



Encontros Bibi: revista eletrônica de
biblioteconomia e ciência da informação

E-ISSN: 1518-2924

bibli@ced.ufsc.br

Universidade Federal de Santa Catarina
Brasil

Lima Gonçalves, Aline

Uso de resumos e palavras-chave em Ciências Sociais: uma avaliação

Encontros Bibi: revista eletrônica de biblioteconomia e ciência da informação, vol. 13, núm. 26, 2008

Universidade Federal de Santa Catarina

Florianópolis, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=14712794006>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

 redalyc.org

Sistema de Informação Científica

Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

Uso de resumos e palavras-chave em Ciências Sociais: uma avaliação

The use of abstracts and keywords in Social Sciences: an evaluation

Aline Lima Gonçalves - alineg@usp.br

Mestranda em Ciência da Informação

Escola de Comunicações e Artes, Universidade de São Paulo

Resumo

Este trabalho tem como objetivo avaliar resumos e palavras-chave de artigos de periódicos publicados em Ciências Sociais por autores brasileiros, a fim de identificar as principais características desses elementos e como isso afeta a representação e recuperação de informação nessa área. Para isso, foi feita uma avaliação de 151 resumos de artigos produzidos por professores da Universidade de São Paulo em três disciplinas, Antropologia, Ciência Política e Sociologia, na tentativa de descobrir como os autores exploraram os tipos de resumo (descritivo, informativo ou descritivo-informativo); num segundo tópico, foram avaliadas as palavras-chave que acompanhavam o resumo e o seu potencial como complemento deste e como ferramenta para relacionar disciplinas. Observou-se que os autores preferem resumos descritivos em detrimento dos informativos, e que, apesar da falta de padronização, as palavras-chave fornecidas pelos autores podem ser usadas como complemento dos resumos e como indicador de algumas relações interdisciplinares. Conclui-se que, apesar da importância de resumos e palavras-chave como elementos de representação de informação, suas funções são pouco exploradas nas Ciências Sociais, o que pode afetar negativamente a recuperação da informação.

Palavras-chave: Resumos. Palavras-chave. Metadados. Representação da informação. Recuperação da informação.

1 INTRODUÇÃO

Representar apropriadamente informações contidas em um determinado documento é de fundamental importância para a recuperação informacional, e os termos de indexação, somados aos resumos, são os principais produtos dessa atividade. Juntos, eles descrevem o conteúdo de um registro, indicando seus pontos principais. Nesse sentido, são elementos que facilitam a comunicação do conhecimento, já que funcionam como “ferramentas de representação”, necessárias em um processo inicial de filtragem, permitindo assim que a informação flua entre o universo dos documentos originais e o dos usuários de informação (PINTO, 2003).

A importância desses dois elementos no processo de representação e recuperação informacional é evidente também no âmbito da comunicação científica. O aumento progressivo do número de publicações e a consequente impossibilidade de ler tudo o que é publicado sobre um assunto tornou necessária a aplicação de mecanismos que pudessem traduzir o conteúdo de um documento de maneira mais condensada, sem que houvesse perda de informação. Ao mesmo tempo, o surgimento de bases de dados com abrangência cada vez mais ampla e mesmo o crescimento de publicações científicas on-line tornou necessário o desenvolvimento de padrões de metadados, ou seja, padrões para a extração de elementos relevantes para tanto para a descrição de um documento quanto para a recuperação de informações. Resumos e palavras-chave (aqui, em substituição aos termos de indexação), assim, assumiram também a função de metadado, e é normal verificar em artigos de periódicos a presença desses dois itens como elementos pré-textuais obrigatórios.

No entanto, a apresentação desses elementos em periódicos nem sempre conta com critérios bem definidos, e é bastante comum observar a exigência de que os autores dos artigos elaborem resumos e formulem palavras-chave para seus próprios trabalhos. Porém, como assinalou Borko (1975), os

autores não necessariamente redigem os melhores resumos para seus próprios artigos, dada a carência de formação especializada e o desconhecimento de regras específicas nessa tarefa. O mesmo também pode ser dito em relação à escolha de palavras-chave: o processo de indexação é altamente complexo, e cada indivíduo pode ter uma maneira diferente de relacionar conceitos, de acordo com suas próprias experiências. Alguns conceitos podem estar relacionados sob verbetes diferentes daqueles que o leitor supunha, e, ademais, conceitos também mudam com o passar do tempo (MEADOWS, 1999).

Se os problemas apontados por Meadows podem ocorrer no próprio processo de indexação, mesmo com critérios bem definidos quanto à escolha de termos, a questão tende a se tornar mais grave com relação às palavras-chave, normalmente empregadas em linguagem natural e sem normas específicas de escolha ou de redação. Como o próprio Meadows colocou, em pouco tempo o vocabulário usado por uma determinada disciplina pode ficar obsoleto. Mais ainda, ele pode ser usado de maneiras diferentes entre áreas distintas ou até dentro de uma mesma área.

No âmbito das Ciências Sociais o tema merece particular atenção, dado o estado fragmentário de sua formação teórica, sob a forma de disciplinas díspares e desconexas, como indicou Nunes (2005). As Ciências Sociais tem como objeto de estudo a realidade social, e cada uma de suas disciplinas adota uma ótica diferente em relação a esse objeto, empregando diferentes abordagens metodológicas. A dificuldade em delimitar um objeto teórico (já que a realidade social é algo extremamente amplo) coloca áreas como a Sociologia, a Economia, a Psicologia Social e a Antropologia dentro do escopo das Ciências Sociais (SILVA E PINTO, 1986 apud FERNANDES, 2002), às quais Nunes (2005) adiciona ainda a Geografia Humana, a Demografia, a Ciência Política a Lingüística e a Etnologia Social. Assim, não é de se surpreender que dentro de uma mesma ciência social seja possível encontrar textos teóricos muito diferentes relativos a um mesmo objeto real, tanto quanto aos conceitos empregados e às relações existentes entre eles quanto aos resultados aos quais esses textos conduzem. Dessa maneira, um mesmo objeto real pode resultar em diferentes objetos científicos, e isto dentro de uma mesma disciplina (NUNES, 2005, p. 45).

Tendo em vista todas essas questões próprias das Ciências Sociais e a relevância de resumos e palavras-chave em termos de representação e recuperação, é possível dizer que esses elementos, quando escritos pelos próprios autores dos artigos, funcionam adequadamente para representação em áreas tão diversificadas disciplinarmente? No amplo universo das Ciências Sociais, palavras-chave serviriam também para demonstrar relações entre campos de estudos? Na tentativa de responder a essas perguntas, este artigo visa avaliar resumos e palavras-chave de artigos de periódicos publicados em Ciências Sociais por autores brasileiros, procurando identificar as principais características desses elementos e como isso pode afetar a representação e recuperação de informação nessa área.

A primeira parte do artigo faz uma pequena caracterização sobre resumos e palavras-chave como instrumentos de recuperação e representação informacional, apontando tipologias, principais funções e uma pequena revisão teórica. Depois, é descrito o método de avaliação de artigos da área de Ciências Sociais. Uma análise dos resultados obtidos é feita a seguir, finalizando, assim, com algumas considerações sobre o uso de resumos e palavras-chave em Ciências Sociais e o seu possível impacto para recuperação de informações.

2 RESUMOS E PALAVRAS-CHAVE COMO FERRAMENTAS DE REPRESENTAÇÃO E RECUPERAÇÃO

2.1 Resumos: revisão teórica, conceitos e funções

Representação e condensação. Esses dois atributos constituem a base sobre a qual se situam

resumos usados com fins documentários. Conforme a própria definição da NBR 6028 (ABNT, 2003a), um resumo nada mais é do que a representação concisa de pontos relevantes de um documento. Em outras palavras, um resumo é tanto uma representação quanto documento condensado (CROSS; OPPENHEIM, 2006), sendo, portanto, um texto independente, e nesse sentido ele pode ser analisado também como uma representação textual autônoma, com base em um texto original (PINTO, 2006). Nesse ponto, como lembra Lancaster (1993, p. 88), o resumo é diferente de um “extrato”, que apenas extrai frases do próprio texto (por exemplo, uma frase da introdução e duas ou três da conclusão).

Temos nessas definições, portanto, a noção de “representação”, em que um conjunto de significados é usado em substituição de outros, e a noção de “condensação”, em que apenas os elementos mais importantes serão destacados. É isso o que permite situar o resumo como uma ferramenta interessante para recuperação de informações, já que representações ou substitutos residem no cerne de sistemas de informação bibliográfica e são essenciais para uma efetiva transferência de informação em todos os campos científicos (TIBBO, 1993).

Embora o uso de resumos seja observado desde a antiga Mesopotâmia, na qual rótulos de conteúdo sobre envelopes de argila preveniam as tábua cuneiformes de qualquer alteração, como bem observaram Tibbo (1993, p. 21), e sempre tenham sido utilizados ao longo da história, foi mesmo no século XX que o uso de resumos ganhou força como ferramenta auxiliar no processo de representação e recuperação informacional. Após a segunda guerra mundial, houve um aumento tão grande no número de trabalhos publicados que ler integralmente todos os documentos relevantes tornou-se impossível. Os resumos, então, passaram a ser itens regulares em boa parte dos periódicos acadêmicos, e serviços especializados de elaboração de resumos vieram à tona. Após o surgimento de bases de dados eletrônicas, enquanto ainda não era possível recuperar texto completo, o uso de resumos tornou-se compulsório. Atualmente, em ambientes eletrônicos em que a recuperação de texto integral já não constitui problema, o papel do resumo como objeto de informação continua válido, já que muitos pesquisadores o utilizam como substituto do texto completo e como identificador para um documento maior. Nesse sentido, ele é um elemento de metadado, um dado útil para a descrição de um documento (WAARD; KIRCHZ).

É, portanto, no período pós-guerra, particularmente após a década de 60, que os resumos se tornam objeto de estudo, tendo como principal foco o quesito “qualidade”. Autores como Borko e Bernier (1975), por exemplo, defendem que a qualidade de resumos está ligada à sua brevidade, exatidão e clareza, o mesmo observado em Lancaster (1993, p. 100), que acrescenta ainda que um resumo deve ser estruturado de acordo com informações contidas no título do item, sem repeti-las.

Num outro contexto, a qualidade de resumos é focada quanto ao seu uso em sistemas de recuperação automática, como bases de dados. Alguns desses estudos se concentram no item “geração automática de resumos”. Nesse ponto, Johnson (1995) apresenta o estado da arte das pesquisas nesse setor, e aponta a grande dificuldade em encontrar uma técnica que seja capaz de formular resumos mais completos. Maria Pinto (2003) aborda algumas tendências de pesquisa (teórica, metodológica e pragmática) a fim de adaptar o uso de resumos em ambientes digitais. No que se refere aos resumos em sistemas de informação, Pinto e Lancaster (1999), ao comparar resumos para leitura por humanos e por máquinas, concluem que resumos ainda são de grande utilidade para serviços de informação, apesar da disponibilidade de buscas por texto completo. Pinto (1999) identifica ainda quatro paradigmas quando da elaboração de resumos para uso em sistemas de recuperação: comunicacional (bastante ligada à teoria matemática da comunicação de Shannon e Weaver, visando à transmissão de informação via resumos), físico (contemplando métodos de extração automática e recuperação da informação), cognitivo (abrangendo processamento de informação, inteligência artificial e também recuperação de informação) e o paradigma sistêmico (gestão da qualidade de resumos).

O grande diferencial dos estudos sobre resumo, no entanto, não está apenas na avaliação teórica de sua formulação, mas também como o usuário de resumos e o próprio resumidor se envolvem o processo de extração dos itens mais relevantes de um texto. A questão vem ganhando importância, o que é observado em estudos recentes como os de Pinto (2006), que foca a questão da qualidade de resumos tendo em vista suas variáveis e atributos, admitindo a subjetividade como um fator que interfere tanto na elaboração quanto na compreensão de um resumo. De acordo com a autora, o atributo da qualidade percebida nos resumos depende exclusivamente dos usuários. Montesi e Urdiciain (2005) também avaliam os resumos da perspectiva dos usuários, numa pesquisa que tenciona apontar os principais problemas indicados por pesquisadores de seis universidades europeias em um conjunto de resumos das áreas de agricultura e educação, com resultados bastante negativos em relação à qualidade da terminologia, nível de condensação, falta ou excesso de informação e outros atributos. A ausência de informações importantes também foi apontada em estudos como os de Cook, Beckman e Bordage (2007), que demonstraram a ausência de uma série de elementos importantes na área de educação médica, e por Cross e Oppenheim (2005), ao analisarem um número de resumos de uma base de dados referencial da área de protozoologia, cuja amostra contou também com resumos elaborados por autores de artigos.

É interessante notar que boa parte dos estudos sobre resumos concentram-se ou na teoria sobre sua elaboração e avaliação, ou na avaliação tendo como objeto bases de dados específicas, sobretudo na área de ciências biológicas. Quase nada é observado nas áreas de ciências humanas ou sociais, das quais pouco se sabe em termos de avaliação de ferramentas de representação. Uma das poucas exceções é o estudo de Tibbo (1993), que avalia resumos da área de História, contemplando inclusive aqueles escritos por historiadores quando escrevem sobre o seu próprio trabalho. É também um dos poucos trabalhos que focam a qualidade de resumos redigidos por autores.

A questão da avaliação de sistemas de extração automática de resumos também é tema de estudos, dos quais o de Ou, Khoo e Goh (2008) serve como um dos exemplos mais recentes. Os autores apresentam um sistema de elaboração de resumos “multi-documento”, com base na análise de discurso, extração e integração de informação, tendo como *corpus* um conjunto de resumos de dissertações em sociologia. O sistema contou com avaliações positivas dos usuários, mas os próprios autores do artigo afirmam que mais estudos precisam ser feitos para avaliar a verdadeira eficácia desses sistemas em comparação à codificação humana.

Vistos esses aspectos em relação à formulação de resumos, quais seriam suas funções e tipos principais? Borko e Bernier (1975) identificaram como funções a capacidade de substituição, auxiliando a atualização de pesquisadores, que em geral não têm tempo de ler tudo o que é publicado sobre suas próprias áreas; a capacidade de poupar o tempo do leitor; a facilidade de seleção; de superação de barreiras lingüísticas (já que resumos são, em geral, escritos em inglês ou no idioma do leitor, o que permite alguma uniformização idiomática); facilitar a pesquisa por literatura; de melhorar a eficiência na indexação; e de auxílio na preparação de revisões.

Semelhante à caracterização acima, porém de forma mais condensada, Waard e Kircz e Johnson (1995) situam as funções do resumo dentro de um ambiente eletrônico. Para os autores, resumos assumem as seguintes funções:

1. Seleção: já que não se pode ler todos os artigos publicados, é preciso facilitar a escolha do leitor
2. Substituição: todas as informações relevantes devem estar no resumo. Ex: resultados experimentais.

3. Recuperação: de fato, o resumo ideal é uma cadeia de palavras-chave ligadas em uma sentença de fácil leitura.
4. Orientação: o resumo serve como suporte àqueles que lêem, e num ambiente eletrônico ele pode ser o eixo de todos os componentes do texto.

Existem vários tipos de resumo, mas três em especial são tidos como os de maior utilidade para periódicos científicos: indicativo (ou descritivo) e o informativo, que são os tipos apontados na norma brasileira para elaboração de resumos (ABNT 6028). O resumo descritivo traz apenas uma visão geral do texto ou uma breve descrição de suas fontes. Contém apenas os tópicos principais, não apresentando dados qualitativos ou informativos, resultados, conclusões ou recomendações. O foco desse tipo de resumo é normalmente o documento, e em geral não dispensa a consulta ao original. (ABNT, 2003; TIBBO, 1993). Já o tipo informativo apresenta mais detalhes, incluindo pormenores sobre metodologia, resultados, discussões e conclusões. É formulado de tal forma que assume mais autonomia e pode ser visto como um substituto do documento original. É apontado como o tipo mais adequado para fins de recuperação. Em algumas outras classificações, existe um outro tipo de resumo que é chamado informativo-indicativo (ou informativo-descritivo) (CROSS; OPPENHEIM, 2006; TIBBO, 1993). Este tipo informativo-indicativo combina as configurações de resumos indicativos e informativos, trazendo características gerais como nos resumos indicativos mas incluindo alguma conclusão breve (CROSS; OPPENHEIM, 2006).

Além desses aspectos relativos ao conteúdo, resumos também podem ser estruturados ou não estruturados em relação à sua forma de apresentação. Resumos estruturados são aqueles que apresentam partes bem definidas, com um cabeçalho que nomeia cada uma dessas partes, como “contexto”, “objetivos”, “método” e “conclusões”, e que substituíram os resumos tradicionais na maior parte dos periódicos da área médica (HARTLEY, 1998). São, porém, pouco utilizados pelas ciências humanas e sociais, cujos resumos são normalmente apresentados como um parágrafo longo, sem distinção formal entre as partes do documento.

2.2 Palavras-chave: conceitos e revisão

A identificação e atribuição de assuntos, tendo como fins representação e recuperação de informações, é uma atividade normalmente realizada por serviços especializados, por meio do processo de “indexação”. Como a própria palavra indica, a indexação de assuntos envolve a geração de “índices” (index), tratando-se de um processo mental que visa analisar o conteúdo informacional de registros de conhecimento e expressa-lo na linguagem de um sistema (Borko, 1978, p. 8). Nesse contexto, os termos de indexação desempenham uma função semelhante à do resumo, já que, enquanto o resumo tem como função indicar de que trata o documento ou sintetizar seu conteúdo, os termos de indexação podem oferecer uma espécie de miniresumo, que serve de ponto de acesso pelo qual um item é localizado e recuperado (LANCASTER, 1993, p. 5). A definição de palavra-chave está bastante ligada a essa idéia. Numa definição simples, palavra-chave é a palavra que traduz o sentido de um contexto, que o torna claro e o identifica, ou ainda palavra que identifica elementos correlatos ou que pertençam à mesma área de interesse para fins de pesquisa (HOUAIS, 2001, p. 2108). Na norma brasileira relativa à redação de artigos científicos, palavras-chave são representativas do conteúdo do documento e são escolhidas em vocabulário controlado (ABNT, 2003b).

No entanto, convém diferenciar “palavra-chave” do termo de indexação propriamente dito: o termo é fruto de atividades mentais complexas, que envolvem os processos de análise conceitual (identificação dos assuntos do documento) e tradução (conversão da análise conceitual num determinado conjunto de termos), os principais aspectos da indexação, como apontou Lancaster (1993). Palavra-chave pode ser extraída de qualquer parte do documento (do título, do resumo ou de qualquer outra parte do texto), e não necessariamente implica controle de vocabulário. Sua atribuição é livre, podendo ser definida pelos autores do documento, por editores, e até ser geradas

automaticamente.

Dentro de um artigo científico, palavras-chave indicam os principais conceitos e a cobertura de um assunto, úteis para a indexação em mecanismos de busca ou categorização de texto (ERCAN; CICEKLI, 2007). Assim, no escopo da divulgação científica em periódicos especializados, palavras-chave normalmente:

1. permitem que os leitores decidam se um artigo possui ou não material relevante de acordo com seus interesses;
2. fornecem aos leitores termos para uso em buscas de documentos similares na *web*;
3. auxiliam indexadores e editores a agrupar materiais relacionados;
4. permitem que editores e pesquisadores registrem mudanças nos assuntos de uma disciplina através do tempo;
5. ligam tópicos específicos de pesquisa dentro de um escopo maior (HARTLEY; KOSTOFF, 2003).

As pesquisas em torno do uso de palavras-chave, dadas as suas características de não padronização, assumem um papel diferenciado em relação aos estudos sobre indexação e terminologia, embora se confundam um pouco com esses dois itens. Borko (1978) situa bem o uso de índices de palavras em contraposição aos índices de assunto (duas coisas bastante distintas), e caracteriza o uso de palavras-chave dentro de um contexto e fora de um contexto. São as duas siglas KWIC e KWOC (*keyword in context* e *keyword out of context*, respectivamente). O primeiro item refere-se a palavras extraídas diretamente do título, que assumiriam uma qualidade informativa desde que situadas com outras palavras do título, formando um contexto identificável. O segundo item é uma forma de evitar um truncamento, ou seja, uma lista de palavras que não entraram no contexto do documento. As duas práticas apresentam sérias desvantagens decorrentes da falta de padronização, mas se mostraram importantes na elaboração dos primeiros sistemas de informação automatizados.

De fato, a descrição automática dos assuntos do documento é um dos grandes desafios tanto em sistemas de recuperação especializados quanto em mecanismos de busca na Internet. Um dos principais métodos usados para esse fim é a indexação automática, ou extração automática de palavras-chave, que conta com técnicas de processamento de linguagem natural, e, a exemplo do que acontece no caso de resumos gerados automaticamente, extrai palavras potencialmente indicativas para representar o conteúdo de um documento. (ERCAN; CICEKLI, 2007).

Diversos estudos sobre o tema tem sido produzidos com diferentes vieses: Anderson e Pérez-Carballo (2001a; 2001b), por exemplo, focam a questão da natureza da indexação, fazendo um contraponto entre a indexação humana e a produzida por máquina, com e sem controle de vocabulário. Boger e outros (2001) também fazem uma comparação semelhante, mas abordando a identificação de palavras-chave por redes neurais artificiais e por usuários de sistemas de filtragem. Um exame de bases de dados e de catálogos on-line de bibliotecas é feito por East (2006), com a intenção de avaliar a indexação de monografias da área de filosofia. Já Voorbij (1998) avalia o grau de recuperação de monografias existentes no catálogo on-line da Biblioteca Nacional da Holanda, comparando a eficiência de descritores de assunto e palavras-chave extraídas do título.

Apesar da importância das palavras-chave para recuperação, nem todos os periódicos possuem critérios de seleção muito explícitos. Em algumas bases de dados e periódicos especializados, pode-se contar com o auxílio de profissionais que farão a escolha de quais palavras-chave são mais adequadas para representação de um determinado artigo. Contudo, dadas as dificuldades financeiras e operacionais de manter esse esquema, é comum observar nas regras para submissão de artigos orientações para que o próprio autor designe as palavras-chave para seu texto, embora esse processo não seja o mais adequado. Apesar de isso ser um fato recorrente, há poucos trabalhos relacionados a

palavras-chave oferecidas pelos próprios autores, apesar de esse elemento ser de apresentação obrigatória na maior parte dos periódicos, inclusive na área de Ciências Sociais.

Uma das poucas exceções são os dois artigos de Gil-Leiva e Alonso-Arroyo (2005, 2007). As autoras analisam nesses dois trabalhos palavras-chaves fornecidas por autores em contraste com os termos fornecidos por indexadores de três bases de dados. Conclui-se que palavras-chave fornecidas pelos autores estavam direta ou indiretamente presentes em descritores de assunto feitos por indexadores profissionais. Portanto, é possível dizer que tais palavras são importantes tanto na recuperação de informação por um pesquisador quanto como orientação ao trabalho de profissionais que trabalham com indexação.

Vistos todos esses aspectos, o que mais interessa ao escopo deste artigo, é o processo em que os próprios autores escrevem os resumos e indicam as palavras-chave a serem usadas em seus textos, considerando que a indexação em bases de dados também tomará como fonte esses elementos já formulados pelos autores.

3 PROCEDIMENTOS METODOLÓGICOS

A pesquisa teve caráter exploratório, com alguns procedimentos quantitativos. Consistiu basicamente na coleta de artigos em Ciências Sociais e na análise de resumos e palavras-chave desses documentos, visando identificar alguns pontos comuns de elaboração. No caso dos resumos (abstracts), os pontos analisados foram:

- a) Tipologia: descritivo, informativo ou descritivo-informativo. Tomou-se como exemplo de categorização o trabalho de Tibbo (1993), já citado, pelo fato de oferecer uma categoria a mais nos casos de resumos não totalmente informativos;
- b) Extensão: número de palavras empregadas para a elaboração do resumo;
- c) Presença de elementos para a identificação do conteúdo (objetivos, métodos, resultados, discussão ou recomendações, conclusões, contexto de pesquisa) e clareza. Neste último ponto, não foi abordada a clareza da linguagem em si, visto que isso demandaria um amplo conhecimento do vocabulário da área, mas sim o quanto a estrutura geral do resumo deixava clara as estruturas e propósitos do artigo.

Quanto às palavras-chave, relembramos dois tópicos apontados anteriormente por Hartley e Kostoff (2003): o uso desses elementos como ferramentas que fornecem aos leitores termos para uso em buscas de documentos similares na *web*; e que ligam tópicos específicos de pesquisa dentro de um escopo maior. Assim, pressupondo que as palavras-chave foram escolhidas pelos próprios autores e que estes têm domínio do vocabulário empregado na área, optou-se por uma forma de análise que situasse as palavras-chaves como instrumento de recuperação complementar ao resumo, e como uma forma de relacionar as informações de cada artigo num contexto maior de pesquisa. Não foi feita uma análise detalhada de aspectos semânticos ou da pertinência das palavras-chave para cada artigo individualmente, já que isso demandaria uma leitura do texto como um todo.

A fim de obter um corpus básico de análise, foram selecionados artigos publicados entre os anos de 2004 e 2007 por professores da Faculdade de Filosofia e Ciências Sociais da Universidade de São Paulo, tomando como fonte os Currículos Lattes dos docentes. Optou-se por esse forma de seleção porque, primeiramente, a produção em Ciências Sociais na Universidade de São Paulo tem representatividade no âmbito nacional (os cursos em Ciências Sociais na universidade têm classificação 6 e 7 pela Capes), e, em segundo lugar, porque dessa maneira seria possível ter uma diversidade de periódicos que talvez não fosse observável na análise de bases de dados específicas. Para este estudo, foram selecionados apenas os artigos dispostos em meio eletrônico, sendo desconsiderados aqueles que só poderiam ser encontrados em papel.

Um outro critério adotado para a seleção foi a escolha de artigos que possuíssem ao menos um resumo (abstract), independentemente do idioma. Assim, foram igualmente desconsiderados artigos que eram tradução de algum trabalho realizado por outro autor, resenhas de obras monográficas e entrevistas, por não possuírem resumos e palavras-chave. Da mesma forma, outros tipos de artigos que não apresentavam ao menos o primeiro elemento não foram abordados nesta pesquisa.

Resta lembrar também que, para o âmbito deste estudo, foi adotada a mesma categorização disciplinar usada na Universidade de São Paulo, ao afirmar que o núcleo das Ciências Sociais é formado por três áreas: Sociologia, Antropologia e Ciência Política. Economia, Direito e Comunicação, situadas como constituintes das Ciências Sociais em algumas classificações, foram consideradas áreas independentes.

Feitas essas ressalvas, foram recuperados e avaliados 151 artigos, sendo 32 da área de Antropologia, 58 de Sociologia e 61 de Ciência Política. Como há uma gama muito variada de títulos indexados por diferentes bases de dados, e alguns títulos que talvez nem sejam indexados, não foi feita nesta pesquisa uma análise profunda do critério de cada periódico em relação à elaboração de elementos pré-textuais, embora isso seja pontuado de forma geral ao longo desta pesquisa, a fim de identificar um quadro mais amplo.

4 RESULTADOS OBTIDOS E DISCUSSÃO

4.1 Resumos

Antes de proceder à descrição dos resultados, é importante ressaltar que a identificação dos tipos de resumo (descritivo, informativo ou descritivo-informativo) não foi simples, pelo próprio fato de o resumo em Ciências Sociais não ser estruturado, somando-se a isso a natureza híbrida dos estudos na área, ora com forte tendência teórica e ensaística, ora com tendência empírica. Resumos que não apresentavam claramente objetivos ou métodos mas apenas uma idéia geral do contexto da pesquisa (mais comum em artigos com o formato de ensaio), ou apenas objetivo e hipótese, foram classificados como descritivos.

4.1.1 Tipologia: descritivo, informativo ou descritivo-informativo

Os resultados obtidos na categorização dos resumos podem ser vistos na tabela abaixo:

Tabela 1. Tipos de resumos

	Descritivo	%	Informativo	%	Descritivo-Informativo	%	Total
Antropologia	20	62	4	13	8	25	32
Sociologia	37	64	15	26	6	10	58
Ciência Política	36	59	14	23	11	18	61
total	93	61	33	22	25	17	151

Os dados mostram que há uma preferência por resumos descritivos em Ciências Sociais (61% do total analisado de 151 resumos), predominância que se revela em todas as sub-áreas avaliadas, principalmente em Antropologia e Sociologia (com 62% e 64%, respectivamente). Por outro lado, os resultados apontam para a presença de resumos descritivo-informativos em 17% do total, e de resumos informativos em 22% do total. Apesar do número de artigos avaliados ter sido inferior em relação à Sociologia e Ciência Política, é possível notar uma tendência maior da área de Antropologia quanto ao uso de resumos descritivo-informativos, ou seja, resumos que não apresentam todos os elementos relevantes do texto mas que incluem alguma conclusão breve. Nesta área, a ocorrência desse tipo foi superior ao uso de resumos informativos, fato que não acontece em Sociologia e Ciência Política, em que resumos informativos aparecem mais vezes.

De qualquer forma, a prevalência de resumos descritivos sobre os outros dois tipos gera um ponto de dúvida quanto à validade dos resultados obtidos, uma vez que os resumos descritivos, que são

informações em Ciências Sociais, já que um dos principais pontos para julgar a qualidade dos resumos é a sua capacidade em passar todos os pontos mais importantes do texto. Assim, resumos informativos deveriam ser mais explorados pela área, tanto em trabalhos teóricos quanto empíricos.

4.1.2 Extensão

A eficiência de um resumo depende mais da forma como é estruturado e de que informações apresenta do que da sua extensão. No entanto, a extensão é importante em bases de dados e na recuperação, pois resumos maiores também podem fornecer uma quantidade maior de palavras-chave para extração automática. Assim, é importante que ele não seja longo demais, mas um resumo extremamente curto não será eficaz para a representação dos conteúdos do documento.

Entende-se que o número de palavras usadas para a elaboração de um resumo num artigo científico deve ser estabelecido nas regras de submissão do periódico. Esse número é bastante variável, mas as normas para elaboração de artigos (dentre elas a NBR 6022 e NBR 6028, já citadas neste estudo) apontam para uma quantia entre 100 e 250 palavras em artigos de periódicos e entre 50 e 100 palavras em indicações breves. Esse padrão é o que foi considerado nesta análise.

Nas três sub-áreas analisadas aqui, a que apresentou resumos mais longos foi a de Antropologia (média de 119 palavras por artigo), seguida pela Sociologia (média de 110 palavras por artigo) e, por fim, Ciência Política (107,6 palavras por artigo). Foi também a área que empregou mais vezes resumos com mais de 100 palavras (66% em relação ao total de 32 resumos avaliados da área). Em Sociologia esse percentual foi de 53% dos 58 resumos avaliados, e em Ciência Política, 52% (dos 61 resumos avaliados). Esses dados parecem estranhos se forem confrontados com os do tópico anterior, já que resumos indicativos (mais comuns em Antropologia) tendem a ser mais curtos. Isso pode apontar para o fato de que a área emprega uma quantidade maior de palavras, mas ainda assim não cobre todos os aspectos relevantes do documento, indicando algum grau de redundância. Contudo, somente uma análise semântica da linguagem empregada nos resumos poderia demonstrar isso com efetividade, o que não foi feito neste estudo.

Discrepâncias como resumos curtos demais (abaixo de 50 palavras) ou com excesso de palavras (acima de 250) também foram observadas. No caso de resumos muito curtos, houve 2 ocorrências em Antropologia, 4 em Sociologia e 4 em Ciência Política. Apenas 1 resumo excessivamente longo foi observado, em Sociologia (273 palavras). Tanto um como outro podem ter impactos negativos em termos de recuperação: um resumo curto não transmitirá todos os tópicos relevantes do artigo, e um resumo muito grande incorre no risco de não ser suficientemente objetivo, perdendo uma de suas principais características: o poder de condensação.

4.1.3 Elementos para identificação de conteúdo

Neste ponto, foi avaliado o quanto o resumo deixava antever as partes principais do artigo, ou seja, objetivos, métodos, resultados/conclusões, contexto de pesquisa e hipóteses. A tabela abaixo mostra a quantidade de ocorrência de cada item, e deve ser vista sempre em relação ao total de resumos coletados de cada área (Antropologia: 32; Sociologia: 58; Ciência Política: 61).

Tabela 2: Identificação de conteúdo/estrutura

	Antropologia	%	Sociologia	%	Ciência Política	%
Objetivo	32	100	56	97	59	97
Método	28	87	45	78	48	79
Resultado	12	38	21	36	27	44
Discussão	0	0	0	0	3	5
Contexto	11	34	22	38	19	31
Hipóteses	1	3	6	10	8	13

De todos os resumos analisados, a única subárea que apresentou uma indicação clara de objetivos foi a de Antropologia, em que 100% dos 32 artigos possuíam uma referência ao objetivo do estudo apresentado. Essa incidência foi levemente inferior tanto Sociologia quanto Ciência Política, em que o objetivo ficava evidente em 97% da amostra coletada em cada área. Isso implica dizer que em cerca de 3% dos resumos avaliados em Sociologia e Ciência Política o objetivo não estava expresso de forma clara. Em ambos os casos, o que se observou foram resumos que descreviam unicamente o contexto da pesquisa, mas não mencionavam os propósitos ou o método empregado, comprometendo as principais funções do resumo para representação e recuperação de informações.

No entanto, é interessante ressaltar que a descrição do contexto de pesquisa é importante nas Ciências Sociais, e apareceu de forma recorrente nas três disciplinas constituintes da área. Alguns resumos apontados aqui como descriptivos incluíam, na verdade, objetivos e contexto, e não objetivos e métodos, conforme a as características comumente apontadas para resumos descriptivos em normas e na literatura.

Descrição dos procedimentos metodológicos ocorreram com mais prevalência na Antropologia (87% da amostra de 32 resumos), mesmo considerando a diferença de amostragem em relação às outras duas disciplinas. Em Sociologia e Ciência Política esse tópico teve um número muito aproximado de ocorrências (78% e 79%, respectivamente). Vale destacar que, tanto em Sociologia quanto em Ciência Política, houve uma incidência razoável (22% e 21%, respectivamente) de resumos que não possuíam descrição de procedimentos metodológicos. Nas Ciências Sociais isso pode trazer consequências ruins para a seleção de informações por parte do leitor, já que o entendimento do objeto teórico nessas áreas pode mudar de acordo com a metodologia empregada, sendo, portanto, fundamental que essa informação esteja presente nos resumos. No entanto, vale reforçar a idéia apresentada anteriormente em relação à importância do contexto em Ciências Sociais, já que em alguns dos resumos a caracterização da metodologia foi substituída pela descrição do contexto de pesquisa.

Resultados e conclusões apareceram mais vezes em Antropologia e Ciência Política (em porcentagem, 38% e 44%, respectivamente) e um pouco menos em Sociologia (36%). Apesar do número de resumos informativos ser mais freqüente nas outras disciplinas, em Antropologia observou-se alguma menção aos resultados/conclusões devido ao emprego de resumos descriptivo-informativos. No entanto, deve-se lembrar que esse tipo de resumo trata esse ponto de forma superficial. A menção de tópicos relativos à discussão (3 ocorrências em Ciência Política, totalizando 5% do total de 61 resumos avaliados da área) e de hipóteses (10% e 13% em Sociologia e Ciência Política, contra apenas 3% em Antropologia) pode indicar uma tentativa ou uma necessidade de caracterização mais pormenorizada das pesquisas nas duas primeiras disciplinas.

Apesar da diferença de amostragem, é possível dizer que grande parte dos resumos em Ciências Sociais caracteriza o texto apenas de forma superficial, excluindo outros pontos importantes para que o leitor possa avaliar a pertinência do artigo, como resultados obtidos (o que seria essencial em pesquisas empíricas) ou conclusões, importantes tanto em estudos experimentais quanto teóricos. Ainda que uma análise específica das normas de submissão de cada periódico não tenha sido feita, pode-se presumir que não há uma preocupação muito grande por pelo menos uma parte dos editores e mesmo dos autores em explorar o resumo como alternativa de representação e de recuperação de informações. Com isso, as funções do resumo como formas de “substituição” do documento principal (todas as informações relevantes devem estar no resumo) e “orientação” (suporte àqueles que lêem, como eixo dos conteúdos do texto) não são plenamente aproveitadas, o que, em última análise, tende a prejudicar a recuperação.

4.2 Palavras-chave como complemento e representação de relações entre áreas

É necessário esclarecer que nem todos os artigos cujos resumos foram avaliados aqui possuíam palavras-chave, e assim o corpus do qual foram extraídas essas palavras foi um pouco reduzido (29 artigos em Antropologia; 52 em Ciência Política; e 53 em Sociologia). Isso deixa antever diferenças nos critérios de certos periódicos, que parecem não ter uma atenção especial a esse detalhe.

O número de palavras-chave adotadas por artigo foi muito variável, mas em nenhuma área foi observado um número inferior a três (com exceção de 1 ocorrência em Ciência Política, com 2 termos). A maior parte dos artigos apresenta entre quatro e cinco palavras-chave (médias de 4,18 em Antropologia; 4,49 em Sociologia; e 4,28 em Ciência Política). Como já era previsto, pelo fato de não contar com vocabulário controlado, não foi observada uma padronização na apresentação da maior parte dos termos (uso de siglas, nomes não padronizados e frases como palavras-chave foram observados); no entanto, houve 18 artigos em Antropologia (em porcentagem, 62%) que apresentaram alguma palavra-chave que coincidia com a de outro artigo da mesma área. O mesmo também foi observado na Sociologia e na Ciência Política, mas com menor freqüência (16 e 20, respectivamente, totalizando 30% para Sociologia e 38% para Ciência Política).

É claro que não se deve tomar esse ponto como critério único para avaliar níveis de relacionamento entre os artigos, e muito menos para medir a “qualidade” da recuperação, pois a simples coincidência de termos não quer dizer que eles de fato sejam sinônimos. No entanto, a coincidência de palavras pode apontar para a existência de um vocabulário comum que não estaria evidente apenas com a leitura dos resumos, e nesse sentido palavras-chave podem ser vistas como um complemento destes. Além disso, em bases de dados, palavras-chave fornecidas pelos autores também funcionam como metadado descritivo, especialmente em bases que não possuem vocabulário controlado. Assim, o fato de haver palavras-chave coincidentes afetará os resultados de busca em maior ou menor escala.

Um exame mais detalhado de palavras-chave também pode auxiliar na identificação de outras características relevantes, como o tipo de pesquisa (teórica, empírica) e um contexto em que o estudo se insere, cuja necessidade de descrição já foi vista nos resumos. Conforme foi possível observar nos dados coletados, o emprego de nomes próprios exercendo a função de descritor foi encontrado em todas as disciplinas abordadas. O uso desses termos dá uma pista sobre o tipo de pesquisa desenvolvida no artigo ou sobre o seu contexto, já que pode indicar:

- a) estudos críticos sobre um autor específico, ou sobre conceitos relacionados a esse autor: citação de nomes como “Tocqueville”, “Gramsci”, “Adorno” e “Kant” reforçam as informações do resumo em artigos teóricos, particularmente quando o resumo é descritivo e não passa alguma idéia de contexto de pesquisa ou sobre sua natureza teórica. Isso foi observado apenas 1 vez na Antropologia (3% do total), mas 8 artigos da Sociologia (15% do total) e 3 (6% do total) da Ciência Política citaram nomes próprios de autores entre suas palavras-chave. A ocorrência desse tipo de descrição passa a idéia de uma pesquisa tendo como base teorias ou métodos desenvolvidos por teóricos específicos. Considerando a influência que determinas correntes de pensamento exercem sobre a pesquisa em Ciências Sociais, esse tipo de palavra-chave é uma forma de o autor do artigo situar o seu próprio estudo dentro de uma corrente específica;
- b) menção a personalidades que atuam num determinado contexto: observada particularmente na área de Ciência Política, em que termos como “Hugo Chávez”, “governo Lula” ou simplesmente “Lula” foram observados em 3 ocasiões (6% do total), e situam o artigo dentro de um contexto local (no caso, de um país específico) e temporal (século XXI, por exemplo);

c) menção a lugares: a citação de países, cidades ou mesmo regiões como palavra-chave ocorreu em 3 artigos da Antropologia (10%), 8 em Sociologia (15%), e 17 na Ciência Política (33%). Quando associadas a outros termos, essas palavras não apenas dão uma dimensão de um contexto local como também delimitam o universo de pesquisa. É importante considerar que o fator local realmente desempenha um papel importante nas Ciências Sociais, já que o objetivo dessas disciplinas é voltado à análise de uma realidade social cujas variáveis se modificam de região para região.

Além de nomes próprios, foi registrada em alguns artigos a necessidade de informar datas como palavras-chave, como “anos 1920”, “Brasil 1988-2006” e “Século XX”, particularmente em Sociologia e Ciência Política. A exemplo dos tópicos anteriores, dar uma noção de tempo ajuda a situar o universo da pesquisa, e também funciona como delimitador.

4.3 Relações entre disciplinas

As Ciências Sociais têm uma constituição multidisciplinar por natureza, o que já fica expresso na divisão aqui adotada, em Antropologia, Ciência Política e Sociologia. No entanto, uma avaliação das palavras-chaves permitiu ilustrar um nível amplo de relacionamento com áreas que, em princípio, não fariam parte das Ciências Sociais.

Na Antropologia, foi possível identificar relações com a Literatura, Artes, Comunicação, Educação e História. Como exemplos mais claros de palavras-chave empregadas: “literatura” (como termo isolado); “drama”, “teatro”, “performance” e “antropologia da performance”, em relações com termos da área de Teatro; “jornalismo”, “televisão” e “telenovelas”, no âmbito da Comunicação. No total, 8 artigos de Antropologia apresentaram termos dos quais se podiam fazer referências claras a outras áreas, totalizando 28% da amostra dessa área.

A área de Sociologia foi a que apresentou conexões com o maior número de disciplinas. Em 12 artigos (23% da amostra) foi possível perceber conexões com a Literatura, Artes, Ciência e Tecnologia, Música, Linguística, Cinema, Filosofia, e História, no uso de termos como “literatura brasileira” ou “literatura argentina” (Literatura), “Decoração” (Artes), “Schumann” e “música erutida” (música), “biotecnologia” e “inovação tecnológica” (Tecnologia), “cinema documental” e “epistemologia da imagem” (Cinema), “Adorno”, “dialética” e “materialismo” (Filosofia) e “história social” (História).

Das três áreas, foi a Ciência Política que menos apresentou relações óbvias com outras disciplinas. As áreas mais recorrentes foram Direito, Administração, Economia e Educação. A área de Filosofia também é importante, sobretudo no emprego de termos que são compartilhados entre Ciência Política e Filosofia Política (uso de palavras como “pragmatismo” e “liberalismo”, apenas para citar dois exemplos).

Esta análise não teve a intenção de ser profunda a ponto de dizer com que especificidade os relacionamentos entre disciplinas acontecem nas Ciências Sociais, nem julgar o grau de pertinência das palavras-chave atribuídas pelos autores. Contudo, quando se usa palavras-chave para tentar enxergar relações amplas entre disciplinas, elas podem trazer uma contribuição. Nesse sentido, é possível dizer que as palavras-chave em Ciências Sociais podem complementar o resumo quando este não for suficientemente informativo e também situar o artigo num contexto, inclusive em relação a outras disciplinas. Esse aspecto só será plenamente contemplado, porém, se houver um mínimo de padronização no uso dos termos, sobretudo na exploração de nomes próprios e de termos referentes a locais ou época de pesquisa. A elaboração de índices ou cabeçalhos de nomes próprios é uma alternativa bastante conhecida na Biblioteconomia e mesmo na Editoração, e poderia ser aplicada com bastante pertinência nesses casos.

5 CONSIDERAÇÕES FINAIS

Resumos e palavras-chave são relevantes como elementos de representação e recuperação em Ciências Sociais, mas não são explorados adequadamente pela área. Ao longo do levantamento de resumos para esta pesquisa, foi possível notar que não existe um padrão dentro das próprias disciplinas quanto ao uso de elementos pré-textuais: ora havia artigos com resumos e palavras-chave, ora artigos apenas com resumos, ora artigos sem resumo. As palavras-chave fornecidas pelos autores são úteis para as funções que foram apontadas neste texto, ou seja, como complemento do resumo e como forma de relacionar o estudo dentro de um contexto e também com outras áreas; no entanto, é difícil julgar a pertinência dessas palavras devido à falta de padronização observada. Esse fato ressalta a importância da indexação propriamente dita, com um processo de análise conceitual e tradução terminológica, e a importância de maior padronização de vocabulário quando do uso de palavras-chave

Considerando resumo e palavra-chave como metadado em sistemas de informação ou mesmo na *web*, pode haver um impacto negativo em termos de recuperação em Ciências Sociais, já que os elementos não são plenamente aproveitados com esse fim. Ainda que muitas bases de dados utilizem recuperação de texto completo, índices e outros mecanismos inclusive para indicar textos semelhantes e oferecer ao leitor resultados mais relevantes, o processo de seleção poderá ser prejudicado, já que a função de elementos pré-textuais é justamente amparar o leitor na seleção dos artigos. Um maior rigor por parte dos editores e dos próprios autores na elaboração desses elementos certamente trará contribuições à área.

Por fim, o uso resumos e palavras-chave como elementos de representação e recuperação de informações no contexto da pesquisa científica ainda carece de estudos, principalmente em áreas altamente interdisciplinares, como é o caso das Ciências Sociais. Mais pesquisas sobre o tema são necessárias, e trarão uma dimensão maior tanto da terminologia empregada por essas ciências quanto do potencial da Ciência da Informação no que se refere a pesquisas voltadas à comunicação científica.

REFERÊNCIAS

ANDERSON, James D.; PÉREZ-CARBALLO, José. The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part I: research and the nature of human indexing. **Information processing and management**, v. 37, p. 231-254, 2001a.

_____, _____. The nature of indexing: how humans and machines analyze messages and texts for retrieval. Part II: Machine indexing, and the allocation of human versus machine effort. **Information processing and management**, v. 37, p. 255-277, 2001b.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6028**. Informação e documentação: resumo: apresentação. Rio de Janeiro: 2003a.

_____. **NBR 6022**. Informação e documentação: artigo em publicação periódica impressa: apresentação. Rio de Janeiro, 2003b.

BOGER, Zvi et al. Automatic keyword identification by artificial neural networks compared to manual identification by users of filtering systems. **Information processing and management**, v. 37, p. 187-198, 2001.

BORKO, Harold; BERNIER, Charles L. **Abstracts concepts and methods**. New York: Academic Press, 1975.

- _____; _____. **Indexing concepts and methods.** New York: Academic Press, 1978.
- COOK, David A.; BECKMAN, Thomas J.; BORDAGE, Georges. A systematic review of titles and abstracts of experimental studies in medical education: many informative elements missing. **Medical education**, v. 41, p. 1074-1081, 2007.
- CROSS, Cate; OPPENHEIM, Charles. A genre analysis of scientific abstracts. **Journal of documentation**, v. 62, n. 4, p. 428-446, 2006.
- EAST, John W. Subject retrieval of scholarly monographs via electronic databases. **Journal of documentation**, v. 62, n. 5, p. 597-605, 2006.
- ERCAN, Gonenc; CICEKLI, Ilyas. Using lexical chains for keyword extraction. **Information processing and management**, v. 43, p. 1705-1714, 2007.
- FERNANDES, Helder et al. A investigação em ciências sociais. Aproximação ao contexto da Educação Física e Desporto. **Lecturas: educación física y deportes**, v. 8, n. 53, oct. 2002. [documento eletrônico]. Disponível em: <http://www.efdeportes.com/efd53/sociais.htm> Acesso em: 17 ago. 2008.
- GIL-LEIVA, Isidoro; ALONSO-ARROYO, Adolfo. La relación entre las palabras clave Aportadas por autores de artículos de revista y su indización en las bases de DATOS ISOC, IME E ICYT, **Revista española de documentación científica**, v. 28, n. 1, 2005.
- _____; _____. Keywords given by authors of scientific articles in database descriptors. **Journal of the American society for information science and technology**, v. 58, n. 8, p. 1175-1187, 2007.
- HARTLEY, James. Is it appropriate to use structure abstracts in non-medical science journals? **Journal of information science**, v. 24, n. 5, p. 359-364, 1998.
- _____; KOSTOFF, Ronald N. How useful are 'key words' in scientific journals? **Journal of information science**, v. 29, n. 5, p. 433-438, 2003
- HOUAISS, Antônio. Palavra-chave. In: _____. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2001, p. 2108
- JOHNSON, Frances. Automatic abstract research. **Library review**, v. 44, n. 8, p. 28-36, 1995.
- LANCASTER, F. W. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos/Livros, 1993.
- MEADOWS, Arthur Jack. **A comunicação científica**. Brasília, DF: Briquet de Lemos/Livros, 1999.
- MONTESI, Michela; URDICIAIN, Blanca Gil. Abstracts: problems classified from the user perspective. **Journal of information science**, v. 31, n. 6, p. 515-526, 2005.
- NUNES, A. Sedas. **Questões preliminares sobre as ciências sociais**. 13. ed. Lisboa: Editorial Presença, 2005.

OU, Shyian; KHOO; Christopher Soo-Guan; GOH, Dion H. Design and development of a concept-based multi-document summarization system for research abstracts. **Journal of information science**, v. 34, n. 3, p. 308-326, 2008.

PINTO, Maria. Paradigms for abstracting systems. **Journal of information science**, v. 25, n. 5, p. 365-380, 1999.

_____; LANCASTER, F. W. Abstracts and Abstracting in Knowledge Discovery. **Library Trends**, v. 48, n. 1, p. 234-248, Summer 1999.

_____. Abstracting/abstract adaptation to digital environments: research trends. **Journal of Documentation**, v. 59, n. 5, p. 581-608, 2003.

_____. A grounded theory on abstracts quality: Weighting variables and attributes. **Scientometrics**, v. 69, n. 2, p. 213-226, 2006.

TIBBO, Helen R. **Abstracting, information retrieval and the humanities:** providing access to historical literature. Chicago: ALA, 1993. (ACRL Publications in Librarianship, 48).

VOORBIJ, Henk. Title keywords and subject descriptors: a comparison of subject search entries of books in the humanities and social sciences. **Journal of documentation**, v. 54, n. 4, set. 1998.

WAARD, Anita de; KIRCH, Joost. Metadata in science publishing. [documento eletrônico]. Disponível em: <http://wwwis.win.tue.nl/infwet03/proceedings/8/> Acesso em: 25 mai. 2008.

ABSTRACT

This study aims to make an evaluation of abstracts and keywords used in periodicals published by Brazilian authors in the area of Social Sciences, in order to identify the main characteristics of these practices and how they can affect the information retrieval and representation. About 151 abstracts written by professors of the University of São Paulo were evaluated, with a view to discovering how the authors use the types of abstracts (descriptive, informative or descriptive-informative); As a second topic, we examined keywords, in an attempt to analyse how these keywords complement the abstracts, and to see if keywords can be used as tools to represent concepts in related disciplines. The results showed that the Brazilian social scientist prefer descriptive abstracts, and that, in spite of the lack of standards, the keywords they supply can be used as a complement to the abstracts, and as indicators of interdisciplinary relations. Our conclusion is that, in spite of the importance of abstracts and keywords as informational tools, they are not properly used in the Social Sciences, and this fact may have negative effects on information retrieval.

KEYWORDS: Abstracts. Keywords. Metadata. Information representation. Information retrieval.

Originais recebidos em: 11/06/2008

Texto aprovado em: 02/09/2008