

Sistema de Informações de Milícias: Proposta de um Arcabouço Analítico para Análise de Textos em Dados do Disque Denúncia

Walkir A.T Brito¹, Claudia L.R. Motta², Paulo R.A.S Filho³, Zeca Borges³

¹ Programa de Pós-Graduação em Informática – PPGI/UFRJ Universidade Federal do Rio de Janeiro – Rio de Janeiro, RJ – Brazil

² Programa de Pós-Graduação em Informática – PPGI, Instituto Tércio Pacitti de Aplicações e Pesquisa Computacionais – NCE/ UFRJ – Universidade Federal do Rio de Janeiro (UFRJ)
Caixa Postal 15.064 – 91.501-970 – Rio de Janeiro – RJ – Brazil

³ Coordenação e Departamento de Computação Disque Denuncia (DD) – Rio de Janeiro

***Abstract.** This article presents an analytical framework for the analysis of texts in complaints in the Hotline database (Disque Denuncia), providing an information system called SIM - Militia Information System. Through this, information was sought and it was possible to make discoveries about personas and entities involved with the Militias in the State of Rio de Janeiro through the case study applied in the period from 2012 to 2019 in complaints.*

***Resumo.** Este artigo apresenta um arcabouço analítico para análise de textos em denúncias na base de dados do Disque Denúncia, fornecendo um sistema de informação denominado SIM – Sistema de Informações de Milícias. Por meio deste, buscamos informações e fizemos descobertas sobre personas e entidades envolvidas com as Milícias no Estado do Rio de Janeiro por meio do estudo de caso aplicado no período de 2012 a 2019 em denúncias.*

1. Introdução

O tema da violência é inegavelmente cada vez mais importante em diferentes contextos, tornando-se um dos grandes desafios para a sociedade e para os órgãos de segurança pública na contemporaneidade, no Brasil e no mundo. Órgãos de inteligência em segurança têm buscado encontrar soluções que utilizem as ferramentas de informação e comunicação para solucionar o problema da grande quantidade de dados. Deste modo, o objetivo desta pesquisa é apresentar a proposta de um arcabouço analítico para análise de textos em denúncias, na base de dados do Disque Denuncia (DD), apresentando o sistema de informação denominado SIM – Sistema de Informações de Milícias. Este espaço analítico é uma cooperação entre a ciência de dados e as ferramentas de coleta, limpeza, gestão e monitoramento de conhecimento. A partir do SIM, é possível buscar informações e fazer descobertas a respeito de personas e entidades envolvidas com as Milícias no Estado do Rio de Janeiro. Este trabalho subsidia estudos para as áreas de inteligência e seus órgãos de segurança através de sua implementação e utilização. Utilizamos uma pesquisa aplicada, com o estudo de caso, com a temática Milícia, no

período compreendido entre 2006 e 2019 em denúncias recebidas pelo DD. Elegemos este corte temporal por ter ocorrido um grande crescimento de áreas sob domínio desses grupos paramilitares no estado, tornando-os hoje um dos maiores desafios enfrentados pelos órgãos de segurança no Rio de Janeiro. Implementamos o SIM para testar as possibilidades do emprego e utilização desta ferramenta na área de inteligência. E a partir de sua implementação, foi possível avaliar os principais resultados obtidos nessa primeira fase. É importante destacar que o SIM pode ser utilizado para diversas análises e temas no banco de dados do DD e não apenas aquelas a que se refere este artigo.

Este trabalho é um recorte da proposta de estudo de doutorado em informática em sistemas complexos, sob a temática da utilização de tecnologias em bases de dados de segurança pública, gestão de conhecimento, mineração de dados textuais, análise de texto e informação. Esse estudo trata da fase de implementação da metodologia própria desenvolvida, bem como a utilização e teste do SIM no estudo de caso Milícia, possibilitando o teste prático do modelo proposto na tese em questão. O emprego de tecnologias, suas aplicações em bases de dados de denúncias e ações que colaborem na perspectiva de redução da criminalidade têm sido um grande desafio enfrentado pelas áreas de inteligência e segurança. Por isso, colaborações deste gênero trazem bem-estar para a sociedade.

Este artigo divide-se quatro seções. A primeira seção apresenta os trabalhos relacionados. A segunda seção apresenta o Disque Denúncia, suas características, premissas e canais de atendimento. A terceira seção apresenta a proposta do modelo do sistema de informação aplicada neste estudo. A quarta seção parte os resultados da aplicação e implementação do SIM, suas considerações finais e trabalhos futuros.

2. Trabalhos Relacionados

Podemos citar como exemplo existente o sistema utilizado pela Secretaria de Estado da Segurança Pública de Santa Catarina, que adquiriu um produto de tratamento de Big Data para apoio à inteligência chamado “Guardião”. Tal sistema é parte de uma plataforma que aplica vários conceitos com a capacidade de tratar de maneira eficiente um número muito maior de dados, onde agentes de segurança podem responder de forma mais efetiva à resolução e investigação de denúncias e crimes e, principalmente, também os prevenir e (ou) antecipá-los.

Encontramos a aplicação de tecnologias utilizando Business Intelligence (BI), para modelos preditivos em criminalidades ou monitoramento de ocorrências policiais [Chaudhuri 2011]. Há trabalhos ligados a terrorismo e organizações criminosas em rede, principalmente com dados não estruturados, utilizando redes sociais empregando processamento de linguagem natural e Data Mining em Twitter [Chen 2015], permitindo previsão de futuros crimes [Chen 2004]. No levantamento da literatura de trabalhos relacionados, percebemos que são utilizadas algumas técnicas de Machine Learning [Wang 2006]. O uso de ontologias semânticas na Web em domínio para extrair automaticamente indicadores de radicalização de mensagens e postagens em redes sociais, utilizadas em domínio da radicalização em mensagens, foram encontrados e pesquisados [Barhamgi 2018]. Há estudos que englobam a detecção automática de

recrutamento de pessoas a grupos extremistas como o estado islâmico em redes sociais que determinem se um usuário pertence a estes grupos ou se está em processo de recrutamento, a partir das informações relacionadas a nomes de usuários, seu tipo de perfil, redes de amigos e conteúdo textual desses usuários [Alvari 2019].

Não encontramos estudos em língua portuguesa sobre análise de texto empregando linguagem natural, aplicados ou comparados, gerando conhecimento sobre o objeto do estudo. Tampouco foram encontradas outras referências de trabalhos com sistemas de denúncias na literatura, com esse foco de transformar os dados em informações, que devem ser usadas pelos analistas de inteligência na composição de um estudo de caso como em um único sistema de informação. Foram encontradas apenas pesquisas que auxiliam análises criminais, modelos preditivos de rotas de patrulhamento, monitoramento de áreas policiais, maximização de efetivo de patrulhamento nessa área. A relevância e necessidade deste trabalho trata da gravidade do tema milícia tão presente em nossas vidas e em nossa sociedade, e da carência em inovação tecnológica para a área de inteligência, de forma a contribuir a melhores práticas e resultados no combate à criminalidade, através de utilização do sistema de informação e da aplicação de suas análises.

Destaca-se ainda que as denúncias recebidas pelo DD em seus canais possuem características de informalidade na escrita e grafias das palavras. São empregadas gírias, expressões, termos, e semânticas específicas e complexas de guetos, com representatividades e simbologias próprias e inerentes aos seus grupos e representações que podem simbolizar poder, fúria, medo, indignação, abandono, e outros sentimentos empregados nas denúncias e que não podem ser ignorados na gestão do conhecimento dessa temática. O conhecimento e a gestão do conhecimento são conceitos complexos e multifacetados, assim como a violência e todos os atores envolvidos. A combinação de análise de texto em denúncias e diversas técnicas de processamento de linguagem natural - PLN [PÁSSARO 2009], medidas de similaridade, mineração de texto, limpeza e tratamento de dados são necessidades que acompanham este estudo, permitindo uma melhor e maior disponibilidade desses recursos para a realização dessas tarefas. A aplicação de técnicas de Inteligência Artificial, mais especificamente de mineração de textos, nas etapas de classificação, extração, difusão e análise das denúncias recebidas são necessárias para resolver os problemas que dificultam o bom funcionamento do serviço, assim como possibilitam um maior potencial analítico com os dados e seu estudo sobre o tema.

3. O Disque Denúncia

O Rio de Janeiro conta com o Disque Denúncia¹, uma central de atendimentos criada para ajudar as polícias Federal, Civil e Militar no esclarecimento de crimes e delitos, mediante o recebimento de denúncias feitas por meio de ligações anônimas da população. Este instrumento tem sido caracterizado, ao longo dos seus 25 anos de funcionamento, por receber as mais variadas demandas que, apesar de não serem apenas criminais, guardam relação direta com os sentimentos de temor, insegurança e medo da população carioca ao longo dos anos.

Concebido como uma alternativa para contribuir para a diminuição da violência na cidade, o DD começou a funcionar em 1995, tendo como objetivo receber

1 Disponível em <<http://www..disquedenuncia.org.br/>>. Acesso em 22 abr 2020

informações da população sobre atividades criminosas e casos de violação aos direitos humanos para transmiti-las aos órgãos de segurança pública. O serviço, que tem como princípio mais importante a garantia do anonimato daquele que fornece as informações, atende desde reclamações sobre a conduta de um vizinho até dados completos sobre quadrilhas e crimes ligados ao tráfico de drogas e de armas. Recebe também chamadas sobre atividades criminosas, sobre o modo de agir de determinadas facções, dentre outras. Assim, no estado do Rio de Janeiro, o DD atua como um canal entre a população e os agentes de segurança pública no combate à criminalidade e na solução de problemas que atingem comunidades, bairros e cidades. Ao receber denúncias anônimas e coletar informações de importância fundamental na apuração de ações criminosas e delitos, contribui consequentemente para a redução da criminalidade. Recebendo mais de 100 mil relatos anônimos de crimes e/ou delitos por ano, este é um programa bastante semelhante ao *American Crime Stoppers*², por exemplo. A informação desses relatórios de crimes é, pois, extremamente importante para reduzir os níveis de criminalidade, ajudando a polícia e a inteligência, assim como outras autoridades, a realizar operações mais eficientes contra criminosos e servindo como um termômetro de atenção com o que está ocorrendo na cidade.

Possui como pilar do Serviço do DD a trilogia da relação entre os agentes de segurança pública, a população e a Imprensa (Figura 1), empregando o uso inteligente da informação, garantindo o anonimato e a difusão, auxiliando o trabalho das autoridades policiais para a diminuição dos índices de violência urbana. Através de seus canais de recebimento das denúncias com a participação da população, o serviço é o elemento de operação entre os órgãos de segurança, através da integração dos dados das denúncias difundidas, e estabelecendo pauta de notícias para a imprensa na resolução dessas denúncias e crimes.

Atualmente, o DD recebe denúncias através de diferentes canais de comunicação com a população, respeitando suas preferências e estabelecendo a mesma metodologia de atendimento em todos os canais. São eles: ligações telefônicas, *facebook*, *whatsapp*, aplicativos e *chats*. Desta forma, foi determinante implementar ferramentas que melhorassem a gestão dos dados coletados bem como uma plataforma de canais de comunicação múltiplos canais nesse estudo. A seguir, apresentaremos o modelo proposto para este estudo de caso.

3.1 Modelo Proposto ao Estudo de Caso

O Sistema de Informações de Milícias é capaz de preparar, combinar, enriquecer e analisar dados de forma "*self-service*". O *self-service analytics* consiste em uma forma de análise de autoatendimento que possui uma abordagem para analisar dados mais avançada. O modelo permite que os profissionais que o utilizam possam manipular dados com maior facilidade na busca de oportunidades e estudos, descobrindo os elementos importantes para seus temas (*insights*). Por ser prático e visual, o profissional não precisa ter conhecimentos avançados em estatística ou tecnologias, basta saber e entender pelo tema ou o que deseja saber sobre os dados. Trata-se de uma plataforma que permite que você descubra, prepare e analise todos os dados coletados, além de facilitar a implantação e compartilhamento dessas análises em grande escala, gerando *insights* e estudos mais detalhados.

2 Disponível em: <http://www.crimestoppersusa.org/>>. Acesso em 22 abr 2020

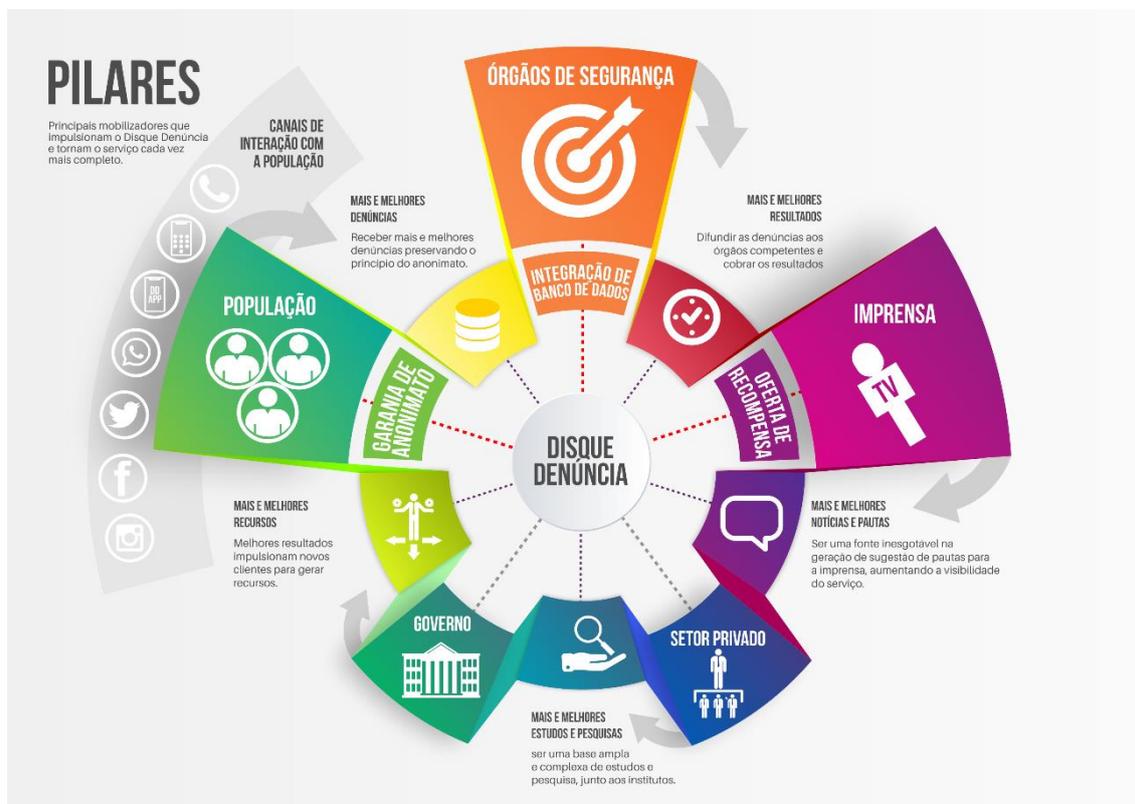


Figura 1: Pilares do Serviço do Disque Denúncia

Inicialmente, foi possível identificar problemas e dificuldades relacionados as bases de dados utilizadas neste estudo, tais como: bases disjuntas e dados heterogêneos, impossibilidade de guardar os históricos sequencialmente para serem empregados na análise de texto, os procedimentos de tratamento e limpeza, padronização dos dados, manter a estrutura e sigilo dos dados, os cruzamentos de informações para montagem dos filtros e *dataset* necessários na composição analítica, a criação dos *dashboard* de monitoramento que serão mostrados nos resultados obtidos, a limpeza e geocodificação das denúncias e dos textos referentes aos dados contidos nas denúncias, a integridade e garantia do anonimato, a acessibilidade e escalabilidade dos dados permitindo acesso e consulta de qualquer lugar, as atribuições de senhas e permissões, por tratar-se de assuntos confidenciais, a salvaguarda dos dados e informações.

A descrição do Modelo Proposto neste estudo (Figura 2) compõe o que denominamos SIM – Sistema de Informações de Milícias. Nele está contido desde da coleta dos dados, limpeza, otimizações, cargas até as análises compreendidas no arcabouço analítico de textos em denúncias. O SIM é um ambiente seguro para coleta e análise de dados utilizadas para a área de inteligência com a temática Milícia, e executado a partir da implementação dos seus 4 módulos principais que são: (1) Módulo de Identificação; (2) Módulo de Interpolação; (3) Módulo de Agrupamento (*clustering*) e (4) Módulo de Análise.

O sistema atua em um formato espiral de retroalimentação dos seus módulos, de maneira contínua a cada carga e atualiza as análises estabelecidas no sistema através dos *dashboard* existentes, atualmente sendo atualizado a cada 30 minutos, diretamente na

base de dados. A metodologia proposta para este trabalho envolveu o ciclo de tratamento dos históricos de denúncias, no período de 2006 a 2019, contemplando para o estudo de caso, um total de 51.333 denúncias no período. Correlacionando a Tag: #MILÍCIA e suas formações desde a extração de seus radicais temáticos, até a construção de palavras-chave e dicionários especializados para apoio à extração de entidades contidas nas denúncias, viabilizando o desenvolvimento de possíveis cenários que compõem os crimes narrados em formato de uma matriz de relacionamentos.

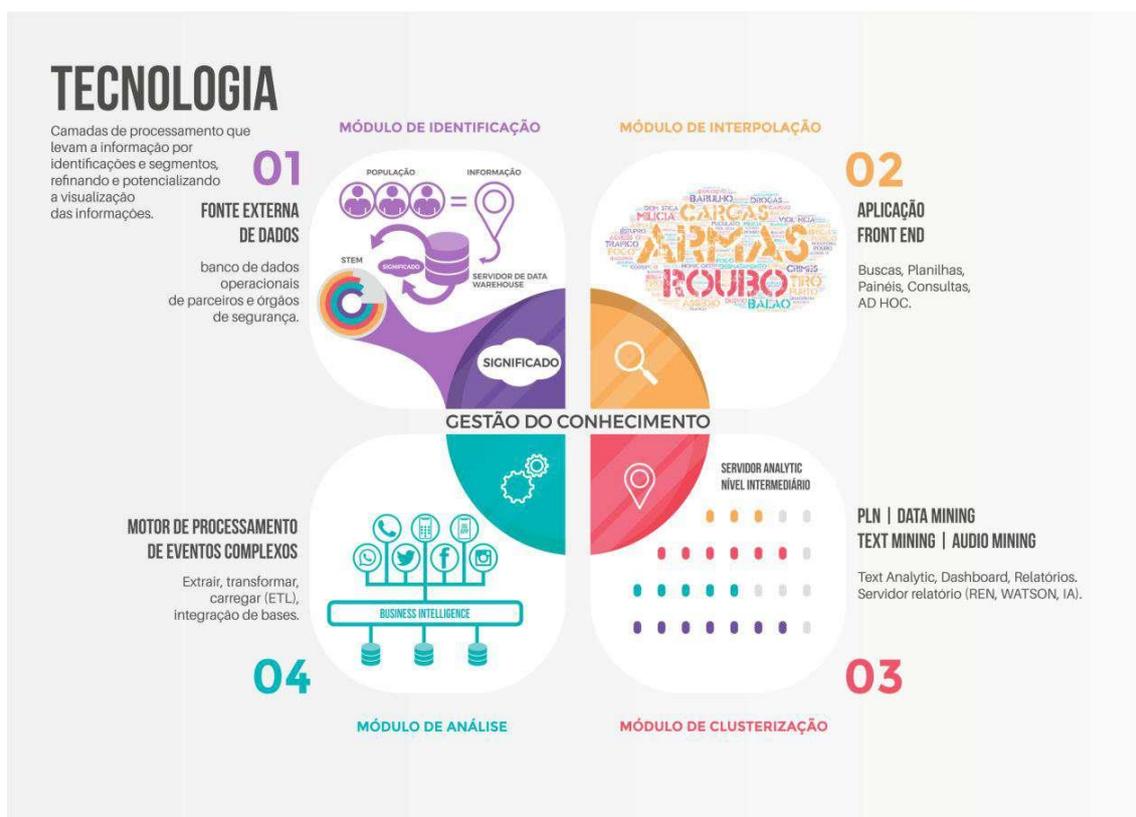


Figura 2 Modelo Proposto SIM – Sistema de Informação de Milícias

1. Módulo de Identificação: Através do texto fonte será produzido uma lista de palavras-chave (*stems*), aplica-se um algoritmo de *stemming* que transformará cada palavra no seu respectivo stem ou forma raiz. Algoritmos de *Stemming* [Hull 1996], ou stemmers tem o objetivo de reduzir uma palavra à sua forma raiz. As palavras-chave de uma consulta ou de um documento são mais bem representadas por *stems* do que pelas palavras originais. Com isso, reduzimos o tamanho da base e destacamos os assuntos, temas e termos importantes, e definimos o número de termos distintos necessários para representar uma lista de documentos representativa. Gera-se uma base de específica para cada estudo e termos específicos [Loeck 2016]. Neste trabalho, escolhemos as palavras envolvendo o contexto milícias, milícia, miliciano, venda, confronto, operação, tiroteio, disparo de armas, violência, entre outras (Figura 5 e 6). E o mesmo processo pode ser aplicado a qualquer outro tema, palavra, termo, semântica, entidade em estudos futuros dentro da base de dados do DD.

2. Módulo de Interpolação: A partir da lista de *stems* de cada denúncia busca-

se a lista de significados mapeados com suas respectivas estatísticas de frequência do banco, i.e., o número de ocorrência de palavras relacionadas a cada categoria. E que foi utilizada neste estudo para o resultado das nuvens de palavras e a criação do dicionário específico por categoria utilizada no DD.

3. Módulo de Agrupamento (*Clustering*): As denúncias são clusterizadas a partir dos significados extraídos de cada uma delas, podendo ser analisadas por *stems*, termo, tema ou expressões. Primeiramente agrupam-se as denúncias similares e adiciona-se informações estatísticas aos temas classificados. Cria-se um sumário executivo de acordo com a ordem de *stems* ou expressões relacionadas a cada categoria para posterior validação. Podemos utilizar diversas técnicas de processamento e análise de textos desse módulo, tais como *text mining*, *data mining*, nuvem de palavras, PLN entre outros.

4. Módulo Análise: A partir dos tópicos classificados e clusterização, utiliza-se a base de conhecimento por categoria que contém uma lista de conhecimento com medidas que servirá para o direcionamento dos tópicos extraídos e similaridades (ex.: Polícia Mineira, Comando Vermelho [cv], 158 [extorsão], chocolate [maconha], etc.) e para a realização de inferências do usuário ou entidades. Todas as etapas contribuem para a criação do dicionário especialista utilizado ao estudo.

Utilizamos nesse contexto a visão de ciência e análise de dados, tornando possível a qualquer analista de inteligência em segurança pública, ou qualquer profissional de dados treinado a realizar tarefas de análise de dados avançadas no ambiente analítico. Este sistema entrega o tratamento de análises mais profundas e complexas, seleções de temas e classificações em amostras ou *data set* específicos, armazenamento e tratamento desses dados, desde a coleta de informações até sua análise, movimentando-se pelo tratamento, armazenamento e filtros em um único ambiente, disponível ao analista da área de inteligência algo não encontrado na literatura, um único ambiente seguro do processo de coleta, tratamento e análise aplicado em denúncias.

Para isso, foi necessário no módulo de identificação, todo o preparo para as cargas das bases, bem como as execuções dos algoritmos da lista de palavras-chaves para encontrar os assuntos, temas e termos importantes. Desta forma, o método permitiu a visualização de indícios ocultos pela complexidade das informações textuais que envolvem a denúncia e propiciou a descoberta de conhecimento entre associações de denúncias com seus temas, áreas e contexto violentos. Deste modo, a denúncia produzida pela central configura-se como uma ferramenta dinâmica, que não se esgota no texto produzido pelo atendente, nem é responsável direta por seus resultados: depende de outros processos para consolidar sua legitimidade como instrumento de produção de informação e conhecimento produzido, sobretudo quando caracterizada como delito, crime ou ato criminal.

Com o agravamento da crise do estado do Rio de Janeiro, novos equipamentos não puderam ser comprados. Os licenciamentos de suprimentos, bem como suas licenças de suporte e manutenção também não puderam ser atualizados. Por isso, optou-se pela construção de um sistema de informação e sua plataforma analítica em um ambiente na nuvem, a fim de reduzir os custos assim como ter acesso restrito e seguro de qualquer lugar. Outros obstáculos de ordem pública ou política surgiram durante realização deste trabalho, como a extinção da secretaria de Segurança Pública, sua

possível mudança de prédio e os acessos aos atuais sistemas de segurança existentes, que poderiam comprometer todo o projeto. Para solucionar estes problemas, definiu-se a criação de um ambiente analítico na nuvem para a composição da proposta deste arcabouço analítico de análise textuais de denúncia demonstrou-se ser a melhor opção novamente. Assim sendo, qualquer mudança estrutural ou operacional na composição envolvida não modificaria ou comprometeria o andamento e implementação do sistema. Foi concebido o ambiente analítico em um formato de *data science* com análise de dados no formato "*self-service*", como dito anteriormente, em uma plataforma que permite que os analistas descubram, preparem e analisem todos os seus dados, além de facilitar a implantação e compartilhamento de análises em grande escala, gerando *insights* mais detalhados para as denúncias recebidas possibilitando desta maneira todo o suporte de análise esperado, bem como dando independência aos analistas e seus estudos, objetivando facilitar suas realizações e viabilizando aprimoramentos conforme novas atribuições e necessidades.

4. Aplicação do SIM

O ambiente analítico necessário ao SIM foi criado, bem como o acesso a todos os seus *dashboards*, configurando-se como a parte analítica utilizada pelo analista de inteligência. Respeitou-se as premissas de gestão do conhecimento empregado na área de inteligência em segurança pública, que estabelece análise do tipo: “Tendências”, “Onde?”, “Quando?”, “Para Quem” e “Nuvem de palavras” por tema específico. Estabelece uma ferramenta de monitoramento em tempo real, das denúncias recebidas comparadas com anos, meses e ciclos anteriores, tornando possível escolher a classificação por tema e observar seus resultados ao longo dos períodos, por áreas, temas ou indivíduos. Trata-se de um ambiente com características de visualização de dados e *data science*, na nuvem, e que por questões de segurança e acessibilidade das informações não podemos detalhar todo o processo. Trata-se de um sistema inovador no âmbito de tratamento e relacionamento, um arcabouço analítico de análise de textos em denúncias, transformando em informações relevantes a temática milícias.

As figuras a seguir apresentam telas do *dashboard* utilizadas no SIM (Figura 3: Onde, Figura 4: Para quem foi feita a Difusão e Figura 5 a nuvem de palavras gerada pelo SIM). Do lado direito, encontramos sempre o mapa do Estado do Rio de Janeiro contendo a quantidade de ocorrências e do lado esquerdo, outras informações relevantes, bem como, os filtros que podem ser mudados nas análises. Os pontos coloridos no mapa refletem a localização geográfica de onde, e quando estão ocorrendo essas denúncias e quais os conteúdos relacionados a Milícia. Serve para entender quando essas ocorrências aconteceram, bem como suas frequências por Dia e Mês, possibilitando desta forma analisar sua sazonalidade, dias da semana, horários entre outros.



SIM - Sistema de Informação de Milícia

Crime por Categoria	Qte
ARMAS E EXPLOSIVOS E MUNIÇÕES	2175
CIDADANIA	8
CRIME ORGANIZADOS	599
CRIMES CONTRA CRIANÇA E O ADOLESCENTE	72
DESAPARECIDOS	22
FALSIFICAÇÃO E PIRATARIA	64
HISTORICO	8
MEIO AMBIENTE	505
ORDEM PÚBLICA	167
OUTROS	29
PRESDIOS	52
PROCURADOS	377
ROUBOS E FURTOS	3233
SERVIÇOS PÚBLICOS	925
TRÂNSITO	11
VIOLÊNCIA CONTRA PESSOA	1040
Total	4852

Bairros (Top 10)	Qte
JACAREPAGUA	662
CAMPO GRANDE	471
SANTA CRUZ	391
PRACA SECA	386
GUARATIBA	220
TAQUARA	210
BANGU	186
CASCADURA	181
RECREIO DOS BANDEIRANTES	180
ITANHANGA	156
Total	3044

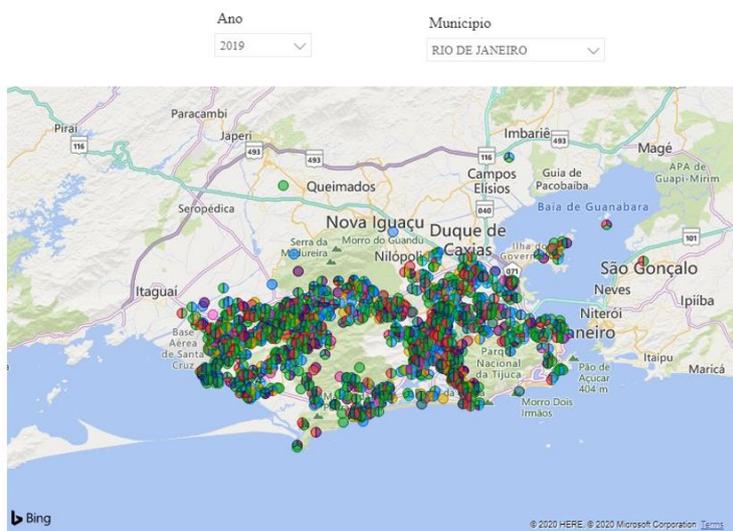


Figura 3: Tela Onde, no Sistema de Informação de Milícias

Na Figura 4, apresentamos quais os assuntos e para quais órgãos de difusão foram enviadas a denúncia e quais os resultados obtidos por esses órgãos. Este tipo de conhecimento permite realizar um melhor planejamento das ações de inteligência, bem como medir e mensurar a eficácia dos resultados de cada órgão. Cabe ressaltar que, nessa etapa de difusão não detalharemos totalmente as especificações adotadas no trabalho por questões de sigilo das informações, porém podemos dizer que, existem órgão, delegacias e batalhões especializados para cada tipo de ocorrência criminal e que foram todos adequados, contemplados e validados no sistema de informação.



SIM - Sistema de Informação de Milícia

Ano: 2019

Municípios: RIO DE JANEIRO

Assuntos: Todos

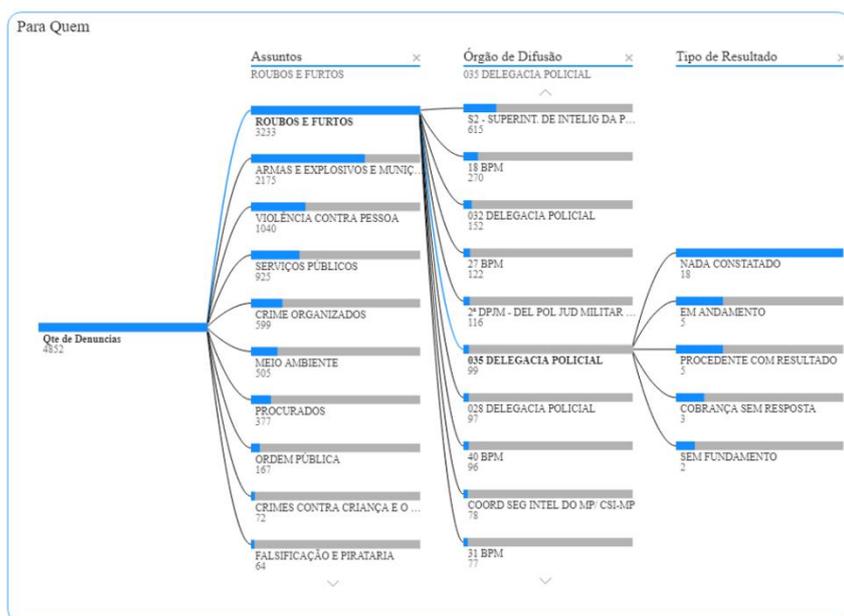


Figura 4: Tela Para Quem foi realizada a Difusão das informações coletadas pelo SIM e quais os resultados obtidos, retroalimentando o sistema de informação

e simbologias utilizadas em denúncias pelos diferentes tipos de crimes que foram utilizadas nas nuvens de palavras. As classificações utilizadas nas denúncias e quais as palavras empregadas, utilizadas por nichos, gueto, grupos, facções, organização criminosa, milícias, temas de crimes entre outros, que se relacionam de alguma maneira entre si ou em conjunto foram identificadas e resolvidas. As tecnologias de *stemming* para a preparação dos dados e clusterização para extração do conhecimento, a partir dos dados classificados, se mostraram adequadas para a abordagem do problema apresentado, assim como a mineração de dados textuais: classificar e relacionar um grande volume de dados textuais de forma que fossem fornecidas informações confiáveis e que tivessem sentido no contexto. Conforme percebemos o impulso analítico de visualização dos dados abarca todo o ciclo de tratamento dos históricos de denúncias, desde a extração de radicais temáticos, formados pelas principais palavras relacionadas às diferentes categorias de denúncia, passando pela construção dos dicionários especializados para apoio à extração de entidades, sua localização e formação das redes relacionais e temas que foram construídos a partir das denúncias. Os cenários foram convertidos em *dashboards*, utilizando técnicas de visualização de dados *Self-Service*, permitindo ao analista da área de inteligência, pesquisar e visualizar toda a análise de vínculos, além de aplicar a sua utilização no ciclo de produção do conhecimento, reconhecido e identificados nas denúncias ao mesmo tempo em que há descoberta de novos conhecimentos e descobertas sobre o tema milícias.

Como exemplo do que foi descoberto em nossas análises, é a representação das palavras dentro de seus respectivos territórios. Assim, é possível extrair mensagens subliminares a partir de suas representações de palavras dentro da denúncia, identificando seus significados e as ações que são capazes de gerar. A partir disso, é possível elaborar um glossário e dicionário de termos que atenda a complexidade sintática e gramatical na base do DD em territórios determinados.

A conversão dos elementos de tendência utilizados na área de inteligência foram respeitados e aplicados: onde, quando, comparativo, variação, e nuvem de palavras encontrando-se apta ao sistema de informação, possibilitando desta maneira uma visão mais profunda da complexidade relacionada ao tema milícias, possibilitando ao analista da área de inteligência com sua utilização ampliar sua visão e possibilidades de análises. Esta é a grande contribuição desse sistema de informação. Um ambiente de coleta e análises de informações, onde os resultados obtidos até o momento são bastante expressivos pois aumentou de maneira significativa a utilização do sistema, bem como os recursos tecnológicos ligados ao seu desenvolvimento aplicado. Desta maneira a metodologia aplicada ao estudo mostra-se relevante e promissora aos objetivos estabelecidos, em como sua aplicação em um mundo real e cheio de desafios para a sociedade e da mesma maneira aos órgãos e áreas de inteligência no enfrentamento da Milícias e da criminalidade como um todo. A partir das análises no decorrer do tempo, foi possível examinar as principais modificações e seus contextos, bem como seus temas de atuação e área geográfica apontados, ocasionando a construção de conhecimento para ao tema Milícias, sendo utilizado em operações de inteligência.

Referências

- Alvari H., Sarkar, S., Shakarian, P. Detection of Violent Extremists in Social Media. Proceedings ICDIS 2nd Int. Conference on Data Intelligence and Security, 2019.
- Barhamgi, M., Masmoudi, A., Lara_cabrera, R., Camacho, D. Social Networking data

- analysys with semantics: application to the radicalization problem. *Journal of Ambient Intelligence and Humanized Computing* (2018).
- Chaudhuri et al 2011, An Overview of Business Intelligence Technology. ACM, N. 8. 2011. Vol. 54, 88-98.
- Chen, H. Chung, W. Xu, J.J. Qin, Gwy, Chau, M. Crime data Mining: A general framework and some examples. *IEEE Computer Society* (2004) 50-56.
- Chen, X., Cho, Y., Jang, S., Crime prediction using Twitter sentiment and weather. *SIEDS Systems and Information Engineering Design Symposium* (2015).
- Furtado, B. A., Sakowski, P. A. M. & Tóvolli, M. H. Modelagem de Sistemas Complexos para Políticas Públicas. IPEA, 2015. Disponível em: <http://www.ipea.gov.br/portal/index.php?option=com_content&id=25860&Itemid=383>. Acessado em 14/09/2019.
- Hull, D.A. Querying Across Languages: A Dictionary-Based Approach to Multilingual Information Retrieval David A. Hull Gregory Grefenstette Rank Xerox Research Centre 6 chemin de Maupertuis, 38240 Meylan France. disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.66.5385&rep=rep1&type=pdf>> Acessado em 14/09/2019.
- Loeck, J.F., Magnus, T.M, Montenegro, D., Mattos, E.V., Gestão da Informação e governos locais: experiências do observatório de Segurança Pública de canoas (RS) e novas possibilidades, *Revista Brasileira de Segurança Pública*, São Paulo, v. 10. n.2, 88-103 Ago/Set 2016. Disponível em:<<http://www.forumseguranca.org.br/publicacoes/gestao-da-informacao-e-governos-locais-experiencias-do-observatorio-de-seguranca-publica-de-canoas-rs-e-novas-possibilidades/>> Acessado em 14/09/2019.
- Pássaro, S., Klein, E. & Loper, E. Natural Language Processing with Python: analyzing text with the natural language toolkit, 2009. Disponível em: https://books.google.com.br/books?hl=en&lr=&id=KGIfiiP1i4C&oi=fnd&pg=PR5&dq=Natural+language+processing+with+Python:+analyzing+text+with+the+natural+language+toolkit&ots=Y3Alv8MGH3&sig=qVmMsCwsls_o9UO5XM22uttVYZ_s#v=onepage&q=Natural%20language%20processing%20with%20Python%3A%20analyzing%20text%20with%20the%20natural%20language%20toolkit&f=false. Acessado em 14/09/2019.
- Storani, P. 2019. Disponível em: <<https://www.bol.uol.com.br/listas/o-que-e-milicia.htm>>, acessado em 08 de setembro de 2019.