

Técnica inovadora para análise de dados qualitativos

Henrique Freitas

Escola de Administração - Universidade Federal do Rio Grande do Sul (UFRGS)
Avenida Washington Luis, 855 - 90010-460 - Porto Alegre - RS - Brazil

hf@ea.ufrgs.br

RESUMO

Propõe-se um modelo para análise de dados qualitativos a partir de técnicas e ferramentas de análises léxica e de conteúdo: 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. O modelo proposto integra as técnicas mais conhecidas, permitindo que o analista registre a subjetividade de sua percepção e tenha ao mesmo tempo uma noção do resultado da análise em curso.

1. INTRODUÇÃO

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 [FREITAS e JANISSEK, 2000].

Discute-se aqui sobre a questão dos dados qualitativos e das técnicas de análise, desenvolvendo-se exemplo a partir do uso de ferramenta apropriada [FREITAS, JANISSEK, MOSCAROLA e BAULAC, 2002]: nosso intuito é 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.

2. ANÁLISE DE DADOS QUALITATIVOS

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.

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.

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; LEE, LIEBENAU e DEGROSS, 1997]. A estratégia de uso de questões abertas ou fechadas num instrumento é debatida por LEBART e SALEM (1994).

Naturalmente, 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?*”. 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.

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. 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’).

A utilidade da combinação de métodos tem sido assunto discutido fortemente na comunidade acadêmica internacional [MASON, 1997; LEE, LIEBENAU e DEGROSS, 1997; KELLE, 1995], 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, 1994], 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.

3. TÉCNICAS DE ANÁLISE TEXTUAL

A análise de documentos, sejam eles originários de pesquisas quali ou quantitativas, inclui análise léxica e análise de conteúdo. 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). Apresentam 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 apresentados por 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], 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) apresenta algumas vantagens da análise de conteúdo, 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 mais espontâneos (e não induzidos ou expressamente perguntados).

Com base em textos europeus e americanos [WEBER, 1990; SILVERMAN, 1993; MILES e HUBERMAN, 1994; BARDIN, 1996; LEBART e SALEM, 1994; GRAWITZ, 1993; GAVARD-PERRET e MOSCAROLA, 1995], FREITAS e JANISSEK (2000) apresentam noções gerais sobre a análise de dados textuais, com 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.

4. INTERFACE PARA ANÁLISE DE DADOS QUALITATIVOS

Para ilustrar as formas de análise de questões abertas nas enquêtes que se tem praticado, usaremos uma pesquisa realizada junto a 508 gerentes de diferentes áreas, todos se manifestando sobre problemas que enfrentam e suas organizações quanto à informação ou à tecnologia da informação (TI). Uma questão, reunindo as respostas do relato do problema em si e do seu tema, foi criada por um recurso do sistema utilizado (o Sphinx Léxica® - cujas características estão apresentadas em <http://www.sphinxbrasil.com>).

Definimos [FREITAS, JANISSEK-MUNIZ e MOSCAROLA, 2004] recentemente um ambiente onde se aplicam algumas 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 e depuração do *corpus*; 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, 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, 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.

Neste artigo propõe-se um modelo integrando essas técnicas todas, recentemente concebido, 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. Do analista ou dos analistas, pois o modelo pode derivar para algo publicado na Web, onde cada analista poderia acessar um formulário e registrar sua análise. Propõe-se um modelo integrando diferentes técnicas, 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 1.

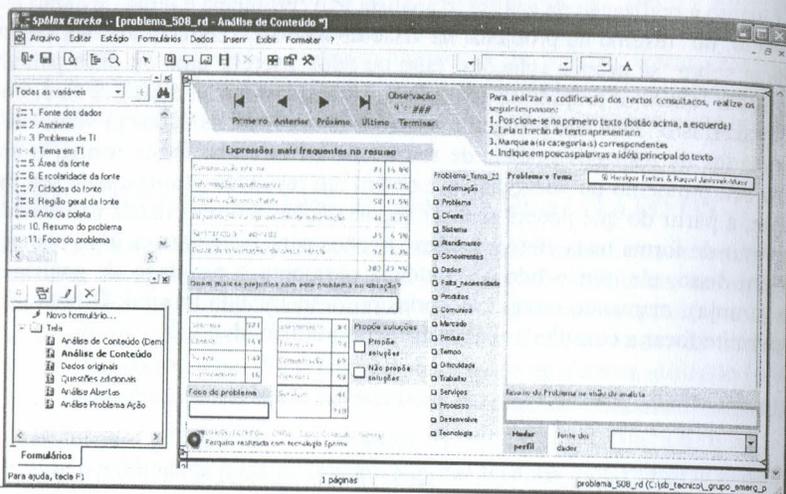
No que se refere à concepção do ambiente: notem as variáveis do estudo na coluna da esquerda, em cima. Na mesma coluna, embaixo, um guia para o analista saber que formulário está ativo. Na janela maior, a montagem da interface ou formulário que permitirá ler o 'problema e tema' (variável 14), escrever 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).

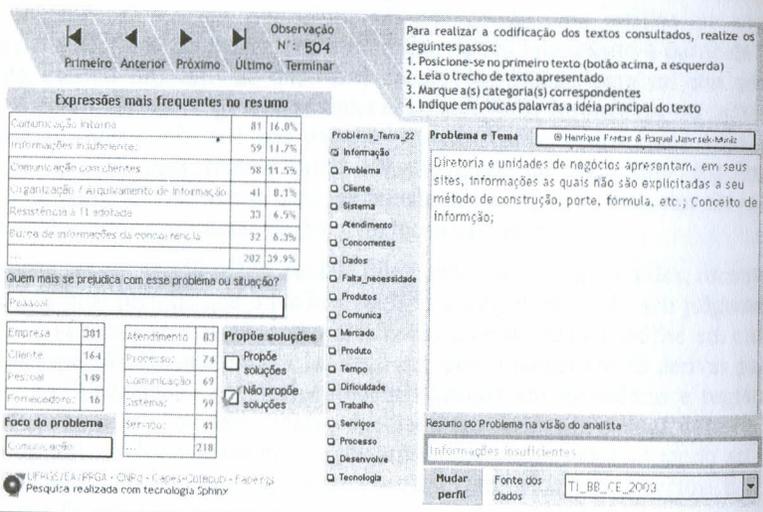
Quanto à realização da análise, o analista lê o 'problema e tema', e então escreve sua reflexão: no 'resumo do problema na visão do analista', bem como clica na questão que indaga sobre 'se propõe soluções' (sim ou não), escreve uma ou duas palavras na questão sobre 'quem se prejudica' com aquela situação 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.

Quadro 1. A análise léxica automatizada, a partir de base-textual e reflexão do analista

Concepção do ambiente¹



Realização da análise



5. REFLEXÃO FINAL

Com o avanço tecnológico da última década, é tempo de revolucionar conceitos, pressupostos, com a precaução metodológica inerente. As técnicas aqui abordadas 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

¹ O estagiário Ricardo Simm Costa, estudante da Escola de Administração da UFRGS, cooperou com o autor na realização operacional deste modelo. Esta interface é fruto de interações do autor com Jean Moscarola (Univ. de Savoie, Annecy, França), Yves Baulac (Le Sphinx Développement, Grenoble, França), e Raquel Janissek-Muniz.

técnicas e métodos para preparação dos dados abertos textuais, propondo-se um modelo integrado dessas técnicas, 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.

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 das idéias aqui abordadas.

6. REFERÊNCIAS BIBLIOGRÁFICAS

BARDIN, L. (1996) "L'analyse de contenu", PUF, 8^{ème} Edition, Paris.

CRESWELL, J. W. (1998) "Qualitative Inquiry and Research Design", Sage Publications.

FRANKFORT-NACHMIAS, C. & NACHMIAS, D. (1996) "Research Methods in The Social Sciences", St. Martin's Press, New York.

FREITAS, H., JANISSEK-MUNIZ, R. e MOSCAROLA, J. (2004) "Análise qualitativa em formulário interativo: rumo a um modelo cibernético conjugando análises léxica e de conteúdo", CIBRAPEQ- 1^a Conferência internacional do Brasil de Pesquisa Qualitativa, Poster e workshop. 10 pág. Anais em CD-ROM, Taubaté/SP.

FREITAS, H. e JANISSEK, R. (2000) "Análise léxica e Análise de Conteúdo: técnicas complementares, sequenciais e recorrentes para análise de dados qualitativos", Sphinx, Porto Alegre/RS.

FREITAS, H.; MOSCAROLA, J. (2000) "Análise de dados quantitativos e qualitativos: casos aplicados usando o Sphinx®", Sphinx, Porto Alegre/RS.

FREITAS, H. (2000) "As tendências em Sistemas de Informação com base em recentes congressos", ReAd (<http://www.adm.ufrgs.br/read>), No.13, Jan., Porto Alegre/RS.

FREITAS, H., CUNHA Jr., M. V. M. e MOSCAROLA, J. (1996) "Pelo resgate de alguns princípios da análise de conteúdo: aplicação prática qualitativa em marketing", Anais do 20º ENANPAD, ANPAD, Marketing, 23 - 25 de Setembro, Angra dos Reis/RJ.

FREITAS, H., CUNHA Jr. M.V.M. e MOSCAROLA, J. (1997) "Aplicação de sistema de software para auxílio na análise de conteúdo", RAUSP, v. 32, n° 3, Jul/Set., São Paulo/SP.

FREITAS, H., JANISSEK, R., MOSCAROLA, J. e BAULAC, Y. (2002) "Pesquisa interativa e novas tecnologias para coleta e análise de dados usando o Sphinx®", Sphinx, Porto Alegre.

- GAVARD-PERRET, M.L. & MOSCAROLA, J. (1995) "De l'énoncé à l'énonciation: relecture de l'analyse lexicale en marketing", GEREg and CERIAM, Université de Savoie, Cahier GEREg, Annecy.
- GRAWITZ, M. (1993) "Méthodologie des Sciences Sociales", Dalloz, 1976 (chap. II, pp. 586-627), and 1993 (9 ed., chap. 2, section 2, pp. 532-565), Paris.
- KELLE, U. (1995) "Computer-Aided Qualitative Data Analysis: theory, methods and practice", Sage Publications.
- KIRK, J. & MILLER, M. (1986) "Reliability and validity in qualitative research", Sage publications.
- KRIPPENDORFF, K. (1980) "Content analysis: an introduction to its methodology", The Sage CommText Series.
- LEBART, L. & SALEM, A. (1994) "Statistique Textuelle", Dunod, Paris.
- LEE, A. S. LIEBENAU, J. ; DEGROSS, J. I. (1997) "Information Systems and Qualitative Research. Philadelphia", Chapman & Hall, Proceedings ICIS, USA.
- MASON, J. (1997) "Qualitative researching" Sage.
- MILES, M. B. & HUBERMAN, A. M. (1994) "Qualitative data analysis", Sage Publications.
- MOSCAROLA, J. (1990) "Enquêtes et Analyse de données avec le Sphinx", Libraire Vuibert, Paris.
- MOSCAROLA, J. (1993) "Analyse de contenu et analyse de données: solutions logiciels pour une intégration quali/quantitative" Actes JADT, Paris.
- MOSCAROLA, J. (1994) "La communication politique vue par l'analyse lexicale", IUP, Université de Savoie, Cahier du GEREg, n° 32, Annecy.
- SILVERMAN, D. (1993) "Interpreting Qualitative Data" Sage Publications.
- WEBER, R.P. (1990) "Basic content analysis", Sage University paper.
- WEITZMAN, E. A. e MILES, M. B. (1995) "Computer programs for qualitative data analysis" Sage Publications.