

Repressão à Criminalidade com o uso de Computação Cognitiva: Classificação de Urgência, Severidade e Relevância de Denúncias

Alexandre Filgueiras Costa¹, Claudia Lage Rebello da Motta¹

¹Programa 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
afcosta@ufrj.br, claudiam@nce.ufrj.br

Abstract. The Disque Denúncia (DD) is a telephone service for receiving anonymous complaints about criminal practices, we have developed a collaborative platform to complement the DD service to the population of Rio de Janeiro by offering a new channel for collecting complaints, DORA, a virtual assistant specialist in collecting complaints. The focus of this research is the classification of complaints received, in three dimensions: Urgency, Severity and Relevance, which aim to validate and improve the diffusion of the complaint to the responsible body with the necessary accuracy and celerity.

Resumo. O Disque Denúncia (DD) é um serviço telefônico de recebimento de denúncias anônimas sobre práticas criminosas, desenvolvemos uma plataforma colaborativa para complementar o atendimento do DD à população do Rio de Janeiro oferecendo um novo canal de coleta de denúncias, a DORA, uma assistente virtual (AV) especialista em coletar denúncias. O foco desta pesquisa é a classificação das denúncias captadas, em três dimensões: Urgência, Severidade e Relevância, que têm como objetivo validar e aprimorar a difusão da denúncia para o órgão responsável com a acurácia e celeridade necessária.

1. Introdução

Segurança pública é um tema preocupante no Brasil, mais especificamente no Rio de Janeiro não seria diferente. Ao longo dos anos o sucateamento das forças policiais tem deixado uma lacuna cada vez maior para ser preenchida pelos nossos governantes, atualmente o uso de tecnologias está cada vez mais em voga para desbancar os braços financeiros e logísticos do crime organizado [KU, Chih-Hao; LEROY, Gondy, 2014] e [CHEN, Hsinchun et al., 2004].

O Instituto MovRio, através do DD, vivencia a mesma crise que toda empresa brasileira vem sofrendo, com a perda de receita e doações, necessitando reduzir custos e otimizar os gastos. Essa redução de receita provoca a diminuição da capacidade de atendimento e conseqüente perda de capital intelectual. Uma denúncia pode ser feita a qualquer instante, e a agilidade no atendimento às demandas do setor público pode fazer toda a diferença no combate à criminalidade.

1.1. Visão Geral

O Instituto MovRio possui um relacionamento de colaboração com diversas áreas da sociedade, instituições públicas e do mundo corporativo, na figura do Disque Denúncia ele apoia os cidadãos na sua jornada de realização de denúncias. Com sua credibilidade, resultados atingidos e a garantia de anonimato do cidadão na captação das denúncias, garante grande presença nos meios de comunicação incluindo o programa de Procurados com recompensas.

Por ser uma organização da sociedade civil de interesse público (OSCI) depende do apoio financeiro das instituições públicas, privadas e de doações realizadas pela população. Além dessa simbiose, podemos destacar as parcerias com os órgãos responsáveis (Polícia Militar, Polícia Civil, Instituto Brasileiro do Meio Ambiente (IBAMA), Polícia Federal, entre outros) pela sequência da execução das denúncias triadas, e o Instituto de Segurança Pública do Rio de Janeiro (ISP) que consome os dados e os indicadores gerados pelo Instituto MovRio. Podemos observar o modelo operacional do Instituto MovRio na Figura 1.

Essa pesquisa, diante desse cenário, tem como objetivo a correta classificação das denúncias captadas. Com a construção da plataforma colaborativa de denúncias, o cidadão, passa a ter mais um canal para combater os problemas de segurança da sociedade atual. O Instituto MovRio, recebe uma oxigenação tecnológica e abrem-se as portas para a captação de novos recursos e parceiros que financiem a operação do DD. A academia representada pela Universidade Federal do Rio de Janeiro é o amálgama que une todos esses parceiros com sua expertise na área de soluções de problemas sociais e de segurança pública com o uso de sistemas complexos adaptativos.

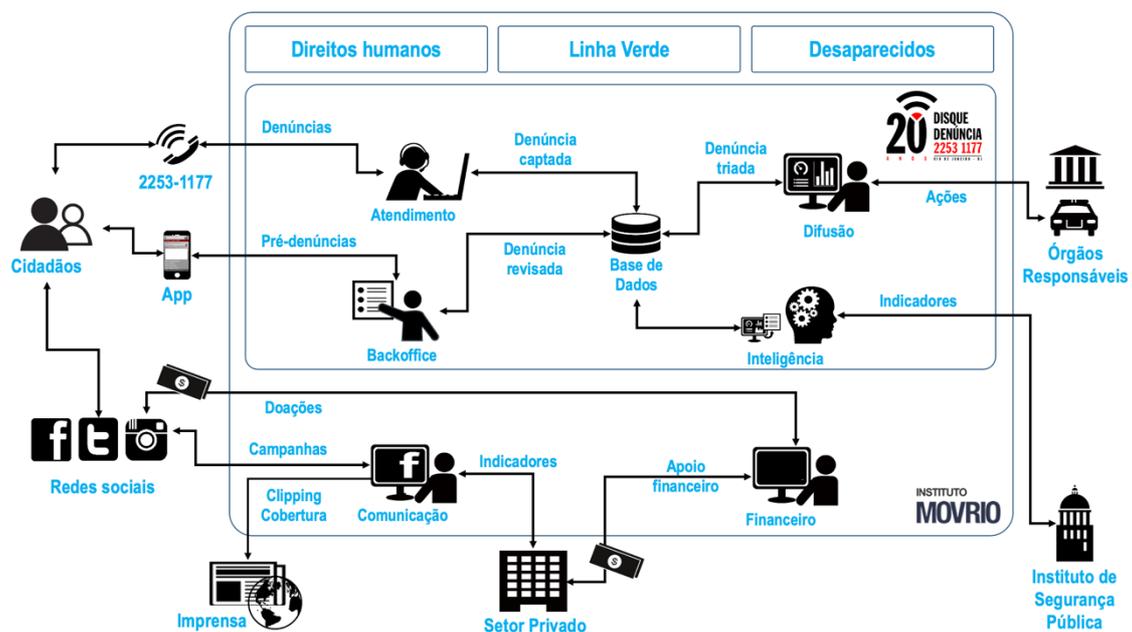


Figura 1. Modelo operacional do Instituto MovRio

1.2. Objetivo

O objetivo dessa pesquisa é garantir que a denúncia recepcionada esteja adequada e seja classificada conforme tabela definida pelos curadores da solução. A pergunta que este trabalho pretende responder é:

- Como garantir o alinhamento da classificação de Urgência, Severidade e Relevância da denúncia com a tabela definida pelos curadores?

1.3. Relevância

O tema segurança pública para o cidadão do estado do Rio de Janeiro é de extrema relevância visto que a nossa sociedade está sendo atacada a todo tempo por crimes como: tráfico de drogas, roubo de veículos, roubo de carga, crimes relacionados às armas. Os quatro principais temas tratados pelo sistema colaborativo (cerca de mais de 80% de todo volume de denúncias capturadas pelo Disque Denúncia)¹.

1.4. Trabalhos Relacionados

A DORA foi construída diferente de outros assistentes que respondem perguntas. Ela guia o cidadão, realizando perguntas pertinentes, para coletar o maior número de informações necessárias para a completude da denúncia [FERRAN, Cristianna Madeira De, 2022].

¹ Dados referentes ao mês de setembro do ano de 2018, obtidos junto ao Instituto MovRio.

Em [BRITO, Walkir, et al., 2020] podemos ver um arcabouço analítico de textos em denúncias com a temática “milícias”, para com diversas técnicas avançadas de mineração de dados e de extração de conhecimento, possam ser retiradas informações para tomada de decisão estratégicas dos gestores responsáveis. A classificação realizada é para a clusterização dos dados e otimização do processamento de textos pelo SIM (Sistema de Informação de Milícias) e não está associada às dimensões definidas nessa dissertação.

Em relação à classificação de denúncias temos em [PINHO, Rafael Adnet et al., 2017] classificações primárias e secundárias relacionadas à dimensão “tipo de crime” reportado com o uso de redes neurais sem peso. Na plataforma colaborativa de denúncias aqui construída, a classificação primária de tipo de crime já é fornecida pelo cidadão no início do reporte da sua denúncia e caso haja alguma sub-denúncia associada a DORA realiza perguntas para uma maior completude da captação da denúncia.

Podemos também encontrar em [RANGEL, Pimentel, Chiote, 2017] classificações relacionadas ao modelo resumido de Urgência, Severidade e Relevância para as denúncias realizadas para a “linha verde” (de cunho ambiental) do DD e alinhadas com o código penal ambiental do estado do Rio de Janeiro.

Essa pesquisa amplia o conceito de classificação das denúncias para outras duas dimensões (severidade e relevância) além da dimensão existente (urgência), que está relacionada a “temas da semana” para os temas: tráfico de drogas, roubo de carga, roubo de veículo e armas.

2. Plataforma Colaborativa de Denúncias

A plataforma colaborativa de denúncias, construída pelo grupo de pesquisa, é um conjunto de soluções orquestradas com o objetivo de prover um ambiente em que o cidadão possa realizar uma denúncia e que após a coleta e classificação da denúncia seja realizada a difusão dela para as áreas competentes, ela é composta por cinco componentes: Bases de conhecimento, API para prover serviços, *Chat* cognitivo, Ferramenta administrativa e um Sistema de classificação das denúncias, sendo este, o foco principal e a minha contribuição nesta pesquisa.

2.1. Abordagem Metodológica

Para construir a plataforma de coleta e classificação de denúncias a abordagem metodológica adotada nessa pesquisa foi o Design Science Research [PEFFERS, Ken et al., 2007] (DSRM) como ilustrado na Figura 2.

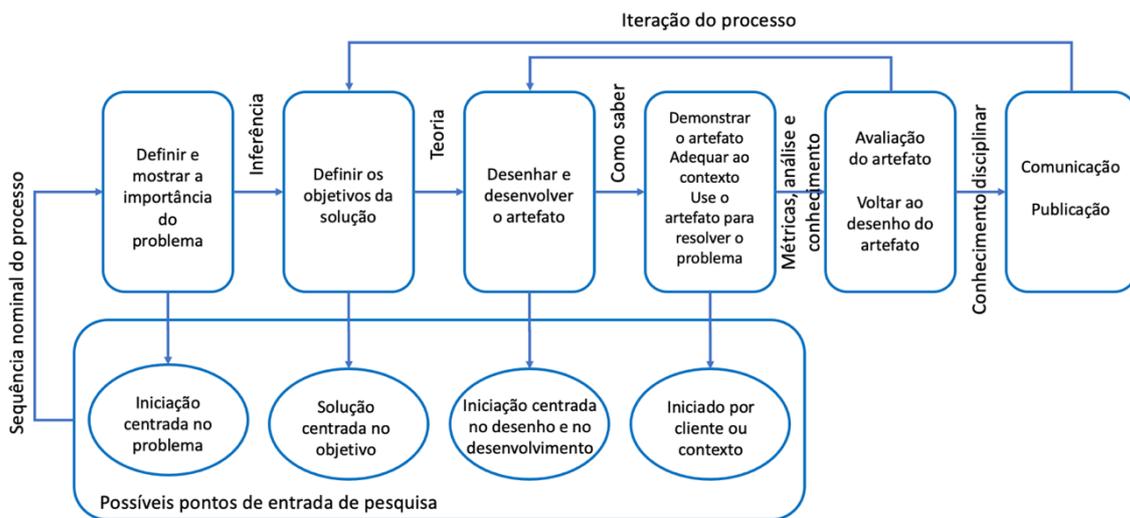


Figura 2. Modelo do processo da DSRM

No ciclo “iteração do processo” ilustrado na Figura 2, elencamos um tipo de denúncia a ser incorporada no Assistente Virtual (AV) e entregue ao time do DD, ao final de um conjunto de sprints, publicamos a *release*, realizamos a cerimônia de demonstração e avaliamos o AV junto com o time do DD. Ao longo do próximo ciclo, enquanto desenvolvíamos a captação de outro tipo de denúncia pelo AV, os atendentes estavam testando e usando o AV, conseqüentemente gerando dados para que a análise e as métricas pudessem ser coletadas e o processo de curadoria do AV fosse realizado. Os itens de melhoria e as análises realizadas para o aprimoramento do AV foram tratadas ao longo dos ciclos subseqüentes e ao final da entrega da última *release*, após o tempo de maturação do AV realizamos ajustes apontados nas curadorias relacionados à captação e à classificação das denúncias.

Durante os dois dias de colaboração nas sessões do Design Thinking (DT) foram aplicadas diversas dinâmicas com as áreas de atendimento, inteligência, difusão, comunicação, gestão e de tecnologia da informação do Instituto, além dos consultores de cidadania corporativa da IBM e os mestrandos e doutorandos da Universidade Federal do Rio de Janeiro. Chegamos na identificação do problema junto ao time do Instituto MovRio que teve como objetivo promover uma discussão única e multidisciplinar sobre os principais pontos de melhorias e projetar uma visão de futuro da aplicação das capacidades de sistemas complexos adaptativos no Instituto MovRio. Ao longo do DT foram geradas 100 ideias que foram categorizadas e 35 delas foram priorizadas, na Figura 3 mostramos as categorias e as 8 principais ideias selecionadas para o roadmap do Instituto MovRio.

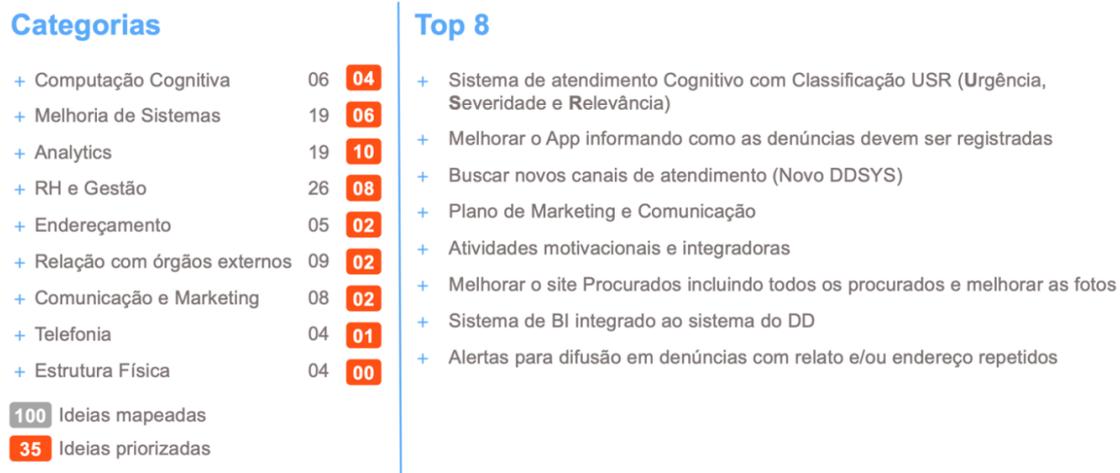


Figura 3: Categorização das ideias do DT

2.2. Definição dos Objetivos da Plataforma

Os objetivos mapeados ao longo da aplicação do DT foram:

1. Promover o debate sobre o atendimento ao cidadão;
2. Definir um Roadmap tecnológico para o Instituto;
3. Listar iniciativas voltadas ao uso de tecnologias cognitivas;
4. Identificar direcionadores de negócio corporativos a curto, médio e longo prazos;
5. Identificar parceiros e alavancar a captação de doações.

O objetivo principal da plataforma colaborativa de denúncias construída é que o Instituto MovRio tenha um produto para engajar o cidadão na jornada de realização de denúncias através de um sistema cognitivo que simule uma conversa, utilizando linguagem natural, entre o ser humano e a máquina (*chatbot*). Além disso, existem outros sistemas satélites que complementam essa plataforma com o objetivo de realizar o suporte dessa jornada e aumentar o índice de sucesso da captura das denúncias e sua difusão.

2.3. Desenho e Implementação da Plataforma

A plataforma teve sua arquitetura baseada na arquitetura de referência de sistemas cognitivos de domínio conversacional (*chatbots*). A metodologia de desenvolvimento utilizada foi a metodologia ágil "*IBM Garage Methodology*" com entregas regulares a cada sprint. As entregas foram apresentadas para as atendentes e corpo técnico do Instituto MovRio em cerimônias de "sprint review" na sede do Instituto MovRio. Essa pesquisa foca no processo de

classificação da denúncia em Urgência, Severidade e Relevância, que é usado para a otimização do processo de difusão da mesma ilustrado na Figura 4 e será explicado em maiores detalhes no Capítulo 3.

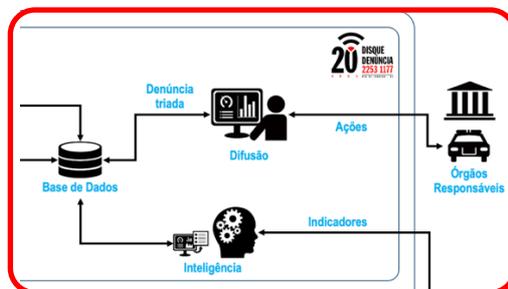


Figura 4: Processo de difusão das denúncias no DD

Diagrama de casos de uso modelam o comportamento de um sistema e ajudam a capturar os requisitos dele. Eles descrevem funcionalidades de alto nível, além disso, identificam as interações entre o sistema e seus atores. Para a plataforma colaborativa de denúncias, dividimos em subsistemas os diagramas de caso de uso: Chatbot, Portal Administrativo e API da plataforma como visto na Figura 5.

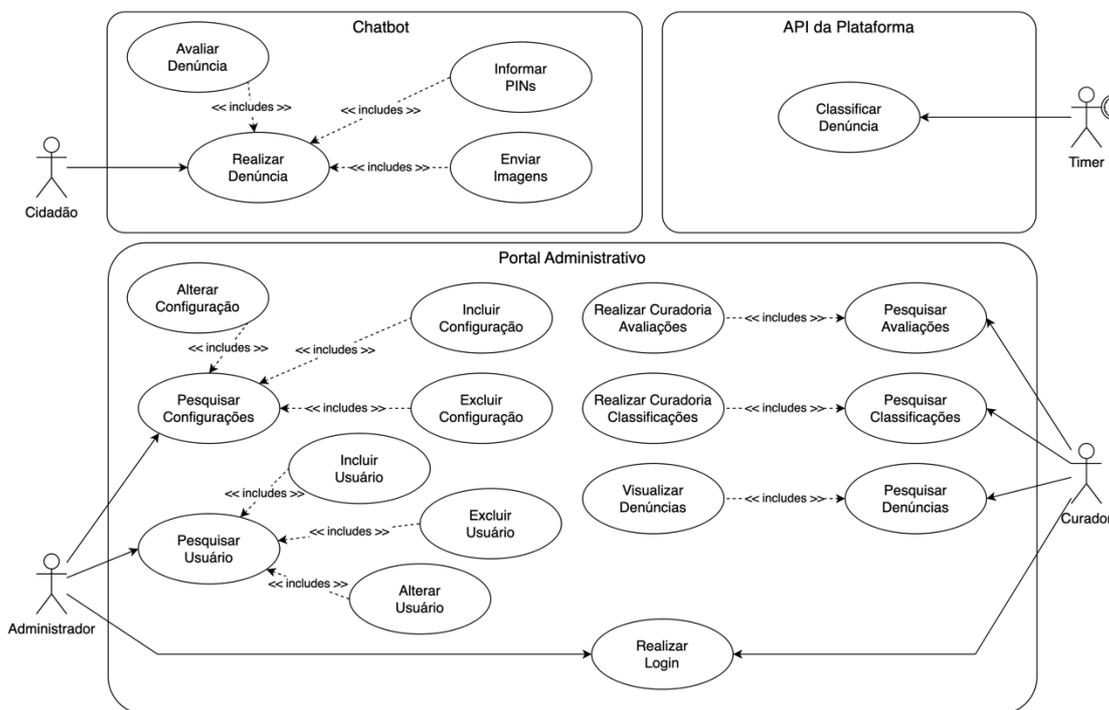


Figura 5: Diagrama de casos de uso

Por essa pesquisa não ter o foco em comparação de arquiteturas para sistemas de *chatbot*, e sim na classificação da denúncia coletada, a arquitetura escolhida, por ser uma arquitetura agnóstica à IA utilizada na plataforma, se baseia na referência de sistemas cognitivos de domínio conversacional

(*chatbots*) [SRIVASTAVA, Saurabh; PRABHAKAR, T. V., 2019] que possui alguns componentes como: canais de entrada, *chat* responsivo, API, serviços cognitivos para PLN, serviços analíticos, base de dados, painel administrador, ferramenta de curadoria, integração com serviços corporativos, serviços públicos e serviços de parceiros, etc. Todos, tendo como base, uma sólida política de governança, monitoramento, segurança e de entrega contínua como ilustrado na Figura 6.

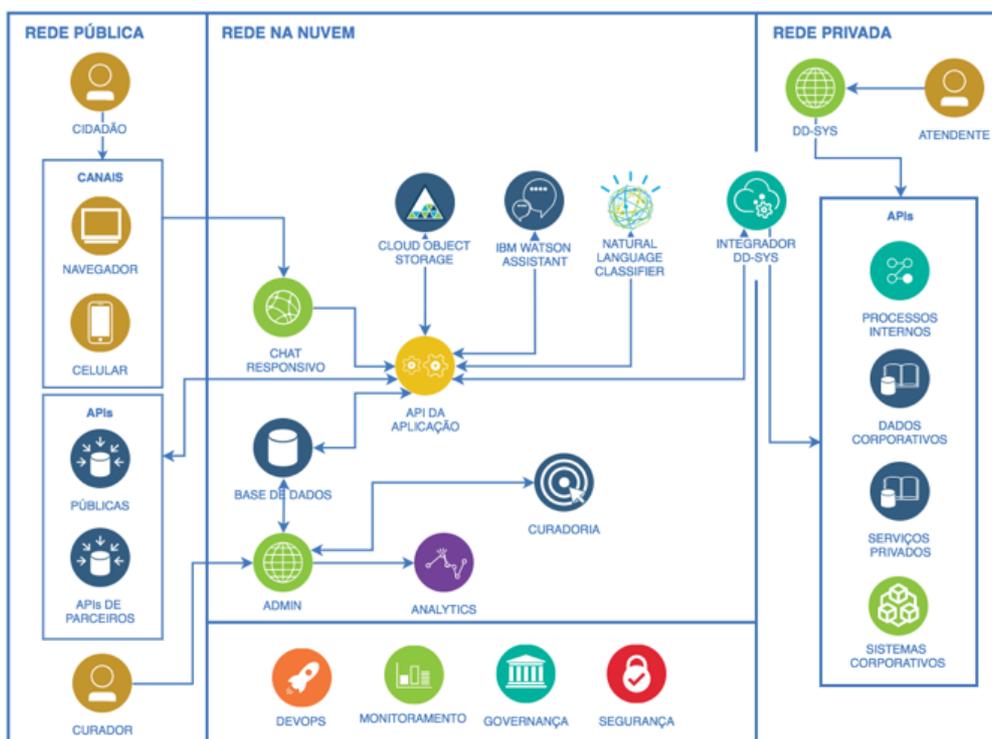


Figura 6: Arquitetura da Plataforma Colaborativa de denúncias

3. Sistema de Classificação De Denúncias

Atualmente o processo de classificação das denúncias coletadas pelos atendentes do DD possui somente uma dimensão, "Urgência" e baseia-se em "temas da semana", por exemplo, se o tema da semana é o traficante "Rogério 157", todas as denúncias que estão relacionadas ao traficante serão marcadas como "urgente" pelo atendente. Com isso, se na semana seguinte tivermos uma denúncia relacionada ao "Rogério 157" e o tema da semana for outro, perderemos essa classificação de "urgente" na denúncia recebida. Outro cenário levantado durante a sessão de DT, é que somente uma dimensão de classificação da denúncia é ineficaz para a classificação correta e um novo modelo de classificação deveria ser proposto.

3.1. Definição da Classificação

Ao longo do DT foi proposto pelo Zeca Borges, fundador do DD, um novo modelo de classificação com mais duas outras dimensões, totalizando três dimensões:

1. **Urgência** – Relacionada a temporalidade da denúncia;
2. **Severidade** – Relacionada ao uso de violência da denúncia;
3. **Relevância** – Relacionada aos envolvidos na denúncia.

Além das dimensões, foram criadas três classes para representar o nível de criticidade delas (baixa, média e alta), a escolha de três classes se dá por conta da simplificação do modelo de treinamento do serviço de classificação. Com esse novo modelo proposto, a denúncia realizada sobre o “Rogério 157” seria classificada com uma urgência média para baixa, porém com uma relevância alta, assim permitindo a difusão correta da denúncia e otimizando os recursos já escassos.

3.2. Visão Macro da Base de Classificação

No núcleo da base de classificação da plataforma é utilizado o serviço cognitivo *IBM Watson Natural Language Classifier*, um serviço disponível na *IBM Cloud* sobre um modelo de tarifação por consumo de chamadas da API. Para o treinamento do modelo de classificação foram utilizados exemplos como os abaixo listados na Tabela 1.

Tabela 1. Resultado da classificação do piloto

Dimensão	Classe	Exemplo
Urgência	Baixo	Policiais são coniventes com o tráfico.
Urgência	Baixo	Gerente do tráfico de drogas.
Urgência	Médio	Menores de idade usando drogas e ameaçando moradores.
Urgência	Médio	Gritos e insultos contra a companheira.
Urgência	Alto	Virou uma cracolândia e já arrombaram a paróquia.
Urgência	Alto	Consumo de drogas no horário escolar quando transitam inúmeras crianças.
Severidade	Baixo	Tráfico de drogas rola solto no ponto de moto táxi.
Severidade	Baixo	Traficantes armados.
Severidade	Médio	Carro abandonado usado para guardar armas e droga.
Severidade	Médio	Dono das bocas e gerente do tráfico aterrorizando os moradores.
Severidade	Alto	Vivendo um inferno, vários fuzis, a comunidade agora anda aterrorizada.
Severidade	Alto	Tráfico de drogas rola solto no ponto de moto táxi.
Relevância	Baixo	Guarda produto de roubo de carga.
Relevância	Baixo	Depósito de carga roubada.
Relevância	Médio	Elementos no galpão com vários maquinários para falsificação das cervejas.
Relevância	Médio	Muito elementos descarregando carreta roubada de botijões de cozinha.
Relevância	Alto	Milicianos tomaram posse do bairro barbarizando moradores e comerciantes.
Relevância	Alto	Rogério 157, com muita arma de guerra, drogas e viciados causando terror.

3.3. Implementação

O mecanismo de classificação de uma denúncia capturada é representado temporalmente pelo diagrama de sequência ilustrado na Figura 7:

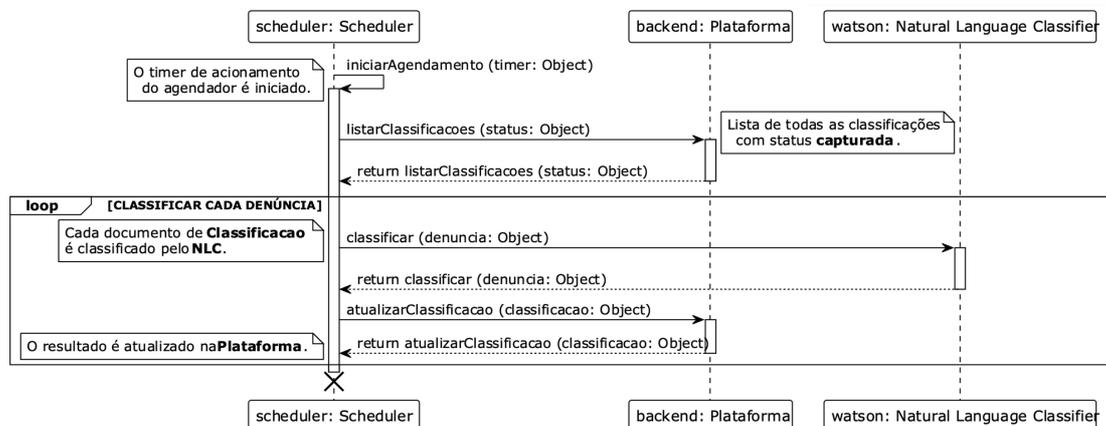


Figura 7: Diagrama de sequência da classificação das denúncias

3.4. Validação

Antes do processo de avaliação da plataforma, separamos 22 denúncias (aproximadamente 20% da amostra total) com a seguinte distribuição: 2 de arma, 3 de roubo de carga, 3 de roubo de veículo, 3 de outros e 11 de tráfico de drogas, oriundas da inserção realizada em fevereiro de 2020 pelos atendentes do DD para realizarmos uma última rodada de classificações. Com os resultados da classificação, calculamos os percentuais das classes: baixa, média e alta dentro das dimensões: Urgência, Severidade e Relevância, e a distribuição segue a Figura 8. Validando com o time de especialistas do DD a distribuição da dimensão Urgência ser classificada em 100% baixa segue o esperado, dado que a dimensão reflete a temporalidade da denúncia. Por conta do termo de confidencialidade assinado, omitimos o texto original das denúncias e estamos somente mapeando os identificadores delas (identificadores da base de dados) nos resultados.

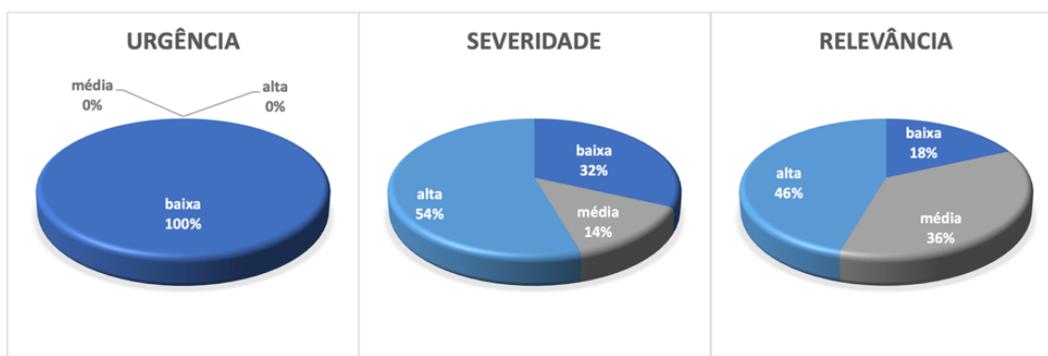


Figura 8: Percentual das classes encontradas na classificação da massa de testes

4. Avaliação do Artefato

Após a validação das classificações das denúncias de teste selecionadas na Seção 3.4 iniciamos a avaliação do artefato construído conforme a metodologia ilustrada na Figura 2 da Seção 2.1. O objetivo da avaliação do artefato construído foi monitorar a captação e validar que a classificação das denúncias inseridas pelas atendentes do DD durante a fase piloto da plataforma colaborativa de denúncias estava de acordo com a classificação do time de especialistas no assunto do DD.

4.1. Separação das Denúncias

Para a avaliação do artefato separamos as 78 denúncias inseridas pelas atendentes do DD no piloto realizado em fevereiro de 2020, com a seguinte distribuição: 7 de arma, 10 de roubo de carga, 12 de roubo de veículo, 9 de outros e 40 de tráfico de drogas, continuamos omitindo o texto original da denúncia e somente deixando os identificadores delas (identificadores da base de dados).

4.2. Execução da Classificação

Como visto na Seção 3.3, a execução da classificação das denúncias é realizada por um mecanismo de tarefas agendadas presente na API da plataforma colaborativa de denúncias que de tempos em tempos seleciona as classificações com o status "capturada", consome o serviço cognitivo responsável pela classificação e atualiza a base de dados movendo a classificação para o status "classificada". Com os resultados da classificação, calculamos os percentuais das classes: baixa, média e alta dentro das dimensões: Urgência, Severidade e Relevância, e a distribuição segue a Figura 9.

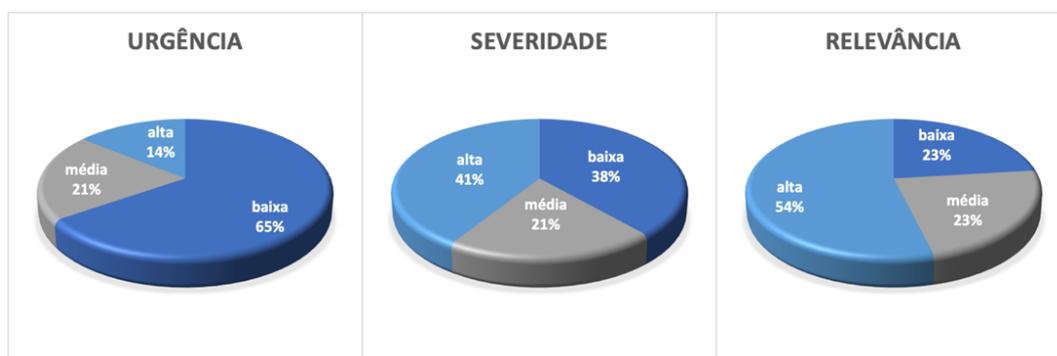


Figura 9: Percentual das classes encontradas na classificação do piloto

Como resultado da classificação das denúncias do piloto, obtivemos os indicadores mapeados na Tabela 2, apesar de termos os valores desmembrados das classes alta, média e baixa nos indicadores, a classificação resultante tem como valor final somente a classe de maior valor de cada dimensão.

Tabela 2. Resultado da classificação do piloto

ID Classificação	Tipo de Denúncia	Urgência			Severidade			Relevância		
		baixa	média	alta	baixa	média	alta	baixa	média	alta
5f39658dc87b007c9a8b0c	armas	0.9945	0.0036	0.0019	0.0161	0.14	0.8439	0.3532	0.2269	0.4199
5f383fb1caf87b007c9a8a58	armas	0.0346	0.9612	0.0042	0.4182	0.5326	0.0493	0.0929	0.8941	0.013
5f370bc7caf87b007c9a893c	armas	0.9647	0.0293	0.006	0.0828	0.6353	0.2819	0.0886	0.8953	0.0161
5f370999caf87b007c9a8920	armas	0.1234	0.7812	0.0955	0.0134	0.0035	0.9831	0.0302	0.0241	0.9457
5f36f5f9caf87b007c9a88d2	armas	0.9519	0.0429	0.0052	0.1261	0.3471	0.5269	0.1968	0.0591	0.7441
5f25bb7108951300872c442e	armas	0.0023	0.0083	0.9894	0.8175	0.0744	0.108	0.0864	0.861	0.0526
5f25b09e08951300872c4417	armas	0.8363	0.1545	0.0092	0.0679	0.5148	0.4173	0.0691	0.0389	0.892
5f3a77a0caf87b007c9a8c7c	outros	0.992	0.0045	0.0036	0.0487	0.0234	0.9279	0.0531	0.109	0.8379
5f3a7577caf87b007c9a8c6c	outros	0.0723	0.9155	0.0122	0.1241	0.015	0.8609	0.011	0.0593	0.9297
5f39a842caf87b007c9a8bba	outros	0.0146	0.0406	0.9448	0.6546	0.0954	0.25	0.0433	0.0128	0.9438
5f27334e08951300872c4494	outros	0.9716	0.0075	0.0209	0.053	0.7479	0.1991	0.0836	0.1028	0.8136
5f27319b08951300872c4484	outros	0.9395	0.0507	0.0098	0.0048	0.0047	0.9905	0.0324	0.3329	0.6347
5f25bd6908951300872c4440	outros	0.9642	0.019	0.0168	0.0492	0.7578	0.193	0.0596	0.0369	0.9035
5f384a2dc87b007c9a8a9a	rouboCarga	0.9532	0.0413	0.0056	0.9267	0.023	0.0503	0.7448	0.2242	0.031
5f29b94c08951300872c4668	rouboCarga	0.0552	0.9359	0.0089	0.3624	0.0235	0.6141	0.6386	0.0142	0.3472
5f29aedf08951300872c462e	rouboCarga	0.0053	0.0233	0.9715	0.0218	0.0091	0.9691	0.8971	0.0169	0.086
5f29a7e908951300872c460a	rouboCarga	0.9627	0.0308	0.0065	0.7902	0.1524	0.0574	0.9291	0.0592	0.0117
5f29a2e708951300872c45dc	rouboCarga	0.9604	0.0347	0.0049	0.8564	0.0479	0.0957	0.3114	0.6498	0.0388
5f2990ea08951300872c45c2	rouboCarga	0.0018	0.0056	0.9926	0.3497	0.541	0.1092	0.122	0.8677	0.0103
5f298dfa08951300872c45ac	rouboCarga	0.0017	0.0034	0.9949	0.218	0.0512	0.7307	0.9889	0.0078	0.0032
5f298bc308951300872c4594	rouboCarga	0.9625	0.0323	0.0052	0.7535	0.0154	0.2311	0.3381	0.0128	0.6491
5f29898008951300872c4582	rouboCarga	0.0066	0.9914	0.002	0.0157	0.0265	0.9578	0.1706	0.0874	0.742
5f29865d08951300872c4568	rouboCarga	0.0516	0.9385	0.0099	0.2713	0.1212	0.6075	0.0367	0.0354	0.9279
5f29848a08951300872c4556	rouboCarga	0.99	0.0067	0.0033	0.9338	0.011	0.0552	0.2541	0.0456	0.7003
5f29825f08951300872c453e	rouboCarga	0.7947	0.1888	0.0166	0.7593	0.0148	0.2259	0.6279	0.0373	0.3348
5f29806308951300872c4528	rouboCarga	0.9478	0.0219	0.0302	0.3117	0.024	0.6643	0.2336	0.0517	0.7147
5f297c4608951300872c450e	rouboCarga	0.9709	0.0252	0.0039	0.3386	0.0151	0.6463	0.6598	0.0335	0.3068
5f29781708951300872c44f2	rouboCarga	0.9936	0.0034	0.003	0.6339	0.0181	0.3479	0.8517	0.1227	0.0256
5f39b297caf87b007c9a8c2e	rouboVeiculo	0.0052	0.0338	0.961	0.456	0.0707	0.4733	0.522	0.3273	0.1507
5f39ab3caf87b007c9a8be8	rouboVeiculo	0.8747	0.0803	0.045	0.0277	0.5473	0.425	0.7222	0.0168	0.261
5f3968f6caf87b007c9a8b2e	rouboVeiculo	0.0054	0.0418	0.9528	0.029	0.4451	0.5259	0.2799	0.7012	0.0189
5f3843a1caf87b007c9a8a7a	rouboVeiculo	0.1372	0.8539	0.0089	0.9767	0.0105	0.0128	0.7178	0.2602	0.0219
5f29c1fe08951300872c4696	rouboVeiculo	0.972	0.0244	0.0036	0.0496	0.4822	0.4683	0.2845	0.0334	0.6821
5f29bec208951300872c4678	rouboVeiculo	0.2029	0.7853	0.0118	0.0623	0.7252	0.2125	0.266	0.0779	0.6561
5f29b67708951300872c4658	rouboVeiculo	0.0351	0.1919	0.773	0.766	0.0141	0.2199	0.1315	0.0579	0.8106
5f29b46608951300872c4648	rouboVeiculo	0.9714	0.0245	0.004	0.9544	0.0208	0.0249	0.4525	0.0183	0.5292
5f29ab9d08951300872c461a	rouboVeiculo	0.9692	0.0245	0.0063	0.4201	0.092	0.4879	0.9325	0.0077	0.0598
5f29a51d08951300872c45f0	rouboVeiculo	0.9701	0.0153	0.0146	0.1717	0.0448	0.7835	0.9856	0.0041	0.0102
5f3a73acc87b007c9a8c5c	trafico	0.9691	0.0269	0.004	0.9099	0.0625	0.0276	0.2641	0.7105	0.0254
5f39c24bc87b007c9a8c44	trafico	0.9921	0.006	0.0019	0.0855	0.0672	0.8474	0.3017	0.6491	0.0491
5f39af4dc87b007c9a8c1e	trafico	0.9657	0.0069	0.0274	0.1515	0.3521	0.4964	0.1341	0.3145	0.5514
5f39ad7ac87b007c9a8c02	trafico	0.0036	0.9944	0.002	0.0245	0.4247	0.5509	0.3158	0.5	0.1842
5f39a9c5caf87b007c9a8b04	trafico	0.008	0.9879	0.0041	0.1046	0.0117	0.8837	0.0076	0.0095	0.9829
5f39a72ac87b007c9a8baa	trafico	0.0303	0.0082	0.9615	0.0084	0.0043	0.9873	0.1731	0.0433	0.7836
5f39a523caf87b007c9a8b90	trafico	0.1447	0.8468	0.0086	0.2383	0.686	0.0757	0.3824	0.4165	0.2012
5f39a2aec87b007c9a8b7c	trafico	0.0019	0.0063	0.9918	0.4227	0.5497	0.0276	0.0493	0.0327	0.9181
5f398403caf87b007c9a8b60	trafico	0.002	0.0026	0.9954	0.2052	0.0257	0.7691	0.3275	0.1671	0.5055
5f398115caf87b007c9a8b4f	trafico	0.7489	0.2365	0.0145	0.0664	0.5895	0.3441	0.0674	0.4391	0.4935
5f3961bcc87b007c9a8b7c	trafico	0.9923	0.0043	0.0035	0.9734	0.0045	0.0222	0.1637	0.0497	0.7865
5f395fe2caf87b007c9a8ad	trafico	0.9907	0.005	0.0043	0.9789	0.0084	0.0126	0.023	0.0398	0.9373
5f395c80caf87b007c9a8ab6	trafico	0.9932	0.0016	0.0052	0.9269	0.027	0.0461	0.2678	0.0693	0.6628
5f383e57caf87b007c9a8a3c	trafico	0.9618	0.0339	0.0043	0.9047	0.0855	0.0099	0.5577	0.4039	0.0384
5f383937caf87b007c9a8a1e	trafico	0.9707	0.0252	0.0041	0.6948	0.2294	0.0758	0.2675	0.6604	0.0721
5f3834c1caf87b007c9a8a06	trafico	0.0028	0.9947	0.0025	0.8978	0.0671	0.0352	0.2051	0.7098	0.0852
5f38324dc87b007c9a89f4	trafico	0.9932	0.0051	0.0017	0.5928	0.1275	0.2797	0.0024	0.9746	0.023
5f382c2cc87b007c9a89da	trafico	0.9295	0.0618	0.0087	0.2062	0.0208	0.773	0.0081	0.0122	0.9798
5f382708caf87b007c9a89c0	trafico	0.1544	0.8297	0.0159	0.0741	0.0363	0.8897	0.0269	0.0153	0.9578
5f381c6bc87b007c9a89aa	trafico	0.9924	0.0058	0.0018	0.0658	0.7171	0.2172	0.2532	0.0304	0.7163
5f381a45caf87b007c9a898a	trafico	0.0036	0.9885	0.0079	0.0389	0.0122	0.9489	0.0593	0.1771	0.7637
5f381798caf87b007c9a8976	trafico	0.9942	0.0043	0.0015	0.2097	0.73	0.0602	0.0647	0.8462	0.0891
5f3815fbc87b007c9a8964	trafico	0.9905	0.0074	0.0021	0.9219	0.0564	0.0216	0.0663	0.8109	0.1228
5f3813a8caf87b007c9a8954	trafico	0.9886	0.0089	0.0025	0.2188	0.0221	0.7591	0.1595	0.0385	0.802
5f36f92ac87b007c9a8900	trafico	0.9947	0.003	0.0023	0.9278	0.0409	0.0313	0.2824	0.0242	0.6935
5f36f7d6caf87b007c9a88e8	trafico	0.0021	0.994	0.0039	0.9169	0.0554	0.0277	0.8846	0.0879	0.0275
5f36f30fc87b007c9a88c0	trafico	0.006	0.9922	0.0018	0.9847	0.0086	0.0066	0.7682	0.175	0.0568
5f36f11ac87b007c9a88ac	trafico	0.9945	0.003	0.0025	0.192	0.0474	0.7605	0.0033	0.9638	0.0329
5f36ec9caf87b007c9a8894	trafico	0.9949	0.0023	0.0028	0.5525	0.1549	0.2926	0.5332	0.1094	0.3574
5f36ea89caf87b007c9a887a	trafico	0.9945	0.0036	0.002	0.6186	0.0303	0.3511	0.023	0.0283	0.9487
5f36e787caf87b007c9a8864	trafico	0.9641	0.0293	0.0065	0.1807	0.0287	0.7907	0.042	0.0109	0.9471
5f36e4d7caf87b007c9a884e	trafico	0.9949	0.0033	0.0018	0.5163	0.0677	0.4159	0.0777	0.0517	0.8706
5f36e0c8caf87b007c9a8836	trafico	0.9944	0.0036	0.002	0.7612	0.1932	0.0456	0.2741	0.6157	0.1102
5f36db54caf87b007c9a881e	trafico	0.9934	0.003	0.0036	0.076	0.0122	0.9118	0.063	0.0302	0.9068
5f36d6ccc87b007c9a880c	trafico	0.9911	0.0061	0.0028	0.0714	0.0207	0.9079	0.0399	0.0455	0.9146
5f273b4608951300872c44d0	trafico	0.9888	0.0077	0.0036	0.6467	0.1478	0.2054	0.2775	0.1233	0.5992
5f27349408951300872c44a8	trafico	0.9924	0.0044	0.0032	0.6986	0.2339	0.0675	0.0472	0.0249	0.9279
5f272fa08951300872c4474	trafico	0.991	0.0069	0.0021	0.2059	0.0207	0.7733	0.3147	0.1634	0.5219
5f25cba208951300872c445c	trafico	0.9937	0.0047	0.0016	0.022	0.5969	0.3811	0.5352	0.3589	0.1059
5f22058908951300872c4401	trafico	0.9926	0.0056	0.0018	0.4105	0.5369	0.0526	0.257	0.633	0.11

4.3. Avaliação dos Especialistas

Após uma análise do time de especialistas do DD foi identificado uma taxa de assertividade de 93% das classificações realizadas na avaliação do produto. Vale ressaltar alguns pontos relacionados às dimensões Urgência, Severidade e Relevância. Com relação a dimensão Urgência, que está relacionada a temporalidade, a maioria das denúncias se passam no passado, são muitas denúncias com termos como: "ontem", "semana passada", "frequentemente" e poucas com termos no presente, a justificativa dada pelo SME é que normalmente os denunciantes não se expõem realizando a denúncia "ao vivo" e as reportam posteriormente em algum lugar seguro. A dimensão Severidade, relacionada ao uso de violência, obteve uma classificação de mais de 60% para as classes "alta" e "média" devido ao uso de armamento pesado (fuzis, pistolas automáticas e metralhadoras) reportado nas denúncias, o que infelizmente se mantém presente no dia a dia do nosso Estado. Como a captação das denúncias realiza o cruzamento das informações coletadas com a entidade "@procurados" mapeada na base de conhecimento da DORA, a nossa assistente virtual, o percentual para as classes "alta" e "média" obteve um índice maior que 75% para a Dimensão Relevância, validando que esse cruzamento com a base de procurados é de extrema importância para a classificação das denúncias.

De maneira geral a base de conhecimento da classificação das denúncias da plataforma colaborativa está muito aderente ao esperado pelo time de especialistas, além disso a plataforma provê um mecanismo de curadoria das denúncias captadas e classificadas, onde o curador pode validar tanto o *feedback* do denunciante como acrescentar detalhes para a classificação realizada.

5. Conclusão

O objetivo dessa dissertação é garantir que a denúncia recepcionada pela plataforma colaborativa de denúncias esteja adequada e seja classificada conforme tabela definida pelos curadores da solução. A correta coleta e classificação da denúncia nas dimensões: Urgência, Severidade e Relevância tem um alto impacto na difusão para o órgão competente e uma difusão correta remete diretamente ao indicador de repressão à criminalidade. Todas as evoluções das bases de conhecimento (captura e classificação) podem ser feitas pelo portal administrativo, na persona dos curadores da solução.

Com isso, podemos dizer que foi atingido com sucesso e com o aval do time de especialistas no assunto do DD, garantimos o alinhamento da classificação das dimensões Urgência, Severidade e Relevância, através da avaliação da classificação descrita na Seção 3.4 e da avaliação do artefato descrita no Capítulo 4.

5.1. Limitações

Uma limitação deste trabalho está relacionada a propriedade intelectual das bases de conhecimento, dos artefatos gerados (incluindo código fonte) e dos exemplos utilizados para o treinamento e a execução dos experimentos, todos os participantes da parceria realizada entre o Instituto MovRio, a UFRJ e a IBM assinaram um termo de confidencialidade e os dados não podem ser disponibilizados para o público. O código fonte da plataforma colaborativa e as bases de conhecimento da captação e da classificação são propriedades intelectuais do Instituto MovRio.

5.2. Trabalhos Futuros

Como trabalhos futuros identificamos algumas oportunidades após a construção da plataforma colaborativa de denúncias:

1. Implementar um mecanismo de identificação de fraudes das denúncias capturadas, como exemplo, uma denúncia que indica um traficante da facção A em uma área de uma facção B, possivelmente é uma fraude para possibilitar que o traficante se locomova ou que prejudique a facção B.
2. Uso da plataforma colaborativa em outras áreas, o sistema de classificação poderia ser customizado para realizar a classificação de outros cenários, como por exemplo a priorização de obras de uma cidade ou a priorização de eventos dentro de uma catástrofe ambiental.
3. Utilizar as APIs e mecanismos de ciência de dados para tirar o que há escondido dentro dos dados coletados, por exemplo relacionando com as manchas de calor mapeadas pelo ISP, ou ainda, realizar a logística de operações de inteligência policiais baseadas na coleta dos PINs, usando a latitude e longitude para melhorar a performance das operações.

Referências

- BRITO, Walkir et al. Sistema de Informações de Milícias: Proposta de um Arcabouço Analítico para Análise de Textos em Dados de Denúncia. In: Anais do XLVII Seminário Integrado de Software e Hardware. SBC, 2020. p. 174-185.
- CHEN, Hsinchun et al. Crime data mining: a general framework and some examples. *computer*, v. 37, n. 4, p. 50-56, 2004.
- FERRAN, Cristianna Madeira De. USO DE COMPUTAÇÃO COGNITIVA NO PROCESSO SEMÂNTICO DE UM ATENDENTE VIRTUAL: UMA APLICAÇÃO NA ÁREA DE SEGURANÇA PÚBLICA, 2022.

- KU, 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.
- PEFFERS, Ken et al. A design science research methodology for information systems research. *Journal of management information systems*, v. 24, n. 3, p. 45-77, 2007.
- PINHO, Rafael Adnet et al. Automatic crime report classification through a weightless neural network. In: *Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. 2017. p. 165-170.
- SRIVASTAVA, Saurabh; PRABHAKAR, T. V. A Reference Architecture for Applications with Conversational Components. In: *2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS)*. IEEE, 2019. p. 1-5.