ORGANIZAÇÃO E DISSEMINAÇÃO DA INFORMAÇÃO ESTATÍSTICA

SILVIA G. ROCHA[1]
sigrocha@seade.gov.br
MARILDA L. G. DE LARA[2]
larama@usp.gov.br

Reconhecendo a informação estatística como representação e medida de realidades sociais, destaca-se o papel da Ciência da Informação na produção de informações que facilitam o acesso às estatísticas, cumprindo assim, não só o papel de distribuir, mas também o de mediar o estoque e o sistema de informações, criando condições de uso de uma informação bastante complexa e propiciando a construção do conhecimento. Trabalha-se com o cenário tecnológico atual, considerando a Internet como principal instrumento de disseminação. Nesse contexto, propõe-se o uso da linguagem documentária para viabilizar a transferência da informação nesse novo modelo de relacionamento com o usuário, onde não existe a intermediação presencial do documentarista. Consideram-se as especificidades da informação estatística, cujas características a diferenciam da informação textual, o que leva a concluir que nem todas as recomendações para construção de linguagem documentária para representar esse tipo de informação se aplicam necessariamente à representação da informação estatística.

Palavras-chave: linguagem documentária; indexação; representação documentária; disseminação da informação; informação estatística.

[1] Mestre em Ciências da Informação (ECA/USP). Gerente de Atendimento e Disseminação de Informação (Fundação Seade)
[2] Doutora em Ciências da Comunicação, Docente (ECA/USP)

ORGANIZAÇÃO E DISSEMINAÇÃO DA INFORMAÇÃO ESTATÍSTICA

1. Introdução

A informação estatística tem sua origem ligada ao Estado e à necessidade de inventariar bens, produções e pessoas, com fins arrecadatórios e militares, pois era preciso saber quantos soldados estavam disponíveis para os nem tão eventuais conflitos, havendo, portanto, mil justificativas para se contar e controlar essa população. Naturalmente, essas contagens eram rudimentares e não necessariamente exaustivas.

Nos dias atuais, considera-se que as estatísticas representam uma leitura da configuração social criada para representar uma realidade. Sob esse ponto de vista, dado estatístico é uma tradução resultante de um trabalho de conceituação, de organização e observação.

No Brasil, 1870 foi o ano da criação da Diretoria Geral de Estatística, que coordenou as primeiras atividades sistematizadas, ligadas a levantamentos censitários, culminando em 1872, com o primeiro Recenseamento Geral do Império do Brasil. Em 1934 foi criado o Instituto Nacional de Estatística, que só existiu de fato em 1936, mudando em 1938 para Instituto Brasileiro de Geografia e Estatística – IBGE.

A produção das estatísticas no Brasil continua fortemente apoiada na produção do IBGE. A principal fonte de dados estatísticos é o Censo Demográfico, realizado em intervalos decenais em todo o território nacional. Além das pesquisas censitárias, as estatísticas vêem sendo produzidas não só pelo IBGE, mas também pelas agências estaduais de estatísticas, através de registros administrativos ou de pesquisas amostrais, concebidas para prover o país de informações demográficas e socioeconômicas, nos intervalos censitários, ou de pesquisas com temática específica como as que medem o mercado de trabalho e uma série de outras investigações realizadas com o intuito de aprofundar a abrangência temática pesquisada e complementar cronologicamente o levantamento censitário.

Ao analisar a estatística produzida no país, desde o começo do século XX, percebe-se claramente uma alteração gradativa quanto ao tipo de informação pesquisada. A produção das estatísticas, nos dias de hoje, é muito diversificada nos quesitos que são investigados e muito detalhada na abrangência geográfica solicitada. Do mesmo modo, nos últimos anos a demanda por informações conhece um acentuado crescimento, que direta ou indiretamente afeta as instituições produtoras de informação estatística.

A partir do momento que essa informação passa a ser demandada por variados segmentos da sociedade, e não só por especialistas, e que os progressos das tecnologias de informação e comunicação começam a ser utilizados por um público muito mais amplo, as instituições produtoras de estatísticas vêem-se frente à tarefa de conseguir disponibilizar suas informações de forma rápida, com baixo custo e com facilidade de acesso. A esse enfrentamento, a Internet se destaca como grande ferramenta para solucionar questões de custo, velocidade de transmissão de grandes volumes de dados e facilidade de acesso.

Mas a discussão que se coloca no momento não está focada no ato simples de tornar públicas as informações. É preciso responder uma série de indagações:

> "Quem utiliza as informações disponibilizadas? Com que facilidade e nível de compreensão? Qual é o significado dessas informações para o público? Qual é a relação entre as informações divulgadas e as demandas informacionais da população? Que aplicação é feita dessas informações? (...) Perguntas do gênero devem ser feitas para que se possa aquilatar a extensão das realizações concretas de disseminação, e muitas delas são válidas mesmo nas situações em que são utilizados meios tradicionais de disseminação" (Lara & Conti, 2003, p. 27).

Mesmo enquanto essas questões estão sendo debatidas, já é possível detectar que a forma de disponibilização da informação estatística pode ser melhorada se usados os recursos tecnológicos disponíveis, para facilitar seu acesso.

No final da década de 1990, a Internet transformou-se no meio principal de disseminação, mas de uma forma geral transpôs-se para o ambiente web a publicação tal qual era no papel. Verifica-se que nos casos em que a informação foi reformatada para se adequar ao novo ambiente, a forma anterior de organização das informações foi mantida, permanecendo, assim, todas as dificuldades de localização da informação e acrescentando-se um sério agravante: o usuário da Internet não conta com um mediador para auxiliá-lo na busca da informação. A isso, soma-se o fato das instituições produtoras de informações estatísticas não terem a prática de organizar estatística para o público leigo.

Além das dificuldades advindas da organização da informação, existe também aquela originada da prática de disponibilizar informação estatística em tabelas, adotada há mais de um século. Embora contando hoje com os recursos trazidos pela informática, essa prática persiste sem que os responsáveis pela formatação da informação percebam o quanto este formato dificulta a busca e a recuperação da informação.

Para facilitar o entendimento dessa colocação, podemos exemplificar com um produto, disponível na Internet, em que se destaca o fato de ser um produto que disponibiliza

informações sobre população, saúde, educação, economia, etc. desagregadas para todos os municípios de um determinado Estado.

Apesar das informações estarem desagregadas por municípios, não se obtém acesso rápido e fácil às informações desejadas se o produto foi concebido em forma de tabelas. Para recuperar o perfil de um município, é necessário abrir todas as tabelas que compõem o produto, anotar a informação que cada uma delas traz sobre o município em estudo e, só depois desse demorado, oneroso e complicado processo é que se consegue a informação demandada.

Se o produto tiver sido concebido como um banco de dados, com as informações disponibilizadas como variáveis, esse processo será realizado com a seleção das informações e do município desejado, e, ai sim, a tabela será montada, com as informações escolhidas pelo usuário.

Chamamos genericamente de variável todas as medidas encontradas nas diversas pesquisas ou consolidadas através dos inúmeros registros administrativos, e também por extensão, indicadores ou índices resultantes das relações entre essas medidas.

"O termo variável tem uma história ambígua. Originário da matemática e da física teórica, o conceito de variável foi apropriado pelas ciências sociais, tomando um sentido cada vez mais amplo. Pouco a pouco foi estabelecendo-se a idéia ou o sentido de compreender, no conceito de variável, o resultado da separação de conjuntos de objetos selecionados, conforme um ou vários critérios específicos, em função de características determinadas, como, por exemplo, sexo, idade etc" (Dencker & Da Viá, 2001, p.127).

Em outras palavras, variável é um fenômeno medido que ganha sentido quando localizado no tempo e no espaço, o que lhe confere uma dimensão. A informação estatística não existe fora dessa tríade. Sempre é composta por um acontecimento mensurável, num determinado lugar, em determinado período.

Quando se pensa em disponibilização de informação estatística em bancos de dados que reúnem variáveis e não tabelas, o uso de uma linguagem de organização de informações seria um auxiliar para a comunicação entre o banco de dados e o público usuário.

Naturalmente a linguagem documentária sozinha não pode resolver todos os problemas de recuperação de informação. Entretanto, trabalhamos a partir da hipótese de que a representação e a recuperação podem ser melhor solucionadas a partir do uso das metodologias de construção de Linguagens Documentárias e das metodologias da Terminologia, historicamente utilizadas na representação de informação textual.

A inclusão de novas tecnologias ao nosso cotidiano e a expansão acelerada da Internet propiciou uma facilidade muito grande na produção e disseminação de informação. Aliadas ao baixo custo e a possibilidade de divulgar amplamente a informação, a Internet tornou-se a mídia preferida.

Vale lembrar, no entanto, que o uso de um bom veículo não significa por si só uma boa transmissão de informação. Não basta disponibilizar a informação, é necessário criar condições para a que a informação seja apropriada e transformada em conhecimento.

O atendimento a esse quadro demanda a criação de mecanismos para proporcionar acesso mais fácil e livre às informações, como também reclama a elaboração de produtos que sejam adequados a distintas necessidades de diferentes tipos de usuários.

2. CONSTRUÇÃO DA LINGUAGEM DOCUMENTÁRIA PARA DADOS ESTATÍSTICOS

Foi com a preocupação de propiciar aos usuários um acesso mais fácil e livre às informações estatísticas, que a Fundação Seade inseriu, nos procedimentos de preparação dos seus produtos a serem divulgados na Internet, uma etapa de tratamento documentário, para que a disponibilização de seus dados se adequassem às necessidades de usuários distintos e para que eles pudessem contar com meios de acessar, muito rapidamente, informações complementares necessárias à compreensão da informação disponível.

A adoção de procedimentos da Análise Documentária – disciplina de caráter metodológico própria para tratar, organizar e disponibilizar a informação – e as metodologias de organização de linguagens documentárias balizaram o trabalho, quer em relação ao desenvolvimento de uma linguagem propriamente dita, quer no que diz respeito à indexação das informações. No entanto, como veremos mais à frente, foi possível constatar diferenças entre o tratamento do texto e o da informação estatística. Chamamos de linguagem documentária, a linguagem construída com o fim de representação das informações que tem como princípio o controle das relações entre os termos. A maior parte das linguagens documentárias está organizada de maneira lógico-semântica. Para efeito de contextualização procuramos, no quadro a seguir, sintetizar brevemente o quadro das relações entre os termos, suas funções e símbolos de representação.

Quadro 1 – Relações básicas de um vocabulário controlado

Tipo de Relação	Definição	Finalidade do Uso	Representação
Equivalência	Relação entre o termo preferido e o não preferido, utilizado quando dois ou mais termos são considerados, para fins de indexação, como referentes ao mesmo conceito.	Indicar qual o termo preferido ou preterido pelo sistema no caso de sinônimos, quase sinônimos ou mesmo no caso de deslocamentos genéricos.	**USE** (indica o termo preferido) **UP** (**u**sado **p**ara - indica os termos considerados equivalentes)
Hierárquica	Relação baseada em graus ou níveis de superordenação/ subordinação: Gênero/Espécie ou relação genérica, ou baseada em níveis de ordenação: Todo/ Parte ou relação partitiva.	Indicar uma noção fundamental que inclui noções específicas. Desse modo, as relações genéricas/partitivas indicam que todo conceito específico é parte de um conceito amplo.	**TG** (termo genérico) **TE** (termo específico)
Associativa	Relação baseada em associações de sentido, num determinado contexto. A ocorrência e utilidade de qualquer associação dependem do modo de organização dos domínios de especialidade e das necessidades do usuário.	Indicar pares de termos que não pertencem à mesma hierarquia, mas que se inter-relacionam. Devem ser utilizados quando a consulta às informações indexadas pelo outro termo complementarem a pesquisa.	**TR** (termo relacionado)

Além do resumo apresentado nesse quadro, podem destacar-se alguns outros recursos, como:

a) Deslocamento genérico – se refere a termos considerados demasiado específicos para adoção como termos de indexação, mas que podem ser mapeados com os termos mais gerais e mais adequados ao sistema.

b) Notas explicativas - dizem respeito ao dispositivo que reduz a ambigüidade na interpretação de um termo. São geralmente identificadas com o prefixo NE no vocabulário. Pode ser uma verdadeira definição de dicionário, pode indicar como um termo deve ser interpretado no contexto daquele vocabulário e /ou distingui-lo de outros termos com os quais possa ser confundido.

c) Definições – recursos utilizados para delimitar a intensão e ou a extensão do termo utilizado. A partir das referências da Terminologia, podemos fazer uma distinção entre as notas explicativas e as definições: as notas são mais adequadas para dirimir questões de aplicação do termo, já as definições se relacionam à circunscrição de seu significado no contexto utilizado.

A partir do momento que incluímos, no nosso sistema de linguagem controlada, um recurso como o do uso de definições, passamos a construir um vocabulário com contribuições oriundas da Terminologia e a primeira noção da terminologia que devemos considerar é a de

“conceito”, descrita pela norma ISO 1087-1:2000 como unidade de conhecimento constituída de:

a) síntese do objeto (elemento da realidade passível de ser percebido ou concebido);

b) predicados necessários a ele atribuídos;

c) signos lingüísticos para fins de comunicação.

O entendimento desta noção é o primeiro e mais fundamental passo para a construção de linguagens documentárias: trabalhamos com unidades de informação e sem o reconhecimento de que organizamos conceitos ou significados, não podemos determinar corretamente quais signos são seus representantes. Da mesma forma, sem o conhecimento das características a eles atribuídas, não os definimos, como também não construímos redes de relações consistentes, o que significa não cumprir o papel de possibilitar a organização e a comunicação documentária. Levando-se em consideração que a linguagem documentária exerce papel de intermediação entre o sistema de informação e os usuários, devemos prever formas de explicitar as noções envolvidas em cada unidade de informação e suas relações, para tentar garantir ao indexador, usuário desse instrumento, e ao receptor, destinatário final do sistema de informações, as mesmas condições para compreender a lógica utilizada na construção da linguagem documentária. Essa necessidade nos leva à segunda compreensão importante das funções dos instrumentos terminológicos: o de fornecer meios para o estabelecimento das relações entre os conceitos. Os conceitos são delimitados a partir das características a eles atribuídos. Uma característica é um aspecto a partir do qual se realiza a representação de um objeto. A identificação de características permite delimitar os conceitos e através de suas oposições e similitudes, as relações entre eles.

As relações entre os conceitos dão forma a um sistema de conceitos, composto principalmente de: conceito genérico; conceito específico; conceito partitivo; conceito coordenado; conceito associado.

Uma vez determinadas as relações entre os conceitos, chegamos à terceira questão importante relativa à apropriação das contribuições da Terminologia para a construção da linguagem documentária: a definição.

A definição é, segundo a norma ISO 1087-1:2000, um “Enunciado que descreve um conceito, permitindo diferenciá-lo nitidamente dos demais conceitos dentro do sistema”.

Na realidade, o processo de identificação de características dos conceitos e do estabelecimento das relações entre eles, faz parte do processo definicional. A fase ora especificada compreende o processo de sua expressão lingüística, de forma a facilitar a

compreensão do conceito. O uso regular de definições é o instrumento que garante a explicitação da lógica da construção da linguagem documentária, apontada como necessária em parágrafo anterior.

No empreendimento de identificar princípios para elaborar um instrumento que represente e recupere dados estatísticos, verificamos que uma linguagem documentária com esse fim deve ser muito mais uma forma de organizar os dados eles mesmos para disponibilização do que, como ocorre na representação de textos, uma forma de prover sínteses dos textos representados. Isso acontece porque, como já dissemos anteriormente, os dados estatísticos são de natureza muito específica no seu recorte da realidade e, por conseguinte, sua representação deve ser realizada num nível mais particularizante do que aquele normalmente utilizado na representação de textos. Diferentemente do que ocorre com a representação de textos, qualquer tentativa de sintetizar as representações de dados estatísticos através, por exemplo, de pistas para recuperação, pode gerar expectativa de informação nem sempre passível de ser satisfeita.

A afirmação acima pode ser confirmada através de exemplos.

Ex.: "Óbitos por acidentes de trânsito"

Se desdobrarmos o enunciado acima e representarmos essa informação sob a rubrica "Acidentes de trânsito", podemos, dentro de um sistema estatístico, gerar no usuário a expectativa de encontrar, como resposta, o número de acidentes de trânsito ocorridos, informação que não necessariamente existe no sistema. É necessário, portanto, representar um dado estatístico num maior nível de especificidade, como por exemplo, sob o termo "Óbitos por acidentes de trânsito" que, nesse caso já é o próprio nome da variável. Cabe aqui a observação de que nem sempre o nome da variável é normalizado na apresentação dos dados, o que às vezes também provoca dificuldade no entendimento da informação. Aliás, esse é um outro aspecto que confirma a necessidade de utilizar os princípios metodológicos previstos para a construção de uma linguagem documentária.

No momento em que os termos utilizados no vocabulário se aproximam quase que literalmente dos nomes das variáveis, resta, para a linguagem documentária, a função de organização dessa informação de modo relacional, envolvendo todos os termos utilizados. Por essa razão, defendemos a hipótese de que, no que tange às linguagens de organização e representação da informação estatística, o nível de especificidade da linguagem coincide com o nível da especificidade do vocabulário utilizado para a indexação. Decorre daí que, diferentemente do que acontece com a indexação de textos, a indexação de dados estatísticos não pode ser generalizante.

É possível, porém, organizarmos conjuntos de informações, classificadas originalmente em temas diferentes, reunindo-as sob um único tema representativo de um recorte específico. Os temas (ou categorias) representam, pela sua própria natureza, um nível de generalização e abstração maior. Em outras palavras, sintetizamos a informação quando a reunimos com outras informações que, no seu conjunto, traduzem um recorte generalizante, como um cenário construído sob uma determinada ótica, tal como acontece, por exemplo, com o termo "Mulheres", onde se despreza a especificidade da informação para oferecer ao usuário uma síntese temática que tem seu significado num universo mais geral. Nesse contexto, as informações de "Óbitos femininos" como a informação de "Eleitores femininos" ou de "Gestação de adolescentes", ou ainda, "População feminina" ou a informação de "Mortalidade de mulheres em idade fértil", podem ser encontradas reunidas genericamente sob o termo Mulheres.

Mantendo bem clara a noção de que o trabalho documentário é por natureza institucional e, portanto, não deve pretender atingir o usuário individual, mas um usuário coletivo que é previsto pela instituição, sugerimos que sejam combinados os dados de produção e de recepção. Sob o foco da produção (emissão), deve-se procurar usar as categorias generalizantes de informação que são razoavelmente uniformes e compartilhadas pelas instituições estatísticas; sob o foco do usuário (recepção), é preciso procurar "acompanhar" as formas de solicitação, as dúvidas e perguntas feitas pelo usuário e introduzir tais referências na organização das informações, seja via uso de equivalências ou de poli-hierarquias, recursos que possibilitam que a linguagem de recuperação se aproxime da linguagem natural usada pelo usuário não especializado.

Utilizando os recursos oferecidos pelas metodologias de organização da linguagem documentária é possível facilitar a consulta por meio de termos utilizados na linguagem coloquial e, ao mesmo tempo, classificar essas informações na hierarquia adotada pela instituição produtora de estatísticas, que geralmente segue a estrutura utilizada pelo órgão responsável pelo registro público que gera os dados.

Além da estrutura de relações, que provê e assegura os significados e a interpretação dos termos, um sistema de recuperação de informação estatística deve contemplar o acesso a definições, pois mais do que dados são representados conceitos. As definições, além de auxiliarem o usuário leigo na tarefa de seleção e recuperação das informações, asseguram a própria estrutura da organização dos dados.

Outro recurso a ser inserido no sistema de recuperação são as notas explicativas que explicam o processo da produção dos dados. Isso auxilia, por exemplo, a compreender

possíveis incompatibilidades entre os totais apresentados para o Estado e a soma dos resultados para municípios, oriundos de problemas relacionados à forma de coleta de dados.

Do mesmo modo, pode ser necessário incluir notas sobre a comparabilidade de séries históricas, alertando sobre mudanças de denominação ocorridas ao longo do tempo, que nem sempre correspondem a alterações muito grandes de conceito.

Outra informação importante a ser disponibilizada é a fonte dos dados, que não só garante a credibilidade das informações como, aliada às notas e definições, explicam eventuais divergências entre informações coletas por pesquisas ou instituições diferentes.

Em síntese, além da indexação propriamente dita, que se distingue da textual, a apresentação de informações estatísticas requer elementos de contextualização como os expostos acima que funcionam como metadados. A produção de definições, notas explicativas e identificação de fontes requerem o trabalho conjunto de documentalistas e especialistas, de modo a garantir a consistência das informações.

O que foi dito acima confirma nossa hipótese de que, o processo de indexação de dados estatísticos é complexo e particular não se restringindo às operações metodológicas de representação utilizadas para o tratamento dos textos. Ao contrário, ele exige agregar, paralelamente, um conjunto de referências que permitam qualificar e contextualizar cada item de informação. Se na elaboração de uma linguagem para o tratamento da informação estatística são utilizados princípios de elaboração de tesauros e da Terminologia, tal como ocorre para o caso dos textos, no processo de indexação são agregadas novas tarefas que se relacionam ao cumprimento das funções contextualizantes e explicativas. Embora constituam processos de natureza diferente, normalmente eles ocorrem de forma paralela. Os dois processos se combinam de modo a garantir maior precisão nas operações de tradução conceitual, orientando readequações na formulação da linguagem. Um exemplo disso é o fato de que ass definições, notas e fontes são também instrumentos importantes utilizados na etapa de construção da linguagem, mesmo que não sejam, posteriormente, divulgados acompanhando a informação disponibilizada.

3. CONSIDERAÇÕES FINAIS

Como pudemos verificar, a informação estatística obedece a regras particulares. Comparada à informação textual, os dados estatísticos – expressões numéricas que se referem a aspectos da realidade – têm características peculiares que não podem deixar de serem

observadas para que se possam constituir sistemas de informação que possibilitem o acesso e a apropriação da informação.

É bastante incomum que esse gênero de informação seja submetido a tratamentos documentários, o que dificulta o trabalho de identificação de procedimentos que permitam sua melhor organização e disseminação, ocorrendo, inúmeras vezes, que as formas utilizadas para sua disponibilização se limitem às normas de apresentação tabular. No entanto, a experiência concreta com o atendimento ao usuário nas buscas por informação estatística evidencia a necessidade de parâmetros para sua organização e tratamento.

Recuperando a literatura documentária, pudemos observar que os princípios de organização e recuperação da informação têm na noção de linguagem suas bases fundamentais. A informação estatística, embora tenha características distintas da informação textual, também se constitui como linguagem, justificativa principal, portanto, para que nossas investigações tivessem aí sua origem.

Em síntese, considerando os aspectos da linguagem, o tratamento e a representação da informação estatística é necessário observar que:

1) O vocabulário deve conter termos muito específicos: cada termo praticamente corresponde a um dado indexado;

2) A linguagem documentária deve ter também a função de organização da informação;

3) Os conjuntos de informações estatísticas podem ser organizados em grandes temas, seguindo a organização tradicional desse tipo de informação, mas outras hipóteses de organização devem ser construídas: através das estruturas específicas, as informações devem contemplar diferentes arranjos, de modo a propor outros cenários informativos pertinentes;

4) É imprescindível construir redes de equivalências e de associações entre os termos com vistas a adequar a linguagem de organização às solicitações dos usuários, bem como a dirimir suas dúvidas;

5) É fundamental construir um conjunto de metadados: definições, notas e fontes são indispensáveis, tanto para a contextualização dos dados estatísticos (na emissão ou na recuperação da informação), quanto para assegurar a própria estrutura da organização dos dados;

6) É preciso ter claro que, mesmo entre os dados estatísticos, poderemos encontrar especificidades que precisam ser levadas em conta na construção da linguagem e no processo de indexação. Essas variações vão surgir de acordo com a forma com que a informação

estatística é disponibilizada: tabulações fixas, variáveis para construção de tabulações personalizadas ou microdados para a livre criação de indicadores.

O arrolamento acima não esgota as observações passíveis de serem feitas em relação às especificidades do trabalho com a informação estatística. Importante dizer que os esforços empreendidos têm de se somar para constituir metodologias. Se a Internet é um bom veículo para disseminar a informação estatística, todos os meios para a potencialização do uso devem ser mobilizados.

REFERÊNCIAS

AUSTIN, D; DALE, P. **Diretrizes para o estabelecimento e desenvolvimento de tesauros monolingües.** Trad. de Bianca Amaro de Melo e rev. de Ligia Maria Café de Miranda. Brasília: IBCTI; SENAI. 1993. 86p.

BARRETO, A. de A. A oferta e a demanda da informação: condições técnicas, econômicas e políticas. **Ciência da Informação**, IBICT, v. 8, n.2, 1999. Disponível em: <http://www.ibict.br>

BARRETO, A. de A. A questão da informação. **São Paulo em Perspectiva**, v.8, n.4, p.3-8, out/dez.1994.

CINTRA, A. M. M; TÁLAMO, M. F. G. M; LARA M.L.G; KOBASHI, N. Y. **Para entender as linguagens documentárias**. 2ed. rev. e ampl. São Paulo: Polis, 2002. 92p.

DENCKER, A. F. M; DA VIÁ, S. C**. Pesquisa empírica em ciências humanas: com ênfase em comunicação**. São Paulo**:** Futura, 2001.190p.

FERREIRA, S. P. Produção e disponibilização de estatísticas. **São Paulo em Perspectiva**, v.17 n.3-4, p.17-25 jul/dez.2003

ISO 1087-1:2000 (E/F). **Terminology work – Vocabulary, part 1: theory and application/ Travaux terinologiques – Vocabulaire – Partie 1: théorie and application**. 1.ed. Genève: International Standard Organization.

ISO 2788/BS5723 (1986/1987). **Guidelines to establishment and development of monolingual thesauri**. Genève; London: International Standard Organization; Britsh Standard.

JANNUZZI, P. M; GRACIOSO, L. S. Produção e disseminação de informação estatística: agências estaduais no Brasil. **São Paulo em Perspectiva**, v.16, n.3, p.92-103, jul/set. 2002

LARA, M. L. G. de. **Conceitos terminológicos para a construção de linguagens documentárias**. (versão preliminar, texto não publicado, 2002).

LARA, M. L. G; CONTI, V. L. . Disseminação da informação e usuários. **São Paulo em Perspectiva**, v.17 n.3-4, p.26-34, jul/dez.2003

NBR 12676: 1992. **Métodos para análise de documentos – determinação de seus assuntos e seleção de termos de indexação**. Rio de Janeiro, Associação Brasileira de Normas Técnicas – ABNT.

ROCHA, S. G. **A representação documentária de informações estatísticas: quando a ordem dos fatores altera o produto**. 2006. 90f. Dissertação (mestrado em Ciência da Informação). Escola de Comunicação e Artes. Universidade de São Paulo. São Paulo.

SENRA, N. de C. Informação estatística: demanda e oferta, uma questão de ordem. **DataGramaZero: Revista de Ciência da Informação**, v.1, n.3, jun.2000.

SMIT, J. W; BARRETO, A. A. Ciência da informação: base conceitual para a formação do profissional. In: VALENTIN, M. L, org. **Formação do profissional da informação**. São Paulo: Polis, p.9-23,2002.

TÁLAMO, M. F. G. M. **Linguagem documentária**. São Paulo: APB, 1997. 12p. (Ensaios APB, n.45).