação G01S; determinação da sensibilidade, granulidade, ou densidade de materiais fotográficos G03C 5/02; teste das peças componentes de um reator nuclear G21C 17/00)
A subclasse continua até a letra W	

Fonte: elaboração própria

#### Nível IV: Grupos e Subgrupos

**Símbolo** - É constituído pelo símbolo da subclasse, seguido de dois números separados por uma barra oblíqua. Ex.: G01N

**Símbolo do Grupo Principal** - É constituído pelo símbolo da subclasse, seguido de um número com um a três dígitos, da barra oblíqua e o número 00. Ex.: G01N 33/00

**Título do Grupo Principal** - Define um campo de matéria dentro do escopo de sua subclasse considerada de utilidade na busca de invenções. Ex.: Investigação ou análise de materiais por métodos específicos não abrangidos pelos grupos G01N 1/00-G01N 31/00 (G01N 33/00)

**Subgrupo Símbolo** - Formam subdivisões sob os grupos principais. É constituído pelo símbolo da subclasse, seguido de um número com um a três dígitos do seu grupo principal, da barra oblíqua e de um número, com pelo menos dois dígitos, que não sejam 00. Ex.: G01N 33/48

**Subgrupo Título** - Define, de forma precisa, um campo de matéria dentro do escopo do seu grupo principal, o qual é considerado de utilidade na busca de invenções. O título é precedido por um ou mais pontinhos, indicando a posição hierárquica desse subgrupo, isto é, indicando que cada subgrupo forma uma subdivisão do grupo mais próximo acima deste, tendo um pontinho a menos. O título do subgrupo é geralmente um termo completo e, neste caso, começa com letra maiúscula. Um

título do subgrupo começa com uma letra minúscula, no caso de ser lido como continuação do título do próximo grupo mais elevado e menos recuado, do qual ele depende, isto é, tendo um pontinho a menos. Em todos os casos, o título do subgrupo deve ser lido como sendo dependente de, e restrito ao título do grupo sob o qual está recuado.

**Quadro 9: Grupos e Subgrupos da CIP referente à classe G01**

<b>Seção: Física</b>	
<b>Subseção: Instrumentos</b>	
<b>Classe: G01 Medição; Teste</b>	
<b>Subclasse: G01N Investigação ou análise dos materiais pela determinação de suas propriedades químicas ou físicas</b>	
<b>Grupo: G01N 33/00 Investigação ou análise de materiais por métodos específicos não abrangidos pelos grupos G01N 1/00-G01N 31/00</b>	
<b>Subgrupos</b>	
O subgrupo começa em G01N 1/02	
G01N 33/48	. Material biológico, por ex., sangue, urina (G01N 33/02-G01N 33/14, G01N 33/26, G01N 33/44, G01N 33/46 têm prioridade; determinação da capacidade germinativa das sementes A01C 1/02); Hemocitrômetros (contagem de glóbulos distribuídos sobre uma superfície realizando a varredura desta superfície G06M 11/02)
G01N 33/483	. . Análise física de material biológico
G01N 33/487	. . . de material biológico líquido
G01N 33/49	. . . . do sangue
G01N 33/493	. . . . da urina
G01N 33/497	. . . de material biológico gasoso, por ex., respiração
G01N 33/50	. . Análise química de material biológico, por ex., sangue, urina; Testes por métodos envolvendo a formação ligações bioespecíficas de ligantes; Testes imunológicos (processos de medição ou testes outros que não os imunológicos envolvendo enzimas ou micro-organismos, composições ou papéis reativos para os mesmos; processos de preparação destas composições, controle de condições responsivas ao meio em processos microbiológicos ou enzimológicos.
G01N 33/52	. . . Utilização de compostos ou composições para investigação colorimétrica, espectrofotométrica ou fluorométrica, por ex., utilização de papel reagente
G01N 33/53	. . . Imuno-ensaio; Ensaio envolvendo ligantes bioespecíficos; Materiais para os mesmos (preparações contendo antígenos ou anticorpos A61K; haptenos em geral, ver os locais relevantes na classe C07; peptídeos, por ex., proteínas em geral C07K)
O subgrupo continua até G01N 35/10	

Fonte: Elaboração própria.

De acordo com as seções da CIP, pode-se perceber que as invenções são classificadas de acordo com os ramos da indústria, da "técnica" ou da atividade humana em relação às quais são caracteristicamente relevantes. Esse enfoque é o comumente designado de "orientação industrial", "orientação técnica", "orientação segundo o pedido de privilégio" (INPI, 2011). Neste caso, pode-se incluir, por exemplo, fiação, tecelagem, malharia, que figuram na CIP na Seção D ("Têxteis e Papel"). Por outro lado, as invenções são classificadas de acordo com as funções

para as quais são caracteristicamente pertinentes, dentre elas estão, por exemplo, transportar, embalar e estocar, que são funções que dizem respeito a quase todos os ramos da indústria, não havendo um campo específico. Esse enfoque é comumente chamado de "orientação segundo a função". (INPI, 2011).

Embora a CIP seja, em princípio, orientada para a função, na realidade ela combina ambos os enfoques. É o resultado da experiência adquirida por pessoas, cujas tarefas diárias consistem na comparação de invenções, para as quais a proteção de patente é reivindicada, com invenções similares já reveladas em documentos de patentes publicados.

A recuperação da informação contida nos documentos em patente é praticamente 100% realizada em bases de dados nacionais e internacionais disponíveis online.

#### 5.4. AS BASES DE PATENTES PÚBLICAS E PRIVADAS

Existem várias bases de patentes tanto públicas (gratuitas) quanto privadas (comerciais) disponíveis no mundo. Geralmente as bases de patentes públicas são as bases disponibilizadas pelos próprios escritórios onde a patente obrigatoriamente é depositada. A seguir apresentam-se algumas das principais bases públicas e privadas, algumas características e os links para o respectivo acesso.

**Quadro 10: Exemplos de bases públicas de patentes**

EXEMPLOS DE BASE PÚBLICAS DE PATENTES	
Base de Patentes	Características
INPI: Instituto Nacional da Propriedade Industrial (Escritório Brasileiro) <sup>2</sup>	Apresenta apenas documentos de patente brasileiros depositados a partir de 1992; Possibilidade de busca em vários campos; Pode utiliza operadores booleanos e truncagem; O resultado da busca é apresentado na forma de uma listagem; A visualização do documento original, em português, é "linkado" no Esp@cenet (nem sempre é disponível); Não há possibilidade de exportação dos dados; Não acessa família de patentes
USPTO: <i>United States Patent and Trademark Office</i> (Escritório Norte Americano) <sup>3 4</sup>	Apresenta apenas documentos de patente norte-americanos. Divisão de bases de dados entre pedidos de patente a partir de 2001 e textos completos de patente (patentes aprovadas) a partir de 1976 e imagens das patentes completas desde 1790; Possibilidade de busca em vários campos; Pode utiliza operadores booleanos e truncagem; O resultado da busca é apresentado na forma de uma listagem; É possível ter informações sobre os inventores e titulares de tais patentes, assim

<sup>2</sup> INPI: <http://pesquisa.inpi.gov.br/MarcaPatente/jsp/servimg/servimg.jsp?BasePesquisa=Patentes>

<sup>3</sup> USPTO – pesquisa em pedidos de patente: <http://appft1.uspto.gov/netahtml/PTO/search-adv.html>

<sup>4</sup> USPTO – pesquisa em patente concedida: <http://patft.uspto.gov/netahtml/PTO/search-adv.htm>

	como observar os relatórios descritivos, os exemplos, as reivindicações e outros; Possibilidade de impressão do documento original; Não há possibilidade de exportação dos dados; Não acessa família de patentes.
<i>EPO - Esp@cenet:</i> European Patent Office (Escritório Europeu <sup>5</sup> )	Apresenta documentos de patentes do mundo todo depositados a partir de 1970. Possibilidade de busca em vários campos; Pode utiliza operadores booleanos e truncagem; O resultado da busca é apresentado na forma de uma listagem; Somente mostra os primeiros 500 documentos; Exporta resultados nos formatos CSV e XLS; Possibilidade de impressão do documento original. Acesso à família de patentes
<i>CIPO: Canadian Intellectual Property Office</i> (Escritório Canadense <sup>6</sup> )	Apresenta apenas documentos de patente canadenses a partir de 1920 Possibilidade de busca em vários campos; Pode utiliza operadores booleanos e truncagem; O resultado da busca é apresentado na forma de uma listagem; Possibilidade de impressão do documento original; Não há possibilidade de exportação dos dados.
<i>WIPO: World Intellectual Property Organization</i> (PCT database) <i>Patentscope® - WIPO</i> <sup>7</sup>	Apresenta documentos de patente depositados via PCT, além de coleções de alguns países, inclusive de pedidos depositados no Brasil; Possibilidade de busca em vários campos; Pode utiliza operadores booleanos e truncagem; O resultado da busca é apresentado na forma de uma listagem; Possibilidade de impressão do documento original; Não há possibilidade de exportação dos dados; Não acessa família de patentes.
<i>Free Patente Online</i> <sup>8</sup>	Apresenta apenas documentos de patentes norte-americanos e europeus, resumos do Japão e PCT; Possibilidade de busca em vários campos; Pode utiliza operadores booleanos e truncagem; O resultado da busca é apresentado na forma de uma listagem; Possibilidade de impressão do documento original; Exporta resultados individualmente para o EndNote; Não acessa família de patentes.
<i>Patentlens</i> <sup>9</sup>	Apresenta patentes concedidas nos Estados Unidos, Europa; Austrália e pedidos de patente nos Estados Unidos, PCT e Austrália; Possibilidade de busca em vários campos; Pode utiliza operadores booleanos e truncagem; O resultado da busca é apresentado na forma de uma listagem; Possibilidade de impressão do documento original; Exporta resultados nos formatos RIS; Acesso à família de patentes

Fonte: elaboração própria

#### Quadro 11: Exemplos de bases privadas de patentes

EXEMPLOS DE BASES PRIVADAS DE PATENTES	
Base de Patentes	Características
INPADOC ( <i>International Patent documentation Center</i> )	Esta é a mais completa base sobre patentes, pois engloba referências sobre 96% das patentes concedidas por 71 países, em todas as áreas do conhecimento. Apesar de não conter os resumos das patentes, fornece as seguintes informações: título; nomes dos inventores, proprietários e depositantes; códigos de classificação nacional e internacional; número, data e país de publicação do documento de patente; data e país de prioridade; família da patente (quando concedida por vários países); status legal (somente para as patentes concedidas pelos organismos internacionais EPO- <i>European Patent Office</i> e WIPO- <i>World Intellectual Property Organization</i> e para patentes concedidas no seguintes países: Áustria, Alemanha,

<sup>5</sup> Esp@cenet: [http://worldwide.espacenet.com/advancedSearch?locale=en\\_EP](http://worldwide.espacenet.com/advancedSearch?locale=en_EP)

<sup>6</sup> CIPO <http://brevets-patents.ic.gc.ca/opic-cipo/cpd/eng/introduction.html>

<sup>7</sup> Patentscope® - <http://www.wipo.int/patentscope/search/en/search.jsf>

<sup>8</sup> Free Patent Online <http://www.freepatentsonline.com/search.html>

<sup>9</sup> Patentlens - <http://www.patentlens.net/patentlens/structured.html>

	Dinamarca, Espanha, Estados Unidos, França, Grã-Bretanha, Holanda, Hungria, Luxemburgo, Mônaco, Suíça, de acordo com o Tratado de Cooperação sobre Patentes, assinado em 1970). Segundo informações do INPI: Valor aproximado – US\$ 15,00 por documento de patente pesquisado
Geneseq	Permite a pesquisa sobre documentos de patente relacionados às sequências de ácidos nucleicos e de proteínas. Abrange sequências biológicas patenteadas desde 1981 em todo o mundo a partir de 41 escritórios de patente, incluindo WO, US, EP, JP, DE, IN e CN.
STN <sup>10</sup>	STN é operado pelo <i>Chemical Abstracts Service (CAS)</i> , uma divisão da <i>American Chemical Society</i> , na América do Norte; FIZ Karlsruhe (FIZ-K) na Europa e Associação Internacional de Química da Informação (JAICI), no Japão. Cobre uma variedade de temas, tais como: agricultura, saúde e segurança, biologia, ciência de materiais, biotecnologia, matemática, negócio, medicina, química, patentes, energia, petróleo, engenharia, farmacologia, alimentação humana e animal, física, geologia, toxicologia e regulamentações do governo.
<i>World Patents Index</i>	É uma das excelentes bases de dados sobre patentes, em várias áreas do conhecimento. Permite a recuperação das seguintes informações: título e demais dados da folha de rosto do documento de patente, inclusive resumo; família de patentes correlatas; referências bibliográficas. Dispõe de <i>Thesaurus</i> (compilações de termos/expressões controlados, usados na indexação dos documentos), para consulta em linha. Também disponibiliza diagramas, em sua maioria relacionados a patentes de engenharia, desde 1988. Segundo informações do INPI: Hora de conexão – US\$ 354,00 + valor do “display” por documento
<i>Derwent Innovation Index</i> <sup>11</sup>	Contém informações sobre documentos de patentes publicados mundialmente. Equivale à base “ <i>World Patents Index</i> ” e apresenta a vantagem de indexar textos completos dos documentos de patentes.

Fonte: elaboração própria.

Uma análise no quadro acima permite verificar que há certas limitações no uso tanto nas bases gratuitas quanto nas comerciais. A maior parte das bases gratuitas abrangem documentos nacionais e as que abrangem documentos depositados internacionalmente ou apresentam alguma limitação ou não possuem uma interface amigável para coleta dos dados para análise. Com relação às bases comerciais, um maior número de campos podem ser acessados, as buscas podem ser feitas no documento completo, os mecanismos de busca são mais flexíveis e em algumas bases os resumos dos documentos são reescritos, o que otimiza a recuperação de documentos através de palavras-chave, porém toda essa eficiência gera custo bem mais elevado.

<sup>10</sup> STN: <https://stnweb.cas.org/>

<sup>11</sup> Derwent <http://portal.isiknowledge.com/portal.cgi?DestApp=DIIDW&Func=Frame>

## 6. MALÁRIA

A malária, também conhecida como paludismo, febre palustre, maleita e sezão é uma doença infecciosa aguda ou crônica causada por protozoários parasitas que pertence ao filo *Apicomplexa*, classe *Sporozoa*, ordem *Coccidiida*, subordem *Haemosporidiidae*, família *Plasmodiidae* e gênero *Plasmodium* que são transmitidos para o ser humano através da picada da fêmea do mosquito *Anopheles* produzindo febre, além de outros sintomas.

Segundo Brasil (2011), cinco espécies de protozoários do gênero *Plasmodium* podem causar a malária humana: *Plasmodium falciparum*, *P. vivax*, *P. malariae*, *P. ovale* e *P. knowlesi*. No Brasil, somente as três primeiras espécies deste parasita estão presentes, sendo as infecções por *P. vivax* predominantes, seguido das infecções por *P. falciparum*. A quarta espécie de protozoário, o *P. ovale*, ocorre apenas no continente africano e o *P. knowlesi* no Sudeste Asiático, porém, ocasionalmente, casos importados de outros países podem ser diagnosticados no Brasil.



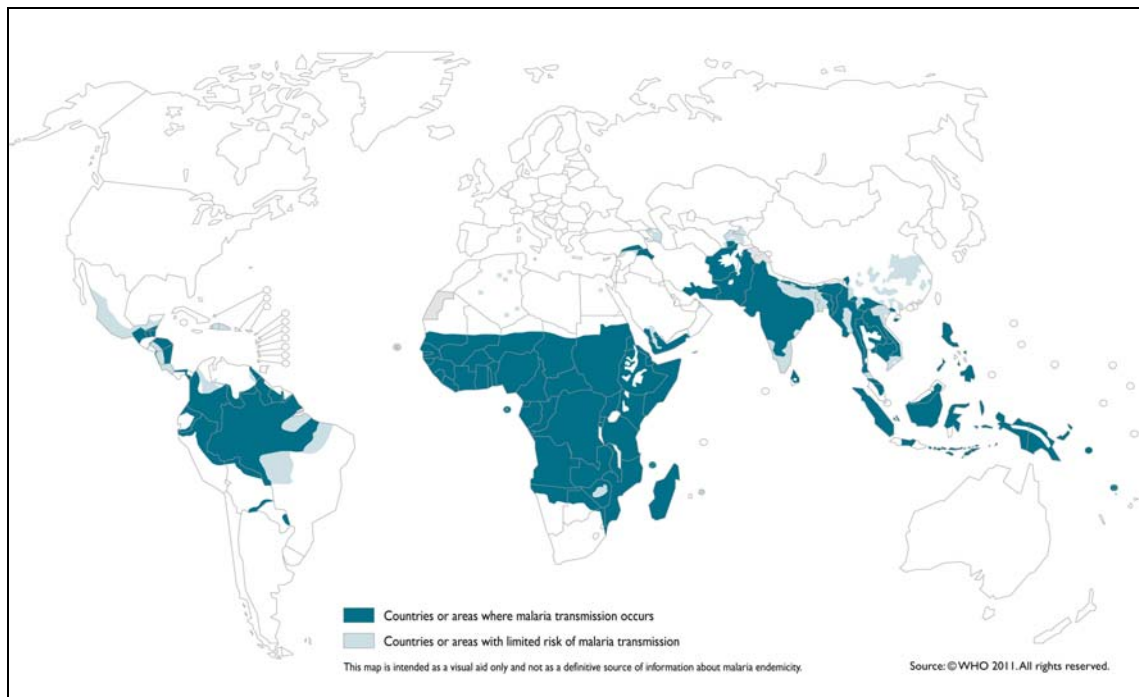
Figura 6 - Fêmea adulta de *Anopheles gambiae* em posição de alimentação de sangue na pele humana. WHO/TDR/Stammers

Além da transmissão pelo mosquito *Anopheles*, a malária pode ser transmitida acidentalmente por transfusão de sangue (sangue contaminado com plasmódio), pelo compartilhamento de seringas (em usuários de drogas ilícitas) ou por acidente com agulhas e/ou lancetas contaminadas. Há, ainda, a possibilidade de transmissão neonatal (Brasil, 2009, p. 14).

Oliveira (2009) aponta que os sintomas são febre paroxística, calafrios, sudorese, anemia e astenia. O *P. falciparum* é o responsável pela maioria dos casos graves e complicados de malária que podem evoluir a óbito, especialmente, ao atingir populações e indivíduos parcialmente imunes ou com ausência de imunidade adquirida contra a doença. A malária por *P. vivax* apresenta evolução clínica mais benigna, porém é responsável por grande perda de dias de trabalho e produtividade.

### 6.1. EPIDEMIOLOGIA DA MALÁRIA

Segundo o Brasil (2010), a malária é reconhecida como grave problema de saúde pública no mundo, ocorrendo em quase 50% da população, em mais de 109 países e territórios. Sua estimativa é de 300 milhões de novos casos e 1 milhão de mortes por ano, principalmente em crianças menores de 5 anos e mulheres grávidas do continente africano.

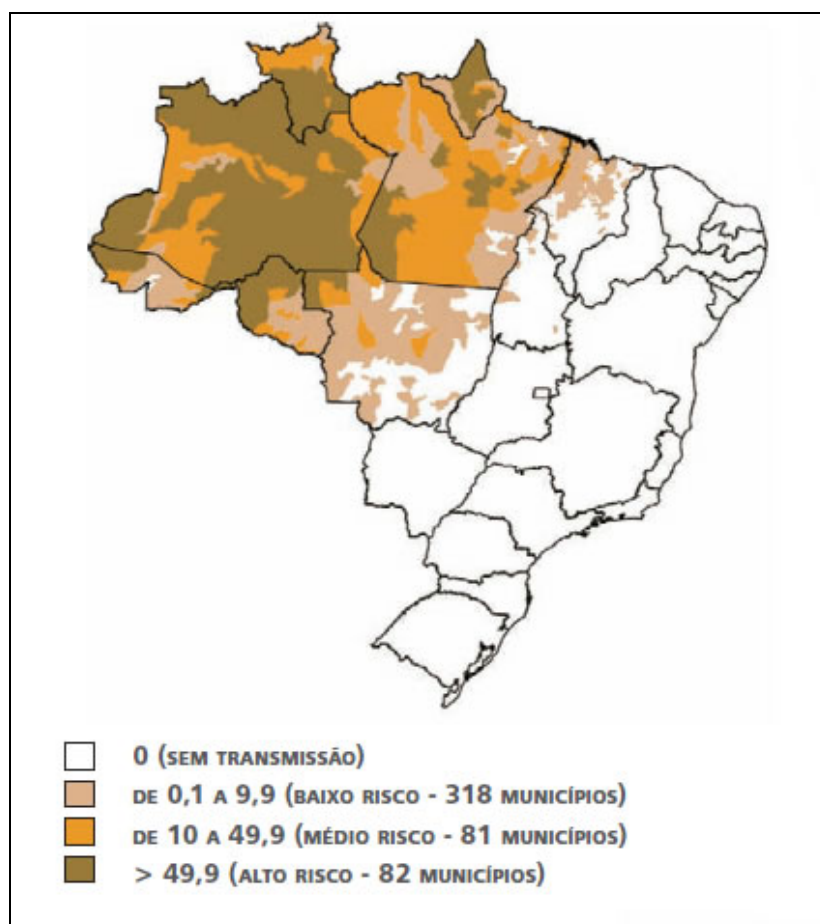


**Figura 7 - Países incidentes e áreas de risco da malária no mundo em 2010. WHO (2011)**

No Brasil, o quadro epidemiológico da malária é preocupante nos dias atuais. No ano de 2006, ainda segundo Brasil (2010), o Brasil registrou 545.696 casos de malária, sendo a causada pela espécie *Plasmodium vivax* a de maior incidência (73,4%). Arcanjo (2004) afirma que, no entanto, a transmissão do *P. falciparum*,

sabidamente responsável pela forma grave e letal da doença, tem apresentado redução importante nos últimos anos.

Embora em declínio, o número absoluto de casos no ano de 2008 ainda foi superior a 300.000 pacientes em todo o país. Dados deste mesmo ano revelam que a região da Amazônia é considerada a área endêmica do país para malária, por concentrar 99% dos casos distribuídos em seis estados: Acre, Amapá, Amazonas, Pará, Rondônia e Roraima. A distribuição espacial do risco de transmissão da doença no Brasil é apresentada na figura a seguir.



**Figura 8 - Distribuição espacial do risco de transmissão da doença no Brasil em 2008. SIVEP-malaria/svs/ms/2007 em Brasil (2009)**

Analisando-se o mapa, pode-se comprovar que a transmissão da malária não é completamente estável. De acordo com a incidência parasitária anual (IPA), costuma-se classificar as áreas endêmicas como de alto risco (IPA > 50/1.000 hab.), médio risco (IPA entre 10-49/1.000 hab.) e baixo risco (IPA < 10/1.000 hab.).

No Brasil, a grande incidência de casos na região amazônica que a torna endêmica está relacionada, segundo o Ministério da Saúde a:



Fatores biológicos (presença de alta densidade de mosquitos vetores, agente etiológico e população suscetível); geográficos (altos índices de pluviosidade, amplitude da malha hídrica e a cobertura vegetal); ecológicos (desmatamentos, construção de hidroelétricas, estradas e de sistemas de irrigação, açudes); e sociais (presença de numerosos grupos populacionais, morando em habitações com ausência completa ou parcial de paredes laterais e trabalhando próximo ou dentro das matas. (BRASIL, 2010)

Na região extra-amazônica onde são notificados apenas 1% do total de casos de malária no Brasil, cerca de 92% são importados dos estados da área endêmica e países da África. Mesmo nas áreas sem registro de casos de malária, a existência do vetor torna-a vulnerável a transmissão quando da presença de um homem infectado e portador de gametófitos, o que explica o significativo número de novos focos de transmissão de malária em área extra-amazônica registrados nos últimos anos.

## 6.2. INVESTIMENTOS EM MALÁRIA

A malária é um dos grandes problemas de saúde pública no Brasil e no mundo, são prevalentes em regiões de baixo poder aquisitivo e por esta particularidade não representa um atrativo econômico para que as indústrias farmacêuticas e biotecnológicas invistam em pesquisa e desenvolvimento de novos medicamentos, vacinas e reagentes de diagnósticos.

De acordo com Mahoney e Morel (2006), isto gera o que se conhece como “falhas” – “falha de ciência”, “falha de mercado” e “falhas dos sistemas e serviços de saúde”.

**Quadro 12: Falhas da Ciência, mercado e sistemas de saúde relacionadas às alternativas e soluções propostas**

<b>Problemas</b>	<b>Alternativas/Soluções</b>
Falhas de ciência (medicamentos inexistentes devido a conhecimento técnico-científico insuficiente, como vacinas contra malária, HIV/AIDS).	Estimular a pesquisa fundamental e induzir o desenvolvimento tecnológico.
Falhas de mercado (medicamentos caros, fora do alcance das populações).	Políticas de redução de preços (ex: licenciamento compulsório; negociações governos-companhias farmacêuticas; importações paralelas) ou de financiamento de medicamentos e vacinas existentes.
Falhas de sistemas de saúde (medicamentos baratos ou mesmo grátis não chegam aos	Melhorar os sistemas de logística e infraestrutura; combater a ineficiência, a corrupção, lutar contra

pacientes).	barreiras culturais ou religiosas que impedem o acesso destes bens pelas populações mais necessitadas.
-------------	--

Fonte: Mahoney; Morel, (2006).

Corroborando com a questão das falhas, Goodlee *et al.* (2004) apontam que apenas 10% das pesquisas em saúde, são direcionadas aos problemas de saúde de 90% das doenças globais o que os autores denominam como “Gap de conhecimento”.

Pode-se dizer que algumas doenças são muitas vezes negligenciadas na definição das agendas de saúde e nos orçamentos, já que os países onde elas ocorrem possuem recursos limitados para investir em saúde. Tornam-se, portanto, essenciais e prioritários a pesquisa e o desenvolvimento de novas ou melhores intervenções e sua incorporação às políticas e ações de saúde. (Morel, 2004, p. 263).

Em 2004, foi estabelecida a Agenda Nacional de Prioridade de Pesquisa em Saúde (ANPPS) como um dos alvos estratégicos da reformulação do papel do Ministério da Saúde no ordenamento do esforço nacional de pesquisa em saúde. O item 19 da agenda contempla as doenças transmissíveis onde se inclui a malária. Inicia-se desta forma a preocupação no combate as doenças negligenciadas no Brasil.

Com a inclusão da malária como prioridade de pesquisa, o Ministério da Saúde passou a promover o financiamento de projetos de pesquisa por meio do lançamento de editais temáticos. Além disso, passou a estimular o investimento das fundações de amparo a pesquisas estaduais, secretarias estaduais de Saúde, assim como secretarias de ciência e tecnologia estaduais e suas participações como gestores.

No âmbito internacional, destaca-se o Fundo Global de Luta contra AIDS, Tuberculose e Malária criado em 2002 como uma Parceria Pública Privada (PPP) dedicada à captação de recursos para a prevenção e tratamento dessas doenças com a finalidade de salvar vidas. A PPP congrega governos, sociedade civil, setor privado e comunidades afetadas, representa uma nova abordagem ao financiamento internacional da saúde. O Fundo Global trabalha em colaboração com outras organizações bilaterais e multilaterais para complementar os esforços já existentes em relação às três doenças.

Em quase 10 anos de existência tornou-se o grande financiador internacional de programas de saúde, o financiamento para a detecção e tratamento de 7,7 milhões de casos de tuberculose, a distribuição de 160 milhões de mosquiteiros tratados com inseticida e a manutenção de mais da metade de todos anti-retrovirais usados em tratamento de cerca de 3 milhões de pessoas no mundo.

Além disso, os programas apoiados pelo Fundo Global têm fornecido tratamento preventivo para a transmissão do HIV de mãe para filho a mais de 1 milhão de mulheres, distribuindo 2.700 milhões de preservativos, e 31 milhões de “sprays” para prevenir a malária dentro das casas.

O World Malaria Reports (2009; 2010), divulgados pela *World Health Organization* (WHO) registram que os fundos comprometidos com o controle da malária a partir de fontes internacionais têm aumentado consistentemente entre 2003 e 2009. Os fundos passaram de US\$ 300 milhões em 2003 para US\$ 1,7 bilhão em 2009. Em 2010 os fundos ficaram em US\$ 1,8 bilhão, porém muito aquém dos recursos necessários para atingir as metas globais estimadas em mais de US\$ 6 bilhões para o ano de 2010.

O aumento dos financiamentos ao combate da malária, tanto no nível nacional quanto internacional, associados aos avanços biotecnológicos e genômicos e a melhora do poder aquisitivo da população dos países em desenvolvimento estimulou empresas e institutos de pesquisa a desenvolverem estudos/produtos relacionados à prevenção/detecção da malária. Como consequência, a quantidade e variedade de testes para diagnóstico da doença disponíveis no mercado tem crescido rapidamente ao longo dos anos, o que tem dificultado a supervisão regulatória e a seleção da qualidade dos testes diagnósticos pelas agências de fomento.

### 6.3. DIAGNÓSTICO DA MALÁRIA

O Diagnóstico rápido e preciso é fundamental para a gestão eficaz da malária. Noppadon *et al.* (2009) apontam que o impacto global da malária tem estimulado interesse em desenvolver estratégias eficazes de diagnóstico, não só para áreas de recursos limitados onde a malária é um fardo substancial em sociedade, mas também nos países desenvolvidos, onde a experiência no diagnóstico da malária é muitas vezes inexistente.

“O diagnóstico clínico da malária é tradicional entre os médicos. Este método é menos dispendioso e mais amplamente praticado” (NOPPADON *et al.*, 2009, p. 94). Para Looareesuwan (1999) o diagnóstico clínico é baseado nos sinais dos pacientes e sintomas além dos achados físicos em exame. Os primeiros sintomas da malária são muito inespecíficos e variáveis, e incluem febre, dor de cabeça, fraqueza, mialgia, calafrios, tontura, dor abdominal dor, diarreia, náuseas, vômitos, anorexia e prurido.

Apesar dos sintomas relatados, o diagnóstico clínico da malária é insuficiente para o correto manejo da doença, que exige diagnóstico precoce e acurado, uma vez que a malária apresenta sintomas clínicos inespecíficos que são comuns a várias doenças infecciosas agudas. (OLIVEIRA, 2009 p. 43). Justifica-se assim, de acordo com a WHO (1995) a necessidade de métodos laboratoriais para confirmação. O diagnóstico definitivo da malária deverá ser baseado em critério clínico e confirmado sua parasitemia com diagnóstico laboratorial.

Arcanjo (2004) aponta que devido às necessidades de controle da malária, o diagnóstico parasitológico é de grande importância para avaliar a eficiência e eficácia dos programas ou medidas de controle, bem como no estabelecimento da terapêutica e evolução das manifestações clínicas, principalmente em áreas rurais e remotas, de alta endemicidade.

O diagnóstico da malária é realizado por meio de diferentes técnicas baseadas em diagnóstico laboratorial e molecular, dentre elas estão a microscopia ótica, QBC (*quantitative buffy coat*), os testes rápidos (RDTs), os testes sorológicos e os métodos de diagnóstico moleculares.

## 7. OBJETIVO GERAL

Identificar a melhor forma possível de elaborar uma estratégia de busca

### 7.1. OBJETIVOS ESPECÍFICOS

- Analisar os documentos de patente quanto à relevância para malária;
- Analisar os documentos de patente quanto ao escopo de proteção;
- Analisar termos utilizados e sua localização na estratégia de busca;
- Analisar o uso da linguagem natural e controlada na estratégia de busca;
- Analisar o uso da truncagem na estratégia de busca.

## 8. CAMINHO DA PESQUISA

A metodologia desenvolvida ao longo da pesquisa para a dissertação baseou-se em princípios bibliométricos para elaborar uma metodologia de busca sobre reagentes e *kits* de diagnóstico para malária em documentos de patente. Trata-se então de estudo exploratório que visou identificar a melhor forma possível de elaborar uma estratégia de busca.

Conforme visto nos capítulos anteriores, é de suma importância o entendimento dos conceitos que norteiam a pesquisa. Outro ponto de grande relevância é a estratégia de busca que, em princípio, é determinada pela natureza da base de dados a ser acessada e pela sua respectiva estrutura de informação, isto é, pela formatação de seus campos de identificação do documento e dos campos de identificação do conteúdo temático do mesmo.

Com base nas informações apresentadas e no escopo da pesquisa que necessita de dados internacionais, optou-se pelo uso da base de dados *Derwent Innovation Index* a qual será mais detalhada a seguir.

### 8.1. DERWENT INNOVATION INDEX

A *Derwent Innovation Index* é uma base comercial, específica para patentes, elaborada pela Thomson Reuters<sup>12</sup> e disponibilizada para comunidade brasileira de pesquisa pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); sendo acessada apenas por computadores internos de das universidades e centros de pesquisa que possuem acesso ao Portal Capes. É composta pelo *Derwent World Patents Index*®, *Patents Citation Index*<sup>TM</sup> e *Chemistry Resource* (assinatura separada), e apresenta mais de 16 milhões de invenções práticas, desde 1963 até os dias de hoje. A informação de patente são coletadas em mais de 40 autoridades emissoras de patente em todo o mundo e são classificadas em três categorias ou seções; *Chemical*, *Engineering* e *Electrical and Electronic* (Química, Engenharia e Elétrico e Eletrônico).

---

<sup>12</sup> A Thomson Reuters apresenta-se como umas bases de dados, no campo científico uma das mais conceituadas é a Web Of Science para recuperação de publicações.

Para a elaboração da estratégia de busca faz-se necessário um maior conhecimento da base, dos campos que podem ser pesquisados e como a base apresenta o resultado das buscas realizadas. A seguir, apresentam-se as características mais detalhadas da base escolhida para busca.

A busca na base *Derwent* pode ser realizada de três formas: busca básica, busca por patentes citadas e busca avançada.

A tela de “pesquisa” básica, figura 11, permite buscar um termo em três ou mais diferentes “campos”, ou tipos de informação, pré-estabelecido no *menu* suspenso. Este *menu* apresenta dez tipos de busca. São eles: padrão, título, inventor, número da patente, nome único de depositante, depositante, classificação internacional de patentes; código de classificação *derwent*, código manual *derwent*, número de primeiro acesso *derwent*.

O padrão da base é a utilização do *menu* suspenso “*Topic*” para “pesquisa” por assunto com o uso de palavras-chave encontradas no título e no resumo. Os títulos e resumos são reescritos, em inglês, por indexadores da *Thomson Scientific* e servem para tornar compreensível e simplificar a busca de patentes. O resumo é conciso, preciso e relevante, cobrindo o escopo mais amplo possível da invenção, conforme declarada na reivindicação principal e pode compreender categorias como Inovação, descrição detalhada, uso e vantagem, dentre outras, dependendo da patente selecionada. Para alguns registros, dispõe-se também de resumo equivalente e resumos de tecnologia/extensão.

A busca por inventor, feita utilizando o *menu* suspenso “*Inventor*”, busca todos os inventores do documento de patente. A busca pode ser feita com o uso do sobrenome e as iniciais do inventor.

O campo número da patente, “pesquisável” pelo *menu* suspenso “*Patent Number*” contém todos os números de patente pertencentes a uma mesma família.

O depositante, “pesquisado” pelo *menu* suspenso “*Assignee – Name Only*” identifica todos os depositantes, já o *menu* suspenso “*Assignee*” apresenta o nome e o código do depositante da patente. Os códigos de depositante são utilizados pelos indexadores da *Thomson Scientific* para unificar nomes ambíguos de consignatários.

A CIP é “pesquisada” pelo *menu* suspenso “*Int. Patent Classification*”.

O código de classe *Derwent* “pesquisado” no *menu* suspenso “*Derwent Class Code*” é uma classificação de patente própria do *Derwent* similar à classificação internacional de patentes. Esta classificação é atribuída por indexadores e dividida

nas seções *Chemical Engineering* e *Electrical and Electronic* (Química, Engenharia e Elétrica e Eletrônica).

Pelo *menu* suspenso “*Derwent Manual Code*”, é possível buscar os códigos estabelecidos pelos indexadores da *Derwent* para indicar os aspectos técnicos inovadores e as aplicações de uma invenção.

Por último, pelo *menu* suspenso “*Derwent Prim. Access N.º.*”, pode-se buscar o número de adesão primário *Derwent* que é um número único de identificação atribuído pelo *Derwent* a cada documento. O número é formado pelo ano de publicação, número de seis dígitos de série e o número de atualização de dois dígitos que indica quando *Derwent* publicou o resumo da patente.

The screenshot displays the 'Derwent Innovations Index' search page. At the top, there's a navigation bar with 'All Databases', 'Select a Database', and 'Derwent Innovations Index'. Below this, there's a search bar with a dropdown menu for 'Topic' and a list of search criteria including 'Title', 'Inventor', 'Patent Number', 'Int. Patent Classification', 'Derwent Class Code', 'Derwent Manual Code', 'Derwent Prim. Access. No.', 'Assignee - Name Only', and 'Assignee'. The interface includes a 'Search' button and a 'Clear' button. Below the search bar, there are 'Current Limits' for 'Timespan' (All Years, updated 2011-10-15) and 'Citation Databases' (Chemical Section, Electrical and Electronic Section, Engineering Section). The page also features a sidebar with 'Maintenance Alert', 'Support, Tools, Tips', and 'My ResearcherID'.

Figura 9: Tela de busca básica na base Derwent.

Finalizada a apresentação da forma de busca básica apresenta-se a busca por patentes citadas (figura 12). A base indexa a família das patentes citadas por examinadores de seis autoridades emissoras de patentes: Estados Unidos, Japão, PCT (WIPO), EPO, Alemanha e Grã-Bretanha. As buscas podem ser realizadas por: Número da patente citada; depositante citado; inventor citado e número de adesão primário *Derwent*.



**WEB OF KNOWLEDGE<sup>SM</sup>** DISCOVERY STARTS HERE **THOMSON REUTERS**

Go to mobile site | Sign In | Marked List (0) | My EndNote Web | My ResearcherID | My Citation Alerts | My Saved Searches | Log Out | Help

**All Databases** | **Select a Database** | **Derwent Innovations Index** | **Additional Resources**

Search | Cited Patent Search | Advanced Search | Search History | Compound Marked List (0)

### Derwent Innovations Index<sup>SM</sup>

**Cited Patent Search.** (Find the patents that cite a patent or patents) View our Cited Patent Search tutorial.

Enter the patent number, assignee, inventor, and/or accession number. Fields are combined with the Boolean AND operator.

in 
  
 Example: EP797246 or US5723945-A

in 
  
 Example: XEROX CORP or XERO

in 
  
 Example: Von Oepin R

[Add Another Field >>](#)

Searches must be in English

---

Current Limits: (To save these permanently, sign in or register.)

- Timespan**
  - All Years (updated 2011-10-15)
  - From 1963-66 to 2011 (default is all years)
- Citation Databases: Chemical Section; Electrical and Electronic Section; Engineering Section
- Adjust your results settings

View in: [简体中文](#) | [English](#) | [日本語](#)

© 2011 Thomson Reuters | [Acceptable Use Policy](#) | [Please give us your feedback on using Web of Knowledge](#)

Figura 10: Tela de busca por patente citada.

A busca avançada, figura 13, usa duas ou mais etiquetas (*tags*), operadores booleanos, parênteses e referências conjunto para criar a sua consulta. Os campos das tags são preestabelecidos pela base como códigos:

**WEB OF KNOWLEDGE<sup>SM</sup>** DISCOVERY STARTS HERE **THOMSON REUTERS**

Sign In | Marked List (0) | My EndNote Web | My ResearcherID | My Citation Alerts | My Saved Searches | Log Out | Help

**All Databases** | **Select a Database** | **Derwent Innovations Index** | **Additional Resources**

Search | Cited Patent Search | Advanced Search | Search History | Compound Marked List (0)

### Derwent Innovations Index<sup>SM</sup>

**Advanced Search**

Use 2-character tags, Boolean operators, parentheses, and set references to create your query. Results appear in the Search History at the bottom of the page.

Example: TS=(nanotub\* SAME carbon) NOT AU=Smalley RE #1 NOT #2 [more examples](#) | [view the tutorial](#)

Searches must be in English

---

Current Limits: (To save these permanently, sign in or register.)

- Timespan**
  - All Years (updated 2011-10-15)
  - From 1963-66 to 2011 (default is all years)
- Citation Databases: Chemical Section; Electrical and Electronic Section; Engineering Section
- Adjust your results settings

---

**Search History**

Set	Results
Open Saved History	

**There are no search sets to display.**  
Use the search options to create new search sets.

**Booleans: AND, OR, NOT, SAME**

**Field Tags:**

TS=Topic	CP=Cited Patent Number
TI=Title	CX=CP + Family
AU=Inventor	CA=Cited Assignee
PN=Patent Number	CN=Cited Assignee Name
IP=Int. Patent Classification	CPC=Cited Assignee Code
DC=Derwent Class Code	CAU=Cited Inventor
MAN=Derwent Manual Code	CD=Cited PAN
PAN=Derwent Prim. Access. No.	
AN=Assignee Name	
AC=Assignee Code	
AE=Assignee Name + Code	

**Combine Sets**  AND  OR

Figura 11: Tela de busca avançada da base Derwent.

**Quadro 13: Códigos de "tags" da Derwent**

Campos das etiquetas (tags):	
<b>TS</b> =Padrão	<b>CP</b> =Número de Patente citada
<b>TI</b> =Título	<b>CX</b> = Número de Patente citada e Família
<b>AU</b> =Inventor	<b>CA</b> =Depositante citado
<b>PN</b> =Número da Patente	<b>CN</b> =Nome de Depositante citado
<b>IP</b> =Classificação Internacional de Patente	<b>CC</b> =Código de Depositante citado
<b>DC</b> =Código de Classe Derwent	<b>CI</b> =Inventor citado
<b>MC</b> =Código Manual Derwent	<b>CD</b> = Número de Adesão Primário <i>Derwent</i>
<b>GA</b> = Número de Adesão Primário <i>Derwent</i>	citado
<b>AN</b> =Nome do Depositante	
<b>AC</b> =Código do Depositante	
<b>AE</b> =Nome e Código do Depositante	

Fonte: Derwent

Os resultados aparecem no histórico de pesquisa na parte inferior da página, por exemplo, #1, #2 e #3. Os resultados podem ser combinados entre operadores booleanos ao final.

A base ainda apresenta como característica nas três formas de busca a possibilidade de refino e ampliação dos resultados. Uma das formas de refino da busca na base é a utilização dos operadores booleanos *AND*, *OR*, *NOT* e *SAME*. O operador *AND* se usa para encontrar registros que contenham *todos* os termos separados pelo operador; *OR* é usado para encontrar registros que contenham qualquer um dos termos separados pelo operador; *NOT* é usado para excluir registros que contenham determinadas palavras na pesquisa e *SAME* é usado para encontrar registros onde os termos separados pelo operador aparecem na mesma frase. Com o uso desses operadores, é possível fazer busca por exemplo, pelo depositante de patentes em associação com os inventores; buscar por uma determinada palavra-chave, associada com uma determinada CIP.

Outra forma de refino é a utilização de escala temporal, ou seja, é possível buscar patentes em um ano específico ou um bloco de anos pelo campo "*Timespan*". Além dessas formas é possível especificar a busca por seção: Chemical, Engineering e Electrical and Electronic sendo possível selecionar uma, duas ou as três seções.

Como forma de ampliar a busca, muito comum em bases de dados, é a utilização de símbolos de truncamentos. Os símbolos são usados para recuperar

plurais e grafias variantes. Nesta base, os símbolos permitidos são: \* para muitos caracteres; ? para um caractere e \$ para um ou mais caracteres. Por exemplo: *diagno\** retornaria as palavras *diagnose, diagnosis, diagnostic, diagnostics* dentre outros.

## 8.2. FORMULAÇÃO DA ESTRATÉGIA DE BUSCA E COLETA

### 8.2.1. Etapa 01

A estratégia da busca, executada em 06 de abril de 2011, foi montada para ser realizada na “pesquisa avançada” da base Derwent, no campo padrão (“*Topic*”), utilizando termos que representassem os conceitos diagnóstico e malária. Pensou-se em refinar a busca interligando *plasmodium, anopheles* e *malaria* com os conceitos de *kit* (truncado), reagente (truncado), *diagnosis, diagnostic* (truncado) e a classificação internacional de patentes G01N (truncada), mencionada ao longo do capítulo de patentes e indicada em algumas publicações sobre diagnóstico. Este código da classificação compreende investigação ou análise dos materiais pela determinação de suas propriedades químicas ou físicas. Vale a pena lembrar que os termos são vertidos para o inglês, idioma utilizado pela maioria das bases de dados internacionais de patentes. A escala temporal definida foi a de recuperar documentos entre 1989 e 2011, pois a interface da base Derwent, disponibilizada pelo Portal Capes permite a consulta segundo a data de inserção da informação na base e não pela data de depósito do documento de patente. Neste sentido, justificase esta escala temporal para garantir a cobertura de 20 anos que é o tempo de duração de uma patente. A limpeza dos dados que permitiu fixar o período como 1990 a 2010 foi feita com o uso do software de mineração de dados.

The screenshot shows the Thomson Reuters Web of Knowledge interface. At the top, it says "WEB OF KNOWLEDGE™ DISCOVERY STARTS HERE" and "THOMSON REUTERS". Below this, there are navigation tabs for "All Databases", "Select a Database", "Derwent Innovations Index", and "Additional Resources". The "Derwent Innovations Index" tab is selected, and the "Advanced Search" option is active. A search bar is present with a "Search" button. Below the search bar, there are "Current Limits" for "Timespan" (All Years, updated 2011-10-15) and "Citation Databases" (Chemical Section, Electrical and Electronic Section, Engineering Section). A "Search History" table is visible at the bottom, showing three sets of results:

Set	Results	Query
# 3	3,107	#2 AND #1 Databases=CDerwent, EDerwent, MDerwent Timespan=1989-2011
# 2	>100,000	TS=(kit* AND reagent* AND diagnosis AND diagnostic*) OR IP=(G01N*) Databases=CDerwent, EDerwent, MDerwent Timespan=1989-2011
# 1	6,392	TS=(plasmodium OR anopheles OR malaria) Databases=CDerwent, EDerwent, MDerwent Timespan=1989-2011

Figura 12: Resultado da etapa 01 da busca.

O resultado obtido nesta estratégia inicial, 3.107 documentos, não foi satisfatório por alguns motivos: a utilização da palavra referente ao nome do vetor fez com que se recuperasse uma grande quantidade de documentos para a eliminação do vetor, no caso um pesticida. O uso do termo *kit* (truncado) recuperou uma grande quantidade de documentos que não representavam um *kit*, além disso, *kit* é considerado um conjunto de utensílios que não necessariamente é para diagnóstico, o mesmo fato se apresentou para o termo reagente.

Este procedimento inicial é relatado na literatura como comumente utilizado pelos buscadores de patentes (BHAVNANI; CLARKSON; SCHOLL, 2011) e visa exatamente “aprender” com os erros para refazer a busca.

### 8.2.2. Etapa 02

Após os erros encontrados, a nova escolha dos termos e da estratégia de buscas foi definida através de uma amostragem dos termos encontrados nos próprios documentos de patentes pertinentes ao assunto. Alguns autores chamam a este procedimento de “*snow ball*”.

A nova estratégia da busca, executada em 28 de maio de 2011, também montada na “pesquisa avançada” da base Derwent, no campo padrão (“*Topic*”), teve como princípio manter a escala temporal e o cruzamento de termos relacionados à malária com os termos relacionados ao diagnóstico, acrescidos da classificação internacional que melhor lhe representa, porém com algumas diferenças com relação à etapa 01.

Na primeira etapa da estratégia, que continha os termos relativos à malária, o termo *plasmodium* foi mantido, porém foi truncado já que foram encontrados documentos com erro de escrita por juntar o gênero e espécie *plasmodiumfalciparum* ao invés de *Plasmodium falciparum*. O termo *anopheles* foi descartado, por se referir ao vetor, conforme explicado na etapa 01. O termo *malaria* foi mantido, porém também foi truncado por ter sido encontrado documentos com erros de escrita. Outros termos foram adicionados como as espécies de *plasmodium* pensando na possibilidade do gênero estar abreviado, como por exemplo, *P. falciparum*. Desta forma *falciparum*, *vivax*, *ovale* e *knowlesi* foram adicionados à estratégia, a espécie *malariae* entrou na estratégia pela truncagem do termo *malaria* (*malaria\**).

Na segunda etapa da estratégia, referente aos termos relativos ao diagnóstico, os termos *kit* e *reagent* foram removidos, os termos *diagnosis* e *diagnostic* (truncado) foram mantidos e o termo *diagnosing* foi acrescido. Para aumentar a qualidade e relevância dos dados, a classificação internacional de patentes sofreu uma pequena alteração. Com base nos documentos de patente identificados e em uma melhor análise da classificação observou-se que os documentos foram classificados no grupo 33 da classificação G01N. Este grupo se refere à Investigação ou análise de materiais por métodos específicos não abrangidos pelos grupos G01N 1/00-G01N 31/00. Dentre os subgrupos pertencentes a este grupo estão a investigação e análise de material biológico, por ex., sangue, urina e outros.

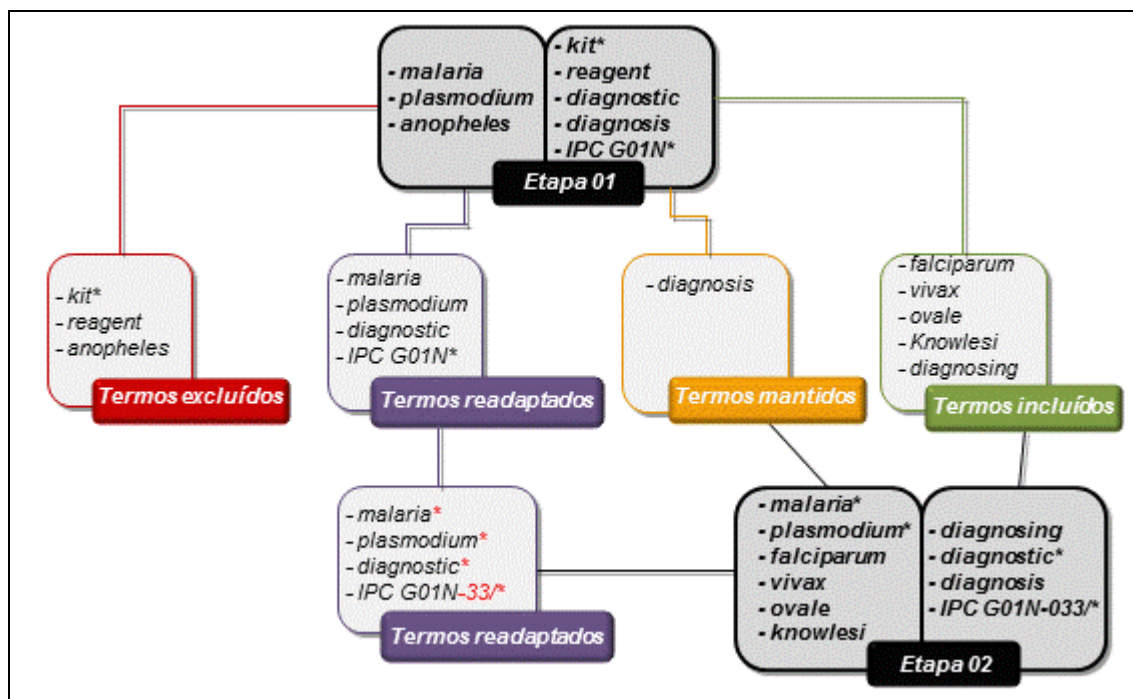


Figura 13: Fluxograma representativo das etapas 01 e 02

Neste sentido a estratégia final ficou formulada da seguinte forma:

**Quadro 14: Estratégia de Busca formulada**

#1	TS=( <i>malaria*</i> OR <i>plasmodium*</i> OR <i>falciparum</i> OR <i>vivax</i> OR <i>ovale</i> OR <i>knowlesi</i> )
#2	TS=( <i>diagnostic*</i> OR <i>diagnosing</i> OR <i>diagnosis</i> ) OR IP=(G01N-033/*)
<b>#1 AND #2</b>	

Fonte: elaboração própria

The screenshot shows the Thomson Reuters Web of Knowledge interface. At the top, it says "WEB OF KNOWLEDGE™ DISCOVERY STARTS HERE" and "THOMSON REUTERS". Below this are navigation links like "Sign In", "Marked List (0)", "My EndNote Web", "My ResearcherID", "My Citation Alerts", "My Saved Searches", "Log Out", and "Help".

The main content area is titled "Derwent Innovations Index™" and "Advanced Search". It includes a search bar with a "Search" button and a note: "Searches must be in English". Below the search bar are "Current Limits" for "Timespan" (All Years, updated 2011-10-15) and "Citation Databases" (Chemical Section, Electrical and Electronic Section, Engineering Section). There are also links for "Save History / Create Alert" and "Open Saved History".

The "Search History" table is as follows:

Set	Results		Combine Sets	Delete Sets
# 3	1.182	#2 AND #1 Databases=CDerwent, EDerwent, MDerwent Timespan=1989-2011	<input type="checkbox"/>	<input type="checkbox"/>
# 2	>100.000	T\$(diagnostic* OR diagnosing OR diagnosis) OR IP=(G01N-033P) Databases=CDerwent, EDerwent, MDerwent Timespan=1989-2011	<input type="checkbox"/>	<input type="checkbox"/>
# 1	6.431	T\$(plasmodium* OR malaria* OR falciparum OR vivax OR knowlesi) Databases=CDerwent, EDerwent, MDerwent Timespan=1989-2011	<input type="checkbox"/>	<input type="checkbox"/>

At the bottom right of the search history table, there are options for "Combine Sets" (AND, OR) and "Delete Sets" (Select All, Delete).

Figura 14: Resultado da etapa 02 da busca.

O resultado da busca, 1.182 documentos, foi descarregado de 500 em 500 documentos, por limitação da base, como registro completo no formato texto plano (txt) para ser tratado e analisado pela ferramenta de mineração de texto denominada *Vantage Point*.

### 8.3. TRATAMENTO

O *Vantage Point* é um software de mineração de dados desenvolvido pelo *Georgia Institute of Technology* e comercializado pela *SearchTechnology*, cuja licença acadêmica o Instituto de Comunicação e Informação Científica e Tecnológica (Icict) da Fundação Oswaldo Cruz dispõe. O *software* permite a extração de dados de bases bibliográficas e textuais, gera listas, matrizes e mapas de conhecimento.

Os arquivos descarregados, no formato txt, da base *Derwent* foram importados para o *Vantage Point* pelo uso de um filtro específico para a base disponível na ferramenta. Neste procedimento os campos são extraídos do texto usando a correspondência de padrões, baseado em regras e técnicas de processamento de LN.

Após a importação dos dados, foi realizado um procedimento para uniformização dos dados através de uma macro denominada “*list cleanup*”. O software usa técnicas de correspondência difusa para identificar, associar e limpar os dados. Dentre as possibilidades estão a associação de nomes de pessoas e empresas, limpeza de erros ortográficos, hifenização para proporcionar ao usuário uma maior qualidade de dados que é essencial para uma boa análise.

Outro tratamento automatizado realizado nos dados foi a identificação do ano de prioridade. A indexação do documento na base *Derwent* é feita pela data de publicação na base e não a data de depósito da patente. Além disso, o campo destinado ao depósito prioritário não é único, apresenta todos os dados de prioridade da patente. Para a solução desse problema foi utilizado uma macro disponível no *Vantage Point* que verifica o ano mais antigo dentre um conjunto de anos em um campo. Desta forma, pode-se criar um novo campo somente com o ano de prioridade, removendo os demais anos disponíveis no campo originalmente.

#### 8.4. ANÁLISE

Conforme informado na parte de formulação da estratégia de busca e coleta, a escala temporal utilizada para a busca foi de 20 anos, porém para a etapa de análise, foi feito um corte temporal com base na obsolescência e vida média de uma tecnologia patenteada que segundo os estudos de Chen *et al.* (2009) é de cinco anos.

Poder-se-ia tentar validar este cálculo, visto que o dado informado não aponta isto para todas as áreas de conhecimento, mas não era propósito deste trabalho calcular a vida média e sim este dado foi utilizado para corte temporal, visto o grande número de documentos encontrados na busca, 1.182 documentos.

Com base no corte temporal foi criado no *Vantage Point* um grupo com base no ano de prioridade do documento de patente de 2005 a 2009 na sua integralidade, resultando um total de 516 documentos.

Como forma de estabelecer uma estratégia para busca de informação de patente para monitoramento, foi realizada a leitura do título e resumo (que inclui campos como novidade, uso, vantagens e descrição detalhada) dos documentos de patente.



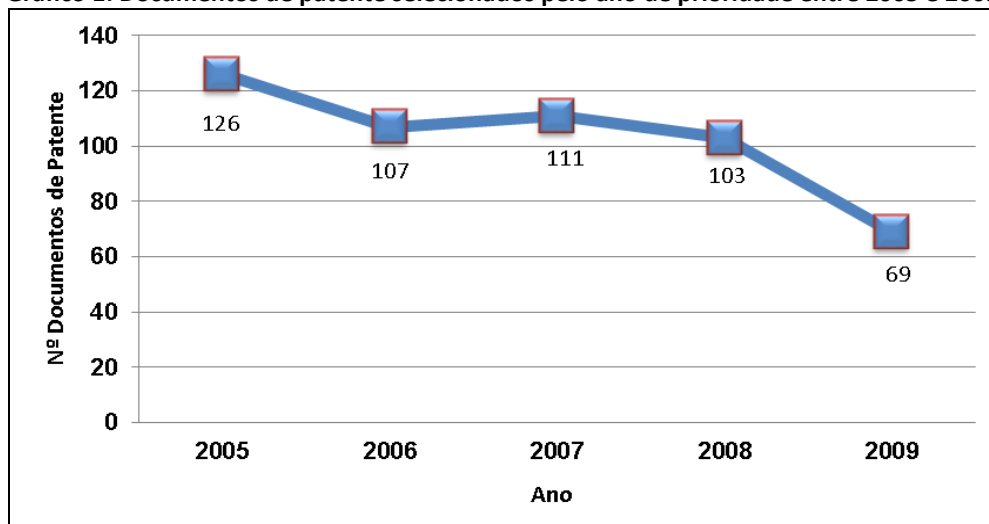
Os critérios para análise dos documentos de patente foram definidos através de uma amostragem dos próprios documentos de patentes pertinentes ao assunto. O primeiro critério definido se refere à relevância para malária, se o documento é específico, se não é específico (utilidade para malária e para outras doenças) ou se não há relevância/aplicabilidade. O segundo critério se refere ao escopo de proteção da patente: tratamento, prevenção, prognóstico, diagnóstico e não categorizado (são os que não apresentaram relevância/aplicabilidade).

As demais análises foram realizadas com base nos documentos de patente que apresentaram alguma relação com diagnóstico da malária, são elas: análise das palavras-chave utilizadas, o uso da CIP, o uso da truncagem, análise da localização dos termos utilizados na estratégia de busca, ou seja, se estão no título, e/ou resumo e/ou foco tecnológico. Neste caso cabe ressaltar que a base *Derwent* em alguns documentos, além do resumo tradicional da base, apresenta um resumo equivalente denominado foco tecnológico. O software de mineração de dados diferencia o campo do resumo como (AB) e campo foco tecnológico (TF) por este motivo foi levado em consideração os campos título (TI), resumo (AB) e foco tecnológico (TF).

## 9. RESULTADOS E DISCUSSÃO

Neste capítulo, serão apresentados os resultados encontrados a partir da metodologia explicitada no capítulo anterior. De acordo com os critérios de busca apresentados na metodologia, foram identificados 1182 documentos de patentes no intervalo de 20 anos. A filtragem realizada pelo corte temporal de cinco anos referente ao ano de prioridade do documento de patente para fins de análise resultou na identificação de 516 documentos depositados no intervalo de janeiro de 2005 a dezembro de 2009 completos.

**Gráfico 1: Documentos de patente selecionados pelo ano de prioridade entre 2005 e 2009.**



A análise de conteúdo dos 516 documentos selecionados pelo recorte temporal, que visou identificar a utilidade para malária e o escopo de proteção do documento mostrou que 89 (17%) documentos são específicos para malária, ou seja, o escopo de proteção se refere exclusivamente para a malária; 325 (63%) documentos apresentam a malária de forma não específica, ou seja, o escopo de proteção se refere tanto para malária quanto para outras doenças e 102 (20%) dos documentos não apresentaram relevância/aplicabilidade para malária, ou seja, apesar de ter sido recuperado pela presença de uma palavra representativa do conceito malária, o escopo de proteção não se refere à malária (ver quadro 15, a seguir).

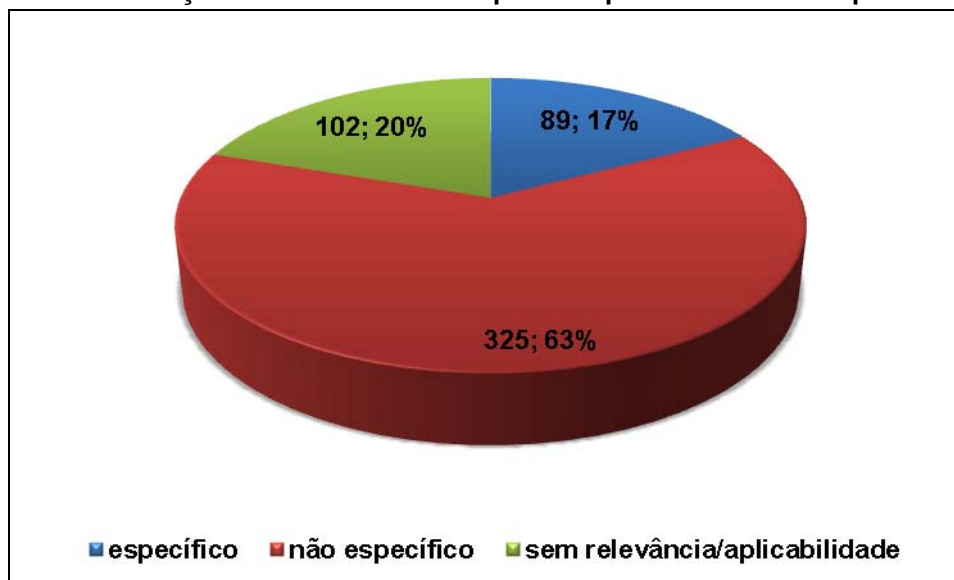
Estes documentos podem ser considerados ruído na recuperação e apresentavam alguma das características relacionadas no quadro a seguir. Algumas

destas ocorrências mereceriam análise mais aprofundada, a depender de um especialista.

**Quadro 15: Ruídos na recuperação da informação e suas características**

Ruídos	Características
Outras doenças	documento é para outra doença e a malária se apresenta como uma condição que não afeta o resultado.
Outras espécies de <i>Plasmodium</i>	documento contém o termo <i>plasmodium</i> , porém a espécie não está entre as cinco referentes à malária humana.
Termo ovale não referente à espécie do <i>Plasmodium</i>	documento contém o termo " <i>Patent Foramen Ovale</i> ", contém o termo ovale mas trata-se de um orifício no septo entre os dois átrios cardíacos direito e esquerdo.
Termo <i>Plasmodium</i> em outras finalidades	documento contém o termo <i>plasmodium</i> , mas refere-se à produtos derivados do <i>plasmodium</i> para outras finalidades.
	documento refere-se a peptídeo ou proteína produzida pelo <i>plasmodium</i> que não é de utilidade para malária.
Outras utilidades com uso do sangue	documento refere-se à esterilização de sangue onde a malária se apresenta como um exemplo.
	documento refere-se a método de obtenção de sangue para teste de glicose onde a malária pode estar presente no sangue.

**Gráfico 2: Avaliação dos documentos de patente quanto à relevância para malária**



Os 102 documentos de patente considerados sem relevância/aplicabilidade foram excluídos das análises posteriores.

A análise dos 414 documentos baseada no escopo de proteção com a atividade relativa à malária, ou seja, tratamento, prevenção, diagnóstico, prognóstico e suas combinações revelou que 257 documentos de patente apresentavam proteção com apenas uma das atividades identificadas: 123 como tratamento, 115 como diagnóstico, 19 como prevenção e 03 como prognóstico. O escopo de

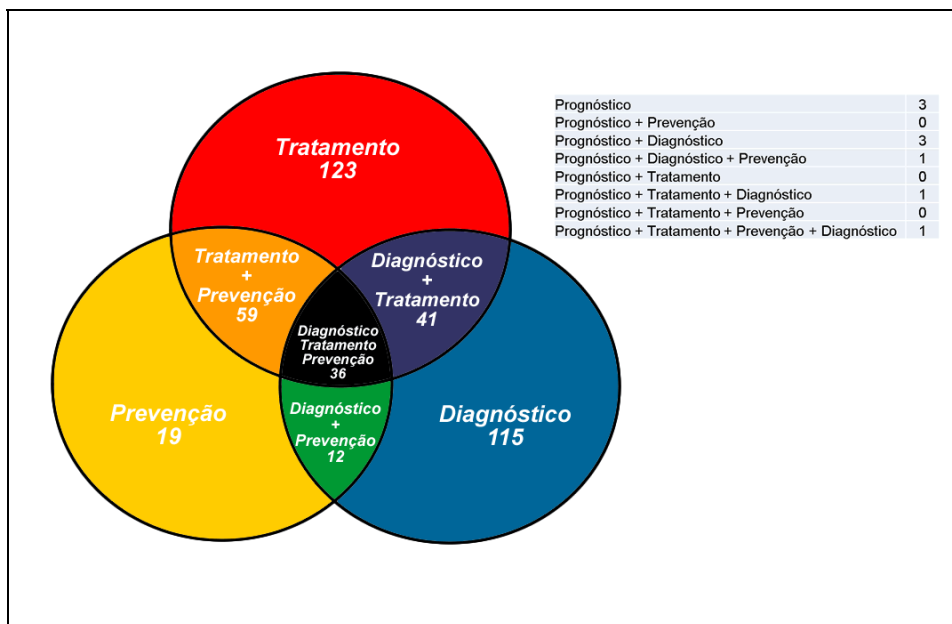
proteção dos 157 documentos restantes foram classificados com duas, três ou quatro combinações entre as categorias estabelecidas: 59 documentos como tratamento e prevenção; 41 como diagnóstico e tratamento; 12 como diagnóstico e prevenção; 03 como diagnóstico e prognóstico; 36 como diagnóstico, tratamento e prevenção; 01 como prognóstico, diagnóstico e prevenção; 01 prognóstico, diagnóstico e tratamento e por fim 01 documento como prognóstico, tratamento, prevenção e diagnóstico.

De acordo com os resultados apresentados nos gráficos 2 e 3, verifica-se que poucos foram os casos onde o documento de patente protege alguma especificidade (quanto à doença e/ou escopo de proteção), isto porque é muito comum que o escopo da proteção seja solicitado de forma ampla, sem especificidade. Existem várias razões para a ampliação do escopo de proteção. Dentre elas pode-se citar: ampliar a abrangência da barreira comercial de forma a impedir a entrada de produtos no mercado de concorrentes; maior possibilidade de licenciamento e, assim, de retorno através do recebimento de *royalties* associados às atividades de Pesquisa & Desenvolvimento; gastos com o registro e a colocação do produto no mercado.

Em suma, muitas patentes são requeridas sem que o depositante conheça o potencial comercial da invenção e se ela se constituirá num meio efetivo de bloquear a competição. Porém, estrategicamente, as indústrias buscam amplos escopos de proteção de modo que a patente crie um nicho de mercado e as proteja de novos entrantes.

Além disso, a ampliação do escopo de proteção é uma estratégia adotada pelos depositantes, já que, na avaliação pelos examinadores dos escritórios oficiais de propriedade intelectual, estes podem restringir o escopo de acordo com a legislação local de cada país.

**Gráfico 3: Avaliação dos documentos de patente quanto ao escopo de proteção com base na malária.**



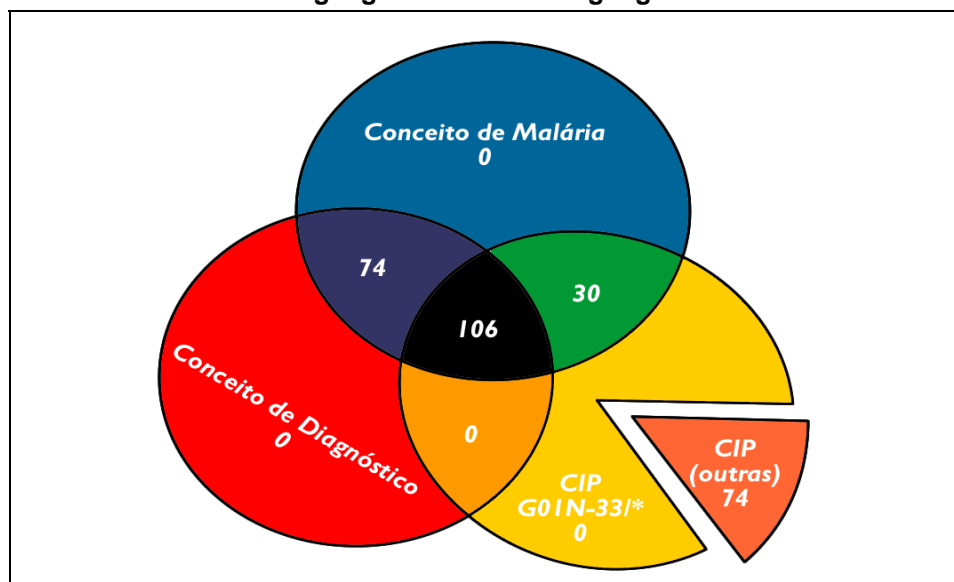
As demais análises foram realizadas com base no somatório dos 210 documentos classificados como diagnóstico, independente de estar associado à outra classificação.

A análise dos termos representativos do conceito malária utilizada na estratégia de busca #1 (*malaria\* OR plasmodium\* OR falciparum OR vivax OR ovale OR knowles*), em associação com o conceito de diagnóstico e a CIP G01N-33/\* utilizada na estratégia de busca #2 (*(diagnostic\* OR diagnosing OR diagnosis) OR IP=(G01N-033/\*)*) revelou que o conceito de malária estava presente em todos os 210 documentos de patente (100%) com alguma utilidade para o diagnóstico da malária. Já o conceito de diagnóstico estava presente em 180 (85,71%) dos documentos, o que significa que os 85,71% dos documentos foram recuperados pela associação dos conceitos de malária com o conceito de diagnóstico. Conseqüentemente, os 30 documentos restantes (14,29%) foram recuperados da associação entre o conceito de malária e a CIP G01N-33/\*.

Estes resultados podem ser visualizados no gráfico 4 e tabela 1, a seguir, e apontam que a recuperação por LN foi mais efetiva quanto ao retorno do número de documentos. Este fato corrobora o estudo de Markey, Aherton e Newton (1980), ao afirmarem que a busca com termos da LN pode, frequentemente, ser a melhor opção quando se deseja alto índice de retorno. Ainda sobre a LN, Lopes (2002) também afirma que pode ser o melhor caminho para o encontro da informação

desejada nas buscas relativas ao levantamento das últimas tecnologias em uma determinada área, ou a um novo assunto, ou a um novo produto e, ainda, na busca em documentos de patentes.

**Gráfico 4: Linguagem Natural X Linguagem Controlada**



**Tabela 1: Análise preliminar da estratégia de busca**

Estratégias de busca e combinações	nº docs. com	Nº dos sem
Conceito de malária	210	0
Conceito de diagnóstico	180	30
CIP G01N-033/*	136	74
Conceito de malária + conceito de diagnóstico	180	30
Conceito de malária + CIP G01N-033/*	30	180
Conceito de malária + conceito de diagnóstico + CIP G01N-033/	106	104
Conceito de malária + conceito de diagnóstico + outras classificações	74	136

Fonte: Elaboração própria.

Porém, a recuperação da informação por uso de LC, a CIP, obteve retorno de documentos que não foram identificados pelo uso somente da LN. Estes resultados vão ao encontro da conclusão dos estudos de Markey, Aherton e Newton (1980) que afirmaram que melhor desempenho da estratégia de busca é aquele que utiliza os dois métodos concomitantemente.

A análise da ocorrência das palavras utilizadas na estratégia de busca, sintetizada no quadro 16, foi feita com base na presença das palavras no título, resumo e no foco tecnológico e revelou que, de uma forma geral dentro da linha #1 da estratégia de busca, (*malaria\** OR *plasmodium\** OR *falciparum* OR *vivax* OR

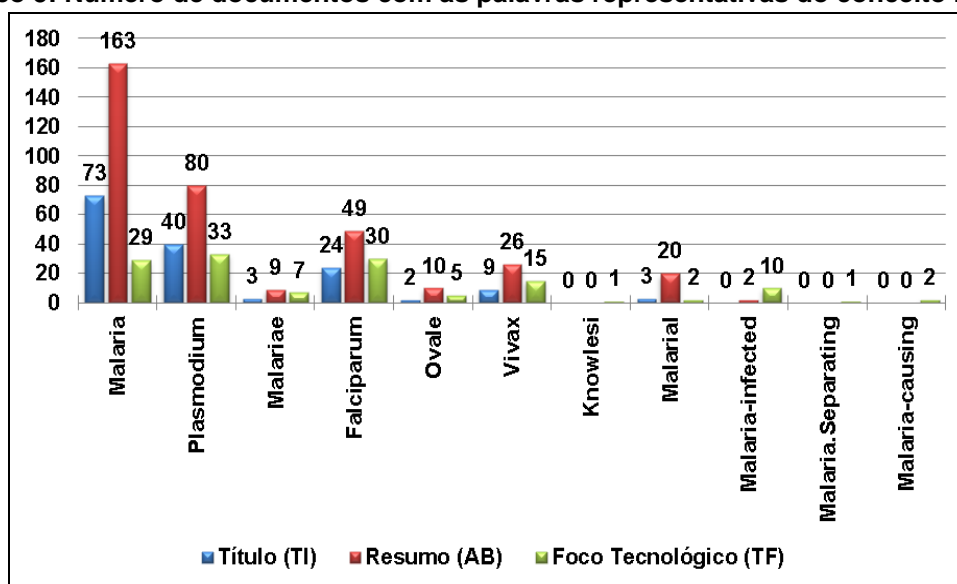
*ovale OR knowlesi*), obteve-se retorno de documentos com as seguintes palavras: *malaria; malariae; malarial; malaria-infected; malaria.separating; malaria-causing; plasmodium; falciparum; ovale; vivax; knowlesi*. Já na linha #2 da estratégia de busca, *((diagnostic\* OR diagnosing OR diagnosis) OR IP=(G01N-033/\*))*, obteve-se retorno de documentos com as palavras-chave *diagnostic; diagnostics; diagnosing e diagnosis*.

**Quadro 16: Análise da ocorrência de palavras utilizadas na estratégia de busca**

#1 ( <i>malaria* OR plasmodium* OR falciparum OR vivax OR ovale OR knowlesi</i> )	#2 ( <i>((diagnostic* OR diagnosing OR diagnosis) OR IP=(G01N-033/*))</i> )
<i>falciparum; knowlesi. malaria.separating; malaria; malaria-causing; malariae; malaria-infected; malarial; ovale; plasmodium; vivax;</i>	<i>diagnosing diagnosis. diagnostic; diagnostics;</i>

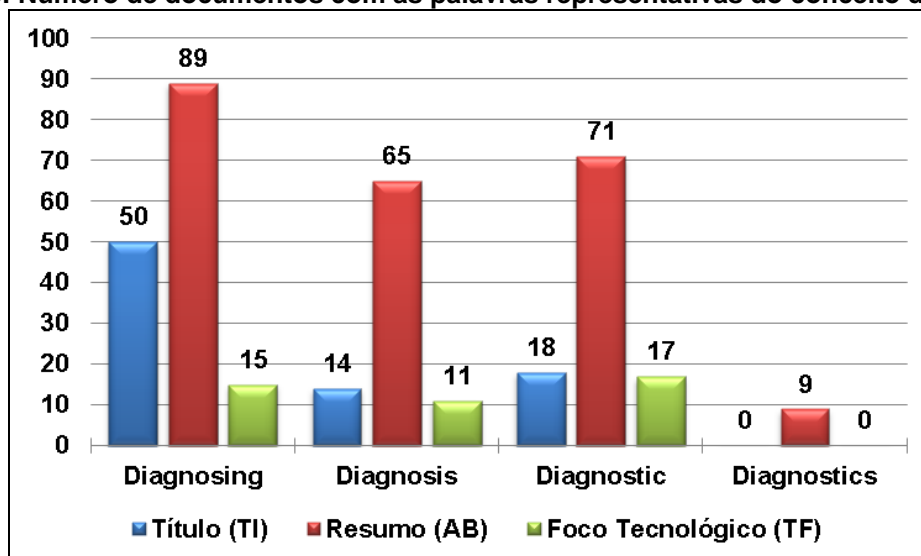
Com relação ao número de documentos contendo as palavras selecionadas no primeiro nível da estratégia representativas do conceito malária e visualizadas no gráfico 5, observou-se que a palavra *Plasmodium* esteve presente em 40 documentos no título, 80 no resumo e 33 no foco tecnológico; a espécie *falciparum* retornou em 24 documentos no título, 49 no resumo e 30 no foco tecnológico; a espécie *ovale* em 02 documentos no título, 10 no resumo e 07 no foco tecnológico; a espécie *vivax* em 09 documentos no título, 26 no resumo e 05 no foco tecnológico; a espécie *knowlesi* retornou apenas 01 documento no foco tecnológico; Malaria esteve presente em 73 documentos no título, 163 documentos no resumo e 29 documentos no foco tecnológico. Já as palavras truncadas *malariae, malarial, malaria.separting e malaria-causing* retornaram da seguinte forma: *malariae* esteve em 03 documentos no título, 09 no resumo e 07 no foco tecnológico; *malarial* em 03 documentos no título, 20 no resumo e 02 do foco tecnológico; *malaria-infected* em 02 documentos no resumo e 10 no foco tecnológico; *malaria.separating* em apenas 01 documento no foco tecnológico e *malaria-causing* com apenas 02 documentos.

Gráfico 5: Número de documentos com as palavras representativas do conceito malária



Quanto ao número de documentos contendo as palavras seleccionadas no segundo nível da estratégia referente ao diagnóstico, observou-se que 50 documentos apresentaram a palavra-chave *diagnosing* no título, 89 no resumo e 15 no foco tecnológico. Outra palavra identificada, *diagnosis*, retornou em 14 documentos no título, 65 documentos no resumo e 11 documentos no foco tecnológico. A palavra *diagnostic* retornou em 18 documentos no título, 71 no resumo e 17 no foco tecnológico. Por fim a truncagem da palavra *diagnostic*, *diagnostics*, retornou em 09 documentos somente no resumo. Estes resultados podem ser visualizados no gráfico 6, a seguir.

Gráfico 6: Número de documentos com as palavras representativas do conceito diagnóstico





Cabe afirmar que nos números encontrados foram identificadas correlações entre as palavras pesquisadas, ou seja, a presença de duas ou mais palavras no mesmo documento, o que impulsionou analisar de forma mais específica o retorno das palavras-chave pesquisadas nos 210 documentos selecionados quanto a sua localização específica.

O resultado desta análise é que a palavra *diagnosing* estava presente em 93 documentos, sendo 03 (3%) somente no título, 34 (37%) somente no resumo, 01 (1%) somente no foco tecnológico e nas combinações, 41 (44%) no título e resumo, 08 (9%) no resumo e foco tecnológico, 06 (6%) no título, resumo e foco tecnológico e nenhum no título e no foco tecnológico.

A palavra *diagnosis* estava presente em 70 documentos sendo 03 (4%) somente no título, 45 (64%) somente no resumo, 02 (3%) somente no foco tecnológico e nas combinações: 11 (16%) no título e resumo, 09 (13%) no resumo e foco tecnológico e nenhum no título e foco tecnológico e no título, resumo e foco tecnológico.

A palavra *diagnostic* estava presente em 73 documentos sendo 01 (1%) somente no título, 44 (60%) somente no resumo, 01 (1%) somente no foco tecnológico e nas combinações: 11 (15%) no título e resumo, 10 (14%) no resumo e foco tecnológico, 06 (8%) no título, resumo e foco tecnológico e nenhum no título e foco tecnológico.

A palavra *diagnostics* estava presente em 09 documentos sendo todas as 9 (100%) somente no resumo.

A palavra *malaria* estava presente em 169 documentos sendo 01 (1%) somente no título, 84 (50%) somente no resumo, 05 (3%) no foco tecnológico e nas combinações: 55 (32%) no título e resumo, 07 (4%) no resumo e foco tecnológico, 17 (10%) no título, resumo e foco tecnológico e nenhum no título e foco tecnológico.

A palavra *plasmodium* estava presente em 96 documentos sendo 02 (2%) somente no título, 36 (37%) somente no resumo, 14 (15%) no foco tecnológico e nas combinações: 25 (26%) no título e resumo, 06 (6%) no resumo e foco tecnológico, 13 (14%) no título, resumo e foco tecnológico e nenhum no título e foco tecnológico.

A palavra *malariae* estava presente em 12 documentos sendo 05 (42%) somente no resumo, 03 (25%) somente no foco tecnológico e nas combinações: 01 (8%) no resumo e foco tecnológico e 03 (25%) no título, resumo e foco tecnológico.

Esta palavra não foi encontrada no título, e nas combinações título e resumo e título e foco tecnológico.

A palavra *falciparum* estava presente em 64 documentos sendo 01 (1%) somente no título, 17 (27%) somente no resumo, 13 (20%) somente no foco tecnológico e nas combinações, 16 (25%) no título e resumo, 10 (16%) no resumo e foco tecnológico, 07 (11%) no título e no foco tecnológico e não houve retorno para título e foco tecnológico.

A palavra *ovale* estava presente em 11 documentos sendo 06 (55%) somente no resumo, 01 (9%) somente no foco tecnológico e nas combinações: 2 (18%) no resumo e foco tecnológico e 02 (18%) resumo e título e foco tecnológico. Esta palavra não foi encontrada no título e nas combinações título e resumo e título e foco tecnológico.

A palavra *vivax* estava presente em 31 documentos sendo 11 (36%) somente no resumo, 05 (16%) somente no foco tecnológico e nas combinações: 05 (16%) no título e resumo, 6 (19%) no resumo e foco tecnológico, 04 (13%) no título, resumo e foco tecnológico. Esta palavra não foi encontrada no título e na combinação título e foco tecnológico.

A palavra *knowlesi* estava presente em apenas 1 documento no foco tecnológico.

A palavra *malaria* estava presente em 21 documentos sendo 16 (76%) no resumo, 01 (5%) no foco tecnológico e nas combinações: 03 (14%) no título e no resumo e 01 (5%) no resumo e no foco tecnológico. Esta palavra não foi encontrada no resumo e nas combinações título e foco tecnológico e título, resumo e foco tecnológico.

O termo *malaria-infected* estava presente em 3 documentos sendo 02 (67%) no resumo e 01 (33%) no foco tecnológico.

Os termos *malaria.separating* e *malaria-causing* estavam presentes em apenas 01 documento e ambas no foco tecnológico.

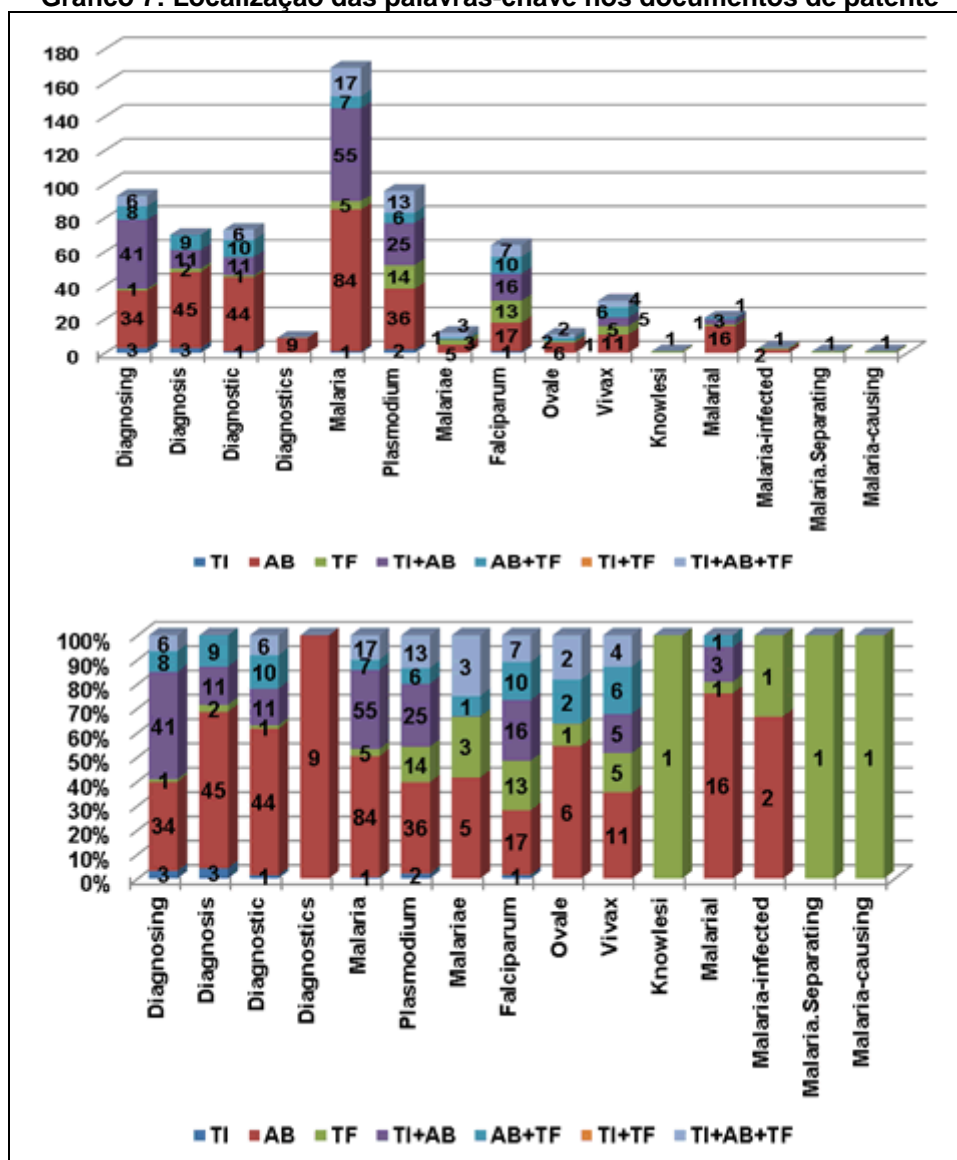
No gráfico 8, a seguir, pode-se visualizar os resultados encontrados e verificar que das 15 palavras encontradas, 12 estão concentradas em sua grande maioria no resumo e as outras 03 no foco tecnológico, ou seja, nenhuma das palavras-chave apresentou relevância no título. Neste ponto há uma grande diferença entre artigos científicos e patentes.

Nos artigos científicos, a principal representação do conteúdo do artigo está no título que conforme Pessoa Jr (2007) afirma que é de suma importância para recuperar a informação desejada. Já no caso dos documentos de patente, o título se apresenta com menor relevância em comparação com os artigos científicos. O título da patente geralmente expressa a natureza da invenção sem, no entanto, conter expressões definidoras de limitações como os artigos científicos.

Vale a pena ressaltar que mesmo na base selecionada para este estudo, a Derwent, que reindexa os títulos de documentos de patentes, confirmou-se o que o título não apresenta relevância para a recuperação de informação contida em documentos de patentes.

A recuperação da informação nas bases de dados de patentes pode ser realizada por vários campos, dentre eles o(s) inventor(es); depositante(s); CIP geralmente associadas ao título e/ou resumo, sendo o mais usual a busca pelo resumo já que, como afirmado anteriormente, o título não é tão relevante. A informação contida no resumo do documento de patente revela a natureza, substância e o(s) objetivo(s) reivindicado(s), geralmente o resumo é harmonioso com a invenção reivindicada com todo e qualquer objeto mencionado na reivindicação.

Gráfico 7: Localização das palavras-chave nos documentos de patente



Ao analisar a relevância do truncamento das palavras-chave, verificou-se que das 06 palavras encontradas, 04 não foram relevantes para a busca por apresentar em sua totalidade, (100%), a presença de outras palavras no mesmo documento, desta forma conclui-se que o documento seria recuperado utilizando a mesma estratégia com a exclusão destas palavras. São elas: *malaria-causing*, *malaria.separating*, *malaria-infected* e *malariae*. Acredita-se que três das palavras em questão somente estavam presentes pela estratégia de busca por um erro de indexação da base ao utilizar sinais de pontuação de forma indevida, no caso, usados para ligar os elementos de palavras compostas e sinais usados para indicar o final de um período, marcando uma pausa absoluta.

Já as palavras *diagnostics* e *malarial* foram consideradas importantes para a estratégia da busca, 44,44% e 28,57% dos documentos recuperados apresentaram as palavras independente da localização, no título e/ou resumo e/ou foco tecnológico

**Quadro 17: Ocorrência de palavras recuperadas a partir da truncagem**

Palavras truncadas	Outras palavras encontradas nos mesmos documentos
<i>Diagnostics</i>	<i>Diagnosing</i> (TI, AB, TF); <i>diagnosis</i> (AB) e <i>diagnostic</i> (AB).
<i>Malaria-causing</i>	<i>Malaria</i> (TI, AB, TF) e <i>plasmodium</i> (TF).
<i>Malaria.Separating</i>	<i>Malaria</i> (TF).
<i>Malaria-infected</i>	<i>Malaria</i> (AB, TF); <i>plasmodium</i> (TI, AB, TF); <i>malariæ</i> (AB, TF); <i>falciparum</i> (AB, TF); <i>ovale</i> (AB, TF); <i>vivax</i> (AB, TF) e <i>malarial</i> (AB).
<i>Malarial</i>	<i>Malaria</i> (TI, AB, TF); <i>plasmodium</i> (TI, AB, TF); <i>malariæ</i> (AB, TF), <i>falciparum</i> (TI, AB, TF); <i>ovale</i> (AB); e <i>vivax</i> (AB, TF).
<i>Malariæ</i>	<i>Malaria</i> (TI, AB, TF); <i>plasmodium</i> (TI, AB, TF); <i>knowlesi</i> (TF); <i>ovale</i> (TI, AB, TF); <i>vivax</i> (TI, AB, TF); <i>malarial</i> (TI, AB) e <i>malaria-infected</i> (TF).

Como apontado no capítulo de estratégia de busca, a truncagem permite a utilização da raiz da palavra para recuperar todas as possibilidades de expansão da mesma que para a formulação de uma estratégia de busca se fez necessário para avaliar a relevância de cada palavra truncada, porém caso não seja esta a finalidade é necessário conhecimento das combinações dos termos para que não seja recuperada informações irrelevantes, denominadas ruídos de informação.

Com base nas informações descritivas e nos resultados, apresentam-se a seguir as conclusões deste trabalho.

## 10. CONCLUSÃO

O monitoramento ambiental, normalmente entendido na literatura como a aquisição e o uso da informação sobre eventos, tendências e relacionamentos que acontecem no ambiente externo da organização, é utilizado para entender as mudanças das forças externas a organização que afetam o seu funcionamento para que seja possível desenvolver respostas rápidas e precisas de forma a manter posições privilegiadas nos mercados em que atuam ao longo do tempo.

Neste contexto, um processo de monitoramento contínuo é fundamental para que as organizações possam identificar e selecionar fontes de informações úteis e confiáveis para a tomada de decisão estratégica.

Os estudos apontam a importância trivial da busca e identificação da informação para o monitoramento, a coleta de informação como elemento crucial já que dela depende a eficácia das outras atividades do Monitoramento Ambiental e que o principal passo é a decisão de quais fontes de informação monitorar.

Para o sucesso de uma busca de informação é necessário planejamento da estratégia de busca que se reflete na decisão da melhor base para o contexto da busca; decisão das palavras-chave; decisão da fórmula lógica da estratégia. Em suma, um conhecimento das técnicas de estratégia de busca, dos prós e contras da estratégia no texto completo dos documentos, da busca em linguagem controlada, da política de indexação, da cobertura de assunto das bases de dados e das características das linguagens de recuperação dos diversos sistemas disponíveis.

Com base nas informações apresentadas, verificou-se que a qualidade da estratégia e a linguagem (vocabulário) são fatores importantes para a recuperação da informação. Estudos comparativos entre a linguagem natural e a linguagem controlada revelaram que nas buscas relativas ao levantamento das últimas tecnologias e, ainda, na busca em documentos de patentes, o uso da estratégia em linguagem natural pode ser o melhor caminho para o encontro da informação desejada já que a terminologia pode ser muito atual, as aplicações ainda não são tão significativas para serem indexadas, portanto os novos termos não foram incorporados a nenhuma lista de linguagem controlada.

Outros estudos apontam que a linguagem natural pode, frequentemente, ser a melhor opção quando se deseja alto índice de retorno, porém é potencialmente ambígua, ou seja, exige muitas vezes a identificação e inclusão de sinônimos dos

termos utilizados. Já a linguagem controlada é “rígida”, inflexível, complica e atrasa o processo de indexação, mas precisa.

Como conclusão dos estudos, principalmente no estudo em bases de patentes revelou que o melhor desempenho da estratégia de busca é aquele que utiliza as duas linguagens concomitantemente incluindo-se, ainda, a Classificação Internacional de Patentes.

Ainda em termos de uso da linguagem verificou-se que para busca em linguagem natural a truncagem pode ser útil para reduzir o número absoluto de palavras em uma estratégia de busca e para se verificar quais as variações quanto aos prefixos truncados, porém caso não seja esta a finalidade é necessário conhecimento das combinações dos termos para que não seja recuperada informações irrelevantes, denominadas ruídos de informação.

Quanto à localização dos termos utilizados na estratégia de busca verificou-se que para busca em documentos de patente o resumo é mais relevante do que o título, que é inversamente proporcional quando comparado aos artigos científicos no qual a principal indexação está no título.

Com relação às bases de patentes públicas e privadas verificou-se que há certas limitações no uso em tais bases. A maior parte das bases gratuitas abrangem documentos nacionais e as que abrangem documentos depositados internacionalmente ou apresentam alguma limitação ou não possuem uma interface amigável para coleta dos dados para análise. Com relação às bases comerciais, um maior número de campos podem ser acessados, as buscas podem ser feitas no documento completo, os mecanismos de busca são mais flexíveis e em algumas bases os resumos dos documentos são reescritos, o que otimiza a recuperação de documentos através de palavras-chave, porém toda essa eficiência gera custo bem mais elevado.

Quanto ao escopo de proteção da patente verificou-se que geralmente a proteção é solicitada de forma ampla que dentre as razões estão: ampliação da abrangência da barreira comercial de forma a impedir a entrada de produtos no mercado de concorrentes; maior possibilidade de licenciamento (troca de royalties); maior retorno dos investimentos e riscos das atividades de Pesquisa & Desenvolvimento, dos gastos com o registro e da colocação do produto no mercado. Além disso, muitas patentes são requeridas sem que o depositante conheça o potencial comercial da invenção e se ela se constituirá num meio efetivo de bloquear

a competição. Outro ponto a considerar é a avaliação pelo analista dos escritórios de concessão que podem restringir o escopo de acordo com a legislação local de cada país.

Não há estratégia de busca que apresente 100% de relevância quanto aos documentos recuperados. A meta de qualquer estratégia de busca seria uma revogação elevada e uma alta precisão, isto é, a recuperação de praticamente todos os documentos relevantes. No entanto, revogação e precisão tendem a se correlacionar inversamente, ou seja, uma maior precisão leva a uma menor revogação e vice-versa, o que significa que o pesquisador de patentes tem de encontrar o equilíbrio adequado entre estes polos de acordo com o tipo de busca escolhida.

De certa forma, esta conclusão vem ao encontro da questão colocada por alguns autores citados de que a lógica da busca para monitoramento tecnológico é de assumir uma busca “imperfeita”, pois se deve trabalhar com o risco de se recuperar documentos na frequência baixa que tanto pode concentrar documentos relacionados à inovação como aqueles que são considerados ruídos estatísticos, por não apresentarem qualquer relação com o que se procura.

Acredita-se que o presente estudo é de grande contribuição para a recuperação da informação na área de patentes, uma vez que são raros os estudos que teorizem a área no aspecto abordado, particularmente na área da saúde. Além disso, a metodologia, os passos para a formulação e execução da estratégia de busca que no caso deste trabalho foi aplicada para malária pode ser utilizada para outras doenças que afetam as pessoas ao redor do mundo o que se torna fundamental para a área da saúde.



## REFERÊNCIAS

- ALMIND, T. C.; INGWERSEN, P. Informetric analyses on the world wide web: methodological approaches to “webmetrocs”. **Journal of Documentation**, v. 53, n. 4, p. 404-426, 1997.
- ALVARADO, R. U. A Bibliometria no Brasil **Ci. Inf.**, Brasília, v. 13, n. 2, p. 91-105, jul./dez. 1984.
- ARAÚJO, C. A. Bibliometria: evolução história e questões atuais. **Em Questão**, Porto Alegre, v. 12, n. 1, p. 11-32, jan/jun. 2006.
- ARAÚJO JUNIOR, R. H.; TARAPANOFF, K. Precisão no processo de busca e recuperação da informação: uso da mineração de textos. **Ci. Inf.**, Brasília, v. 35, n. 3, p. 236-247, set./dez. 2006
- ARAÚJO, V. M. R. H. de A patente como ferramenta da informação. **Ci, Inf.**, Brasília, v. 10, n. 2, p. 27-32, 1981.
- ARCANJO, A. R. L **Estudo da aplicabilidade dos testes imunocromatográficos como diagnóstico para malária na atenção básica de saúde no município de Manaus**. 2004. 67 f. Dissertação (Mestrado) – Universidade do Estado do Amazonas, Manaus
- AUSTER, E.; CHOO, C. W. Environmental scanning by CEOs in two canadian industries. **Journal of the American Society for Information Science**, v. 44, n. 4, p. 194-203, 1993.
- BARROSO, G. B. W. **Élaboration et mise à disposition d’une base de données de documents brevet tombés dans le domaine public**. 2003. Tese (Doutorado) - Universidade de Toulon, Toulouse.
- BHAVNANI, S.K., CLARKSON, G. SCHOLL, M. Collaborative search and sensemaking of patents. In: Conference on Human Factors in Computing Systems, 2008. **Proceedings**, p. 2799-2804. Disponível em: <<http://www.scopus.com/inward/record.url?eid=2-s2.0-57049176573&partnerID=40&md5=8870346fe647a1db2ecd7c3a624b72fe>>. Acesso em 30 maio. 2011.
- BRAGA, Fabiane dos Reis. **Um modelo de monitoramento ambiental (environmental scanning) orientado para o planejamento estratégico da CNEN**. 2008. 96f. Dissertação (Mestrado)-IBICT/UFRJ/ECO.
- BRAGA, G. M. Relações Bibliométricas Entre a Frente de Pesquisa (Research Front) e Revisões da Literatura: Estudo Aplicado a Ciência da Informação. **Ci. Inf.**, Rio de Janeiro, v. 2, n. 1, p. 9-26, 1973.
- BRANÍCIO, S. A. R. ; PEIXOTO, Manoel Otelino da Cunha ; CARPINETTI, L. C. R. . O monitoramento de informações tecnológicas externas para o desenvolvimento de

novos produtos. In: **XXI ENEGEP - Encontro Nacional de Engenharia de Produção**, 2001, Salvador. Anais do XXI ENEGEP, 2001.

BRASIL. Lei nº 9279, de 14 de maio de 1996. Regula direitos e obrigações relativos à propriedade industrial. **Diário Oficial da União**, Brasília, 15 maio. 1996.

\_\_\_\_\_. Ministério da Saúde. Portal da Saúde. **Malária**. Disponível em: <[http://portal.saude.gov.br/portal/saude/profissional/visualizar\\_texto.cfm?idtxt=31086&janela=2](http://portal.saude.gov.br/portal/saude/profissional/visualizar_texto.cfm?idtxt=31086&janela=2)>. Acesso em 01 setembro 2010

\_\_\_\_\_. Ministério da Saúde. Portal da Saúde. **Conhecendo mais sobre a Malária** Disponível em: <[http://portal.saude.gov.br/portal/saude/profissional/visualizar\\_texto.cfm?idtxt=31082&janela=1](http://portal.saude.gov.br/portal/saude/profissional/visualizar_texto.cfm?idtxt=31082&janela=1)>. Acesso em: 22 junho. 2011.

\_\_\_\_\_. Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia. **Por que pesquisa em saúde?** Brasília : Ministério da Saúde, 2007. 20 p. - (Série B. Textos Básicos de Saúde) (Série Pesquisa para Saúde: Textos para Tomada de Decisão)

\_\_\_\_\_. Ministério da Saúde. Secretaria de Vigilância em Saúde **Manual de diagnóstico laboratorial da malária**. Brasília: Ministério da Saúde, 2009. 116 p. – (Serie A. Normas e Manuais Técnicos)

BROOKES, B. C. Biblio, sciento, infor-metrics? What are we talking about? In: EGGHE, L.; ROUSSEAU, R. (Ed.). **Informetrics 89/90**. Amsterdam : Elsevier, 1990. p. 31-43.

CABRAL, L. Patente como Fonte de Informação Tecnológica para as Empresas de Base Tecnológica. In: ENCONTRO DE PROPRIEDADE INTELECTUAL E COMERCIALIZAÇÃO DE TECNOLOGIA, 7., 2004, Rio de Janeiro, **Anais**. Rio de Janeiro, Rede de Tecnologia, 2005, p. 257-270.

CANONGIA, C.; MARIA DE NAZARÉ, F.; ANTUNES, A. Gestão da informação e monitoramento tecnológico: o mercado dos futuros genéricos. **Perspectivas em Ciência da Informação**, v. 7, n. 2, p. 155-166, 2002.

CHAVES, G. C. *et al.* A evolução do sistema internacional de propriedade intelectual: proteção patentária para o setor farmacêutico e acesso a medicamento. **Cad. Saúde Pública**, Rio de Janeiro, v. 23, n. 2, p. 257-267, Fev. 2007.

CHEN, D. Z. *et al.* Constructing a new patent bibliometric performance measure by using modified citation rate analyses with dynamic backward citation Windows. **Scientometrics** v 82, p.149–163, 2010.

COELHO, G.M. **Prospecção tecnológica**: metodologias e experiências nacionais e internacionais. Rio de Janeiro: INT/Finep/ANP Projeto CT-Petro, 2003.

CRONIN, B.; MCKIM, G. Science and scholarship on the World Wide Web: a North American perspective. **Journal of Documentation**, v. 52, v. 2, 1996, p. 163-171.

DEERWESTE, S. *et al.* Indexing by Latent Semantic Analysis. **Journal Of The American Society For Information Science**, v. 41, n. 6, p. 391-407, 1990.

DICIONÁRIO AURÉLIO ELETRÔNICO. Rio de Janeiro: Nova Fronteira, 1996. versão 2.0.

DIRNBERGER, D. A guide to efficient keyword, sequence and classification search strategies for biopharmaceutical drug-centric patent landscape searches - A human recombinant insulin patent landscape case study. **World Patent Information**, v. 33, p. 128-143, 2011.

EMERICK, M. C. **Gestão Tecnológica como Instrumento para a Promoção do Desenvolvimento Econômico-Social: uma proposta para a Fiocruz**. 2004. 239f. Dissertação (Mestrado) – Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rio de Janeiro.

FAIBISOFF, S. G.; ELY, D. P. Information and information needs. In: KING, D. W., ed. **Key papers in the design and evaluation of information systems**. White Plains, Knowledge Industry Publications, c1978. 270 p..

FELDMAN, R.; HIRSH, H. Exploiting background information in knowledge discovery from text. **Journal of Intelligent Information Systems**, v. 9, n. 1, p. 83-97, July/Aug. 1997.

FERREIRA, A. S.; ABREU, M. L. T. de. Desconstruindo um artigo científico. **R. Bras. Zootec**. v. 36, suppl., p. 377-385, Jul, 2007.

FONSECA, E. N. **Bibliometria: teoria e prática**. São Paulo: Ed. Cultrix, 1993. p.12-13.

FORESTI, N. A. B. Contribuição das revistas brasileiras de biblioteconomia e ciência da informação enquanto fonte de referência para a pesquisa. **Ci. Inf.**, Brasília, v. 19, n. 1, p. 53-71, jan./jun, 1990

FRANÇA, R. O. A patente. In CAMPELLO, B. S.; CENDÓN, B. V.; KREMER, J. M. **Fontes de Informação para Pesquisadores e Profissionais**. 2nd ed. Belo Horizonte: Editora UFMG, 2007. cap. 12, p. 153-182.

GOODRICH, R. S. Monitoração do ambiente externo: uma necessidade para as organizações tecnológicas. **Revista de Administração Empresarial**, Rio de Janeiro, v27, n1, 5-10, janeiro/março 1987.

GUIMARÃES, M. C. S. **Avaliação em Ciência e Tecnologia: um estudo prospectivo em química**. 1992. Dissertação (Mestrado). CNPq/IBICT-UFRJ/ECO, Rio de Janeiro. 2 v.

KABACK, S. M. Online patent information: who needs indexing? We do, naturally. **World Patent Information**, v. 14, n. 3, p. 198-199, Aug. 1992.

HAWKINS, D. T. Multiple database searching: techniques and pitfalls. **Online**, v. 2, n. 2, p. 1-15, 1978.

\_\_\_\_\_. Online Information retrieval systems. In: WILLIAMS, M. E., (Ed). **Annual Review of Information Science and Technology**. Washington, American Society for Information Science, c1981. V. 16, p. 182-183.

HAWKINS, D. T.; WAGERS, R. Online bibliographic search strategy development. **Online**, v. 6, n. 3, p. 10-15, May 1982.

KIRKBRIDE, P. Full text, free text and controlled vocabulary strategic search planning. In: ONLINE/CD-ROM 91. **Chicago: Online**, 1992. p. 73-78.

KNOX, D. R.; HLAVA, M. K. Effective search strategies. **Online Review**, v. 3, n. 2, p.148-152, 1979.

KREMER, J. M. Estratégia de Busca. **R. Esc. Bibliotecon. UFMG**, Belo Horizonte, v. 14, n. 2, p. 187-220, set, 1985

KUPFER, D.; TIGRE P. B.. Modelo SENAI de Prospecção: Documento Metodológico. Capítulo 2: Prospecção Tecnológica. In: **Organizacion Internacional Del Trabajo CINTERFOR**. Papeles de La Oficina Técnica n.14, Montevideo: OIT/CINTERFOR; 2004.

INPI. **Apresentação - Portal INPI**. Disponível em: <<http://www.inpi.gov.br/menu-esquerdo/informacao>>. Acesso em 09 julho 2011.

\_\_\_\_\_. **Classificação Internacional de Patentes (IPC) – Portal INPI**. Disponível em: <[http://www.inpi.gov.br/menu-esquerdo/patente/pasta\\_classificacao/index\\_html](http://www.inpi.gov.br/menu-esquerdo/patente/pasta_classificacao/index_html)>. Acesso em 08 julho 2011.

\_\_\_\_\_. **INPI – Instituto Nacional de Propriedade Industrial**. Disponível em: <<http://www6.inpi.gov.br/patentes/instituicoes/convencao.htm>>. Acesso em 08 julho 2011.

\_\_\_\_\_. **Lei 9279/96 - Portal INPI**. Disponível em: <[http://www.inpi.gov.br/menu-esquerdo/patente/pasta\\_despacho/tabela1\\_html](http://www.inpi.gov.br/menu-esquerdo/patente/pasta_despacho/tabela1_html)>. Acesso em 10 julho 2011.

\_\_\_\_\_. **O que é patente? – Portal INPI**. Disponível em: <[http://www.inpi.gov.br/menu-esquerdo/patente/pasta\\_oquee](http://www.inpi.gov.br/menu-esquerdo/patente/pasta_oquee)>. Acesso em 07 julho 2011.

\_\_\_\_\_. **Princípios Básicos – Portal INPI**. Disponível em: <[http://www.inpi.gov.br/menu-esquerdo/patente/pasta\\_classificacao/principios\\_html](http://www.inpi.gov.br/menu-esquerdo/patente/pasta_classificacao/principios_html)>. Acesso em 09 julho 2011.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. Brasília: Brique de Lemos, 1993.

LANCASTER, F. W. **Information retrieval systems: characteristics, testing and evaluation**. 2 ed. New York: Wiley – Interscience, c1979. 400 p.

LONGA, L. C. D. **O Gerenciamento da Informação Tecnológica Contida na Literatura Patentária: uma proposta para a FIOCRUZ**. 2007. 176 f. Dissertação (Mestrado) – Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Rio de Janeiro.

LOOAREESUWAN, S. Malaria. In: Loareesuwan, S.; Wilairatana, P., **Clinical Tropical Medicine**. 1st ed. Bangkok, Thailand. Medical Media. 1999, p 5-10.

LOPES, I. L. Estratégia de busca na recuperação da informação: revisão da literatura. **Ci. Inf.**, Brasília, v. 31, n. 2, p. 60-71, maio/ago. 2002

\_\_\_\_\_. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. **Ci. Inf.**, Brasília, v. 31, n. 1, p. 41-52, jan./abr. 2002.

MACEDO, M. F. G.; Barbosa, A. L. F. **Patentes, Pesquisa & Desenvolvimento: um manual de propriedade intelectual**. Rio de Janeiro: Fiocruz, 2000. 164 p.

MACEDO, M. F. G.; MULLER, A. C. A.; MOREIRA, A. C. **Patenteamento em biotecnologia. Um guia prático para os elaboradores de pedidos de patente**. Brasília: Embrapa Comunicação para Transferência de Tecnologia, 2001. 200 p.

MACIAS-CHAPULA, C. A. O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. **Ci. Inf.**, Brasília, v. 27, n. 2, p. 134-140, maio/ago. 1998.

MARKEY, K.; ATHERTON, P. **Online training and practice manual for ERIC database searchers**. Syracuse: Syracuse University, 1978.

MUDDAMALLE, M. R. Natural language versus controlled vocabular in information retrieval: a case study in soil mechanics. **Jasis**, v. 49, n. 10, p. 881-887. Oct. 1998.

NOPPADON, T. *et al.* Malaria Diagnosis: A Brief Review **Korean J Parasitol**. v. 47, n. 2, p. 93-102, Jun, 2009

OLDROYD, B. K; CITROEN, C. L. Study of strategies used in online searching. **Online Review**, v. 1, n. 4, p. 295-310, 1977.

OLIVEIRA, P. H.; PAULA NETO, W. de; OLIVEIRA, F. H. R. de. **Monitoramento ambiental e fontes de informação para a tomada de decisão estratégica: o caso dos varejos formais instalados nas imediações do shopping popular Oiapoque em Belo Horizonte**. *Ci. Inf.* [online]. 2008, vol.37, n.3, pp. 103-114. ISSN 0100-1965.

OLIVEIRA, L. G. de *et al.* Informação de patentes: ferramenta indispensável para a pesquisa e o desenvolvimento tecnológico. **Quim. Nova**, v. 28, Suplemento, S36-S40, 2005

OLIVEIRA, M. R. F. **Análise de custo-efetividade de teste rápido para o diagnóstico de casos novos de malária em doze municípios endêmicos do Estado do Pará.** 2009. 216 f. Tese (Doutorado) – Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo.

PALOP, F; VICENTE GOMILA, J.M.A.. **Vigilância Tecnológica e Inteligência competitiva: su potencial para la empresa espanola.** Madrid, Espana: Fundación COTEC para la inovación Tecnológica, 1999. 107 p.

PESSOA JR, A. **Preparo de Artigo Científicos.** São Paulo: Universidade de São Paulo, 2007. 92 slides, color, 25,4 cm x 19,05 cm

PORTER, A. L. et al. **Forecasting and management of technology.** New York: John Wiley, 1991. 464 p.

\_\_\_\_\_. Technology Futures analysis: toward integration of the field & new method. **Technol Forecast Soc Change.** 2004;71 (3):287-303.

QUONIAM, L. **Les productions scientifiques em bibliometrie et dossier de travaux.** Marseille, Universite Aix-Marseille III, 1995 (Habilitation a diriger recherches). 182 p.

SANCHO, Rosa. Indicadores bibliometricos utilizados en la evaluacion de la ciencia y la tecnologia: revision bibliografica. **Revista Española de Documentación Científica,** v. 13, n. 3/4, p. 842-865, 1990.

SANTOS *et al.* Prospecção de tecnologias de futuro: métodos, técnicas e abordagens. **Parcerias estratégicas,** n.19, Dez. 2004.

SKUMANICH, M.; SILBERNAGEL, M. **Foresighting around the world:** a review of seven bent-un-kind programs. Seattle: Battelle, 1997.

SILVA, C. H. **Services d'information dans un monde globalisé: tendances et stratégies.** 2002. 342 f. Tese (Doutorado) - Universite d'Aix-Marseille III, Marseille.

SMITH, A. A tale two web spaces: comparing sites using web impact factors. **Journal of Documentation,** v. 55, n. 5, p. 577-592, Dec. 1999.

SPINAK, E. **Diccionario enciclopédico de bibliometría, ciencia métrica e informetría.** Montevideo, 1996. 245 p.

SPINK, A.; SARACEVIC, T. Dynamics of search term selection during mediated online searching. In: ASIS ANNUAL MEETING, 56th, 1993, Columbus. **Proceedings.** New York, 1993. v. 30, p. 63-72.

STREHL, L. O fator de impacto do ISI e a avaliação da produção científica: aspectos conceituais e metodológicos **Ci. Inf.,** Brasília, v. 34, n. 1, p.19-27, jan./abr. 2005

SUTTON, H. **Competitive intelligence.** New York: The Conference Board, 1988.

TARAPANOFF, K. **Técnicas para tomada de decisão nos sistemas de informação**. 3ª ed. Brasília: Thesaurus, 2002. 163 p.

VAN RAAN, A. F. J. Scientometrics: state-of-art. **Scientometrics**, v. 38, n. 1, p. 205-218, 1997.

VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**, Brasília, v. 31, n. 2, Aug. 2002.

VITZTHUM *et al.* Proteomics: from basic research to diagnostic application. A Review of requirements & needs. **Journal of Proteome Research**, v.4, p. 1086-97, 2005.

WORD HEALTH ORGANIZATION. **A rapid dipstick antigen capture assay for the diagnosis of falciparum malária**. Geneva: 1995.

WORLD INTELLECTUAL PROPERTY ORGANIZATION. **Classificação Internacional de Patentes**. Disponível em: <<http://www.wipo.int/classifications/ipc/en/general/preface.html>>. Acesso em 08 julho 2011.

\_\_\_\_\_. **Preface to the International Patent Classification**. v. 5, 8 ed. 2006, 50 p.

\_\_\_\_\_. **WIPO-Administered Treaty**. Disponível em: <[http://www.wipo.int/treaties/en/ShowResults.jsp?lang=en&treaty\\_id=2](http://www.wipo.int/treaties/en/ShowResults.jsp?lang=en&treaty_id=2)>. Acesso em 08 julho 2011

WORMELL, I. Informetria: explorando bases de dados como instrumentos de análise. **Ciência da Informação**, Brasília, v. 27, n. 2, 1998.