

Referência completa para citação

FREITAS (H.), JANISSEK (R.) e MOSCAROLA (J.). Análise qualitativa em formulário interativo: rumo a um modelo cibernético conjugando análises léxica e de conteúdo. CIBRAPEQ-Congresso Internacional de Pesquisa Qualitativa, 24 a 27 de março, Taubaté/SP, 2004. Poster e Workshop. 16 p..

Análise qualitativa em formulário interativo: rumo a um modelo cibernético conjugando análises léxica e de conteúdo

Henrique FREITAS (hf@ea.ufrgs.br)
Professor do PPGA/EA/UFRGS, Pesquisador CNPq
Doutor em Gestão (UPMF, Grenoble, França)

Raquel JANISSEK-MUNIZ (janissek@tiscali.fr)
Doutoranda em Gestão (UPMF, Grenoble, França) - Bolsista CAPES

Jean MOSCAROLA
Professor da Université de Savoie, GEREG/CNRS (Annecy, França)
Doutor em Economia Aplicada (Paris, França)

Resumo: Com alguns anos de experiência em tecnologias computacionais e método, faz-se um mix de idéias, conceitos e práticas e concebe-se um modelo para análise de dados qualitativos a partir da combinação de técnicas e ferramentas de análises léxica e de conteúdo, passando por alguns estágios, e podendo mesmo ser realizado via web. Ao mesmo tempo em que lê e analisa, 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 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; bem como a produção de um julgamento a partir da própria leitura do mesmo processo de análise de conteúdo, usando um protocolo definido a priori; e ainda uma técnica alternativa, a qual, para criar o protocolo emergindo do texto lido um a um, parte não do marco zero, 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 de sua percepção e tenha ao mesmo tempo uma noção do resultado da análise em curso, em todos seus aspectos.

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

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 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. Esta iniciativa nasceu da compreensão de que diversos pesquisadores pensam, de forma equivocada, que a escolha deve recair sobre uma ou outra técnica. Ora, a riqueza de análise só pode ser 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, revisando a literatura inerente, bem como procura-se desenvolver exemplos de certas abordagens consideradas de interesse do analista, a partir do uso da ferramenta Sphinx Léxica® (FREITAS, JANISSEK, MOSCAROLA e BAULAC, 2002). Um caso é desenvolvido passo a passo, a partir de uma pesquisa essencialmente qualitativa. Nosso intuito é, sobretudo, fornecer aos gerentes, pesquisadores, professores e alunos de diversos níveis de formação, 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. É o que se segue, nas seções 2 a 7.

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 por exemplo 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 certo grupo por algum produto, sistema, serviço, etc.

Como o pesquisador, o analista, ou mesmo o gerente ou executivo 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, a de help desk ou outro, e se dar conta das principais reclamações ou sugestões atuais dos seus clientes? Não seria este tipo de dado, que fica esquecido, engavetado ou nem mesmo pensado como um recurso, uma fonte nobre de dados gerados sem nenhum custo adicional para a organização? Ou como o pesquisador interpretativista poderia, antes de iniciar toda sua labuta, ter uma noção do ‘terreno’ que começa a tentar compreender?

Como poderia aquele mesmo gerente ousar pensar nisso se ele nem imagina que existam métodos e técnicas para tal? Ele pode muito bem 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 talvez que poderia saber disso sem mesmo ter que indagar aos clientes. De fato, a técnica quantitativa e supostamente bem 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, ainda por cima a partir da exploração de um dado completamente espontâneo, não induzido de forma alguma.

Não são só as técnicas de coleta e de análise quantitativa de dados que foram aperfeiçoadas com o tempo. Hoje em dia, são factíveis e diversas as formas e tecnologias de se capturar automática e praticamente gratuitamente dados qualitativos (subjetivos, textos enfim) que nos denotem as preferências e comportamentos de um certo grupo de pessoas. Os procedimentos, métodos e ferramentas que possibilitam isso são cada vez mais presentes na literatura e no mercado. Há todo um leque de possibilidades que pode ser utilizado (MOSCAROLA, 1990; WEITZMAN e MILES, 1995; FREITAS, CUNHA e MOSCAROLA, 1997).

O que existe, contudo, é uma confusão conceitual ou mesmo prática sobre em que consistem e mesmo de como essas técnicas e ferramentas poderiam ser utilizadas. Em que implica cada uma delas? Como se poderia aplicar e explorar em benefício próprio cada uma delas? Enfim, há mesmo uma ignorância sobre a existência da análise léxica e da análise de conteúdo¹, sobre as distintas finalidades de uma e outra. Há ainda uma idéia equivocada de que se deve escolher uma OU outra para uma dada análise ou exploração de dados, quando na verdade procura-se aqui demonstrar que pesquisadores acadêmicos ou analistas em geral DEVEM ter uma clara visão das vantagens em explorar estes recursos de análise de forma sequencial, recorrente e sobretudo complementar.

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. É pois tempo de método múltiplo de pesquisa, onde o pesquisador ou o analista comercial, financeiro, etc, tem cada vez mais e mais clara a necessidade de estar em contato com diferentes visões das situações reais. E isto utilizando-se de fontes de dados e de técnicas diversas, levando em consideração o rigor das pesquisas. Isto só reforça o valor,

¹ 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.

utilidade e oportunidade de se abordar de forma mais organizada e mesmo didática as técnicas e ferramentas no que se refere a dados qualitativos.

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.

Tudo isto valoriza a questão de como se tratar questões dtas qualitativas. Tais debates foram objetos de congresso internacional (LEE, LIEBENAU e DEGROSS, 1997). A estratégia de uso de questões abertas ou fechadas num instrumento é debatida por LEBART e SALEM (1994, p.25-28). Além disso, nas pesquisas, de forma geral, uma etapa pouco valorizada ou um tanto negligenciada é a de preparação de dados: há todo um esforço para que os dados sejam coletados, de forma que não se dispõe de muito tempo quando os dados poderiam e deveriam ser analisados a fundo, em termos de significado e implicações. Ocorre que, antes de analisar os dados e cruzá-los das mais diferentes formas na busca por relações causais ou possíveis explicações para hipóteses previamente formuladas, é necessário que sejam preparados certos tipos de dados (idade em faixa etária, por exemplo).

Da mesma forma, é necessário um investimento (mais de tempo de análise do que outro recurso) para que sejam criadas categorias pertinentes e claras a partir de dados tipo texto, ou seja, efetivamente qualitativos. Um exemplo pode ser a criação de um dado novo a partir de uma questão texto que poderia ser "*quais as sugestões que você apresentaria para a melhoria dos nossos serviços?*" ou então mesmo "*quais as principais reclamações que você poderia apresentar?*". Normalmente, pesquisadores e analistas desavisados ou talvez despreparados ou inadvertidos sobrecarregam suas pesquisas com questões abertas, criando uma expectativa em quem responde, a qual se torna 'falsa expectativa', visto que tais respostas nunca serão tabuladas e avaliadas como se deveria. Logo, o aprendizado que tal conteúdo permitiria não será jamais produzido. Cuidado especial é necessário no sentido de não se abusar de questões abertas num instrumento de pesquisa, pois seu 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. Ou seja, ao invés de perguntar se o cliente está muito ou pouco satisfeito, o avaliador poderia ler cada uma das respostas emitidas e registradas e iria ele próprio julgar se o respondente está ou não satisfeito. 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. Este seria um motivo forte para que se repense cada decisão ou solicitação de codificação a terceiros: a pessoa está implicada no processo? Conhece o *métier* ou contexto pesquisado? Tem engajamento em relação ao objeto pesquisado? Sobretudo, tem como associar a sua tarefa isolada de codificação a todo resto do processo e atividades desse projeto? Terá a iniciativa de sugerir novas dimensões de análise identificadas no decorrer da própria atividade de codificação? Talvez seja melhor arrumar uma resposta adequada a estas questões ou então de fato decidir que o pesquisador se responsabilizará pela realização da tarefa ou pela supervisão estreita desse processo.

Normalmente, uma codificação derivada de uma análise vai gerar um novo dado de múltipla escolha (das diversas categorias criadas, algumas serão marcadas para cada resposta emitida por cada respondente). Este novo dado retratará de forma bem mais objetiva o conteúdo de cada opinião ou resposta. A construção das opções ou categorias desse novo dado gerado deve ser realizada segundo certas regras, de forma a se poder ter confiança na sua coerência, consistência, força enfim. Entretanto, muito da prática empresarial que se tem podido observar (em contato com institutos de pesquisa por ocasião de assessorias de tecnologia ou método) mostra que pessoas não tão bem preparadas realizam tal função. Ocorre que tal novo dado vai ser certamente a base para decisões estratégicas ou operacionais importantes para a organização envolvida. Logo, é de extrema importância a adequada e mesmo rigorosa

formulação destas categorias. De fato, cada categoria pode implicar em ações empresariais, que exigirão alocação de recursos, etc., ou seja, as implicações das corretas definições de categorias podem ser enormes para a organização. Caso as categorias não sejam pertinentes, todas análises, decisões e ações estarão igualmente comprometidas. Não é esta uma atividade para ser realizada com adivinhações ou amadorismos de qualquer sorte. Conceitos, técnicas, métodos e ferramentas são em muito necessários. É fundamental que se desenvolva um conjunto de diretrizes ou procedimentos que possam ser seguidos.

A preparação e análise de dados provenientes de pesquisas ou captura de dados, sejam eles qualitativos ou quantitativos, passa então pela identificação e categorização adequada dos seus conteúdos, na busca pela produção de conhecimentos e identificação de relações que nos permitam avançar na compreensão dos fenômenos investigados. Esta busca por informações consistentes, relevantes e fidedignas requer tempo e dedicação por parte do analista ou pesquisador, no intuito de gerar resultados que traduzam a realidade ou contexto estudado. Entretanto, não raro estas análises levantam dúvidas quanto à efetividade dos termos e resultados obtidos, especificamente pela não sistematização da sequência de análise dos dados ou por outros motivos, como a inadequação do próprio perfil ou preparação de quem analisa, como recém evocado.

Ainda, há toda uma polêmica sobre a legitimidade de se ter somente uma pessoa responsável por uma dada codificação: seria válida tal categorização? Por quê? Seria necessário se ter mais de uma pessoa a codificar? Será que realmente esta segunda pessoa aportaria um ganho de qualidade na codificação realizada? Será que ela estaria engajada o suficiente no mesmo processo de forma a contribuir? Será ainda que este engajamento seria condição importante? Será que tal esforço e mesmo gasto de tempo seria necessário e mesmo compensador? Estas questões foram abordadas em 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, puramente quantitativos, tal qual os imagina-se - 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, trechos de livros, 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.

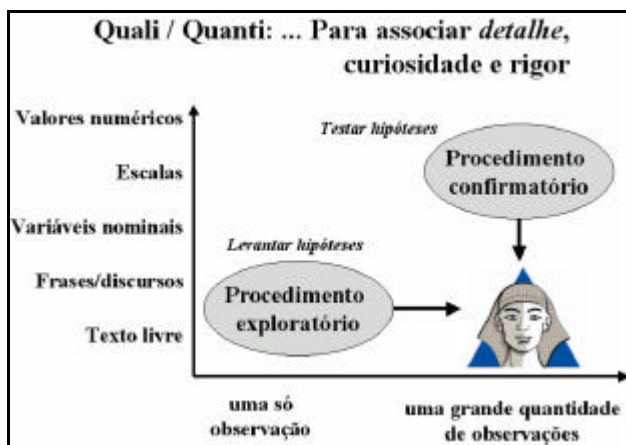


Figura 1 - O desafio de associar qualitativo e quantitativo (FREITAS e MOSCAROLA, 2000)

Já a abordagem mais científica, dita **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".

A utilidade da combinação destes métodos tem sido assunto discutido fortemente na comunidade acadêmica internacional (MASON, 1997; LEE, LIEBENAU e DEGROSS, 1997), apresentando-nos grande variedade de aspectos a serem levados em conta quando da formulação de questões com vistas à obtenção e análise de dados, sejam eles qualitativos x quantitativos, sejam eles diretos x indiretos, abertos ou fechados. É tempo de seguir em frente com mais estudos qualitativos (MILES e HUBERMAN, 1997), e educar nossos gerentes, começando pelas nossas crianças, que o mundo não é somente quantitativo, e sim qualitativo. Pelo menos, que bom estudo quantitativo não deveria ser precedido por um qualitativo? CRESWELL (1998), e também KIRK e MILLER (1986) oferecem-nos alguns conceitos e discussões a respeito de pesquisa qualitativa e principalmente sobre confiabilidade (ou fidedignidade) e validade desse tipo de estudo. Da mesma maneira, a revista MIS Quarterly (v.21, n.º.3, September 1997) publicou 3 papers baseados em pesquisa qualitativa deste mesmo assunto.

MASON (1997, p.9-34) oferece-nos alguns *insights* sobre essa discussão: o que deveria ser uma pesquisa qualitativa? De que forma relacionar aspectos quali e quanti em uma pesquisa? Estas pesquisas, seja na obtenção dos dados seja no tratamento destes, deveriam ser sistemática, rigorosa e estrategicamente conduzidas, ainda flexíveis e contextuais, envolver críticas auto-construtivas pelo pesquisador, ou provocar reflexões. Deveriam produzir explicações sociais para *charadas* ou *quebra-cabeças* intelectuais permitindo a generalização de alguma forma ou então, permitindo ampla aplicação. Não ser vistas como um corpo unificado de filosofia e prática, cujos métodos podem simplesmente ser combinados sem problemas, nem ser vistas como necessariamente opostas e não complementares. Ainda, deveriam ser conduzidas como uma prática ética levando em consideração o contexto político.

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, 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 conteúdo²

A análise de documentos, sejam eles 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.

Com o uso conjunto destas técnicas de análise de textos 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, ou seja, o uso destas duas técnicas encobrem diversas das possibilidades que dali poderiam surgir ou fazer emergir. Ao final de um esforço de análise de dados, poder-se-ia dispor de resultados significativos aplicáveis a uma dada realidade. Como isso deve ser feito? De que maneira combinar o uso destas duas formas de análise? Os dois tipos de análise de questões abertas são representados na Figura 2 (FREITAS e MOSCAROLA, 2000).

A Análise Léxica (FREITAS e MOSCAROLA, 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 (FRANKFORT-NACHMIAS e NACHMIAS, 1996, p.324-330), que consiste em uma leitura aprofundada de cada uma das respostas, onde, codificando-se cada uma, obtém-se uma idéia sobre o todo (FREITAS, 2000). WEBER (1990, p.10) apresenta algumas vantagens da análise de conteúdo,

² Estes tipos de análises de dados são abordados por KRIPPENDORFF (1980); WEBER (1990); GRAWITZ (1993); LEBART e SALEM (1994); MOSCAROLA (1990, 1993 e 1994); GAVART-PERRET e MOSCAROLA (1995); BARDIN (1996); FREITAS, CUNHA e MOSCAROLA (1996, 1997).

destacando sua aplicabilidade na análise de textos de comunicação de toda natureza, bem como o fato de permitir combinar métodos quanti e qualitativos, e mesmo explorar séries longitudinais de documentos ou fontes múltiplas, e enfim o fato de poder tratar com dados ditos mais espontâneos (e não induzidos ou expressamente perguntados).

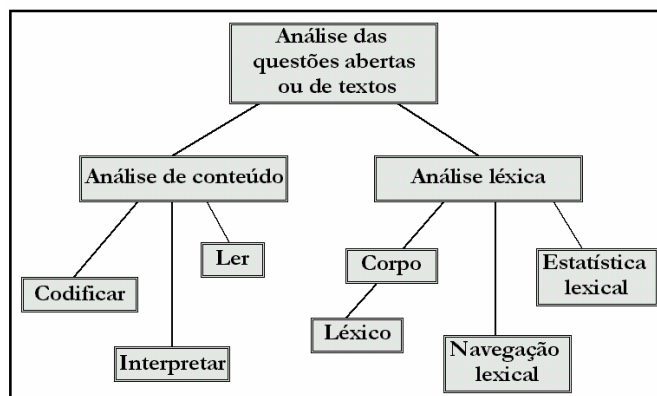


Figura 2 - A análise de conteúdo e a análise léxica (FREITAS e MOSCAROLA, 2000)

Com base em textos europeus (GRAWITZ, 1993; GAVARD-PERRET e MOSCAROLA, 1995), FREITAS, CUNHA e MOSCAROLA (1997) apresentam noções gerais sobre a análise de dados textuais, com sete diferentes níveis de aplicação e desenvolvimento da análise léxica, ilustrando a aplicação e uso dessas técnicas. Este tipo de análise contribui para a interpretação das questões abertas ou textos, a partir da descrição objetiva, sistemática e quantitativa do seu conteúdo. Diversas referências podem ser consultadas para um maior aprofundamento nas questões da análise de dados qualitativos. WEBER (1990) publicou o livro “Basic Content Analysis”, SILVERMAN (1993) descreve métodos para analisar discursos, textos e interações, no livro “Interpreting Qualitative Data”, MILES e HUBERMAN (1994) publicaram o livro “Qualitative Data Analysis”, o *Journal of Applied Management Studies*, inglês, publicou dois papers com aplicação de análise de conteúdo e análise léxica: “The Structure Of Managerial Thinking”, e “Managerial Discourses...”. A *MIS Quarterly* também publicou pesquisa qualitativa durante 1997, mostrando disposição da comunidade (e dos editores!) rumo a estudos qualitativos. A literatura francesa também está recebendo novos textos, da parte de BARDIN (1996), e de LEBART e SALEM (1994).

A análise léxica oferece uma maneira científica de desenvolver investigações em áreas que tradicionalmente são objetos de uma única abordagem literária e subjetiva. São diversas as fontes e motivações que podem viabilizar a coleta de dados textuais (todo dado cujo conteúdo é texto ou alfanumérico): questões abertas de pesquisas de opinião de mercado e outras; o conteúdo de respostas de entrevistas; mensagens recebidas e enviadas; livros, artigos e textos quaisquer; campos tipo texto nas bases de dados corporativas ou outras; o registro automatizado de telefonemas profissionais ou mesmo pessoais (como os de serviço de atendimento ao cliente); o registro automatizado das operações de consulta de sistemas corporativos como o registro de logs Internet, Intranets e outros sistemas, como por exemplo transações bancárias, suporte técnico, todas operações das caixas de supermercado, etc. Todas estas fontes de dados devem ser exploradas via quantificação e resumo, de forma a permitir a melhor compreensão e interpretação do seu conteúdo, visando daí depreender idéias, ações, enfim, produzir uma informação mais rica para embasar as decisões.

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

Para ilustrar as formas de análise de questões abertas nas enquetes 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).

FICHA TÉCNICA DA ENQUETE		
POTENCIAIS INTERESSADOS	Federação das Indústrias, ANPAD, PPGA/EA/UFRGS Sistemas SEBRAE, SOFTEX, SEPRO, ASSEPRO, 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, a partir disso, construir.

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 ao lado.

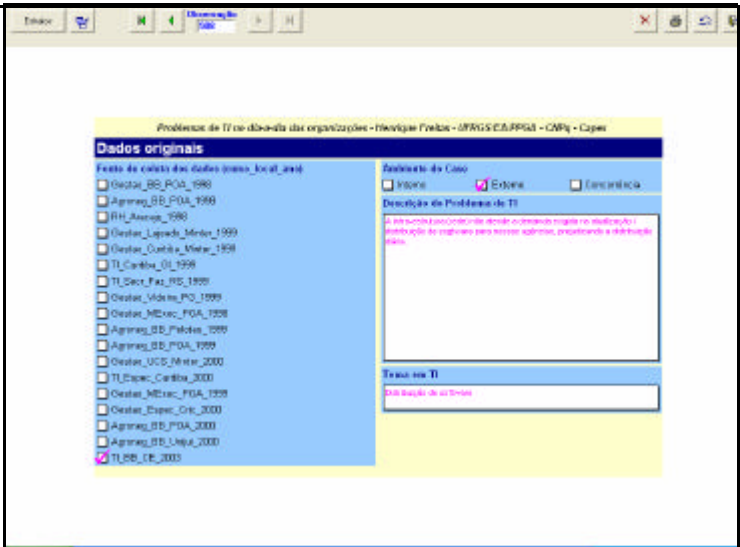


Figura 4 – A definição da base de dados inicial

A partir disso, foram derivados dados novos com base na identificação do grupo, como ilustrado. Com uma coleta simplificada, passa-se a dispor de uma base com elementos 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.

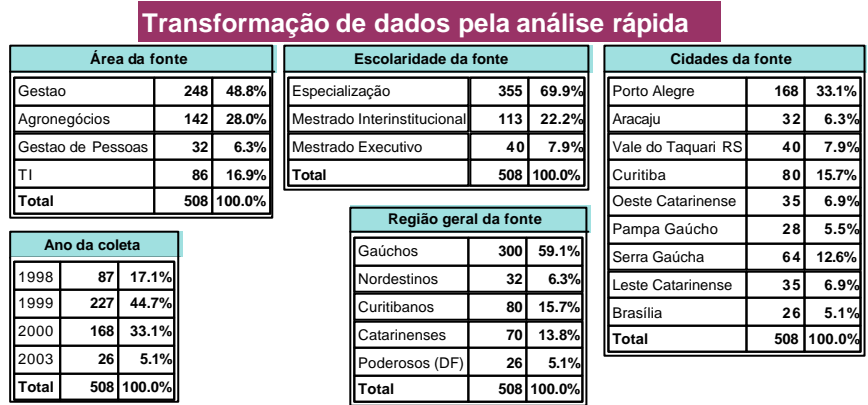


Figura 5 – Dados derivados do dado de identificação anotado na coleta da dados

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.

Figura 6 – Novas questões, desafio do analista

Nosso intuito passa a ser demonstrar algumas técnicas para preparação dos dados abertos textuais, em especial (6.1) a produção de um protocolo rápido a partir da análise léxica, com contagem de palavras e agregação em palavras-chave e depuração do *corpus*; (6.2) a produção de um protocolo que emerge de uma análise de conteúdo, depoimento a depoimento; bem como (6.3) 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, ou seja, algo que se tinha em mente, uma curiosidade ou mesmo uma teoria que se quer confrontar aos depoimentos coletados; assim como uma técnica alternativa (6.4), 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 mais sofisticado (6.5), 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.

6.1. 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), de onde recuperamos sumariamente alguns princípios. 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?

Figura 7 – Ambiente de análise de dados do Sphinx Léxica®

Essa questão, que reunia as respostas do problema em si e do seu tema ou título, foi criada por um recurso do sistema utilizado (o Sphinx Léxica® - cujas características estão apresentadas em <http://www.sphinxbrasil.com>).

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 6) deveriam ser preenchidas pelo analista, para o que diferentes técnicas serão descritas, da mais clássica à mais 'cibernética' (como se enfatizou no título deste artigo), com recursos via Web.

O ambiente de análise léxica evocado é o ilustrado na Figura 8 ao lado. Ele disponibiliza 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.

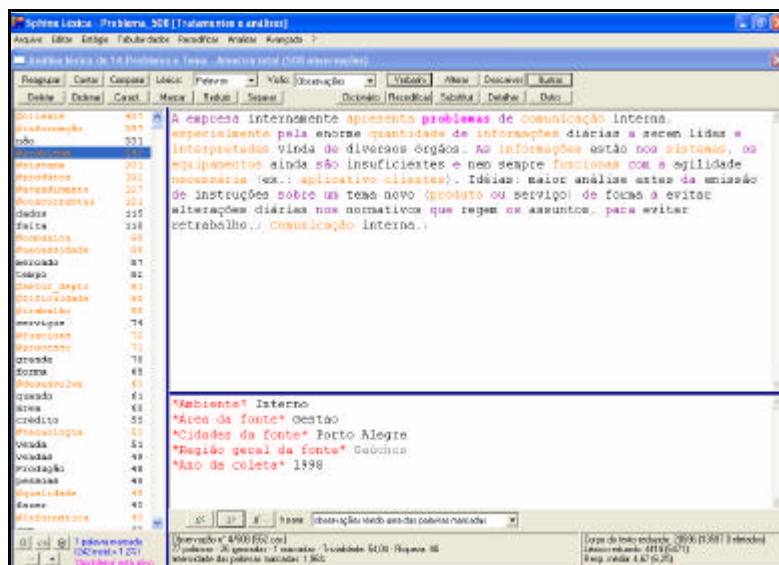


Figura 8 – Ambiente de análise léxica do Sphinx Lexica®

Na Figura 8 (acima), pode-se rapidamente 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 ³, 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 o que possibilitam a navegação e 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 ou conteúdo no sentido da análise realizada, como as preposições, os artigos, etc.) e chegou-se às palavras de conteúdo ou significado. Essas palavras podem ser marcadas no *corpus* e dali origina-se um novo dado, ilustrado na Figura 9 ao lado, onde registraram-se as 22 principais (as mais frequentes). A Figura 16 mais adiante ilustra esse procedimento.

Toma-se conhecimento do que evocam 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? 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 10). Este arquivo é chamado **verbatim**, que significa citação. A navegação lexical irá permitir a substituição das respostas, organizando-as de acordo com seu conteúdo, possibilitando conhecer em que contexto e com que sentido foi empregada a palavra em questão. Pode-se mesmo segmentar os depoimentos ou relatos com base em outro dado objetivo (a região da pessoa, por exemplo).

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%
Total	1841	100.0%

Figura 9 – Palavras ou idéias-chave

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 da Figura 10 apenas alguns dos ‘gaúchos’ estão presentes. Dois cliques de mouse permitem realizar esse tipo de operação. Nosso intuito é colocar ênfase no avanço metodológico, porém viabilizado pela tecnologia disponível.

³ Palavras seriam consideradas como uma forma gráfica, uma seqüência de caracteres que não comporta separador.

Textos extraídos de '14.Problema e Tema'

Lista das 1 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:

n° 4 : A empresa apresenta PROBLEMAS de comunicação interna, especialmente pela enorme quantidade de informações diárias a serem lidas e interpretadas vinda de diversos órgãos.

n° 8 : Precisamos saber se o PROBLEMA é rejeição do cliente por não possuir cultura voltada para o produto ou se nosso marketing está inadequado - PROBLEMAS na venda/aceitação.

n° 13 : PROBLEMAS de acessar sistemas nas horas de pico.

Figura 10 - 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 11. 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 ao lado). 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 12 mostra o tipo de uso e exploração que pode ser feito a partir do dado novo criado com as palavras ou idéias mais frequentes.

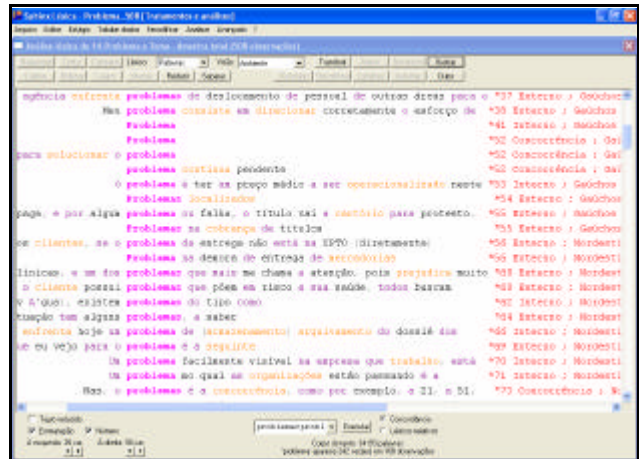
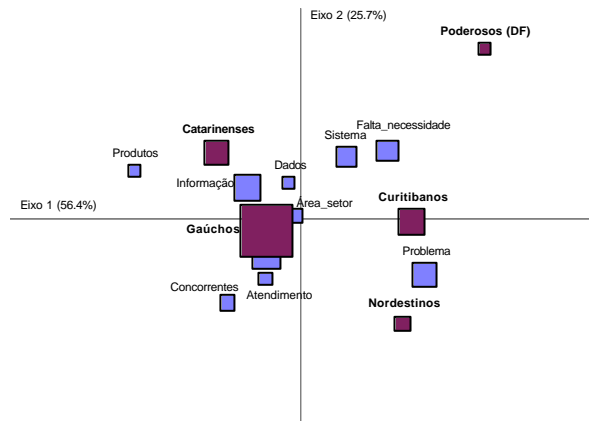


Figura 11 – Contexto da palavra 'problema' e sua ilustração

Grupo n°1						
	Gaúchos	Nordestinos	Curitibanos	Catarinenses	Poderosos (DF)	Total
...	365	51	97	65	31	609
Cliente	145	16	38	36	4	239
Informação	127	9	31	39	7	213
Problema	91	18	37	13	9	168
Falta_necessidade	69	10	28	22	11	140
Sistema	74	3	23	13	9	122
Concorrentes	51	4	10	8	0	73
Área_setor	42	5	11	11	3	72
Dados	41	5	9	12	4	71
Produtos	46	2	3	14	2	67
Atendimento	44	6	7	8	2	67
Total	1095	129	294	241	82	1841

Figura 12 – As 10 principais idéias-chave e a região e a 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 ilustra a Figura 12 e o que se pode rapidamente depreender.

Uma curiosidade que se pode apresentar em complemento consiste em averiguar ou medir a **dimensão das respostas**: as pessoas responderam de forma extensa ou concisa? Cria-se uma hipótese de que aqueles que deram respostas extensas têm um interesse maior do que os demais. Uma hipótese discutível que dará, porém, uma indicação.

Basta criar, por um botão especial, um novo dado quantitativo, que registra a quantidade de palavras por relato, organizá-lo por exemplo em função da média de palavras por depoimento (no caso = 61 palavras, com desvio-padrão de 40 palavras), e depois, por exemplo, cruzar este novo dado com um dado objetivo, como a 'região do respondente'. É o que ilustra a Figura 13, onde se observa que 58,9% dos relatos foram com quantidade de palavras abaixo da média. Contudo, poder-se-ia especular que os 'poderosos' são mais objetivos, que os 'nordestinos' explicam em detalhe, como de certa forma os 'catarinenses' e ainda os 'gaúchos'. São pistas a investigar.

Pode-se considerar que se tem um indicador, grosseiro, do interesse das pessoas de diferentes regiões ao relatar os problemas. No entanto, esta análise necessita maiores investigações e cuidado nas conclusões, visto que se tem distorções de amostra e que o conteúdo em si pode levar a outras pistas.

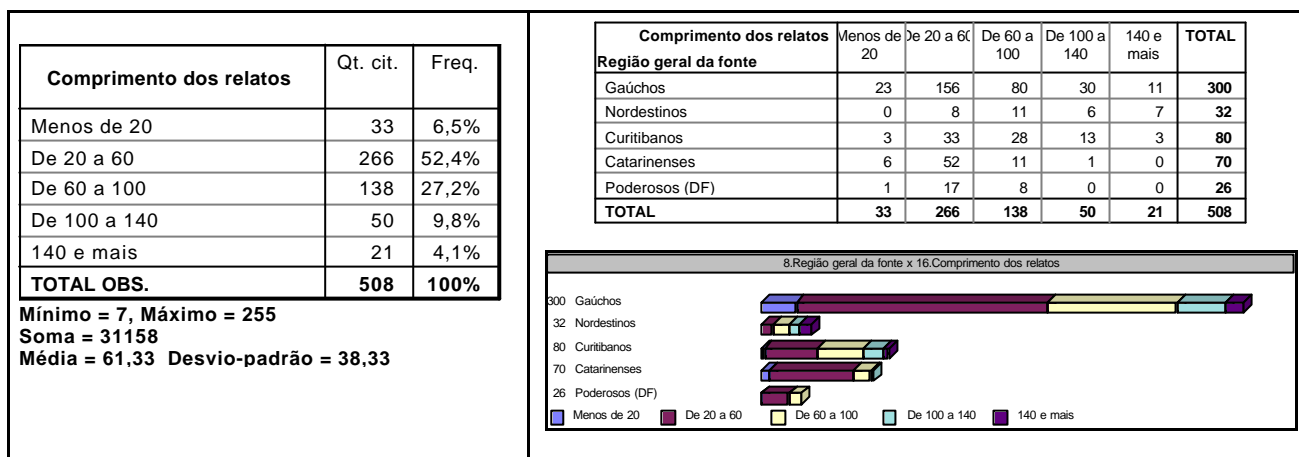


Figura 13 – Quantidade de palavras por depoimento: os relatos são eles longos?

6.2. A produção de um protocolo (idéias-chave) que emerge de análise de conteúdo, relato a relato

A Análise de Conteúdo consiste em uma leitura aprofundada de cada uma das respostas, onde, 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 em profundidade 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 14), 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 14 apresenta o início de uma codificação que evolui naturalmente à medida que se avança na leitura de cada resposta e conseqüente criação de novos códigos ou categorias de respostas a marcar.

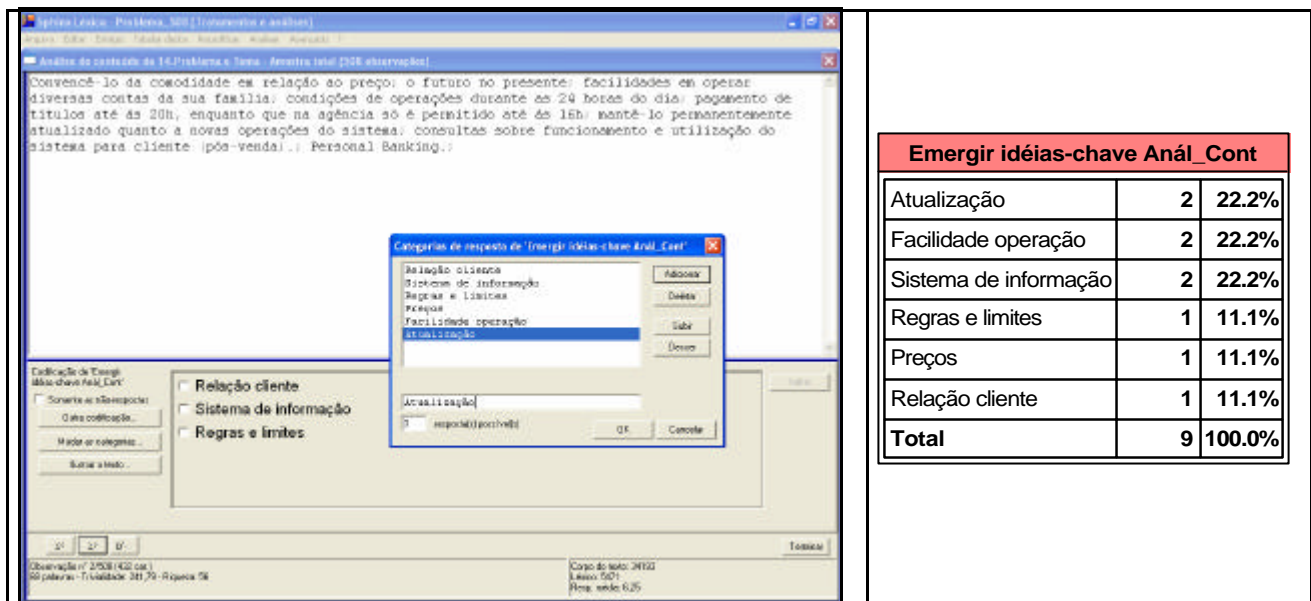


Figura 14 - A análise de conteúdo, com a codificação inicial fazendo 'emergir' protocolo e o resultado derivado tabulado

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 14 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 (FREITAS e JANISSEK, 2000). 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.

6.3. A produção de um julgamento a partir da própria leitura do mesmo processo de análise de conteúdo, usando um protocolo definido a priori

A atividade do analista por ocasião de uma análise de conteúdo demanda paciência, foco, atenção. É de certa forma vista como um 'luxo ao qual o pesquisador não poderia se permitir'. Contudo, é enorme a riqueza de recursos que o próprio analista passa a ter, realizando ele próprio tal atividade. 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. Ou seja, algo que se tinha em mente, uma curiosidade ou mesmo uma teoria que se quer confrontar aos depoimentos coletados. 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 ou evoca 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, como esse dos que 'apenas choram seus problemas' e daqueles que, mesmo com dificuldades, 'vislumbram soluções'. Isto somente a título de ilustração, retrata-se na Figura 15. Os dados da tabulação apresentada na Figura 15 referem-se a uma codificação parcial, mas servem para ilustrar o resultado da operação de análise realizada.

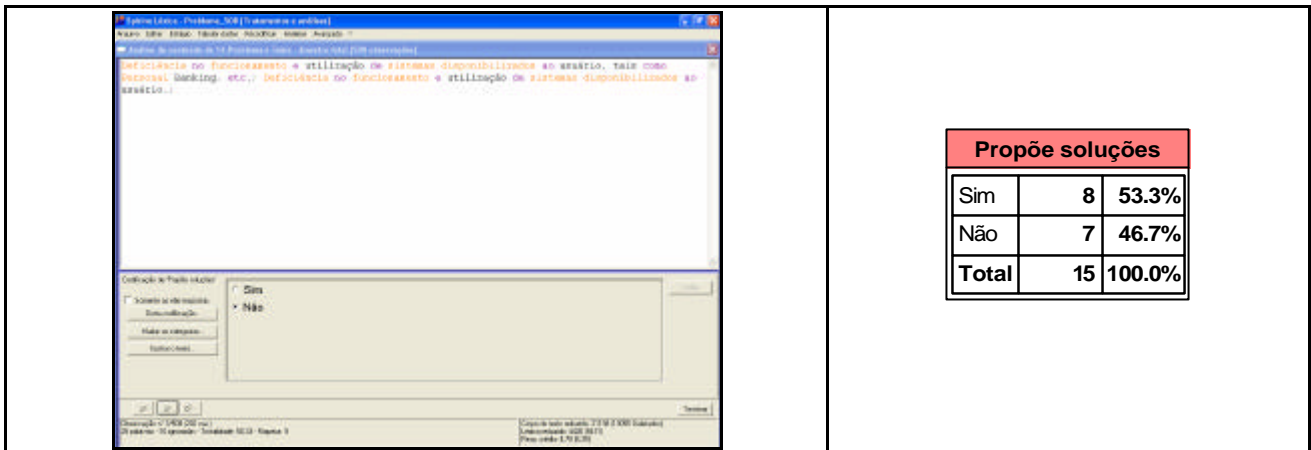


Figura 15 - A análise de conteúdo, com a codificação baseada em protocolo 'a priori' e o resultado derivado tabulado

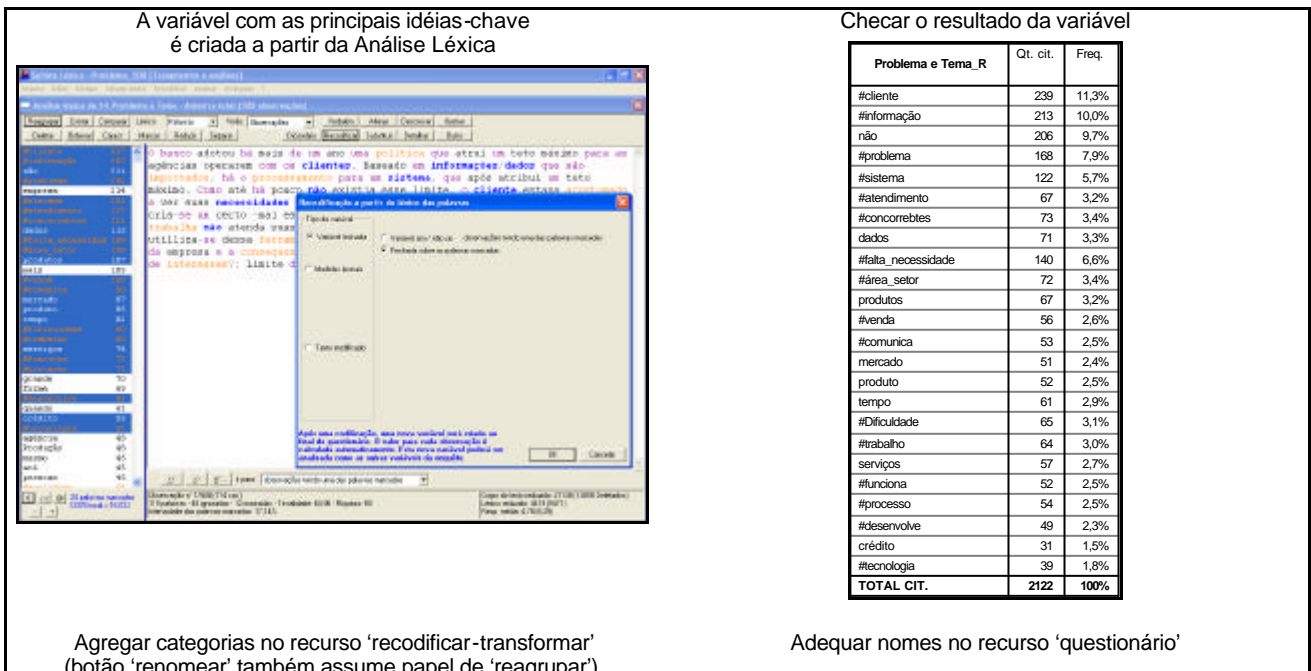
6.4. O protocolo emergindo do texto lido depoimento a depoimento, partindo do resultado bruto da análise léxica

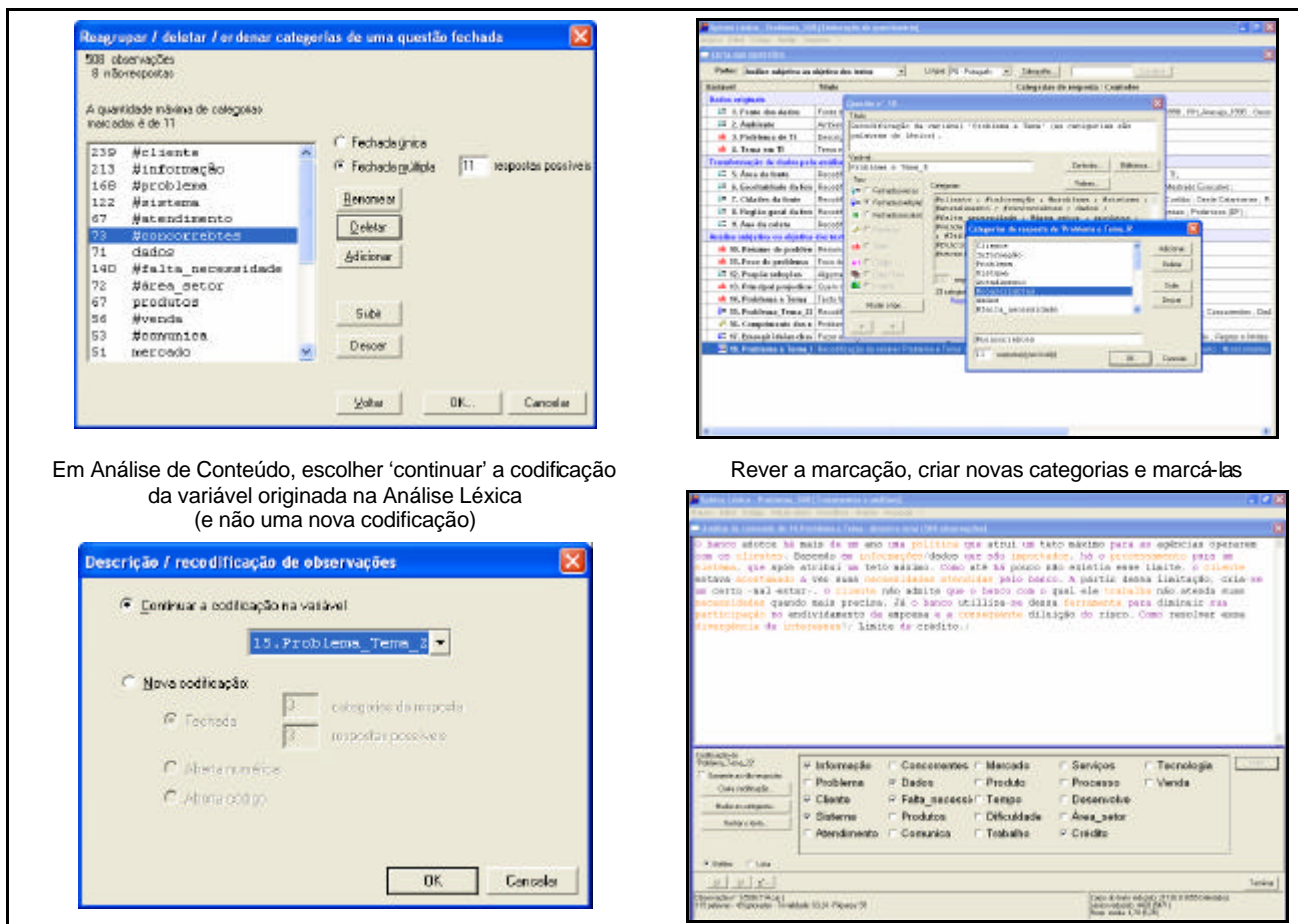
Uma astúcia possível, interessante, é a seguinte: combinar o resultado de uma análise léxica (visto em 6.1) com a atividade de fazer emergir um protocolo a partir da leitura do texto, ou seja, a análise de conteúdo para geração de idéias-chave (vista em 6.2).

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.

Trata-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 (ilustrado e definido em 6.1).

A Figura 16 a seguir ilustra todo esse procedimento.





Em Análise de Conteúdo, escolher 'continuar' a codificação da variável originada na Análise Léxica (e não uma nova codificação)

Rever a marcação, criar novas categorias e marcá-las

Figura 16 - 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.

6.5. Um modelo integrando técnicas para facilitar a expressão e análise automática da subjetividade do julgamento ou percepção dos analistas: rumo ao espaço cibernético!

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 17.

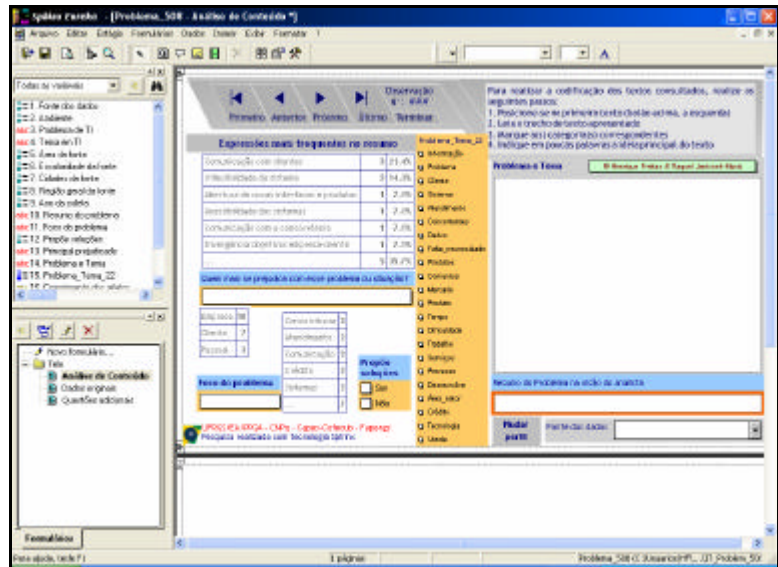
Concepção do ambiente ⁴

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.

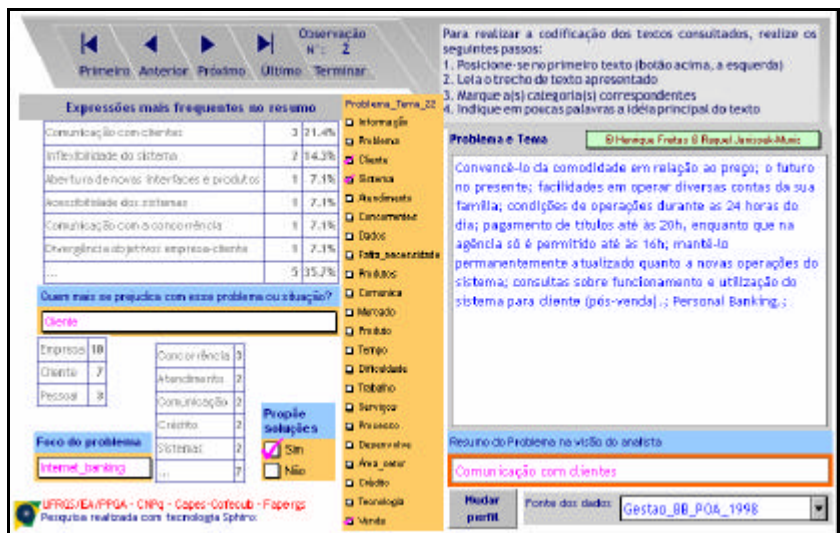


Figura 17 - A análise léxica automatizada, a partir dos textos de base e da reflexão do analista

⁴ O estagiário Ricardo Simm Costa, estudante da Escola de Administração da UFRGS, cooperou com os autores na realização operacional deste modelo.

7. Conclusão: idéias simples de preparação de dados para uma análise mais rica e efetiva

Este artigo já bate na centena de artigos de nossa equipe. 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 – reforce-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.

Referências Bibliográficas

- BARDIN, L. L'analyse de contenu. Paris: PUF, 8^{ème} Edition, 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**. Sphinx, Porto Alegre, 2000.
- FREITAS, H.; MOSCAROLA, J. Análise de dados quantitativos e qualitativos: casos aplicados usando o Sphinx®. Porto Alegre, Sphinx, 2000, 176 p.
- FREITAS, H. As tendências em Sistemas de Informação com base em recentes congressos. Porto Alegre: ReAd (<http://www.adm.ufrgs.br/read>), No.13, Jan. 2000, 20 p.
- FREITAS, H., CUNHA Jr., M. V. M. e MOSCAROLA, J. Pelo resgate de alguns princípios da análise de conteúdo: aplicação prática qualitativa em marketing. Angra dos Reis - RJ: Anais do 20º ENANPAD, ANPAD, Marketing, 23 - 25 de Setembro 1996, pp. 467 - 487.
- FREITAS, H., CUNHA Jr. M.V.M. e MOSCAROLA, J. Aplicação de sistema de software para auxílio na análise de conteúdo. São Paulo: RAUSP, v. 32, nº 3, Jul/Set. 1997, pp. 97-109.
- FREITAS, H., JANISSEK, R., MOSCAROLA, J. e BAULAC, Y. (2002) – Pesquisa interativa e novas tecnologias para coleta e análise de dados usando o Sphinx®. Porto Alegre, Sphinx, 2002, 381p.
- GAVARD-PERRET, M.L. & MOSCAROLA, J. De l'énoncé à l'énonciation: relecture de l'analyse lexicale en marketing. Annecy, France: GERE and CERIAM, Université de Savoie, Cahier GERE, 1995, 34 p.
- GRAWITZ, M. Méthodologie des Sciences Sociales. Paris: Dalloz, 1976 (chap. II, pp. 586-627), and 1993 (9 ed., chap. 2, section 2, pp. 532-565). 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, USA: Proceedings ICIS, Chapman & Hall, 1997.
- LESCA, H., FREITAS, H. e CUNHA Jr., M. **Instrumentalizando a Decisão Gerencial**. Rio de Janeiro - RJ: Revista Decidir, ano III, n. 25 Agosto 96, pp. 6 - 14.
- MASON, J. Qualitative researching. Sage, 1997, 180 p.
- MILES, M. B. & HUBERMAN, A. M. Qualitative data analysis. Sage Publications, 1994, 338 p.
- MOSCAROLA, J. Enquêtes et Analyse de donnés avec le Sphinx. Paris: Libraire Vuibert, 1990.
- MOSCAROLA, J. Analyse de contenu et analyse de données: solutions logiciels pour une intégration quali/quant. Paris: Actes JADT, 1993.
- MOSCAROLA, J. La communication politique vue par l'analyse lexicale. Annecy, France: IUP, Université de Savoie, Cahier du GERE, 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.