

Uso de Computação Cognitiva no Processo Semântico de um Atendente Virtual: Uma Aplicação na área da Segurança Pública

Cristianna M. De Ferran, Claudia L. R. Motta

Programa de Pós-Graduação em Informática (PPGI)
Universidade Federal do Rio de Janeiro (UFRJ)
Caixa Postal 68.530 – 21.941-590 – Rio de Janeiro – RJ – Brasil
cristiannadeferran@yahoo.com.br, claudiam@nce.ufrj.br

Abstract. Approaches to reporting crimes have been widely discussed in recent years, as many cities face a significant increase in violence. Thus, many efforts to encourage citizens to become more active whistleblowers are based on improving the crime reporting model. In this work, we propose the construction of a collaborative reporting platform using a virtual attendant, which was designed to collect confidential information related to public security, simulating human-like behavior, including the identification of intentions through the semantics found in the reports of crimes.

Resumo. Abordagens para denunciar crimes têm sido amplamente discutidas nos últimos anos já que muitas cidades enfrentam um aumento significativo de violência. Assim, muitos esforços para encorajar os cidadãos a tornarem-se denunciadores mais ativos baseiam-se na melhoria do modelo de denúncia de crimes. Nesse trabalho, propomos a construção de uma plataforma colaborativa de denúncias com o uso de um atendente virtual, que foi projetado para coletar informações confidenciais relacionadas à segurança pública, simulando um comportamento análogo ao humano, incluindo a identificação de intenções através da semântica encontrada nos relatos de crimes.

1. Introdução

Desde o século passado, a busca de interação entre humanos e sistemas computacionais que se aproxime da linguagem natural humana vem despertando interesse em diversas áreas do conhecimento, sobretudo entre matemáticos e cientistas da computação. Ao longo dos anos, à medida em que a interatividade homem-máquina vem se aperfeiçoando, o desempenho dos sistemas computacionais que utilizam esta interação vem melhorando, assim como o comportamento das pessoas que interagem com estes sistemas computacionais também vem sendo influenciado.

A nossa pesquisa busca entender como se dá a comunicação através de um diálogo entre duas pessoas de determinada região em uma situação específica, e tentar estabelecer esta mesma comunicação só que, desta vez, entre um indivíduo e uma máquina. Mais especificamente entre um cidadão e um *chatbot*. Nosso contexto são as denúncias de crimes ou delitos relatados pelos cidadãos do Estado do Rio de Janeiro ao Disque Denúncia (DD). O DD é uma organização da sociedade civil de interesse público (OSCIP), sem fins lucrativos, criado em agosto de 1995, quando o Rio de Janeiro enfrentava uma dramática onda de violência. É, não somente, a arma do cidadão na luta contra o crime, mas também uma forma de que este exerça sua cidadania, colaborando

para o êxito das ações da Polícia na segurança pública. As denúncias são feitas anonimamente pela população através de canal telefônico (*Call Center*) e de um aplicativo para *Smartphone (APP)*, construído para atender, inicialmente, ao evento das Olimpíadas realizado no Rio de Janeiro em agosto de 2016. Mensalmente são recebidas aproximadamente 10 mil denúncias, sendo mais de 90% delas recepcionadas pelo *Call Center* e os 10% restantes pelo aplicativo.

Neste contexto, o cidadão liga para o *Call Center* do DD e relata o crime ou delito que testemunhou a um atendente humano. Este atendente entende o que está sendo relatado e formula perguntas ao denunciante para enriquecer a denúncia com informações que sejam relevantes para que as polícias possam atuar com o máximo de eficiência possível para a elucidação dos crimes. Desta mesma forma, nosso *Chatbot*, ou Atendente Virtual (AV), deverá interpretar o relato que está sendo feito pelo denunciante para que, dependendo do tipo de crime que está sendo relatado, possa fazer as perguntas pertinentes para tornar a coleta de informações da denúncia a mais eficaz possível. Esta abordagem é bem diferente das implementações de *chatbots* já conhecidas e presentes em nosso cotidiano, onde o AV é concebido e treinado para auxiliar os humanos em suas tarefas, respondendo a perguntas sobre uma determinada área de conhecimento, consultando bases de dados pré-definidas. Alguns exemplos destes *chatbots* são a Joice da companhia telefônica Oi e a Bia do Banco Bradesco. Estes *chatbots* foram desenvolvidos para sanar dúvidas e dar informações aos clientes destas empresas.

Abordagens para denunciar crimes ou suspeitas da evidência de um crime foram amplamente discutidas nos últimos anos. Os relatos criminais são informações críticas para os investigadores e policiais, que levaram à elucidação de uma série de crimes (Kun e Leroy, 2014; Chainey et al., 2008). Assim, muitos esforços para encorajar os cidadãos a se tornarem “repórteres de crime” mais ativos baseiam-se na melhoria dos benefícios do modelo de denúncia de crimes (Lasley e Palombo, 1995). Kanable (2008) destaca que, atualmente, as plataformas de redes sociais, como o Nextdoor, fornecem mecanismos para que os membros da comunidade compartilhem informações tanto com a aplicação da lei quanto entre os cidadãos. Mais recentemente, Mohler e Brantingham (2018) afirmaram que as denúncias da população continuam a ser a melhor fonte colaborativa de informação sobre crime e desordem social. Um aspecto importante é a introdução de métodos de relatos de alta tecnologia, como sistemas de resposta a emergências, alarmes de vigilância e o uso de telefones celulares (Lasley e Palombo, 1995). Os autores afirmaram que a tecnologia de comunicação continuará impactando na decisão dos cidadãos em denunciar crimes à polícia. Eles também preveem que o uso de ferramentas computacionais interativas deverá ser substancialmente maior do que usar o modo de chamada de voz padrão.

Nesta pesquisa propusemos o uso de *chatbots* e Computação Cognitiva (CC) para melhorar a interação com o cidadão e disponibilizar um novo canal de comunicação aplicado à segurança pública. Fundamentamos nossa proposta no uso bem-sucedido de *chatbots* no gerenciamento do relacionamento com os clientes (CRM), bem como o uso frequente de *chatbots* para ajudar os usuários a obter assistência e informações oportunas e eficientes, além de motivações relacionadas a entretenimento, fatores sociais e relacionais (Brandtzæg e Følstad, 2017). As pessoas preferem interagir com serviços por meio de *chatbots* disponíveis para aplicativos populares de mensagens. Os robôs são uma solução oportuna para uma geração que prefere mensagens de texto aos meios tradicionais de comunicação (Ciechanowski et al., 2018).

Nosso chatbot foi projetado para coletar informações confidenciais relacionadas à segurança pública, através do diálogo entre o denunciante e o AV que chamamos de Dora. Para tal, foi necessário usar técnicas de psicologia para estabelecer um vínculo de confiança e cooperação com o denunciante. A solução permite simular um comportamento análogo ao humano, incluindo a identificação de intenções através da semântica encontrada nos relatos. Estudamos o padrão de fala, os jargões e gírias usadas por uma pessoa que faz denúncias de crimes, principalmente os relacionados ao Tráfico de Drogas, Armas, Roubo de Carga e Roubo de Veículos. Nesse sentido, a CC é vista como uma das soluções proeminentes para desenvolver aplicações do tipo chatbot, permitindo que estas sejam mais assertivas (Rodríguez et al., 2014). A técnica usada para estabelecer o diálogo é chamada Rapport, que representa a maneira de criar uma conexão empática, estabelecer um vínculo de confiança com o interlocutor, promovendo assim a possibilidade de receber informações confidenciais e permitindo que o denunciante se sinta protegido. Os elementos importantes de Rapport que foram usados em nosso diálogo foram: apresentação inicial ao denunciante, deixando claro que ele está dialogando com um robô; perguntas para completar a compreensão do que está sendo denunciado; confirmação da informação com o denunciante; linguagem adaptada ao interlocutor; garantia do anonimato da denúncia.

A solução foi avaliada por meio de um estudo de caso realizado no DD. Os resultados apresentaram evidências de que a AV consegue identificar as intenções dos relatos e, principalmente, consegue interagir fazendo perguntas ao denunciante para enriquecer as denúncias com informações relevantes. As contribuições desta pesquisa são duas: primeiro, o artefato e a tecnologia aplicada provaram resolver o problema de como estabelecer uma nova forma de interação dentro do domínio estudado; segundo, este é o ponto de partida para a construção de um corpus de conhecimento que possa apoiar desenvolvimentos relacionados a denúncias de crime.

1.1. Domínio

O contexto específico desta pesquisa é o programa DD, que recebe denúncias anônimas dos cidadãos a respeito de ocorrências de crimes no Estado do Rio de Janeiro. Este escopo foi determinante em nos guiar a pesquisar um grupo específico de pessoas, as que realizam denúncias ao DD e que denunciam os tipos de crime que resolvemos estudar.

1.2. Problemas Identificados e Questões de Pesquisa

Dentro do domínio que estabelecemos identificamos três principais problemas que nortearam nossa pesquisa e listamos a seguir:

Redução da disponibilidade do atendimento às denúncias no DD.

O DD recebia verbas e doações de empresários, do governo do estado do Rio de Janeiro e da Federação das Indústrias do Estado do Rio de Janeiro (FIRJAN), dentre outras fontes de financiamento e em 2016 teve estas fontes de financiamento drasticamente reduzidas por causa da crise econômica e financeira pela qual passava o Rio de Janeiro. Como consequência, o DD precisou reduzir o atendimento do *Call Center* eliminado o atendimento nas madrugadas e fins de semana. A redução das operações impactou a apuração de crimes que acontecem primordialmente nestes dias e horários.

Dificuldade em registrar as denúncias sem o viés do atendente do *Call Center*.

As denúncias que entram pelo canal *Call Center* são recepcionadas por um atendente humano e registradas em um formulário eletrônico (DDSys) segundo o entendimento do atendente. Ele segue o roteiro predefinido no DDSys, faz as perguntas que julga necessárias para enriquecer o relato da denúncia e preenche o campo do formulário relativo ao relato com suas próprias palavras e segundo seu próprio entendimento.

Falta de padronização das informações relevantes para cada tipo de crime.

Não há padronização das perguntas que deverão ser feitas ao denunciante por tipo de crime para que se obtenha o máximo de informações relevantes para sua elucidação. A qualidade das informações capturadas depende da experiência e percepção do atendente do *Call Center*. No caso de denúncias feitas através do *App*, o denunciante preenche um formulário eletrônico onde não há nenhuma interação para que este se sinta estimulado a fornecer informações sobre outros crimes além do que está sendo relatado.

A partir dos problemas identificados levantamos as seguintes questões de pesquisa:

QP1 – É possível construir um canal digital que interaja com o denunciante, respeitando e utilizando seu vocabulário próprio, para melhoria na captação de denúncias?

QP2 – É possível estabelecer uma comunicação com o denunciante para extrair, a partir do relato, informações sobre sub-denúncias?

1.3. Relevância

Entendemos como aspectos relevantes para a realização desta pesquisa:

- Dar voz a pessoas em situação de risco, adaptando-se à linguagem do usuário.
- Disponibilizar um canal alternativo de denúncias com baixo custo operacional, preservando o anonimato do denunciante.
- Oferecer mais um canal para se fazer denúncia, ampliando desta forma o atendimento ao cidadão e atraindo um novo público, que deseja denunciar ocorrências de violências e crimes.

2. Referencial Teórico

Chatbot é uma ferramenta projetada para simular uma conversa, utilizando linguagem natural, entre o ser humano e a máquina. Através dos anos, esse tipo de ferramenta tem sido aprimorado e aplicações podem ser encontradas em ambientes científicos, comerciais, de entretenimento, dentre outros (RESHMI, S.; BALAKRISHNAN, Kannan, 2016). De maneira genérica, os *chatbots* são preparados para responder a perguntas em uma área específica de conhecimento. Entre as atividades em que pode ser aplicado estão: assistência virtual, tutoria, e-commerce, redes sociais (RESHMI, S.; BALAKRISHNAN, Kannan, 2016). Assim, os *chatbots* são usados por milhares de usuários da web para mediar o acesso a dados ou bases de conhecimento, como também para realizar conversas genéricas (GAGLIO, Salvatore et al., 2019).

Um dos primeiros *chatbots* desenvolvidos, o ELIZA (WEIZENBAUM, Joseph, 1966), foi baseado em técnicas simples de correspondência de padrões. Ele foi o primeiro software para simulação de diálogos. O programa possuía 204 linhas de código-fonte e utilizava técnicas de inteligência artificial para simular um diálogo entre paciente e

psicólogo. Sua inteligência consistia em dar respostas baseadas nas frases ditas pelos pacientes. O ELIZA teve sua importância pelo pioneirismo, embora as ferramentas atuais realizem análises bem mais profundas dos dados e misturem técnicas estatísticas tais como teorias bayesianas (ABDUL-KADER, Sameera A.; WOODS, John C, 2015).

Brandtzæg e Følstad (2017) defenderam a “Teoria de Usos e Gratificações” (U&G) (Rubin, 2009) como uma base teórica para entender a motivação das pessoas para usar os *chatbots*. O autor argumenta que a U&G explica por que e como as pessoas usam mídias específicas para atender suas necessidades. A teoria afirma que o uso específico de um meio depende da gratificação esperada e experimentada que ele proporcionará ao usuário. Neste contexto, usos e gratificações referem-se à motivação para o uso de um meio e à satisfação que as pessoas ganham com o seu uso. Muitos tipos de gratificações podem ser motivadores de um uso da mídia. Sundar e Limperos (2013) sugerem a necessidade de informação, entretenimento, interação social e autoexpressão. De acordo com Brandtzæg e Følstad (2017), como os *chatbots* são uma tecnologia recente, eles são usados principalmente por inovadores e pioneiros, que, por sua vez, provavelmente têm necessidades e gratificações diferentes do restante da população. Então, outra teoria que pode ajudar neste caso é a difusão de inovações. Isso esclarece como uma população adota inovações. Assim, Brandtzæg e Følstad (2017) argumentam que a combinação de U&G e difusão de teorias de inovações poderiam fornecer uma visão sobre a compreensão das várias necessidades e gratificações dos usuários entre os diferentes segmentos da população. “Adotantes precoces” são geralmente mais orientados para o risco e curiosos sobre novas tecnologias, enquanto os tardios e retardatários são mais conservadores e avessos ao risco”. Brandtzæg e Følstad (2017) concluíram que os projetistas deveriam se concentrar no desenvolvimento de *chatbots* que são percebidos como úteis para fornecer ajuda ou informações necessárias de maneira efetiva e eficiente. Portanto, os projetistas do *chatbot* devem identificar casos em que os *chatbots* atendam à necessidade de produtividade dos usuários com mais eficiência do que seria através de outros métodos de interação. Argumentamos que nossa aplicação preenche esses critérios. Além disso, destacamos a importância da comunicação por meio do texto escrito, que pode favorecer a segurança do denunciante, uma vez que pode haver outras pessoas ao seu redor, o que pode constrangê-lo ou até assustá-lo.

3. Trabalhos Relacionados

Os avanços tecnológicos têm sido eficazes para ajudar a sociedade com esse tipo de problema ao fornecer ferramentas e técnicas inovadoras para o relato de crimes e suspeitas de crimes (Chaudhuri et al., 2018; Jimoh et al., 2014; Varsha et al., 2015). Alguns exemplos desse tipo de iniciativa são: Crime Stoppers, de Los Angeles, mensagens SMS (Song et al. 2009), aplicativos para *iPhone*, *iPad* e *Android*, *Twitter Crime Stoppers* da polícia de Newark e o próprio Disque-Denúncia no Brasil. No entanto, os sistemas que oferecem suporte a esses serviços ainda se concentram em métodos tradicionais de interação com os cidadãos por telefonemas, interface com o computador por meio de preenchimento de formulário ou mensagens de texto simples. Assim, identificamos a necessidade de melhorar esses serviços, utilizando ferramentas tecnológicas mais inovadoras, capazes de lidar com a interação, com menores custos tanto para o cidadão quanto para a administração pública.

Há alguns trabalhos correlatos considerando classificação de textos em português (Cesar et al., 2019; Ferreira, 2019; Nascimento, 2019; Andrade, 2015; Rossi, 2015),

inclusive trabalhos para português coloquial (Stiilpen, 2016). No entanto, nenhum tão específico como os trabalhos mais próximos ao problema descrito acima, no qual os autores (Pinho et al., 2017) desenvolveram um modelo para classificar denúncias feitas através do *call center* do Disque-Denúncia(DD) utilizando o algoritmo Weightless Neural Network e (Brito et al. 2018), aplicando análise semântica a partir de sistemas complexos na base de dados do DD. Em ambos os trabalhos foram criadas tabelas de palavras e tipo de classificação das denúncias, bem como as relações dessas classificações dentro de cada denúncia. Deste modo, a classificação das denúncias recebidas pelo call center é usada para a organização dos dados e para estipular a importância de uma determinada palavra, expressão ou contexto na denúncia e sua relação com as outras denuncia utilizando, determinando e criando uma rede neural sem peso aplicada a base de dados do DD. Importante ressaltar que o registro dessas denúncias é feito por atendentes treinados, que transcrevem a denúncia de forma padronizada, gerando um texto com qualidade muito superior e livre dos possíveis erros gramaticais. Destaca-se que a classificação das denúncias desse trabalho é feita pelo atendente, o que dispensaria a classificação automatizada. As Informações contidas nas denúncias por temas permitiu a criação de um dicionário específico e único de palavras, termos e semânticas utilizadas em diferentes tipos de crimes denunciados, nos permitindo empregá-los na construção das nuvens de palavras relacionadas às milícias (Brito et al, 2020).

4. Construção da Plataforma de Denúncias

Como tínhamos o objetivo de buscar uma solução coletiva para um problema real que o DD enfrentava e aumentar nosso conhecimento sobre o domínio a ser estudado, decidimos adotar a pesquisa-ação como metodologia da nossa pesquisa.

Para fazer a pesquisa de linguagem que é necessária para viabilizar a interação entre o denunciante e o nosso atendente virtual, nós iniciamos com um Design Thinking (DT), realizado em agosto de 2018 junto à equipe do DD. No segundo momento fizemos uma imersão no DD para acompanhar a rotina dos atendentes do *call center* e da área de difusão das denúncias. Em seguida, analisamos os dados da base de denúncias. E por fim, utilizamos metodologia ágil para desenvolver a ferramenta computacional para recepção das denúncias, e a cada entrega de funcionalidade a ferramenta era homologada pelos atendentes do DD. Os defeitos reportados assim como as sugestões de melhorias eram desenvolvidos na sprint seguinte. A curadoria do corpus também era uma atividade contínua de cada sprint.

4.1. Design Thinking

O DT está ancorado em três principais pilares: a empatia, a colaboração e a experimentação. Nossa intenção ao realizar a dinâmica de DT era a de proporcionar um ambiente de colaboração centrado nas pessoas, com o objetivo de conhecer para quem estaríamos construindo uma solução. Buscando uma solução coletiva para um problema real que o DD enfrentava e aumentando nosso conhecimento sobre o domínio a ser estudado. Dos artefatos gerados no DT destacamos três principais que foram direcionadores da nossa pesquisa: o mapa de empatia das personas Dora, a atendente do *call center* e Janete, a denunciante, o mapa de ideias e o mapa de priorização das ideias. As informações geradas na dinâmica de DT são de propriedade intelectual do DD. Por isso, nas imagens a seguir, fazemos um resumo das informações, mas não fornecemos pleno acesso.

Esta parte da dinâmica nos ajuda a conhecer as personas que serão o foco da nossa pesquisa entendendo o que as motiva, quais são suas necessidades, o que elas pensam, falam, fazem e, principalmente, sentem. A Figura 1 mostra o mapa de empatia da Dora, atendente do *call center* do DD, após realizada a convergência com o time.

+ Empathy Map Dora (Atendimento)

Fala

O anonimato está garantido.
Fala pouco e ouve muito.
Acalma o denunciante.

Pensa

Se o sistema cair agora, já era!
Registrar a denúncia com qualidade.
Preciso atingir minha meta diária.

Faz

Atende a muitas denúncias e digita muito.
Classifica denúncias e sinaliza as mais importantes.
Preenche roteiro e encaminha denúncia p/ difusão.

Sente

Raiva, ansiedade, cansaço, pressão, preocupação.
Sente a dor do denunciante.
Indignação com os crimes que ouve.

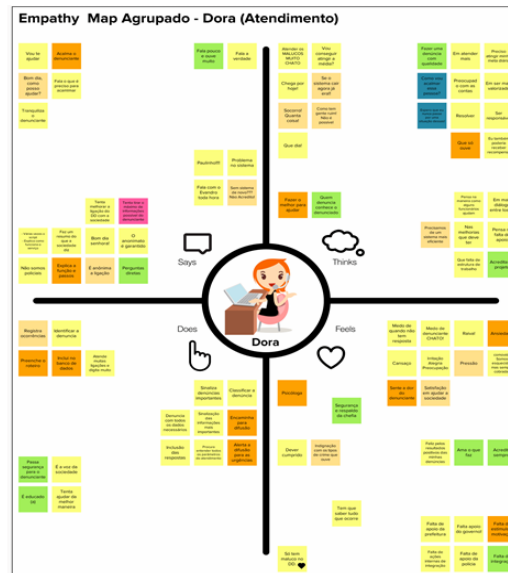


Figura 1. Mapa de empatia da Dora.

A Janete, por outro lado, foi a persona mapeada como a cidadã que realiza as denúncias e ao final da dinâmica obtivemos o resultado mapeado na Figura 2.

+ Empathy Map Janete (Cidadã)

Fala

Vocês (DD) são a última ajuda!
Para os filhos estudarem para mudar de vida.
Sobre os crimes que presencia.

Pensa

Em sair do RJ e voltar para lugar de origem.
A polícia não resolve nada, é corrupta.
Ter uma vida melhor e mais segurança p/ a família.

Faz

Trabalha duro para sobreviver.
Compra produtos piratas ou roubados.
Se submete às leis do tráfico.

Sente

Medo que os filhos virem bandido
Insegurança, cansaço, dor, desamparo e esperança
Medo do companheiro



Figura 2. Mapa de empatia da Janete.

Depois de a jornada da Dora, a visão dos próximos passos (roadmap) a curto, médio e longo prazo, os principais objetivos da sessão de DT e a frase final que iríamos perseguir; aplicamos a dinâmica de grandes ideias (Big Ideas), uma excelente oportunidade dos participantes promoverem um debate (brainstorming) sobre ideias possíveis. Neste momento não existem ideias ruins ou impossíveis, é a hora de liberar a

criatividade em busca de soluções. Após a convergência das ideias tivemos o painel de ideias ilustrado na Figura 3.

+ Big Ideas

- Utilização de Computação Cognitiva
- Melhorias dos sistemas, App e Site
- Georeferenciamento
- Analytics – Sistema de Alertas
- Ações de Comunicação e Marketing
- Ações de RH e Gestão
- Melhoria na relação com Órgãos Externos
- Melhorias na Estrutura física do DD
- Telefonia

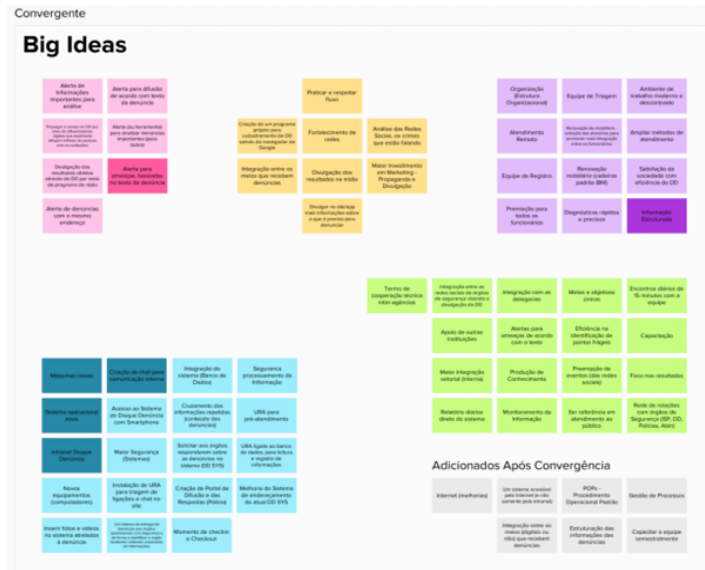


Figura 3. Resultado da dinâmica Big Ideas no DT.

Ao longo do DT foram geradas 100 ideias que foram categorizadas e 34 delas foram priorizadas. A dinâmica de priorização das ideias ajuda o time concentrando as discussões na importância e viabilidade pela ótica da persona final. Para isso é apresentado um gráfico onde ficam dois eixos, importância para o cidadão e a viabilidade daquela ideia, que leva em consideração diversos aspectos como tecnologia, orçamento, ética etc. São fornecidos ainda uma quantidade de “pontos” diferenciados por cor (vermelho para o eixo importância e azul para o eixo viabilidade) para cada participante votar nas ideias convergidas e apresentadas. Após a priorização das ideias tivemos o mapa de ideias ilustrado na Figura 4.



Figura 4. Resultado da dinâmica Mapa de Priorização no DT.

4.2. Imersão no Disque-Denúncia

Após a dinâmica de DT fizemos uma imersão no *call center* do DD para observar a abordagem e estratégias que os atendentes utilizam para capturar informações relevantes de cada denúncia. Durante esta imersão identificamos que o denunciante liga para fazer um tipo de denúncia e o atendente, baseado no que está sendo relatado, identifica outros tipos de denúncias. Deste modo o atendente faz perguntas relacionadas a estas outras denúncias e registra todas as informações coletadas sobre cada uma delas no sistema com o identificador de sub-denúncia, relacionando-as à chamada denúncia principal. A partir deste achado vimos a necessidade de analisarmos a base de denúncias para mapear as sub-denúncias mais frequentes para cada tipo de denúncia com o objetivo de criar as perguntas pertinentes para as sub-denúncias, simulando a ação dos atendentes humanos.

4.3. Análise da Base de Dados do Disque-Denúncia

Na primeira parte deste mergulho na base de dados do DD, analisamos 1500 denúncias reportadas através do *App*, que são aquelas reportadas diretamente pelo cidadão com suas próprias palavras e montamos nuvens de palavras por tipo de crime que estava sendo relatado. Nosso objetivo era conhecer melhor como o cidadão se expressa ao realizar um relato de crime. Fizemos a separação de denúncias por tipo de crime e verificamos que o vocabulário e expressões se alteravam de um tipo de crime para outro. Precisamos fazer este mapeamento para refletir na construção do diálogo, das entidades e intenções do *Chatbot*. Fazer as perguntas certas é fundamental para obter respostas mais assertivas. Para Els Dragt e Jeroen Timmer (2020), fazer perguntas nos torna humanos, ajuda a estabelecer conexões, aprender e transformar.

4.4. Construção das Intenções

Uma intenção representa a finalidade do texto inserido na interface de conversação com o usuário. Em nosso contexto é o tipo de crime que o denunciante deseja relatar. Definimos uma intenção para cada tipo de relato que coletamos na solução. Além das intenções que representam a coleta de relatos, criamos intenções auxiliares para orientar o diálogo com o denunciante, como possíveis sub-denúncia (baseado no mapeamento que fizemos de sub-denúncias mais frequentes por tipo de denúncia principal), denúncias validadas, intenção não reconhecida e outras denúncias. Uma intenção é definida pela estrutura: #nome da intenção e exemplos que se referem a essa intenção. A Tabela 1 mostra alguns exemplos de Intenções cadastradas.

Tabela 1. Lista de algumas Intenções com 2 exemplos cada uma

Intenção	Exemplo 1	Exemplo 2
#Sobre_consumo_de_drogas	consumindo drogas	usando entorpecentes
#Sobre_foragido_da_Justiça	o foragido da justiça	o fugitivo foi visto circulando livremente
#Sobre_menor_envolvido	aliciando crianças	recrutando menores
#Sobre_obstrução_via_publica	montaram barricadas	criaram uma trincheira
#Sobre_posse_de_armas	exibindo fuzis	homens armados
#Sobre_violencia_contra_a_mulher	o marido sempre bate na mulher	meu ex-marido está me perseguindo
#tráfico_de_drogas	distribuindo drogas	vendendo entorpecentes
#Veiculo_roubado	desmonte de carros	carro supostamente roubado

4.5. Construção das Entidades

As Entidades são termos relevantes para a identificação de uma Intenção, para o propósito da denúncia e para o contexto do diálogo realizado. Ao reconhecer as entidades no texto escrito pelo denunciante, podemos enriquecê-lo e tomar as decisões necessárias na condução do diálogo. Definimos várias entidades para enriquecer as intenções e auxiliar no fluxo de coleta de dados necessários para um relato eficaz. Para cada intenção teremos entidades relacionadas a ela além das entidades que são comuns a várias intenções diferentes. A entidade é definida por: nome, ocorrências e sinônimos. A Tabela 2 mostra algumas Entidades cadastradas com suas ocorrências e 2 sinônimos para cada ocorrência.

Tabela 2. Lista de algumas Entidades com ocorrência e 2 sinônimos cada uma

Entidade	Ocorrência	Sinônimo 1	Sinônimo 2
@arma_de_fogo_de_grande_porte	Fuzil	fuzis	AR-15
@arma_de_fogo_pequeno_porte	revolver	revólver	draga
@cor_de_pele	preto	preta	negro
@droga	Ecstasy	balinha	extase
@envolvido_estatura	alto	alta	grandão
@envolvido_tipo_fisico	forte	fortinho	fortinha
@facao	Comando Vermelho	CV	C.V.
@frequencia	semanal	semanalmente	de 2a a 6a
@menor	moleque	mike	mulek
@procurado_em_potencial	foragido	foragida	foragidos
@turno	noite	anoitecer	anoitecendo

4.6. Construção do Diálogo

O diálogo é fluxo ramificado de conversação, que define como a AV responde quando reconhece as Intenções e Entidades definidas na conversa com o denunciante, composta por vários nós de diálogo. A Figura 5 ilustra o tronco principal do nosso diálogo.

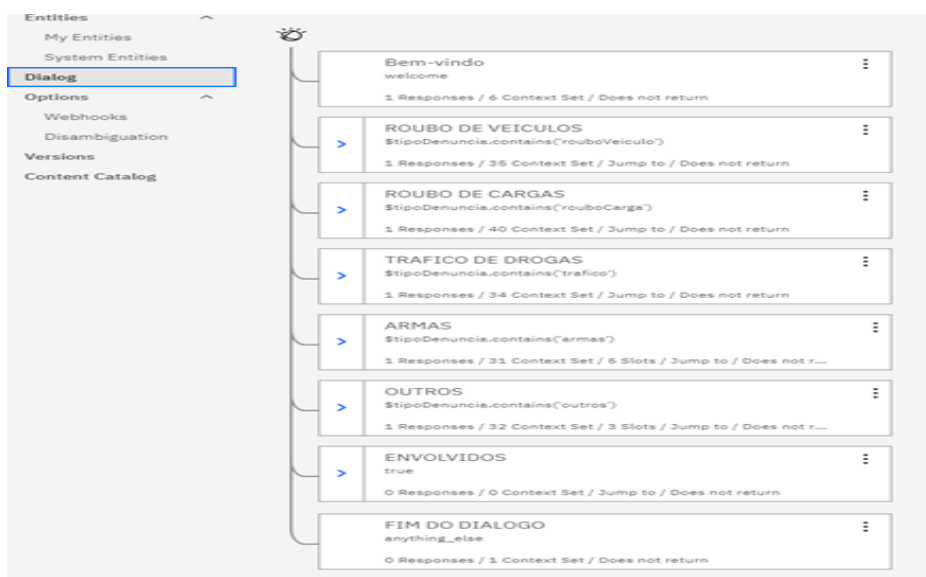


Figura 5. Tronco principal com os nós do diálogo da Dora

No nosso contexto o diálogo tem um tronco principal composto por nós principais que são: nó para a intenção de saudação, um para cada intenção dos tipos de crime principais (tráfico de drogas, armas, roubo de carga, roubo de veículo, outros), nó para intenção não identificada e nó para intenção de fim de diálogo. Para uma resposta humanizada, devemos criar vários nós de diálogo encadeados horizontalmente para cada nó de diálogo de Intenção capturado (nós filhos). A caixa de diálogo é definida por: nome, gatilho e respostas (que podem ser condicionais ou aleatórias). Além dessa estrutura, existe o conceito de saltos (“*jump to*”) usados para ir de um nó de diálogo para outro após a resposta do denunciante.

4.7. Interface com a Atendente Virtual Dora

Iniciamos a concepção da Dora com um design sofisticado para proporcionar uma boa experiência ao usuário e estimulá-lo a utilizar o novo canal de denúncias a ponto de preferir realizar suas denúncias através da Dora. Afinal, este canal tem um custo operacional bem mais baixo do que o *call center*, tem alta disponibilidade (atende 24 horas por dia, 7 dias na semana) e as denúncias são registradas exatamente como relatadas pelo denunciante, com a vantagem sobre o atual *App* do DD de interagir com o denunciante para extrair mais informações sobre o crime que está sendo relatado e identificar possíveis sub-denúncias relacionadas. Quando começamos os testes unitários, percebemos que um dos membros do time, que é morador de favela, não se sentia confortável em realizar suas denúncias em locais públicos por ter receio de ser identificado como um denunciante pelas pessoas a sua volta. Tivemos que “voltar para a prancheta” e refazer o desenho da Dora para que o diálogo da denúncia ficasse, para quem olha de fora, muito parecido com um diálogo feito através do WhatsApp, por exemplo. A Figura 6 mostra o início do diálogo com a Dora, começando com uma saudação, deixando claro para o denunciante que se trata de uma AV e não um humano, que o anonimato está garantido, oferecendo as opções denúncias que poderão ser feitas e pedindo para que o denunciante realize sua denúncia.



Figura 6. Saudação da Dora para iniciar a denúncia

5. Validação das Questões de Pesquisa e Resultados

A aplicação que construímos é a Dora, AV de denúncias de crimes para o DD. Para validar a pesquisa foram feitos dois experimentos. No primeiro experimento verificamos se a Dora era capaz de atender a denúncias de tráfico de drogas, armas, roubo de carga, roubo de veículo e outros tipos de crimes de maneira análoga ao atendente do *call center*. Para isso era preciso que a Dora fosse capaz de fazer as perguntas pertinentes ao tipo de crime que estava sendo relatado e coletasse as informações necessárias para que a denúncia pudesse ser encaminhada, pelo time de Difusão de denúncias, para os órgãos competentes. Foi possível atender a 78 denúncias através da Dora, durante o período de 10 dias, distribuídas da seguinte forma: 7 de armas; 10 de roubo de carga; 12 de roubo de veículo; 40 de tráfico de drogas e 9 de outros tipos de crimes. Na avaliação do time de Difusão, das 78 denúncias apenas 2 não foram compreendidas pela Dora, ambas do tipo Outros, como ilustrado na tabela 3.

Tabela 3. Quantidade de denúncias recebidas versus denúncias aprovadas

Tipo de Denúncia	Qtd Denúncias Recebidas	Qtd Denúncias Aprovadas
Armas	6	6
Roubo de carga	10	10
Roubo de veículos	14	14
Tráfico de drogas	39	39
Outros	9	7
Totais	78	76

No segundo experimento queríamos verificar se a Dora era mais eficaz do que o *App* do DD por ser capaz de identificar e reconhecer sub-denúncias através do relato e estabelecer um diálogo fazendo perguntas pertinentes para obter informações sobre as sub-denúncias. Selecionamos 123 denúncias da base de dados de denúncias do *App* no período de novembro a dezembro de 2019, distribuídas da seguinte forma: 11 de armas; 16 de roubo de carga; 20 de roubo de veículo; 62 de tráfico de drogas; 14 de outros tipos de crimes. A tabela 4 mostra a quantidade de denúncias com potencial sub-denúncia e dentre elas quantas foram identificadas pela Dora.

Tabela 4. Quantidade de sub-denúncias identificadas por tipo de denúncia

Tipo de Denúncia	Total de denúncias	Potencial sub-denúncia	Sub-denúncias identificadas	% Sub-denúncias Identificadas
Armas	11	6	5	83,3%
Outros	14	1	0	0,0%
Tráfico de Drogas	62	58	53	91,4%
Roubo de Veículos	20	12	10	83,3%
Roubo de Carga	16	9	7	77,8%
Totais	123	86	75	87,2%

Em média 87,2% das denúncias onde havia sub-denúncias, estas foram identificadas pela Dora e ela fez perguntas pertinentes relacionadas aos tipos de crime destas sub-denúncias para enriquecer o relato e possibilitar o encaminhamento através do time de Difusão de denúncias para os órgãos competentes. As sub-denúncias enriquecem a denúncia principal com as informações coletadas pela Dora, porém a tipificação da denúncia mantém-se a da denúncia principal.

6. Conclusão

Esta pesquisa propõe a AV Dora, um *chatbot* implementado para ser um canal de comunicação alternativo para denunciar crimes e delitos testemunhados pelos cidadãos do Rio de Janeiro, garantindo o anonimato do denunciante e preservando a integridade do relato. Implementamos a proposta no contexto das denúncias feitas ao DD. A solução mostrou-se viável e eficiente como pudemos constatar através dos resultados obtidos nos experimentos descritos nos resultados. Os relatos coletados pela Dora são registrados no mesmo banco de dados que os coletados pelo *call center* e pelo *App*. Assim, os resultados obtidos encorajam possíveis desdobramentos desta pesquisa.

6.1. Contribuições da Pesquisa

Como principal contribuição vemos a construção de um modelo ancorado na tríade Intenções, Entidades e Diálogo, para acompanhar o dinamismo com que as denúncias se modificam, tornando-se um modelo vivo e de fácil manutenção. As denúncias refletem o momento em que a sociedade está passando e há uma grande diversificação e surgimento de novos crimes e delitos além da alteração nas suas frequências. O vocabulário utilizado pelo denunciante também é vivo. Novas gírias, expressões, abreviações surgem a cada dia e o modelo precisa ser evoluído nesta mesma direção. A construção da Dora vem com o intuito de sanar ou neutralizar problemas que elencamos: a) Redução da disponibilidade do atendimento às denúncias no DD; b) Dificuldade em registrar as denúncias sem o viés do atendente do *Call Center*; c) Falta de padronização das informações relevantes para cada tipo de crime.

Deste modo, também vemos como contribuição desta pesquisa a disponibilização de um canal de comunicação de denúncias eficaz, que funcione 24 horas por dia, 7 dias na semana, e a possibilidade de registrar a denúncia com as palavras do denunciante, sem o viés do atendente do *call center*. Durante a pesquisa observamos a riqueza de informações que os relatos das denúncias trazem e que atualmente é pouco explorada pelo DD e pelos órgãos de inteligência e segurança pública. A melhoria contínua na captura de informações relevantes para cada tipo de denúncia também se mostrou uma contribuição importante. Ela é possível uma vez que gravamos na base de dados do sistema de atendimento de denúncias do DD todas as perguntas que são feitas pela Dora e as respectivas respostas do denunciante e disponibilizamos para a curadoria. Desta forma, é possível analisar cada diálogo, identificar algum ponto de melhoria e atualizar o corpus da aplicação de forma contínua e caso seja necessário.

Percebemos também como contribuição para o DD a nossa forma de trabalho colaborativo que trouxe luz para a necessidade de maior interação entre as áreas de Atendimento, Difusão e Análise e da sistematização de um processo de melhoria contínua na recepção das denúncias do *call center*. Eles já faziam uma curadoria informal toda vez que o time de Difusão identificava uma denúncia que não tinha os dados mínimos para ser encaminhada para as polícias ou quando havia um tema “quente” na mídia, mas esta comunicação ficava dependente da senioridade de cada profissional. A sistematização deste processo aproximou mais os times de Atendimento e Difusão e trouxe ganho de produtividade na medida em que os atendentes passaram a ser mais assertivos nas perguntas do script de atendimento e menos denúncias são descartadas pela Difusão.

6.2. Pesquisas Futuras

Apontamos a seguir algumas sugestões de trabalhos futuros que julgamos relevantes.

Aprofundar outros tipos de denúncias mais frequentes: Vislumbramos como uma continuidade natural deste trabalho mapear outros tipos de denúncias que se mostrem mais frequentes com o objetivo de detalhar as Intenções e Entidades e evoluir o corpus da aplicação, buscando o aprofundamento em temas que conversam entre si do mesmo modo que fizemos para temas relacionados ao tráfico de drogas.

Verificar se a disponibilização de um novo canal de comunicação de denúncias engaja novos denunciante: Pesquisar se o novo canal de comunicação de denúncias tem uma distribuição percentual parecida com a dos tipos de denúncias que são reportadas ao *call center* e enviadas pelo *App* e se as localizações das ocorrências também são semelhantes. Caso não sejam, este dado enseja a hipótese de que teríamos um novo denunciante reportando ocorrências ao DD que não é o denunciante usual. Estaríamos atingindo um novo público, com hábitos e vocabulário próprio que precisaria ser estudado para garantir o sucesso do novo canal.

Desenho de modelo preditivo de ocorrências: O cruzamento de informações contidas nos relatos das diversas denúncias pode ser usado para detectar padrões de atuação de quadrilhas e facções criminosas possibilitando o desenho de um modelo preditivo de ocorrências, os chamados modelos inteligentes. Estas informações são ferramentas importantíssimas para pautar as ações das polícias investigativas e na tomada de decisão dos gestores da segurança pública e das Agências de Inteligência.

Análise de sentimento nos relatos das denúncias: Existem informações relevantes entre centenas ou milhares de textos das denúncias que nos permitiriam aplicar uma análise de sentimentos para entender melhor as reações dos denunciante e melhorar sua interação com a Atendente Virtual (Bird et al., 2019).

Glossário de expressões de facções criminosas: Analisar as formas de escrita dos denunciante para identificar se as expressões utilizadas se modificam de acordo com a localidade da ocorrência X facção criminosa que domina a área, possibilitando a criação de um glossário de cada região e das facções criminosas. Este glossário pode ser utilizado pelas Agências de Inteligência para interceptação de mensagens em redes sociais.

Reconhecimento facial das imagens enviadas: Os denunciante enviam fotos e vídeos através do *App* do DD e da Dora. Esta fonte relevante de informação é pouco explorada porque a quantidade de imagens é muito grande e sua análise é feita de forma manual. Poderia ser implementado um processo automático utilizando IA para reconhecimento facial nas imagens que são enviadas, fazendo o cruzamento com a base de Procurados do DD ou das delegacias para identificação de foragidos da justiça.

Latitude e Longitude: Outra perspectiva futura é a utilização dos dados de latitude e longitude, coletados nos PINs de localização das ocorrências e nos PINs auxiliares, para serem plotados num mapa por tipo de denúncia para que os órgãos de segurança possam planejar suas ações em determinadas regiões baseando-se nas informações precisas fornecidas pelos denunciante.

Aplicativo para Registro de Ocorrências (R.O): Poderia ser desenvolvido um aplicativo como a Dora para o RO, melhorando a interação com os cidadãos para a apuração dos casos com maiores detalhes.

Referências

- ABDUL-KADER, Sameera A.; WOODS, John C. Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications*, v. 6, n. 7, 2015.
- ANDRADE, P.H.M.A. (2015) *Aplicação de Técnicas de Mineração de Textos para Classificação de Documentos: um Estudo da Automatização da Triagem de Denúncias na CGU*, Dissertação em Computação Aplicada da UnB.
- BRANDTZÆG, Petter; FØLSTAD, Asbjørn. Why People Use Chatbots. *The 4th International Conference on Internet Science*. 2017.
- BRANTINGHAM, Jeffrey; MOHLER, George. Does Predictive Policing Lead to Biased Arrests? Results from a Randomized Controlled Trial. *Statistics and Public Policy*, n. 5, p. 11-17, 2018.
- BRITO, Walkir; MOTTA, Claudia. Arcabouço Analítico de Textos em Denúncias: Novas Tecnologias no Tratamento e Análise de Denúncias Estudo de Caso Milícias. <https://doi.org/10.5753/semish.2020.11327> , v. 47., p. 174-185, 2020.
- BRITO, Walkir A.T.; MOTTA, Claudia L.R; BORGES, Zeca. Análise Semântica a partir de Sistemas Complexos na Base do Disque Denúncia. *Cadernos de Segurança Pública* , v. n 10, p. 1-16, 2018.
- CESAR, M. V. G. Classificação de falhas de equipamentos de unidade de intervenção em construção de poços marítimos por meio de mineração textual. 2017. Dissertação (Mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2017.
- CHANEY, Spencer et al. The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime. *Security Journal*, n. 21, p. 4-28, 2008.
- CHAUDHURI et al. An Overview of Business Intelligence Technology. *ACM*, N. 8., Vol. 54, 2011, 88-98.
- CIECHANOWSKI, Leon et al. In the shades of the Uncanny Valley: An Experimental Study of Human–Chatbot Interaction. *Future Generation Computer Systems*. v. 92, p. 539-548. 2018.
- DRAGT, Els; TIMMER, Jeroen. Dare to Ask: Learn to Ask Questions Like a Pro. 2020
- FERREIRA, H.H. (2019) *Processamento de Linguagem Natural e Classificação de textos em Sistemas Modulares*, Monografia do Departamento de Ciência da Computação da UnB.
- GAGLIO, Salvatore et al. Smart assistance for students and people living in a campus. In: 2019 IEEE International Conference on Smart Computing (SMARTCOMP). IEEE, 2019. p. 132-137.
- JIMOH, Rasheed et al. A Scalable Online Crime Reporting System, *African Journal of Computing & ICT*, v. 7, n. 1, 2014.
- KANABLE, R. Talking to tipsters: Websites and text messages allow valuable, anonymous communication, *Law Enforcement Technology*, v. 35, n. 11, p. 10-12, 2008.

- KUN, Chih Hao; LEROY, Gony. A Decision Support System: Automated Crime Report Analysis and Classification for E-Government. *Government Information Quarterly*, v. 31, n. 4, p. 534-544, 2014.
- LASLEY, James; PALOMBO, Bernadette Jones. When Crime Reporting Goes High-Tech: An Experimental Test of Computerized Citizen Response to Crime. *Journal of Criminal Justice*, v. 23, n. 6, p. 519-529, 1995.
- NASCIMENTO, R.M.F. (2019). Classificação automática de discursos de ódio em textos do twitter. 2019. 47 f. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação) – Unidade Acadêmica de Serra Talhada, Universidade Federal Rural de Pernambuco, Serra Talhada.
- PINHO, R. A., BRITO, W. A. T., MOTTA, C. L. R. & LIMA, P. V. Automatic Crime Report Classification through a weightless Neural Networks. ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017. Disponível em: <https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2017-42.pdf>. Acessado em 29/11/2021.
- RESHMI, S.; BALAKRISHNAN, Kannan. Implementation of an inquisitive chatbot for database supported knowledge bases. *sādhanā*, v. 41, n. 10, p. 1173-1178, 2016.
- RODRÍGUEZ, Juan Manuel; MERLINO, Hernán; FERNÁNDEZ, Enrique. Comportamiento Adaptable de Chatbots Dependiente del Contexto. *Revista Latinoamericana de Ingeniería de Software*, v. 2, n. 2, p. 115-136, 2014.
- ROSSI, R. G. (2015). Classificação automática de textos por meio de aprendizado de máquina baseado em redes. Tese de Doutorado, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos. doi:10.11606/T.55.2016.tde-05042016-105648.
- RUBIN, Alan. Uses-and-gratifications perspective on media effects. *The Sage handbook of media processes and effects*. p. 147-159, 2009.
- SONG, Wonsong et al. Using IM and SMS for Emergency Text Communications. 3rd International Conference on Principles, Systems and Applications of IP Telecommunications. n.4. p. 1-7.
- STIILPEN Jr, M. (2016). Um Arcabouço de Processamento de Textos Informais em Português Brasileiro para Aplicações de Mineração de Dados, Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Ouro Preto
- SUNDAR, Shyam; LIMPEROS, Anthony. Uses and Grats 2.0: New Gratifications for New Media. *Journal of Broadcasting & Electronic Media*, v. 57, 2013.
- WEIZENBAUM, Joseph. ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, v. 9, n. 1, p. 36-45, 1966.
- VARSHA, Naikwade et al. Grid Based Authentication for Online Crime Reporting System, *International Journal of Computer Science and Mobile Computing*, v. 4, n. 3, p. 308-313, 2015.