

Novas possibilidades da pesquisa qualitativa via sistemas CAQDAS

ALEX NICHE TEIXEIRA E FERNANDO BECKER*

Apresentação



Argumentar sobre as possibilidades de operacionalização da pesquisa social quando utilizamos metodologia qualitativa via *CAQDAS* consiste em um desafio. Via de regra, os manuais de metodologia de pesquisa qualitativa são muito ricos ao mostrar as técnicas e dificuldades que enfrentamos nos trabalhos de campo ou análises de documentos. Obstáculos na interação face a face, técnicas de estranhamento, de observação, de análise de dados, etc. Mas os manuais de metodologia qualitativa, não trazem nenhuma informação direta sobre o uso dos pacotes para a análise de dados qualitativos. Devemos considerar também que tais pacotes são relativamente desconhecidos no Brasil, tendo seu uso difundido recentemente. Acreditamos que o uso desses programas torna-se imprescindível para a pesquisa sociológica devido à grande economia de recursos e tempo, mas também pela possibilidade de gerar análises mais ricas e aprofundadas.

A proposta deste artigo não é discutir técnica de pesquisa, mas apontar os recursos oferecidos por esses programas de análise de dados qualitativos e como funcionam. Trata-se de contribuir para o debate, ainda incipiente, sobre como operacionalizar a pesquisa sociológica utilizando esses pacotes, que servem como uma poderosa ferramenta para as diferentes técnicas de pesquisa qualitativa.

* Mestrandos do Programa de Pós-graduação em Sociologia da UFRGS

A partir disso, em um primeiro momento, mostramos a lógica de funcionamento dos dispositivos *CAQDAS*, que está fundamentada no princípio da codificação do texto. Logo após, apontamos os recursos de codificação oferecidos por ferramentas de busca automática em um destes programas: o NUD*IST. Em seguida, são demonstrados os mecanismos de exploração dos conteúdos codificados, que permitem realizar um refinamento da análise. Seguem considerações finais.

1 O que é CAQDAS ?

Kelle (1997, p.3) define o termo *CAQDAS* (*Computer-aided qualitative data analysis software*) como uma série de programas de computador orientados para o auxílio na análise de dados qualitativos¹. Esses programas começam a ser desenvolvidos na década de 80, essencialmente nos Estados Unidos e na Inglaterra, por e para pesquisadores da área de Ciências Humanas ligados à pesquisa qualitativa². Esse desenvolvimento visa suprir a demanda por programas aplicados à pesquisa qualitativa que, até então, oferecia uma série de obstáculos em sua viabilização operacional, tais como: grande gastos de tempo, custos, perda em relação a dados quando se trabalhava com grandes massas de dados, entre outros.

Segundo Barry (1998), o uso de programas para análise de dados qualitativos apresenta dois aspectos dicotômicos: por um lado, temos as vantagens que se expressam através da economia de tempo e de custos, a possibilidade de explorar de forma acurada o relacionamento entre os dados e, por outro, as vantagens em termos de uma estrutura formal que auxilia na construção conceitual e teórica dos dados. Entretanto, são apontadas algumas dificuldades em relação ao uso desses programas. Ocorreria

1 Programas como: The Ethnograph, Hyperqual, Winmax, Atlas/TI, NUD*IST, Kwalitan, Hyper Research, Alceste, NVivo.

2 Atualmente existe uma rede de trabalho unindo pesquisadores que utilizam estes programas denominada Caqdas Networking Project, fundada pela **UK Economic and Social Research Council**, rede esta sem fins lucrativos que visa a disseminação e discussão destes sistemas.

um distanciamento entre pesquisadores e dados; dados qualitativos passariam a ser analisados de forma quantitativa e, por fim, ocasionariam uma homogeneidade entre os métodos de análise de dados, inibindo a criatividade do pesquisador.

Não obstante, podemos pensar os programas de análise qualitativa de dados como novas ferramentas incorporadas ao processo de pesquisa, que possibilitam ao pesquisador dedicar-se à exploração e investigação dos dados, dispensando tarefas como, por exemplo, a tediosa seleção de blocos temáticos em fichas de entrevista ou documentos. Uma maneira de evitar o tratamento quantitativo de dados qualitativos é uma postura de vigilância constante do processo de pesquisa, visando, assim, não incorrer nas desvantagens que as técnicas de operacionalização via programa possam trazer. Isso, de certa maneira, gera um maior controle sobre o processo de pesquisa.

Cabe ressaltar que esses programas oferecem uma inovação extremamente importante: a possibilidade de testar e relacionar hipóteses valendo-se dos recursos e benefícios trazidos pela informática, até então privilégio da pesquisa quantitativa, utilizando qualquer tipo de técnica qualitativa, com qualquer material que possa ser transformado em texto.

2 Como operacionalizar?

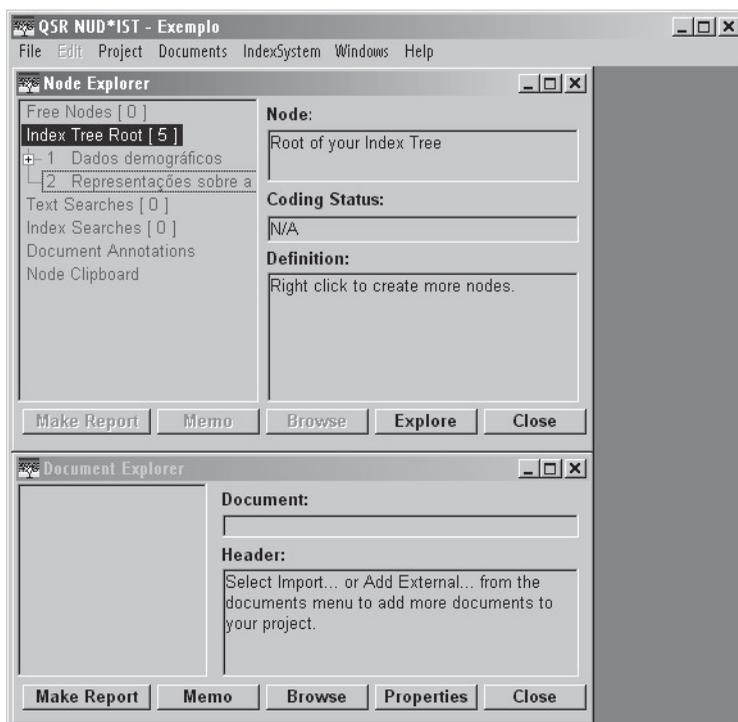
Tomaremos a partir de agora, o programa QSR NUD*IST como referência básica para fins de explanação. A operacionalização dos dados no NUD*IST, assim como na maioria dos programas orientados para a pesquisa

qualitativa, tem por base o princípio da codificação. Strauss define codificação como *o termo geral para conceitualização de dados; assim, os códigos abrangem questões nascentes e oferecem respostas provisórias sobre categorias e seus relacionamentos* (1987, p. 21).

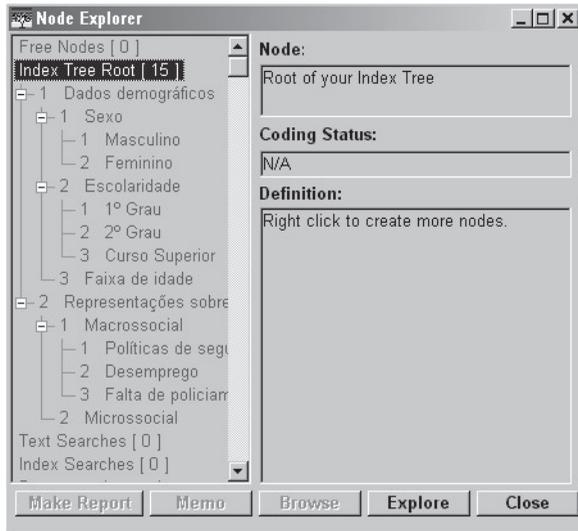
Desta maneira, após a leitura minuciosa dos documentos (entrevistas, documentos históricos, artigos de jornal, transcrições de vídeo, enfim, qualquer material qualitativo) os temas passam a ser codificados. Strauss (1991, p.72) e Corbin (1991, p.3) apontam três vias para a codificação dos documentos. A primeira compreenderia uma análise linha a linha na qual seriam examinadas palavras, frases e outras evidências. Outra possibilidade é a análise de sentenças ou parágrafos, na qual se busca uma categorização a respeito desses blocos. Por último, uma análise do documento como um todo.

Os códigos são arquivados em um sistema denominado *nodes-nós* – que vão conter as referências para os conceitos, categorias e hipóteses. Os “nós” são recipientes que armazenam a codificação, ou seja, os nós irão conter a referência a uma porção de texto codificado. Este é o princípio básico de ação do NUD*IST: a codificação do texto e o armazenamento dessas referências em nós específicos. O conjunto dos nós formam a *index tree root*, ou uma árvore onde todos os nós estão dispostos de forma hierarquizada e relacional. O NUD*IST trabalha com um sistema de janelas, como podemos ver no quadro abaixo. Na janela superior (*Node Explorer*) ficam armazenados os “nós”, e na janela inferior (*Document Explorer*) ficam armazenados os documentos que estamos analisando.

Poderíamos realizar, por exemplo, uma pesquisa sobre violência urbana na cidade de Porto Alegre, utilizando entrevistas estruturadas com 500 pessoas, entre homens e mulheres, na faixa de 20 a 60 anos, diversos graus



de escolaridade, querendo investigar as representações e percepções a respeito do aumento das taxas de violência. Bem, as percepções sobre o aumento da violência podem assumir diversas formas – relacionadas a condições macrosociais, ou microsociais e suas variantes. Nossos entrevistados poderiam atribuir o aumento da violência a causas individuais – idéia de que quem agride, o patrimônio ou pessoas, é intrinsecamente mau – ou a causas macrosociais – o aumento da violência estaria relacionado a fatores como desemprego, políticas de segurança pública, etc. Teríamos, então, essas categorias distribuídas em nós. O conjunto desses nós formam a *index tree root* – ou um sistema hierárquico de codificação conforme o exemplo abaixo:



Assim, codificaríamos todas as entrevistas com pessoas do sexo masculino no nó 1.1.1. (Dados demográficos / Sexo / Masculino); as entrevistas com pessoas com 2º grau no nó 1.2.2. (Dados demográficos / Escolaridade / 2º Grau) e sucessivamente até codificarmos todas as entrevistas.

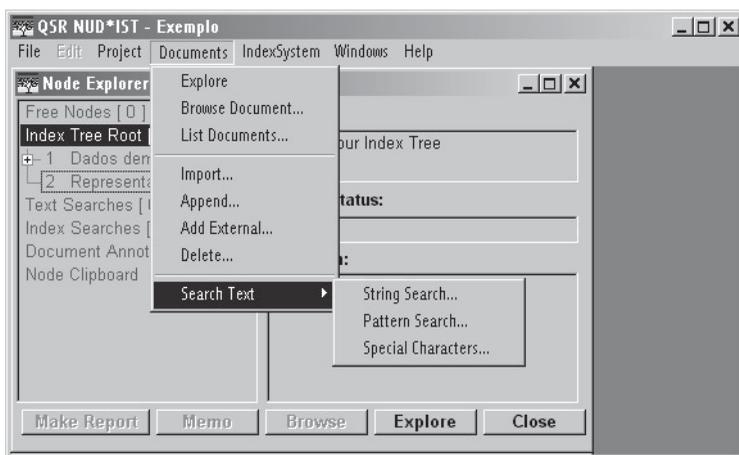
As referências às falas das pessoas que atribuem o aumento das taxas de violência às políticas de segurança pública estariam localizadas em 2.1.1. (Representações sobre a violência / Macrossocial / Políticas de segurança pública), já as que consideram o aumento da violência como decorrente da falta de um policiamento ostensivo estariam em 2.1.3. (Representações sobre a violência / Macrossocial / Falta de policiamento).

Vejamos então quais são os recursos oferecidos pelo NUD*IST para codificação e como funcionam.

3 Buscas automáticas na base de dados – *Text Searches*

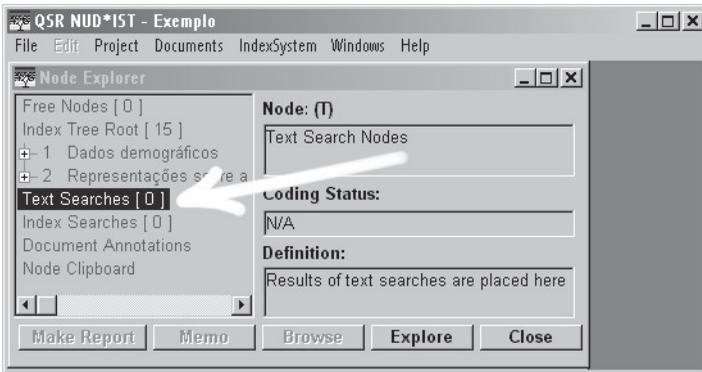
Inicialmente, o NUD*IST oferece duas possibilidades para a codificação. A primeira propicia uma construção qualificada dos dados através da leitura direta dos documentos, na qual poderíamos usar a idéia da codificação em três planos proposta por Strauss e Corbin (1991). A segunda está baseada na busca automática por palavras ou padrões lingüísticos (*Search Text*), como, por exemplo, encontrar a ocorrência de uma palavra em todas as entrevistas, ou buscar por palavras com padrões, sufixos, prefixos semelhantes. Esta busca automática oferece a vantagem da velocidade, principalmente quando são utilizadas grandes massas de dados, mas, por outro lado, depende do caráter da pesquisa, pode levar a uma codificação mais dispersa.

No NUD*IST há duas formas de usar o *search text* ou busca textual para explorar os documentos visando posterior codificação: por intermédio das ferramentas *String Search* ou *Pattern Search* (figura 3), que podem ser acessadas a partir do menu *Documents*:



A *String Search* é uma busca simples por uma seqüência de caracteres. O NUD*IST encontrará toda a unidade de texto³ contendo a exata seqüência informada na caixa de diálogo do *Search text*. Essa seqüência pode ser uma palavra ou frase. Já a *Pattern Search* é uma busca por um padrão de caracteres especificados com o auxílio de caracteres especiais. Com esta ferramenta, pode-se fazer uma busca para diferentes significados de um mesmo termo, diversas ocorrências com um mesmo radical, etc.

Quando uma busca por *string* ou por *pattern* é feita, um nó *Text Search* contendo a codificação é criado na área *Text Searches* (figura 4) do *Node Explorer*.



O resultado da busca também é copiado temporariamente no *Node Clipboard*, e é emitido um relatório que pode ou não ser salvo no formato *.txt. O fato de uma busca de texto ser armazenada em um nó torna-a disponível como um componente para posterior construção de idéias e buscas subseqüentes. Este é o princípio do *system closure*, ou refinamento da pesquisa, pois, como qualquer outro nó, esta busca pode ser utilizada para recodificação, novas buscas de texto ou em operações do tipo *index search* que serão detalhadas no ponto 4.

3 Para o NUD*IST, uma unidade de texto é igual a um parágrafo, ou uma porção de texto encerrada por um

* Caracteres especiais utilizados em *Pattern Searches*

O NUD*IST utiliza até 14 caracteres especiais a fim de especificar uma busca por padrões. Esses caracteres podem ser combinados a fim de sofisticar a pesquisa automática na base de dados. A seguir são apresentados alguns dos caracteres especiais mais importantes, assim como algumas combinações e seus efeitos:

^ (*Left Anchor*) - O acento circunflexo, quando usado no início do *pattern*, faz com que a busca encontre a palavra especificada no início de uma unidade de texto. Quando usado em outro lugar, que não no início, é considerado como um caractere comum. É útil, entre outras coisas, quando queremos saber o que determinado entrevistado dentro do grupo está falando, já que, via de regra, as transcrições trazem à frente da resposta o nome do emissor do discurso.

[] (*Alternation start and end*) - Os colchetes servem para delimitar o início e o fim de um *pattern* com alternativas de *strings* ou gama de caracteres. As *strings* ou gama de caracteres entre colchetes tornam-se grupos diante de *repetition wildcards* (tipo de caracteres especiais que será explicado mais adiante).

| (*String Alternation dividir*) - A barra vertical separa as alternativas de busca em um *pattern*. A busca é sempre exata, ou seja, não se pode usar *paterna* como alternativas somente *strings*. Os caracteres especiais [] e | devem ser utilizados conjuntamente. Dessa forma, pode-se fazer uma busca de tipo *string alternation*, bastante útil quando procuramos por diferentes significados de um mesmo termo, mas que pode variar de acordo com o emissor. Ao se fazer uma busca de tipo *pattern search* para a expressão **[funcionário|empregado|trabalhador|colaborador|associado]**, será encontrada, nos documentos trabalhados, qualquer ocorrência de um destes *strings*. Com a utilização desses caracteres em conjunto, também pode-se buscar por um radical e suas alternâncias:

empres[a|ária|ário|ariado]. A busca encontrará as ocorrências de “empresa”, ou “empresária”, ou “empresário”, ou “empresariado”.

. (*Spanning wildcard*) - O ponto serve para encontrar qualquer caractere que esteja em seu lugar, seja no início no meio ou no fim de uma palavra: uma busca por **empre...** encontra “EMPRESA”, “EMPRESÁRIO”, “EMPRESÁRIA”, “EMPRESARIADO”, “EMPREGO”, “EMPREGADO”, “EMPREGADA”, “EMPREGADOR”, etc.

(*Zero or one repetition wildcard*) - O sinal suspenso torna o seu caractere ou grupo à esquerda opcional. Pode ser usado no meio de uma palavra: uma busca por **fus#ca** recupera todas as ocorrências de “fusca” e de “fuca”. Neste caso é uma ferramenta muito útil quando se lida com diferentes pronúncias de uma mesma palavra. Pode-se, ainda, utilizar este caractere para encontrar diferentes formas de transcrição para uma mesma palavra: uma busca por **[Senhor|Sr.]#João** encontra “Senhor João” ou “Sr. João”.

4 Possibilidades de análise através do *Index Search*

Através de uma série de mecanismos incorporados ao NUD*IST, podemos testar e/ou explorar nossas hipóteses. Este mecanismo denomina-se *index search*, ou pesquisa nos índices da árvore principal (*index tree*), e é uma maneira de conduzir a análise de padrões de codificação dos nós no contexto do *Index System*, ou seja, no conjunto daquilo que foi codificado.

Para proceder às operações de *Index Search*, o NUD*IST dispõe de 17 operadores (figura 5) que permitem formular e responder praticamente qualquer questão sobre relacionamentos entre categorias e padrões de codificação dos documentos.



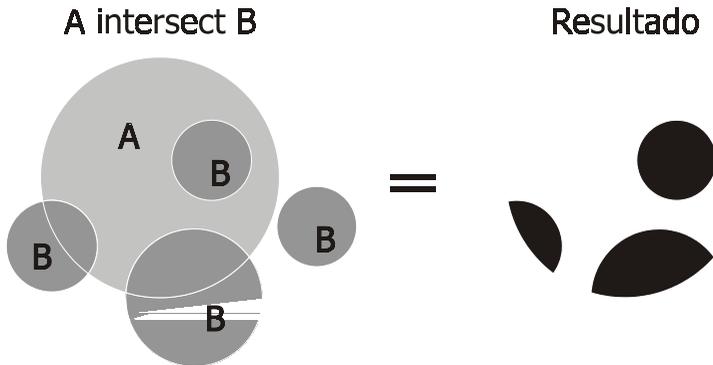
Neste ponto, imagina-se que a pesquisa já está mais avançada. Já foram feitas as buscas exploratórias utilizando o *Search Text*, e os nós básicos já estão formados. Isso porque este tipo de pesquisa somente funciona utilizando nós com informação, o que não impede que novos nós possam ser adicionados à árvore depois deste procedimento. Mas uma *Index Search* não pode partir do nada. É preciso que haja, na árvore, nós com codificação, ou seja, não apenas categorias, mas categorias “cheias” com índices que remetem a porções de texto.

Apresentamos a seguir os principais tipos de buscas por *Index Search*

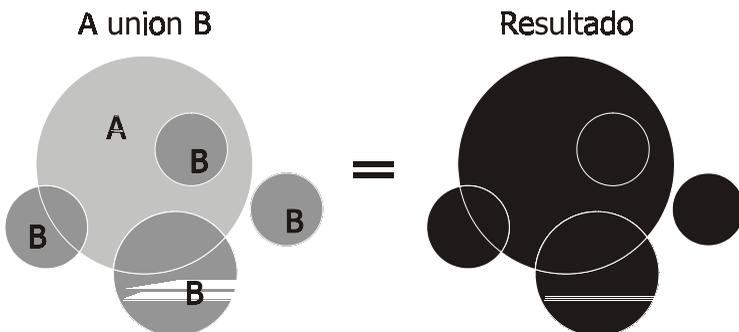
4.1 Collation Searches

Também chamadas operações Booleanas, são buscas que encontram todos os textos que estão codificados por um ou mais nós escolhidos de uma determinada maneira. Seus operadores são:

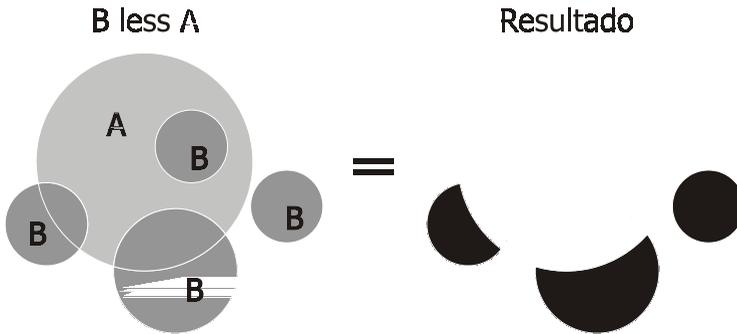
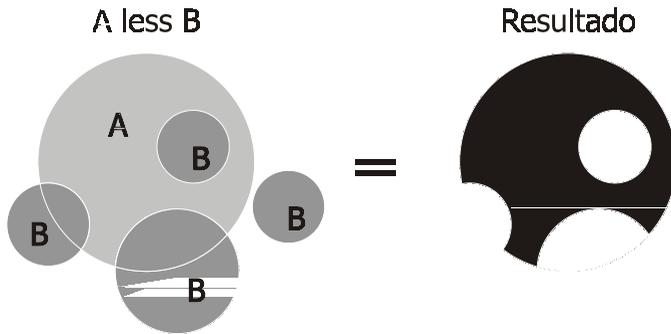
Intersect - encontra todas unidades de texto codificadas por todos os nós de um conjunto de dois ou mais nós indicados:



Union - Encontra todas as unidades de texto codificadas por qualquer um dos nós de um conjunto de dois ou mais nós indicados:

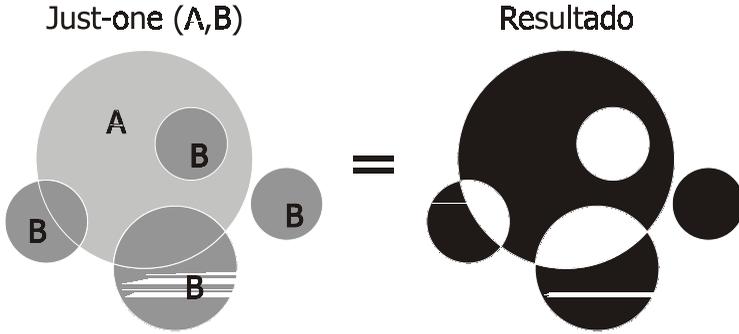


Less - Encontra todas as unidades de texto codificadas pelo primeiro nó indicado, mas não por nenhum outro nó (um ou mais):

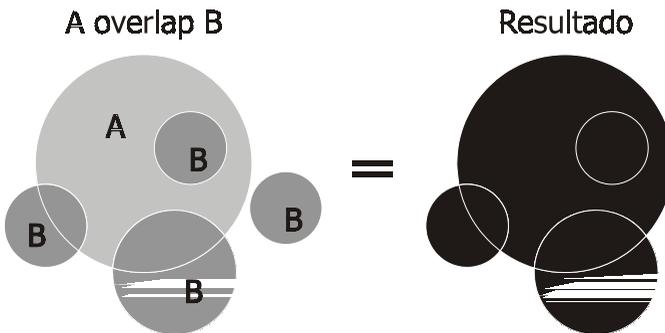


Just-one - Encontra todas as unidades de texto codificadas por somente um dos nós de um conjunto de dois ou mais nós indicados:

Sociologias, Porto Alegre, ano 3, nº 5, jan/jun 2001, p.94-114



Overlap - Como a operação *union*, une todas as unidades de texto codificadas por qualquer um dos nós, a diferença é que, se a busca de uma extensão de unidade de texto contiver referências de apenas um dos nós indicados, a busca é desconsiderada:



4.2 Tree-structured-searchs

São buscas que exploram a estrutura-árvore (*index tree root*) de um *index system*. Um operador recolhe todos os índices de referência de todos os nós na sub-árvore abaixo de um nó escolhido. Outros dois operadores cruzam qualitativamente, em uma matriz ou um vetor, os índices de referência nos nós “afiliados” de uma categoria com os nós “afiliados” de uma outra. Não se podem usar estas buscas com um *index system* vazio. Seus operadores estão divididos em dois subgrupos:

4.2.1 Operadores taxonômicos

Inherit - junta todos os índices de referência nos nós ascendentes de um nó informado, ou seja, aqueles acima no caminho que vai dele até a raiz do *index system*.

Collect - junta todos os índices de referência de um nó dado e em todos os nós nas ramificações da árvore abaixo dele.

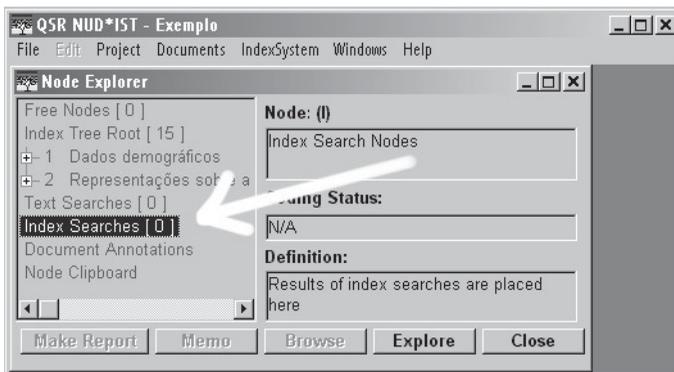
4.2.2 Operadores matriciais

Matrix – Este operador deve ser utilizado em conjunto com outro operador. Aplicando-se *matrix* com *intersect*, por exemplo, para os nós “violência” e “gênero”, obtém-se como resultado uma rigorosa distribuição da codificação das percepções de violência de acordo com o gênero em cada uma das subcategorias dos dois nós. O resultado também pode ser visualizado como uma tabela do tipo linha por coluna, com o número de documentos e unidades de texto codificadas em cada cruzamento de “violência” e “gênero”. Este tipo de operação guarda certa relação com o cruzamento entre variáveis (*cross tabs*), bastante utilizado em análises quantitativas,

e pode ser útil para a análise de conteúdo clássica, já que esta tabela pode ser exportada para um programa estatístico como o SPSS.

Vector - o mesmo que a *matrix*, com a diferença de que o operador (*intersect* ou qualquer outro) é aplicado ao próprio endereço do primeiro nó (não às suas subcategorias) e às subcategorias do segundo nó informado. Disto resulta uma “matriz” de várias colunas e uma só linha, ou seja, um vetor.

Da mesma forma que no *Search Text*, o NUD*IST grava automaticamente todas buscas por índices de referência em sua área específica no *Index System*, chamada *Index Searches* (figura 6) e no *Node Clipboard*, de forma que é possível fazer outra pergunta a partir do novo nó gerado usando outro operador para explorar seu relacionamento com qualquer outro nó na árvore. Assim, é possível continuar ininterruptamente perguntando, até que se tenha a reposta necessária para a análise.



5 Considerações finais

As contribuições aqui apresentadas nos permitem compreender a possibilidade do uso de programas na pesquisa qualitativa e as vantagens que trazem para o processo de investigação. Vimos as potencialidades de

uso dos sistemas *CAQDAS* no que diz respeito à operacionalização da pesquisa qualitativa. Na pesquisa quantitativa, programas como SPSS possibilitam o teste de hipóteses e a exploração dos dados através de operações como correlação, regressão, análise de variância entre outros. Facilidades estas que decorrem da própria natureza dos dados quantitativos.

O uso dos sistemas *CAQDAS* oferecem facilidades que, até então, eram privilégios da pesquisa quantitativa. No entanto, tratar dados qualitativos com o auxílio de um programa não significa um processo mecânico e padronizado, muito menos operacionalizar material qualitativo como se fosse quantitativo. É através da atitude de vigilância do pesquisador e tendo em mente a própria natureza dos dados qualitativos que podemos evitar que isso ocorra.

Devemos entender esses sistemas como ferramentas, como catalisadores do processo de pesquisa e usar todas as possibilidades que são oferecidas. Cabe salientar que um programa é somente um meio facilitador, não um fim em si mesmo. Além disso, um programa jamais substituirá a criatividade, o bom senso e o olhar sociológico do pesquisador. O que temos em mãos é uma ferramenta de trabalho muito rica, que nos auxilia a testar hipóteses com material qualitativo, explorar grandes massas de dados, e não uma espécie de “oráculo” no qual todas as perguntas são respondidas.

As técnicas de pesquisa qualitativa também não mudam em função desses programas, o que muda é a maneira como os dados são tratados e processados. A agilidade proporcionada pelo programa permite utilizar um maior volume de dados e libera tempo ao pesquisador para concentrar-se na pesquisa, e não mais em marcas ou tiras de papel em meio a uma pilha de entrevistas. Assim, abre-se todo um leque de possibilidades e inovações que tornam a pesquisa, a exploração, o teste de hipóteses e a análise na investigação qualitativa muito mais flexíveis e ágeis.

Referências bibliográficas:

BARRY, Cristine. Choosing Qualitative Data Analysis Software: Atlas/ti and Nudist Compared. **Sociological Research Online**, vol. 3, no. 3, 1998. Disponível na internet: <<http://www.socresonline.org.uk/socresonline/3/3/4.html>>. Em 22.01.99.

BECKER, Howard. **Métodos de Pesquisa em ciências Sociais**. São Paulo: HUCITEC, 1993.

KELLE, Udo. **Capabilities for Theory Building & Hypothesis Testing in Software for Computer Aided Qualitative Data Analysis**. Data Archive Bulletin May, 1997 No. 65, 1997. Disponível na internet: <<http://www.soc.surrey.ac.uk/caqdas/kelle.doc>>. Em 15.04.99.

NUD*IST 4.0. **User Guide**. Sidney: QSR, 1996.

RUDIO, Franz Victor. **Introdução ao projeto de pesquisa científica**. Petrópolis: Vozes, 1983.

SCHRADER, Achim. **Introdução a Pesquisa Social Empírica**. Porto Alegre: Ed. Globo, 1974.

SIEGEL, Sidney. **Estatística não-paramétrica para as ciências do comportamento**. São Paulo: McGraw-Hill, 1975.

STRAUSS, Anselm. **Qualitative Analysis for social Scientists**. Cambridge: Cambridge, 1987.

STRAUSS, Anselm & CORBIN, Juliette. **Basics of qualitative research Grounded theory Procedures and Techniques**. London: Sage Publications, 1990.

Resumo

Os manuais de metodologia qualitativa, não trazem nenhuma informação direta sobre o uso de pacotes para a análise de dados qualitativos, que, além disto são relativamente desconhecidos no Brasil. De outro lado, o uso destes programas torna-se imprescindível para a pesquisa sociológica devido a grande economia de recursos e tempo, mas também pela possibilidade de gerar análises mais ricas e aprofundadas. Este texto visa apontar os recursos oferecidos por estes programas de análise de dados qualitativos, especialmente o QSR NUD*IST, e como funcionam. Importa contribuir para a discussão sobre como operacionalizar a pesquisa sociológica utilizando estes pacotes, que servem como uma poderosa ferramenta para as diferentes técnicas de pesquisa qualitativa.

Palavras-chave: pesquisa qualitativa / metodologias informacionais / CAQDAS / NUD*IST.

