

### **Referência para citação**

FREITAS, H.; JANISSEK-MUNIZ, R.; MOSCAROLA, J. Modelo de formulário interativo para análise de dados qualitativos. Revista de Economia e Administração, São Paulo-SP, v. 4, nº 1, p. 27-48, Janeiro/Março 2005.

---

## **Modelo de formulário interativo para análise de dados qualitativos Interactive form for qualitative data analysis**

### **Henrique Freitas**

Endereço: Universidade Federal do Rio Grande do Sul – UFRGS, Escola de Administração – EA, Av. Washington Luis, 855, sala 307, 90010-460, Porto Alegre, RS, Brasil.

E-mail: [hf@ea.ufrgs.br](mailto:hf@ea.ufrgs.br)

### **Raquel Janissek-Muniz**

Endereço: Universidade Federal do Rio Grande do Sul – UFRGS, Escola de Administração – EA, Av. Washington Luis, 855, sala 307, 90010-460, Porto Alegre, RS, Brasil.

E-mail: [janissek@tiscali.fr](mailto:janissek@tiscali.fr)

### **Jean Moscarola**

Endereço: Universidade Federal do Rio Grande do Sul – UFRGS, Escola de Administração – EA, Av. Washington Luis, 855, sala 307, 90010-460, Porto Alegre, RS, Brasil. Fone:

(51)3316-3482. E-mail: [jean.moscarola@univ-savoie.fr](mailto:jean.moscarola@univ-savoie.fr)

### **Resumo:**

A partir de experiência conceitual e prática em tecnologias computacionais e método de pesquisa e de análise de dados, concebe-se um modelo (de formulário) para análise de dados qualitativos a partir da combinação de técnicas e ferramentas de análises léxica e de conteúdo. Ao mesmo tempo em que lê, analisa e codifica, o pesquisador pode ver o resultado de suas análises mais ou menos objetivas ir tomando forma, regulando, por aí mesmo, o seu próprio protocolo ou vocabulário de análise. Demonstra-se técnicas para preparação dos dados abertos textuais, em especial propõe-se um modelo integrando essas técnicas, o qual permite que o analista registre a subjetividade de sua percepção e tenha ao mesmo tempo uma noção do resultado da análise em curso, em todos seus aspectos. O modelo proposto pode ser realizado mesmo via web.

**Palavras-chave:** Dados qualitativos – Análise Léxica – Análise de Conteúdo – Protocolos de análise

**Abstract:**

Based on years of experience in computational technologies and method, one mix of ideas, practical concepts becomes and we offer a model for qualitative data analysis combining techniques and tools of lexical and content analyses, which can even be carried through the web. As the analyst is reading and coding, he or she can see the results of its more or less objective analysis taking form, so regulating, for there exactly, its proper protocol or vocabulary of analysis. A model is considered integrating these techniques, which allows the analyst to register the subjectivity of its perception and has at the same time a notion of the results of the analysis in course.

**Key words:** Qualitative data - Lexical analysis – Content analysis – Protocols

**1. Introdução: conceitos ‘dominados’, mas com prática inovadora**

O diferencial proporcionado pelos métodos e a isenção conferida pelo pesquisador que respeita estes métodos retratam as diferenças das competências e das intenções em uma pesquisa. Todo profissional deve desenvolver habilidades e dominar sistemas, técnicas e métodos inerentes à necessidade por construção de conhecimento a partir de dados disponíveis no seu contexto de atuação. Com o auxílio de instrumental, demonstra-se a viabilidade de explorar dados quanti-qualitativos e produzir informações que possam trazer respostas ágeis a muitos questionamentos que surgem no dia-a-dia de uma organização e mesmo no trabalho do profissional de pesquisa. Com o uso de dados qualitativos, pode-se identificar ou antecipar oportunidades e problemas de forma bem mais pontual, precisa e com um custo operacional bem reduzido. Existem várias técnicas de coleta e de análise de dados que permitem capturar automática e quase gratuitamente (algum custo operacional sempre se manifestará ou será inerente) dados qualitativos. Preconiza-se que o pesquisador atue direto,

não delegue a seus auxiliares a tarefa de entender o que está ‘por trás dos dados’. Esse tipo de dado pode ser explorado mais de uma vez pelo pesquisador, sendo fonte para a geração de novos, diferentes e curiosos dados.

Este artigo recorre a algumas técnicas para realizar análise de dados textuais, em especial as análises léxica e de conteúdo, buscando repassar ao leitor condições para uma investigação prática e eficaz. Estas técnicas são apresentadas enfatizando o seu uso em conjunto. O pesquisador ou analista tem, nelas, diferentes recursos que permitem a exploração adequada dos dados, através de procedimentos sistematizados. Diferentes interfaces e formas de organizar a tarefa serão ilustradas. Em particular, são apresentadas as idéias de complementaridade, recorrência e sequencialidade do uso das análises léxica e de conteúdo, abordando diferentes aspectos metodológicos para a exploração mais criteriosa da análise de dados qualitativos. Ao invés de uma ou outra técnica, a riqueza de análise seria melhor atingida com a variação e a aplicação conjunta de diferentes técnicas (Freitas e Janissek, 2000).

A seguir, discute-se sobre a questão dos dados qualitativos e das técnicas de análise, procurando-se desenvolver exemplos de interesse do analista, a partir do uso de software (Freitas, Janissek, Moscarola e Baulac, 2002). Um caso é usado para ilustração. Nosso intuito é fornecer recursos que permitam a exploração adequada de dados qualitativos, através de procedimentos sistematizados que assegurem a qualidade e mesmo originalidade das descobertas, a partir de uma visão integrada e um tanto inovadora.

## **2. Explorar dados textuais, um desafio possível**

A realização de pesquisas acadêmicas ou profissionais tem cada vez mais desafiado os analistas e pesquisadores visto que a objetividade dos dados coletados em uma pesquisa não é mais condição suficiente para a compreensão de um fenômeno (como a opinião de um certo

público, satisfação do cliente, resistência dos usuários finais de uma tecnologia recentemente adotada). De fato, a subjetividade é que vai permitir explicar ou compreender as verdadeiras razões do comportamento ou preferência de um grupo por um produto, serviço, etc.

Como o pesquisador, o analista, ou mesmo o gerente poderia realizar tal tipo de análise ou exploração de dados? Qual seria o gerente que não gostaria de – rapidamente – percorrer todos os dados captados através de uma linha 0800, como a do serviço de suporte ou outro, e se dar conta das reclamações ou sugestões atuais dos seus clientes? Não seria este tipo de dado, que fica esquecido e nem mesmo pensado como um recurso, uma fonte nobre de dados gerados sem custo adicional para a organização? Como o pesquisador poderia, antes de iniciar sua labuta, ter noção do ‘terreno’ que vai tentar compreender?

Ele pode imaginar uma pesquisa onde os clientes dizem se estão ou não satisfeitos com os serviços prestados numa escala de totalmente satisfeitos a totalmente insatisfeitos, mas não se dá conta que poderia saber disso sem mesmo ter que indagar aos clientes. A técnica quantitativa e melhor dominada aporta um tipo de resposta que é em diversos casos superficial no sentido que não permite uma qualificação ou um refinamento da opinião. No final, isto pode mesmo ser mais oneroso, pois quando os procedimentos de pesquisa são estabelecidos de maneira formal, há um viés natural na resposta oferecida pelas pessoas ou clientes, e o que se tem de volta são quantidades. Ao passo que, usando dados qualitativos, opiniões mais abertas, espontâneas ou mesmo – e porque não principalmente – aquelas coletadas de forma indireta (como a do setor de pós-venda ou de atendimento ao cliente), pode-se ter a chance de identificar ou antecipar oportunidades e problemas de forma bem mais pontual, precisa e com um custo operacional bem menor, principalmente a partir da exploração de um dado completamente espontâneo, não induzido de forma alguma.

Hoje, dispõe-se de diversas formas e tecnologias de se capturar automaticamente dados qualitativos (textos enfim) que nos denotem as preferências e comportamentos de um

certo grupo de pessoas: os métodos e ferramentas que possibilitam isso estão na literatura e no mercado (Moscarola, 1990; Weitzman e Miles, 1995; Freitas, Cunha e Moscarola, 1997).

Existe, contudo, uma confusão conceitual ou mesmo prática sobre em que consistem e mesmo de como essas técnicas e ferramentas poderiam ser utilizadas: pesquisadores acadêmicos ou analistas em geral deveriam ter uma clara visão das vantagens em explorar estes recursos de análise de forma sequencial, recorrente e sobretudo complementar (Freitas e Janissek, 2000).

### **3. Dados qualitativos, sua coleta e análise usando métodos múltiplos**

Tentando olhar de forma mais precisa um processo de pesquisa qualquer, vai-se notar que dificilmente ele deixa de recorrer a mais de uma técnica de coleta de dados, bem como mais de uma fonte, fazendo inclusive recurso a diferentes técnicas de análise de dados. Cada vez mais forte é a tendência de que uma pesquisa quantitativa, mais objetiva, deve ser precedida de uma atividade mais subjetiva, qualitativa, que permita melhor definir o escopo e a forma de focar o estudo. Também tem sido consenso que, mesmo sendo objetiva em sua essência, sempre se pode recorrer a algum tipo de opinião mais espontânea ou aberta, de forma a captar ‘um algo mais’ da parte do respondente.

Nas pesquisas, de forma geral, uma etapa pouco valorizada ou um tanto negligenciada é a de preparação de dados: além do esforço para que os dados sejam coletados (que não deixa muito tempo para que sejam analisados a fundo), é necessário que sejam preparados certos tipos de dados (idade em faixa etária, por exemplo). Da mesma forma, é necessário um investimento de tempo de análise) para que sejam criadas categorias pertinentes e claras a partir de dados tipo texto. Um exemplo é a criação de um dado novo a partir de uma questão texto que poderia ser “*quais as sugestões que você apresentaria para melhorar nossos serviços?*” ou então “*quais as principais reclamações que você apresentaria?*”. Pesquisadores (talvez despreparados) sobrecarregam suas pesquisas com questões abertas, criando uma

expectativa em quem responde, visto que tais respostas poderão nunca ser tabuladas e avaliadas como se deveria. Logo, o aprendizado que tal conteúdo permitiria não será jamais produzido. Cuidado especial é necessário, pois esse ‘excesso’ exige que o respondente se concentre bem mais que o normal: ele poderá pois ficar gradativamente desatento nas suas respostas e mesmo se desinteressar pelas questões posteriores.

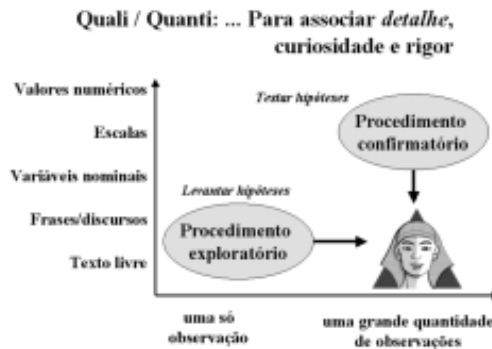
Cria-se pois uma falsa de idéia de que um dado espontâneo como esse dos exemplos recém evocados deve ser analisado uma só vez, quando ele poderia ser uma fonte bastante grande e diferenciada para a geração de novos, diferentes e curiosos dados. Esta geração de novos dados a partir de um dado texto pode emergir do próprio texto de cada respondente, como uma lista mais objetiva de sugestões ou de reclamações (no caso dos dois exemplos citados). Mas tal fonte de dados também pode ser objeto de um julgamento a partir de critérios a priori adotados pelo gerente, analista ou pesquisador, como por exemplo gerar um indicador de satisfação a partir da simples leitura das sugestões ou das reclamações emitidas pelo respondente. Isto é claro de forma subjetiva, contudo, muitas vezes este dado poderá ser melhor considerado para fins de análise. Ou seja, a um dado aberto e espontâneo, podem corresponder  $n$  dados objetivos, gerados seja pela análise e identificação gradativa de um protocolo (como a lista de sugestões que se faz emergir do texto), seja pela avaliação da opinião de cada pessoa em relação a um protocolo ou mesmo escala de medida preparada pelo analista (como por exemplo ‘satisfeito’ ou ‘insatisfeito’).

É importante a definição de quem deveria realizar as análises (e quem finalmente as realiza!): normalmente, solicita-se a um terceiro (auxiliar, bolsista, etc) para realizar esta *única* codificação, quando ela poderia ser rica em aprendizado caso feita pelo analista ou pesquisador. Esta atividade exige a leitura de cada uma das respostas, gerando, neste processo, novas idéias de análise e uma riqueza de compreensão e percepção sem iguais para o analista. Diversos aspectos inerentes são abordados por Freitas e Janissek (2000).

#### **4. Dados textuais: análise qualitativa ou quantitativa?**

A busca por informações consistentes e válidas não deve se deter *somente* nos dados estruturados, quantitativos, na forma de clássicas planilhas, relatórios volumosos, números, percentuais e gráficos: cada vez mais precisa-se recorrer aos dados de natureza qualitativa, como textos, discursos, entrevistas, reportagens, etc. Dados estes que envolvem elementos que muitas vezes desafiam a astúcia do pesquisador ou do homem de negócios, pois escondem em suas entrelinhas posicionamentos, opiniões, perfis, que exigem uma leitura atenta e ferramentas que possibilitem chegar com maior rapidez (condição de sobrevivência) às informações realmente pertinentes (Lesca, Freitas e Cunha, 1996). Deve-se poder ir do dado bruto ou puro ao dado elaborado, via interpretação, análise e síntese, e, a partir disso, por uma constatação ou curiosidade, poder rapidamente aprofundar a investigação, eventualmente voltando à fonte e ao dado bruto como recurso mesmo de sustentação de argumento ou simplesmente de ilustração.

É importante explorar e sobretudo cruzar de todas as formas possíveis dados quantitativos e qualitativos (Figura 1) para a geração de idéias, a verificação de hipóteses, a elaboração de conclusões ou indicação de planos de ação, etc. O uso de técnicas qualitativas x quantitativas, tanto para coleta quanto análise de dados, permitem, quando combinadas, estabelecer conclusões mais significativas a partir dos dados coletados, conclusões estas que balizariam condutas e formas de atuação em diferentes contextos. A abordagem literária, dita **qualitativa**, mais associada às técnicas de análise léxica e de conteúdo, pressupõe a análise de poucas fontes ou dados, num procedimento exploratório ou de elaboração de hipóteses.



*Figura 1 - O desafio de associar qualitativo e quantitativo (Freitas e Moscarola, 2000)*

Já a abordagem mais **quantitativa** pressupõe grande quantidade de dados num procedimento de confirmação de hipóteses. Há necessidade de se tratar do quantitativo, enriquecendo-o com informações qualitativas em grande número, de forma a ganhar força de argumento e qualidade nas conclusões e relatórios: o desafio é a busca da associação entre o quantitativo e o qualitativo, onde, por exemplo, o procedimento exploratório ganha força, visto que se poderá multiplicar os dados tratados, reforçando sobremaneira (e mesmo garantindo o ‘bom caminho’) o procedimento confirmatório.

Segundo Kelle (1995, p.15), "espera-se que as forças de ambas abordagens possam se reforçar mutuamente: a intersubjetividade e a fidedignidade ou confiabilidade providas pela informação padronizada derivada de amplas amostras, por um lado, e o conhecimento íntimo de um simples caso ou passagem de um texto adquirido pela análise interpretativa, por outro lado. A questão essencial nesse tipo de análise seria como transformar o significado da análise textual numa matriz de dados quantitativa, isto de maneira didática e sistemática".

Sejam quais forem os fenômenos, entidades ou realidade objetos de investigação, o uso conjunto destes métodos qualitativos e quantitativos permitiriam um maior aprofundamento no conhecimento dos dados (Mason, 1997; Lee, Liebenau e DeGross, 1997; Miles e Huberman, 1997; Creswell, 1998; Kirk e Miller, 1986), evidenciando-se aspectos do que se deseja investigar e, da mesma forma, possibilitando focar o pensamento sobre o assunto, decidir e executar. O uso conjunto destes dois tipos de análise permite estabelecer



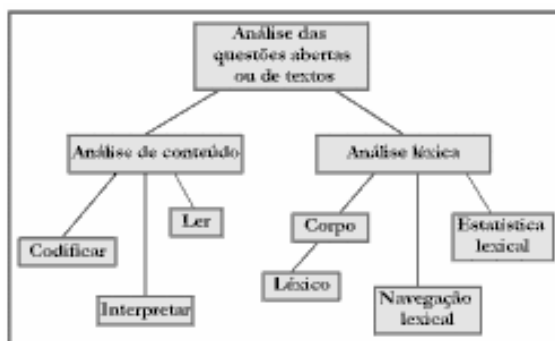
conclusões: é importante, pois, diferenciar ambos enfoques, identificando as possibilidades de cada um. No escopo de uma análise de dados qualitativos, pode-se ter uma abordagem mais qualitativa, com base na presença ou ausência de uma dada característica, ou então uma abordagem mais digamos 'quantitativa', onde se busca identificar a frequência dos temas, palavras, expressões ou símbolos considerados. A noção de importância deve ser clara em cada uma destas duas abordagens: o que aparece com mais frequência é o que importa na abordagem quantitativa, enquanto a abordagem qualitativa valoriza a novidade, o interesse e os aspectos que permanecem na esfera do subjetivo. Tem-se assim um dilema de análise: adotar categorias específicas, retratando fielmente a realidade, mas com uma lista de temas cuja frequência será fraca, ou então reagrupar deliberadamente os dados num pequeno número de categorias, não sem sacrificar informação talvez essencial, a qual estará eventualmente perdida no resultado final (Freitas e Moscarola, 2000).

### **5. As análises léxica e de conteúdo**

A análise de documentos, sejam originários de pesquisas quali ou quantitativas, inclui análise léxica e análise de conteúdo. Ela apresenta um conjunto de características racionais, sendo mais ou menos intuitiva, pessoal e subjetiva. Como outros métodos, apresenta problemas de validade, como autenticidade do texto, validade de interpretação e veracidade dos fatos. Tem ainda, em muitos casos, o defeito do trabalho não sistematizado, dependendo fortemente do valor e competência do pesquisador. Assim, é possível produzir novos dados que podem, por sua vez, ser confrontados especialmente com dados sócio-demográficos, como por exemplo um elenco de reclamações ou sugestões agora vistas por sexo, por faixa etária, por renda, por departamento ou qualquer outro dado mais objetivo ou quantitativo.

Antecipar a análise léxica à de conteúdo faria com que a análise de dados se desse de uma maneira plena: o uso destas 2 técnicas encobrem diversas das possibilidades que dali poderiam emergir - Figura 2. A Análise Léxica (Freitas e Janissek, 2000) consiste em se

passar da análise do texto para a análise do léxico (o conjunto de todas as palavras encontradas nos depoimentos ou respostas). Já a Análise de Conteúdo, que consiste em uma leitura aprofundada de cada uma das respostas, onde, codificando-se cada uma, obtém-se uma idéia sobre o todo (Krippendorff, 1980; Grawitz, 1993; Gavard-Perret e Moscarola, 1995; Freitas, Cunha e Moscarola, 1997; Weber, 1990; Silverman, 1993; Miles e Huberman, 1994; Bardin, 1996; Lebart e Salem, 1994; Frankfort-Nachmias e Nachmias, 1996, p.324-330; Moscarola, 1990, 1993 e 1994; Gavart-Perret e Moscarola, 1995; Freitas, 2000). WEBER (1990, p.9) define Análise de Conteúdo como um método de pesquisa que utiliza um conjunto de procedimentos para tornar válidas inferências a partir de um texto.



*Figura 2 - A análise de conteúdo e a análise léxica (Freitas e Moscarola, 2000)*

## **6. Um caso ilustrativo: pesquisa sobre “Problemas em Informação e Tecnologia”**

Para ilustrar as formas de análise de questões abertas nas enquêtes que se tem praticado, apresenta-se uma pesquisa realizada junto a gerentes de diferentes áreas, todos se manifestando sobre problemas que enfrentam e suas organizações quanto à informação ou à tecnologia da informação (TI).

<b>FICHA TÉCNICA DA ENQUÊTE</b>		
POTENCIAIS INTERESSADOS	Federação das Indústrias, ANPAD, PPGA/EA/UFRGS Sistemas SEBRAE, SOFTEX, SEPRO, ASSESPRO, ACIC, etc.	
OBJETIVOS	Verificar a percepção gerencial sobre os problemas e situações atualmente enfrentados quanto à gestão da informação e da tecnologia da informação nas organizações.	
AMOSTRA	Tipo de público-alvo	Quantidade
	Gerentes grandes e médias empresas de diversas áreas	508 gerentes

*Figura 3 - Ficha técnica da enquête “Problemas de TI”*

Esta pesquisa (Figura 3) teve por objetivo analisar a percepção gerencial acerca dos problemas e situações enfrentados em relação à gestão da informação, os sistemas de informação em si, e a tecnologia da informação. A idéia é que se possa mergulhar na lógica da pesquisa, do seu desenho simples, e acompanhar o que se pode construir a partir disso.

Inicialmente, em cada grupo pesquisado, simplesmente anotava-se o grupo, cidade e época, e 3 perguntas eram colocadas: pense num problema afeto à questão da informação ou TI; defina se este problema se refere ao cliente, ao cliente interno ou ao concorrente; relate o problema e atribua-lhe um título. É o que ilustra a Figura 4. A partir disso, foram derivados dados novos com base na identificação do grupo. Com uma coleta simplificada, passa-se a dispor de uma base com elementos chave consistentes para ‘pivotar’ a análise, ou seja, para cruzar os dados qualitativos que ainda vão se originar das análises léxica e de conteúdo que serão relatadas, no exemplo são: área, escolaridade, ano, região, local.

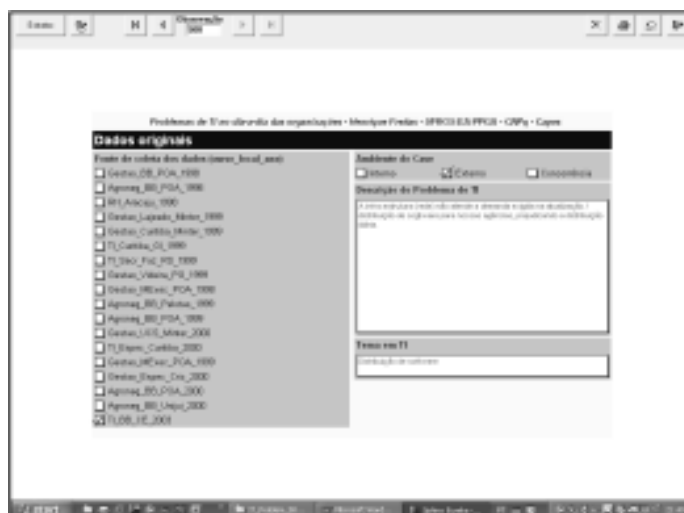


Figura 4 – A definição da base de dados inicial

Foram então criadas uma questão texto agregando o título e o relato do problema ou situação, de forma a permitir que as análises léxica ou de conteúdo fossem efetuadas em cima ou com base em um só dado, de uma só vez.

Essa questão, que reunia as respostas do problema em si e do seu tema ou título, foi criada por um recurso do sistema utilizado (Freitas, Janissek, Moscarola e Baulac, 2002). Com base na teoria ou interesses de investigação, foram criadas outras questões, que dariam lugar à reflexão do analista (de forma aberta, subjetiva, espontânea, ou então, quando possível, fechada). As outras questões (Figura 5) deveriam ser preenchidas pelo analista, para o que diferentes técnicas serão descritas, da mais clássica à com recursos via Web.

*Figura 5 – Novas questões, desafio do analista*

Nosso intuito é demonstrar técnicas para preparação dos dados abertos textuais, em especial a produção de um protocolo rápido a partir da análise léxica, com contagem de palavras e agregação em palavras-chave; a produção de um protocolo que emerge de uma análise de conteúdo, depoimento a depoimento; bem como a produção de um julgamento a partir da própria leitura do mesmo processo de análise de conteúdo, usando para tal um protocolo definido a priori (algo que se tinha em mente, uma curiosidade ou mesmo uma teoria que se quer confrontar); e ainda uma técnica alternativa, a qual, para criar o protocolo emergindo do texto lido depoimento a depoimento, parte não do marco zero (sem nenhuma categoria), mas sim do resultado bruto da análise léxica resultante da 1ª técnica proposta.

Ainda, propõe-se um modelo mais sofisticado, integrando essas técnicas todas, recentemente concebido por nossa equipe, o qual permite que o analista registre a subjetividade do seu julgamento ou percepção e vá tendo ao mesmo tempo uma noção do resultado da análise em curso, em todos seus aspectos.

## **7. Produção de um protocolo rápido e de outras medidas a partir da análise léxica**

A produção de um protocolo rápido a partir da análise léxica, com contagem de palavras e agregação em palavras-chave e depuração do *corpus* é uma técnica descrita em detalhe em Freitas e Janissek (2000). A Análise Léxica consiste em se passar da análise do texto para a análise do léxico (o conjunto de todas as palavras encontradas nos depoimentos ou respostas). Como analisar esse tipo de dados qualitativos? A respeito desta questão surgem outras, entre elas: (1) será que o problema ou tipo de problema varia segundo o tempo? Segundo a região? Segundo a área gerencial? (2) Será que se consegue depreender uma tipologia clara dos problemas em TI?

A partir de um ambiente onde diversos recursos de análise textual estão disponíveis, pode-se produzir diversas medidas (quão comum é cada texto em relação ao vocabulário do conjunto de textos, quão original, quão longo ou curto, etc.), bem como extrair textos de interesse peculiar em determinada situação, ou então, compreender o ambiente de uma palavra ou idéia-chave, ou ainda, recorrer às técnicas disponíveis para análise léxica ou de conteúdo (a base de nossa ilustração será a questão que reúne o problema de TI relatado e o título atribuído a este relato).

O ambiente de análise léxica evocado é o ilustrado na Figura 6. Ele disponibiliza n recursos, como por exemplo: percorrer as respostas com determinado perfil, ou então ilustrar o texto ativo (janela maior) com outros dados (como ambiente, área, cidade, etc.), resgata na coluna da esquerda todas as palavras diferentes com sua frequência de incidência nos textos, permite deletar e reagrupar (por cliques, por raiz, etc.) essas mesmas palavras, para reduzi-las a idéias-chave, e permite ainda selecionar textos e extrai-los para um arquivo texto, ou então gravar novos dados na base a partir das idéias-chave marcadas na lista da coluna da esquerda.

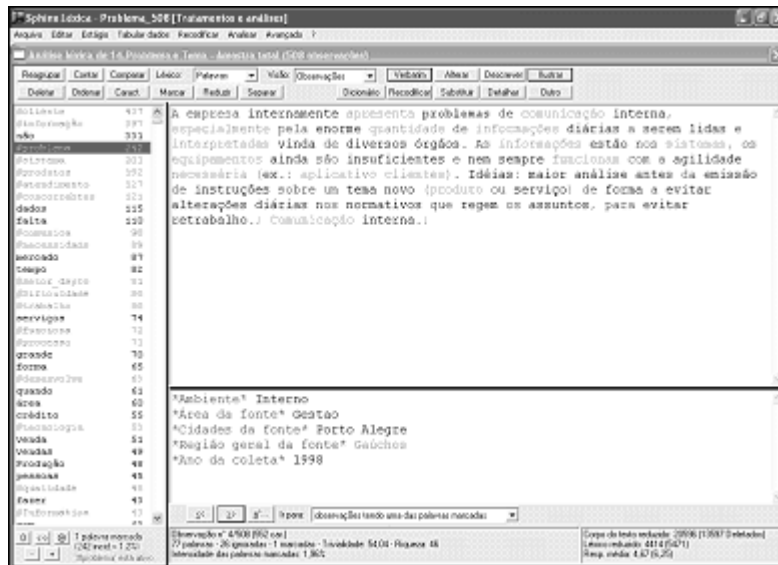


Figura 6 – Ambiente de análise léxica do Sphinx Léxica®

Na Figura 6, pode-se verificar que no total as 508 respostas representam mais de 34 mil palavras, as quais se transformaram em 20.596 após a exclusão de 13.597 palavras do tipo artigo, preposição, etc, e vê-se ainda que, entre todas as palavras, 4414 são diferentes. Aprende-se rapidamente sobre os problemas em TI, estes estando afetos a cliente, informação, falta ou necessidade, sistemas, produtos, atendimento, concorrentes, dados, etc: é isso que possibilita a análise léxica. Na verdade, fez-se a chamada **aproximação lexical**, ou seja, retirou-se da lista ou “léxico” todas as palavras chamadas “instrumentais” (todas aquelas que não comunicam diretamente uma idéia no sentido da análise realizada) e chegou-se às palavras de conteúdo ou significado. Essas palavras podem ser marcadas no *corpus* e dali origina-se um novo dado, onde registraram-se as 22 principais (as mais frequentes).

Toma-se conhecimento do que evocam (Figura 7) os gerentes ao relatarem seus problemas com informação e TI. Nota-se que 168 pessoas falam de “problema” em seus relatos, mas em que sentido?

Problema_Tema_22		
Cliente	239	13.0%
Informação	213	11.6%
Problema	168	9.1%
Falta_necessidade	140	7.6%
Sistema	122	6.6%
Concorrentes	73	4.0%
Área_setor	72	3.9%
Dados	71	3.9%
Produtos	67	3.6%
Atendimento	67	3.6%
Dificuldade	65	3.5%
Trabalho	64	3.5%
Tempo	61	3.3%
Serviços	57	3.1%
Processo	54	2.9%
Comunica	53	2.9%
Produto	52	2.8%
Mercado	51	2.8%
Desenvolve	49	2.7%
Tecnologia	39	2.1%
Venda	33	1.8%
Crédito	31	1.7%
<b>Total</b>	<b>1841</b>	<b>100.0%</b>

Figura 7 – Palavras ou idéias-chave

Faz-se a chamada **navegação lexical**, que consiste em ir diretamente às respostas onde foi citada a palavra em questão, podendo-se gerar um arquivo contendo a lista das respostas do léxico; pode-se mesmo exportar estas respostas para um documento em editor de texto (Figura 8). Este arquivo é chamado **verbatim** (citação). Isso permite a restituição das respostas, organizando-as de acordo com seu conteúdo, possibilitando conhecer em que contexto e com que sentido foi empregada a palavra. Pode-se mesmo segmentar os depoimentos ou relatos com base em outro dado objetivo (a região da pessoa, por exemplo).

Isto é apenas um exemplo de uma extração, onde a idéia-chave PROBLEMA foi marcada e então selecionados todos os textos de relato de problema de TI que mencionavam ‘problema’ ou similar agrupado por raiz, e isto, no caso, segmentado pela ‘região geral da fonte’, no caso da ilustração apenas alguns dos ‘gaúchos’ estão presentes. Dois cliques permitem realizar esse tipo de operação. Nosso intuito é colocar ênfase no avanço metodológico, porém viabilizado pela tecnologia disponível.



Textos extraídos de '14.Problema e Tema'  
 Lista das palavras marcadas do léxico (com as quantidades de ocorrências para cada palavra)

#problema 242

- As respostas envolvidas são as observações tendo uma das palavras marcadas
- As respostas são reagrupadas segundo as categorias da variável '8.Região geral da fonte'
- As extrações são reduzidos aos frases.

8. Região geral da fonte = Gaúchos:

nº 4 : A empresa apresenta **PROBLEMAS** de comunicação interna, especialmente pela enorme quantidade de informações diárias a serem lidas e interpretadas vinda de diversos órgãos.

nº 8 : Precisamos saber se o **PROBLEMA** é rejeição do cliente por não pssuir cultura voltada para o produto ou se nosso marketing está inadequado - **PROBLEMAS** na venda/aceitação.

nº 13 : **PROBLEMAS** de acessar sistemas nas horas de pico.

Figura 8 - Algumas respostas extraídas a partir do “Problema de TI” relatado

Com um duplo-clique em cima da palavra ‘problema’ na lista de palavras (ou *corpus*), pode-se acessar o dito ‘ambiente’ da palavra, como ilustra a Figura 9: tem-se ali uma boa noção dos termos vizinhos (antes e depois da palavra selecionada), sabe-se qual a observação em que aquilo foi dito, e pode-se mesmo ilustrar com outras variáveis ou dados do estudo (ilustrado em vermelho na figura).

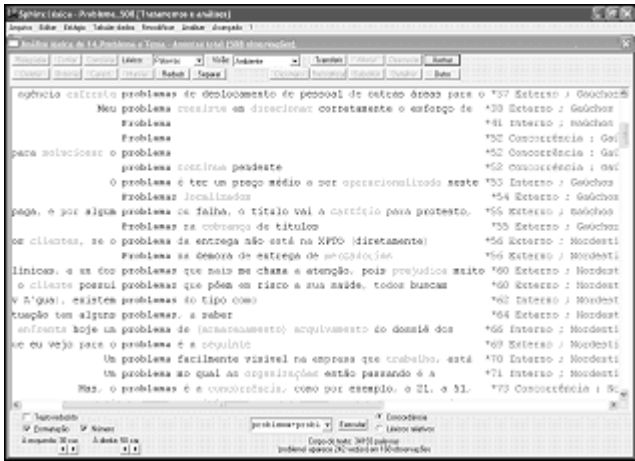


Figura 9 – Contexto da palavra ‘problema’ e sua ilustração

A partir desse tipo de listagem, pode-se novamente recorrer ao recurso chamado verbatim e extrair listas para uso posterior ou ilustração de argumentos. Diversas são as

situações em que isso seria um excelente recurso facilitador ao trabalho do analista ou pesquisador. A Figura 10 mostra o tipo de uso e exploração que pode ser feito a partir do dado novo criado com as palavras ou idéias mais frequentes.

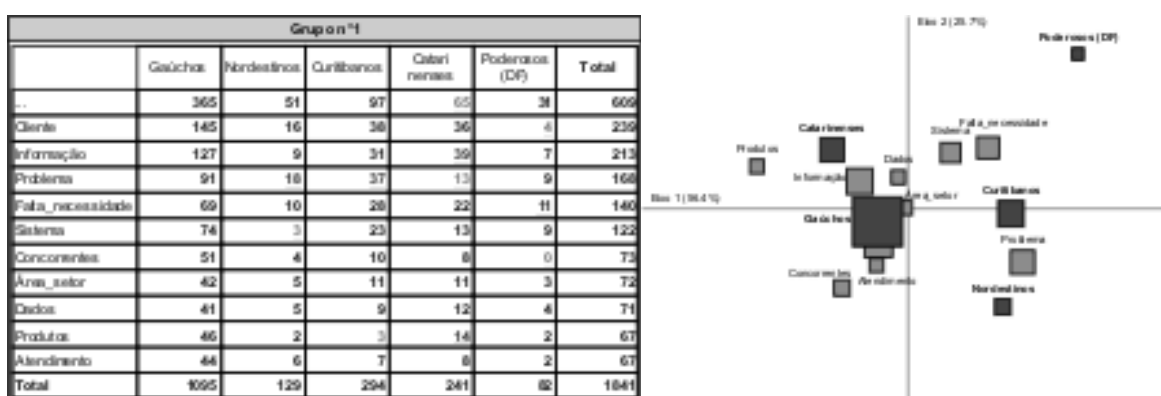


Figura 10 – As 10 principais idéias-chave e a região e o mapa da análise de correspondência

Tendo destacado as 22 idéias-chave, pode-se reter somente as 10 mais frequentes para cruzá-las com outras do estudo (a área de atuação do gerente, ou sua região de atuação), e assim descobrir pistas para análise ou para decisões a tomar. Aqui, parece haver pelo menos, numa análise muito sumária (uma vez que nosso intuito neste texto não é o caso dos problemas de TI em si, mas sim a ilustração metodológica do que se pode fazer com esse tipo de dado e com tecnologia adequada!), certa distinção entre a preocupação de catarinenses de um lado, e de nordestinos, dos curitibanos e os ditos ‘poderosos’ (apenas para o folclore da análise, naturalmente!) de outro lado, estes últimos 3 tendo mesmo entre eles alguma distinção de foco: é o que se pode rapidamente depreender da Figura 10.

## 8. A produção de um protocolo (idéias-chave) emergindo da análise de conteúdo

A Análise de Conteúdo consiste em uma leitura aprofundada de cada uma das respostas (relato a relato): codificando-se cada uma, obtém-se uma idéia bem mais precisa sobre o todo: é o que chamamos de fazer emergir um protocolo ou lista de idéias-chave. Aplica-se esta técnica inicialmente para analisar cada expressão específica de uma pessoa ou

grupo envolvido. Ela consiste em ir lendo o texto correspondente a cada uma das empresas entrevistadas (parte superior da imagem à esquerda na Figura 11), e, a seguir, interpretar o conteúdo de forma a codificar as respostas dadas, como se pode ver na parte inferior da mesma imagem (Freitas e Janissek, 2000). À medida que se lê o conteúdo da resposta em análise, pode-se criar as categorias de respostas, as quais se deve marcar antes de passar ao próximo respondente. A tabela de resultados à direita na Figura 11 apresenta o início de uma codificação que evolui naturalmente à medida que se avança na leitura de cada resposta e consequente criação de novos códigos ou categorias de respostas a marcar.

Categoria	Frequência	Porcentagem
Atualização	2	22.2%
Facilidade operação	2	22.2%
Sistema de informação	2	22.2%
Regras e limites	1	11.1%
Preços	1	11.1%
Relação cliente	1	11.1%
<b>Total</b>	<b>9</b>	<b>100.0%</b>

Figura 11 - A análise de conteúdo (codificação inicial faz 'emergir' protocolo) e resultado tabulado

Uma escolha que deve ser feita antes de iniciar uma codificação é se deverá se tratar de um dado novo com respostas simples ou múltiplas (no caso do exemplo é uma resposta de múltipla escolha, pois um respondente poderia citar diversos aspectos de um problema de TI), já uma leitura preocupada em associar cada respondente a um só tipo de resposta, por exemplo “propõe soluções” ou “não propõe soluções”, seria uma de resposta única. Os dados da tabulação apresentada na Figura 11 referem-se a uma codificação parcial, mas servem para ilustrar o resultado da operação de análise realizada.

Este é um contexto no qual se pode ler e anotar o conteúdo do que é lido, com a possibilidade de enriquecer os temas que se está interpretando. É como se fosse uma tentativa de “fechar” a questão posteriormente. Muito embora esse método possa parecer um trabalho maçante e demorado, ele pode também ser gratificante. De fato, é um método clássico de análise de conteúdo, que envolve leitura e muita atenção. Mas a codificação resultante de todo esse esforço será expressiva, caracterizando a realidade inerente àquela amostra, além de possibilitar ao analista o pleno domínio sobre os dados e as sutilezas das entrelinhas do texto analisado, além de um novo dado excelente para cruzamentos, segmentações, análises, enfim.

Como evocado, para otimizar o processo e seus resultados ou produtos, a idéia que se pode agregar é a seguinte: além de fazer emergir idéias-chave a partir da leitura de cada relato ou depoimento ou entrevista, o analista pode igualmente emitir certos julgamentos ou percepções. Um julgamento a priori que se poderia imaginar no caso em questão (apenas de forma a ilustrar o que aqui se preconiza), é saber se, em cada problema relatado, a pessoa apenas relata a situação, ou se igualmente ela propõe soluções a esse mesmo problema! Assim, numa só leitura, o pesquisador pode produzir dados novos que possibilitarão, depois, na análise, uma idéia melhor, mais completa, argumentos mais consistentes, seleção de depoimentos mais interessantes sob um certo ângulo.

### **9. O protocolo emergindo do texto lido depoimento a depoimento, partindo do resultado bruto da análise léxica**

Uma astúcia possível, interessante, é a seguinte: combinar o resultado de uma análise léxica com a atividade de fazer emergir um protocolo a partir da leitura do texto, ou seja, a análise de conteúdo para geração de idéias-chave. Alguns cuidados deverão ser observados, entre eles: deixar para ajustar os nomes das categorias ou modalidades de resposta apenas na estrutura do questionário, bem como apenas deletar categorias no recurso dito de

'transformação'. Ou seja, na hora da codificação em si, apenas agregar novas categorias e marcá-las ou então marcar algo que o sistema não fez na oportunidade da análise léxica.

Tratas-se então de criar o protocolo emergindo do texto lido depoimento a depoimento, contudo partindo não do marco zero (sem nenhuma categoria), mas sim do resultado bruto da análise léxica resultante da primeira técnica proposta (Figura 12).

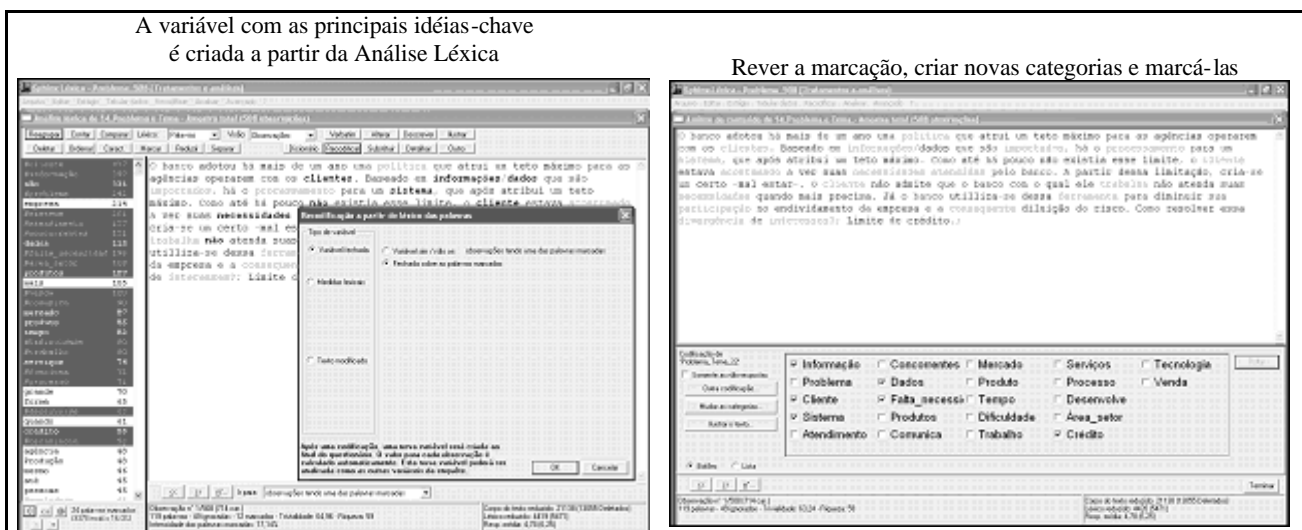


Figura 12 - A análise de conteúdo, com a codificação partindo do resultado da análise léxica

Tal procedimento permite que seja lido o texto (note que o próprio texto trás já marcado em cor diferente todas as categorias agrupadas em idéias-chave por ocasião da análise léxica), e confirmadas as idéias-chave, bem como, com o recurso de mudar as categorias, podem ser acrescentadas outras que o analista porventura deseje, a partir de sua leitura mais atenta de cada relato ou depoimento. Assim, o fruto da análise de conteúdo é facilitado, certamente aliviando a tarefa do analista e enriquecendo o resultado, posto que ele se preocupa não mais em fazer emergir completamente o protocolo de idéias-chave, mas sim em revisá-lo e complementá-lo.

## **10. Modelo integrando técnicas para facilitar a subjetividade do julgamento dos analistas**

Do analista ou dos analistas! Sim, o modelo pode mesmo derivar para algo publicado na Web, onde cada analista poderia acessar um formulário e registrar sua análise. Propõe-se um modelo mais sofisticado, integrando essas técnicas todas (vistas nas seções anteriores), muito recentemente concebido por nossa equipe, o qual permite que o analista registre a subjetividade do seu julgamento ou percepção e vá tendo ao mesmo tempo uma noção do resultado da análise em curso, em todos seus aspectos. É o que se ilustra com a Figura 13.

**Concepção do ambiente**<sup>1</sup>: notem as variáveis do estudo na coluna da esquerda, em cima. Na mesma coluna, embaixo, um guia para o analista saber que formulário está ativo. Na janela maior, a montagem da interface ou formulário que permitirá ler o ‘problema e tema’ (variável 14), escrever o resumo do analista, o qual será imediatamente tabulado na janela ‘expressões’, o mesmo acontecendo com o ‘quem se prejudica’ e o ‘foco’ de fato, todos 3 novos dados subjetivos partindo do analista, mas já sendo tabulados digamos ‘online’! Restará marcar se o respondente propõe ou não soluções em seu relato. Para tal concepção, basta arrastar com o mouse as variáveis para dentro da janela maior, depois organizá-las e melhor ajustá-las (recursos com o mouse direito). Note-se ainda que a variável fruto da análise léxica ali se encontra para ser corroborada pelo analista (coluna laranja).

**Realização da análise**: o analista lê o ‘problema e tema’, e então escreve sua reflexão: no ‘resumo do problema na visão do analista’, bem como clica na questão que indaga sobre ‘se propõe soluções’ (sim ou não), escreve uma ou duas palavras na questão sobre ‘quem se prejudica’ com aquela situação relatada, e qual de fato, em uma ou duas palavras o ‘foco’ do problema relatado. Esses 3 novos dados já vêm, cada um, acompanhados da respectiva tabela de resultados ali inserida. Desta forma, o analista coordena seu protocolo ou vocabulário, e obtém um resultado otimizado de seu esforço de análise, a partir do quê poderá realizar

segmentações em seus dados e análises, bem como extrair de forma mais efetiva e eficaz trechos para sustentar sua argumentação ou tese. Além disso, ele, tendo lido o problema, corrobora o resultado da análise léxica (coluna laranja), marcando outras categorias ou desmarcando mesmo. O botão ‘mudar perfil’ permite focar a consulta e análise em certo público-alvo.

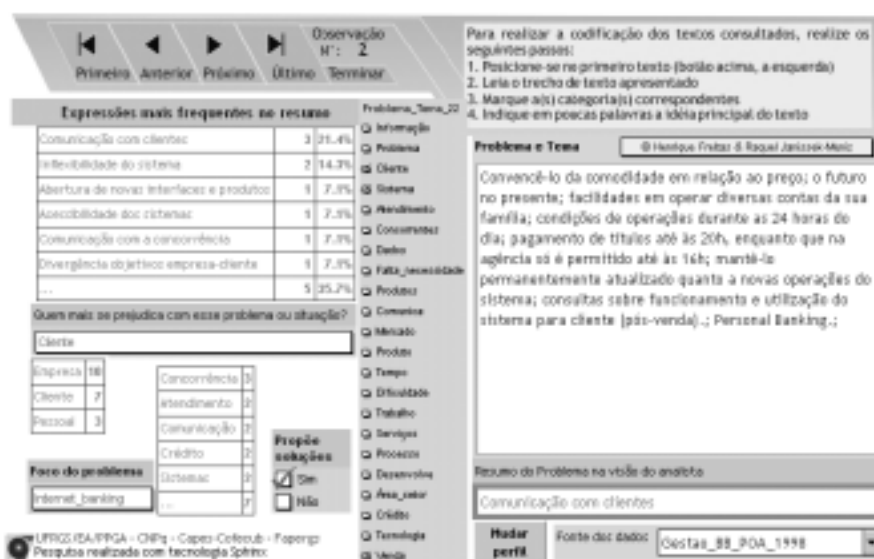
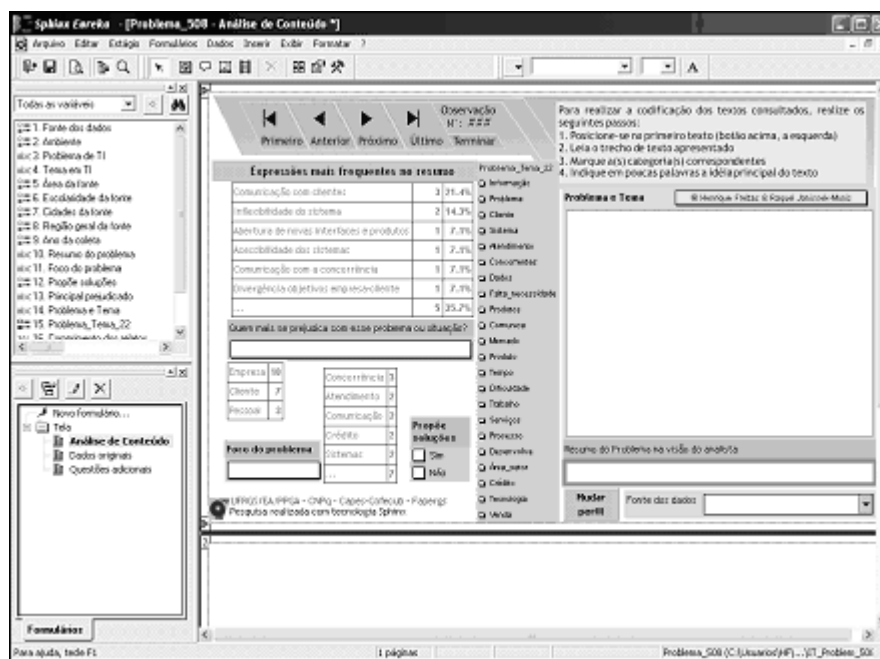


Figura 13 - A análise léxica automatizada, a partir dos textos de base e da reflexão do analista

## **11. Conclusão: idéias simples de preparação de dados para uma análise mais efetiva**

Nossa preocupação ao escrever é que os textos produzidos possam ter efetiva utilidade para os leitores. Com todo avanço tecnológico da última década sobretudo, é mais que tempo de revolucionar conceitos, preconceitos, pressupostos, quebrar de fato paradigmas. Não sem a precaução ou o cuidado metodológico inerente. As técnicas abordadas neste artigo, sobretudo na seção 6, onde mostramos como praticá-las, são de certa forma um instrumento na mão dos analistas para que estes possam empreender e se desafiar a produzir novos dados, mais ricos, mais consistentes, mais ligados à realidade, mais agregadores de valor à tomada de decisão!

De fato, demonstra-se técnicas e métodos para preparação dos dados abertos textuais, em especial a produção de um protocolo rápido a partir da análise léxica; a produção de um protocolo que emerge de uma análise de conteúdo, depoimento a depoimento; bem como a produção de um julgamento a partir da própria leitura do mesmo processo de análise de conteúdo, usando para tal um protocolo definido a priori; e ainda uma técnica alternativa, a qual, para criar o protocolo emergindo do texto lido depoimento a depoimento, parte não do marco zero (sem nenhuma categoria), mas sim do resultado bruto da análise léxica resultante da primeira técnica proposta. Por fim, propõe-se um modelo integrando essas técnicas todas, o qual permite que o analista registre a subjetividade do seu julgamento ou percepção e vá tendo ao mesmo tempo uma noção do resultado da análise em curso, em todos seus aspectos.

Nossa proposta – reforça-se - é de que as duas técnicas sejam utilizadas de forma **sequencial** (uma após a outra), **recorrente** (pode-se ir e vir, deve-se mesmo ir e vir de uma a outra) e **complementar** (elas não são excludentes, ou seja, não se deve escolher uma ou outra, deve-se adquirir finalmente a visão, a consciência de que os recursos de ambas são excelentes ferramentas na mão do analista e que ele deve fazer bom uso e não isolar uma em detrimento de outra).



Enfim, quando as pessoas e as organizações começam a prestar atenção nos telefonemas que recebem, nas sugestões e reclamações da clientela ou de fornecedores, e quando a academia começa a valorizar bem mais as questões subjetivas em harmonia com aquelas em demasia objetivas, temos firme convicção da potencial contribuição deste trabalho.

---

<sup>i</sup> O estagiário Ricardo Simm Costa, estudante da Escola de Administração da UFRGS, cooperou com os autores na realização operacional deste modelo.

## **12. Referências Bibliográficas**

- BARDIN, L. **L'analyse de contenu**. PUF, 8<sup>ème</sup> Edition, Paris:, 1996.
- CRESWELL, J. W. **Qualitative Inquiry and Research Design**. Sage Publications, 1998.
- FRANKFORT-NACHMIAS, C. & NACHMIAS, D. **Research Methods in The Social Sciences**. New York: St. Martin's Press, 1996.
- FREITAS, H. e JANISSEK, R. **Análise léxica e Análise de Conteúdo: técnicas complementares, seqüenciais e recorrentes para análise de dados qualitativos**. Porto Alegre: Sagra, 2000.
- FREITAS, H.; MOSCAROLA, J. **Análise de dados quantitativos e qualitativos: casos aplicados**. Porto Alegre: Sagra, 2000, 176 p.
- FREITAS, H. As tendências em Sistemas de Informação com base em recentes congressos. **ReAd** (<http://www.ea.ufrgs.br/read>), Porto Alegre, no.13, 20 p, Jan. 2000.
- FREITAS, H., CUNHA Jr., M. V. M. e MOSCAROLA, J. Pelo resgate de alguns princípios da análise de conteúdo: aplicação prática qualitativa em marketing. **Anais do 20º ENANPAD, ANPAD**, Angra dos Reis - RJ:, Marketing, p. 467 – 487, 23-25 Set. 1996.

- FREITAS, H., CUNHA Jr. M.V.M. e MOSCAROLA, J. Aplicação de sistema de software para auxílio na análise de conteúdo. **RAUSP**, S. Paulo, v. 32, nº 3, p. 97-109, Jul/Set. 1997.
- FREITAS, H., JANISSEK, R., MOSCAROLA, J. e BAULAC, Y. **Pesquisa interativa e novas tecnologias para coleta e análise de dados**. Porto Alegre: Metrópole, 2002, 381p.
- GAVARD-PERRET, M.L. & MOSCAROLA, J. De l'énoncé à l'énonciation: relecture de l'analyse lexicale en marketing. **Université de Savoie Cahier GERE**G, França, 34 p, 1995.
- GRAWITZ, M. *Méthodologie des Sciences Sociales*. **Dalloz, 1976 (chap. II, pp. 586-627), and 1993 (9 ed., chap. 2, section 2, pp. 532-565)**, Paris: 870 p.
- KELLE, U. **Computer-Aided Qualitative Data Analysis: theory, methods and practice**. Sage Publications, 1995.
- KIRK, J. & MILLER, M. **Reliability and validity in qualitative research**. Sage publications, 1986.
- KRIPPENDORFF, K. **Content analysis: an introduction to its methodology**. The Sage CommText Series, 1980, 191 p.
- LEBART, L. & SALEM, A. **Statistique Textuelle**. Paris: Dunod, 1994. 342 p.
- LEE, A. S. LIEBENAU, J. ; DEGROSS, J. I. *Information Systems and Qualitative Research*. Philadelphia, **Proceedings ICIS**, Chapman & Hall, USA, 1997.
- LESCA, H., FREITAS, H. e CUNHA Jr., M. **Instrumentalizando a Decisão Gerencial**. Rio de Janeiro - RJ: Revista Decidir, ano III, n. 25, p. 6 - 14 Agosto 96.
- MASON, J. **Qualitative researching**. Sage, 1997, 180 p.
- MILES, M. B. & HUBERMAN, A. M. **Qualitative data analysis**. Sage Publ., 1994, 338 p.
- MOSCAROLA, J. **Enquêtes et Analyse de donnés**. Libraire Vuibert, Paris, 1990.
- MOSCAROLA, J. *Analyse de contenu et analyse de données: solutions logiciels pour une intégration quali/quant*i. **Actes JADT**, Paris, 1993.

MOSCAROLA, J. La communication politique vue par l'analyse lexicale. Annecy, **IUP, Université de Savoie, Cahier du GERE**G, France, n° 32, 1994.

SILVERMAN, D. **Interpreting Qualitative Data**. Sage Publications, 1995. 224 p.

WEBER, R.P. **Basic content analysis**. Sage University paper, 1990. 96 p.

WEITZMAN, E. A. e MILES, M. B. **Computer programs for qualitative data analysis**. Sage Publications, 1995. 372 p.