

Uso do software IRaMuTeQ para tomada de decisão em priorização de problemas

Patricia M. Milhomem - Universidade de Brasília - patriciamotamilhomem@gmail.com

Ari Melo Mariano- Universidade de Brasília- arimariano@unb.br

Vitor C. F. R. Almeida - Universidade de Brasília - victordacunha@gmail.com

Gabriel N. Aprá - Universidade de Brasília - gabriel.apra@live.com

Simone B. Simão Monteiro - Universidade de Brasília - simone_simao@yahoo.com.br

Resumo: O presente artigo objetivou evidenciar a usabilidade do software IRaMuTeQ, como suporte na tomada de decisão, em relação a priorização de problemas para melhoria. Considera-se esta etapa a mais crítica em melhoria de processos, pois os problemas selecionados nesta fase, definirá as entregas, ou seja, as propostas de melhoria. Por essa razão, a seleção dos problemas deve ser realizada com impessoalidade e critérios bem definidos. Para tanto, o software IRaMuTeQ têm se apresentado como uma ferramenta de altíssima importância para o gerenciamento e apoio à análise de dados. A partir de seus relatórios, é possível agrupar conjuntos de informações que se correlacionam, e com isso classificar e identificar padrões, para projetar as tendências desejadas de acordo com o enfoque. Para obter os dados de análise desta pesquisa, foi realizada coleta de informações por meio de entrevistas semiestruturadas com os executores dos processos das seções em estudo. A partir dos dados coletados, utilizou-se o software IRaMuTeQ para subsidiar a identificação dos problemas raízes das seções, consequentemente, as oportunidades de melhoria, que possibilitaram ditar os rumos do aprimoramento e do desenvolvimento gerencial da organização militar em análise.

Palavras-chaves: Software IRaMuTeQ; Priorização de problemas, Melhoria de processos.

IRaMuTeQ software used to prioritize decision-making problems

Abstract: This article aimed to highlight the usability of IRaMuTeQ software, as a support in decision making, regarding the prioritization of problems for improvement. This step is considered the most critical in process improvement, because the problems selected in this phase will define the deliverables, ie the proposals for improvement. For this reason, problem selection should be done with impersonality and well-defined criteria. To this end, IRaMuTeQ software has been presented as a very important tool for managing and supporting data analysis. From your reports, you can group sets of correlating information, thereby classifying and identifying patterns, to project desired trends according to focus. To obtain the analysis data of this research, information was collected through semi-structured interviews with the process executors of the sections under study. From the collected data, the IRaMuTeQ software was used to support the identification of the root problems of the sections, consequently, the opportunities for improvement, which allowed us to dictate the direction of improvement and managerial development of the military organization under analysis

Key-words: IRaMuTeQ Software; Problem prioritization, Process improvement.

1. Introdução

Em melhoria de processos é recorrente a necessidade de tomada de decisão nas fases de priorização, seja em relação a processos que precisam ser analisados ou problemas que precisam ser selecionados. O CBOK (2013) acrescenta que nesta fase a análise interfuncional deve incluir uma governança que estabeleça critérios de priorização e ordenação dos processos a serem analisados. Situação semelhante ocorre na priorização de problemas, é imprescindível definir critérios e parâmetros para selecionar os problemas que serão estudados, uma vez que estes problemas nortearão todas as entregas de melhoria.

Na fase de priorização de problemas, um dos fatores que podem ser considerados são as evidências, dados e fatos para análise das causas-raiz. Contudo, a tomada de decisão pode ser um processo complexo, envolver alguma incerteza e decisões subjetivas. Desta forma, a literatura apresenta algumas ferramentas que podem auxiliar na escolha dos problemas que serão investigados, Matriz GUT, diagrama de Moody, Árvore de Realidade Atual (ARA), AHP (Analytical Hierarchy Process). Porém todas elas necessitam de dados estruturados na entrada da análise. Neste ponto, a análise via Text Mining com IRaMuTeQ (*Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*) garante o uso de Texto, dado do tipo não estruturado.

O software IRaMuTeQ pelo seu rigor estatístico, desenvolvido na linguagem Python e utiliza funcionalidades providas pelo software estatístico R. Pelas diferentes possibilidades de análise, interface simples e compreensível, e, sobretudo por seu acesso gratuito, tem sido amplamente utilizado nos estudos das ciências humanas e sociais, que têm o conteúdo simbólico proveniente dos materiais textuais como uma fonte importante de dados de pesquisa (CAMARGO & JUSTO, 2013).

No Brasil, ele começou a ser utilizado em 2013 em pesquisas de representações sociais, entretanto, outras áreas também se apropriaram do seu uso, e contribuem para a divulgação das várias possibilidades de processamento de dados qualitativos, visto que permite diferentes formas de análises estatísticas de textos, produzidas a partir de entrevistas, documentos, entre outras (CAMARGO & JUSTO, 2013).

O software IRaMuTeQ é uma ferramenta que objetiva realizar análises estatísticas sobre corpus textuais e sobre tabelas indivíduos/palavras coletada nas reuniões de levantamento de problemas com os gestores e executores das seções. A análise textual é um tipo específico de análise de dados, na qual tratamos de material verbal transcrito, ou seja, de textos (NASCIMENTO-SCHULZE & CAMARGO, 2000). Essa análise tem várias finalidades, sendo possível analisar textos, entrevistas, documentos, redações etc.

Portanto, diante do desafio de priorizar problemas em macroprocessos de melhoria, este artigo tem como objetivo descrever o uso do IRaMuTeQ na análise de corpus textuais de várias entrevistas com executores dos processos de duas Seções (A e B) de uma organização militar para selecionar as oportunidades de melhoria. As análises possibilitadas pelo software IRaMuTeQ, sendo elas Lexográfica, Especificidades, Análise Fatorial de Correspondência (AFC), Método da Classificação Hierárquica Descendente (CHD), Análise de Similitude e Nuvem de Palavras. Subsidiaram a tomada de decisão, uma vez que centralizaram as informações mais importantes relatadas pelos executores.

2. Metodologia

A pesquisa refere-se a uma análise qualitativa, a qual utilizou o software IRaMuTeQ, como ferramenta de apoio na tomada de decisão para priorização de problemas de duas seções, denominadas Seção A e Seção B de uma determinada organização militar.

A coleta dos dados foi realizada por meio de entrevistas, nas quais foi aplicado um roteiro semiestruturado, com participação de 3 executores de cada seção. A entrevista foi realizada individualmente e teve duração média de 50 minutos. A transcrição da fala foi realizada em paralelo com as entrevistas.

Após realizar todas as entrevistas, o próximo passo foi preparar o corpus textuais de acordo com o estabelecido para análise no software IRaMuTeQ, para tal foram seguidos os procedimentos listados abaixo:

- a) Colocar o conteúdo a ser analisado em um único arquivo de texto sem parágrafos, no caso desta pesquisa, foram realizadas entrevistas com diversos atores. Por isso, as perguntas e as intervenções realizadas pelo pesquisador foram suprimidas para não entrar na análise.
- b) Preparar os textos para serem analisados no software. Essa preparação requer retirar acentos, uniformizar as siglas, os verbos devem ser colocados na forma de próclise, pois o dicionário não prevê as flexões verbo-pronominais. Ex No lugar de “tornei-me”, a escrita deve ser: “me tornei”.
- c) Separar os textos com linhas de comando (com asteriscos). Como exemplificado abaixo
*** *Entrevista_1 *S_Sec_A / **** *Entrevista_1 *S_Sec_A
- d) Revisar todo o arquivo, para que os erros de digitação ou outros não sejam tratados como palavras diferentes. Observar se o texto está justificado, se tem palavras em negrito, itálico, com: aspas ("), apóstrofo ('), hífen (-), cifrão (\$), percentagem (%) e nem asterisco (*). Este último é usado somente nas linhas que antecedem cada texto (linhas de comando).
- e) Inserir o texto preparado no software IRaMuTeQ, processar e analisar os resultados. Para a Seção A foram analisadas no total 1930 palavras e para a Seção B 1934.

3. Resultados e Discussões

Os resultados obtidos na análise textual possibilitaram realizar análises do tipo Lexográfica, Especificidades, Análise Fatorial de Correspondência (AFC), Método da Classificação Hierárquica Descendente (CHD), Análise de Similitude e Nuvem de Palavras como visto mais adiante.

3.1 Análise Classificação Hierárquica de Descendente (CHD)

Os segmentos de texto são classificados em função dos seus respectivos vocabulários, e o conjunto deles é repartido em função da frequência das formas reduzidas. A partir de matrizes cruzando segmentos de textos e palavras, aplica-se o método de CHD e obtém-se uma classificação estável e definitiva (REINERT, 1990). Esta análise visa obter classes de segmentos

de texto que, ao mesmo tempo, apresentam vocabulário semelhante entre si, e vocabulário diferente dos segmentos de texto das outras classes (CAMARGO, 2005). A partir dessas análises em matrizes o software organiza a análise dos dados em um dendograma da CHD, que ilustra as relações entre as classes. A Figura 1 mostra o dendograma da CHD para a Seção A.

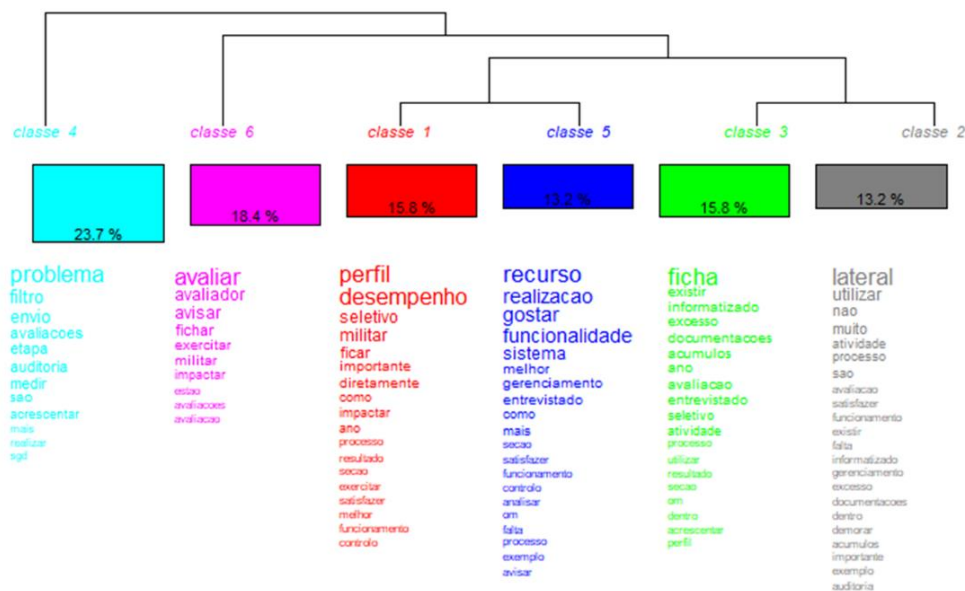


Figura 1 - Dendograma da CHD para a Seção A.

O dendograma da Seção A foi dividido em dois corpus e dois subcorpus considerados suficientes/estáveis para análise do vocabulário semelhante, cada uma com sua respectiva cor. No primeiro, obteve-se as classes 4 e 6 que corresponder a 42,15% do total. O corpus da classe 6 se subdividiu em dois subcorpus. O primeiro subcorpus corresponde as classes 1 e 5 com 29% do total e o segundo referentes as classes 3 e 2 também com 29% do total. Desta forma, percebe-se que as palavras que obtiveram maior porcentagem quanto à frequência média foram os referentes a classe 6 e 4. A leitura das palavras que aparecem nessas classes e de suas inserções nos segmentos os textos, foi possível compreender que o assunto em questão se refere as oportunidades de melhoria da avaliação de desempenho da organização, sejam em relação aos filtros de análise e as fichas utilizadas. A Figura 2 mostra o dendograma da Seção B.

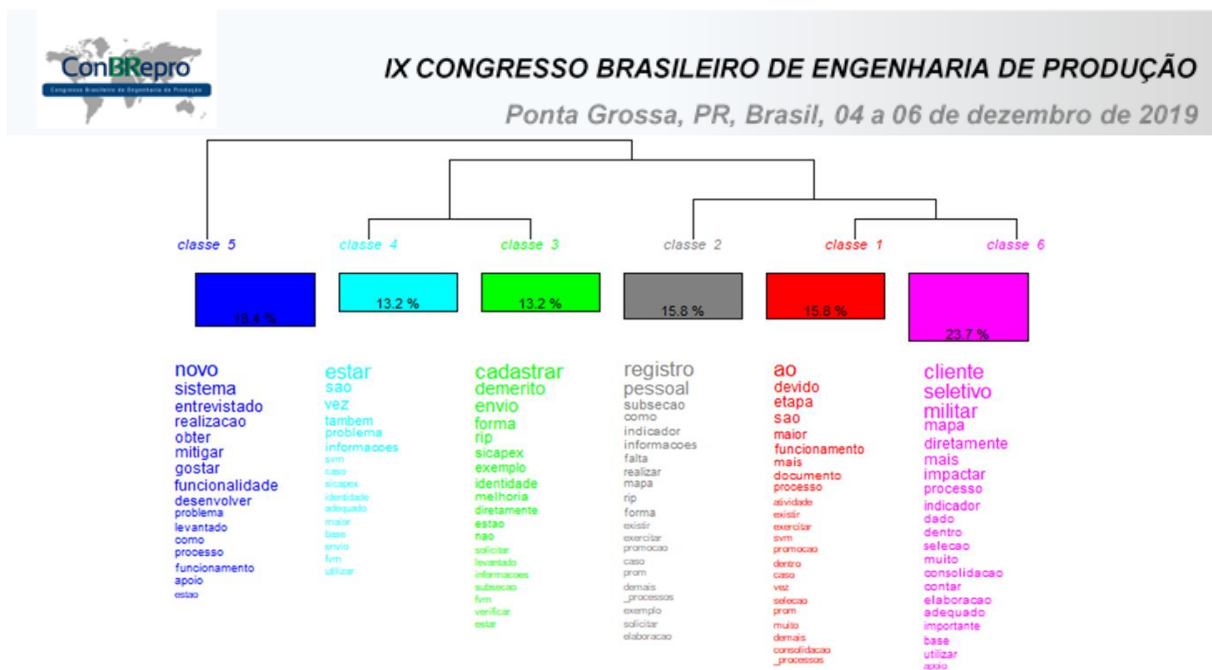


Figura 2 - Dendrograma da CHD para a Seção B.

Para a Seção B a divisão do dendrograma foi diferente da Seção A, pois contempla um corpus com 18,4% e o subcorpus 4 e 3 com 26,4% e os três subcorpus 2, 1 e 6 juntos possuem 55,3% da frequência total. O corpus enfatiza a importância de um novo sistema, afirmado *in loco* e os subcorpus que englobam a classe 2, 1 e 6 dizem a respeito do gerenciamento das informações e da importância destes para os clientes, por isso, necessitam ser melhorados, principalmente, com a implantação de um novo sistema.

3.2 Análise de Similitude

A Análise de Similitude baseia-se na teoria dos grafos e possibilita identificar as coocorrências entre as palavras e seu resultado traz indicações da conexidade entre as palavras, auxiliando na identificação da estrutura da representação (MARCHAND e RATINAUD, 2012). A Figura 3 mostra a análise de similitude da Seção A.

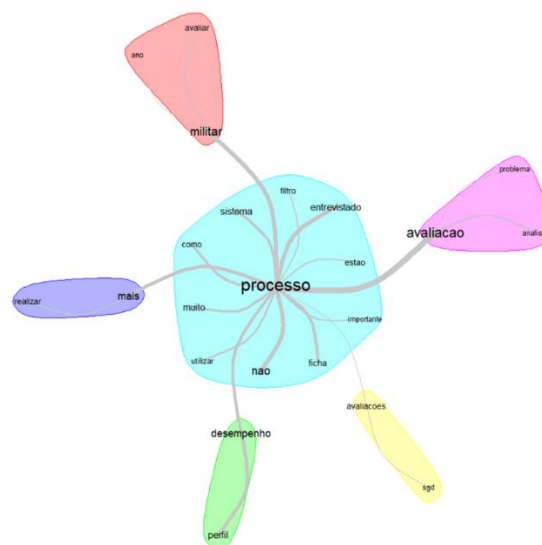


Figura 3 - Análise de similitude Seção A

Pode-se identificar na estrutura, que o núcleo central se refere a processos que está diretamente relacionado com o processo de avaliação gerenciado por esta seção, como ratificado pelo sistema periférico da análise de similitude, no qual, contém as seguintes palavras: avaliação, desempenho, perfil, militar. Destacando que esses são os processos mais importantes da seção, necessitando de maior atenção, pois uma proposta de melhoria neste impactará diretamente na avaliação de desempenho. Para a Seção B temos a análise de similitude representada na Figura 4 abaixo.

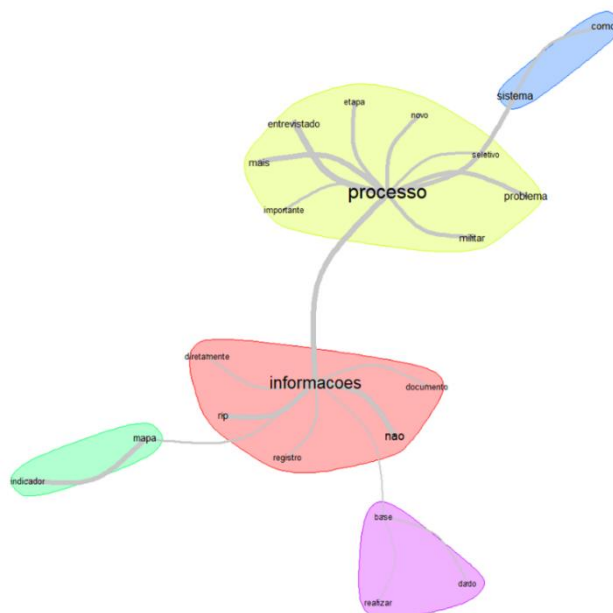


Figura 4 - Análise de similitude Seção B

Ao contrário do observado na Seção A para a B temos dois núcleos principais, um referente ao processo e outra as informações, o que está coerente com o identificado *in loco*. Uma vez que as informações são correspondes aos documentos elaborados por essa seção, como o documento de registro de informações e um mapa com indicadores de qualidade mencionados nas extremidades.

Em relação ao núcleo de processos, identifica uma derivação com os sistemas, bem relatado nas reuniões devido a necessidade de um novo sistema e melhor integração com outros sistemas da diretoria para os existentes.

3.3 Análise Fatorial de Correspondência

A Análise Fatorial retoma as frequências e os valores de correlação Qui2 de cada palavra do Corpus e representa num plano cartesiano as diferentes palavras e variáveis associadas a cada uma das classes do dendograma da CHD (MENDES *et al*, 2015). A Figura 5 apresenta o plano cartesiano de palavras da Seção A.

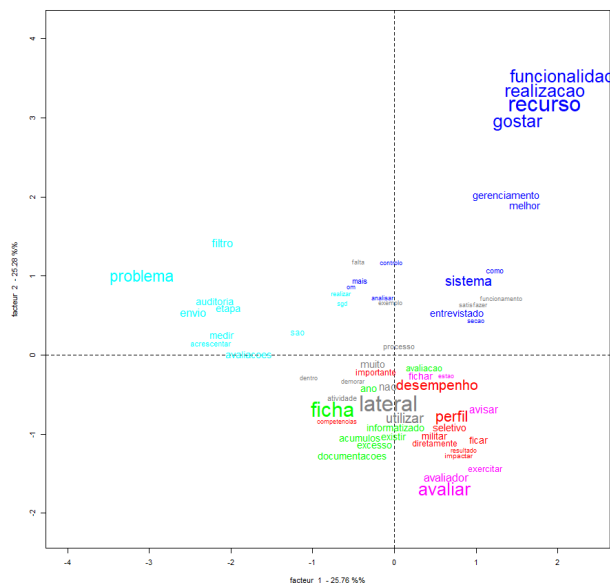


Figura 5 - AFC Seção A

Verifica-se que na classe 1 (primeiro quadrante da direita) a palavra central é AVALIAR, seguida de PERFIL e DESEMPENHO. A classificação é coerente, pois estão todas no mesmo contexto, uma vez que o perfil de desempenho é o produto da avaliação de desempenho, utilizado como auxílio para selecionar o militar em processos seletivos.

Entre a classe 1 e 4 temos LATERAL e UTILIZAR, as duas palavras aparecem simultaneamente devido a melhor necessidade de utilização deste tipo de avaliação, mencionado pelos executores. Na classe 2 temos RECURSO, REALIZAÇÃO, FUNCIONALIDADE e GOSTAR. Na classe 3, um pouco mais distante temos a palavra PROBLEMAS em evidência e na classe 4 temos FICHA em realce, essas últimas com uma maior dificuldade para definir o contexto. Análise semelhante foi realizada para a Seção B e pode ser observada na Figura 6.

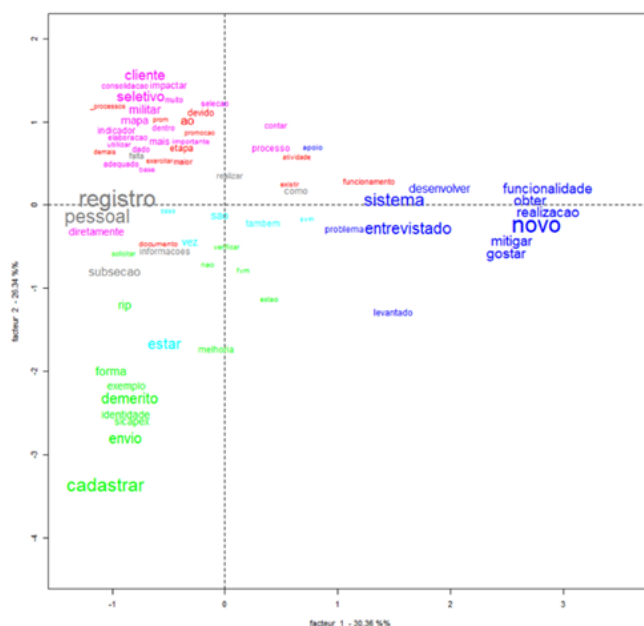


Figura 6 - AFC Seção B

Para a Seção B temos na classe 1 palavras como NOVO, ENTREVISTADO, REALIZAÇÃO. Na classe 2 temos SISTEMAS, DESENVOLVER, FUNCIONALIDADES, OBTER diretamente condizente com as necessidades da seção, necessidade de mais sistemas, ou mais funcionalidades nos existentes. Na classe 3 temos CLIENTE, SELEÇÃO, MAPA, o que está direcionado ao documento de Indicadores, elaborado diretamente para as outras seções responsáveis pela seleção do militar, por exemplo, cursos e movimentação. E, por fim, temos na classe 4: REGISTRO, PESSOAL, CADASTRAR, ENVIO, DEMÉRITO todas essas informações são registradas em documento apropriado e utilizadas para tomada de decisão.

3.4 Nuvens de Palavras

Para auxiliar e finalizar a tomada de decisão quantos aos problemas que serão investigados, foram elaboradas nuvens de palavras com o texto das entrevistas da Seção A e B. O diagrama de nuvem de palavras, assim também chamado, destaca as palavras com maior frequência no texto, no qual o tamanho da fonte de cada uma das palavras é proporcional ao número de vezes que ela foi repetida. A Figura 7 apresenta a nuvem de palavras da Seção A.



Figura 7 - Nuvem de palavras Seção A

As palavras relevam características próprias de cada entrevista, permitindo conectar os estudos e classificá-los (MARIANO *et al.*, 2011). Como resultado, destacaram-se as palavras: PROCESSO, MILITAR, DESEMPENHO, AVALIAÇÃO, SISTEMA. Tais palavras enfatizam o observado nas análises anteriores, ou seja, para a Seção A, temos o processo da avaliação de desempenho necessitando de maiores cuidados. A Figura 8 mostra a nuvem de palavras da Seção B.



Figura 8 - Nuvem de palavras Seção B

Para a Seção B, nota-se que as palavras: PROCESSO, INFORMAÇÕES e SISTEMAS são as mais evidentes. Indicando que a oportunidade de melhoria se refere a falta de sistema e para os que existem a integração, corroborando o que foi percebido *in loco*.

Portanto, a análise de todos os gráficos acima facilitou a identificação das oportunidades de melhoria, possibilitando priorizar os problemas mais significativos, problemas raiz, ou seja, aqueles que desencadeiam e mais impactam em outros problemas para cada uma das seções. Para a Seção A, percebeu-se que o problema raiz refere-se a metodologia da avaliação de desempenho e para a Seção B a problemática está mais relativa a um novo sistema e uma melhor integração dos existentes com outros da diretoria, o que facilitaria o gerenciamento/cadastro das informações e, conseqüentemente, uma entrega mais rápida e completa para o cliente.

4. Conclusão

A implementação do software IRaMuTeQ na fase de priorização de problemas possibilitou a identificação dos problemas-raízes e evidenciou a aplicabilidade das análises de corpus textuais para subsidiar essa tomada de decisão. Conseqüentemente, identificar as oportunidades de melhoria, que possibilitaram ditar os rumos do aprimoramento e do desenvolvimento gerencial da organização em análise.

Para a Seção A em análise identificou-se que a equipe de melhoria poderia voltar suas atenções para a metodologia de avaliação de desempenho oportunizando na resolução de outros pontos viáveis de melhoria identificados. Para a Seção B, enfatizou-se a necessidade de integração dos sistemas e até mesmo o desenvolvimento de outros, observado *in loco*.

Portanto, a aplicação do software IRaMuTeQ possibilitou realizar vários tipos de análises, dentre elas as do tipo Lexográfica, Especificidades, Análise Fatorial de Correspondência (AFC), Método da Classificação Hierárquica Descendente (CHD), Análise de Similitude e Nuvem de Palavras. O que permitiu uma visão mais clara dos problemas raízes, ratificando a utilização dessa ferramenta para essa funcionalidade.

Referências

ABPMP BPM CBOK. **Guia para o Gerenciamento de Processos de Negócios Corpo Comum de Conhecimento ABPMP BPM CBOK V3.0.** ABPMP International, 2013.

CAMARGO BV, JUSTO AM. **Tutorial para uso do software de análise textual IRAMUTEQ.** Disponível em: <<http://www.iramuteq.org/documentation/fichiers/tutoriel-en-portugais>> Acesso em: 10 set. 2019.

CAMARGO BV, JUSTO AM. **IRAMUTEQ: um software gratuito para análise de dados textuais.** Disponível em: < <http://pepsic.bvsalud.org/pdf/tp/v21n2/v21n2a16.pdf> > Acesso em: 10 set. 2019.

NASCIMENTO-SCHULZE, C. M.; CAMARGO, B. V. **Psicologia social, representações sociais e métodos.** Temas psicol. [online]. 2000, vol.8, n.3, pp. 287-299. ISSN 1413-389X.

REINERT M. **Alceste une méthodologie d'analyse des données textuel les e tune application: Aurelia de Gerard de Nerval.** Bull Methodol Sociol.1990;26(1):24-54.

CAMARGO, B. V. **ALCESTE: Um programa informático de análise quantitativa de dados textuais.** Editora da Universidade Federal da Paraíba, 2005.

RATINAUD, P., & MARCHAND, P. **Application de la méthode ALCESTE à de "gros" corpus et stabilité des "mondes lexicaux" : analyse du "CableGate" avec IraMuTeQ.** Em: Actes des 11eme Journées internationales d'Analyse statistique des Données Textuelles (835–844). Presented at the 11eme Journées internationales d'Analyse statistique des Données Textuelles. JADT 2012, Liège.

MENDES FRP, ZANGAO MOB, GEMITO MLGP, SERRA ICC. **Social Representations of nursing students about hospital assistance and primary health care.** Rev Bras Enferm [Internet]. 2016;69(2):321-8. DOI: <http://dx.doi.org/10.1590/0034-7167.2016690218i>

MARIANO, A. M.; CRUZ, R. G.; GAITÁN, J. A. **Meta Análises como instrumento de pesquisa: uma revisão sistemática da bibliografia aplicada ao estudo das alianças estratégicas internacionais.** CONGRESSO INTERNACIONAL DE ADMINISTRAÇÃO - GESTÃO ESTRATÉGICA: INOVAÇÃO COLABORATIVA E COMPETITIVIDADE. Ponta Grossa, UEPG, 2011.